



UNIVERSITAT DE
BARCELONA

Estudio epigenómico de los factores de riesgo y pronóstico en el ictus isquémico

Natalia Cullell Fornés

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITAT DE
BARCELONA



Fundació
Docència i Recerca
MútuaTerrassa

**ESTUDIO EPIGENÓMICO DE LOS FACTORES DE RIESGO Y PRONÓSTICO EN EL ICTUS
ISQUÉMICO**

Memoria de tesis doctoral presentada por

NATALIA CULLELL FORNÉS

Para optar al grado de doctora por la Universidad de Barcelona

Dirigida por:

JERZY KRUPINSKI

Fundació Docència i Recerca MútuaTerrassa / Hospital Universitari MútuaTerrassa)

ISRAEL FERNÁNDEZ CADENAS

Hospital de la Santa Creu i Sant Pau / Fundació Docència i Recerca MútuaTerrassa

Tutor:

ÁNGEL MARÍA CHAMORRO

Hospital Clínic de Barcelona

Programa de Doctorado de Medicina e Investigación Traslacional.

Facultad de Medicina y Ciencias de la Salud. Universitat de Barcelona

Junio de 2023

AGRADECIMIENTOS

La presentación de esta tesis doctoral es la culminación de unos años de formación, de mucho aprendizaje y trabajo intenso pero llenos de momentos de satisfacción, que no hubieran sido lo mismo sin las personas que me rodean.

La casualidad, si es que existe, me llevó a formar parte del grupo de Farmacogenómica y genética neurovascular. He tenido la gran suerte de tener como directores de tesis a Israel y a Jurek, con quienes nunca tendré suficientes palabras de agradecimiento por todo lo que me han enseñado y las oportunidades que me han dado. GRACIAS, de verdad, por todos estos años.

También quería agradecer a todos los compañeros de la FMT, mencionando especialmente a MJ, por haberme enseñado mucho durante estos años y por su soporte. También por el tiempo dedicado a hablar sobre baloncesto y más aún, por haber hecho posible que volviera a jugar!

Gracias a los compañeros de Sant Pau porque, aunque los años de tesis son “por definición” años intensos y duros, estos han sido mucho más llevaderos gracias al soporte de todos ellos. A pesar de estar distanciados un tiempo por la pandemia, os he sentido siempre cerca. ¡GRACIAS!

Para finalizar, pero no menos importante, quería dar millones de gracias a mi familia. A mis padres, por ser la guía durante todos mis años de formación, por ser mi soporte y porque he llegado donde estoy gracias a ellos. ¡No puedo ser más afortunada de tenerlos como padres! También a mi marido. Nunca podré estar suficiente agradecida al destino que hizo que, gracias a formar parte de este grupo, pudiera conocer al amor de mi vida. Sin ese congreso europeo de ictus, nada sería lo mismo. Gracias por todo el soporte durante estos años y por todo el amor. También a Sofía, que, sin ese congreso, tampoco existiría. Escribir la tesis doctoral contigo en mi barriga es lo más especial que podría haber ocurrido y, culminarla contigo a mi lado, no tiene palabras, ¡Te quiero infinito, princesa!

FINANCIACIÓN

Los estudios de esta tesis doctoral han estado financiados a través de diferentes proyectos concedidos en convocatorias públicas y privadas. A continuación, se detallan los diferentes proyectos:

Instituto de Salud Carlos III:

- Proyecto EPIGENESIS: PI17/02089,

- Proyecto MAESTRO: PI18/01338

- Proyecto EPINEXO: PI20/00678

Marató TV3:

- Proyecto EPIGENESIS

Fundació MútuaTerrassa:

- Proyecto EPIGENESIS

Eranet-Neuron, European research grants:

- Proyecto iBioStroke

Boehringer Ingelheim:

- Proyecto SEDMAN

Pfizer/ Bristol-Myers Squibb:

- Proyecto APHAS

Fondos de desarrollo europeo regional (FEDER)

AGAUR:

- 2017SGR-1427

ÍNDICE

AGRADECIMIENTOS	I
FINANCIACIÓN	III
ÍNDICE	V
ABREVIATURAS Y ACRÓNIMOS	17
PUBLICACIONES	19
RESUMEN.....	21
INTRODUCCIÓN	25
a. El ictus	25
i. Descripción	25
ii. Epidemiología	25
iii. Clasificación	29
1. Ictus isquémico	29
2. Ictus hemorrágico	30
iv. Factores de riesgo en el ictus	32
1. Factores de riesgo modificables	32
2. Factores de riesgo no modificables	35
iv. Pronóstico.....	36
1. Escalas de evaluación del pronóstico post-ictus	37
2. Correlación entre las escalas de evaluación del pronóstico post-ictus.....	38
3. Factores demográficos y clínicos asociados con el pronóstico después de un ictus.....	39
b. Genética	40
i. Conceptos básicos	40
ii. Bases de la genómica moderna: Proyecto Genoma Humano y ENCODE	42
iii. Estudios de asociación mediante rastreo genómico o <i>Genome-wide Association Studies (GWAS)</i>	44
1. Conceptos básicos	44
2. Origen de los GWAS.....	45
3. Logros y carencias de los GWAS	46
c. Estudios GWAS en el ictus	47

d.	
i.	Estudios GWAS y riesgo de ictus48
1.	GWAS y riesgo de ictus isquémico48
2.	GWAS y riesgo de ictus hemorrágico58
ii.	Estudios GWAS y pronóstico del ictus isquémico59
1.	GWAS y endofenotipos de la fase aguda60
2.	GWAS y la evolución del ictus a largo plazo62
e.	Epigenética63
i.	Conceptos básicos63
ii.	Mecanismos epigenéticos63
iii.	Metilación del ADN66
1.	Descripción66
2.	Herramientas para el estudio de la metilación del ADN68
a.	Estudios de genes candidatos.....70
b.	Estudios epigenómicos73
i.	<i>Epigenome-wide association studies (EWAS)</i>73
f.	Estudios epigenómicos en el ictus77
i.	Estudios de ARNnc y de modificaciones de histonas en el ictus77
1.	Estudios en relación con el riesgo de ictus isquémico77
a.	<i>ARN no codificante (ARNnc)</i>77
b.	<i>Modificación de histonas</i>82
2.	Estudios en relación con el ictus hemorrágico83
a.	<i>ARN no codificante (ARNnc)</i>83
b.	<i>Modificación de histonas</i>84
3.	Estudios en relación con el pronóstico del ictus85
a.	<i>ARN no codificante (ARNnc)</i>85
b.	<i>Modificación de histonas</i>88
i.	Estudios EWAS de metilación del ADN en pacientes con ictus90
1.	Estudios en relación con el riesgo de ictus.....90
2.	Estudios en relación con la evolución post-ictus.....95
g.	Interpretación y utilidad de los resultados genéticos y epigenéticos mediante herramientas bioinformáticas97
i.	Repositorios y bases de

datos.....	97
ii. Estudios de causalidad: Randomización Mendeliana	99
iii. Estudio de correlación genética	102
iv. Estudio de vías metabólicas	104
v. Estudios de asociación de transcriptoma y proteoma de genoma completo (<i>Transcriptome-Wide Association Studies</i> y <i>Proteome-Wide Association Studies</i>).	106
vi. Reposicionamiento farmacológico.....	108
 HIPÓTESIS	113
OBJETIVOS	117
MATERIAL, MÉTODOS y RESULTADOS (ARTÍCULOS)	121
DISCUSIÓN	313
CONCLUSIONES.....	327
BIBLIOGRAFÍA	331

ABREVIATURAS Y ACRÓNIMOS

ABREVIACIÓN	TÉRMINO
450K	Infinium HumanMethylation450
5MC	5-metilcitosina
ADN	Ácido desoxirribonucleico
AHA	The Stroke Council of the American Heart Association
AHSA	Ácido hidroxámico suberoilánilida
AIMS	Amplificación de puntos inter-metilados
AIT	Accidente isquémico transitorio
ARNI	ARN de interferencia
ARNLINC	ARNnc intergénico largo
ARNM	ARN mensajero
ARNNC	ARN no codificante
ARNNCL	ARNnc largos
ARNR	ARN ribosomal
ARNSNO	ARN nucleolar
ARNT	ARN de transferencia
ASA	American Stroke Association
AVAD/DALY	Años de vida ajustados por discapacidad
BSPP	Bisulphite padlock probes
CHARGE	Cohorts of Heart and Aging Research in Genomic Epidemiology
COMPASS	Consortium of Minority Population genome-wide Association Studies of Stroke
CPG	Dinucleótidos de citosinas y guaninas
CTMP	Carboxy-Terminal Modulator Protein
DM	Diabetes mellitus
DMH	Hibridación de Metilación Diferencial
DMR	Regiones diferencialmente metiladas
DNMTs	ADN metiltransferasas
DREAM	Digital Restriction Enzyme Analysis of Methylation
ENCODE	ENCyclopedia of DNA Elements
EQTL	Expression quantitative trait loci
ESUS	Embolic Stroke Undetermined Source
EWAS	Epigenome-wide association studies
FDR	False-Discovery Rate
GBD	Global Burden of Diseases, Injuries, and Risk Factors
GENISIS	Genetics of Early Neurological Instability after Ischemic Stroke
GISCOME	Genetics of Ischaemic Stroke Functional Outcome
GNOVA	GeNetic cOVariance Analyzer
GODS	Genetic contribution to functional Outcome and Disability after Stroke
GREML	Genomic restricted maximum likelihood
GSEA	Gene Set Enrichment Analysis
GWAS	Genome-Wide association studies
GWPA	Genome-wide Pathway Analysis
HATS	Acetiltransferasas de histonas

HDACS	Deacetilasas de histonas
HIC	Ictus causado por hemorragia intracerebral
HKMTS	Lisinas metiltransferasas de histonas
HP	Hematoma parenquimatoso
HR	Hazard Ratio
ICTUS	International Citicoline Trial on Acute Stroke
ISGC	International Stroke Genetics Consortium o Consorcio Internacional de Genética del Ictus
ISLET CELL AUTOANTIGEN 1 LIKE	ICA1L
L1000 CONNECTIVITY MAP	LINCS
LCR	Líquido cefalorraquídeo
LDSC	Linkage disequilibrium score regression
LMM	Linear mixed model
LSD-1	Demetilasa 1
MCAM	Amplificación de Islas CpG Metiladas
MIARN	Micro ARN
MQTL	DNA methylation quantitative trait loci
MRS	Modified Rankin Scale
MS-AP-PCR	PCR arbitrariamente sensible a la metilación
MTAG	Multi-Trait Analysis of GWAS
MVP	Posición variable de metilación
NAT	Transcritos naturales antisentidos
NIHSS	National Institutes of Health Stroke
NINDS	U.S. National Institute of Neurologic disorders and Stroke
OCSP	Oxfordshire Community Stroke Project
ORA	Over-representation Analysis
PCR	Proteína C reactiva
PIARN	ARN que interacciona con Piwi
PRMTS	Argininas metiltransferasas
PWAS	Proteome-wide association studies
REML	Restricted aximum likelihood
RLGS	Restriction Landmark Genoma Scanning
RM	Randomización Mendeliana
SIGN	Stroke Genetics Network
SIRT	Sertuinas
SMR	Summary data-based Mendelian randomization
SNC	Sistema nervioso central
SNV O SNP	Single nucleotide variant or polymorphism
TET	Ten-eleven translocation
TOAST	Trial of Org 10172 in Acute Stroke Treatment
TRANS-PHAR	Integration of Transcriptome-wide association study and Pharmacological database
TSARN	Derivados de ARN de transferencia
TWAS	Transcriptome-wide association studies
VISP	Vitamin Intervention for Stroke Prevention
WGSBS	Whole-genome shotgun bisulphite sequencing
WHO-ICF	World Health Organization International Classification of Function
WTCCC2	Consortio de Control de Casos Wellcome Trust o Wellcome Trust Case Control Consortium 2
YLD	Años vividos con discapacidad
YLL	Años de vida perdida

PUBLICACIONES

Tesis en formato de compendio de artículos.

La tesis consta de cuatro objetivos principales y cuatro artículos:

Artículo 1:

Autores: **Natalia Cullell**, Carolina Soriano-Tárraga, Cristina Gallego-Fábrega, Jara Cárcel-Márquez, Nuria P Torres-Águila, Elena Muiño, Miquel Lledós, Laia Lluçà-Carol, Manel Esteller, Manuel Castro de Moura, Joan Montaner, Alba Fernández-Sanlés, Roberto Elosua, Pilar Delgado, Joan Martí-Fábregas, Jerzy Krupinski, Jaume Roquer, Jordi Jiménez-Conde, Israel Fernández-Cadenas.

Título: DNA Methylation and Ischemic Stroke Risk: An Epigenome-Wide Association Study.

Revista: Thrombosis and Haemostasis

Año: 2022. Volumen: 122(10):1767-1778

Factor de impacto: 6.830. Q1 (Categoría/Especialidad: HEMATOLOGY Y PERIPHERAL VASCULAR DISEASES)

Artículo 2

Autores: **Natalia Cullell**, Carolina Soriano-Tárraga, Cristina Gallego-Fábrega, Jara Cárcel-Márquez, Elena Muiño, Laia Lluçà-Carol, Miquel Lledós, Manel Esteller, Manuel Castro de Moura, Joan Montaner, Anna Rosell, Pilar Delgado, Joan Martí-Fábregas, Jerzy Krupinski, Jaume Roquer, Jordi Jiménez-Conde, Israel Fernández-Cadenas.

Título: Altered methylation pattern in EXOC4 is associated with stroke outcome: an Epigenome-Wide Association Study.

Revista: Clinical Epigenetics.

Año: 2022;14(1):124

Factor de impacto: 7.280. Q1 (Categoría/Especialidad: GENETICS & HEREDITY Y ONCOLOGY)

Artículo 3:

Autores: **Natalia Cullell**, Cristina Gallego-Fábrega, Jara Cárcel-Márquez, Elena Muiño, Laia Lluçia-Carol, Miquel Lledós, Jesús M Martín-Campos, Jessica Molina, Laura Casas, Marta Almeria, Israel Fernández-Cadenas, Jerzy Krupinski.

Título: ICA1L Is Associated with Small Vessel Disease: A Proteome-Wide Association Study in Small Vessel Stroke and Intracerebral Haemorrhage.

Revista: International Journal of Molecular Sciences.

Año: 2022. 15;23(6):3161

Factor de impacto: 6.208. Q1 (Categoría/Especialidad: BIOCHEMISTRY & MOLECULAR BIOLOGY)

Artículo 4:

Autores: **Natalia Cullell**, Cristina Gallego-Fábrega, Jara Cárcel-Márquez, Elena Muiño, Laia Lluçia-Carol, Miquel Lledós, Jesús M Martín-Campos, Israel Fernández-Cadenas, Jerzy Krupinski.

Título: Identification of genetic regulation at the proteomic level in the long-term stroke outcome: A Proteome-Wide Association study.

Revista: No publicado aún.

RESUMEN

Título: Estudio epigenómico de los factores de riesgo y pronóstico en el ictus isquémico

Introducción: El ictus es una de las principales causas de muerte a nivel mundial, siendo la primera causa para las mujeres en España. También se asocia a discapacidad, siendo su principal causa en personas adultas en España. A nivel genético, el ictus es una enfermedad compleja, en la que múltiples genes participan en su riesgo y pronóstico. Además, es una enfermedad heterogénea ya que participan diferentes mecanismos fisiopatológicos, razón por la cual la enfermedad presenta diferentes subtipos. Actualmente, mediante la aplicación de nuevas tecnologías *ómicas*, que analizan de forma agnóstica todo el genoma, ha aumentado el conocimiento sobre los factores implicados en el riesgo y pronóstico del ictus. Esta información tiene una implicación directa en la generación de medidas para mejorar tanto la prevención como su tratamiento.

Entre las diferentes tecnologías *ómicas* aplicadas en el ictus, los estudios de *Genome-wide association study* (GWAS) han identificado hasta 89 loci independientes asociados al riesgo de ictus y nueve loci asociados al pronóstico después del ictus. Por otro lado, estudios epigenómicos de metilación del ADN (*Epigenome-wide association study*; EWAS) también han permitido la identificación de 24 loci asociados al riesgo de aparición de esta patología. Sin embargo, no se han realizado estudios epigenómicos analizando el pronóstico neurológico después de un ictus.

Hipótesis y objetivos: Las técnicas *ómicas* que analizan la metilación del ADN junto con la integración de datos genómicos y proteómicos puede permitir conocer los mecanismos y las vías metabólicas por los que se produce el ictus y los mecanismos que expliquen la variabilidad en la evolución neurológica a corto/medio plazo, además del pronóstico funcional a largo plazo. Los objetivos planteados en la presente tesis fueron el de identificar nuevos genes cuya metilación del ADN se asociara al riesgo de ictus y al pronóstico neurológico a corto/medio plazo, así como estudiar su relación de causalidad. También estudiar las proteínas reguladas genéticamente asociadas al riesgo de ictus lacunar y al pronóstico funcional a largo plazo mediante un análisis

bioinformático de integración de diferentes aproximaciones *ómicas* con el objetivo de priorizar genes identificados mediante GWAS y encontrar nuevas proteínas que puedan ser utilizadas en el futuro como biomarcadores y/o dianas terapéuticas.

Métodos: Para la identificación de genes con metilación diferencial asociada al riesgo de ictus y a la evolución neurológica, se utilizó una aproximación mediante *Epigenome-Wide Association study* (EWAS). Para la interpretación de los resultados se realizaron análisis adicionales de randomización mendeliana (RM), de metilación específica por subtipos celulares y de vías metabólicas. Para poder integrar datos genómicos con datos proteómicos de tejido cerebral, se realizó un análisis bioinformático de *Proteome-wide association study* (PWAS). Los resultados del PWAS se validaron mediante estudios de colocalización y RM.

Principales resultados: Como resultado de esta investigación, se identificaron un total de 821 dinucleótidos citosina/guanina (CpGs) con una metilación diferencial en tejido sanguíneo (hipometilación en el 99,7% de las CpG) asociada al riesgo de ictus. Las principales modificaciones en la metilación del ADN se identificaron en CpGs localizadas en los genes *ZFHX3* y *MAP3K1*. Según el estudio de metilación diferencial por subtipos celulares en sangre, las células Natural Killer (NK) fueron las principales mediadoras de los resultados de metilación. El estudio de causalidad mediante RM para los genes *ZFHX3* y *MAP3K1*, mostró que la metilación era causa de ictus cardioembólico y aterotrombótico para *MAP3K1* y también ictus lacunar para *ZFHX3*. En el estudio EWAS de pronóstico neurológico, cuantificado mediante la diferencia entre la NIHSS basal y la NIHSS al alta (Δ NIHSS), la CpG cg00039070 en el gen *EXOC4* presentó una mayor metilación asociada con un peor pronóstico neurológico. La metilación de *EXOC4* se asoció a cambios a nivel proteico (estudiado con el *array* SOMAscan), identificándose un enriquecimiento de proteínas pertenecientes a la vía de la activación de las células NK.

La integración de datos genómicos y proteómicos de cerebro mediante PWAS, permitió priorizar ICA1L como posible biomarcador y diana terapéutica en el ictus de tipo lacunar así como en los ictus hemorrágicos no lobares, dos fenotipos con mecanismo fisiopatológico común. En concreto, se encontró que la regulación genética de ICA1L promovía el aumento de los niveles de la proteína que conllevaban a una disminución

del riesgo de ambos tipos de ictus. El estudio PWAS de pronóstico funcional post-ictus a largo plazo permitió identificar las proteínas THEM4 y APOL2 cuyos niveles, regulados genéticamente, se asociaron a un peor (THEM4) o mejor pronóstico (APOL2) a los tres meses medido con la escala modificada de Rankin (mRS).

Conclusiones: En esta tesis se ha realizado la integración de diferentes datos *ómicos*. Este análisis ha permitido identificar genes regulados epigenéticamente y proteínas cuya expresión es controlada genéticamente, asociados al riesgo y pronóstico del ictus. Debido al carácter modificable de la regulación epigenética y proteómica, tanto los genes regulados mediante metilación de ADN como las proteínas identificadas en los estudios de esta tesis, podrían ser buenas dianas terapéuticas para prevenir y mejorar el pronóstico de los pacientes después de un ictus.



INTRODUCCIÓN

INTRODUCCIÓN

a. El ictus

i. Descripción

La palabra “ictus” es ampliamente utilizada en el ámbito de la salud. Sin embargo no fue hasta 2013 que se hizo pública una definición consensuada por parte de un grupo de trabajo organizado por el Consejo de Ictus de la Asociación Estadounidense del Corazón (*The Stroke Council of the American Heart Association (AHA)*) y la Asociación Americana de Ictus (*American Stroke Association (ASA)*) (1). El ictus se considera un déficit neurológico causado por un daño focal local en el sistema nervioso central (SNC) de origen vascular. Dentro del ictus se incluye el infarto o isquemia cerebral, la hemorragia intracerebral (HIC) y la hemorragia subaracnoidea (1).

ii. Epidemiología

El ictus es una de las principales causas de muerte y discapacidad a nivel mundial (2), lo que supone un gran impacto sanitario y social. Los principales parámetros epidemiológicos como la incidencia, prevalencia, mortalidad, años de vida perdida (YLL), años vividos con discapacidad (YLD) y los años de vida ajustados por discapacidad (AVAD o DALY, *Disability Adjusted Life Years*, en inglés) varían según el territorio. En concreto, la prevalencia más elevada de ictus estandarizado por edad se encuentra en Oceanía, Sureste de Asia, Norte de África, Oriente Medio y Asia Oriental (3).

En 2020 se publicaron los resultados más recientes disponibles acerca de la epidemiología global del ictus. Estos resultados fueron obtenidos como resultado del estudio *Global Burden of Diseases, Injuries, and Risk Factors (GBD) 2019* en el que se analizaron datos comprendidos entre el 1990 y el 2019 de un total de 204 países y 369 enfermedades, incluyendo el ictus (4). En 2019, la prevalencia global de ictus fue de 101,5 millones de personas (77,2 millones de ictus isquémicos, 20,7 millones de HIC y 8,4 millones de hemorragias subaracnoideas). En cuanto al ictus isquémico, aunque hubo una disminución del 1,9% en la prevalencia estandarizada por edad entre el 1990 y el 2019, se observó un incremento del 3,6% de entre los años 2010 y 2019 (3).

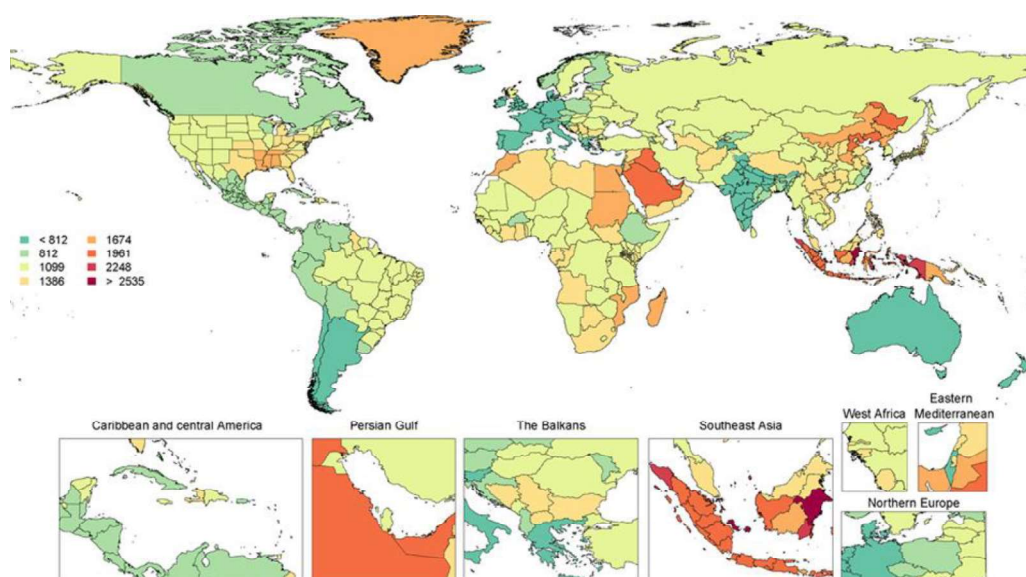


Figura 1: Mapa de la prevalencia estandarizada por la edad del ictus en el mundo

Prevalencia por 100 000 habitantes en personas de ambos sexos.

Fuente: Datos derivados del Estudio GBD2019. Imagen extraída de referencia (3).

En la Unión Europea, el ictus es la segunda causa de mortalidad y la primera de discapacidad en el adulto. Afecta aproximadamente a 1,1 millón de habitantes anualmente y causa 440,000 defunciones. En el estudio GBD 2017 (5) se realizó una proyección de 30 años de la epidemiología del ictus (5). En la proyección hecha para el año 2047 se estimó que habría un 3% de aumento en la incidencia y un 27% de aumento en la prevalencia del ictus. Por el contrario, se estimó una disminución del 17% en la mortalidad y del 33% en los AVAD.

En cuanto a los datos epidemiológicos del ictus en España, se considera la segunda causa de muerte global (la tercera en hombres y la primera en mujeres). En 2019 se publicó el Atlas del Ictus con datos obtenidos a través de fuentes primarias de encuestas y registros oficiales con el objetivo de describir la epidemiología del ictus en España. La incidencia en 2018 en población española de más de 18 años fue de 187,4 casos por 100.000 habitantes. La prevalencia estimada fue del 1,7%. Además, Los pacientes con ictus presentaron una baja calidad de vida, dos de cada tres supervivientes a un ictus presentaron algún tipo de secuela (un 62,4% de los pacientes tuvo algún problema de movilidad y un 59,1% algún problema para realizar sus actividades cotidianas) (6).

iii. Clasificación

El ictus es una enfermedad heterogénea para la que se conocen más de 150 posibles causas (7). La clasificación de los ictus en subtipos es muy útil en la práctica clínica, debido al diferente manejo que tiene cada uno de ellos, pero también es necesaria para llevar a cabo estudios epidemiológicos y genéticos, ensayos clínicos y otros tipos de estudios (7). Las dos grandes subentidades en las que se puede clasificar el ictus son el ictus isquémico (aproximadamente el 80% de los ictus) y el ictus hemorrágico (aproximadamente el 20% de los ictus) (8,9).

1. Ictus isquémico

Es el subtipo más común de ictus y se caracteriza por una hipoperfusión cerebral de causa cardioembólica, aterotrombótica (enfermedad de gran vaso), lacunar (enfermedad de pequeño vaso) u otras causas. De manera simple, la proporción de cada una de estas causas se dice que se rige por la regla de los cuartos, es decir, cada una de ellas supone un 25% del total de ictus isquémico. Sin embargo, esta proporción es muy variable y depende de cada población (10).

Se han desarrollado diferentes clasificaciones que permiten diferenciar los diferentes subtipos de ictus isquémico. Las dos más ampliamente utilizadas son la *Trial of Org 10172 in Acute Stroke Treatment* (TOAST) y la *Oxfordshire Community Stroke Project* (OCSP) (11). El criterio TOAST permite identificar el mecanismo fisiopatológico que de forma más probable ha causado el ictus, utilizando los hallazgos clínicos y estudios diagnósticos auxiliares para su clasificación (12). Es el criterio más ampliamente utilizado y permite clasificar el ictus en 5 categorías diferentes: aterosclerosis de grandes arterias (por causa embólica o trombosis), cardioembolismo, oclusión de pequeño vaso (lacunar), ictus de otra causa determinada (diseccción, vasculitis, enfermedades genéticas y otras), ictus de causa indeterminada o criptogénico (dos o más causas identificadas, evaluación negativa o evaluación incompleta) (9,11). En el caso de los ictus criptogénicos, se pueden dividir en ictus con probable causa embólica (ESUS) o sin causa embólica (No ESUS) (13) (Figura 2).

Por el contrario, la clasificación OCSP se basa exclusivamente en los hallazgos clínicos, principalmente utilizando técnicas de neuroimagen, para clasificar a los pacientes en

base a la extensión y región cerebral en la que se ha producido el infarto. De esta manera, clasifica el ictus isquémico en infarto lacunar (LACI), infarto total en la circulación anterior (TACI), infarto parcial en la circulación anterior (PACI) e infarto en la circulación posterior (POCI) (7) .

2. Ictus hemorrágico

Este subtipo de ictus se produce por la ruptura de un vaso sanguíneo en el cerebro (14) y está asociado a una gran mortalidad (un 80% en países desarrollados). Los dos principales tipos de hemorragias cerebrales son la hemorragia intraparenquimatosa o HIC (alrededor del 15% de los ictus) y la hemorragia subaracnoidea (el 5% de todos los ictus) (Figura 2).

Los principales mecanismos causantes del ictus hemorrágico son la hipertensión, la angiopatía amiloide cerebral, el uso de fármacos anticoagulantes orales y las malformaciones arteriovenosas (11). La hemorragia subaracnoidea se puede atribuir a un aneurisma en el 85% de los casos (15).

La angiopatía amiloide cerebral se caracteriza por la deposición del péptido β -amiloide, el cual provoca cambios degenerativos en capilares, arteriolas y arterias. Esto produce la ruptura esporádica de estos vasos en personas de edad avanzada o en personas jóvenes, debido a mutaciones en el gen codificante de la proteína precursora de amiloide (14).

La HIC causada por el uso de anticoagulantes orales afecta en mayor grado a pacientes con vasculopatías relacionadas con la hipertensión o con angiopatía amiloide cerebral (14).

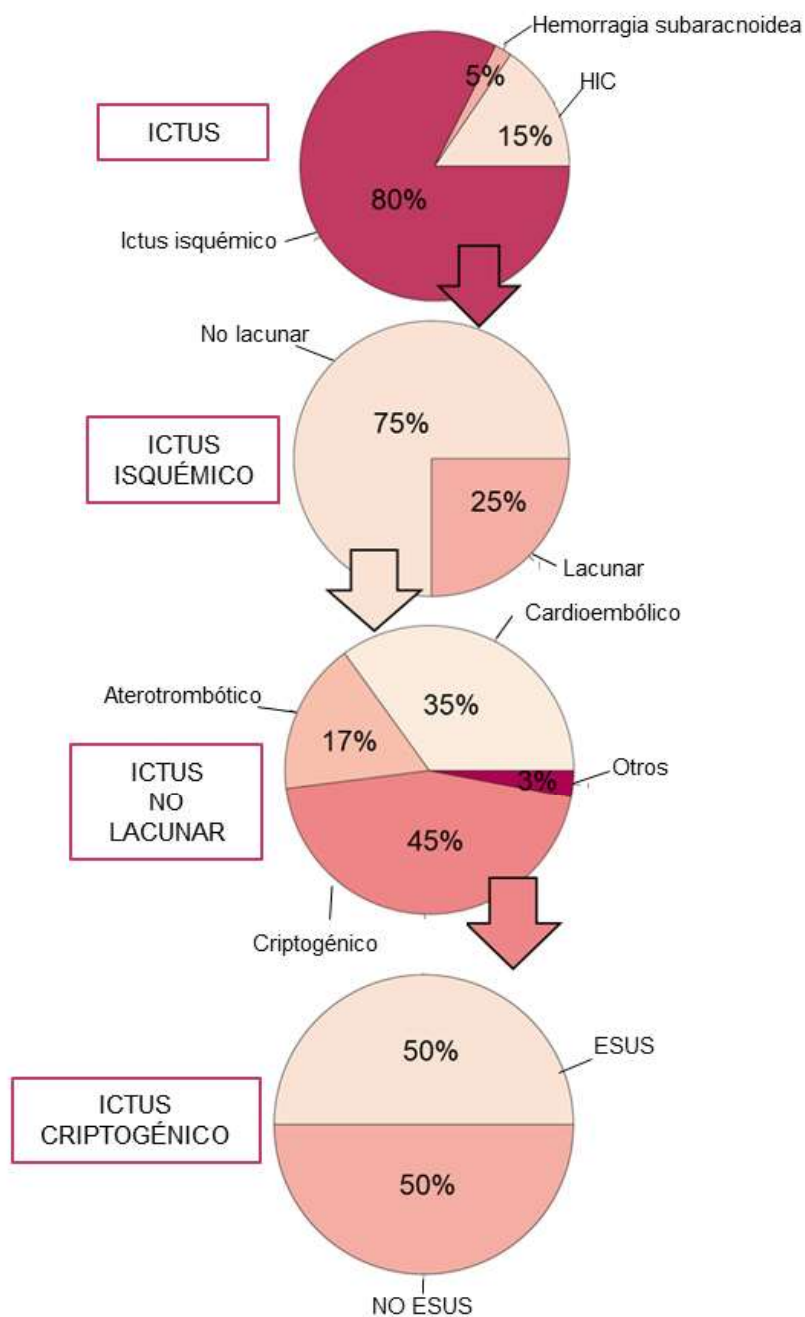


Figura 2: Clasificación de los ictus según *Trial of Org 10172 in Acute Stroke Treatment (TOAST)*. Los porcentajes son aproximados según se han descrito en (16,17)

ESUS: Ictus embólico de causa indeterminada (Embolic stroke of undetermined cause)

Figura original de la doctoranda basada en el estudio (13).

iv. Factores de riesgo en el ictus

Se han descrito diferentes factores de riesgo para el ictus. Sin embargo, debido a que es una enfermedad muy heterogénea, los factores de riesgo también dependerán del mecanismo por el cual se produce el ictus. Algunos de los factores de riesgo para el ictus son modificables mientras que otros no lo son. La identificación de los factores de riesgo modificables es importante para reducir la carga del ictus. Además, los factores de riesgo también se pueden clasificar según si se producen a corto plazo (o inmediatos), los cuales son desencadenantes de un ictus, o si se producen a medio o a largo plazo. También se pueden diferenciar los factores de riesgo para ictus en personas jóvenes y en personas de edad avanzada (9).

La combinación de los diferentes factores de riesgo puede ser útil para predecir el riesgo individual de tener un ictus. De esta manera se puede reducir el riesgo a través de la modificación de aquellos factores modificables (9).

Actualmente existen algunas herramientas para calcular el riesgo de aparición de un ictus. Por ejemplo, el Perfil de riesgo para ictus del estudio Framingham (*Framingham Stroke Risk Profile*), el cual puntúa el riesgo de ictus en los próximos 10 años estratificado por sexo y combinando diferentes predictores (18).

En un estudio publicado en 2010 llamado INTERSTROKE (19) se identificó un total de 10 factores de riesgo modificables capaces de explicar el 90% del riesgo de ictus. Este estudio demostró la relevancia de hacer una prevención primaria a través de la reducción de factores de riesgo modificables.

1. Factores de riesgo modificables

Los factores de riesgo modificables se pueden clasificar según si son condiciones médicas o factores conductuales.

Los principales factores de riesgo modificables son la hipertensión, diabetes mellitus (DM), fibrilación auricular (FA) y cardiopatía atrial, dislipidemia, el sedentarismo, la nutrición, obesidad, síndrome metabólico, consumo de alcohol, drogas, tabaco, inflamación e infecciones. En el estudio INTERSTROKE, el primer gran estudio

internacional diseñado para evaluar la relevancia de los factores de riesgo en el ictus, los siguientes factores se asociaron al riesgo de ictus: la hipertensión, el hábito de fumar, la dieta, la obesidad (medido con la ratio entre la cintura y la cadera), la actividad física regular, la DM, el consumo de alcohol, el estrés, la enfermedad cardíaca y la dislipidemia (la ratio entre la apolipoproteína B y A) (9).

La hipertensión es el factor de riesgo que tiene un efecto mayor, lineal y directo con el riesgo de ictus. Los valores de presión sanguínea tienen una asociación directamente proporcional con el riesgo de ictus, incluso cuando los valores se encuentran dentro de la normalidad. Debido a que es un factor de riesgo modificable mediante el cambio en el estilo de vida o con el uso de fármacos, la reducción de la presión arterial es una de las estrategias más efectivas para reducir el riesgo de ictus (9). Además, estudios de causalidad utilizando datos genéticos (estudios de Randomización Mendeliana (RM)), han demostrado que el aumento de la presión arterial es un factor causal en la aparición de un ictus (20).

En cuanto a la DM, se considera un factor de riesgo causal independiente para el ictus (20). Un 20% de las muertes provocadas por ictus se producen en personas diabéticas. Además, el número de años con diabetes también se asocia de manera directa con el riesgo de ictus (9).

La FA una patología cardíaca prevalente en personas de edad avanzada. Esto hace que, con el envejecimiento de la población, esta enfermedad esté más presente y aumente el número de ictus producidos por ella. En concreto, esta enfermedad se asocia de forma causal con el riesgo de ictus cardioembólico (21), producido por el émbolo de un trombo del corazón a un vaso cerebral (9).

La relación de la dislipidemia con el riesgo de ictus es más compleja. El aumento del colesterol unido a lipoproteínas de baja densidad (LDL), aumenta el riesgo de ictus isquémico, aunque disminuye el riesgo de ictus hemorrágico, mientras que el aumento del colesterol unido a lipoproteínas de alta densidad (HDL) disminuye el riesgo de ictus isquémico pero aumenta el riesgo de ictus hemorrágico (20). El papel de los triglicéridos en el riesgo de ictus es variable en función del tipo de ictus. Se ha encontrado una asociación causal entre los niveles de triglicéridos y el riesgo global de ictus isquémico y

específicamente de ictus aterotrombótico, pero no de otros subtipos de ictus como el ictus cardioembólico (20). Además se ha observado que los niveles elevados de triglicéridos junto con niveles elevados de colesterol se asocian a un mayor riesgo de ictus aterotrombótico (9).

Los hábitos de vida se asocian también al riesgo de ictus. Realizar actividad física de forma mantenida en el tiempo reduce el riesgo de sufrir muchas enfermedades, incluido el ictus. Esta asociación puede estar causada por la disminución de algunos factores de riesgo como la presión arterial o la DM en las personas que tienen una vida activa. De igual manera, la dieta se asocia al riesgo de ictus, probablemente por la influencia que tiene sobre otros factores de riesgo. En general, la dieta mediterránea, con un alto consumo de fruta y verdura, se asocia con un menor riesgo de ictus. Relacionado con la actividad física y la dieta, se encuentra la obesidad, asociada a un aumento de riesgo de ictus. En concreto, el incremento de la adiposidad abdominal (medido con la ratio cadera-cintura) es el factor que se asocia principalmente al riesgo de ictus. De hecho, se ha encontrado una asociación causal mediante RM del ratio cadera-cintura con los diferentes subtipos de ictus, excepto el de tipo cardioembólico (20). Finalmente, el concepto síndrome metabólico que une la obesidad, dislipidemia, prehipertensión y prediabetes se asocia con el riesgo de ictus (9).

En muchos estudios se ha descrito una asociación en forma de U o de J entre el consumo de alcohol y el riesgo de ictus. Sin embargo, se trata de estudios observacionales. En cambio, estudios basados en MR han identificado una relación lineal causal entre el consumo de alcohol y el riesgo de ictus (20,22,23). Además, la asociación del alcohol también se relaciona con el subtipo de ictus. En concreto, la asociación es más directa con el ictus de tipo hemorrágico (9).

El abuso de drogas como la cocaína, heroína, anfetaminas y éxtasis también aumenta el riesgo de ictus, tanto isquémico como hemorrágico. El tabaco dobla el riesgo de ictus de una manera dependiente de la dosis y además esta asociación es causal para los diferentes subtipos de ictus, excepto el ictus cardioembólico y el ictus hemorrágico (20). Algunos estudios han observado que el tabaco contribuye al 15% de las muertes de ictus anuales (9).

Finalmente, diferentes marcadores de inflamación e infección se han encontrado asociados con la aparición de ictus. Una posible explicación es que los diferentes marcadores de inflamación estén indicando la presencia de aterosclerosis (enfermedad en la que se forman placas de ateroma en las paredes de las arterias con un gran componente inflamatorio), que se asocia al riesgo de ictus aterotrombótico. Por ejemplo, se ha estudiado ampliamente la relación de los niveles de la proteína C reactiva (PCR), un marcador de inflamación, con el riesgo de ictus. Sin embargo, estudios de causalidad con análisis de RM, no han demostrado que los niveles de PCR se asocien causalmente con el riesgo de ictus. Sin embargo, otros marcadores plasmáticos de inflamación, como los niveles de IL6 o de MCP-1 sí que se han asociado de forma causal en estudios de RM con un mayor riesgo de ictus (20). Las infecciones agudas pueden desencadenar a corto plazo un ictus. Algunos estudios han demostrado que el riesgo de ictus aumenta después de una hospitalización debido a una infección y disminuye cuando aumenta el tiempo transcurrido tras la hospitalización (24).

2. Factores de riesgo no modificables

Los factores de riesgo no modificables incluyen la edad, el sexo, la etnia y la genética. En cuanto a la edad, esta tiene una relación directamente proporcional con la incidencia de ictus, siendo el doble por cada década que pasa a partir de los 55 años. Para el riesgo de ictus hemorrágico, este incrementa a partir de los 45 años.

La influencia que el sexo tiene en el riesgo de ictus está condicionada y varía en función de la edad. Globalmente, la incidencia de ictus es mayor en mujeres, debido a su mayor esperanza de vida. Sin embargo, en edades avanzadas, los hombres tienen un ligero mayor riesgo de sufrir un ictus que las mujeres. Esta tendencia es diferente en personas jóvenes. En ese momento, las mujeres tienen un riesgo mayor de sufrir un ictus. Este riesgo está asociado a factores hormonales (embarazo, post-parto, uso de anticonceptivos, etc.) (9).

En cuanto a la etnia, también se encuentran diferencias de riesgo, algunas específicas de subtipos concretos de ictus. Las personas de raza negra y los hispanos americanos tienen un riesgo mayor. Especialmente, los adultos de raza negra tienen un riesgo mucho más elevado de padecer una hemorragia subaracnoidea o HIC (9).

La genética también se ha identificado como factor de riesgo asociado al ictus. Sin embargo, su estudio es complicado debido a la presencia de diferentes subtipos de ictus y a la variabilidad entre poblaciones. Diferentes estudios han encontrado que la variabilidad genética puede contribuir en el riesgo de ictus a través de diferentes mecanismos. Por un lado, existen algunos síndromes causados por la mutación de un gen (enfermedades monogénicas) que tienen como principal manifestación el ictus. Por otro lado, hay enfermedades monogénicas que una de sus manifestaciones puede ser el ictus. Además, variaciones genéticas comunes (polimorfismos) en algunos genes se asocian con el riesgo de ictus. En este caso, cada uno de los polimorfismos aporta un riesgo bajo o moderado al riesgo de ictus. Sin embargo, la combinación de varios de ellos puede aumentar de forma considerable el riesgo. Finalmente, mutaciones asociadas a patologías que son factor de riesgo para ictus también contribuyen al aumento en el riesgo de ictus. Aunque la genética se considera un factor no modificable, los efectos que causan las mutaciones genéticas pueden ser tratables. Además, la terapia génica es actualmente una realidad y es aplicable a algunas enfermedades de causa monogénica (9).

v. Pronóstico

El ictus es una de las principales causas de discapacidad a nivel mundial. La severidad y la evolución de los síntomas clínicos en la fase aguda del ictus son potenciales predictores del pronóstico a largo plazo (25). Por eso, la correcta evaluación de la severidad de un ictus durante la fase aguda es crucial para predecir cuál será el pronóstico de los pacientes (26).

La severidad cambia durante las primeras horas y días después de un ictus. En las primeras horas, los déficits son inestables y pueden presentar cambios repentinos, tanto de mejoría (debido a la recanalización del territorio vascular, por ejemplo) o empeoramiento (debido a la propagación del coágulo, por ejemplo). Pasadas las 6-10 primeras horas, la lesión está asentada y no se producen cambios tan bruscos. En el periodo subagudo, entre las 12 horas y los siete días después del ictus, la mayoría de los pacientes sin complicaciones, presentan mejoría de sus déficits (26).

1. Escalas de evaluación del pronóstico post-ictus

El estado funcional se puede categorizar según la World Health Organization International Classification of Function (WHO-ICF) en términos de empeoramiento, actividad, participación y calidad de vida (27).

Hay diferentes escalas diseñadas para evaluar el estado funcional después de un ictus en cada uno de los niveles de la WHO-ICF (27).

La escala modificada de Rankin o *modified Rankin Scale* (mRS) fue la primera utilizada para evaluar el estado funcional de los pacientes en un ensayo de ictus. Es una de las escalas más utilizadas actualmente y su uso está recomendado por sociedades profesionales y cuerpos regulatorios para su uso en ensayos clínicos. Consiste en una escala ordinal del 0 al 6 para medir la independencia funcional a nivel global con especial énfasis en la función motora (28). Dentro de los niveles de la WHO-ICF se clasifica como una escala de participación ya que es más amplia que las escalas de actividad que se focalizan en actividades de la vida cotidiana. En este caso, la escala mRS permite al clínico tener en cuenta atributos que afectan a la discapacidad ya que alteran el bienestar de la persona siendo independientes del estado físico, como el lenguaje o cognición, funcionamiento social o la depresión (28). El mRS tiene una buena correlación con la localización, el tipo y la extensión de la lesión producida por el ictus (28). Su puntuación a corto y medio plazo se asocia con los cuidados requeridos a largo plazo después del ictus (27).

La escala de la National Institutes of Health Stroke (NIHSS) se diseñó para evaluar las intervenciones en ensayos clínicos. Sin embargo, actualmente su uso está muy extendido en la práctica clínica del ictus agudo. Esta escala se divide en 15 mediciones del empeoramiento neurológico que tienen en cuenta la conciencia, movimiento ocular, visión, coordinación, lenguaje, función sensitiva, fuerza de los miembros superiores e inferiores, función de los músculos faciales y heminegligencia (27). Esta escala permite la evaluación de la severidad del ictus en una fase temprana y permite predecir la evolución de los pacientes a largo plazo (hasta los tres meses post-ictus). El 80% de los pacientes que puntúan bajo en esta escala (< 5) reciben la alta domiciliaria. Aquellos pacientes que tienen un valor entre 6 y 13 en la escala, suelen necesitar rehabilitación.

Finalmente, las puntuaciones superiores a 14 se asocian con dependencia a largo plazo (27). Sin embargo, como la escala NIHSS se basa en el empeoramiento, no es una buena medida de la discapacidad global que resulta del ictus. Para ello, se han diseñado otras escalas enfocadas en otras medidas del estado funcional de los pacientes con ictus (27).

Otra escala que se usa frecuentemente es el Índice Barthel. Esta escala fue diseñada para cuantificar la independencia. Se tienen en cuenta 10 parámetros, con una puntuación total de 100, en los que se considera el grado de asistencia que necesitan los pacientes al realizar actividades cotidianas. Como más elevada es la puntuación en esta escala, mayor independencia presenta el paciente. Una puntuación baja en la escala Barthel correlaciona con discapacidad futura, mayor tiempo de recuperación y requerimiento de mayor cuidado. Sin embargo, la capacidad predictiva de esta escala no es del todo óptima si se realiza durante los primeros cinco días después del ictus (27). Su uso en ensayos clínicos es limitado, especialmente en los que no se espera discapacidad moderada o severa (27).

2. Correlación entre las escalas de evaluación del pronóstico post-ictus

La severidad del déficit neurológico, medido con la escala NIHSS, predice el estado de discapacidad a largo plazo. La NIHSS medido entre 1 y 3 horas después del ictus contabiliza un cuarto de la varianza de la discapacidad a largo plazo (mRS a los 90 días). La NIHSS medido a las 24 horas después del ictus explica el 50% del mRS al tercer mes (26).

Sin embargo, los cambios producidos de forma temprana en la severidad (calculando la diferencia entre la NIHSS basal y la NIHSS a las 24h, Δ NIHSS24) pueden ser mejores predictores de la evolución ya que tienen en cuenta los cambios producidos durante las primeras horas después del ictus, pudiendo revelar los mecanismos implicados en la variabilidad neurológica post-ictus (29). En el estudio de Heitsch et al., la Δ NIHSS24 se asoció de forma significativa e independiente al pronóstico favorable a largo plazo, explicando más del 30% de la varianza del mRS al tercer mes (29).

3. Factores asociados con el pronóstico después de un ictus

Diferentes factores de tipo demográfico, clínico y genético se han asociado al pronóstico temprano y tardío después de un ictus.

En la fase aguda (entre las 24 y 48 horas post-ictus) y subaguda (hasta los 10 días post-ictus), los factores más asociados con el pronóstico son factores demográficos (especialmente la edad y el sexo), algunas comorbilidades (DM, hipertensión y FA), el tratamiento previo con estatinas, variables medidas de forma temprana (en las primeras horas) post-ictus (los niveles de leucocitos, la severidad, el tratamiento recibido para el ictus), complicaciones ocurridas después del ictus como las transformaciones hemorrágicas (TH) o las infecciones y factores genéticos (30,31). Las TH consisten en el trasvase de sangre periférica hacia el cerebro a través de la disrupción de la barrera hematoencefálica (32) y se asocian (en un porcentaje bajo) al tratamiento con rTPA (activador del plasminógeno tisular recombinante), el único tratamiento aprobado para la fase aguda del ictus, el cual cataliza el paso de plasminógeno a plasmina. Otros factores como la dislipidemia o la disfunción renal también se han asociado en algunos estudios con el pronóstico post-ictus. Sin embargo, existen resultados contradictorios al respecto, por lo que serían necesarios más estudios para clarificar su efecto.

Como se ha comentado, la Δ NIHSS24 puede ser mejor predictora del pronóstico después de un ictus. En el estudio de Heitsch, et al., el primer estudio en evaluar factores asociados a la variable dinámica Δ NIHSS24, diferentes variables se asociaron de forma independiente a la Δ NIHSS24: la NIHSS basal, el tratamiento endovascular (rTPA) para el ictus, la edad, los niveles de glucosa y la presión sanguínea. Por lo tanto, es importante destacar que esta es una variable dinámica que se asocia significativamente y de forma independiente al pronóstico a largo plazo (a los 90 meses post-ictus) (29).

Heitsch, et al., también identificó diferentes variables adquiridas y/o cuantificadas durante las primeras horas después de un ictus asociadas a tener una evolución favorable a los 90 días tras el ictus: la edad, la diabetes, la FA y la etnia africana (29).

Con el objetivo de predecir el pronóstico de un ictus, se han diseñado diversas escalas o *scores*. Por ejemplo, el score *ASTRAL* utiliza diferentes variables clínicas asociadas a un pronóstico desfavorable post-ictus ($mRS > 2$) como son la edad, la NIHSS basal, el tiempo

entre el ictus y el ingreso en el hospital, el rango de los campos visuales, la glucosa basal y el nivel de consciencia. La precisión de predicción conseguida calculada con el área bajo la curva (AUC) es de 0.85 (33).

También se han diseñado *scores* para predecir la mortalidad de los pacientes con ictus. Por ejemplo, el *score SOAR (Stroke Subtype, Oxford Community Stroke Project Classification, Age, Prestroke Modified Rankin)*, utilizando datos de edad, sexo, discapacidad previa, etiología y territorio del ictus es capaz de predecir la mortalidad a los 7 días con una AUC de 0.80 (34).

b. Genética

i. Conceptos *básicos*

El concepto de genética se propuso por primera vez en la década de los 1860 a partir de los experimentos de Gregor Mendel. La palabra “genética” se podría definir como el estudio de los genes, aunque, como se verá más adelante, actualmente este concepto es más amplio. Los genes son la región funcional de la molécula llamada ácido desoxirribonucleico (ADN). Los genes se estructuran en exones e intrones y cifran la información para producir una secuencia de aminoácidos que conforman las proteínas. El conjunto de genes determina las propiedades de cada especie. A su vez, cada especie presenta variaciones interindividuales, debidas a las variaciones que pueden presentar los genes, denominados alelos, los cuales determinan cambios a nivel de las proteínas (35).

El genoma es el conjunto de ADN que forma parte de un organismo. El ADN es una estructura lineal en forma de doble hélice formada por dos cadenas entrelazadas y compuesta por nucleótidos (un grupo fosfato, una molécula de azúcar desoxirribosa y una base nitrogenada (adenina, guanina, citosina o timina)). La unión entre las dos cadenas de ADN se produce a través de enlaces entre bases nitrogenadas concretas (complementarias). El ADN se organiza en una estructura conocida como cromatina, la cual se une a un octámero de proteínas llamadas histonas para formar el nucleosoma (36). Posteriormente, el nucleosoma se empaqueta en cromosomas. En el caso de los humanos, el genoma está organizado en dos dotaciones de 23 cromosomas. 22 de los

23 pares de cromosomas se denominan autosomas mientras que el par restante conforma los cromosomas sexuales (37).

El principal producto de los genes son las proteínas, que determinan la estructura y fisiología de los organismos. Para poder generar proteínas, el primer paso que se produce es la transcripción de los genes por parte de la ARN polimerasa para formar una molécula de ácido ribonucleico (ARN). El ARN que contiene las secuencias provenientes de los exones se denomina ARN mensajero (ARNm) y será el producto inicial del segundo paso, denominado traducción. En este proceso, los ribosomas, unidos al ARN de transferencia (ARNt), se unen al ARNm para que se pueda realizar la traducción gracias a su función de transporte de aminoácidos, y catalizan su ensamblaje a partir de la lectura de los nucleótidos para formar la estructura primaria de la proteína. Cuando el ribosoma llega al codón de terminación, este se despega del ARNm para finalizar la traducción (35).

La base de la variación hereditaria se encuentra en la variación genética, que a su vez se produce por la variación alélica de los genes. Se pueden distinguir dos tipos de variaciones genéticas, la discontinua y la continua. La variación discontinua se produce cuando un carácter puede tener dos o más formas (fenotipos) diferentes en una especie. Los fenotipos diferentes suelen estar determinados por alelos diferentes. La constitución alélica de un organismo se denomina genotipo. Entre los diferentes tipos de variantes genéticas, las más frecuentes son la sustitución de un nucleótido que puede conducir a la sustitución de un aminoácido o a la generación de un codón de parada de la traducción. Además, también existen las deleciones o inserciones de bases nucleotídicas. Las mutaciones pueden modificar la composición de aminoácidos de las proteínas provocando cambios en su tamaño y forma que pueden llevar a cambio o pérdida de su función biológica. La variación continua da lugar a un espectro ininterrumpido de fenotipos, en la que los fenotipos intermedios son más frecuentes que los extremos. Este espectro de fenotipos suele estar determinado frecuentemente por un componente genético unido a un componente ambiental (35).

ii. Bases de la genómica moderna: Proyecto Genoma Humano y ENCODE

El Proyecto Genoma Humano (PGH) se considera la base de la genómica moderna tal y como la conocemos hoy y es uno de los primeros proyectos a gran escala de la biología (38). A parte de significar un avance a nivel biológico y tecnológico, el PGH estableció algunas bases que son claves en la investigación actual como es la creación de grandes consorcios con múltiples investigadores, la estandarización de hacer públicos los datos generados, la priorización del desarrollo tecnológico o la consideración de aspectos éticos, legales y sociales (38). El proyecto fue iniciado oficialmente en 1990 y finalizó en 2003. Sin embargo, mucho antes, desde 1984, el Departamento de Energía de los Estados Unidos ya se reunió para idear una técnica para poder secuenciar el genoma humano y en 1988, el Instituto Nacional de Salud americano reunió a expertos para planear el PGH. El principal objetivo del PGH fue mapear el genoma humano y determinar la secuencia de sus 3,2 mil millones de bases. Sin embargo, se consideraron otros objetivos dentro del Proyecto como mapear y secuenciar el genoma de otros organismos importantes en el estudio de la biología, así como desarrollar una tecnología apropiada para el análisis del ADN. El objetivo global del Proyecto era amplio y el mayor hito que se quería alcanzar era el de poder tratar enfermedades gracias a la información de la secuenciación del genoma humano. En 1994 se publicaron resultados del mapa de ligamiento del genoma humano, que era el primero de los objetivos del Proyecto. En diciembre de 1999, se consiguió la secuencia de los 33,5 millones de bases del cromosoma 22. En abril de 2003 se anunció la finalización del PGH y en 2004 es cuando finalmente se publicó la secuencia del genoma humano (39).

Uno de los hitos más importantes del PGH fue el de desarrollar y mejorar las técnicas de secuenciación del ADN. El método utilizado en el PGH fue la secuenciación de ADN mediante método Sanger y el Proyecto permitió la mejora de la automatización, desarrollo de matrices o *arrays* de capilaridad y la introducción de conceptos como la cobertura profunda del genoma (40). La primera secuencia del genoma humano fue generada mediante instrumentos de secuenciación de ADN basados en capilaridad (38). Además, es importante que el PGH abrió las puertas al desarrollo de otras tecnologías

de secuenciación del ADN por síntesis que permitieron disminuir el tiempo de secuenciación del genoma (40).

Todos estos descubrimientos permitieron que el área científica de la genética evolucionara mucho. Por ejemplo, se describió que existía una variación de 1 entre 1,200-1,500 bases de ADN entre personas no relacionadas. Esta variación se explicó por la presencia de *single nucleotide variants* o polimorfismos de nucleótido simple (SNV o SNP) y por *copy number variations* o variaciones en el número de copia (CNV). También se identificó la gran similitud que existía a nivel genético entre diferentes organismos.

Uno de los descubrimientos que más ha marcado en el desarrollo de estudios posteriores fue el de que los genes no se distribuyen de forma uniforme a lo largo de los cromosomas, si no que existen regiones ricas y regiones pobres en genes. Estas regiones desérticas corresponden al 20% del genoma. Además, contrariamente a lo que se pensaba, solo alrededor del 2% del genoma humano da lugar a la síntesis de proteínas, lo que corresponde a unos 20,000 genes codificantes de proteína, similar a la de organismos considerados menos complejos. Esto permitió entender que la complejidad de los organismos venía dada por la codificación de los genes a diferentes proteínas y no por el número de genes. Además, algunos investigadores empezaron a pensar que las secuencias repetitivas del genoma podrían ser importantes para entender la estructura de los cromosomas y su dinámica (41).

Con toda esta información revelada por el PGH junto con todas las cuestiones que se habían quedado por responder, en 2003 se lanzó el proyecto ENCyclopedia of DNA Elements (ENCODE). Este proyecto se inició con el objetivo de entender los componentes funcionales y estructurales del ADN, como genes, promotores, potenciadores o *enhancers*, genes represores, exones, orígenes de replicación, etc. El proyecto ENCODE se dividió en tres fases. En la primera de ellas, la fase piloto (2003-2007), se generaron los datos relacionados con las regiones codificantes de proteínas. En la segunda fase (2007-2012) se analizó el 99% del genoma considerado no funcional o no codificante. Finalmente, en la tercera fase (2012-2016), se refinaron los resultados previos y se aplicó el conocimiento a cuestiones biológicas básicas para realizar estudios de enfermedades a gran escala (41).

iii. Estudios de asociación mediante rastreo genómico o *Genome-Wide association studies* (GWAS)

1. Conceptos básicos

Los estudios de asociación mediante rastreo genómico o *Genome-Wide Association Studies* (GWAS) fueron propuestos por primera vez por Risch y Merikangas en 1996 (42) y el primer GWAS se publicó en 2005 (43). Este tipo de estudio se centra principalmente en el análisis de cientos a miles de variaciones genéticas comunes (frecuencia alélica > 1%), también conocidos como polimorfismos, SNV o SNP (43,44), distribuidas a lo largo de todo el genoma. La hipótesis inicial en la que se centran los GWAS es que las enfermedades comunes son poligénicas, es decir, muchos polimorfismos en diferentes genes confieren un riesgo para el desarrollo de estas enfermedades, contrariamente a variaciones raras o mutaciones *de novo* (44). Los GWAS estudian las diferencias en las frecuencias alélicas de variantes genéticas en individuos que tienen un ancestro similar pero con un fenotipo diferente (45). Los resultados significativos (considerados con un valor de P inferior a 5×10^{-8}) en los GWAS suelen conllevar que un bloque de SNPs correlacionados en la misma región se asocian de manera estadísticamente significativa con el fenotipo analizado. El principal objetivo de analizar enfermedades comunes mediante GWAS es el de obtener datos sobre la base genética de estas enfermedades con el fin de predecir el riesgo y de identificar los mecanismos fisiopatológicos causantes de la enfermedad. Por lo tanto, los resultados de los GWAS pueden ayudar a desarrollar medidas preventivas para las enfermedades estudiadas pero también a la identificación de genes que pueden servir como diana terapéutica en el tratamiento de enfermedades comunes (44). Sin embargo, actualmente se ha expandido la utilidad de los GWAS y ha permitido su aplicabilidad en otras áreas como en la estimación de la heredabilidad de enfermedades o fenotipos, cálculo de correlaciones genéticas e inferencia de relaciones causales entre factores de riesgo y enfermedades (45).

2. Origen de los GWAS

Como se ha comentado anteriormente, el PGH permitió asentar y desarrollar nuevas tecnologías para la secuenciación del ADN. Además, los resultados que mostraron que el 0.1% de variación genética es lo que hace a cada persona única, cambió la visión de la genética humana para analizarla desde el punto de vista de las diferencias y no de las similitudes. Gracias a los descubrimientos del PGH, el área de la medicina genómica se centra actualmente en encontrar la base genética causante de las variaciones en el fenotipo. Para ello, se desarrollaron plataformas de *arrays* de ADN de alta densidad para genotipar millones de posiciones del genoma de cada uno de los individuos analizados (44). Esta tecnología permite conocer la variación genética común en cada genoma con el uso de grandes tamaños muestrales. Cabe destacar que su precio ha ido disminuyendo hasta llegar a menos de cien euros por cada muestra. Paralelamente, esto ha permitido que el número de estudios analizando las variaciones genéticas haya aumentado exponencialmente (44). Sin los datos de un genoma de referencia de alta calidad como el generado por el PGH, el uso de estos *arrays* no sería viable.

El proyecto internacional HapMap también fue clave para el desarrollo de los GWAS. Este proyecto, iniciado en 2002 y finalizado en 2010, fue diseñado para generar una base de datos de patrones de variación común en la secuencia de todo el genoma humano (46,47). Debido a que existe una fuerte asociación entre SNPs dentro de muchas regiones cromosómicas, no es necesario el genotipado de todos los SNPs pertenecientes a esta región. Con la selección de unos SNPs determinados que sean marcadores de esa región (conocidos como “tag SNPs”) es posible inferir cual será el genotipo en el resto de SNPs en la región. Este proceso es conocido como imputación y es posible debido a la publicación de mapas de patrones de desequilibrio de ligamiento (LD), el primero de ellos publicado por HapMap. El LD es la asociación no aleatoria de alelos de diferentes loci (o lugares específicos del genoma) y difiere entre ancestros diferentes (48). En el proyecto HapMap se genotiparon 270 muestras pertenecientes a población africana, europea, china y japonesa. Las primeras dos fases del proyecto permitieron publicar datos de LD de estas poblaciones para cubrir de un cuarto a una tercera parte de todos los SNPs con frecuencias de más de 5% (49). Posteriormente se publicaron datos con

muestras adicionales de las cuatro poblaciones anteriores y siete más (un total de 1,301 sujetos). Este nuevo mapa de LD obtenido a partir de más muestras de diferente ancestro permitió una mejor selección de tag SNPs para realizar análisis más detallados en diferentes poblaciones. Por lo tanto, el proyecto HapMap ha permitido identificar hasta cuanto se puede reducir el número de tag SNPs para conseguir la cobertura del genoma deseada (50).

3. Logros y carencias de los GWAS

Actualmente se han realizado más de 5,700 GWAS en más de 3,300 fenotipos (45). Los estudios GWAS han permitido encontrar muchas variantes genéticas asociadas al riesgo de enfermedades comunes. Este riesgo suele ser conferido por un número muy grande de polimorfismos que confieren un riesgo pequeño cada uno de ellos, lo que requiere de grandes tamaños muestrales para su identificación (44). El número de pacientes incluidos en muchos de los estudios actuales llega a ser de millones de participantes, permitiendo así un aumento del poder estadístico para encontrar un número mayor de asociaciones significativas. Los GWAS han permitido entender mecanismos complejos de enfermedades comunes y ha ayudado a desarrollar terapias de precisión, especialmente en cáncer. Los GWAS también han permitido un avance médico gracias a la translación de los resultados biológicos.

Uno de los problemas en los GWAS es que parte de los SNPs encontrados asociados al riesgo de enfermedades se han identificado en regiones no codificantes del genoma. Por lo tanto, encontrar el gen causante de la enfermedad y el mecanismo por el que éste confiere un mayor riesgo ha supuesto un reto. Sin embargo, actualmente se están desarrollando herramientas basadas en anotaciones bioquímicas o estudios de expresión, entre otros, que permiten la anotación funcional y mapeo fino para sobrellevar este problema (44).

En los estudios GWAS los resultados genéticos que se obtienen son específicos de cada población y esto puede hacer que en muchos casos los resultados no puedan ser extrapolados a otros individuos con un ancestro diferente. Actualmente, la mayoría de los GWAS están limitados a población europea, aunque gran parte de los esfuerzos actuales están dedicados a realizar análisis en diferentes poblaciones (45).

Otro punto flojo de los GWAS es que las variantes genéticas comunes identificadas en muchos de estos estudios han permitido explicar solo una parte pequeña de la heredabilidad de las enfermedades. Parte de la heredabilidad no encontrada en los GWAS se puede explicar por variaciones genéticas de baja frecuencia, por variaciones comunes que no llegan al nivel de significancia establecido en los GWAS, pero también por otros mecanismos heredables como la epigenética.

c. Estudios GWAS en el ictus

Antes de la era de los estudios GWAS, gracias a estudios en gemelos y a la historia familiar, se observó que el ictus es una enfermedad compleja con predisposición genética importante en la que múltiples genes están implicados en su riesgo (51,52). Esta característica del ictus motivó su estudio a través de GWAS. Además, debido a que los estudios genéticos epidemiológicos previos a los GWAS parecían indicar una gran heterogeneidad en la genética de los diferentes subtipos etiológicos de ictus, muchos de los GWAS han estudiado el componente genético asociado al riesgo específico de cada subtipo de ictus (53).

En 2007 se creó el *International Stroke Genetics Consortium* o Consorcio Internacional de Genética del Ictus (ISGC) (<https://www.strokegenetics.org/>) con el fin de agrupar diferentes grupos de expertos en genética del ictus a nivel internacional y así investigar y dar a conocer los factores genéticos de esta enfermedad. El Consorcio de Control de Casos Wellcome Trust o *Wellcome Trust Case Control Consortium 2* (WTCCC2) (<https://www.wtccc.org.uk/cc2/>), creado en 2008 también fue constituido con el fin de financiar la realización de un amplio catálogo de estudios GWAS en diferentes enfermedades, incluido el ictus. El WTCCC2 es una continuación del WTCCC, formado por un grupo de investigadores de UK establecido en 2005 con el objetivo de determinar los patrones de variación genética mediante el uso de tecnologías de genotipado de alto rendimiento, así como explorar el uso de GWAS. Posteriormente, en 2013, se publicó la creación de la red de genética en ictus o *Stroke Genetics Network* (SiGN) dentro del Instituto Nacional de Enfermedades Neurológicas e Ictus de Estados Unidos o *U.S. National Institute of Neurologic disorders and Stroke* (NINDS) (54). El objetivo del SiGN fue el de diseñar estudios genéticos específicos para cada subtipo etiológico de ictus

isquémico en los que hubiera acuerdo entre centros en cuanto a cómo realizar la asignación del subtipo (54).

Actualmente, se han publicado diversos estudios GWAS investigando tanto los factores genéticos asociados al riesgo de ictus en general y a sus subtipos etiológicos, así como las variaciones genéticas relacionadas con el pronóstico y evolución a corto y largo plazo de los pacientes con ictus.

i. Estudios GWAS y riesgo de ictus

1. GWAS y riesgo de ictus isquémico

El primer estudio GWAS en ictus fue publicado en 2010 utilizando una cohorte de 249 pacientes de ictus y 268 controles en el que se analizaron 400,000 SNPs (55,56). No se encontró ninguna asociación estadísticamente significativa al realizar corrección por pruebas múltiples, probablemente debido al bajo poder estadístico por presentar un tamaño muestral pequeño. Sin embargo, sí que permitió ver diferentes polimorfismos, situados en diferentes regiones cromosómicas, con una asociación nominal, por lo que se confirmó que no era solo un gen el que estaba asociado al riesgo de ictus. La primera asociación significativa y validada identificada mediante GWAS fue en el gen *PRKCH*, en un GWAS realizado por Kubo et al. en población japonesa (57) (Tabla 1). En concreto, este gen se asoció al riesgo de ictus lacunar. Posteriormente, en un metaanálisis en el que se incluyó una muestra mayor (1,544 ictus y 18,058 controles), la región 12p13 (cercana al gen *NINJ2* y *WNK1*) se encontró asociada al riesgo de ictus (52,58). Este mismo estudio replicó los resultados en dos nuevas cohortes. Sin embargo, en un nuevo estudio que unió los datos de diferentes miembros del ISGC con el objetivo de replicar estos hallazgos previos, no se pudo volver a replicar la asociación de esta región con el riesgo de ictus o con algún subtipo etiológico específico (59), estimando que la asociación observada en la región 12p13 podría ser un falso positivo. Posteriormente, se publicaron algunos estudios en los que se observó que algunas asociaciones genéticas identificadas en otras enfermedades cardiovasculares, también se asociaban con el riesgo de ictus. Este es el caso de los genes *PITX2* y *ZFHX3*, previamente asociados a FA, con los que también se encontró una asociación con el riesgo de ictus en dos GWAS realizados por el mismo grupo en población islandesa (60,61) (Tabla 1). En concreto, la asociación más potente de estos genes se encontró para el ictus de subtipo

cardioembólico. De forma similar, una variante en el cromosoma 9 (en la región 9p21), previamente asociada a infarto de miocardio y a enfermedad coronaria arterial (62), también se asoció al riesgo de ictus en diferentes poblaciones (63,64).

En el 2012 se empezaron a publicar más datos de GWAS que ya incluían un tamaño muestral mayor, uno como parte del estudio WTCCC2 (53) y el segundo como parte de la colaboración denominada METASTROKE (51), formada por miembros del ISGC. En el primero de los GWAS, en el que se incluyeron 3,548 pacientes que habían sufrido un ictus y 5,971 controles, todos ellos de ancestro europeo (53), se identificó una nueva asociación en el gen *HDAC9* con el riesgo de ictus aterotrombótico. Además, se replicaron las asociaciones en *PITX2* y *ZFHX3* con ictus cardioembólico y en la región 9p21 con ictus de tipo aterotrombótico (53). El segundo de estos GWAS consiguió reunir un número de sujetos considerablemente mayor (todos de ancestro europeo) comparado con los GWAS publicados en ictus hasta el momento. En total, se combinaron datos de 15 cohortes para alcanzar un tamaño muestral de 12,389 pacientes con ictus y 62,004 controles (51). Este estudio permitió validar las asociaciones previas de los genes *PITX2* y *ZFHX3* y del gen *HDAC9* (Tabla 1). La región 9p21 presentó una asociación nominal con el riesgo de ictus aterotrombótico. Se identificaron 12 loci con asociación nominal, pero no pudieron ser replicados (51). Los resultados de estos dos GWAS fueron metaanalizados junto a datos de una nueva cohorte con 3,420 casos de ictus isquémico y 6,821 controles, consiguiendo llegar a tener más de 15,000 casos de ictus. En este análisis se consiguió identificar una región (12q24), cercana al gen *ALDH2*, significativamente asociada al riesgo de ictus isquémico (65) (Tabla 1). Posteriormente, como parte del estudio METASTROKE, se publicó un segundo GWAS con el objetivo de estudiar variaciones raras (con una frecuencia inferior al 5%), a parte de las variaciones comunes. Esto fue posible gracias a la imputación con datos del panel de SNPs del estudio 1000 Genomas, que a diferencia de los paneles del HapMap (utilizados en los GWAS previos), cubría un número mayor de SNPs de baja frecuencia. En este nuevo GWAS del consorcio METASTROKE, se realizó un primer análisis en cohorte europea y un metaanálisis trans-étnico que incluyó un total de 26,127 pacientes con ictus y 53,788 controles (66). Se identificó el gen *ABO* asociado al riesgo de ictus isquémico en población caucásica y se confirmaron las asociaciones en los genes *PITX2*,

ZFHX3 y *HDAC9* (Tabla 1). Este estudio también permitió identificar un enriquecimiento de variantes de baja frecuencia asociadas al riesgo de ictus lacunar y aterotrombótico mientras que para el ictus de tipo cardioembólico se encontró un enriquecimiento de variantes de alta frecuencia (66).

Posteriormente, se publicó un estudio en el que se usó una aproximación diferente, utilizando la información de la edad de inicio del ictus, para encontrar asociaciones genéticas asociadas al ictus o a sus subtipos etiológicos. La idea surgió del concepto de que los pacientes jóvenes con ictus debían tener una predisposición genética mayor. Se identificó el gen *MMP12* asociado al riesgo de ictus de tipo aterotrombótico (67) (Tabla 1).

El NINDS-SiGN también publicó un GWAS en 2016, con un tamaño muestral total de 37,702 casos de ictus y 397,209 controles, que permitió identificar una región en 1p13.2 cercana al gen *TSPAN2* asociada con el ictus de tipo aterotrombótico (68) (Tabla 1). Además, la región 12q24, identificada en el metaanálisis realizado previamente entre el WTCCC2 y el METASTROKE, se asoció en este nuevo estudio específicamente con el riesgo de ictus lacunar. De nuevo se confirmaron las asociaciones previas en los genes *PITX2*, *ZFHX3* y *HDAC9* (68). Posteriormente, un nuevo GWAS con 84,961 participantes incluidos del consorcio *Cohorts of Heart and Aging Research in Genomic Epidemiology* (CHARGE) permitió encontrar un nuevo locus, en el cromosoma 6p25, cercano al gen *FOXF2* asociado con el riesgo de todos los ictus (69). Este resultado fue validado en cohortes adicionales. En este estudio también se observó que la delección del gen *Foxf2* en ratones adultos llevaba a que estos padecieran infartos cerebrales y microhemorragias (69).

En el caso específico de ictus de tipo lacunar, no se había descrito ningún gen específico asociado a este subtipo de ictus, aunque la región 12q24 sí que parecía tener más relevancia en ictus lacunares, también se había asociado con todos los ictus isquémicos. En un estudio GWAS de Traylor M et al., se identificó una asociación significativa en 16q24.2 con ictus de tipo lacunar (70). Además, el SNP más significativo en esta región se encontró que era regulador de la expresión génica de *ZCCHC14* y se asoció a tener hiperintensidades de sustancia blanca, un fenotipo característico en los ictus lacunares y otras manifestaciones de enfermedad de pequeño vaso (70). Posteriormente, se

publicó un nuevo estudio incluyendo 2,987 pacientes con ictus de tipo lacunar (71). Se encontraron cinco loci asociados *ICA1L-WDR12-CARF-NBEAL1*, *ULK4*, *SPI1-SLC39A13-PSMC3-RAPSN*, *ZCCHC14* y *ZBTB14-EPB41L3*, tanto en el análisis con pacientes europeos como en el trans-étnico. Además, los autores de este estudio realizaron un análisis adicional, conocido como *Multi-Trait Analysis of GWAS* (MTAG) (71). El MTAG consiste en el metaanálisis de GWAS de dos fenotipos diferentes pero correlacionados con el fin de incrementar el poder estadístico combinando el tamaño de efecto de ambas patologías. Se combinaron los pacientes con ictus lacunar con datos de un GWAS de hiperintensidades de sustancia blanca (muy común en la enfermedad de pequeño vaso). Este análisis permitió identificar otros siete loci asociados con ictus de tipo lacunar: *SLC25A44-PMF1-BGLAP*, *LOX-ZNF474-LOC100505841*, *FOXF2-FOXQ1*, *VTA1-GPR126*, *SH3PXD2A*, *HTRA1-ARMS2* y *COL4A2* (71).

La gran revolución en el campo de los GWAS en ictus llegó de la mano del grupo de trabajo MEGASTROKE, dentro del ISGC (72). En el estudio publicado por este grupo se incluyeron un total de 512,612 participantes (67,162 pacientes de ictus y 454,450 controles). Gracias a este incremento significativo en el número de pacientes estudiados y a que se realizaron análisis trans-étnicos, se identificaron un total de 32 loci asociados al riesgo de ictus, siendo 22 de ellos nuevas asociaciones no identificadas previamente (Tabla 1). Dieciocho de estas nuevas asociaciones fueron identificadas en el análisis trans-étnico y las demás en el análisis de participantes europeos. A parte del estudio GWAS, se realizaron análisis adicionales. Entre ellos, se hizo un estudio de genes en el que en lugar de estudiar la asociación de SNPs con el riesgo de ictus, se estudió la asociación de la suma de polimorfismos dentro de un mismo gen con el riesgo de ictus. Este análisis permitió identificar una nueva asociación en la región de los genes *ICA1L-WDR12* con el riesgo de ictus lacunar. La gran mayoría de los loci se asociaron con los ictus isquémicos (20 loci) o con el riesgo de ictus independientemente de su subtipo, tanto isquémicos como hemorrágicos (18 loci). Un total de 6 loci se asociaron con el riesgo de ictus aterotrombótico y 4 loci con los ictus cardioembólicos. En el caso de los ictus lacunares, las únicas asociaciones significativas fueron las del análisis de asociación por genes. El estudio MEGASTROKE también permitió encontrar un solapamiento de las variantes asociadas al riesgo de ictus con variantes genéticas previamente asociadas a

factores de riesgo vascular. Esta asociación también se encontró al construir puntuaciones o *scores* de riesgo genético para diferentes fenotipos vasculares los cuales se asociaron con el riesgo de ictus. La asociación más importante fue para el *score* de enfermedad coronaria arterial con el riesgo de ictus aterotrombótico, debido a que ambas comparten la misma fisiopatología. Finalmente, es importante indicar que el estudio MEGASTROKE demostró que el uso de GWAS puede ser una buena estrategia para la identificación de potenciales tratamientos. En concreto, se encontró que un 11% de los genes identificados en este GWAS estaban registrados como dianas de fármacos aprobados para su uso clínico, entre ellos fármacos antitrombóticos utilizados en el ictus (72). Estos resultados son muy importantes ya que han permitido seguir evaluando la estrategia de los GWAS para el reposicionamiento farmacológico, del cual se hablará más adelante.

Tabla 1: Resumen de los loci asociados al riesgo o evolución del ictus mediante GWAS

LOCUS	FENOTIPO ICTUS	POBLACIÓN	REFERENCIA
<i>PRKCH</i>	Riesgo: LAC	Japoneses	54
<i>PITX2</i>	Riesgo: Ictus, CE	Europeos	57,58,50,48,63,65,69,72
<i>ZFHX3</i>	Riesgo: Ictus, CE	Europeos	57,58,50,48,63,65,69,72
9p21	Riesgo: Ictus, AT	Trans-étnico	60,61,50
<i>HDAC9</i>	Riesgo: AT	Europeos	50,48,63,65,69,72
12q24 (<i>ALDH2</i> , <i>SH2B3</i>)	Riesgo: IS, LAC	Europeos	62,65,69,72
<i>ABO</i>	Riesgo: IS	Europeos	63,69,72
1p13.2 (<i>TSPAN2</i>)	Riesgo: AT	Europeos	64,69
6p25 (<i>FOXF2</i>)	Riesgo: Ictus, LAC	Europeos	66,68,69,72
16q24.2	Riesgo: LAC	Europeos	67,69
<i>MMP12</i>	Riesgo: ATE	Trans-étnico	64,69,72
<i>ICA1L-WDR12</i>	Riesgo: LAC, HIC no lobar	Trans-étnico	68,69,70,72,74
1p36(<i>CASZ1</i>)	Riesgo: Ictus, IS	Trans-étnico	69,72
1p13(<i>WNT2B</i>)	Riesgo: Ictus, IS	Trans-étnico	69,72

1q43(<i>RGS7</i>)	Riesgo: CE	Europeos	69
2p23(<i>KCNK3</i>)	Riesgo: Ictus, IS	Trans-étnico	69,72
3q25(<i>TM4SF4– TM4SF1</i>)	Riesgo: ATE	Europeos	69
4q25(<i>ANK2</i>)	Riesgo: IS	Europeos	69
4q31(<i>EDNRA</i>)	Riesgo: ATE	Trans-étnico	69
4q31(<i>FGA</i>)	Riesgo: IS, CE	Trans-étnico	69,72
5q23(<i>LOC100505841</i>)	Riesgo: IS, LAC	Trans-étnico	68,69,72
5q35(<i>NKX2-5</i>)	Riesgo: CE	Trans-étnico	69
6p21(<i>SLC22A7– ZNF318</i>)	Riesgo: Ictus	Trans-étnico	69,72
7q21(<i>CDK6</i>)	Riesgo: IS	Trans-étnico	69,72
9p21(Chr9p21- <i>CDKN2B</i>)	Riesgo: Ictus, IS	Trans-étnico	69,72
9q31(<i>LINC01492</i>)	Riesgo: ATE	Europeos	69
10q24(<i>SH3PXD2A</i>)	Riesgo: Ictus, IS, LAC	Trans-étnico	68,69,72
12p12(<i>PDE3A</i>)	Riesgo: IS	Trans-étnico	69,72
12q24(<i>TBX3</i>)	Riesgo: Ictus, IS	Trans-étnico	69,72
13q14(<i>LRCH1</i>)	Riesgo: Ictus, IS	Trans-étnico	69,72
15q26(<i>FURIN–FES</i>)	Riesgo: IS	Trans-étnico	69,72
17p13(<i>PRPF8</i>)	Riesgo: IS	Trans-étnico	69
19p13(<i>ILF3–SLC44A2</i>)	Riesgo: IS	Trans-étnico	69,72
19p13(<i>SMARCA4– LDLR</i>)	Riesgo: Ictus, IS	Trans-étnico	69,72
<i>NOS3</i>	Riesgo: Ictus	Europeos	70,72
<i>COL4A1</i>	Riesgo: Ictus	Europeos	70
<i>DYRK1A</i>	Riesgo: Ictus	Europeos	70
<i>HNF1A</i>	Riesgo: IS	Africanos	71
<i>ULK4</i>	Riesgo: LAC	Europeo, Trans-étnico	68

<i>SPI1-SLC39A13- PSMC3-RAPSN</i>	Riesgo: LAC	Europeo, Trans-étnico	68
<i>ZBTB14-EPB41L3</i>	Riesgo: LAC	Europeo, Trans-étnico	68
<i>LOX-ZNF474</i>	Riesgo: LAC	Europeo, Trans-étnico	68
<i>VTA1-GPR126</i>	Riesgo: LAC	Europeo, Trans-étnico	68
<i>FGF5</i>	Riesgo: IS	Trans-étnico	72
<i>GRK5</i>	Riesgo: IS	Trans-étnico	72
<i>CENPQ</i>	Riesgo: IS	Trans-étnico	72
<i>PRRX1</i>	Riesgo: CE	Trans-étnico	72
<i>F11</i>	Riesgo: CE	Trans-étnico	72
<i>LINC00924</i>	Riesgo: IS	Trans-étnico	72
<i>PROCR</i>	Riesgo: IS	Trans-étnico	72
<i>ATP2B1</i>	Riesgo: Ictus	Trans-étnico	72
<i>PTPN11</i>	Riesgo: IS	Trans-étnico	72
<i>HOXC4</i>	Riesgo: IS	Trans-étnico	72
<i>PLSCR5</i>	Riesgo:	Trans-étnico	72
<i>MTMR4</i>	Riesgo: IS	Trans-étnico	72
<i>ANKRD33</i>	Riesgo: Ictus	Trans-étnico	72
<i>DEFB1</i>	Riesgo: CE	Trans-étnico	72
<i>SWAP70</i>	Riesgo: IS	Trans-étnico	72
<i>EVX1</i>	Riesgo: Ictus	Trans-étnico	72
<i>LAMC1</i>	Riesgo: Ictus	Trans-étnico	72
<i>PRDM16</i>	Riesgo: LAC	Trans-étnico	72
<i>RPRML</i>	Riesgo: Ictus	Trans-étnico	72
<i>EBF2</i>	Riesgo: Ictus	Trans-étnico	72
<i>THADA</i>	Riesgo: Ictus	Trans-étnico	72
<i>LSP1</i>	Riesgo: IS	Trans-étnico	72
<i>LPA</i>	Riesgo: ATE	Trans-étnico	72

<i>MESDC1</i>	Riesgo: Ictus	Trans-étnico	72
<i>FIGN</i>	Riesgo: Ictus	Trans-étnico	72
<i>THAP5</i>	Riesgo: IS	Trans-étnico	72
3p12	Riesgo: IS	Africanos	72
<i>COBL</i>	Riesgo: IS	Asiáticos	72
<i>HTRA1</i>	Riesgo: LAC	Trans-étnico	68,72
<i>PTCH1</i>	Riesgo: LAC	Africanos	72
<i>BNC2</i>	Riesgo: IS	Europeos	72
<i>TAP1</i>	Riesgo: IS	Europeos	72
<i>PIK3CG</i>	Riesgo: IS	Trans-étnico	72
12q21.1	Riesgo: HIC lobares	Europeos	73
<i>APOE</i>	Riesgo: HIC lobares	Europeos	74 63
1q22 (<i>PMF1</i>)	Riesgo: HIC no lobar, LAC	Europeos	68,72,73,74
2q33	Riesgo: HIC no lobar	Europeos	74
13q34 (<i>COL4A2</i>)	Riesgo: HIC no lobar, LAC	Europeos	68,72,74
<i>ZCCHC14</i>	Riesgo: HIC no lobar, LAC	Europeos	68,72,74
<i>ZBTB46</i>	Endofenotipo fase aguda: Hematoma parenquimatoso (rTPA)	Europeos	77
<i>RP11-362K2.2:RP11-767I20.1</i>	Endofenotipo fase aguda: Hematoma	Europeos	78

	parenquimatoso (reperfusión)		
2p25.1	Endofenotipo fase aguda: Δ NIHSS24h	Trans-étnico	79
2q31.2	Endofenotipo fase aguda: Endofenotipo fase aguda: Δ NIHSS24h	Trans-étnico	79
2q33.3	Δ NIHSS24h	Trans-étnico	79
4q34.3	Endofenotipo fase aguda: Δ NIHSS24h	Trans-étnico	79
5q33.2	Endofenotipo fase aguda: Δ NIHSS24h	Trans-étnico	79
6q26	Endofenotipo fase aguda: Δ NIHSS24h	Trans-étnico	79
7p21.1	Endofenotipo fase aguda: Δ NIHSS24h	Trans-étnico	79
14q24.3	Endofenotipo fase aguda: Contaje leucocitario	Europeos	80
<i>PATJ</i>	mRS 90 días	Europeos	81
<i>LOC105372028</i>	mRS 60-190 días	Europeos	82

IS: ictus isquémico; LAC: ictus lacunar; AT: ictus aterotrombótico; CE: ictus cardioembólico; HIC: Hemorragia intracerebral; mRS: modified rankin scale; Δ NIHSS24h = NIHSS basal – NIHSS 24h

Tabla original de la doctoranda

Poco después, el estudio MEGASTROKE fue combinado con datos del UK Biobank en un metaanálisis de 72,147 casos y 823,869 controles. En este nuevo análisis se identificaron tres nuevos loci asociados al riesgo de ictus en *NOS3*, *COL4A1* y *DYRK1A*. Sin embargo, como limitación, en este estudio no se pudo realizar un análisis específico por subtipos etiológicos de ictus, porque la cohorte del UK Biobank no dispone de estos datos (73).

Debido a que se conoce que el riesgo genético para las enfermedades complejas varía en función del ancestro y a que la mayoría de los estudios GWAS en ictus se habían realizado en población europea o en análisis trans-étnicos, Keene et al. llevó a cabo un estudio GWAS en población africana (74). Este estudio forma parte de un consorcio denominado Consorcio de estudios de asociación de todo el genoma en ictus en poblaciones minoritarias o *Consortium of Minority Population genome-wide Association Studies of Stroke* (COMPASS). En él se incluyeron 22,000 individuos de ancestro africano de 13 cohortes diferentes. En este metaanálisis se identificó una asociación significativa cercana al gen *HNF1A* y 24 loci nominalmente asociados al riesgo de ictus. La asociación en *HNF1A* y 16 de los 24 loci nominales se validaron en diferentes cohortes de población europea (74).

Muy recientemente, en el año 2022 se publicaron los resultados del Consorcio GIGASTROKE, que forma parte del ISGC. Este es el mayor estudio GWAS en ictus publicado hasta el momento, con 110,182 casos de ictus y 1,468,220 controles incluidos en un metaanálisis trans-étnico. Este aumento del tamaño muestral ha permitido un aumento significativo del poder estadístico y con ello la identificación de 60 loci asociados (34 de ellos no conocidos previamente) asociados al riesgo de ictus (Tabla 1). Se encontró evidencia en 47 genes que podrían tener una potencial implicación funcional en el ictus. Estos datos, aparte de permitir revelar mecanismos por los que se produce la enfermedad, ha ayudado a identificar posibles dianas terapéuticas y a

mejorar la predicción del riesgo de ictus. Entre estos resultados destacan los obtenidos mediante el análisis de asociación de un *score* de riesgo genético construido con los loci asociados a cualquier subtipo de ictus y utilizado para la predicción del riesgo de ictus isquémico en cinco ensayos clínicos de enfermedades cardiometabólicas. Se observó una Hazard Ratio (HR) de 1.35 de desarrollar un ictus isquémico en los pacientes con un valor en el tercil más alto del *score* (75).

2. GWAS y riesgo de ictus hemorrágico

Comparado con el número de estudios GWAS realizados en ictus isquémico, la cantidad de estudios focalizados exclusivamente en ictus hemorrágico es mucho menor. Aunque en muchos de los GWAS realizados en pacientes con ictus descritos anteriormente se realizaron análisis combinando tanto pacientes con ictus isquémico como pacientes con ictus hemorrágico, estos estudios no realizaban análisis específicos en pacientes con ictus hemorrágico.

Actualmente solo hay dos GWAS publicados que hayan analizado pacientes con ictus hemorrágico de forma independiente. El primero fue publicado por Woo et al., en 2014 (76). En este GWAS se incluyeron tanto pacientes con HIC lobar ($n = 664$) como pacientes con HIC no lobar ($n = 881$) (76). Esta es una clasificación de las HIC según el territorio que afectan. Ambos tipos de HIC difieren en la patología vascular subyacente. Previamente se había estimado mediante estudios de heredabilidad un alto componente genético en el riesgo de HIC, aunque se esperaba que este componente genético asociado a cada subtipo de HIC fuera diferente. La variación genética en estos casos se comparó con la de 1,481 controles sanos. Se identificó un loci asociado a las HIC lobares en la región 12q21.1 y otro para las no lobares en la región 1q22. Sin embargo, solo pudo ser replicada la asociación identificada para las HIC no lobares (76).

Cinco años más tarde, en 2019, se publicó un nuevo estudio GWAS en HIC (77). En este caso, primero se utilizó la misma estrategia que en el GWAS previo, realizando un análisis por subtipos de HIC, incluyendo 755 pacientes con HIC lobar y 1005 pacientes con HIC no lobar, además de 1,711 controles sanos. En este análisis se encontraron variantes genéticas del gen *APOE* significativamente asociadas al riesgo de HIC lobar.

También se identificó una región nominalmente asociada al riesgo de HIC no lobar en la región 13q34 pero ninguna asociación en 1q22. Los autores describieron que estas diferencias con el estudio previo podían ser causadas por un diferente tamaño muestral pero también por la imputación con diferentes paneles (el HRC en este caso comparado con el de 1000 Genomas en el estudio previo). Sin embargo, los autores también realizaron un MTAG (77). En este caso concreto, se combinaron los GWAS de ictus lacunar con el de HIC, ya que ambos tipos de patología son manifestaciones agudas de la enfermedad de pequeño vaso. Este análisis permitió aumentar el tamaño muestral hasta un total de 241,024 participantes (6,255 casos y 233,058 controles). En este análisis sí que se encontró una asociación significativa para el riesgo de HIC no lobar en la región previamente identificada por Woo et al. en 1q22. Además, los loci 2q33 y 13q34 también se asociaron al riesgo de HIC no lobares. Mediante estudios *in silico* de determinación funcional se observó que los polimorfismos significativos de la región 1q22 se asociaban a la expresión de *PMF1* y los de la región 2q33 a *NBEAL1*, *FAM117B* y *CARF*. Finalmente, los de la región 13q34 lo hacían con los genes *COL4A2* y *COL4A1*. Además, utilizando una estrategia para encontrar asociación por genes, en lugar de por SNPs, se encontró el gen *ICA1L* y el gen *ZCCHC14* asociados con el riesgo de HIC no lobar (77).

ii. Estudios GWAS y pronóstico del ictus isquémico

El pronóstico después de un ictus es muy variable, tanto a corto plazo (durante las primeras 24-48 horas), a medio plazo (hasta los 10 días) o a largo plazo (a partir de los 10 días). La variación en el pronóstico está influido por muchas variables, descritas previamente, incluida la genética (30). El estudio de los factores que influyen en la evolución después del ictus nos permite aumentar el conocimiento en las vías implicadas en los cambios después del ictus y también poder predecir la evolución de cada paciente y con ello poder aplicar terapias o tratamientos como la fibrinólisis o la trombectomía. Además, estos tratamientos también están directamente relacionados con la recuperación de los pacientes después del ictus. Por lo que el estudio de los factores asociados a la respuesta a estos tratamientos también es importante en la determinación de las vías involucradas en la evolución del ictus (30).

Actualmente se han realizado estudios GWAS con el objetivo de identificar factores genéticos asociados a la evolución temprana y a largo plazo del ictus. A parte de estudiar directamente la genética de la evolución de los pacientes mediante variables cuantitativas de la evolución (NIHSS y mRS), también se ha estudiado el componente genético de endofenotipos de los ictus determinantes de la evolución en la fase aguda del ictus.

1. GWAS y endofenotipos de la fase aguda

La determinación de los factores genéticos implicados en la evolución neurológica de los pacientes en la fase aguda, medida con la escala NIHSS, es muy relevante debido al potencial predictivo de la evolución inicial de los pacientes en el estado funcional a largo plazo. Además, la fluctuación en la escala NIHSS durante las primeras 24h, cuantificada por la diferencia entre la NIHSS basal y a las 24 horas (Δ NIHSS24h) es también predictivo de la evolución a largo plazo del paciente (29).

Entre las principales variables asociadas al estado funcional durante las primeras horas del paciente, se encuentran los niveles de glucosa, la severidad inicial del ictus, la presión sanguínea, la FA, el tratamiento con estatinas previo al ictus, la realización de trombectomía y trombólisis, la TH y los niveles de células del sistema inmunológico (30). Como ya se ha comentado, entre las variables más asociadas al Δ NIHSS24h se encuentran la mayoría de las variables asociadas a la evolución en la fase aguda (la NIHSS inicial, el tratamiento con rtPA, la edad, la glucosa y la presión sistólica) (29).

Todas estas variables se pueden considerar endofenotipos de fase aguda del ictus. Un endofenotipo se considera una variable cuantificable que se asocia a la patología o fenotipo estudiado y no puede ser observada a simple vista (78). En este caso concreto, un endofenotipo sería cualquier factor cuantificable que se relacione con el pronóstico en la fase aguda del ictus. Por ejemplo, diferentes parámetros de neuroimagen como la formación de edema, los niveles de leucocitos, la recanalización después del tratamiento endovascular, etc. Además, en el caso del ictus, algunos endofenotipos son generales para todos los subtipos etiológicos de ictus mientras que otros son específicos de ciertos subtipos (78).

El ISGC creó un grupo de trabajo, llamado El grupo de trabajo de endofenotipos agudos o *The acute endophenotypes working group* con el objetivo de poder establecer proyectos de genómica en el área de los endofenotipos de la fase aguda y así capturar mecanismos implicados en la evolución aguda del ictus isquémico (79). Algunos de los proyectos llevados a cabo por miembros de esta iniciativa incluyen estudios de farmacogenética para estudiar los factores genéticos implicados en las TH en pacientes tratados con trombólisis y trombectomía, el estudio del cambio neurológico temprano (mediante la Δ NIHSS24h) o el estudio del conteo celular de leucocitos, entre otros.

En el caso del estudio mediante GWAS de las TH en pacientes tratados con trombólisis (rtPA) o trombectomía mecánica, existen dos artículos publicados (80,81). Las transformaciones hemorrágicas, especialmente la más severa de ellas (el hematoma parenquimatoso (HP)), se asocia con discapacidad y mortalidad a los tres meses post-ictus. Una variante en el gen *ZBTB46* se identificó y replicó mediante GWAS encontrándose una asociación a HP en 1,324 pacientes tratados con rtPA (80). Por otro lado, se identificó una asociación significativa de un polimorfismo en el gen *RP11-362K2.2:RP11-767I20.1* con los HP en un metaanálisis de 2,034 pacientes que recibieron terapia de perfusión (trombólisis o trombectomía mecánica) (81). Además, mediante estudios de correlación genética, se encontró que las HP comparten bases genéticas con la enfermedad de Alzheimer y con la presencia de hiperintensidades de sustancia blanca en neuroimagen (81).

Otro endofenotipo de pronóstico de fase aguda cuantificable en el ictus isquémico es la variación en la NIHSS durante las primeras 24 horas estudiado con la variable Δ NIHSS24h. En un GWAS trans-étnico realizado con esta variable en 5,876 participantes de diferentes países como parte del estudio *Genetics of Early Neurological Instability after Ischemic Stroke* (GENISIS), se identificaron siete loci diferentes (2p25.1, 2q31.2, 2q33.3, 4q34.3, 5q33.2, 6q26 y 7p21.1) que explicaban el 2,1% de la variación de la NIHSS en la fase aguda. Se realizaron estudios de anotación funcional para identificar los genes causantes de estas asociaciones. Con esta metodología, se encontraron los genes *ADAM23* y *GRIA1*, ambos reguladores de vías de excitabilidad neuronal (82).

El conteo celular de leucocitos durante las primeras 24 horas después de un ictus también es un endofenotipo que se ha asociado al pronóstico de los pacientes. En el

GWAS publicado por Torres-Águila N, et al., se incluyeron 2,064 pacientes con ictus isquémico y con datos de conteo celular de leucocitos (83). En el estudio principal, el locus 14q24.3 se asoció con estos niveles leucocitarios y también con el pronóstico agudo del ictus (NIHSS a las 24 horas). Esta región es modificadora de la expresión de los genes *ACOT1* y *PTGR2*. Además, combinando los datos del estudio principal con los de la replicación, se identificaron otras regiones asociadas con los niveles de leucocitos (83).

2. GWAS y la evolución del ictus a largo plazo

Se han realizado estudios genéticos de GWAS para determinar los factores genéticos asociados con la evolución a largo plazo en el ictus y en concreto con la discapacidad post-ictus con el objetivo de analizar las vías metabólicas implicadas y encontrar potenciales tratamientos.

El primero de ellos, el estudio *Genetic contribution to functional Outcome and Disability after Stroke* (GODS) analizó como variable de evolución funcional post-ictus el valor en la escala mRS (como variable cuantitativa) a los tres meses. En el metaanálisis realizado con 1,791 pacientes (pacientes de la fase inicial de descubrimiento y de la fase de replicación), se identificó el gen *PATJ* asociado a un peor pronóstico funcional a los tres meses (84).

El segundo de los estudios, publicado por parte de la red *Genetics of Ischaemic Stroke Functional Outcome* (GISCOME), también analizó la variable mRS, pero en este caso medida entre los 60 y 190 días después del ictus (85). En el análisis principal se consideró la variable mRS como dicotómica (0-2 vs 3-6), lo que indica pacientes independientes vs pacientes dependientes funcionalmente. Se identificó el gen *LOC105372028* asociado a la evolución funcional de los pacientes con ictus. Este gen regula la expresión de *PPP1R21*. Sin embargo, los resultados de este estudio no fueron replicados en ninguna cohorte adicional (85).

d. Epigenética

i. Conceptos básicos

La epigenética es un concepto que se describió por primera vez en 1942 por Waddington (86) para explicar cambios en el fenotipo sin modificaciones del genotipo. Actualmente, se puede definir la epigenética como los cambios en la expresión génica sin modificación del código genético y que funcionan como un punto intermedio entre el genoma y el ambiente (87). Las principales características de la epigenética es que es heredable, muy flexible y se altera según el estilo de vida o las condiciones ambientales (88). Estas características que pueden parecer contradictorias se explican por diferentes tipos de epigenética, la “directa” y la “indirecta”. La epigenética directa consiste en todas las modificaciones que ocurren durante la vida de un individuo mientras que la indirecta son todas las modificaciones epigenéticas que se heredan de los padres (87). Gracias a la epigenética, se pueden establecer los diferentes subtipos celulares, lo que conlleva que cada célula concreta pueda tener un patrón epigenético diferente. La combinación de la información genética junto a la epigenética permite obtener mucha información acerca de los procesos implicados en las diferentes enfermedades y explicar parte de la “heredabilidad perdida”, que no puede ser explicada solo a través de estudios genéticos (86,88). Pero además, como la mayoría de modificaciones epigenéticas son reversibles, se consideran potenciales dianas terapéuticas (86,88). Actualmente, la epigenética clínica se utiliza en oncología (88). Existen diferentes biomarcadores de diagnóstico, pronóstico y respuesta a terapia y también tratamientos basados en la epigenética aprobados en cáncer (88).

El avance tecnológico ha permitido el estudio del epigenoma hasta la máxima resolución (a nivel de pares de bases) (86). Actualmente existen diferentes mapas epigenómicos de enfermedad y salud en humanos (88).

ii. Mecanismos epigenéticos

Los principales mecanismos epigenéticos se pueden clasificar en modificación de las histonas, ARN no codificante y metilación del ADN. Sin embargo, otros procesos como el *splicing* alternativo o la edición del ADN también se han descrito como procesos epigenéticos (87).

Modificaciones de histonas: Las histonas, que forman parte del nucleosoma, pueden ser modificadas covalentemente en las colas amino-terminales. Estas modificaciones influyen sobre la transcripción y otros procesos como la reparación, replicación y recombinación del ADN. Este efecto es posible debido a la alteración directa de la estructura de la cromatina, pero también al efecto sobre la unión de moléculas efectoras (89). Actualmente se sabe que la combinación de modificaciones de histonas es relevante para determinar diferentes patrones biológicos y en la regulación de la accesibilidad a la cromatina (36,86).

Las modificaciones más relevantes de las histonas son la acetilación, fosforilación y metilación. La primera de las modificaciones, descrita a mitades de 1960 por Allfrey, fue la acetilación de las lisinas. Esta acetilación es muy dinámica y está mediada por acetiltransferasas de histonas (HATs) y deacetilasas de histonas (HDACs) (89). Normalmente, la acetilación de las histonas se asocia con un aumento de la expresión génica mientras que la deacetilación se asocia con una inhibición de la expresión (90).

La fosforilación de las histonas tiene lugar principalmente en serinas, treoninas y tirosinas, en la región N-terminal de las histonas. La fosforilación está regulada por las enzimas quinasas y fosfatasas. Tanto la fosforilación como la acetilación de las histonas altera la carga positiva de la proteína y de esta manera afecta a la estructura de la cromatina, haciéndola menos compacta y por lo tanto facilitando el acceso de diferentes proteínas efectoras (89).

Las histonas también pueden ser metiladas, principalmente en lisinas y argininas. En este caso, la carga de la proteína no se ve alterada. Algunos residuos específicos de lisina son metilados por diferentes lisinas metiltransferasas de histonas (HKMTs). Las lisinas pueden ser mono-, di- o tri-metiladas (91). En el caso de la arginina, son las argininas metiltransferasas las encargadas de esta función (PRMTs). Esta metilación puede ser mono- o di-metilación (91). Ambos tipos de metilación pueden ser reversibles gracias a la función de las demetilinas (89). Según el residuo afectado y la valencia, esta modificación puede conllevar una inhibición o activación de la transcripción, debido a los diferentes patrones y combinaciones que puede adoptar (90,91).

ARN no codificante (ARNnc): Del 98% de las regiones no codificantes a proteínas del genoma, un 80% se transcribe a ARN. Este ARN es considerado no codificante (92). Por lo tanto los ARNnc son moléculas de ARN que no se traducen a proteínas y tienen diferentes funciones reguladoras como el control de la transcripción, el procesamiento del ARN, el control de la arquitectura nuclear y la regulación de la cromatina (87). Los ARNnc pueden ser regulados por modificaciones epigenéticas pero también son efectores de la epigenética, pudiendo transmitir información epigenética reguladora (87). El ARNnc se puede clasificar según su tamaño en ARNnc cortos (cuando tiene menos de 200 nucleótidos) o largos (más de 200 nucleótidos).

Los ARNnc largos (ARNncl) actúan mediante la regulación de la transcripción, realizando principalmente su función como moléculas de señalización. También regulan la estabilidad del ARN y el reclutamiento de complejos proteicos (87). Los ARNncl se suelen transcribir y procesar de forma similar a los ARN codificantes. Se considera que más de un 68% del transcriptoma humano son ARNncl (92). Estos pueden ser (87):

- Transcritos naturales antisentidos (NAT): complementarios a un ARN codificante.
- ARNnc intergénico largo (ARNlinc): codificado por intrones de regiones intergénicas.
- ARN solapado en el mismo sentido: transcrito por la misma cadena de ADN que otro transcrito.
- ARN intrónico en el mismo sentido: originado de las regiones intrónicas de genes codificantes.
- Transcrito procesado: un transcrito de ARN que ha sufrido *splicing*.

Los ARNnc cortos pueden ser (87):

- ARN que interacciona con Piwi (piARN): Son moléculas de entre 24 y 31 nucleótidos que forman complejos con las proteínas Piwi. Su principal función es la de inhibir la actividad de los transposones durante el desarrollo de la línea germinal pero también actúan en células somáticas (93). Tienen la capacidad de promover la metilación de la histona 3 y de esta manera inhibir la transcripción de sus ARN diana (87).

-micro ARN (miARN): Son moléculas de ARN de una sola cadena de 20-24 nucleótidos que tienen como diana el ARN con el cual son complementarios en una región de 6 nucleótidos en el 3'UTR (90). Los miARN regulan actividades de silenciamiento de alrededor del 50% de los genes a nivel post-transcripcional (93). Un solo miARN puede estar implicado en la regulación de diferentes ARNm diana. Los miARN ejercen su acción inhibidora mediante la degradación del ARNm o el bloqueo de la traducción (94).

-ARN de interferencia (ARNi): Consiste en una doble cadena de ARN linear que se procesa a ARN de cadena simple de entre 20-24 nucleótidos. Tiene una función de silenciamiento post-transcripcional similar a la de los miARN (93).

-ARN nucleolar (ARNsno): Están implicados en modificaciones para convertir el ARN en ribosomal (87).

-Derivados de ARN de transferencia (tsARN): Los tsARN derivan de los ARN de transferencia y tienen una función similar a los miARN (87).

Metilación del ADN: Consiste en la adición de un grupo metilo en las citosinas del ADN (86). Esta modificación se explicará en el siguiente apartado con más detalle.

iii. Metilación del ADN

1. Descripción

El ADN puede ser modificado mediante un proceso químico de adición de un grupo metilo en la posición 5 de las citosinas, conocido como 5-metilcitosina (5mC). Las enzimas ADN metiltransferasas (DNMTs) son las encargadas de transferir el grupo metilo. Además la *ten-eleven translocation* (TET), una enzima metilcitosina dioxigenasa puede transferir grupos hidroximetilo a la citosina, siendo otra forma posible de metilación del ADN (95). Frecuentemente estas citosinas se encuentran en un contexto de dinucleótidos de citosinas y guaninas (CpG) (96). Se estima que en el genoma humano hay alrededor de millones de CpGs. Las regiones del genoma ricas en CpGs se denominan 'islas CpG'. Se definen como regiones de al menos 500 pares de bases con un contenido de citosinas y guaninas superior al 55% (97). En los genomas de los vertebrados se encuentran islas CpG en aproximadamente la mitad de los genes. La densidad de CpGs en los promotores suele tener una distribución bimodal (98). Actualmente se sabe que la metilación también puede ocurrir en citosinas en otros

contextos (citosina + citosina/adenina/timina) de una forma más frecuente de lo que se pensaba (95).

La influencia de esta modificación en la regulación génica fue descrita a mitad de los años 70 por Holliday y Pugh. Más adelante, en los años 80, se encontró una asociación entre la metilación del ADN y la represión génica (86). Actualmente se sabe que la metilación del ADN se asocia a represión génica cuando se produce en regiones promotoras ricas en CpGs y en elementos transposables mientras que la metilación intra-génica da lugar a la activación génica (96). Esta regulación de la transcripción viene determinada por el control que ejerce la metilación del ADN en la accesibilidad a puntos de unión de factores de transcripción en el ADN (96). La distribución específica de las marcas de metilación determina información importante para la represión o activación de la expresión génica (99). Normalmente existe una hipermetilación en zonas densas en CpGs (los promotores de los genes, regiones intra-génicas y puntos de inicio de transcripción). Los genes que se encuentran más hipermetilados son aquellos implicados en morfogénesis de tejidos, desarrollo de sistema hematológico o respuesta efectora inmunológica (100). Estos cambios en la expresión génica causados por la metilación del ADN son relevantes para la regulación de procesos celulares, impronta genética, inactivación del cromosoma X y reparación del ADN, entre otros. Además, en diferentes enfermedades existe un patrón diferente de metilación que determina el riesgo o la evolución de esa enfermedad (101).

El área clínica donde se ha descrito más ampliamente la metilación del ADN es en cáncer. Primero, se observó una hipometilación global en cáncer y posteriormente una hipermetilación focalizada en genes supresores de tumores (86). Posteriormente, la influencia de la metilación del ADN se ha estudiado en muchas otras enfermedades, como la diabetes, obesidad, aterosclerosis, hipertensión arterial y enfermedad coronaria (102). Por ejemplo, en el caso de la aterosclerosis, la hipertensión y la enfermedad coronaria también se ha observado alteración de la metilación global (103). Con el tiempo, se han ido desarrollando diferentes fármacos que afectan el estado de metilación. El primero de ellos fue la 5-azacitidina, la cual es capaz de bloquear la metilación del ADN y de esta manera afectar a la expresión génica y a su vez al fenotipo de líneas celulares de fibroblasto (86).

Al estudiar la función que la metilación del ADN tiene en diferentes enfermedades es importante tener en cuenta que la metilación puede producirse como consecuencia de la patología o puede ser causante de ella. En algunos casos, la variación en la metilación se encontrará previa al desarrollo de la enfermedad. En esos casos, esa metilación específica se encuentra en la línea germinal o se puede producir de forma estocástica en el útero y afectar a todas las células somáticas o producirse después del nacimiento y afectar solo a tejidos concretos. También puede ser consecuencia de un determinado ambiente (95). Además, la metilación global del genoma es cambiante a lo largo de la vida. Esta remodelación en la metilación del ADN es especialmente importante en los leucocitos durante los primeros cinco años de vida. En personas mayores de 50 años, la metilación tiende a reducirse con los años.

La variación genética también puede influir la metilación del ADN. Aquellos polimorfismos que comportan un cambio en la metilación son denominados loci de rasgos cuantitativos de metilación del ADN o *DNA methylation quantitative trait loci* (mQTL) (95). La mayoría de los mQTL regulan la metilación del mismo loci en el que se encuentran (cis-mQTL), aunque algunos de ellos también lo pueden hacer en otros genes (trans-mQTL).

2. Herramientas para el estudio de la metilación del ADN

Como ya se ha comentado, la metilación se distribuye de una manera específica en puntos concretos del genoma (99). La determinación de este patrón es importante para entender como la metilación del ADN afecta a la expresión génica durante procesos como el desarrollo o en diferentes enfermedades (99).

Hay información importante para tener en cuenta al hacer análisis de metilación del ADN. Por un lado, debido a que la epigenética es específica de cada célula, la metilación del ADN puede ser estudiada a nivel de una sola molécula de ADN o como la media de metilación de una región en todas las moléculas de ADN presentes en la muestra (99). Por otro lado, el grupo 5mC se pierde durante la amplificación del ADN por lo que técnicas como la reacción en cadena de la polimerasa (PCR) no pueden detectarlo. Por este motivo, los diferentes métodos para el análisis de la metilación del ADN utilizan tratamientos previos antes de la amplificación o hibridación. Este pre-tratamiento se

puede realizar mediante digestión por endonucleasa, enriquecimiento por afinidad o conversión de bisulfito.

En el primero de ellos se hace uso de las endonucleasas, unas enzimas de restricción que se unen a una secuencia específica del ADN y la digieren. Estas endonucleasas pueden ser sensibles o insensibles a la metilación. Según este efecto, su función como endonucleasa se activará o inactivará en función del estado de metilación. En ambos casos, esto permitirá distinguir entre una región metilada y no metilada (104).

En el caso del enriquecimiento por afinidad, se utilizan anticuerpos o proteínas que se unen específicamente a las regiones metiladas del genoma (99).

Por último, la conversión de bisulfito, descubierta en los años 90, supuso una revolución en las técnicas de análisis de la metilación del ADN. Esta técnica se basa en el uso de bisulfito sódico, el cual elimina el grupo amino a las citosinas, convirtiendo la citosina en uracilo. En el caso de las citosinas que no están metiladas, esta reacción se produce de una forma mucho más rápida que en el caso de las citosinas que sí que lo están. De esta manera, las citosinas metiladas seguirán siendo citosinas mientras que las no metiladas habrán cambiado a uracilo (99).

Cada uno de estos tratamientos previos a la detección de la metilación del ADN tiene ciertas ventajas y desventajas. La principal desventaja del uso de enzimas de restricción es que no se puede obtener información de la metilación para cada CpG. En el caso del enriquecimiento por afinidad, no se pueden obtener datos de la metilación a la resolución de una sola base. Por último, la conversión de bisulfito reduce la complejidad del genoma de cuatro pares de bases a solo tres, complicando su análisis con muchas técnicas (99).

El uso de todos estos pre-tratamientos se ha combinado con diferentes técnicas de biología molecular con mucho éxito, como las PCR, hibridación o la secuenciación para poder detectar y diferenciar citosinas metiladas de las no metiladas. Los diferentes métodos disponibles actualmente para el análisis del ADN son fruto de la combinación de diferentes métodos de pre-tratamiento junto a diferentes técnicas de análisis de la metilación (99).

a. Estudios de genes candidatos

Cuando se comenzó a estudiar la metilación del ADN, los análisis se centraban en regiones concretas de interés del genoma. Este tipo de estudio se conoce como estrategia de análisis de genes candidatos, en los que solo un gen o varios genes de interés son seleccionados para su análisis. Aunque en la actualidad los estudios de genes candidatos son menos comunes, en algunos casos concretos se siguen realizando.

A principios de los años 80 se publicó el primer estudio que analizó la metilación de un locus concreto mediante un análisis de metilación del ADN. En este análisis se utilizó la digestión con endonucleasas como pre-tratamiento previo al análisis de metilación, el cual se realizó mediante gel de electroforesis e hibridación con *southern blot* (una técnica de biología molecular que permite la detección de una secuencia específica de ADN en el gel de electroforesis). Posteriormente, las diferentes técnicas de pre-tratamiento mencionadas se fueron combinando con los *southern blot* con el objetivo de determinar la metilación del ADN. Sin embargo, este método fue sustituido por el análisis con PCR, mucho más sensible y específico para el análisis de la metilación del ADN en un solo locus (101).

La conversión de bisulfito también ha sido ampliamente utilizada en estudios de genes candidatos. Una de las aproximaciones ha sido el análisis de los productos de ADN obtenidos tras la conversión mediante secuenciación de Sanger (con una amplificación previa mediante PCR) (101).

Otro método utilizado en ensayos de genes candidatos es el de la combinación de la espectrometría de masa (MALDI-TOF) desarrollado por Sequenom con la plataforma conocida como EpiTYPER. En este caso, se requiere una amplificación del gen de interés después de la conversión de bisulfito. Este método permite obtener resultados cuantitativos fiables (101).

b. Estudios epigenómicos

Actualmente está más extendido el uso de una aproximación agnóstica para identificar patrones de metilación del ADN asociados a cierto fenotipo o enfermedad. En este tipo de estudios, contrariamente a los estudios de genes candidatos, se analizan un amplio espectro de genes o incluso todo el genoma, sin tener una hipótesis previa.

A lo largo de los años, se han ido desarrollando diferentes técnicas con este objetivo. Las primeras de ellas se centraron en el uso de enzimas de restricción como pre-tratamiento antes de la detección de la metilación del ADN. Una de las primeras fue el Escaneo de genoma de referencia de restricción o *Restriction Landmark Genoma Scanning* (RLGS). Este método permitía detectar diferencias en la metilación debido a diferencias en el patrón de fragmentos de restricción generados por una endonucleasa sensible a la metilación utilizando gel de electroforesis. También ayudaba a conocer el estado de metilación global del ADN genómico. Otras técnicas que también han utilizado como primer paso la digestión por endonucleasas han sido la PCR arbitrariamente sensible a la metilación (MS-AP-PCR) y la amplificación de puntos inter-metilados (AIMS). Sin embargo, estas técnicas se han ido sustituyendo por otras que no requieren el uso de geles de electroforesis y se han adaptado técnicas utilizadas en el análisis de la expresión génica basadas en el análisis por matrices o *arrays* (99). Los *arrays* consisten en un sustrato sólido al que se le unen diferentes oligonucleótidos que hibridan con cadenas de ADN complementarias (100). Algunos de estos métodos que unen la digestión por endonucleasa junto a un *array* son la Amplificación de Islas CpG Metiladas (MCAM) o la Hibridación de Metilación Diferencial (DMH) (99).

El uso de la conversión de bisulfito con el análisis posterior mediante *arrays*, ha supuesto un reto debido a que tras la conversión se obtienen secuencias de ADN menos complejas, lo que no permite adaptar correctamente las técnicas de hibridación. Sin embargo, existen dos claras excepciones en el uso de la conversión de bisulfito junto a los *arrays* por parte de la compañía Illumina. Ambas son adaptaciones de dos *arrays* de genotipado (el GoldenGate BeadArray y el Infinium) para permitir el análisis de la metilación mediante el genotipado del ADN convertido por bisulfito. El primero de ellos, permitió el análisis de hasta 1,536 CpG en un máximo de 96 muestras simultáneamente. Con este método se podía distinguir hasta el 17% de diferencias en la metilación global entre dos muestras (105). Posteriormente, en 2009, Illumina adaptó otro de sus ensayos, el de la plataforma Infinium para poder analizar la metilación del ADN de más de 27,000 CpGs con el *HumanMethylation27 DNA Analysis BeadChip*. Este fue un gran avance en el análisis de la metilación del ADN a nivel de todo el genoma ya que permitía determinar la metilación de 27,578 CpGs de 14,495 promotores de genes codificantes

por proteína y 110 promotores de miARN. Más adelante, se desarrolló el *Infinium HumanMethylation450* (450K) que permitía analizar más de 450,000 CpGs distribuidas a lo largo del genoma. En la actualidad está disponible el *Infinium HumanMethylation850* (EPIC), permitiendo el análisis de más de 850,000 CpGs. Aunque se han ido introduciendo pequeñas modificaciones, como la cobertura de más regiones reguladoras inter-génicas, los tres *arrays* de Infinium se basan en la misma tecnología. Esta técnica se basa en el genotipado cuantitativo de los polimorfismos C/T generados por conversión de bisulfito mediante hibridación con oligómeros de ADN, pasando por un paso de amplificación de todo el genoma (99,106).

Actualmente los avances en las tecnologías de secuenciación permiten el análisis de la metilación del ADN usando esta técnica y obteniendo información a nivel de un solo par de base. Es posible la combinación del método enzimático de digestión con el uso posterior de técnicas de secuenciación de nueva generación. Su principal ventaja es que no requiere de un *array* diseñado previamente, cubriendo una parte mayor del genoma con una cantidad inicial menor de ADN y permitiendo un análisis de la metilación del ADN específico por alelo (99). Un ejemplo de esta metodología es el Análisis digital de enzimas de restricción de metilación o *Digital Restriction Enzyme Analysis of Methylation* (DREAM). Con este método se pueden llegar a cubrir 50,000 CpGs únicas con una alta sensibilidad a niveles bajos de metilación (107). Los otros métodos de pre-tratamiento también se han unido a la posterior secuenciación para la detección de la metilación del ADN. Actualmente, los métodos que se basan en el enriquecimiento por afinidad se han unido a técnicas de secuenciación de nueva generación. Este método es muy óptimo cuando el objetivo es encontrar la metilación global del genoma, sin embargo no sirve cuando se necesita información de CpGs específicas (99). La conversión de bisulfito también se ha unido a las técnicas de secuenciación del ADN. El mayor problema en este caso también es la baja complejidad de la cadena de ADN después de la conversión de bisulfito. Para ello, una solución ha sido la unión de la técnica de captura en *arrays* o de captura en candado previo a la secuenciación. El problema del uso de la técnica de captura en *arrays* es que es relativamente ineficiente en regiones del ADN ricas en citosinas y guaninas y para que funcione correctamente necesita grandes cantidades de ADN o la realización de amplificación previa a la

hibridación, con lo que existe el riesgo de eliminar la metilación del ADN. La técnica de captura en candado permite incrementar la eficiencia de enriquecimiento y se ha utilizado junto a secuenciación de nueva generación en la técnica conocida como sondas de candado de bisulfito o *bisulphite padlock probes* (BSPP). Sin embargo, el estado de metilación también puede afectar a la eficiencia de captura y alterar los resultados (99).

Finalmente, la última técnica diseñada para analizar la metilación del ADN de todo el genoma es la Secuenciación de bisulfito de escopeta de genoma completo o *Whole-genome shotgun bisulphite sequencing* (WGSBS) que permite el análisis de la metilación del ADN a una resolución de un solo par de bases gracias a la secuenciación de alto rendimiento (108).

i. Epigenome-wide association studies (EWAS)

En 2008 se propuso el uso de los estudios de asociación del epigenoma o *Epigenome-wide association studies* (EWAS) como el equivalente a los GWAS para el estudio de la epigenética. El objetivo es obtener datos a nivel de todo el genoma de como los factores epigenéticos pueden influir en la complejidad de las enfermedades para poder saber en qué nivel la variación epigenómica afecta a la variación fenotípica. La marca epigenética más estudiada en los EWAS es la metilación del ADN ya que es la más estable y accesible. Además, la metilación del ADN correlaciona con otras variaciones epigenéticas, por lo que se puede llegar a hacer una extrapolación. En los EWAS, contrariamente a los estudios de genes candidatos, el objetivo es obtener datos con una cobertura de todo el genoma en un tamaño muestral adecuado con el uso de la tecnología mencionada anteriormente. La elección de la tecnología en un estudio EWAS concreto depende de una serie de factores como la cobertura, resolución, exactitud, rendimiento y coste deseados (95). Actualmente, el método más utilizado es el uso de la conversión de bisulfito junto a *arrays*. En concreto, el *array* 450K de Illumina ha sido el más utilizado en los estudios publicados, seguidos del EPIC, con menos publicaciones debido a su aparición posterior.

Los primeros estudios EWAS se realizaron en cáncer y permitieron descubrir que los tumores se asociaban a una ganancia de metilación del ADN en islas CpG, una pérdida

de la impronta genética y una remodelación epigenética de los elementos de repetición (109). Actualmente se han estudiado mediante EWAS alrededor de 230 enfermedades, identificándose más de 620,000 asociaciones (110).

La variación en la metilación del ADN que se estudia en los EWAS, equivalente a los SNPs de los GWAS, es la variación de la metilación en una sola CpG y se conoce como posición variable de metilación (MVP). También se puede estudiar la metilación diferencial en una región (DMR), en la que en lugar de estudiar solo una CpG se analizan diferentes CpG adyacentes, típicamente en menos de 1 kilobase (kb) de distancia. Las MVP y DMR se analizan principalmente en unos puntos concretos del genoma: en promotores, islas CpG (de unos 1,000 pares de bases) y regiones de impronta. Aunque también pueden ser analizadas en “orillas de islas CpG” o *CpG shore* (a una distancia máxima de 2kb de una isla CpG) o “estantería de CpGs” o *CpG shelf* (entre 2kb y 4kb de distancia de una isla CpG). Por último se pueden analizar también las CpG en “mar abierto” u *open sea*, que significa que están en el resto de genoma, lejos de islas CpG (109).

Algunas características propias de la epigenética hacen que el diseño de los análisis EWAS tenga que ser cuidadoso. Por ejemplo, el hecho de que la epigenética sea diferente en cada célula o tejido, que sea cambiante a lo largo del tiempo o que pueda ser causa o consecuencia de un fenotipo. También hay que tener en cuenta el tipo de cohorte que se incluirá en el estudio. Se puede elegir una cohorte de casos y controles, de solo casos (principalmente para el análisis de variables cuantitativas), de familias, de gemelos monocigóticos discordantes en un fenotipo o una cohorte prospectiva con seguimiento longitudinal. El tipo de cohorte más fácil de conseguir es la de casos/controles o solo casos. Sin embargo, con estas cohortes será más complicado ajustar por factores confusores como la genética o el ambiente. En las cohortes de familias, se puede estudiar la herencia. En las cohortes de gemelos se puede controlar por la genética y en las cohortes de seguimiento longitudinal se puede establecer causalidad. Sin embargo, son tipos de cohorte difíciles de conseguir (95).

En el caso de la elección del mejor tejido para realizar el estudio epigenético, hay que tener en cuenta a que tipo celular afecta el fenotipo que se quiere estudiar y también la accesibilidad que este tiene. La mayoría de EWAS en humanos se realizan en personas vivas, por lo que no todos los tejidos son fáciles de obtener. En el caso de las variaciones

en la metilación del ADN que son heredadas o producidas durante las primeras fases de la embriogénesis, se van a encontrar en todas las células somáticas y, por lo tanto, cualquier tejido, como la sangre, será óptimo para su análisis. La sangre es el tejido más estudiado debido a su accesibilidad, a pesar de que en una enfermedad determinada el principal tejido afectado no sea ese. Se ha observado en algunos estudios que existe una correlación entre la metilación del ADN de diferentes tejidos. También hay que tener en cuenta que en la mayoría de los tejidos existe variabilidad celular, que cada uno de estos tipos celulares puede tener unas variaciones epigenéticas específicas y que una enfermedad concreta puede estar afectando un tipo celular específico. Sin embargo, se han desarrollado diferentes métodos para estimar y ajustar los análisis de EWAS por los diferentes tipos celulares presentes en la muestra estudiada (95). Además, en 2018 se publicó un nuevo algoritmo, Epidish, el cual, además de estimar el conteo celular permite predecir el tipo celular específico asociado a los niveles diferenciales de metilación encontrados en el EWAS (111).

Otra consideración importante en los estudios EWAS para el cálculo del tamaño muestral y el poder estadístico, es que la metilación es una variable cuantitativa que considera la media de la metilación en las diferentes células presentes en una muestra, de la metilación en los diferentes alelos y la metilación en cada cadena de ADN (95). Esta metilación se puede cuantificar con los valores beta o valores M. Los valores beta puntúan del 0 (no metilación) a 1 (metilación total) y son la ratio entre la intensidad de las sondas metiladas y la intensidad total de las sondas analizadas. Los valores beta se transforman en valores M cuando se considera la ratio entre la intensidad de las sondas metiladas y la intensidad de las sondas que no están metiladas. En este caso, los valores obtenidos pueden ser negativos (baja metilación) o positivos (alta metilación) (100).

También es importante contar con un buen poder estadístico en los estudios EWAS. Este se puede definir como la probabilidad de que una prueba estadística rechace la hipótesis nula cuando la hipótesis alternativa es la correcta, evitando falsos positivos (error tipo I) y falsos negativos (error tipo II). Para tener un buen poder estadístico es importante considerar de forma correcta el umbral de significancia estadística y tener un correcto tamaño muestral. Al considerar el umbral de significancia estadística, hay que tener en cuenta el factor de testeo múltiple (se consideran muchas CpG en un solo análisis) y la

correlación existente entre las diferentes CpG analizadas. Se han realizado diferentes estudios de simulación para estimar el mejor umbral de significancia en los estudios EWAS y se considera que debería ser de un valor de $P < 9.4 \times 10^{-8}$ en análisis con el *array* EPIC y $P < 2.4 \times 10^{-7}$ en los análisis con un número de CpGs equivalente al del *array* 450k. En cuanto al tamaño muestral, se recomienda incluir aproximadamente una muestra de 1,000 sujetos para poder detectar DMPs y DMRs significativos en los estudios caso-control. Aunque no hay una recomendación clara, en estudios longitudinales o en los que incluyen solo casos y se analiza una variable cuantitativa, los estudios publicados demuestran que el tamaño muestral puede ser menor (100).

Los principales factores confusores en los EWAS son los factores técnicos (conocidos como efecto de *batch*) y los factores ambientales, que pueden afectar tanto a la metilación del ADN como al fenotipo. También hay variables que se asocian de forma independiente al fenotipo analizado. Es importante en los estudios EWAS ajustar los análisis por aquellas variables que pueden ser confusoras como el *batch*, el tabaco, la edad y el sexo (95).

Aunque algunos estudios EWAS se han realizado en una sola fase, es importante incluir una segunda fase de replicación, con una cohorte independiente pero similar. Los diferentes problemas que se han mencionado, como la existencia de factores confusores o la causalidad reversa, hacen que sea recomendable realizar los estudios EWAS en dos fases (95).

Para poder analizar los EWAS de una forma eficiente, se necesitan herramientas bioinformáticas potentes que permitan el análisis e interpretación de los datos generados (100). Existen diferentes paquetes en el lenguaje estadístico R que permiten realizar el análisis EWAS de inicio a fin a partir de los *arrays* de Illumina. Algunos de los más utilizados son Minfi, ChAMP y Meffil. Estos paquetes permiten realizar los diferentes pasos estándares en el análisis de los EWAS: importación de datos crudos, controles de calidad de las muestras y de las sondas analizadas, normalización, corrección por efecto de *batch*, ajuste por diferentes tipos celulares (este paso se puede realizar posteriormente) y análisis de DMP y DMR (ajustando por variables confusoras y por la proporción de los diferentes tipos celulares, si no se ha hecho previamente).

Además, con estas herramientas también es posible realizar algunos análisis funcionales *in silico*.

e. Estudios epigenómicos en el ictus

i. Estudios de ARNnc y de modificaciones de histonas en el ictus

La relación entre los ARNnc y el ictus se ha estudiado mediante modelos *in vitro*, *in vivo* en animales y en humanos. Los ARNnc más estudiados y para los cuales se ha encontrado asociación con la patogénesis del ictus son los ARNnc l y los miARN (94). Los ARNnc son considerados buenos biomarcadores diagnósticos, pronósticos y terapéuticos por ser estables, específicos y reproducibles. Además, se consideran que pueden ser buenas dianas terapéuticas. Es por este motivo que han sido ampliamente estudiados en el ictus (112).

En el caso de la asociación entre las modificaciones de las histonas y el ictus, los principales estudios publicados se han realizado en modelos animales y celulares. Además, son estudios que han analizado en estos modelos el efecto de la isquemia en las modificaciones de las histonas poniendo principalmente el enfoque en su efecto en la evolución después del ictus (113).

1. Estudios en relación con el riesgo de ictus isquémico

a. *ARN no codificante (ARNnc)*

Los resultados obtenidos de los diferentes estudios con ARNnc e ictus demuestran que existe una expresión diferencial de diferentes ARNnc l y miARN entre ictus y controles. Estos datos han ayudado a establecer un papel de los ARNnc en la patogénesis del ictus pero también sugieren que pueden ser buenas dianas terapéuticas (94).

En el caso de los ARNnc l, están bien descritos en diversos estudios un total de nueve ARNnc l (ANRIL, MALAT1, N1LR, MEG3, H19, C2dat1, FosDT, SNHG14 y TUG1) asociados al riesgo de ictus. Su asociación con el ictus isquémico parece debida a su acción en la apoptosis celular, inflamación, muerte celular y angiogénesis. Sin embargo, la mayoría de estos estudios se han realizado en modelos celulares y animales de isquemia y no todos ellos se han encontrado asociados en estudios en humanos (92,114).

A continuación, se expondrán resultados obtenidos en estudios en humanos que han analizado la diferente expresión de ARNncl entre casos de ictus y controles mediante una aproximación sin hipótesis previa, con el uso de *arrays* o secuenciación de ARN mediante ARN-seq. Estos análisis sugieren que la regulación por parte de los ARNncl en el ictus se puede producir a través de la modulación de los miARN, ARNm y de vías asociadas con el ictus isquémico. Un estudio realizado por Dykstra-Aiello et al., identificó 300 y 97 ARNncl con diferente expresión entre casos de ictus y controles masculinos y casos de ictus y controles femeninos, respectivamente. Estos ARNncl diferencialmente expresados se encontraron en genes previamente conocidos por ser de riesgo para el ictus como lipoproteínas, *ABO* o prostaglandinas (115). En el estudio de Guo et al., en el que utilizó un *array* de ARNncl, se identificaron 560 y 690 ARNncl que tenían niveles inferiores o superiores, respectivamente, en casos de ictus vs controles. Tres de estos ARNncl fueron validados mediante PCR: lncRNA-ENST00000568297, lncRNA-ENST00000568243, y lncRNA-NR_046084 (94,116).

En otro estudio que utilizó la técnica de RNA-seq en muestras de exosomas (vesículas extracelulares que contienen ARN, entre otros) se encontró una expresión diferencial de los ARNncl lnc-CRKL-2, lnc-NTRK3-4, RPS6KA2-AS1 y lnc-CALM1-7 asociada al riesgo de ictus isquémico leve (con un NIHSS inferior a 4) (94,117). Es importante el estudio de la composición de los exosomas en relación al ictus, debido al papel que estos juegan en la comunicación inter-celular y su participación en procesos como la remodelación cerebral (118).

En estudios en humanos también se ha observado que los ARNncl están implicados en la transición de la fase aguda a la sub-aguda del ictus mediante la regulación del sistema inmunológico, el cual pasa de estar activado a suprimirse sistémicamente. Se identificaron 3,009 y 2,034 ARNncl diferencialmente expresados a las 24 horas y a los siete días post-ictus, respectivamente. Mediante un análisis de enriquecimiento de vías metabólicas se encontró un aumento en la expresión de genes pertenecientes a la vía de la presentación y procesamiento de antígenos mientras que esta vía se normalizaba a los siete días después del ictus. Otras vías metabólicas, como la GABAérgica, se alteraba a los siete días después del ictus (119).

Además, algunos estudios han analizado la diferente expresión de ARNncl en algunos tipos celulares concretos, como las células mononucleares de sangre periférica (120). Se ha encontrado un aumento de linc-DHFRL1-4, SNHG15, y linc-FAM98A-3 en pacientes con accidente isquémico transitorio (AIT) (una interrupción del flujo sanguíneo cerebral breve, con una duración de los síntomas < 24h) y controles. Es importante que algunos de estos ARNncl se asocian con la dinámica del ictus, como se identificó en el estudio previo, siendo altos solo en algunas fases después del ictus. Por ejemplo, el linc-FAM98A-3 volvía a niveles normales a partir del día 7. Utilizando ARN-seq, otro estudio identificó que la expresión de los ARNncl era diferente a las 24h y a los 7 días (94).

Más recientemente, se han realizado análisis utilizando una aproximación diferente. Wang et al., publicó un estudio de redes de co-expresión génica para encontrar módulos de ARNncl asociados al ictus. De entre todos los ARNncl analizados (4,838), los ARNncl más significativos pertenecían a las vías de la infección por el virus HTLV-1 y la vía de señalización mTOR (121).

En el caso de los miARN, se han realizado muchos estudios que los han asociado al riesgo de ictus isquémico. Sin embargo, si nos centramos en estudios que hayan usado una tecnología de *arrays* o ARN-seq en humanos, el número de estudios es más limitado (114).

Uno de los primeros estudios que evaluó mediante *array* el perfil de expresión de diferentes miARN en ictus fue publicado por Tan et al., en 2009. En este estudio se identificaron 157 miARN (de un total de 836 que se analizaron) con diferente expresión entre casos de ictus jóvenes (con edad inferior a 50 años) y controles. Además, se analizó la asociación de estos miARN por subtipos etiológicos de ictus y también se encontró asociación. En concreto, se identificaron 8 miARN con menor expresión en los tres subtipos de ictus (aterotrombótico, lacunar y cardioembólico) y 17 miARN con mayor expresión en todos los subtipos (122). El mismo autor publicó más tarde otro estudio en el que se incluyeron pacientes con ictus isquémico que eran jóvenes, pero en este caso, no presentaban factores de riesgo. Se identificaron 21 miARN con una expresión similar en todos los casos de ictus. De ellos, cuatro miARN presentaron niveles más bajos en casos respecto a muestras controles: miR-25*, miR-34b, miR-483-5p y miR-498. En este estudio también se encontraron diferencias en la expresión de miARN por subtipos

específicos de ictus que sugieren que el uso de miARN podría ser válido para distinguir y clasificar a los pacientes con ictus en base a su subtipo etiológico (123). Además, este estudio también permitió demostrar que los pacientes con ictus con factores de riesgo y sin ellos pueden tener mecanismos moleculares diferentes implicados en su patogénesis (123).

Otro estudio publicado por Jickling et al., identificó 8 miARN con expresión diferencial en sangre periférica en pacientes con ictus isquémico agudo comparado con controles. En concreto, los miR-122, miR-148a, let-7i, miR-19a, miR-320d y miR-4429 se encontraron disminuidos y los miR-363 y miR-487b se encontraron sobre-expresados en pacientes con ictus. Estos miARN regulan la expresión de genes involucrados en el sistema inmunitario, extravasación de leucocitos y procesos trombóticos (124). Otros estudios han encontrado una disminución del miR-320 en pacientes con ictus isquémico (114). En el caso del estudio publicado por Wang et al., en el que se usó un *array* de miARN, también vieron una disminución del miARN circulante miR-320d en pacientes con ictus agudo comparado con controles. Además, en este estudio también se encontró el miR-320e con niveles más bajos en pacientes con ictus. Otros dos miARN circulantes, miR-106b-5p y miR-4306, se encontraron sobre-expresados en pacientes con ictus agudo (125).

En otro estudio, el cual también utilizó *arrays* de miARN, se identificaron 115 miARN con diferente expresión en suero en pacientes con ictus isquémico comparado con controles. Entre los miARN con más diferencia en la expresión entre grupos, tres miARN fueron validados: miR-32-3p, miR-106-5p, y miR-1246. Cabe destacar que el miR-106b-5p había sido identificado en el estudio realizado por Wang et al. Mediante el análisis de vías metabólicas, se encontró que los genes regulados por estos miARN parecían estar implicados en la vía de señalización de MAPK (como el miARN let-7e-5p identificado en el estudio anterior), guía de axones, glioma, etc. (94,126).

Mediante el uso de *array*, el miARN let-7e-5p se encontró incrementado en pacientes con ictus isquémico respecto a controles. Estudios posteriores realizados con modelos celulares y predicción bioinformática sugirieron una inhibición de la expresión de genes de la vía de las MAPK por parte del miARN let-7e-5p (94,127), igual que el estudio previo, el cual también identificó que esta vía estaba alterada por miARN (94,126).

También se han publicado estudios diseñados para identificar miARN que puedan diferenciar el riesgo de diferentes etiologías de ictus. En el estudio de Sepramaniam et al., se demostró que un total de 32 miARN eran capaces de distinguir entre los diferentes subtipos etiológicos de ictus isquémico (94,128). Además, se encontraron 58 miARN con una expresión disminuida y 47 miARN con una expresión aumentada en pacientes de ictus respecto controles. También se observó que la regulación de estos miARN era diferente en la fase aguda y la fase de recuperación del ictus. La asociación independiente de cinco de estos miARN con el riesgo de ictus fue validada (miR-125b-2*, miR-27a*, miR-422a, miR-488 and miR-627).

La expresión diferencial de miARN en asociación al ictus isquémico también ha sido evaluada en otros tejidos, en células específicas, como las células mononucleares de sangre periférica y en vesículas concretas como los exosomas. El perfil de miARN pudo ser evaluado en líquido cefalorraquídeo (LCR) usando un *array* de 378 miARN. Se identificaron 183 miARN diferencialmente expresados entre pacientes con ictus y pacientes con otro tipo de enfermedades neurológicas. En total, dos miARN (let-7c y miR-221-3p) estaban sobre-expresados en pacientes con ictus. En este mismo estudio también se analizó el tejido sanguíneo y los resultados fueron diferentes: un total de dos miARN se encontraron sobre-expresados (miR-151a-3p y miR-140-5p), mientras que uno de ellos estaba disminuido en pacientes con ictus (miR-18b-5p) (129). En el estudio realizado en células mononucleares de sangre periférica, se identificaron 29 miARN con diferente expresión en pacientes con ictus respecto a controles en este tipo celular. En concreto tres miARN se encontraron sobre-expresados en todos los pacientes con ictus: miR-4656, miR-432, miR-503 y uno infra-expresado en todos los pacientes: miR-874 (130). En el caso de los miARN presentes en exosomas, el miR-146a se encontró que regulaba la microglía y los macrófagos en el ictus isquémico, además de proteger enfrente el daño cerebral.

Teniendo en cuenta la utilidad que los miARN pueden tener como biomarcadores no invasivos para predecir el riesgo de ictus, un estudio realizado por Sonoda et al., en 2019 desarrolló un modelo en el que combinaba tres miARN (miR-1268b, miR-4433b-3p, y miR-6803-5p) identificados en su estudio. Este modelo permitía diferenciar a pacientes

con ictus respecto a controles con una sensibilidad del 84% y especificidad del 98% (131).

b. *Modificación de histonas*

Como no existen estudios analizando modificaciones de histonas en pacientes con ictus, a continuación, se detallarán algunos de los resultados obtenidos en modelos animales de ictus. Diferentes estudios han identificado que después de producirse una isquemia cerebral, se reduce la expresión génica debido a una disminución de la acetilación de las histonas H3 y H4. La acetilación de las histonas tiene un papel muy importante en el ictus, concretamente en la regulación de la plasticidad de la cromatina y de la expresión génica. En estos análisis se ha observado que existe una deacetilación de las histonas H3 y H4 en las neuronas de la penumbra (región cerebral en la que se observa una disminución del flujo sanguíneo debido al ictus, pero con una afectación potencialmente reversible) debido a una sobre-expresión de HDACs.

La metilación de las histonas es otra modificación frecuente que conlleva cambios en la expresión génica. Sin embargo, en relación con el ictus, hay menos estudios que hayan evaluado el papel de esta modificación. En concreto, se ha identificado una asociación entre la metilación en las histonas H3 en tres residuos de lisina diferentes (4, 9 y 27) asociados con diferentes procesos implicados en el ictus. La metilación en H3K4 está principalmente regulada por la demetilasa 1 (LSD-1). En modelos de isquemia en rata existe una regulación específica de LSD-1 después de la isquemia y esta juega un papel en la función neuronal. A parte de la metilación de H3K4 en tejido cerebral, se ha encontrado que tras una isquemia en modelo animal de ictus, los niveles de metilación de H3K4 disminuyen significativamente en riñón (132). En cuanto a la metilación en H3K9, se ha asociado con represión de la transcripción. En modelos de privación de oxígeno y glucosa en células, la inhibición de la metilación de H3K9 se ha observado que conlleva un aumento de la transcripción asociada con supervivencia neuronal. La regulación de la metilación de H3K9 también se ha asociado con el ictus por su control de la respuesta inflamatoria. En concreto, una reducción de la metilación de esta histona en la lisina 9 se encontró que se asociaba a un aumento de la expresión de genes pro-inflamatorios. Esta regulación de la inflamación en modelos de ictus también está regulada por la metilación en la lisina 27 de la H3. Utilizando modelos animales de rata

se encontró una demetilasa de H3K27, JMJD3, sobre-expresada después de la isquemia, disminuyendo así la metilación de la lisina 27 de las H3 en los promotores de los genes Caspasa-3 y Bax. Esto se asoció con un aumento de la apoptosis y los déficits neurológicos después de la isquemia.

2. Estudios en relación con el riesgo de ictus hemorrágico

a. *ARN no codificante (ARNnc)*

Mediante el uso de *arrays*, Hanjin et al. analizó el perfil de expresión de ARNnc en un modelo de ratas de HIC. Se encontraron 625 ARNnc diferencialmente expresados, los cuales, según el análisis de vías metabólicas, podrían estar implicados en los cambios fisiopatológicos provocados por la HIC a largo plazo (133).

En otro estudio en el que se evaluó el perfil de expresión de miARN en sangre en pacientes con HIC y controles, se identificaron un total de 59 miARN disminuidos en pacientes con HIC. En este estudio también se analizaron específicamente los niveles de miR-21-5p en sangre, pero también en muestras de hematoma. Se vio que estaba disminuido en ambos tipos de muestra (134). Los niveles de miR-21-5p también se han correlacionado con los valores de NIHSS y la evolución de los pacientes con HIC (135).

Guo et al., utilizó un *array* para identificar miARN asociados con la HIC. En este estudio se identificaron 30 miARN con niveles más elevados en plasma de pacientes con HIC, con una sobrerrepresentación de la regulación de la inflamación entre estos miARN (136). En este estudio también se compararon los niveles de miARN entre pacientes con HIC e ictus isquémico. Esta diferenciación es muy importante en la fase inicial del ictus, previo a la realización de pruebas de neuroimagen, para poder administrar el tratamiento adecuado de la forma más rápida posible. Se encontraron cuatro miARN (miR-27a, miR-365, miR-150, y miR-34c-3p) diferencialmente expresados entre ambos tipos de ictus (136).

Leung et al., también comparó pacientes con ictus hemorrágico vs pacientes con ictus isquémico. En su estudio, se encontraron niveles más elevados en plasma del miR-124-3p (que eran más elevados cuando más severa era la HIC) y niveles disminuidos de miR-16 en la fase aguda en pacientes con ictus hemorrágico comparado con ictus isquémico.

Los niveles de miR-124-3p también correlacionaron con el tamaño de la hemorragia (94,137).

En el caso de miARN en ictus hemorrágico, recientemente, en 2022, se publicó una revisión con los diferentes artículos existentes en relación a miARN e HIC en humanos (138). En esta revisión se describen un total de siete artículos que han incluido 305 pacientes con HIC y 165 controles. Esto ha permitido la identificación de 152 miARN que podrían ser biomarcadores para HIC. También se han publicado dos artículos de cohortes que han incluido 1,959 individuos y ha permitido identificar 47 miARN diferencialmente expresados en pacientes con HIC. En uno de estos estudios de cohortes, se incluyeron pacientes sin ictus y se les siguió durante un periodo de 9,8 años de media. Se identificaron tres miARN (miR-6124, miR-5196-5p, y miR-4292) capaces de predecir el riesgo de sufrir un ictus hemorrágico en la población general (139). Finalmente, un estudio transversal identificó en 1,523 participantes un total de 10 miARN alterados (138).

En un estudio realizado en exosomas, se encontró que el miR-146a-5p estaba disminuido en casos de HIC, con un efecto protector en este subtipo de ictus. El miARN-126 exosomal también se encontró disminuido en pacientes con HIC (118,140).

b. *Modificación de histonas*

Como en el caso del ictus isquémico, los diferentes estudios publicados en relación con la regulación de las modificaciones de las histonas en el ictus hemorrágico son en modelos animales. Además, la mayoría de estos estudios se han centrado en la asociación de las modificaciones de histonas con la evolución de las HIC, por lo que los diferentes resultados obtenidos se expondrán en el apartado de estudios relacionados con el pronóstico del ictus.

En un estudio realizado por Wang et al., se encontró que la HDAC10, inhibidora de la acetilación de histonas de tipo II, era protectora frente a las HIC y sus niveles estaban aumentados en tejido cerebral humano y de rata después de una HIC (141). Su inhibición en un modelo de rata produjo un empeoramiento neurológico, edema cerebral, permeabilización de la barrera hematoencefálica, inflamación, etc (141).

3. Estudios en relación con el pronóstico del ictus

a. ARN no codificante (ARNnc)

Algunos estudios en humanos han analizado la relación entre ARNnc y el pronóstico del ictus. Sin embargo, estos estudios se han realizado utilizando la estrategia de ARNnc candidatos. Wang et al., encontró una correlación positiva entre los niveles de H19 y la NIHSS a los siete, 30 y 90 días después del ictus. Estudios posteriores en animales observaron que este ARNnc evitaba la neurogénesis después de un ictus a través de la regulación de la vía de p53/Notch1 (142). El ARNnc H19 también se asoció específicamente con el riesgo de ictus aterotrombótico, a través de un incremento de ACP5 (94).

Los niveles de otro ARNnc en suero, ANRIL, se correlacionaron con la NIHSS inicial y el mRS a los 6 meses después del ictus (143). Este ARNnc se había asociado con el riesgo de ictus isquémico y de ictus isquémico en pacientes con FA (143). Sin embargo, hay resultados contradictorios acerca de si los pacientes con ictus presentan un aumento o disminución de la expresión de ANRIL (94).

Zhu et al., analizó los niveles del ARNnc MIAT en asociación con la NIHSS basal y el mRS a los tres meses después del ictus. Se encontró una correlación positiva de los niveles de este ARNnc con la NIHSS basal y con el mRS. Mediante análisis multivariado, se encontró que este ARNnc era un marcador pronóstico independiente de la evolución funcional y de la muerte después de un ictus. Además, comparado con controles sanos, MIAT estaba sobre-expresado en pacientes con ictus isquémico (94,144).

Los miARN también se han estudiado en relación con la evolución después de un ictus. Los estudios previos en modelos animales sugerían una asociación de diferentes miARN en la regulación de la expresión génica para promover plasticidad neuronal, por eso, era importante hacer también estudios en humanos. Los diferentes estudios sugieren un papel de estas moléculas en procesos relevantes que ocurren después de un ictus como la proliferación, hematopoyesis, metabolismo, función inmunológica y depresión (114).

En el estudio de Tan et al., comentado previamente, se hizo un análisis para ver el perfil de miARN en los pacientes con ictus, teniendo en cuenta los diferentes subtipos y la evolución del ictus en base al mRS al alta. Los pacientes con ictus aterotrombótico y

cardioembólico que tenían una buena evolución post-ictus ($mRS < 2$) se agruparon conjuntamente según la expresión de los diferentes miARN analizados. De forma similar, los pacientes con ictus de los diferentes subtipos con una mala evolución post-ictus ($mRS > 2$), tenían niveles de expresión de miARN similares. Como excepción, se encontró que los pacientes que tuvieron una buena evolución, pero tenían un ictus de tipo lacunar, se agruparon de forma independiente a los demás (122). En el estudio publicado posteriormente por el mismo autor, se identificaron 27 miARN diferencialmente expresados entre pacientes con buena evolución ($mRS \leq 2$) y pacientes con mala evolución ($mRS > 2$). Cabe destacar que el tamaño muestral era muy pequeño y solo había un paciente con un $mRS > 2$ ($mRS = 4$) (123).

En el estudio de Edwardson et al., se analizó la expresión de miARNs en plasma en relación con la mejoría funcional de un déficit motor en las extremidades superiores. Se encontró que los miR-371-3p, miR-524, miR-520g, miR-1255A, miR-453, y miR-583 disminuían su expresión en pacientes que presentaban mejoría, mientras que los miR-941, miR-449b, y miR-581 aumentaban su expresión (94,145).

El miR-210, un miARN que se sabía que tenía asociación con isquemia cerebral en modelos animales, fue estudiado en humanos por Zeng et al. para encontrar si existía correlación con la evolución de los pacientes después del ictus. En humanos se encontró una disminución de este miARN en pacientes con ictus, especialmente a los 7 y 14 días después del evento. Además, este miARN correlacionó positivamente con la evolución de los pacientes después del ictus (114,146).

A parte de estos estudios, se han publicado otros análisis con el objetivo de identificar miARN asociados con otros fenotipos relacionados con la evolución del ictus, como el deterioro cognitivo o la depresión post-ictus.

El deterioro cognitivo después del ictus es una complicación común que también marca la evolución del ictus. En el estudio publicado por Huang et al., se estudió un miARN concreto, el miR-132, ya que previamente se había asociado con empeoramiento cognitivo y con la plasticidad sináptica. Se encontró una elevación significativa de los niveles de miR-132 en los pacientes con ictus que presentaban deterioro cognitivo comparado con controles y el AUC para este miARN fue de 0.961. Además, los niveles

de miR-132 se encontraron correlacionados con los resultados obtenidos de un test neurocognitivo (el test MoCA) (147).

La depresión después de los ictus es una complicación común que puede ocurrir tanto en la fase aguda (las dos primeras semanas después del ictus) o incluso años después (148). En un estudio que analizó la asociación de 1,347 miARN con el riesgo de depresión post-ictus (a las 2 semanas), se encontraron un total de 25 miARN diferencialmente expresados entre los pacientes con depresión temprana. Solo cuatro de ellos se encontraron con mayor expresión en pacientes con depresión (miR-22-3p, miR4476, miR-486-5p y miR-92a-3p). Entre los genes diana de los miARN diferencialmente expresados había un enriquecimiento de las vías del cáncer y de las MAPK (148). En otro estudio realizado en pacientes con depresión más tardía (a los 3 meses post-ictus), se encontraron otros tres miARN diferentes significativamente asociados con este fenotipo (miR-140-5p, miR-221-3p y miR-1246). El miR-221-3p y el miR-140-5p se habían encontrado previamente asociados al riesgo de ictus en LCR y sangre, respectivamente, en el estudio de Sorensen et al. (129). En un análisis de predicción de las dianas de estos miARN, se identificaron las vías de señalización de Ras y de desarrollo axonal (149).

En cuanto a la evolución de pacientes con HIC, también se han realizado diferentes estudios con ARNnc. En un estudio se compararon los niveles de miARN a nivel de todo el genoma entre pacientes con HIC y crecimiento del hematoma y pacientes sin este crecimiento. Un total de 30 miARN se identificaron con capacidad de diferenciar a los pacientes con crecimiento o sin crecimiento del hematoma (150). De forma similar, también se ha evaluado la asociación de miARN con la formación de edema alrededor del hematoma. Se identificó al miR-130a elevado en suero de pacientes con HIC y además como un predictor independiente de la formación de edema. Este miARN también se asoció con la NIHSS a los 14 días y el mRS a los tres meses en pacientes con HIC profunda (151). En otros dos estudios, el miR-23a-3p (152) y el miR-126 (153) también correlacionaron con el volumen del edema de alrededor del hematoma, lo que predispone a los pacientes a una peor evolución post-HIC.

Actualmente hay ensayos clínicos en marcha para estudiar la eficacia de la regulación de miARNs específicos para mejorar la evolución de los pacientes con ictus (154).

b. Modificación de histonas

La modificación de las histonas se ha asociado con el control de diferentes funciones relacionadas con el ictus isquémico como es el estrés oxidativo, la inflamación, la supervivencia neuronal, o procesos como la función de aprendizaje. En estos estudios con modelos animales se ha encontrado que cuando se produce un aumento en la acetilación de la histona H4, esta se asocia a regeneración de tejido y crecimiento axonal lo que lleva a una resistencia de las neuronas a la isquemia y favorece la recuperación funcional después del ictus.

Diferentes estudios sugieren que la inhibición de las HDAC puede dar lugar a un aumento de la neurogénesis y angiogénesis que ayude a promover la recuperación funcional después de la isquemia. Se ha sugerido que los inhibidores de deacetilasas podrían ser buenos fármacos neuroprotectores (113).

Se han realizado diferentes estudios en modelos animales que han servido para probar el efecto de inhibidores de las HDAC para el tratamiento del ictus. En roedores, se ha estudiado como este tratamiento protege del daño isquémico, apoptosis e inflamación, entre otros procesos que suceden en el ictus. Gracias a estos estudios, se han desarrollado ensayos clínicos basados en el uso de inhibidores de la acetilación de histonas (inhibidores de HDACs), principalmente del tipo I y II (155). Los primeros ensayos se realizaron con inhibidores no selectivos de HDACs, como el ácido valproico, y aunque demostraron ser neuroprotectores y actuar en contra de reacciones inflamatorias nocivas producidas en los ictus, también se encontró una alta toxicidad debido a una demetilación del ADN así como una acetilación de las histonas a nivel global (132). Por eso, posteriormente se fueron desarrollando estudios utilizando inhibidores selectivos de tipos concretos de HDAC. El problema en este caso, es que muchos de estos inhibidores selectivos de HDAC no traspasan la barrera hematoencefálica (90). Encontrar fármacos que inhiban específicamente algunas HDAC es importante para evitar la toxicidad, pero también porque cada HDAC actúa de forma diferente. Por ejemplo, se ha observado que la HDAC4 es reguladora positiva de la formación de memoria y puede ayudar a prevenir el deterioro cognitivo y los problemas de memoria en pacientes con ictus. Por lo tanto, no sería conveniente la inhibición de esta HDAC (91). Otro tipo de deacetilasas de histonas que también se han estudiado son

las sirtuinas (SIRT), que son deacetilasas de tipo III. Estas enzimas están implicadas en diferentes procesos como el metabolismo, la apoptosis, autofagia o crecimiento celular. Dos SIRT, la SIRT1 y la SIRT6 están implicadas en la regeneración del cerebro después de la isquemia (113). En un ensayo clínico (*International Citicoline Trial on Acute Stroke (ICTUS)*) en el que se administró un activador de SIRT1 (la citicolina) se encontró que podía ser eficaz en pacientes con ictus moderado sin tratamiento con rTPA (156). La citicolina induce neuroprotección mediante el aumento de los niveles de la proteína SIRT1. Esta activación de SIRT1 podría ser útil en el desarrollo de terapias en el ictus isquémico (156).

Además, también se ha observado que la metilación de las histonas podría estar implicada en los mecanismos que regulan la evolución después de un ictus. En el caso de la H3K4, su metilación podría estar relacionada con el pronóstico del ictus relacionado con la edad. En estudios realizados en ratas, se encontró que los animales de edad más avanzada presentaban tamaño de infarto menor con una disminución de la función de los astrocitos. Estos astrocitos en ratas mayores tenían más metilación de H3K4, relacionada con mayor activación transcripcional y más actividad astrocítica después de la isquemia (132). En pacientes con ictus se ha observado que la metilación en H3K4 junto a la acetilación de H3K9 se asocia con los niveles de TNF, una molécula pro-inflamatoria implicada en el pronóstico del ictus. El papel de la regulación epigenética mediante las modificaciones de histonas de la inflamación también ha sido ampliamente estudiado, debido a todos los procesos pro-inflamatorios que ocurren después de un ictus. La neuroinflamación se sabe que es un mecanismo que contribuye a la muerte neuronal, causando daño al tejido cerebral y empeorando la regeneración del tejido.

En otros estudios con modelos de isquemia cerebral se ha encontrado que un aumento de la metilación de H3K9 se asocia con una disminución de la expresión del receptor de glutamato 2, implicado en la progresión después de un ictus isquémico. Se están empezando a realizar estudios con inhibidores de metilasas o demetiladoras en modelos animales. Por ejemplo, el uso de GSK-126, un inhibidor de la actividad de EZH2, una histona metiltransferasa de H3K27, parece tener especificidad alta y podría ser una

buena opción terapéutica, incluso mejor que los inhibidores de HDAC. Sin embargo, hacen falta más estudios en esta área (132).

En el caso de la HIC, se han realizado estudios en los que se ha observado que la administración de inhibidores de HDAC reduce la inflamación y el daño cerebral en modelos de rata (91). Por ejemplo, se encontró que un inhibidor no específico de las HDAC, el ácido hidroxámico suberoilánilida (AHSa) o vorinostato, era capaz de mejorar el daño neurológico y la evolución funcional en ratón (157). De forma similar, otro inhibidor de HDAC (scriptaide), que modula la polarización de la microglía, se encontró que reducía el daño de la sustancia blanca, el tamaño del hematoma y mejoraba la función neurológica en un modelo de HIC de ratón (158,159).

Un estudio reciente sugiere que el uso de inhibidores de las HDAC de clase 1, concretamente de entinostato, puede ser efectivo en la reducción del volumen del hematoma y mejoría de la evolución después de una HIC en un modelo de ratón. Este efecto parece producirse por una disminución del número de neuronas degenerativas y por una reducción de las citoquinas pro-inflamatorias (160).

ii. Estudios EWAS de metilación del ADN en pacientes con ictus

1. Estudios en relación con el riesgo de ictus

El primer estudio que analizó la metilación global del genoma en relación con el ictus fue publicado en 2010. En este estudio, se analizó la metilación de unos elementos repetitivos del genoma (LINE-1) que se suelen encontrar muy hipermetilados. Alrededor del 55% del genoma consiste en elementos repetitivos, entre los que se encuentra LINE-1, por eso el análisis de la metilación de este elemento se ha utilizado como indicador de la metilación global del ADN (161). En este estudio se encontró una hipometilación global del genoma en pacientes con ictus. Además, se encontró que la hipometilación de LINE-1 se asociaba con un aumento en la incidencia de ictus. También identificaron que la hipometilación se asociaba con un aumento del riesgo de muerte por ictus (161). Posteriormente, mediante un ensayo luminométrico (LUMA) se comparó la metilación en muestras de sangre de pacientes con diferentes subtipos de ictus. LUMA es un método que permite conocer la metilación de citosinas que se encuentran en el motivo

-CCGG, el cual supone un 8% de todas las CpGs y se usa como estimación de la metilación global. En este estudio no se observaron diferencias en la metilación entre ictus aterotrombótico, lacunar y cardioembólico (103). Sin embargo, sí que se encontró que los pacientes con ictus presentaban una hipometilación global comparado con controles sanos, igual que en el estudio previo (103).

En 2014 se publicó un estudio en el que se analizó la metilación del ADN mediante el uso del *Illumina methylation 27 BeadChip array* en dos poblaciones: la primera consistió en 12 voluntarios con ictus y 12 sin ictus y posteriormente en una segunda que incluyó 60 ictus y 55 controles. En este estudio se encontró una CpG en el gen *PM20D1* hipermetilada en pacientes con ictus respecto a controles. Además, otra CpG en el gen *CALD1* se encontró diferencialmente metilada y presentaba interacción entre el riesgo de ictus y la obesidad (Tabla 2). Estos resultados muestran cierto solapamiento en la regulación epigenética de la obesidad y el riesgo de ictus (162).

En 2016 se publicó el primer estudio en pacientes con ictus utilizando el *array* de Illumina 450K. Sin embargo, en este estudio se analizaron pacientes con diferentes enfermedades cardiovasculares, entre ellas el ictus, las cuales se estudiaron de forma independiente con 491 controles sanos. No se encontró ninguna CpG cuya metilación fuera diferencial en los pacientes con ictus comparado con los controles. Sin embargo, el número de pacientes con ictus fue muy pequeño ($n = 27$), por lo que probablemente el poder estadístico no era suficiente para encontrar asociaciones significativas (163).

En 2020 se publicó el primer estudio EWAS que analizó la metilación del ADN usando el *array* 450K junto al *array* EPIC en pacientes con ictus, incluyendo el análisis por subtipos etiológicos vs controles. Se analizaron 401 muestras y los resultados obtenidos fueron posteriormente replicados en dos cohortes independientes ($N = 226$ y $N = 166$). Se identificaron 22 CpGs diferencialmente metiladas entre pacientes con ictus y controles (Tabla 2). Estas CpGs se encontraban en 21 loci, incluido el locus donde se localiza el gen *ZFHX3*, previamente asociado al riesgo de ictus cardioembólico mediante GWAS (164). En el análisis por subtipos de ictus se encontró que los ictus de tipo indeterminado, seguido por los de tipo lacunar, eran los subtipos que presentaban mayores cambios a nivel epigenético (164).

Tabla 2: Resumen de los loci asociados al riesgo o evolución del ictus mediante EWAS

LOCUS	FENOTIPO ICTUS	ARRAY	REFERENCIA
<i>PM20D1</i>	IS	27K	159
<i>CALD1</i>	IS-obesidad	27K	159
<i>ZFHX3</i>	IS	450K+EPIC	161
<i>PIM3</i>	IS	450K+EPIC	161
<i>MAPK1</i>	IS	450K+EPIC	161
<i>LRRC26</i>	IS	450K+EPIC	161
<i>HIF1A</i>	IS	450K+EPIC	161
<i>RNF126</i>	IS	450K+EPIC	161
<i>SENP3</i>	IS	450K+EPIC	161
<i>ANAPC11</i>	IS	450K+EPIC	161
<i>PLBD2</i>	IS	450K+EPIC	161
<i>CCNL2</i>	IS	450K+EPIC	161
<i>PUM1</i>	IS	450K+EPIC	161
<i>ZFHX3</i>	IS	450K+EPIC	161
<i>ITPKB</i>	IS	450K+EPIC	161
<i>NAPA</i>	IS	450K+EPIC	161
<i>IL15RA</i>	IS	450K+EPIC	161
<i>ACSL1</i>	IS	450K+EPIC	161
<i>JMY</i>	IS	450K+EPIC	161
<i>PUF60</i>	IS	450K+EPIC	161
<i>CHSY1</i>	IS	450K+EPIC	161
<i>BAMBI</i>	IS	450K+EPIC	161
<i>MTRNR2L8</i>	ATE	450K	162
<i>CDH2</i>	IS	EPIC	163
<i>PCDHB10/PCDHB11/PCDHB14/PCDHB16/PCDHB3/PCDHB9</i>	IS	EPIC	163
<i>ATP9B</i>	HIC	EPIC	169

<i>LPCAT1</i>	HIC	EPIC	169
<i>ADAP1</i>	HIC	EPIC	169
<i>CD247</i>	HIC	EPIC	169
<i>LOC360030</i>	HIC	EPIC	169
<i>CLMP</i>	HIC	EPIC	169
<i>S1PR1</i>	HIC	EPIC	169
<i>GMDS</i>	HIC	EPIC	169
<i>MOB3B</i>	HIC	EPIC	169
<i>BCL7A</i>	HIC	EPIC	169
<i>FLJ10213; PPP4R2</i>	HIC	EPIC	169
<i>HHEX</i>	HIC	EPIC	169
<i>TBC1D12</i>	HIC	EPIC	169
<i>CAT</i>	HIC	EPIC	169
<i>PLEC1</i>	HIC	EPIC	169
<i>CORO1C</i>	HIC	EPIC	169
<i>IL6R</i>	HIC	EPIC	169
<i>B3GAT1</i>	HIC	EPIC	169
<i>IL10</i>	HIC	EPIC	169
<i>WDFY4</i>	HIC	EPIC	169
<i>ASB10</i>	Recurrencia ictus	450K	170
<i>TTC37</i>	Recurrencia ictus	450K	170
<i>ZDHH6</i>	Tiempo recurrencia ictus	450K	171
<i>TRAF3</i>	Recurrencia ictus (Clopidogrel)	450K	172
<i>PPM1A</i>	Recurrencia ictus (Aspirina)	450K	173

<i>INSR</i>	Isquemia cerebral retardada (hemorragia subaracnoidea)	EPIC	177
<i>CDHR5</i>	Isquemia cerebral retardada (hemorragia subaracnoidea)	EPIC	177

IS: Ictus isquémico; ATE: Ictus aterotrombótico; HIC: Hemorragia intracerebral; 450k: Illumina HumanMethylation 450K array; EPIC: Illumina HumanMethylation 850K array

Tabla original de la doctoranda

También se publicó un EWAS realizado en pacientes que habían sufrido específicamente un ictus del subtipo aterotrombótico. Mediante el *array* 450K, se estudió una cohorte de 12 casos y 12 controles y los resultados fueron validados en dos cohortes mayores (una de 110 casos y 122 controles y otra de 191 casos y 191 controles). Se pudo replicar la metilación diferencial en el promotor del gen *MTRNR2L8*, codificante para un péptido señalizador y neuroprotector (165).

Recientemente, en abril de 2022, se publicó un estudio EWAS realizado en una cohorte de descubrimiento pequeña de 4 casos y 4 controles, y una replicación (realizada mediante MethylTargetTM, basada en la secuenciación dirigida a puntos de metilación concretos) de 188 casos y 188 controles, que permitió identificar, mediante el *array* EPIC siete genes diferencialmente metilados asociados al riesgo de ictus: *CDH2/PCDHB10/PCDHB11/PCDHB14/PCDHB16/PCDHB3/PCDHB9* (166).

La metilación del ADN influye en lo que se conoce como edad biológica, la cual difiere en muchas ocasiones de la edad cronológica. La edad biológica se puede calcular con diferentes métodos, como el de Hannum y Horvath (167,168) utilizando la información

de la metilación de determinadas CpGs. Según diferentes estudios, la edad biológica es un mejor indicador del estado real de envejecimiento de las personas. La edad biológica se ha estudiado en relación con el riesgo de ictus observándose que los pacientes con ictus presentan una mayor edad biológica, lo que equivale a un mayor envejecimiento biológico, respecto a controles sanos (169). Posteriormente, se encontró una asociación entre la aceleración de la edad (la diferencia entre la edad biológica y la edad cronológica) y el riesgo de ictus. Aquellos pacientes más envejecidos biológicamente respecto a su edad cronológica, tenían un riesgo más elevado de padecer un ictus (170).

Recientemente también se han encontrado diferencias en la edad biológica entre hombres y mujeres con ictus. A pesar de que los hombres y las mujeres difieren en la edad de inicio del ictus (las mujeres tienen ictus a edades más tardías), la edad biológica en el momento del ictus se encontró que era la misma entre ambos sexos. Esto ocurre porque las mujeres son biológicamente más jóvenes de lo esperado en el momento del ictus (lo que se conoce como desaceleración de la edad biológica) mientras que a los hombres les ocurre el efecto inverso. De esta manera, la edad biológica en el momento del ictus se iguala entre ambos sexos (171)

Solo hay un estudio que haya analizado la metilación del ADN en pacientes con HIC. En este análisis se usó el *array* EPIC de Illumina para analizar el patrón de metilación del ADN en sangre. Se incluyeron un total de 30 pacientes con HIC y 34 controles sanos. Se encontraron 1,377 CpGs hipermetiladas y 153 CpGs hipometiladas de forma significativa en pacientes con HIC respecto controles (Tabla 2: 20 CpGs más significativas). Mediante análisis de vías metabólicas se vio que la metilación diferencial se encontraba enriquecida en genes pertenecientes a vías inflamatorias (172).

2. Estudios en relación con la evolución post-ictus

Los estudios EWAS que han analizado la metilación del ADN en relación con la evolución del ictus, se han centrado en procesos que pueden ocurrir después del ictus y que contribuyen a la evolución y pronóstico de este.

Un determinante importante de la evolución después de un ictus es la recurrencia. La recurrencia tras un ictus es un evento frecuente que se asocia con discapacidad. La metilación del ADN asociada a la recurrencia de ictus se estudió en una cohorte de

pacientes participantes en un ensayo clínico (*Vitamin Intervention for Stroke Prevention, VISP*), en el que pacientes con ictus isquémico y niveles elevados de homocisteína se aleatorizaron a recibir una dosis diaria alta o baja de vitamina B6, B12 y ácido fólico. Se analizó la metilación asociada a la recurrencia de estos pacientes. Se encontraron dos CpGs significativas, una en el gen *ASB10* y la otra en el gen *TTC37* (173). Posteriormente, el mismo grupo publicó un estudio en el que analizó la metilación asociada al tiempo de recurrencia. Se identificó y replicó un gen, *ZDHHC6*, cuyo grado de metilación se asoció de forma positiva con el tiempo de recurrencia de un ictus pero también de cualquier evento de tipo vascular, tanto en población europea como africana (174).

Por otro lado, se sabe que algunos pacientes recurren a pesar de estar tratados con fármacos antiagregantes, como el clopidogrel y la aspirina. La resistencia a estos tratamientos es un factor conocido para el cual se han encontrado variantes genéticas asociadas. En el estudio de Gallego-Fabrega et al., se identificaron dos genes, *TRAF3* y *PPM1A*, cuya metilación estaba asociada al riesgo de recurrencia a un año en pacientes tratados con clopidogrel (175) y aspirina (176), respectivamente.

La edad biológica también ha sido estudiada en relación con la evolución tardía de los pacientes con ictus (a los tres meses). Se ha encontrado una asociación positiva entre la edad biológica y el valor de la escala mRS a los tres meses post-ictus. Por lo tanto, pacientes más envejecidos biológicamente, presentan una peor evolución. Este estudio permitió determinar que la edad biológica puede predecir mejor que la edad cronológica cual será esta evolución (177). Por otro lado, en otro estudio publicado por el mismo grupo se encontró una asociación significativa y positiva entre la mortalidad a los tres meses y la edad biológica, de forma independiente a otros factores conocidos asociados a la mortalidad, como la NIHSS, la edad cronológica, el mRS previo o los factores de riesgo vascular. Sin embargo, al realizar el análisis por los diferentes subtipos etiológicos de ictus, se encontró que esta asociación era debida a los ictus de tipo aterotrombóticos, los únicos que de forma independiente presentaban esa asociación (178).

La edad biológica también se ha evaluado en relación con el riesgo de recurrencia de ictus a los tres meses tras el primer evento. Se ha encontrado una asociación significativa entre una mayor edad biológica y un mayor riesgo de recurrencia, siendo la edad biológica 2,7 años superior en aquellos pacientes con recurrencia (179).

En cuanto al ictus hemorrágico, se ha estudiado mediante EWAS un factor asociado a mal pronóstico clínico, la isquemia cerebral retardada en pacientes con hemorragia subaracnoidea. Se incluyeron un total de 40 pacientes con hemorragia subaracnoidea, de los cuales 13 presentaron isquemia cerebral retardada. Se encontraron los genes *INSR* y *CDHR5* diferencialmente hipermetilados en pacientes que presentaron isquemia cerebral retardada respecto a los que no lo hicieron. Además, la metilación en estos genes se asoció con una disminución de su expresión (180).

f. Interpretación y utilidad de los resultados genéticos y epigenéticos mediante herramientas bioinformáticas

i. Repositorios y bases de datos

Tanto los GWAS como los EWAS permiten la obtención de una gran cantidad de datos. Por lo tanto, es necesario el análisis de estos datos de una forma sistemática, siguiendo protocolos establecidos y usando herramientas bioinformáticas avanzadas. Sin embargo, más allá del análisis de los datos crudos obtenidos en análisis GWAS e EWAS, es necesaria la interpretación posterior de estos datos y un mapeo funcional. Esto es de gran relevancia debido a que más de un 90% de las variantes obtenidas en los estudios genéticos se encuentran en regiones no codificantes o se identifican en regiones que se encuentran ligadas a otras que son realmente las causantes de la asociación (181). Por eso es muy importante la anotación funcional y la interpretación de los datos en este tipo de estudios, los cuales se pueden hacer, en gran medida, gracias al acceso libre a repositorios y bases de datos.

Actualmente, es complicado entender el concepto de genómica y epigenómica sin pensar en el libre intercambio de los datos obtenidos mediante bases de datos controladas que aseguran la cumplimentación de la normativa actual, como por ejemplo la privacidad. Este intercambio de datos permite maximizar la utilidad y el beneficio público obtenido de los datos generados así como evitar solapamiento (182).

Existen muchas bases de datos con almacenamiento de diferentes tipos de datos genómicos que permiten un uso diferente durante el análisis genómico. Algunas bases tienen información sobre todos los datos conocidos para los diferentes SNPs (dbSNP

(183), Ensembl (184)), genes (Ensembl, GeneCards (185), Gene de NCBI (186)), proteínas (Uniprot (187)) o fármacos (drugbank (188), Therapeutic Target Database (TTD) (189) y PharmGKB (190)). También existen bases de datos específicas para enfermedades concretas. Además, hay disponibles repositorios que permiten conocer los estudios GWAS/EWAS realizados previamente como *GWAS catalog* (191) y el *EWAS datahub* (192). El *GWAS catalog*, fue fundado en 2008 y reúne la información de una forma fácil de visualizar de los diferentes GWAS publicados que tienen las condiciones para ser publicados en esta base de datos. Para algunos de los GWAS, se encuentra también disponible el resumen de las estadísticas obtenidas en ese GWAS concreto para cada uno de los polimorfismos analizados, lo que es de gran utilidad para la comparación con datos obtenidos posteriormente y para su uso con herramientas que permiten analizar este tipo de datos estadísticos. El equivalente al *GWAS catalog* pero en los estudios EWAS de metilación del ADN es el *EWAS Datahub*. En esta base de datos se guardan y comparten los resultados de numerosos EWAS. Actualmente, se encuentran datos de más de 75,000 muestras y permite obtener perfiles de metilación del ADN de referencia de 81 tejidos diferentes, seis ancestros y 67 enfermedades (192). Por otro lado, tanto en el *GWAS catalog* como en el *EWAS datahub* se pueden hacer búsquedas no solo de estudios sino también de asociaciones genotipo-fenotipo. Para la obtención de este tipo de datos también existen otros repositorios como el *PhenoScanner* (193) o el *OpenTargets* (194). Ambos contienen resultados disponibles de forma pública de GWAS. En ambos casos, a partir de los datos de variantes genéticas, genes, regiones genómicas o fenotipos, se obtiene información del resto de datos. Por ejemplo, al realizar una búsqueda por genes, se obtienen las regiones genómicas y variantes genéticas en ese gen con sus respectivas asociaciones con fenotipos obtenidos en análisis GWAS. En el caso de *PhenoScanner*, contiene información de más de 65 billones de asociaciones y más de 150 millones de variantes genéticas. Además de contener esta información, en el caso de *OpenTargets*, se obtiene información suficiente para la identificación y priorización de variantes causales de las asociaciones obtenidas en GWAS y dianas de fármacos a partir de datos genéticos.

Otros repositorios muy útiles a la hora de anotar, priorizar, visualizar e interpretar datos de GWAS son HaploReg (181) o Ensembl. En el caso de HaploReg, este permite la

anotación de variantes del genoma no codificante para encontrar las variantes causantes dentro de un bloque de haplotipo de la asociación identificada mediante GWAS. También permite identificar el tipo celular más probable causante de los resultados obtenidos, usando anotaciones regulatorias y epigenómicas (181).

Otras bases de datos muy utilizadas una vez obtenidos los resultados de los GWAS e EWAS son las que tienen datos sobre eQTLs (GTEx (195), QTLbase (196), eQTLGen (197) y eQTL Catalogue (198), entre otras) y mQTLs (mQTLdb (199)). En ellas se recoge la información sobre SNPs que alteran la expresión génica (*expression quantitative trait loci* (eQTLs)) o la metilación de CpGs en el ADN (mQTLs). Esta información puede ser combinada con la de SNPs asociados a un fenotipo obtenidos en un GWAS o con la de CpGs obtenidos en un EWAS.

ii. Estudios de causalidad: Randomización Mendeliana (RM)

Muchos estudios observacionales han permitido identificar asociaciones entre factores de riesgo y diferentes variables asociadas con el riesgo de ictus o con el pronóstico de la enfermedad. Sin embargo, estos estudios están sesgados y no permiten establecer relaciones de causalidad, debido al potencial efecto de variables confusoras que no se tienen en cuenta y a la posibilidad de la existencia de causalidad reversa.

Cuando se encuentra una asociación entre dos variables, en muchos casos es apropiado el establecimiento de la existencia de una relación de causalidad entre ambas. El método clásico ha sido, y sigue siendo, el desarrollo de ensayos clínicos. En estos estudios se modifica una variable y se estudia su efecto, controlando todos los otros factores. Por ejemplo, se divide una cohorte en dos y se aleatoriza la administración de un fármaco, pero todos los demás tratamientos y factores están controlados y son iguales entre ambos grupos, por lo tanto, si se observa un efecto concreto en el grupo en el que se le ha administrado el fármaco, pero no en el grupo control, será debido a ese fármaco.

Sin embargo, en muchas ocasiones es complicado y costoso el desarrollo de ensayos clínicos. Una aproximación epidemiológica diferente pero basada en los mismos principios que los ensayos clínicos es la RM. En este caso, se utiliza la genética como instrumento para analizar la relación de causalidad entre dos variables (200). La genética se considera una variable aleatorizada, ya que, durante la concepción, de forma

aleatoria, se establece la genética concreta de cada individuo. Los instrumentos basados en la genética se consideran independientes de factores confusores. De esta manera, si, por ejemplo, se quiere estudiar si la hipertensión (variable “exposición” o factor de riesgo) causa ictus (variable “resultado” o enfermedad), se seleccionarán los factores genéticos (instrumentos) asociados a la variable exposición. Normalmente, estos factores genéticos se seleccionan a partir de estudios GWAS. En este ejemplo, se seleccionaría un estudio GWAS (normalmente el más reciente, con un tamaño muestral más grande y realizado en el grupo étnico de interés en el estudio), y se elegirían aquellos SNPs asociados de forma significativamente estadística (normalmente, un valor de $P < 5 \times 10^{-8}$) con la presión sanguínea. Estos datos seleccionados del GWAS contendrán la información de polimorfismos asociados con mayor o menor riesgo de tener hipertensión, pero de forma aleatoria. Es decir, esos pacientes tendrán un riesgo aumentado o disminuido de tener presión sanguínea elevada en base a su genética, la cual es aleatoria. En el siguiente paso, lo que se hará será analizar esos mismos polimorfismos seleccionados asociados a la hipertensión (en este ejemplo), en un GWAS de la variable “resultado”, el ictus, en este caso. Por lo tanto, es importante tener disponible un GWAS para la variable “exposición” y al menos un GWAS para la variable “resultado”, ambos realizados en cohortes del mismo grupo étnico. Mediante diferentes métodos estadísticos se analizará la proporción de cambio explicada por la genética en el factor “exposición” y el cambio explicado por los mismos factores genéticos en la variable “resultado” (201,202). De esta manera, se podrá definir cómo los cambios en un factor de riesgo resultan en cambios en una enfermedad y así se podrá establecer la relación de causalidad. Sin embargo, es necesario que se cumplan algunas asunciones para que las variantes genéticas se consideren válidas. Por un lado, tienen que ser capaces de predecir el factor de riesgo estudiado. Además, tienen que asociarse a la variable resultado a través de su relación con el factor de riesgo. Por último, no pueden relacionarse con confusores de la asociación factor de riesgo-enfermedad (20). Es importante elegir el método estadístico para el análisis de la RM más apropiado en cada caso. Además, es necesaria la realización de análisis adicionales, llamados pruebas de sensibilidad, para descartar pleiotropía (cuando una variante genética influencia más de un fenotipo) o heterogeneidad en los instrumentos genéticos seleccionados (203).

Existen numerosos estudios de RM en el ictus. Estos análisis han sido posibles debido a la publicación del MEGASTROKE, ya que este estudio proporcionó los datos genéticos como recurso muy útil para aplicar estudios de RM en el riesgo de ictus y sus subtipos etiológicos. Por ejemplo, gracias a la RM se ha podido establecer que la presión sanguínea está causalmente asociada con el riesgo de ictus isquémico, especialmente en ictus aterotrombóticos y lacunares e ictus hemorrágico. Por otro lado, se han identificado los niveles de LDL como causantes de un aumento de riesgo de ictus aterotrombótico y los niveles de HDL como causantes de una disminución del riesgo de ictus lacunar. También se ha evaluado la relación de causalidad entre fenotipos metabólicos, exposición a tabaco y alcohol, con el riesgo de ictus. Se ha encontrado que el índice cintura-cadera se asocia con un mayor riesgo de ictus aterotrombótico y lacunar. Por otro lado, el consumo de alcohol y de tabaco se asocian causalmente con el riesgo de ictus. Se han estudiado muchos otros fenotipos en relación con el riesgo de ictus, como variables de conteo celular, dieta, sueño, infecciones, etc. Todos estos análisis demuestran la aplicabilidad de esta metodología para encontrar factores asociados causalmente con el riesgo de ictus y sus subtipos (20).

A parte de ser una herramienta muy útil para establecer relaciones de causalidad entre posibles factores de riesgo y enfermedades, la RM está empezando a utilizarse en otras áreas, como para establecer relaciones de causalidad en estudios de epigenética, especialmente de metilación del ADN (204,205). Como ya se ha comentado, uno de los inconvenientes en el estudio de la epigenética es que es complicado saber si el efecto que se está viendo es causa o consecuencia en una enfermedad concreta. La RM puede ayudar a establecer esta relación. Por ejemplo, si se identifica una CpG diferencialmente metilada asociada al riesgo de ictus, pero no se sabe si es causa o consecuencia, con la RM se identificarán mQTLs (polimorfismos que se asocian a cambios en la metilación) para esa CpG. Estos mQTLs serán SNPs que se utilizarán como instrumentos para establecer la relación de causalidad. Como ejemplo, Portilla-Fernández et al., identificó la metilación de una CpG en el gen *AHRR* asociada al grosor de la íntima-media carotídea. A través de RM encontró que la metilación de esta CpG se asociaba causalmente con el riesgo de ictus (205).

Otra aplicación importante de la RM es la de evaluar fármacos como posibles tratamientos en una enfermedad concreta y poder priorizar los tratamientos que podrían ser estudiados en ensayos clínicos (20). Si se conoce la diana del fármaco y cómo este tratamiento afecta a esa diana, se pueden utilizar instrumentos genéticos asociados a la expresión de esa proteína / gen diana para evaluar el posible efecto del fármaco en la enfermedad. Este análisis se realizó en el estudio GIGASTROKE (75). Se analizaron 218 proteínas que son dianas de fármacos aprobados. Para todas estas proteínas existían GWAS analizando los polimorfismos asociados a sus niveles. Utilizando esta información, se pudo establecer el efecto que estas proteínas pueden tener sobre el ictus. Si el efecto de la proteína sobre el ictus es el contrario que el del fármaco sobre la proteína, los fármacos podrían ser utilizados como posibles tratamientos en el ictus. En otro estudio, se analizó el efecto de dos tipos de fármacos utilizados para la disminución de la presión sanguínea sistólica, los β -bloqueantes y los inhibidores del canal de potasio. La MR mostró resultados similares a ensayos clínicos, viendo que los inhibidores del canal de calcio tenían un efecto más potente sobre el riesgo del ictus que los β -bloqueantes (206). Utilizando esta aproximación, también se ha evaluado el efecto de fármacos antitrombóticos, anticoagulantes e hipolipemiantes (20).

iii. Estudio de correlación genética

Los análisis de correlación genética permiten encontrar solapamiento genético entre dos fenotipos o enfermedades complejas. Este solapamiento puede indicar vías fisiopatológicas comunes entre dos enfermedades estudiadas y ayudar a discernir la arquitectura poligénica de las enfermedades complejas (207,208).

Actualmente se utilizan dos métodos para encontrar el solapamiento o correlación genética: el *linkage disequilibrium score regression* (LDSC) y el *genomic restricted maximum likelihood* (GREML). La exactitud del método GREML parece ser superior, sin embargo, computacionalmente es más rápido el LDSC y además no requiere de datos del genotipo individual de cada paciente. El método de GREML utiliza *linear mixed model* (LMM) combinado con el algoritmo *restricted maximum likelihood* (REML) para calcular la varianza genética a partir de la estimación de la heredabilidad utilizando múltiples SNPs. El LDSC realiza una regresión entre el valor estadístico de asociación del SNP obtenido en el GWAS con su puntuación de LD. La puntuación de LD de un SNP se

determina por la suma de todos los r^2 de LD de ese mismo SNP con todos los demás. La puntuación de LD es específica para cada población étnica. En el análisis de LDSC se puede generar la puntuación de LD específica para la cohorte analizada (cuando se disponen de los genotipos para cada paciente) o utilizando una puntuación de LD de referencia calculada en una población de la misma etnia que la analizada. El análisis mediante LDSC se realiza utilizando el resumen estadístico de los GWAS, los cuales provienen de cohortes independientes. El signo positivo o negativo en los resultados obtenidos de correlación genética indican la dirección de solapamiento (207). Sin embargo, LDSC no permite incluir la anotación funcional del genoma dentro de los cálculos de correlación genética. Por eso, se desarrolló el *GeNetic cOVariance Analyzer* (GNOVA), un método que permite calcular la covarianza genética utilizando los datos de resumen estadístico de GWAS con LDSC e integrando esta información con anotación funcional del genoma (208). Este método ha demostrado ser robusto al solapamiento de muestras. GNOVA permite realizar la estratificación de la correlación genética utilizando tres anotaciones funcionales: según cuartiles de la frecuencia del alelo menor, según predicción de la funcionalidad de los SNPs, y en base a siete tejidos diferentes (208). Posteriormente, en 2021 se lanzó SUPERGNOVA, una herramienta que permite el cálculo de la correlación genética en regiones específicas del genoma (209). Otros métodos, como LOGOdetect, se han desarrollado paralelamente para identificar las regiones concretas del genoma contribuyentes a la correlación genética entre dos fenotipos (210).

En el caso del ictus, se ha estudiado la correlación genética entre el riesgo de ictus y sus subtipos etiológicos con diferentes enfermedades o fenotipos. Un estudio publicado en 2020 por Cai et al., analizó la correlación genética entre un total de 728 fenotipos con el riesgo de ictus y sus subtipos (211). De todos ellos, se encontró correlación genética con 99 fenotipos (después de aplicar corrección por múltiples pruebas con Bonferroni). Entre ellos, se identificaron factores de riesgo conocidos, como la obesidad, la hipertensión, la diabetes tipo II, etc. con una correlación genética significativa con el riesgo de ictus. Estos resultados indican que todos estos factores comparten mecanismos fisiopatológicos con el ictus (211).

iv. Estudio de vías metabólicas

Como ya se ha comentado, los GWAS y EWAS permiten obtener información acerca de la asociación de un SNP o CpG con un fenotipo. Para evitar errores de tipo I, de forma estándar, se seleccionan como significativos aquellos SNPs que tienen un valor estadístico de P más bajo de 5×10^{-8} o CpGs con un valor de $P < 2.4 \times 10^{-7}$ (con el *array* 450K). Sin embargo, aunque esta aproximación es correcta, es posible que muchas asociaciones que tienen un tamaño de efecto más pequeño, aunque sean reales, no sean detectadas. Por eso, una solución propuesta es la organización de los datos obtenidos de los GWAS e EWAS (cada asociación SNP/CpG-fenotipo) en genes para posteriormente agrupar estos genes según su significado biológico (212). Con esta aproximación, el objetivo es encontrar un enriquecimiento de ciertas vías metabólicas entre los resultados obtenidos de los GWAS o EWAS, respecto a lo esperado por azar. Este tipo de análisis, que son sin hipótesis previa, se conocen como análisis de enriquecimiento de vías metabólicas o *Genome-wide Pathway Analysis* (GWPA) (212).

Las vías metabólicas constituyen una manera de organizar las proteínas o genes según su función biológica. Todos aquellos genes que estén implicados en una función determinada pertenecerán a la misma vía metabólica. Las vías metabólicas se pueden describir según la región en la célula donde se desempeña la función o el tipo de proceso que realiza (212).

Para poder realizar análisis de enriquecimiento de vías metabólicas a partir de los datos de los GWAS o EWAS, es necesario tener bases de datos disponibles que contengan los genes organizados en vías metabólicas. Existen diferentes bases de datos, la mayoría de ellas de libre acceso como KEGG, GO, Reactome o Biocarta. También hay algunas que requieren de licencia para su acceso, como es el caso de la *Ingenuity Knowledge Base*. Entre ellas difieren en el foco funcional (vías metabólicas, vías de señalización, procesos metabólicos, nivel celular de actuación, interacciones proteína-proteína, etc.), el tipo de mantenimiento y evaluación de la entrada de información a la base de datos (manual o computacional) y la estructura de la base de datos (ontología o discreta) (212).

Gracias al uso de diferentes algoritmos, se agrega la información obtenida en los GWAS o EWAS con la información recogida en estas bases de datos. Al calcular la sobre-

representación se han de tener en cuenta una serie de consideraciones como el tamaño del gen o de la vía metabólica, el número de CpGs por gen, el patrón de LD o el número de SNPs/CpGs analizados. Por ejemplo, debido al ratio de error de tipo I de $\alpha = 0.05$ para cada SNP o CpG, para aquellos genes que tienen más SNPs o CpGs, la probabilidad de cometer error de tipo I y tener un falso positivo es mayor. Además, si se comete error de tipo I, en el caso de los GWAS, este se extiende a todos los SNPs que están en LD. De manera similar ocurre con las vías metabólicas que incluyen un número mayor de genes. También hay que tener en cuenta que para obtener un listado de genes que pueda ser analizado para encontrar enriquecimiento de vías metabólicas, es necesario agrupar los SNPs y CpGs en genes y seleccionar aquellos genes que deben ser analizados por vías metabólicas. Existen diferentes algoritmos que realizan esta agrupación y que tienen en cuenta estas consideraciones. Algunos de los algoritmos que permiten pasar de datos de SNPs obtenidos en GWAS a resultados de enriquecimiento de vías metabólicas son FUMA, GSA-SNP2, MAGMA, VEGAS2, GSEA4GWAS, etc. Cada uno de ellos se diferencia en la manera en la que se hace la asignación de SNPs a genes, la base de datos de vías metabólicas que utiliza de partida y las correcciones que aplican sobre las consideraciones mencionadas anteriormente. En cuanto a la asignación de SNPs a genes concretos, el método más clásico es hacerlo por posición. Es decir, asignar a cierto gen todos los SNPs que se encuentran en la región intra-génica. Con esta aproximación, también es conveniente asignar a ese gen los SNPs que están en LD con los que se han incluido. Sin embargo, muchos SNPs en los GWAS se encuentran en regiones inter-génicas y, además, SNPs intra-génicos pueden estar regulando la expresión de genes más alejados. Por eso, herramientas más nuevas como FUMA, incluyen información de eQTLs y de secuenciación de captura de cromatina (Hi-C), que informan sobre la expresión afectada por un SNP así como de la disposición tridimensional de la cromatina (212).

En el caso de los EWAS, existen herramientas diseñadas específicamente para el análisis de vías metabólicas a partir de los resultados obtenidos. Estas herramientas, como el MethylGSA (213), disponible en R, permiten analizar los datos de EWAS teniendo en cuenta los posibles sesgos debidos a un diferente número de CpGs por gen.

Los datos obtenidos de agrupación de SNPs en genes se pueden analizar mediante *Over-representation Analysis (ORA)* o *Gene Set Enrichment Analysis (GSEA)*. La diferencia entre ambos es que ORA compara los genes obtenidos con la base de datos de vías metabólicas elegida, sin asignarle un peso específico a cada gen, mientras que GSEA, sí que lo hace. De esta manera, con este método se le asigna una puntuación a cada gen, la cual se tendrá en cuenta en el análisis de vías metabólicas (214).

Una herramienta muy utilizada en publicaciones recientes para el análisis de vías metabólicas es WebGestalt ya que se puede utilizar a través de su web o computacionalmente a través del lenguaje R y contiene información de 12 organismos, 342 genes y 155,175 categorías funcionales. WebGestalt permite elegir diferentes parámetros, como la base de datos funcional, para hacer análisis de ORA y GSEA, entre otras opciones (215).

- v. Estudios de asociación de transcriptoma y proteoma de genoma completo (*Transcriptome-wide association studies* y *Proteome-wide association studies*)

Con el objetivo de interpretar los datos y priorizar genes identificados en GWAS, los *Transcriptome-Wide Association Studies (TWAS)* y los *Proteome-Wide Association Studies (PWAS)* permiten integrar la genómica funcional con los resultados de los GWAS. Ambas aproximaciones utilizan herramientas bioinformáticas para encontrar genes asociados a una enfermedad mediante la regulación de la transcripción (TWAS) o de los niveles de proteínas (PWAS) (216). Los dos realizan un análisis de asociación a nivel de gen o proteína utilizando datos de GWAS, considerando las diferentes variantes genéticas de ese gen o proteína. Para ello se tienen en cuenta las variantes genéticas que tienen la capacidad de regular la transcripción (eQTLs) o la traducción (pQTLs). En el caso del TWAS, se analiza cómo la alteración de la transcripción mediada por eQTLs es capaz de alterar un fenotipo. De forma muy similar, los PWAS analizan lo mismo, pero teniendo en cuenta pQTLs en lugar de eQTLs (216).

Tanto los PWAS como los TWAS presentan ciertas ventajas enfrente a métodos que analizan la genómica basándose en variantes concretas, como los GWAS. Por un lado, permiten extender el conocimiento obtenido de los GWAS de forma que se entiendan

los mecanismos que producen una enfermedad de una manera más funcional. Además, al basarse en genes y proteínas, y no en variantes concretas, estadísticamente se necesita un umbral inferior para encontrar resultados significativos. Además, los TWAS y los PWAS son específicos de tejidos, ya que se utilizan datos de transcriptómica o proteómica de un tejido concreto. Esto permite realizar los análisis en los tejidos más relevantes en una enfermedad concreta (216).

Existen diferentes softwares diseñados para realizar este tipo de análisis. La elección dependerá del tipo de datos del que se disponga. Principalmente, se debe tener en cuenta que se puede trabajar tanto con datos genéticos de GWAS individualizados por paciente o con el resumen estadístico de resultados GWAS. También hay diferentes modelos estadísticos utilizados para identificar eQTLs y pQTLs y diferentes métodos para realizar el análisis de asociación entre genes/proteínas y el fenotipo (216).

Sin embargo, es posible obtener falsos positivos en los resultados de los PWAS y TWAS. En la Figura 3 se pueden ver las situaciones en las que puede ocurrir un falso positivo. Por este motivo, es importante realizar análisis adicionales para comprobar la relación de causalidad en la asociación entre el gen/proteína y el fenotipo. De hecho, algunos de los softwares, como es el caso de *FUSION* (<http://gusevlab.org/projects/fusion/>) (217) ya tienen incorporadas algunas de estas opciones. Una de las aproximaciones es realizar un análisis bayesiano de colocación en el que se evalúa la probabilidad posterior de que los niveles del gen o proteína y el fenotipo estén regulados por la misma variante genética (217). Por otro lado, también se puede testar la pleiotropía mediante un test de *summary data-based Mendelian randomization* (SMR) en el que se evalúa si una variante genética está asociada con un fenotipo solo a través de la modificación de sus niveles de expresión génica o proteica (218).

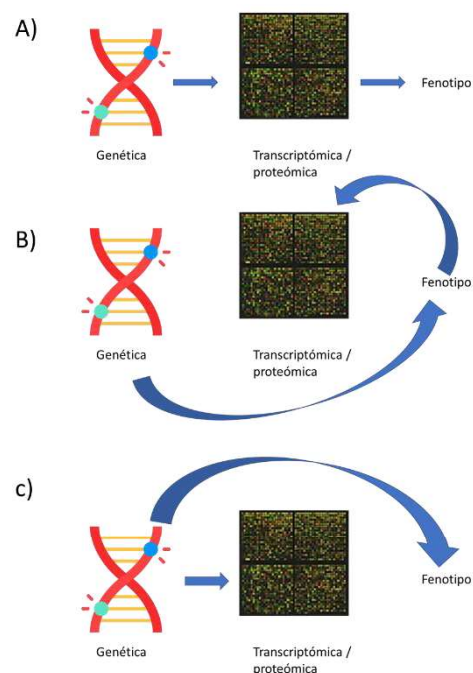


Figura 3: Asociaciones significativas en PWAS/TWAS

Los TWAS/PWAS son significativos en las condiciones A, B y C, siendo solo una significancia real la A. En la opción A, un gen se asocia a un fenotipo concreto a través de cambios producidos en la transcripción o traducción. Las opciones B y C muestran falsos positivos de los TWAS/PWAS.

Figura original de la doctoranda

Actualmente se han realizado muy pocos estudios TWAS o PWAS en relación al ictus. Un estudio realizado por Yang et al., (219) identificó 19 genes asociados con el riesgo de ictus a través de su expresión en diferentes tejidos, principalmente sangre periférica y tejido adiposo. Además, el estudio de Traylor et al., analizó mediante TWAS solo un subtipo de ictus (ictus de tipo lacunar). En este estudio se identificaron seis genes asociados con el ictus lacunar a través de la alteración de su expresión (71).

vi. Reposicionamiento farmacológico

El reposicionamiento farmacológico consiste en el estudio de fármacos ya aprobados o en investigación con una indicación diferente a la aprobada o estudiada. Esta estrategia

surgió por la dificultad en la obtención de nuevos fármacos usando un método clásico, especialmente por el tiempo, tasas de fallo y dinero invertido que supone. En cambio, con el uso del reposicionamiento farmacológico, debido a que la elección del fármaco a testar se realiza de una forma sistemática, bien por métodos computacionales o bien por métodos experimentales, la tasa de éxito es mayor. Por otro lado, como son fármacos ya aprobados o parcialmente estudiados, los fármacos ya han superado, al menos, las primeras fases de estudio de un fármaco, especialmente las fases preclínicas y las fases I-II, por lo que han pasado las fases de seguridad y, por lo tanto, tienen menos riesgo de deserción. Además, esto supone una reducción en el tiempo de salida al mercado y una reducción del coste asociado (220).

Dentro de los métodos computacionales utilizados para elegir fármacos candidatos a ser reposicionados, el uso de GWAS es una estrategia que está resultando exitosa. Los GWAS permiten identificar genes asociados con el mecanismo y la biología de las enfermedades y en muchas ocasiones, estos genes también se encuentran asociados con la fisiopatología de otras enfermedades para las cuales existe un tratamiento eficaz. Sin embargo, es importante el uso de las herramientas disponibles para la correcta anotación funcional de las asociaciones identificadas mediante GWAS para poder encontrar las dianas reales para el tratamiento de enfermedades. Según un estudio de Saneau et al., se encontró que los genes identificados mediante GWAS tienen el doble de probabilidad de que puedan ser dianas de fármacos comparado con el resto del genoma (221). En concreto, más adelante, Nelson et al., identificó que la proporción de fármacos obtenidos con un soporte de la genética aumentaba en las fases de investigación, siendo de un 2% en los estudios preclínicos y de un 8,2% entre los fármacos aprobados (222). Como se ha comentado, el uso de las bases de datos es de gran utilidad cuando se quieren analizar con más profundidad los datos de GWAS (223). En el caso del reposicionamiento farmacológico, el uso de bases de datos que contienen información de fármacos y sus dianas es de gran utilidad para este propósito (220).

A partir de los datos obtenidos mediante GWAS, existen diferentes estrategias para realizar el reposicionamiento farmacológico, las cuales a veces se combinan. En una revisión reciente, estos métodos se han clasificado en aproximaciones basadas en:

1) Genes candidatos: La aproximación se basa en la selección de los SNPs más significativos identificados en el GWAS y su posterior anotación funcional para identificar los genes realmente asociados (222).

2) Mapeo de vías metabólicas/genes: A partir de los datos individuales por SNP en un GWAS, se combinan para tener una visión más global y se consideran las vías metabólicas o un conjunto de genes implicados en la enfermedad. De esta manera se buscan fármacos que actúan en esa vía para poder hacer el reposicionamiento (222).

3) Investigación de similitudes entre fármacos y entre enfermedades: Se considera la similitud entre pares de fármacos y de enfermedades. La ventaja es que no se requiere profundizar en los datos genéticos asociados con la enfermedad, ya que no solo se consideran los genes más significativos (222).

4) Patrón de expresión inversa entre fármacos y enfermedades: Para utilizar esta aproximación, se debe conocer el patrón de expresión que provoca un fármaco. Para obtener esta información, está disponible la base de datos *LINCS (L1000 Connectivity Map)*, la cual incluye 1,328,098 perfiles de expresión génica provocados por alrededor de 42,000 componentes. Esta información se puede combinar con la obtenida a partir de los GWAS. Con los datos genéticos se puede imputar el perfil de expresión génica utilizando herramientas como PrediXcan. Lo que se busca con esta aproximación son fármacos que tengan el efecto contrario sobre la expresión génica que los polimorfismos identificados en un GWAS concreto (222).

5) Redes: Esta aproximación es más global, ya que su objetivo es la integración de muchos datos biológicos para descubrir relaciones entre fármacos y enfermedades. Se basa en el concepto de la organización de moléculas (proteínas, genes, fármacos y enfermedades) en redes. Cada molécula es un nodo que se conecta mediante puentes de unión (222,224). Con la conexión de los datos genéticos obtenidos de GWAS, junto a información de proteínas, enfermedades y fármacos, se pueden descubrir conexiones desconocidas previamente (222).

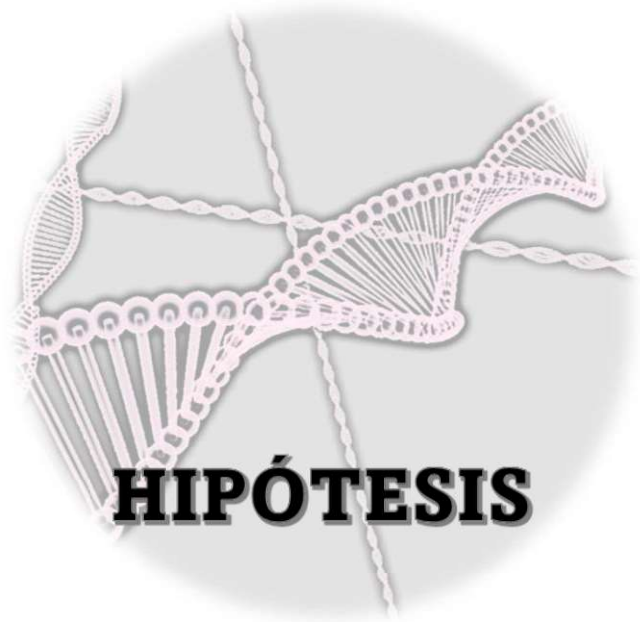
Existen diferentes herramientas bioinformáticas para el análisis sistemático de genes identificados en GWAS con grandes bases de datos de fármacos que contienen información sobre sus dianas con el fin de realizar reposicionamiento de fármacos. Una

de las aproximaciones más recientes es el software *Trans-Phar*. Esta herramienta se basa en la aproximación de patrón de expresión inversa entre fármacos y enfermedades y permite integrar información obtenida de los GWAS con datos de transcriptómica y posteriormente con información de la base de datos farmacológica *L1000 Connectivity Map*. De esta manera, este software busca fármacos que tengan un efecto en la expresión génica inverso al de la variante genética identificada mediante GWAS. Por ejemplo, el programa identificará fármacos que disminuyan la expresión de un gen concreto para el cual la variante identificada mediante GWAS se asocia a un aumento de expresión (221).

En el caso del ictus, recientemente, el estudio GIGASTROKE ha realizado diferentes aproximaciones de reposicionamiento farmacológico utilizando los datos obtenidos de su GWAS. Por un lado, realizó mapeo de vías metabólicas/genes. Se combinó la información de los SNPs para obtener datos de genes. Los genes identificados de riesgo en el ictus se buscaron entre las dianas de fármacos aprobados mediante la herramienta GREP. Se encontró un enriquecimiento de genes asociados al riesgo de ictus que son dianas de fármacos clasificados como “fármacos para tratar la sangre y los órganos formadores de sangre” según el *Anatomical Therapeutic Chemical Classification System [ATC] B drugs*. Por ejemplo, el gen *PDE3A* y *FGA*, los cuales codifican dianas para un fármaco antiplaquetario (cilostazol) y fármaco trombolítico (alteplasa) (75).

Por otro lado, se realizó la aproximación de patrón de expresión inversa entre fármacos y enfermedades utilizando la herramienta *Integration of Transcriptome-wide association study and Pharmacological database* (Trans-Phar) (221) para encontrar posibles fármacos para el tratamiento del ictus. En este caso se identificó un fármaco antagonista del receptor de tromboxano A₂, estudiado como posible antiplaquetario para su uso en el ictus, y que presentó un perfil de expresión diferente al encontrado para el riesgo de ictus cardioembólico. Otro fármaco, BRD.A22514244, con un mecanismo de acción desconocido, se encontró con un perfil de expresión contrario al del ictus de tipo lacunar (75). Con un objetivo similar a Trans-Phar, pero utilizando datos de expresión de proteínas, se realizó RM. El objetivo era encontrar causalidad entre diferentes proteínas que son dianas de fármacos conocidos y el riesgo de ictus. Para ser candidatos a fármacos en el ictus, el método de actuación del fármaco sobre la

expresión de la proteína debía tener un efecto inverso al identificado para el riesgo del ictus. De un total de 218 proteínas estudiadas, se identificaron y confirmaron cinco proteínas (PROC, F11, KLKB1, MMP12 y GP1BA) causalmente asociadas al riesgo del ictus, las cuales eran dianas de fármacos con un mecanismo de acción inverso a la asociación identificada. Fármacos que tienen como diana F11 y KLKB1 ya se encontraban previamente en investigación para el tratamiento del ictus (75). Estos datos permiten pensar en el potencial descubrimiento de nuevos fármacos para el tratamiento del ictus. Además, al identificar genes dianas de fármacos ya estudiados para el tratamiento del ictus demuestran y prueban la gran utilidad del uso de la genética para el descubrimiento de nuevos fármacos.



HIPÓTESIS

El análisis del riesgo y pronóstico del ictus mediante técnicas *ómicas* que analizan la metilación del ADN junto con la integración de datos genómicos y proteómicos, puede ser una buena aproximación para encontrar genes asociados con el ictus y realizar una priorización funcional de genes previamente identificados en estudios genéticos de tipo GWAS.

Este tipo de análisis nos puede permitir conocer los mecanismos y las vías metabólicas por los que se produce el ictus y explicar la variabilidad en la evolución neurológica a corto/medio plazo y el estado funcional a largo plazo tras un ictus, así como identificar nuevos biomarcadores y dianas terapéuticas.



OBJETIVOS

1-Identificar variantes epigenéticas asociadas con el riesgo de ictus:

1.1-Investigar los cambios provocados por las alteraciones epigenéticas en las vías metabólicas humanas.

1.2-Estudiar la metilación diferencial por subtipos celulares presentes en tejido sanguíneo y su relación con el riesgo de ictus.

1.3-Identificar la dirección de causalidad existente entre las modificaciones epigenéticas y el riesgo de ictus.

2-Identificar variantes epigenéticas asociadas con la evolución neurológica a corto/medio plazo (al alta) después del ictus:

2.1-Estudiar la metilación diferencial por subtipos celulares asociada a la evolución neurológica post-ictus.

2.2-Investigar los cambios provocados por las alteraciones epigenéticas relacionadas con la evolución neurológica en las vías metabólicas humanas.

3-Estudiar los factores de riesgo asociados con el ictus de tipo lacunar mediante un análisis bioinformático multi-ómico.

4-Identificar proteínas expresadas en cerebro reguladas genéticamente en asociación al pronóstico funcional a largo plazo tras un ictus mediante integración de datos *ómicos*.



MATERIAL, MÉTODOS Y RESULTADOS

MATERIAL, MÉTODOS Y RESULTADOS

Artículo 1:

Título: DNA Methylation and Ischemic Stroke Risk: An Epigenome-Wide Association Study.

Revista: Thrombosis and Haemostasis.

Objetivo: Identificar variantes epigenéticas asociadas con el riesgo de ictus y estudiar su relación de causalidad.

Resumen:

Antecedentes y objetivo:

La heredabilidad del riesgo de ictus isquémico (IS) se explica en parte por la genética. Otros factores hereditarios, como la epigenética, podrían explicar una proporción desconocida del riesgo de IS. El objetivo de este estudio es evaluar la asociación de la metilación del ADN con el IS mediante el estudio de asociación de todo el epigenoma (EWAS).

Métodos:

Se realizó un EWAS en dos etapas en un total de 1.156 sujetos. Las posiciones de metilación diferencial (DMP) y las regiones metiladas diferencialmente (DMR) se evaluaron con Infinium 450K y EPIC BeadChip en la cohorte de descubrimiento (252 IS y 43 controles). Se replicaron los DMP significativos en una cohorte independiente (618 IS y 243 controles). También se evaluaron las asociaciones de subtipos de ictus isquémico. Se analizaron las CpG replicadas en un análisis de subtipos celulares diferencialmente metilados (DMCT) utilizando Epidish. Además, se realizó un análisis de enriquecimiento de vías metabólicas y un análisis de causalidad con randomización mendeliana (RM) para las CpG replicadas.

Resultados:

Se identificaron 957 CpG significativas ($p \leq 10^{-7}$) en la cohorte de descubrimiento, siendo las CpG en los genes ZFX3 y MAP3K1 las principales asociaciones

($\log FC=0.058, P=2.35 \times 10^{-22}$; $\log FC=0.035, P=3.22 \times 10^{-22}$, respectivamente). ZFH3 y MAP3K1 se encontraron entre los DMR más significativos. Se replicaron 697 CpG considerando valores de P corregidos por Bonferroni ($p < 5.22 \times 10^{-5}$). Todas las DMP replicadas se asociaron con riesgo de ictus cardioembólico, aterotrombótico e indeterminado. El análisis de DMCT demostró que las asociaciones significativas fueron identificadas por su asociación en las células *Natural Killer*. El análisis de enriquecimiento de vías mostró una sobrerrepresentación de genes pertenecientes a ciertas vías, incluido el estrés oxidativo. La metilación de ZFH3 y MAP3K1 se asoció causalmente con el riesgo de subtipos específicos de ictus.

Conclusiones:

El patrón específico de metilación del ADN se asocia causalmente con el riesgo de IS. Estos resultados podrían ser útiles para predecir de forma específica la aparición de ictus y potencialmente podría evaluarse su papel terapéutico.

DNA Methylation and Ischemic Stroke Risk: An Epigenome-Wide Association Study

Natalia Cullell^{1,2,3,*} Carolina Soriano-Tárraga^{4,5,*} Cristina Gallego-Fábrega¹ Jara Cárcel-Márquez¹
 Nuria P. Torres-Águila^{1,6} Elena Muiño¹ Miquel Lledós¹ Laia Lluçà-Carol^{1,7} Manel Esteller^{8,9,10,11}
 Manuel Castro de Moura⁸ Joan Montaner¹² Alba Fernández-Sanlés^{13,14} Roberto Elosua^{13,15,16}
 Pilar Delgado¹⁷ Joan Martí-Fàbregas¹⁸ Jerzy Krupinski^{2,19} Jaume Roquer⁴ Jordi Jiménez-Conde⁴
 Israel Fernández-Cadenas^{1,2}

¹Institut d'Investigació Biomèdica Sant Pau (IIB SANT PAU), Sant Quintí, Barcelona, Spain

²Department of Neurology, Hospital Universitari MútuaTerrassa/ Fundació Docència i Recerca MútuaTerrassa, Barcelona, Spain

³Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain

⁴Neurovascular Research Group, Department of Neurology, Hospital del Mar, IMIM, Universitat Autònoma de Barcelona/ DCEXS-Universitat Pompeu Fabra, Barcelona, Spain

⁵Department of Psychiatry, NeuroGenomics and Informatics, Washington University School of Medicine, St. Louis, Missouri, United States

⁶Evolutionary Developmental Genomics Research Group, The Scottish Oceans Institute, University of St Andrews, St Andrews, United Kingdom

⁷Department of Brain Ischemia and Neurodegeneration, Institut d'Investigacions Biomèdiques de Barcelona, Consejo Superior de Investigaciones Científicas (CSIC), Barcelona, Spain

⁸Josep Carreras Leukaemia Research Institute, Barcelona, Spain

⁹Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain

¹⁰Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

Address for correspondence Israel Fernández-Cadenas, PhD, Stroke Pharmacogenomics and Genetics, IIB-Sant Pau. c/Sant Antoni M^a Claret, 167, 08025 Barcelona, Spain
 (e-mail: israelcadenas@yahoo.es).

¹¹Centro de Investigación Biomédica en Red Cancer, Barcelona, Spain

¹²Department of Neurology, Hospital Universitario Virgen Macarena, Institute of Biomedicine of Seville/Hospital Universitario Virgen del Rocío/CSIC/University of Seville, Seville, Spain

¹³Cardiovascular Epidemiology and Genetics Research Group, IMIM, Barcelona, Spain

¹⁴Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom

¹⁵CIBER Cardiovascular Diseases, Instituto Carlos III, Barcelona, Spain

¹⁶School of Medicine, University of Vic-Central University of Catalonia, Barcelona, Spain

¹⁷Neurovascular Research Laboratory, Vall d'Hebron Institute of Research, Barcelona, Spain

¹⁸Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

¹⁹Centre for Bioscience, School of HealthCare Science, Manchester Metropolitan University, Manchester, England

Thromb Haemost

Abstract

Keywords

- stroke
- DNA methylation
- epigenetics
- epigenome-wide association studies
- EWAS

Background Ischemic stroke (IS) risk heritability is partly explained by genetics. Other heritable factors, such as epigenetics, could explain an unknown proportion of the IS risk. The objective of this study is to evaluate DNA methylation association with IS using epigenome-wide association studies (EWAS).

Methods We performed a two-stage EWAS comprising 1,156 subjects. Differentially methylated positions (DMPs) and differentially methylated regions (DMRs) were assessed using the Infinium 450K and EPIC BeadChip in the discovery cohort (252 IS and 43 controls). Significant DMPs were replicated in an independent cohort (618 IS and 243 controls). Stroke subtype associations were also evaluated. Differentially methylated cell-type (DMCT) was analyzed in the replicated CpG sites using EpiDISH.

* These authors contributed equally to this work.

received

September 21, 2021

accepted after revision

April 12, 2022

© 2022. Thieme. All rights reserved.
 Georg Thieme Verlag KG,
 Rüdigerstraße 14,
 70469 Stuttgart, Germany

DOI <https://doi.org/10.1055/s-0042-1749328>.
 ISSN 0340-6245.

We additionally performed pathway enrichment analysis and causality analysis with Mendelian randomization for the replicated CpG sites.

Results A total of 957 CpG sites were epigenome-wide-significant ($p \leq 10^{-7}$) in the discovery cohort, being CpG sites in the top signals ($\log_{2}FC = 0.058$, $p = 2.35 \times 10^{-22}$; $\log_{2}FC = 0.035$, $p = 3.22 \times 10^{-22}$, respectively). *ZFX3* and *MAP3K1* were among the most significant DMRs. In addition, 697 CpG sites were replicated considering Bonferroni-corrected p -values ($p < 5.22 \times 10^{-5}$). All the replicated DMPs were associated with risk of cardioembolic, atherothrombotic, and undetermined stroke. The DMCT analysis demonstrated that the significant associations were driven by natural killer cells. The pathway enrichment analysis showed overrepresentation of genes belonging to certain pathways including oxidative stress. *ZFX3* and *MAP3K1* methylation was causally associated with specific stroke-subtype risk.

Conclusion Specific DNA methylation pattern is causally associated with IS risk. These results could be useful for specifically predicting stroke occurrence and could potentially be evaluated as therapeutic targets.

Introduction

Ischemic stroke (IS) is a complex disease with an estimated heritability of 37.9%.¹ Different genome-wide association studies (GWAS) have evaluated the association of genetic variants with stroke,^{2–4} identifying 30 loci associated with any stroke or with stroke subtypes.³ Recently, the largest GWAS meta-analyzing data from MEGASTROKE and the UK Biobank identified three novel loci associated with stroke.⁵

However, a proportion of stroke heritability is not yet accounted for⁵ and could be explained by epigenetics, which are heritable changes that do not affect the DNA sequence. Moreover, DNA methylation is dynamic and is affected by genetic and environmental factors. Methylation in DNA CpG sites is the most studied epigenetic variation and is associated with modulation of gene expression.⁶

Epigenome-wide association studies (EWAS) are similar to GWAS and allow an agnostic evaluation of DNA methylation in a genome-wide approach. An increasing amount of knowledge is being gained using EWAS in many complex diseases, such as diabetes mellitus and myocardial infarction.^{7,8} In the stroke field, two EWAS have found an altered methylation pattern in *PPM1A* and *TRAF3* associated with stroke recurrence in patients treated with aspirin or clopidogrel, respectively.^{9,10} Another EWAS has found methylation modifications associated with stroke recurrence¹¹ and with atherothrombotic stroke risk¹² in a small cohort of patients. Moreover, global methylation is associated with stroke¹³ and an EWAS evaluating the association of methylation with IS risk has recently been published.¹⁴ The authors used the established “limma” and “minfi” protocols and identified 21 loci associated with IS.

However, none of the previous published EWAS related with stroke have evaluated the causality of methylation over stroke. Given the dynamic nature of DNA methylation, it can be a risk factor or a consequence for stroke. Moreover,

none of those previous EWAS have used Meffil,¹⁵ a new method for EWAS analysis which allows a reduction in false-positive rates and improves the statistical power.

Our objective was to evaluate the association of DNA methylation with IS risk using the recently published Meffil strategy for EWAS analysis and to establish causality between methylation changes and stroke risk using a Mendelian randomization (MR) analysis.

Methods

We included 252 patients with IS and 43 population-based controls in the discovery cohort from the GRECOS study¹⁶ and the EPIGENESIS and Geno-tPA projects.¹⁷ The replication cohort included 618 stroke patients from the BASICMAR study¹⁸ and 243 controls from the REGICOR project.¹⁹

The study was approved by the ethics committees of all the participant hospitals and all participants provided written informed consent to participate.

A complete description of the cohorts of the study is included in the ► **Supplementary Material** (available in the online version).

Samples and DNA Extraction

Whole blood was obtained for all patients during the first 24 hours after stroke. DNA was extracted and bisulfite converted prior to the EWAS evaluation.

DNA from the discovery cohort was extracted using the Centra Puregene Blood Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. DNA from cases from the replication cohort was extracted using manual salt precipitation in the Banco Nacional de ADN (Spanish National DNA Bank, Instituto de Salud Carlos III). DNA from controls from the replication cohort was extracted in Institut Hospital del Mar d'Investigacions Mèdiques as described in Sayols-Baixeras et al.¹⁹

EWAS

The methylation pattern was evaluated using the 450K-BeadChip (Illumina), which analyzes more than 450,000 CpG sites, and the Infinium EPIC-BeadChip (Illumina) (more than 850,000 CpG sites) in the discovery cohort. The 450K-BeadChip was used for controls in the replication cohort and both arrays were used for the cases from the replication cohort.

CpG quality controls (QCs) consisted of removal of: CpG sites with a detection p -value higher than 0.05, CpG sites from chromosomes X and Y, CpG sites located near a single nucleotide polymorphism (SNP),²⁰ and promiscuous CpG sites with affinity for more than one probe. Sample QCs consist of the removal of individuals with a percentage of missing CpG sites of $>1\%$ and samples with discordant phenotypic and genotypic sex. Beta values were obtained for each CpG site. These values were normalized using the Meffil normalization.¹⁵ Cell-type proportion in blood was estimated using Meffil for whole blood tissue and was considered as a covariate in the regression analysis. Batch effect was corrected using the Combat function from ChAMP.²¹

Statistical Analysis

We used pwrEWAS²¹ in R to estimate the power to obtain epigenome-wide significant results with a difference in methylation for stroke patients and controls of 0.05 considering a sample size between 260 and 305.

All the preprocessing, correction, and normalization steps were done using R scripts (cran.r-project.org), Bioconductor packages (bioconductor.org), Meffil package,¹⁵ and ChAMP packages.^{20,22}

DMP Analysis

We used multiple linear regression for the differentially methylated position (DMP) analysis, where β values for each CpG site were analyzed for the association with subject status (case or control). We adjusted the model for sex, age, smoking status, and cell type proportion because they are known factors that alter methylation. Meffil also adjusted the results by surrogate variables. We used MethylToSNP package in R to evaluate for possible SNP patterns in the significant CpG sites that could be causing false-positive results.²³

We also analyzed the association of the significant and replicated CpG sites by stroke subtypes based on the Trial of Org 10172 in Acute Stroke Treatment (TOAST) classification²⁴ in the discovery cohort. To check for the association of replicated CpG sites with infarct volume, we used surrogate markers for infarct size, and National Institutes of Health Stroke Scale (NIHSS) at baseline and at 24 hours, because we did not have available data for infarct volume.

We estimated the inflation factor and the bias using Bacon.²⁵ This package has been designed to evaluate inflation and bias specifically in epigenome- and transcriptome-wide results.

All the CpG sites with a p -value of $\leq 10^{-7}$ in the discovery cohort were considered significant and were analyzed in the replication cohort. CpG sites with p -values <0.05 in the

replication cohort were considered significant and p -values $<5.22 \times 10^{-5}$ were considered significant after multiple comparison correction by Bonferroni.

DMR Analysis

We also used DMRcate²⁶ to study differentially methylated regions (DMRs) in the discovery cohort. A DMR was considered to be a region that included a minimum of two consecutive significant CpG sites separated by less than 1,000 nucleotides.²⁶

Differentially Methylated Cell-Type Analysis

We used EpiDISH²⁷ to estimate the proportion of myeloid and lymphoid lineage cell types using an algorithm for whole blood samples. We considered the reference matrix defined over seven blood cell subtypes (B cell, CD4+ and CD8+ T cells, natural killer [NK], neutrophils, monocytes, and eosinophils).²⁷ The estimated cell-type fractions were used to identify differentially methylated cell types associated with stroke risk using the CellDMC function in EpiDISH. We selected the significant and replicated CpG sites associated with stroke risk in the DMP analysis to find enrichment of differentially methylated cell type (DMCT).

Pathway Enrichment Analysis

We found enrichment in gene ontology (GO) from DMP results using clusterProfiler in R.²⁸ We selected genes where replicated CpG sites were located. We included pathways with a minimum of 10 significant genes within the pathway. We considered those significantly enriched pathways with a p -value <0.05 .

Analysis of Significant Stroke Genes from GWAS

We selected the 35 loci associated with stroke at a genome-wide significant level ($p < 5 \times 10^{-8}$) from the MEGASTROKE study²⁹ and the MEGASTROKE + UK Biobank study.⁵ These loci overlap with 43 different genes. All the 43 genes were considered in our analysis.

We selected from the EWAS analysis all the CpG sites that passed the QCs and that overlapped with the 43 genes previously described in MEGASTROKE + UK Biobank data⁵ to be associated with stroke or specific stroke subtypes.

The CpG sites with a p -value <0.05 in the EWAS analysis of IS and stroke subtypes were considered nominally and p -value $<4.31 \times 10^{-05}$ significantly associated with stroke after correcting for the total number of CpG sites identified in the 43 genes.

Methylation Quantitative Trait Loci Analysis and MR Analysis

We performed MR analysis for the two top CpG sites in the discovery because (1) they were replicated and (2) they were also among the first results in DMR analysis.

First, we selected 179 patients with genotypic and epigenetic data from the discovery cohort. For each of the CpG sites we used SNPTTEST, applying the method “expected” to find SNPs associated with methylation status. We studied causality between methylation and stroke

Table 1 Descriptive table

	Discovery			Replication		
	Cases	Controls	p-Value	Cases	Controls	p-Value
Subjects (n)	252	43		618	243	
Sex						
Men, n (%)	149 (59.1%)	21 (48.8%)	0.20	341 (55.2%)	120 (49.4%)	0.13
Women, n (%)	103 (40.8%)	22 (51.2%)		277 (44.8%)	123 (50.6%)	
Age in years, median $\pm$ SD	76 $\pm$ 10.89	66 $\pm$ 7	3.8×10^{-3a}	77 $\pm$ 11.35	63 $\pm$ 8.61	2.6×10^{-26a}
Presence of HTN, n (%)	161 (59.9%)	12 (60%)	0.66	445 (72%)	143 (58.8%)	4.8×10^{-13a}
Smoking, n (%)	34 (17.34%)	2 (10%)	0.40	145 (23.5%)	76 (31.3%)	4.5×10^{-3a}
Presence of DM, n (%)	70 (28.7%)	6 (30%)	0.90	NA	NA	
TOAST		N/A			N/A	
CES, n (%)	77 (31.2%)			248 (40.1%)		
LAS, n (%)	74 (30%)			158 (25.6%)		
SVS, n (%)	26 (10.5%)			204 (33%)		
UND, n (%)	67 (26.6%)			4 (0.6%)		
Other, n (%)	2 (0.8%)			4 (0.6%)		
NA	N/A			N/A		

Abbreviations: N/A, not applicable; NA, not available; TOAST, Trial of Org 10172 in Acute Stroke Treatment.

Note: Main clinical characteristics of patients included in the discovery and replication cohorts.

^aStatistically significant p-values for differences between cases and controls.

risk with a two-sample MR approach. We used as instruments the methylation quantitative trait loci (mQTLs) with p -values $<10^{-7}$ identified in the subset of 179 subjects.

The results of the MEGASTROKE project³ were used as the outcome for stroke and stroke subtype risk.

We used the TwoSampleMR package²⁹ in R to perform SNP pruning, harmonization of summary statistics, MR analysis, and sensitivity (heterogeneity and horizontal pleiotropy) analysis.

We considered the inverse variance weighted (IVW) method to infer causality. Moreover, as sensitivity analyses, we performed MR analyses using MR-Egger and weighted-median (WM) approaches. We calculated MR-Egger regressions to estimate the effect of horizontal pleiotropy (MR-Egger relaxes the assumption of the exclusive effect of instruments on outcome through exposure factor). To test for heterogeneity, we calculated the Cochran's Q.

Moreover, to assess for reverse causality, we performed IVW regression using significant polymorphisms ($p < 10^{-8}$) associated with IS and stroke subtypes from the MEGASTROKE project as instruments and the methylation of the two most significant CpG sites as outcome.

Results

Discovery Analysis

In the discovery cohort 252 IS patients and 43 population-based controls were analyzed. We calculated the statistical power using the specific tool designed to calculate power in EWAS. Based on this analysis, the statistical power to detect

differences in methylation ($\Delta\beta$) of 0.05 in a case/control study with a sample size of 295 subjects is estimated to be higher than 0.8 (**►Supplementary Fig. S1**, available in the online version). Cases and controls presented similar characteristics in terms of clinical and demographic variables. Only age differed between cases and controls (**►Table 1**). However, the analyses were adjusted for age.

After QC, a total of 384,509 CpG sites were analyzed in the discovery cohort ($n = 295$). We identified 957 DMPs ($p \leq 10^{-7}$) (**►Table 2**, **►Supplementary Table S1**, and **►Supplementary Fig. S2**, available in the online version). These CpG sites corresponded to 836 different loci. Some genes had more than one significant CpG site: 16 different genes had between two and five significant CpG sites. All but two CpG sites were hypomethylated in stroke patients compared with controls (**►Fig. 1**). Based on Bacon results, the EWAS did not show inflation (1.1), and the estimated bias was low (0.18). No SNP pattern was observed among the significant results that could have driven to false-positive results.

The two top CpG sites with p -value $<10^{-21}$ were cg07786668 in the 5' UTR of *ZFX3* ($p = 2.35 \times 10^{-22}$, logFC (confidence interval, CI) = 0.058 (0.069–0.048)) and cg10115490 in the first exon of *MAP3K1* ($p = 3.22 \times 10^{-22}$, logFC (CI) = 0.035 (0.041–0.029)) (**►Table 2** and **►Supplementary Table S1**, available in the online version). Moreover, *MAP3K1* loci presented an additional epigenome-wide significant CpG site: cg10206657, also in the first exon of the gene ($p = 4.11 \times 10^{-16}$, logFC (CI) = 0.056 (0.069–0.044)).

We identified 148 significant DMRs (96 remained significant after Bonferroni correction ($p < 7.6 \times 10^{-4}$)). The most significant DMR overlapped with *MAP3K1* genes

Table 2 Top epigenome-wide significant CpG sites replicated

CpG	CHR	BP	GENE	feature	cgi	Discovery		Replication	
						logFC (CI)	p-Value	logFC (CI)	p-Value
cg07786668	16	73092391	ZFHX3	5'UTR	Island	0.058 (0.069–0.048)	2.35E-22	0.017 (0.023–0.01)	3.24E-07
^a cg10115490	5	56111596	MAP3K1	1stExon	Island	0.035 (0.041–0.029)	3.22E-22	0.028 (0.032–0.023)	9.89E-32
^a cg15591678	10	64134155	ZNF365	5'UTR	Island	0.036 (0.042–0.029)	3.76E-21	0.022 (0.026–0.018)	1.12E-21
^a cg13432391	19	10828930	DNM2	1stExon	Island	0.028 (0.034–0.023)	1.74E-19	0.019 (0.023–0.015)	4.69E-21
^a cg09000178	16	67063319	CBFB	1stExon	Island	0.055 (0.066–0.044)	2.58E-19	0.023 (0.029–0.016)	1.38E-10
^a cg10849092	15	40574697		IGR	Island	0.029 (0.035–0.023)	1.66E-18	0.016 (0.019–0.012)	1.49E-18
cg23681311	22	22221878	MAPK1	1stExon	Island	0.045 (0.054–0.036)	2.21E-18	0.024 (0.03–0.019)	1.02E-18
^a cg07667295	14	64320118	SYNE2	5'UTR	Island	0.046 (0.056–0.037)	3.43E-18	0.049 (0.057–0.041)	1.10E-30
^a cg09238598	5	14871908	ANKH	TSS200	Shore	0.057 (0.069–0.045)	7.30E-18	0.031 (0.038–0.025)	2.73E-19
^a cg02516155	17	4046031	CYB5D2	TSS1500	Island	0.029 (0.035–0.023)	7.44E-18	0.023 (0.028–0.018)	1.51E-20
^a cg22512322	3	64009096	PSMD6	5'UTR	Island	0.038 (0.046–0.03)	7.73E-18	0.005 (0.009–0.002)	4.75E-3
^a cg06642177	6	134496341	SGK1	Body	Island	0.07 (0.085–0.055)	1.17E-17	0.013 (0.025–0.002)	2.32E-2
^a cg17662034	8	74207518	RDH10	1stExon	Island	0.063 (0.076–0.049)	1.99E-17	0.029 (0.038–0.021)	1.79E-11
^a cg14119263	10	5708327	ASB13	Body	Island	0.048 (0.058–0.038)	2.00E-17	0.027 (0.032–0.022)	1.53E-24
^a cg14864852	22	39102110	GTPBP1	1stExon	Island	0.028 (0.034–0.022)	2.99E-17	0.007 (0.01–0.004)	1.94E-5
^a cg25869317	15	101792241	CHSY1	TSS200	Island	0.045 (0.054–0.035)	5.03E-17	0.025 (0.031–0.019)	1.72E-16
^a cg08867893	10	64134160	ZNF365	5'UTR	Island	0.031 (0.038–0.024)	5.68E-17	0.011 (0.014–0.008)	4.85E-14
^a cg14078059	12	65174660		IGR	Island	0.077 (0.094–0.06)	1.21E-16	0.034 (0.042–0.026)	3.79E-16
^a cg10206657	5	56111587	MAP3K1	1stExon	Island	0.056 (0.069–0.044)	4.11E-16	0.023 (0.028–0.018)	1.37E-16
^a cg16588137	4	186064512	SLC25A4	1stExon	Island	0.024 (0.029–0.018)	6.42E-16	0.011 (0.014–0.008)	4.63E-13
^a cg01668174	2	106810661	UXS1	1stExon	Island	0.025 (0.031–0.02)	7.56E-16	0.014 (0.019–0.01)	1.69E-9
^a cg08327423	16	84150746	MBTPS1	TSS1500	Island	0.018 (0.023–0.014)	7.75E-16	0.011 (0.014–0.008)	4.22E-12
^a cg18829263	5	129240272	CHSY3	TSS1500	Island	0.034 (0.042–0.026)	7.91E-16	0.033 (0.038–0.029)	1.83E-42
^a cg26257177	19	10828928	DNM2	1stExon	Island	0.01 (0.012–0.008)	1.29E-15	0.006 (0.007–0.005)	4.44E-21
^a cg09643587	3	107809710	CD47	1stExon	Island	0.046 (0.057–0.036)	1.68E-15	0.046 (0.052–0.039)	1.06E-40
cg04759220	5	78532560	JMY	1stExon	Island	0.028 (0.035–0.022)	1.69E-15	0.011 (0.016–0.007)	2.78E-6
^a cg26587051	9	123476683	MEGF9	TSS200	Island	0.027 (0.034–0.021)	1.94E-15	0.018 (0.022–0.013)	5.96E-14
^a cg03036064	1	165797085	UCK2	1stExon	Island	0.017 (0.02–0.013)	2.01E-15	0.01 (0.012–0.008)	5.67E-15
^a cg17664568	10	98346460	TM9SF3	Body	Island	0.022 (0.027–0.017)	2.08E-15	0.021 (0.025–0.017)	1.63E-22
^a cg04794887	4	1858231	LETM1	TSS1500	Island	0.036 (0.044–0.028)	2.39E-15	0.022 (0.027–0.017)	2.48E-16
cg04192862	10	28966472	BAMBI	1stExon	Island	0.037 (0.046–0.029)	2.85E-15	0.023 (0.028–0.017)	2.84E-17
^a cg18751306	15	81294292	MESDC1	1stExon	Island	0.029 (0.035–0.022)	3.00E-15	0.015 (0.019–0.01)	6.12E-9
^a cg08377314	19	56186477	EPN1	TSS200	Island	0.03 (0.037–0.023)	3.96E-15	0.012 (0.017–0.008)	7.70E-08
^a cg03852091	2	242577219	THAP4	TSS1500	Island	0.02 (0.025–0.016)	5.18E-15	0.014 (0.017–0.01)	2.67E-15
^a cg05668205	19	16738998	MED26	1stExon	Island	0.02 (0.024–0.015)	5.63E-15	0.016 (0.019–0.012)	6.71E-17
^a cg03536846	8	120886254	DEPDC6	Body	Island	0.02 (0.024–0.015)	6.40E-15	0.011 (0.014–0.008)	3.80E-14
^a cg11299854	5	132083184	CCNI2	1stExon	Island	0.04 (0.05–0.031)	7.16E-15	0.022 (0.027–0.017)	4.41E-16
^a cg05135536	9	128509002	PBX3	TSS1500	Island	0.026 (0.032–0.02)	7.27E-15	0.012 (0.015–0.008)	8.19E-11

Note: Top results ($p < 10^{-14}$) presented as logarithm of the fold change (confidence interval: 97.5–2.5%) (logFC (CI)) and p -values (p) for the discovery cohort and replication cohorts.

^aIndicates new CpG sites not previously associated with stroke.

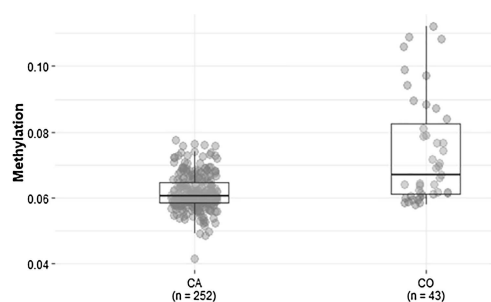


Fig. 1 Boxplot for results in the discovery cohort. Boxplot for the median of the replicated CpG sites in IS patients and controls. Y-axis for the methylation level (β values), and X-axis for cases and controls. CA: IS patients; CO: controls. IS, ischemic stroke.

($p = 4.42 \times 10^{-12}$). DMR overlapping with *ZFH3* was among the most significant DMR results ($p = 3.66 \times 10^{-8}$) (**►Table 3** and **►Supplementary Table S2**, available in the online version).

Replication Analysis

From the 957 significant DMPs in the discovery cohort with the intention to replicate, 938 were available in the replication cohort. Among these CpG sites, 869 CpG sites were significant ($p < 0.05$) in the replication cohort, but 48 were in the inverse effect size direction. A total of 821 CpG sites were consistent with the discovery results in effect size direction and significance level. A total of 697 CpG sites remained significant after accounting for the 957 CpG sites included in

Table 3 DMR in the discovery cohort

CHR	Start	End	no.cpgs	p-Value	Genes
chr5	56110833	56112222	8	4.42E-12	MAP3K1
chr15	101792098	101792838	9	6.20E-9	CHSY1
chr2	242576411	242577535	8	6.70E-9	ATG4B
chr9	123476683	123477254	2	1.77E-8	MEGF9
chr16	73091917	73092823	4	3.66E-8	ZFH3
chr16	67062875	67064157	10	6.48E-8	CBFB
chr8	74206531	74208178	8	7.66E-8	SNORA7
chr17	80477010	80477962	7	5.80E-7	SNORA69
chr17	19770908	19770910	2	8.32E-7	SNORA69
chr5	14871469	14872551	10	8.88E-7	CTB-40H15.4
chr12	44229512	44230058	5	9.76E-7	TMEM117
chr5	78280882	78281103	2	1.02E-6	snoZ6
chr9	139010494	139010884	2	1.21E-6	C9orf69
chr21	44394436	44395084	7	1.54E-6	PKNX1
chr9	89560881	89561881	5	2.56E-6	GAS1
chr10	5707981	5708674	3	3.30E-6	ASB13
chr16	4852372	4852830	4	3.50E-6	ROGDI
chr3	12598618	12598824	3	3.92E-6	MKRN2
chr1	224370909	224371029	2	4.11E-6	snoU13
chr6	155316380	155316492	2	4.45E-6	TIAM2
chr10	64564801	64565336	3	4.62E-6	ADO
chr1	14075870	14076293	5	5.09E-6	snoU13
chr15	40574544	40574925	3	5.13E-6	PLCB2
chr22	39101979	39102110	3	5.46E-6	GTPBP1
chr16	88772396	88772985	6	6.38E-6	CTU2
chr5	132083105	132083184	2	7.14E-6	CCNI2
chr4	1857817	1858231	2	7.21E-6	LETM1
chr10	64133674	64134403	10	7.64E-6	ZNF365
chr10	92617562	92617735	2	9.06E-6	SNORA17

Note: Summary statistics for the most significant ($p < 10e^{-7}$) DMR results in the discovery cohort. Columns from left to right: CHR: chromosome; Start: initial base-pair position for a specific DMR; End: final base-pair position for a specific DMR; no.cpg: number of CpG sites included in the DMR.

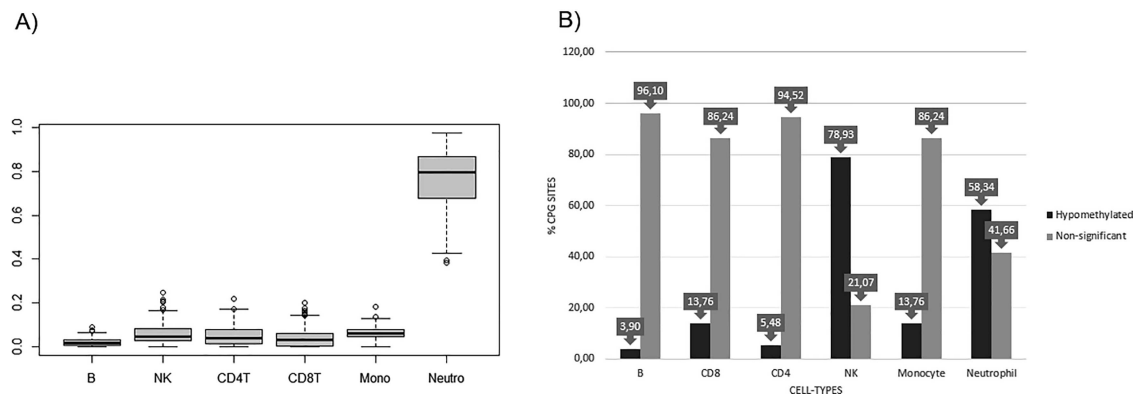


Fig. 2 DMCT for significant CpG sites. (A) Estimated proportion of cell types. (B) Proportion of significant and nonsignificant CpG sites (from the replicated CpG sites associated with stroke) in the different cell types. TSS1500: transcription start site (TSS) between 200 and 1,500 nucleotides from TSS; TSS200: transcription start site (TSS) between 0 and 200 nucleotides from TSS. DMCT, differentially methylated cell-type.

the “intention to replicate” analysis (►Table 2 and ►Supplementary Table S1, available in the online version).

Differentially Methylated Cell Types

The estimated proportion of different cell types was consistent with that expected in EWAS from whole blood,²⁷ being neutrophils the major cell type (►Fig. 2A). The estimated proportion for eosinophils was null. We evaluated the 821 significant and replicated CpG sites using the CellDMC algorithm to identify DMCT, specific cell types driving the observed associations. NK cells were the cells identified to be more strongly driving the identified epigenetic associations with stroke risk, with 78.9% of the evaluated CpG sites being associated ($p < 0.05$) with stroke risk for this cell type (►Fig. 2B). In neutrophils, 58.34% of the CpG sites were associated with stroke risk. For the other cell types, less than 15% of the CpG sites were associated with stroke risk.

Features for the Replicated CpG Sites

From the 821 replicated CpG sites, 99.8% were significantly hypomethylated in stroke patients versus controls, contrasting with the 58% when all the analyzed CpG sites were considered ($n = 384,509$).

When considering location for all the CpG sites analyzed, the higher percentage (33.3%) was in the body of genes, followed by intergenic regions (24.65%) (►Supplementary Fig. S3A, available in the online version). However, when only the 821 replicated DMPs were considered, the higher percentage (32.06%) was in the first exon followed by the promotor regions of genes (34.37%). Moreover, while the analyzed CpG sites were distributed in islands (31%), open seas (36%), shelf (10%), and shore (23%), almost all (96%) of the significant and replicated CpG sites were in islands (►Supplementary Fig. S3B, available in the online version).

Stroke Subtype Sub-analysis

Evaluation of methylation pattern of the significant and replicated DMPs by stroke subtypes showed a specificity trend toward atherothrombotic, cardioembolic, and unde-

termined stroke subtypes (►Supplementary Table S2, available in the online version). All the 821 replicated CpG sites were significantly associated with the three stroke subtypes ($p < 0.05$). In total, 546, 648, and 721 CpG sites for atherothrombotic, cardioembolic, and undetermined stroke subtypes, respectively, remained significant after multiple comparison correction by Bonferroni ($p < 6.09 \times 10^{-5}$). More than half of the replicated CpG sites ($n = 696$) were significantly associated ($p < 0.05$) with lacunar stroke. However, only 15 were significant after Bonferroni correction (►Supplementary Table S3, available in the online version).

To discard that the identified stroke-subtype associations were influenced by infarct volume, we used NIHSS at baseline and at 24 hours as surrogate markers of infarct size. We analyzed the correlation of replicated CpG sites with baseline and 24-hour NIHSSs. None of the analyzed CpG sites correlated with NIHSS after Bonferroni correction.

Pathway Enrichment Analysis and Network Analysis

All the genes for the replicated CpG sites were analyzed for GO pathway enrichment (►Fig. 3). The most significant enriched pathways were related with phagocytosis: immune response-regulating cell surface receptor signaling pathway involved in phagocytosis and Fc-gamma receptor signaling pathway involved in phagocytosis ($p = 1.1 \times 10^{-4}$ in both cases). Other significant pathways enriched in the results were negative regulation of hormone secretion, response to epidermal growth factor, hematopoietic stem cell differentiation, and cellular response to oxidative stress (►Fig. 3). All these pathways included more than 15 genes from the replicated results.

Analysis of Genes from Stroke GWAS

A total of 1,160 CpG sites were located in any of the 43 genes considered in the analysis (►Supplementary Table S4, available in the online version). A total of 129 CpG sites located in 30 different loci were associated with IS ($p < 0.05$) and eight CpG sites from six loci (in *ZFXH3*, *SH2B3*, *SMARCA4*, *TSPAN2*, *ILF3*, and *CDK6* genes) remained significant after Bonferroni adjustment (5×10^{-5}). All of them presented

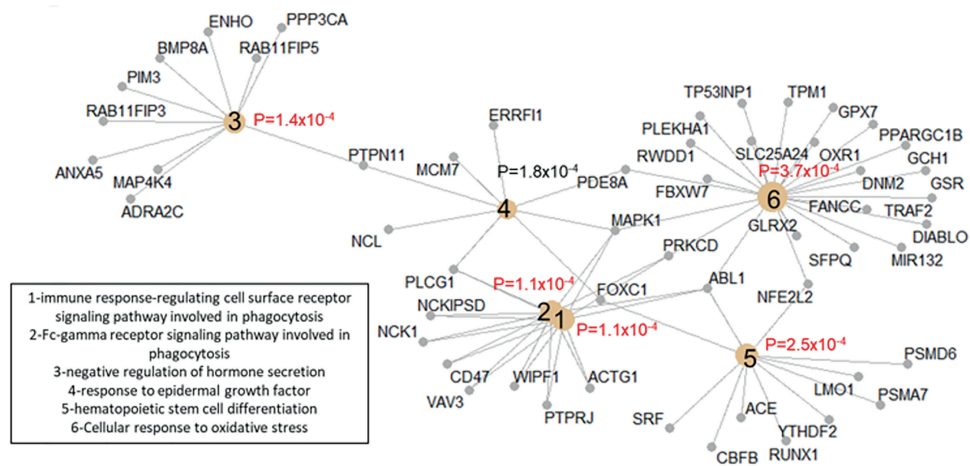


Fig. 3 Pathway analysis for genes from significant and replicated CpG sites. The five more significant clusters of significant GO-enriched pathways with the genes belonging to each cluster.

hypomethylation in stroke patients compared with controls (► **Supplementary Table S5**, available in the online version).

For the other stroke subtypes, 93, 76, 53, and 106 CpG sites were nominally associated with LAS, CES, SVS, and UND stroke, respectively. These CpG sites corresponded to 30, 24, 23, and 5 different loci, respectively (► **Supplementary Table S6**, available in the online version). We specifically checked the association by stroke subtype of the CpG sites located in the six loci associated with all IS after Bonferroni correction. *ZFHX3*, *SH2B3*, *SMARCA4*, and *TSPAN2* were significant in LAS and CES subtypes while CpGs in *ILF3* and *CDK6* were not significantly associated in any specific stroke subtype (► **Supplementary Table S6**, available in the online version).

MR Analysis

To establish causality between the methylation pattern of the two top CpG sites (cg07786668 and cg10115490) and

stroke risk, we performed MR analysis. First, we looked for mQTLs for each of the CpG sites (► **Supplementary Table S7**, available in the online version). We found 95 and 56 mQTLs with p -values $\leq 10^{-7}$ for *MAP3K1* and *ZFHX3* CpG sites, respectively. We used these mQTLs as instruments in the MR analysis. IVW showed evidence of methylation in *MAP3K1* to be the cause of cardioembolic and large-artery stroke ($p = 0.0224$ and 0.0331 , respectively). Moreover, MR-Egger and WM test showed the same effect direction. Neither horizontal pleiotropy ($p = 0.16$) nor heterogeneity ($p = 0.95$) was found (► **Fig. 4**). For *ZFHX3*, IVW was significant for cardioembolic, large artery, and small-vessel stroke subtypes ($p = 0.0491$, 0.0331 , and 0.0104 , respectively). MR-Egger and WM were in accordance with IVW. In this case, nonsignificance was found for horizontal pleiotropy ($p = 0.12$, 0.29 , and 0.47) and heterogeneity ($p = 0.47$, 0.75 , and 0.76) (► **Fig. 4**). Moreover, we evaluated for possible reverse causality of stroke causing changes in methylation of *MAP3K1* and

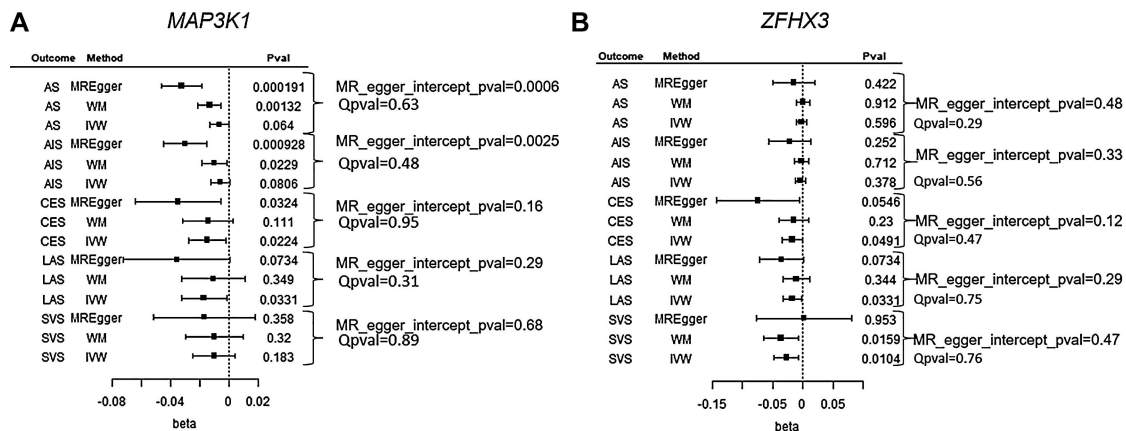


Fig. 4 MR results for the *MAP3K1* and *ZFHX3* CpG sites. MR results for (A) *MAP3K1* CpG site and (B) *ZFHX3* CpG site. Outcome indicates the different GWAS for stroke evaluated: all stroke (AS), all ischemic stroke (AIS), cardioembolic stroke (CES), large-artery stroke (LAS), and small vessel stroke (SVS). The method indicates the MR test used: MR-Egger, weighted-median (WM), and inverse-variance weighted (IVW). Horizontal pleiotropy (MR-Egger intercept) and heterogeneity (Qpval) results are illustrated at the right of each figure. MR, Mendelian randomization.

ZFHX3 CpG sites. No significant signal for reverse causality was obtained (► **Supplementary Fig. S4**, available in the online version).

Discussion

We have identified and replicated a high number of CpG sites ($N = 821$) located in 836 different loci associated with IS. In comparison with the published EWAS for IS, we have expanded the number of CpG sites significantly associated with IS from 22 to 803 and the number of significant loci from 21 to 709.¹⁴ Moreover, the inflation and bias calculated for our study using a specific method developed for EWAS analysis showed lower levels for both parameters. Almost all the DMPs were highly significantly hypomethylated in IS patients compared with controls. The same was observed in a manuscript analyzing global methylation in IS patients.¹³ Our results suggest that NK cells mostly drive the observed associations. NK cells are immune T cells that are able to infiltrate and reside into the brain in different neurological conditions. In stroke, NK cells have been described to have many different roles, involving inflammation, immunodepression, and infections poststroke.^{30,31} It has been observed that NK cells are associated with the stroke outcome by accelerating brain infarction.³²

In our results, the most proportion of the significant CpG sites was in the first exon (32.06%) of genes. These results contrast with the location of CpG sites not associated with stroke risk, with only 4.68% of the CpGs located in the first exon. Moreover, published data revealed that, despite exonic methylation being much more common than previously expected, methylation in the first exon is spared compared with downstream exons.³³ Methylation in the downstream region of promoters, at the first exon, is much more linked to gene repression compared with methylation in the upstream of promoters.³³ Thus, our results seem to support an increase in expression of genes that are related to stroke.

Moreover, our stroke subtype analysis showed that this methylation is not exclusive for one specific stroke subtype. All the replicated CpG sites are associated with risk of cardioembolic, atherothrombotic, and undetermined stroke subtypes. However, these associations are weaker for lacunar stroke, probably showing the different physiopathology for this specific stroke subtype.³⁴ As lacunar stroke normally affects small vessels, it is possible that DNA methylation affects the risk of this stroke subtype in a different way. To ensure that the stroke-subtype methylation results were not masked by infarct sizes, we used two surrogate markers of infarct volume, NIHSS at baseline and at 24 hours poststroke. We did not identify any methylation correlation for the replicated CpG sites with NIHSS. These results indicated that methylation, in contrast to results from genomic data, globally affect a common mechanism for all the stroke subtypes.^{35,36} However, to confirm this hypothesis, stroke volume quantification would be needed.

Two EWASs have been published related to stroke risk. The first one was focused on atherothrombotic stroke and included 12 patients with this stroke subtype and 12 con-

trols.¹² The authors identified one CpG site (cg05740793) specifically associated with atherothrombotic stroke.¹² In the subgroup of atherothrombotic stroke patients and controls from our discovery cohort ($n = 47$ and 13, respectively), we obtained a p -value of 0.73 for this CpG. In the study, that region was replicated using Sequenom EpiTYPER, a technique that allows the determination of DNA methylation of multiple CpGs in a genomic region of 100 to 600 base pairs³⁷ instead of a specific analysis of a CpG site. The second EWAS analysis studied the methylation of IS and controls and identified 22 CpG sites differentially methylated in stroke versus controls.¹⁴ However, the difference with our results could be explained by the different time of blood extraction (12 vs. 24 hours) and the different methods for analysis used.

The genetic influence on a disease could be mediated by different processes at a different level: epigenetic, transcriptomic, and proteomics. In stroke, 35 loci have been identified using GWAS to be associated with risk of stroke. The lead variant or a variant in linkage disequilibrium ($ld > 0.8$) for 18 out of the 35 loci was found to overlap with an expression, methylation, or proteomic quantitative trait locus (eQTL, mQTL, or pQTL, respectively).⁵ However, whether the effect of these polymorphisms at the methylation level is causative for the stroke risk is not known. Two loci whose lead variant overlapped with a mQTL (*SH2B3* and *CDK6*) were identified in our EWAS to be associated with stroke risk. These results suggest that methylation in these genes could be mediating the relationship between the genetic variation and the stroke risk. We found four additional stroke risk loci identified in GWAS whose methylation is associated in our EWAS. In this case, the lead variant for each locus has not been previously associated with methylation level, suggesting that the genetic and epigenetic association of these loci with stroke is independent. The independence of these associations is confirmed by the different regulations that genetics and epigenetics exert over specific stroke subtypes.

ZFHX3 was clearly associated with cardioembolic stroke at the genetic level,⁵ but at the methylation level it is also associated with atherothrombotic and lacunar stroke. In contrast, *TSPAN2* was associated with atherothrombotic stroke at the genetic level but with cardioembolic stroke at the epigenetic level. The most significant CpG in our study was cg07786668, located in the *ZFHX3* gene. This gene encodes a transcription factor that regulates myogenic and neuronal differentiation. Our stroke subtype analysis identified hypomethylation of this gene associated with cardioembolic stroke and MR analysis showed a causal relation between cg07786668 and cardioembolic stroke. These results are in accordance with published GWAS²⁻⁴ that found genetic variations in this gene associated with an increased risk of cardioembolic stroke. Additionally, this CpG site has been previously associated with myocardial infarction in an EWAS with almost 400 subjects.³⁸

The second most significant CpG site in our results was in *MAP3K1* (MEKK1), a member of the MAPK signaling pathway. Moreover, another CpG site in this gene showed epigenome-wide significance. MEKK1 is an upstream MAP kinase-kinase that activates p38 and JNK pathways in response

to ischemia and other stimuli.³⁹ One in vitro study demonstrated that hypoxia in cerebral endothelial cells induced the activation of MEKK1 and the transduction pathway to respond to this stimulus.³⁹

Other members of the MAPK signaling pathway are significant in our results, which are involved in the transmission of signals in response to different stimuli such as ischemia or inflammation.⁴⁰ Recently, one bioinformatic analysis including different genes has identified the involvement of *MAPK1*, a significant gene in our results from the MAPK pathway, in the signaling pathway of IS.⁴¹ We have found in DrugBank that *MAPK1* is a target for approved drugs which have indication (acetylsalicylic acid) or are being studied in clinical trials (minocycline) to be used in stroke patients. Moreover, we identified another gene from our results (*EGLN1*) which is a target for acetaminophen with several clinical trials investigating its indication in stroke patients.

The enrichment pathway analysis showed that phagocytosis is over-represented among significant genes in our results. Phagocytosis mediates immuno-response in the brain and is important in preventing atherosclerosis.⁴² Other pathways are enriched in our results: response to epidermal growth factor, which has been seen in animal models to stimulate endogenous neural precursor cell and then promoting tissue regeneration.⁴³ Additionally, hormone secretion, hematopoietic cell differentiation, and oxidative stress are over-represented in our results and all of them have been described to be related with stroke.^{44,45}

Moreover, as far as we know, this is the first time that DMCT analysis is applied in an EWAS of stroke. Furthermore, no previous MR analysis is applied in epigenetic studies of stroke to determine the causality of specific methylation sites over stroke risk. Methylation studies have the limitation of understanding what is occurring first: methylation or the disease. Our results suggested that altered DNA methylation in *MAP3K1* and *ZFHX3* is a risk factor for cardioembolic and atherothrombotic stroke occurrences. In addition, *ZFHX3* methylation is also a risk factor for lacunar stroke.

Our study has some limitations. First, the sample size in the discovery cohort is small for the sub-analysis of stroke subtypes. However, considering the discovery and replication cohorts, it is the higher sample size for any EWAS in IS. Moreover, the statistical power calculated to obtain effect sizes similar to those published in previous EWAS was higher than 0.8, enough to obtain statistically significant results ($p \leq 10^{-07}$). Additionally, we have replicated the associations in an independent cohort. Finally, functional analysis will be needed to truly assess the effect of the differential methylation pattern on these genes.

Conclusion

We have identified a total of 803 CpG sites corresponding to 709 different loci not previously associated with stroke. Moreover, 99.8% of these CpG sites were hypomethylated in stroke patients compared with controls. Causative assessment with MR analysis indicated causal association of methylation of the

top CpG sites over some IS subtypes. All these results strongly support a role of epigenetics in stroke risk that could be useful for specifically predicting stroke occurrence and could potentially be useful for future therapeutic targets.

What is known about this topic?

- DNA methylation analysis using a genome-wide approach (EWAS) is analogous to a genome-wide association study (GWAS) but in epigenetics that is very useful to identify differentially methylated positions (DMPs) associated with a disease or phenotype.
- Different studies using EWAS have found that DNA methylation is associated with different pathologies and conditions (e.g., cancer, smoking, cardiovascular disease, metabolic syndrome, rheumatoid arthritis, etc.)
- In stroke, epigenetic changes through DNA methylation alteration have been associated with general risk of stroke, with subtype-specific stroke risk, with stroke recurrence, and with stroke recurrence in patients treated with aspirin or clopidogrel.

What does this paper add?

- We have expanded the total number of loci whose DNA methylation is associated with stroke risk from 21 to 709.
- We have applied differentially methylated cell-type (DMCT) analysis in an EWAS of stroke for the first time. The results indicate that the observed DNA methylation associations are mostly driven by natural killer (NK) cells.
- No previous Mendelian randomization (MR) analysis was previously applied in an epigenetic study of stroke to determine the causality of methylation over stroke risk. Our results suggested that altered DNA methylation in *MAP3K1* and *ZFHX3* is causally associated with stroke risk.

Funding

The EPIGENESIS project (Carlos III Institute—PI17/02089, Marató TV3 and Fundació MútuaTerrassa), the MAESTRO project—PI18/01338 (Carlos III Institute), the iBioStroke project (Eranet-Neuron, European research grants), the EPINEXO project—PI20/00678 (Carlos III Institute), the SEDMAN Study (Boehringer Ingelheim), the APHAS project (Pfizer/Bristol-Myers Squibb), and the European Regional Development Fund (ERDF)/Fondo Europeo de Desarrollo Regional (FEDER) and 2017SGR-1427 (AGAUR). I.F.-C is the recipient of a research contract from the Miguel Servet Program (CP12/03298) from the Carlos III Institute. J.C.-M. is supported by an AGAUR contract (Agència de Gestió d'Ajuts Universitaris i de Recerca; FI_DGR 2019, grant number 2019_FI_B 00853) co-financed by Fons Social Europeu (FSE). M.L. is supported by a PFIS contract (Contratos Predoctorales de Formación

en Investigación en Salud): FI19/00309. C.G.-F is supported by a Sara Borrell contract (CD20/00043) from the Carlos III Institute and Fondo Europeo de Desarrollo Regional (ISCIII-FEDER). E.M. is supported by a Río Hortega contract (CM18/00198) from the Carlos III Institute.

Conflict of Interest

None declared.

Acknowledgments

We thank the International Stroke Genetics Consortium, the Spanish Stroke Genetics Consortium, and the RETICS Network INVICTUS (RD16/0019/0002, RD16/0019/0010, RD16/0019/0011, RD16/0019/0021).

References

- Bevan S, Traylor M, Adib-Samii P, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke* 2012;43(12):3161–3167
- Traylor M, Farrall M, Holliday EG, et al; Australian Stroke Genetics Collaborative, Wellcome Trust Case Control Consortium 2 (WTCCC2) International Stroke Genetics Consortium. Genetic risk factors for ischaemic stroke and its subtypes (the META-STROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol* 2012;11(11):951–962
- Malik R, Chauhan G, Traylor M, et al; AFGen Consortium Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium International Genomics of Blood Pressure (iGEN-BP) Consortium INVENT Consortium STARNET BioBank Japan Cooperative Hospital Group COMPASS Consortium EPIC-CVD Consortium EPIC-InterAct Consortium International Stroke Genetics Consortium (ISGC) METASTROKE Consortium Neurology Working Group of the CHARGE Consortium NINDS Stroke Genetics Network (SiGN) UK Young Lacunar DNA Study MEGASTROKE Consortium. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* 2018;50(04):524–537
- NINDS Stroke Genetics Network (SiGN) International Stroke Genetics Consortium (ISGC) Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol* 2016;15(02):174–184
- Malik R, Rannikmäe K, Traylor M, et al; MEGASTROKE consortium and the International Stroke Genetics Consortium. Genome-wide meta-analysis identifies 3 novel loci associated with stroke. *Ann Neurol* 2018;84(06):934–939
- Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology* 2013;38(01):23–38
- Chambers JC, Loh M, Lehne B, et al. Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol* 2015;3(07):526–534
- Rask-Andersen M, Martinsson D, Ahsan M, et al. Epigenome-wide association study reveals differential DNA methylation in individuals with a history of myocardial infarction. *Hum Mol Genet* 2016;25(21):4739–4748
- Gallejo-Fabrega C, Carrera C, Reny JL, et al. TRAF3 epigenetic regulation is associated with vascular recurrence in patients with ischemic stroke. *Stroke* 2016;47(05):1180–1186
- Gallejo-Fabrega C, Carrera C, Reny JL, et al. PPM1A methylation is associated with vascular recurrence in aspirin-treated patients. *Stroke* 2016;47(07):1926–1929
- Davis Armstrong NM, Chen WM, Brewer MS, et al. Epigenome-wide analyses identify two novel associations with recurrent stroke in the vitamin intervention for stroke prevention clinical trial. *Front Genet* 2018;9:358
- Shen Y, Peng C, Bai Q, et al. Epigenome-wide association study indicates hypomethylation of MTRNR2L8 in large-artery atherosclerosis stroke. *Stroke* 2019;50(06):1330–1338
- Soriano-Tárraga C, Jiménez-Conde J, Giralte-Steinhauer E, et al; GeneStroke “The Spanish Stroke Genetics Consortium” Global DNA methylation of ischemic stroke subtypes. *PLoS One* 2014;9(04):e96543
- Soriano-Tárraga C, Lazcano U, Giralte-Steinhauer E, et al. Identification of 20 novel loci associated with ischaemic stroke. *Epigenetics* 2020;15(09):988–997
- Min JL, Hemani G, Davey Smith G, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics* 2018;34(23):3983–3989
- Fernández-Cadenas I, Mendióroz M, Giralte D, et al; GRECOS Study Group. GRECOS project (genotyping recurrence risk of stroke): the use of genetics to predict the vascular recurrence after stroke. *Stroke* 2017;48(05):1147–1153
- Carrera C, Cullell N, Torres-Águila N, et al; Spanish Stroke Genetic Consortium. Validation of a clinical-genetics score to predict hemorrhagic transformations after rTPA. *Neurology* 2019;93(09):e851–e863
- Roquer J, Rodríguez-Campello A, Gomis M, et al. Acute stroke unit care and early neurological deterioration in ischemic stroke. *J Neurol* 2008;255(07):1012–1017
- Sayols-Baixeras S, Lluís-Ganella C, Subirana I, et al. Identification of a new locus and validation of previously reported loci showing differential methylation associated with smoking. *The REGICOR study*. *Epigenetics* 2015;10(12):1156–1165
- Morris TJ, Butcher LM, Feber A, et al. ChAMP: 450k chip analysis methylation pipeline. *Bioinformatics* 2014;30(03):428–430
- Graw S, Henn R, Thompson JA, Koestler DC. pwrEWAS: a user-friendly tool for comprehensive power estimation for epigenome wide association studies (EWAS). *BMC Bioinformatics* 2019;20(01):218
- Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res* 2017;45(04):e22
- LaBarre BA, Goncarenco A, Petrykowska HM, et al. MethylToSNP: identifying SNPs in Illumina DNA methylation array data. *Epigenetics Chromatin* 2019;12(01):79
- Adams HJr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in acute stroke treatment. *Stroke* 1993;24(01):35–41
- van Iterson M, van Zwet EW, Heijmans BTBIOS Consortium. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biol* 2017;18(01):19
- Peters TJ, Buckley MJ, Statham AL, et al. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin* 2015;8:6
- Zheng SC, Breeze CE, Beck S, Teschendorff AE. Identification of differentially methylated cell types in epigenome-wide association studies. *Nat Methods* 2018;15(12):1059–1066
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16(05):284–287
- Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenotype. *eLife* 2018;7:e34408
- Chen C, Ai QD, Chu SF, Zhang Z, Chen NH. NK cells in cerebral ischemia. *Biomed Pharmacother* 2019;109:547–554

- 31 Gan Y, Liu Q, Wu W, et al. Ischemic neurons recruit natural killer cells that accelerate brain infarction. *Proc Natl Acad Sci U S A* 2014;111(07):2704–2709
- 32 Wang Y, Liu J, Wang X, et al. Frequencies of circulating B- and T-lymphocytes as indicators for stroke outcomes. *Neuropsychiatr Dis Treat* 2017;13:2509–2518
- 33 Brenet F, Moh M, Funk P, et al. DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS One* 2011;6(01):e14524
- 34 Caplan LR. Lacunar infarction and small vessel disease: pathology and pathophysiology. *J Stroke* 2015;17(01):2–6
- 35 Tong DC, Yenari MA, Albers GW, O'Brien M, Marks MP, Moseley ME. Correlation of perfusion- and diffusion-weighted MRI with NIHSS score in acute (<6.5 hour) ischemic stroke. *Neurology* 1998;50(04):864–870
- 36 Fink JN, Selim MH, Kumar S, et al. Is the association of National Institutes of Health Stroke Scale scores and acute magnetic resonance imaging stroke volume equal for patients with right- and left-hemisphere ischemic stroke? *Stroke* 2002;33(04):954–958
- 37 Suchiman HED, Slieker RC, Kremer D, Slagboom PE, Heijmans BT, Tobi EW. Design, measurement and processing of region-specific DNA methylation assays: the mass spectrometry-based method EpiTYPER. *Front Genet* 2015;6:287
- 38 Nakatochi M, Ichihara S, Yamamoto K, et al. Epigenome-wide association of myocardial infarction with DNA methylation sites at loci related to cardiovascular disease. *Clin Epigenetics* 2017;9:54
- 39 Lee SR, Lo EH. Interactions between p38 mitogen-activated protein kinase and caspase-3 in cerebral endothelial cell death after hypoxia-reoxygenation. *Stroke* 2003;34(11):2704–2709
- 40 Sun J, Nan G. The mitogen-activated protein kinase (MAPK) signaling pathway as a discovery target in stroke. *J Mol Neurosci* 2016;59(01):90–98
- 41 Eyileten C, Wicik Z, De Rosa S, et al. MicroRNAs as diagnostic and prognostic biomarkers in ischemic stroke-a comprehensive review and bioinformatic analysis. *Cells* 2018;7(12):E249
- 42 Kojima Y, Volkmer JP, McKenna K, et al. CD47-blocking antibodies restore phagocytosis and prevent atherosclerosis. *Nature* 2016;536(7614):86–90
- 43 Wang Y, Cooke MJ, Lapitsky Y, et al. Transport of epidermal growth factor in the stroke-injured brain. *J Control Release* 2011;149(03):225–235
- 44 Sohrabji F, Okoreeh A, Panta A. Sex hormones and stroke: beyond estrogens. *Horm Behav* 2019;111:87–95
- 45 Courties G, Frodermann V, Honold L, et al. Glucocorticoids regulate bone marrow B lymphopoiesis after stroke. *Circ Res* 2019;124(09):1372–1385

SUPPLEMENTARY MATERIAL

Altered methylation pattern in EXOC4 is associated with stroke outcome: An Epigenome-Wide Association study

Natalia Cullell, MSc^{1,2,3}, Carolina Soriano-Tárraga, PhD⁴, Cristina Gallego-Fábrega, PhD¹, Jara Cárcel-Márquez, MSc¹, Elena Muiño, MD¹, Laia Lluçia-Carol MSc¹, Miquel Lledós MSc¹, Manel Esteller, PhD^{7,8,9}, Manuel Castro, PhD⁷, Joan Montaner, MD, PhD^{5,10}, Anna Rosell, PhD⁵, Roberto Elosúa, MD¹¹, Alba Fernández-Sanlés, PhD¹¹, Joan Martí-Fàbregas, MD, PhD¹², Jerzy Krupinski, MD, PhD^{2,6}, Jaume Roquer, MD⁴, Jordi Jiménez-Conde, MD, PhD^{4*}, Israel Fernández-Cadenas, PhD^{1*}

1- Stroke Pharmacogenomics and Genetics, Institut de Recerca de Sant Pau, Hospital Sant Pau, Barcelona, Spain.

2- Neurology. Hospital Universitari MútuaTerrassa/ Fundacio Docència i Recerca MutuaTerrassa, Terrassa, Spain

3- Facultat de Medicina. Universitat de Barcelona, Barcelona, Spain

4- Neurology, Hospital del Mar; Neurovascular Research Group, IMIM; Universitat Autònoma de Barcelona / DCEXS-Universitat Pompeu Fabra, Barcelona, Spain

5-Department of Psychiatry, Washington University School of Medicine, Saint Louis, Missouri, US

6-NeuroGenomics and Informatics, Washington University School of Medicine, Saint Louis, Missouri, US

7-Cancer Epigenetics & Biology Program (PEBC), L'Hospitalet, Barcelona, Spain

8-Department of Physiological Sciences II, School of Medicine, Universitat de Barcelona, Barcelona, Spain

9-Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

10- Neurovascular Research Laboratory, Vall d'Hebron Institut de Recerca (VHIR), Barcelona, Spain

11-Instituto de Biomedicina de Sevilla, IBiS/Hospital Universitario Virgen del Rocío / CSIC. Universidad de Sevilla & Department of Neurology, Hospital Universitario Virgen Macarena Sevilla, Spain

12- Neurology, Hospital de la Santa Creu i Sant, Barcelona, Spain

13- Centre for Bioscience, School of HealthCare Science, Manchester Metropolitan University, Manchester, UK

1-Supplemental Materials and Methods

Metabolic pathway enrichment and feature enrichment analysis

We looked for enrichment of metabolic pathways using as input all the CpG-sites and the CpG-sites that were nominally associated ($p\text{-value} < 10^{-06}$) with ΔNIHSS in the meta-analysis. The analysis was done with MethylGSA (1), a R package specifically designed for pathway analysis from EWAS results. We applied the three functions from the package (all of them designed to adjust for the number of CpGs in each gene to reduce possible bias). Briefly, the first function, methylglm, performs a logistic regression using the number of CpG-sites per gene as covariate. We applied this function for all the CpG types but also restricting the analysis to CpG-sites from promoters (TSS1500 or TSS200) or located in gene bodies. This function does not allow to indicate a p-value threshold for CpG-sites to be included, so we did the analysis including the list of all the CpG-sites analyzed with their meta-analysis p-value. The second function, methylRRA, applies an over-representation analysis (ORA) or Preranked version of Gene Set Enrichment Analysis (GSA-Preranked) in the gene pathway analysis (after correcting for the different p-values in each gene using Robust Rank Aggregation). We performed the analysis with all the CpG-site types, with all the CpG-sites in promoters and with all the CpG-sites in gene bodies. The third function, methylgometh, uses the weighted resampling and Wallenius non-central hypergeometric approximation to adjust for the number of CpGs in each gene. This function does not allow to select CpG-sites only from promoters or gene bodies to be analyzed. With the second and third function we included the list of all the CpG-sites but also the list of nominally associated CpG-sites from the meta-analysis. We included in the analyses the three available pathway databases in MethylGSA: Gene Ontology, KEGG and Reactome. We

considered significant associations when the p-value was < 0.05 and the Q-value (False-Discovery rate (FDR) adjusted p-value) was < 0.05 .

Tissue-specific signal detection

We used eFORGE (experimentally-derived Functional element Overlap analysis of ReGions from EWAS) (2) to estimate tissue specific signals from the significant and nominal results from the EWAS meta-analysis. Briefly, eFORGE looks for enrichment of functional elements in the target CpG-sites compared with background DMPs. This tool provides functional data from ENCODE, Roadmap Epigenomics and Blueprint project. All the three projects describe DNase I hypersensitivity sites as functional element. Moreover, 15-state chromatin marks and histone 3 (H3) marks are available from the Roadmap Epigenomics project. We included analyses with all the different functional elements. A p-value < 0.05 was considered a nominal association and a Q-value < 0.05 according to FDR adjustment was considered statistically significant.

Blood–Brain Epigenetic Correlation

In order to compare the methylation from the significant findings in the meta-analysis between blood and brain, we used the Blood Brain DNA Methylation Comparison Tool (3). This tool correlates methylation level from the 450K Illumina array CpG-sites between blood and four brain regions (prefrontal cortex, entorhinal cortex, superior temporal gyrus and cerebellum). We analyzed the correlation of cg00039070 methylation and the four brain regions included in the tool.

Using *Blood–Brain Epigenetic Concordance* (BECon) (4), we also investigated the concordance in the cg00039070 methylation between three brain regions (Brodmann area (BA) 10, BA20 and BA7) and blood using the three metrics available in this tool: DNA methylation variability between samples in the specific CpG-site, correlation (Pearson correlation and percentile of correlation when comparing with all the CpG-sites) and the effect of cell composition adjustment in methylation.

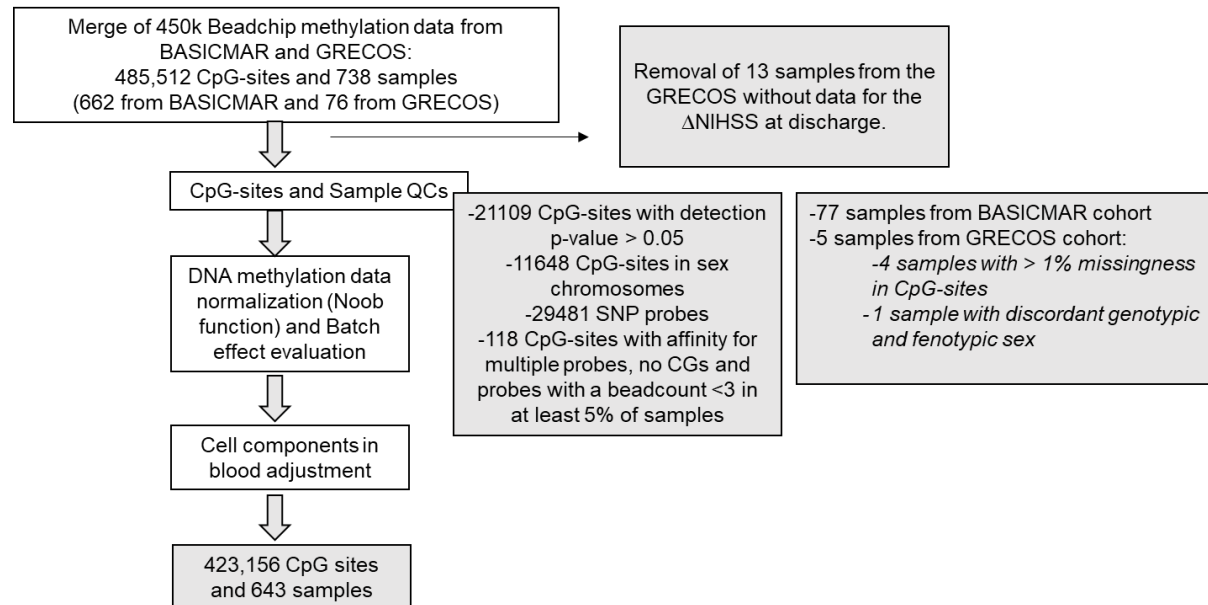
Blood–Brain Epigenetic Correlation

In order to compare the methylation from the significant findings in the meta-analysis between blood and brain, we used the Blood Brain DNA Methylation Comparison Tool (46). This tool correlates methylation level from the 450K Illumina array CpG-sites between blood and four brain regions (prefrontal cortex, entorhinal cortex, superior temporal gyrus and cerebellum). We analyzed the correlation of cg00039070 methylation and the four brain regions included in the tool.

Using *Blood–Brain Epigenetic Concordance* (BECon) (47), we also investigated the concordance in the cg00039070 methylation between three brain regions (Brodmann area (BA) 10, BA20 and BA7) and blood using the three metrics available in this tool: DNA methylation variability between samples in the specific CpG-site, correlation (Pearson correlation and percentile of correlation when comparing with all the CpG-sites) and the effect of cell composition adjustment in methylation.

2- Supplemental e-FIGURES

Figure I: Workflow for CpG-sites and sample QCs



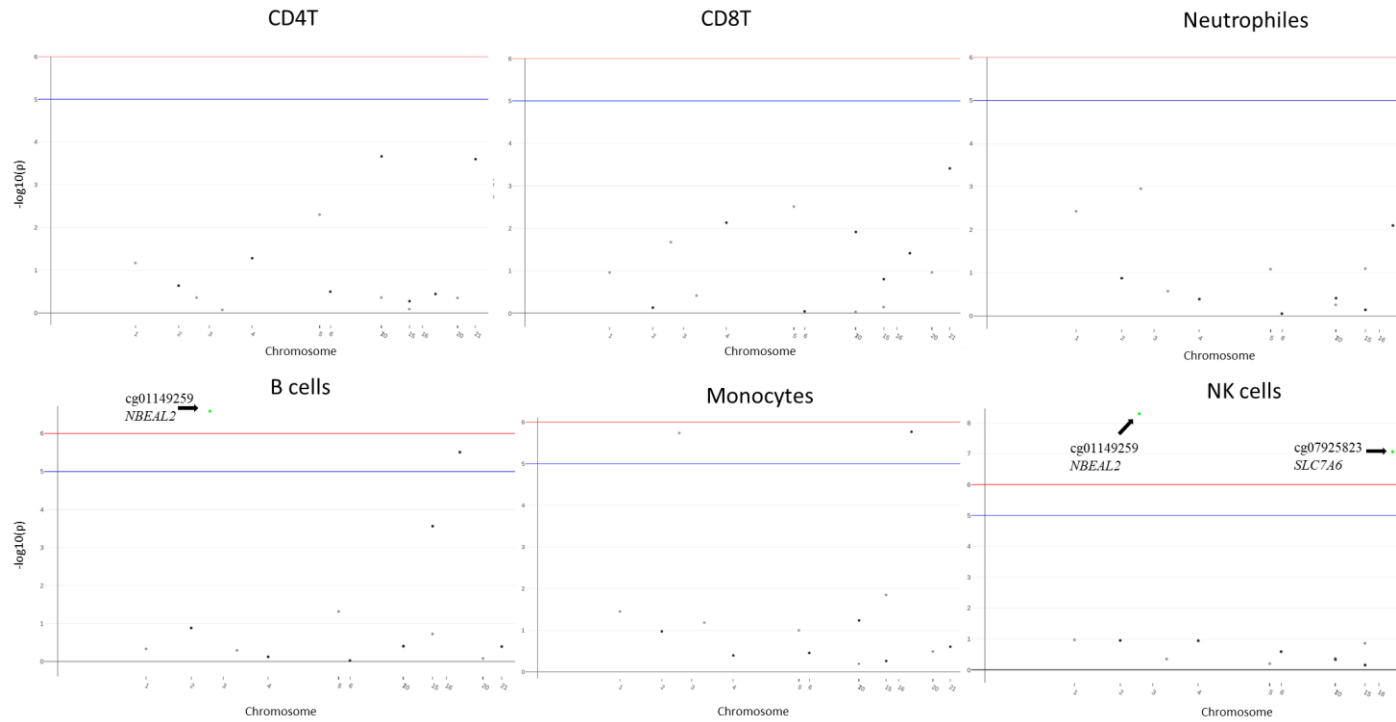
Description for the QCs applied to the Discovery cohorts.

Figure II: Batch effect evaluation with MDS and SVD plots



A) MDS plot stratifying samples with colours for batch effect (two different batch for the BASICMAR cohort and two batch for the GRECOS cohort) differentiation. Each colour identifies a different batch included in the Discovery cohort. B) Single Value Decomposition Analysis (SVD) to study the effect of principal component 1 and 2 (PC1 and PC2) and the batch (BASICMAR vs GRECOS cohorts) in our datasets.

Figure III: Manhattan plots for DMCT



Manhattan plot for the nominal CpG-sites in the meta-analysis stratified by cell-types: CD4+ T-cells, CD8+ T-cells, Neutrophils, B-cells, Monocytes and natural killer (NK) cells. Green dots represent CpG-sites epigenome-wide associated with Δ NIHSS in the specific cell-type. Red and blue lines represent the epigenome-wide and nominal significance threshold, respectively.

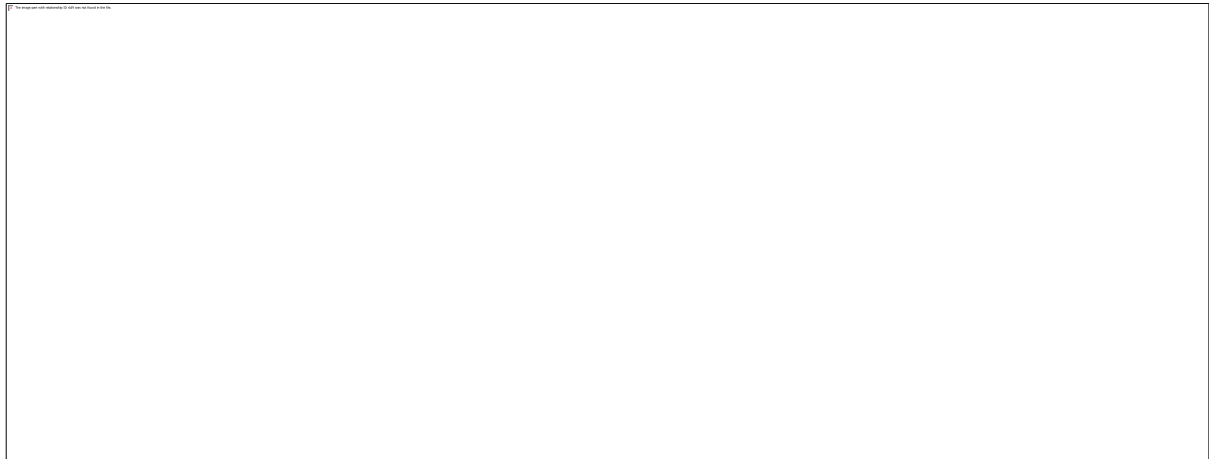
Figure IV: cg00039070 methylation correlation between blood and brain



Correlation coefficients for the DNA methylation in cg00039070 between four brain regions and blood.

PFC: prefrontal cortex; EC: entorhinal cortex; STG: superior temporal gyrus, CER: cerebellum

Figure V: Blood–Brain Epigenetic Concordance (BECon) results for cg00039070



Interpretation of cg00039070 methylation in blood in the context of brain using BECon. Results are subdivided in three brain regions: Brodmann area (BA) 10, BA20 and BA7. Three different measures were used to assess the concordance between methylation in the *EXOC4* CpG-site methylation in blood and brain. 1st) Variability of the methylation in the CpG-site between the analysed samples in BECon. The colour annotation (described in the legend) indicates whether the CpG methylation is variable or not; 2nd) The Pearson correlation coefficient between the brain tissue and blood. The colour annotation indicates the correlation percentil of the specific CpG-site compared with all the CpG-sites analysed and 3th) The measure of the cell composition strength on the CpG in blood and brain. The colour shows how much the beta values for methylation change when adjusting by cell composition.

3- Supplemental e-TABLES

Table I: Analysis of variables associated with Δ NIHSS in bivariate and regression analysis.

Δ NIHSS		
	Bivariate analysis	Backward Stepwise Regression
B_NIHSS	$<2.2 \times 10^{-16*}$	$<2.2 \times 10^{-16*}$
mRS 3	$8.35 \times 10^{-07*}$	$<2.2 \times 10^{-16*}$
Age	0.77	-
Sex	0.90	-
B_glucose	0.45	-
DM	0.09	-
DL	0.07	-
tPA	$3.2 \times 10^{-09*}$	$2.5 \times 10^{-02*}$
AF	$2.7 \times 10^{-04*}$	$1.75 \times 10^{-02*}$
HTN	0.46	-
SMK	0.13	-

Δ NIHSS: difference between baseline NIHSS and NIHSS at discharge; B_NIHSS: NIHSS at baseline; B_glucose: glucose levels measured at baseline; DM: diabetes mellitus; DL: dyslipidemia; tPA: treatment with rtPA; AF: atrial fibrillation; HTN: hypertension; SMK: smoking habit.

Table II: Analysis of variables associated with mRS at 3 months in bivariate and regression analysis.

mRS 3		
	Bivariate analysis	Backward Stepwise Regression
Δ NIHSS	$8.35 \times 10^{-07*}$	$< 2 \times 10^{-16*}$
B_NIHSS	$< 2.2 \times 10^{-16*}$	$< 2 \times 10^{-16}$
Age	$< 2.2 \times 10^{-16*}$	$4.27 \times 10^{-10} *$
Sex	$9.11 \times 10^{-09*}$	$1.3 \times 10^{-02*}$
B_glucose	$1.2 \times 10^{-03*}$	-
DM	0.198	-
DL	$4.7 \times 10^{-03*}$	-
tPA	0.6149	-
AF	$7.67 \times 10^{-11*}$	-
HTN	0.08	-
SMK	$4.098 \times 10^{-04*}$	$4.4 \times 10^{-02*}$

Δ NIHSS: difference between baseline NIHSS and NIHSS at discharge; B_NIHSS: NIHSS at baseline; B_glucose: glucose levels measured at baseline; DM: diabetes mellitus; DL: dyslipidemia; tPA: treatment with rtPA; AF: atrial fibrillation; HTN: hypertension; SMK: smoking habit.

Table III: Summary statistics for the discovery EWAS adjusted by batch

Discovery (batch correction)				
CpG	CHR	BP	COEFF (UCI.LCI)	P-value
cg00039070	7	133482303	-4.72(-6.67.-2.77)	2.96E-06
cg05979619	2	111947000	-6.21(-8.15.-4.26)	9.80E-10
cg10977834	4	77075687	4.61(6.57.2.66)	4.82E-06
cg08119231	1	220747874	4.92(6.87.2.96)	1.13E-06
cg21574204	6	28081180	4.57(6.52.2.61)	5.96E-06
cg01149259	3	46979906	4.28(6.23.2.32)	2.21E-05
cg18831371	5	176610324	5.02(6.97.3.07)	6.65E-07
cg26050512	20	58981011	4.74(6.69.2.79)	2.67E-06
cg21404878	15	72286264	4.76(6.71.2.81)	2.40E-06
cg09548897	16	291277	-4.69(-6.64.-2.74)	3.34E-06
cg14659930	3	114409454	4.43(6.38.2.48)	1.10E-05
cg00939347	10	606320	-5.09(-7.04.-3.15)	4.63E-07
cg07925823	16	68264601	4.43(6.39.2.48)	1.10E-05
cg20383948	21	45478223	5.06(7.01.3.11)	5.48E-07
cg09741713	12	62078134	4.45(6.4.2.49)	1.03E-05
cg12103149	6	30213396	4.41(6.36.2.46)	1.21E-05
cg10156941	17	1563002	4.54(6.49.2.59)	6.82E-06
cg15765398	21	44990079	-5.08(-7.03.-3.13)	4.87E-07
cg12349416	11	4184868	4.26(6.21.2.3)	2.39E-05
cg14482313	12	52233105	4.39(6.34.2.43)	1.35E-05
cg02996131	6	152637463	4.54(6.49.2.59)	6.81E-06
cg19935850	4	41751422	5.07(7.02.3.12)	5.16E-07
cg10363284	4	8005411	-5.09(-7.04.-3.14)	4.82E-07

cg18707780	15	99733585	4.29(6.25.2.34)	2.03E-05
cg20648899	6	93416423	4.4(6.35.2.44)	1.30E-05
cg07987148	20	46690251	4.5(6.46.2.55)	7.99E-06
cg07475390	1	14114109	4.79(6.74.2.84)	2.05E-06
cg00347584	11	47261474	4.43(6.38.2.48)	1.11E-05
cg13114315	21	36174169	4.25(6.21.2.3)	2.45E-05
cg14414100	9	19547532	4.5(6.45.2.54)	8.28E-06
cg18795809	4	10456907	4.51(6.46.2.56)	7.80E-06
cg03732020	11	47261417	4.42(6.37.2.46)	1.18E-05
cg24978805	7	3687557	-4.64(-6.59.-2.68)	4.33E-06
cg08526825	16	2752228	4.4(6.36.2.45)	1.26E-05
cg25354926	4	818823	-4.24(-6.2.-2.29)	2.54E-05
cg04886221	1	27343238	4.43(6.39.2.48)	1.09E-05
cg22363670	7	86643853	4.34(6.3.2.39)	1.64E-05
cg06933752	11	65122316	4.99(6.94.3.04)	7.94E-07
cg04311230	6	159693649	2.55(4.51.0.58)	0.011083
cg04349420	8	48501719	4.34(6.3.2.39)	1.63E-05
cg07797073	1	2051352	4.66(6.62.2.71)	3.81E-06
cg11491381	20	3082857	-3.91(-5.87.-1.95)	0.000102
cg18862005	2	177076135	3.8(5.75.1.84)	0.000162
cg25794823	18	63949347	4.34(6.29.2.38)	1.68E-05

Summary statistics for the EWAS analysis adjusted by batch of the 44 CpG sites with nominal association ($p\text{-value} < 10^{-6}$) in the Discovery Analysis .

CpG: CpG site ID; CHR: chromosome where the CpG site is located; BP: specific chromosomal position for the CpG site; COEFF (UCI,LCI): Effect size for the association of

the CpG-site with the Δ NIHSS, with information for the upper (UCI) and lower (LCI) 95% confidence intervals; P-value: p-value for the association of the CpG-site with the Δ NIHSS.

Table IV: Feature enrichment analysis

	EWAS analysis	Hypermethylation → Δ NIHSS > 0							Hypomethylation → Δ NIHSS > 0							Enrichment analysis	P_value
		1stExon	3'UTR	5'UTR	Body	IGR	TSS1500	TSS200	1stExon	3'UTR	5'UTR	Body	IGR	TSS1500	TSS200		
DISCOVERY	SIGN. (N = 5)	0.0%	0.0%	0.0%	60.0%	20.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	20.0%	VS ALL	0.48
	NOMINAL (N = 44)	0.0%	0.0%	0.0%	18.2%	2.3%	0.0%	0.0%	2.3%	2.3%	15.9%	15.9%	11.4%	11.4%	20.5%	VS ALL	9.23×10^{-03}
	ALL (N = 423,156)	1.7%	1.5%	3.2%	13.8%	10.1%	5.2%	3.5%	3.3%	1.9%	5.6%	19.2%	14.0%	9.2%	7.8%		

Feature enrichment analysis considering the classification of hypermethylated and hypomethylated CpG-sites associated with Δ NIHSS improvement. Different feature comparisons were done depending on the CpG-sites considered.

The features for the significant (N = 5) and significant+nominal (n = 44) CpG-sites from the Discovery analysis were compared with the features from all the CpG-sites included in the array (after QCs; N = 423,156).

EWAS analysis column: analysis considered for CpG-site selection; Hypermethylation and Hypomethylation columns: percentage of CpG-sites belonging to each feature depending on their methylation status regarding Δ NIHSS. Enrichment analysis column: CpG-site background used for comparison (all CpG-sites or all nominal CpG-sites from the Discovery). P_value: p-value for the enrichment analysis.

Δ NIHSS: difference between baseline NIHSS and NIHSS at discharge; 3'UTR: 3' untranslated region (UTR); 5'UTR: 5' untranslated region (UTR); TSS200: transcription start site (TSS) between 0 and 200 nucleotides from the TSS; TSS1500: transcription start site (TSS) between 200 and 1500 nucleotides from the TSS; IGR: intergenic region

Table V: Differentially methylated region (DMR) results

Chr	Start	End	Value	Width	P-value
7	27183133	27184521	0,015	1388	1,87E-04
20	36148699	36149455	-0,012	756	7,46E-04
1	248100276	248100614	0,027	338	1,62E-03
6	30039132	30039801	0,015	669	2,06E-03
12	75784541	75785232	0,022	691	2,21E-03
22	22221599	22221878	0,028	279	2,29E-03
6	28956226	28956731	0,013	505	3,14E-03
17	77924582	77924733	0,025	151	3,66E-03
6	32847548	32847845	0,015	297	4,01E-03
6	33280052	33280518	0,012	466	4,79E-03
5	149546331	149546471	0,021	140	6,10E-03
6	32118295	32118457	0,015	162	6,21E-03
7	1080473	1081250	-0,021	777	6,55E-03
7	56515510	56516255	-0,023	745	8,83E-03
17	7832680	7833237	0,017	557	9,70E-03
15	31515750	31516481	0,018	731	9,85E-03
1	161008462	161008826	-0,016	364	1,11E-02
19	13875014	13875329	0,017	315	1,27E-02
22	45809596	45810043	0,017	447	1,33E-02
18	61616369	61616718	0,016	349	1,42E-02
17	33759512	33760249	0,017	737	1,43E-02
7	27142535	27143287	0,015	752	1,54E-02
12	14926572	14927099	0,012	527	1,71E-02
18	14132207	14132366	0,017	159	1,93E-02

5	1594579	1594863	0,020	284	2,16E-02
20	50418952	50418959	0,014	7	2,16E-02
22	42394590	42394853	0,019	263	2,17E-02
6	29911334	29911558	0,017	224	2,20E-02
1	205818956	205819463	0,017	507	2,39E-02
15	81426347	81426820	0,016	473	2,62E-02
12	108078821	108079012	0,011	191	3,03E-02
1	230415225	230415668	0,017	443	3,15E-02
1	210000883	210001082	0,015	199	3,16E-02
11	70507825	70508410	0,018	585	3,30E-02
11	67383377	67383862	0,012	485	3,34E-02
7	27225811	27226329	0,014	518	3,34E-02
6	31734147	31734580	0,014	433	3,38E-02
5	78985484	78985592	0,012	108	3,51E-02
1	76189707	76189770	0,014	63	3,55E-02
20	61340827	61340885	0,018	58	3,69E-02
15	22833149	22833335	-0,017	186	4,04E-02
12	29302016	29302035	0,013	19	4,10E-02
10	124638756	124639110	0,017	354	4,25E-02
3	45077369	45077920	0,014	551	4,46E-02
15	63340581	63340702	0,014	121	4,53E-02
17	17109936	17110353	0,016	417	4,60E-02
2	157184816	157184978	0,018	162	4,62E-02
20	36322069	36322169	0,013	100	4,92E-02
16	30441151	30441498	0,016	347	4,92E-02
15	41952827	41953061	0,017	234	4,92E-02

Significant DMR results (p-value < 0.05) in the Discovery cohort. Chr: chromosome where the DMR belongs to; Start-End: start position and final position of the DMR; Value: Average methylation variation in the DMR; Width: number of CpG-sites included in the DMR; P-value: p-value for the association of the DMR with Δ NIHSS.

Table VI: Differentially methylated block (DMB) results

Chr	Start	End	Value	P-value
2	212609384	212618994	0,00351132	1,47E-05
2	213248884	213292541	0,00268735	4,30E-05
20	61158255	61167926	-0,00114466	2,95E-04
4	92879355	92879355	0,00164698	5,85E-04
5	97645526	97728044	-0,00103882	6,00E-04
22	22843648	22843648	-0,00160559	6,60E-04
6	62596341	62737607	-0,00120416	7,11E-04
13	67479146	67721697	0,0009132	8,03E-04
8	91227102	91411453	0,00117496	8,29E-04
9	71795999	71819910	0,00099709	2,03E-03
1	99774415	99774415	-0,00120781	2,20E-03
22	21089054	21090192,5	0,00100978	3,18E-03
22	22874756,4	22918718	0,00089951	3,50E-03
3	21426059	21447688	-0,00083278	3,61E-03
6	98744658	99007865	0,00077624	3,86E-03
11	89232216	89322779	0,00087646	3,97E-03
3	177778301	181556886	-0,00031368	4,05E-03
18	22033229	22039676,5	-0,00095528	4,09E-03
9	72158091	72347081	0,00094816	4,24E-03
20	48178928	48252667	0,00093139	4,62E-03
12	10208551	11662075	-0,00029217	4,81E-03
5	97315689	97445281	0,00091453	5,02E-03
4	96051461	96102203	0,00078167	5,07E-03
5	91988522	92582155	-0,00056859	5,52E-03
8	136439150	136803393	-0,00061732	6,10E-03

4	187037287	187637871	-0,00033556	6,29E-03
4	169418137	170328730	-0,00034564	6,65E-03
7	19408343	21788417,5	-0,0003552	6,70E-03
3	81533970	81816600	-0,00073733	6,82E-03
16	34214005,7	34265361	0,00078642	6,86E-03
14	34385840	34488059	0,00064496	6,86E-03
3	171283606	172325619	-0,0003173	7,05E-03
4	125660016	125857940	0,00071598	7,87E-03
4	119910700	120726319	-0,00040147	7,87E-03
15	25414750,6	25726500	-0,00029	8,05E-03
6	85298261	85445250	0,00062238	8,27E-03
7	81475720	81475720	0,00088104	8,30E-03
12	15501337	16512966	-0,00038982	8,35E-03
11	27453034	28985617	-0,00035097	8,38E-03
9	71669714	71691305	-0,00081309	8,53E-03
4	163661836	163907607	0,00069887	8,85E-03
6	154406753	155432132	-0,00033527	9,34E-03
6	143474125	144294847	-0,00033565	9,60E-03
2	77747049	77856944	-0,00068491	9,75E-03
8	87404176	87632363	-0,00062299	9,77E-03
4	7750448	8145692	-0,00026979	9,82E-03
4	36162678	36257279	0,00068168	9,97E-03

Top block results ($p\text{-value} < 10^{-04}$) in the Discovery cohort. Chr: chromosome where the DMR belongs to; Start-End: start position and final position of the block; Value: Average methylation variation in the block; P-value: p-value for the association of the block with Δ NIHSS.

Table VII: EWAS summary statistics in the meta-analyses for dichotomic Δ NIHSS.

CpG	CHR	BP	Gene	Feature	COEFF (UCI,LCI)	P
cg00039070	7	133167057	EXOC4	Body	-3.67(-3.66,-3.67)	4.56E-04
cg00347584	11	47283025	NR1H3	Body	3.66(3.67,3.65)	2.29E-04
cg00939347	10	652260	DIP2C	Body	-3.07(-3.07,-3.07)	2.16E-03
cg01149259	3	47021396	NBEAL2	1stExon	2.60(2.60,2.60)	4.44E-03
cg02996131	6	152958598	SYNE1	TSS1500	4.18(4.19,4.17)	3.04E-05
cg03732020	11	47282968	NR1H3	Body	4.26(4.27,4.25)	2.27E-05
cg04886221	1	27669729	SYTL1	5'UTR	2.33(2.33,2.33)	1.37E-02
cg05979619	2	112704577	MERTK	Body	-2.61(-2.61,-2.61)	1.07E-02
cg06933752	11	64889788	FAU	TSS200	2.83(2.83,2.83)	3.00E-03
cg07475390	1	14440604		IGR	2.38(2.39,2.38)	1.89E-02
cg07925823	16	68298504	SLC7A6	5'UTR	2.75(2.75,2.75)	2.65E-03
cg07987148	20	45318890	TP53RK	TSS1500	3.26(3.27,3.25)	1.02E-03
cg08119231	1	220921216	MOSC2	TSS1500	3.19(3.20,3.19)	1.29E-03
cg08526825	16	2802229	SRRM2	TSS200	2.50(2.50,2.50)	1.80E-02
cg09548897	16	341277	AXIN1	Body	-3.00(-3.00,-3.00)	5.27E-03
cg09741713	12	62471915	FAM19 A2	5'UTR	2.90(2.91,2.90)	3.82E-03
cg10156941	17	1466296	PITPNA	TSS200	5.00(5.00,5.00)	4.93E-04
cg10363284	4	8007138	ABLIM2	Body	-4.09(-4.08,-4.09)	4.60E-05
cg10977834	4	77996840	CCNI	5'UTR	4.33(4.33,4.33)	8.58E-06
cg12103149	6	30181173	TRIM26	TSS200	3.00(3.00,3.00)	1.21E-04
cg12349416	11	4206098		IGR	3.53(3.53,3.52)	4.36E-04
cg13114315	21	37546467	DOPEY2	Body	1.98(2.00,1.97)	4.55E-02

cg14414100	9	19547530	SLC24A 2	Body	2.28(2.29,2.27)	2.22E-02
cg14482313	12	52626889	KRT7	TSS200	4.25(4.26,4.24)	2.79E-05
cg14659930	3	114128301	ZBTB20	5'UTR	3.32(3.34,3.30)	8.71E-04
cg15765398	21	46409994		IGR	-3.58(-3.57,-3.59)	3.36E-04
cg18707780	15	100273790	LYSMD4	TSS200	3.50(3.50,3.50)	6.31E-05
cg18795809	4	10458531	ZNF518 B	5'UTR	2.71(2.72,2.71)	7.17E-03
cg18831371	5	176037325	GPRIN1	TSS200	3.00(3.00,3.00)	3.37E-03
cg19935850	4	41753439		IGR	3.51(3.52,3.50)	4.88E-04
cg20383948	21	46898137	COL18A 1	Body	2.75(2.75,2.75)	7.97E-03
cg20648899	6	94126141	EPHA7	Body	2.95(2.96,2.93)	3.28E-03
cg21404878	15	72578605	BRUNOL 6	3'UTR	2.08(2.08,2.07)	3.66E-02
cg21574204	6	28048958	ZNF165	5'UTR	4.00(4.00,4.00)	3.32E-05
cg22363670	7	86273169	GRM3	TSS200	1.73(1.73,1.73)	6.79E-02
cg24978805	7	3727189	SDK1	Body	-2.44(-2.43,-2.44)	1.37E-02
cg25354926	4	812611	CPLX1	Body	-1.75(-1.75,-1.75)	8.79E-02
cg26050512	20	57556066	TH1L	TSS1500	3.33(3.33,3.33)	7.55E-05

Summary statistics for EWAS results in the meta-analysis considering Δ NIHSS as a dichotomic variable (only for the 38 CpG sites with nominal association in the Discovery EWAS analysis of continuous Δ NIHSS and available in the Discovery and Replication cohorts).

CpG: CpG site ID; CHR: chromosome where the CpG site is located; BP: specific chromosomal position for the CpG site; Gene: Gene annotation (from the Illumina Manifest File); Feature: Genomic location of the CpG site. It could reside in the 5' untranslated region (UTR), between 0 and 200 nucleotides from the transcription start site (TSS), TSS200, or between 200 and 1500 nucleotides from the TSS, TSS1500. It could be also located in the body of the gene, in the 3'UTR or in an intergenic region (IGR); COEFF (UCI,LCI): Effect size for the association of the CpG-site with the dichotomic Δ NIHSS, with information for the upper (UCI) and lower (LCI) 95% confidence intervals; P: p-value for the association of the CpG-site with the dichotomic Δ NIHSS.

Table VIII: Demographic and clinical data for the subjects included in the analysis with SOMAscan

	Stroke patients	Controls
Subjects(n)	26	20
Sex		
<i>Male, n(%)</i>	10 (38.5%)	11 (55%)
<i>Female, n(%)</i>	16 (61.5%)	9 (45%)
Age in years, median (IQR)	77 (13)	64 (5)
Presence of HTN, n(%)	12 (46.15%)	20 (100%)
Smoking, n(%)	3 (11.5%)	3 (15%)
Treatment with rtPA, n (%)	26 (100%)	N/A
TOAST		
<i>CES, n(%)</i>	16 (35.9%)	N/A
<i>LAS, n(%)</i>	5 (27.8%)	N/A
<i>SVS, n(%)</i>	0 (0%)	N/A
<i>Und, n(%)</i>	3 (3.9%)	N/A
<i>NA, n(%)</i>	2 (7.7%)	N/A

Demographic and clinical features for subjects included in the proteomic analysis with SOMAscan.

IQR: interquartile range; HTN: Hypertension; CES: cardioembolic stroke; LAS: large artery stroke; SVS: small vessel stroke; und: undetermined stroke.

Table IX: eFORGE analysis

Project	Data type	Cell	Tissue	P-value	Q-value
Roadmap Epigenomics	DHS	E109 Small Intestine	Small Intestine	0.0265	1
	Chromatin mark: Enh	E071 Brain Hippocampus Middle	Brain	0.0115	1
	Chromatin mark: BivFlnk	E081 Fetal Brain Male	Brain	0.0143	1
	Chromatin mark: TxWk	E092 Fetal Stomach	Digestive	0.0158	1
	Chromatin mark: TxWk	E090 Fetal Muscle Leg	Muscle	0.0172	1
	Chromatin mark: Enh	E074 Brain Substantia Nigra	Brain	0.0177	1
	Chromatin mark: Enh	E006 H1 Derived Mesenchymal Stem Cells	ES-derived	0.0277	1
	Chromatin mark: TssAFlnk	E082 Fetal Brain Female	Brain	0.0284	1
	Chromatin mark: TxWk	E063 Adipose Nuclei	Adipose	0.0304	1
	Chromatin mark: TxWk	E061 Foreskin Melanocyte	Epithelial	0.0307	1

		Primary Cells			
		skin03			
	Chromatin mark:	E098 Pancreas	Pancreas	0.0342	1
	EnhG				
	Chromatin mark:	E112 Thymus	Thymus	0.0373	1
	TssAFlnk				
E054 Ganglion					
Eminence					
	Chromatin	derived	Neurosphere	0.0408	1
	mark: Enh	primary			
		cultured			
		neurospheres			
	Chromatin mark:	E052 Muscle	Muscle	0.0469	1
		Satellite			
	Enh	Cultured			
E059 Foreskin					
	Chromatin mark:	Melanocyte	Epithelial	0.0474	1
		Primary Cells			
	EnhG	skin01			
E006 H1					
	H3K4me1	Derived	ES Cell	0.0339	1
		Mesenchymal			
		Stem Cells			
ENCODE	DHS	hESCT0	ES Cell	0.0267	1
Blueprint	DHS	CD14+ CD16- monocyte	Venous Blood	0.0422	1
	DHS	CD14+ CD16- monocyte	Cord Blood	0.0439	1

Significant associations in the eFORGE analysis to identify tissue specific signal from the significant and nominal associations from the EWAS metaanalysis. DHS from Roadmap, ENCODE and blueprint projects was selected in the analysis.

DHS: DNase I hypersensitive sites; ES: Embryonic stem; Enh: Enhancer; BivFlnk: Flanking bivalent transcription start site (TSS)/Enh; TxWk: weakly transcribed states; EnhG: Genic Enhancer; TssAFlnk: Flanking Active TSS

4- Supplemental References

1. Ren, X.; Kuan, P.F. MethylGSA: A Bioconductor Package and Shiny App for DNA Methylation Data Length Bias Adjustment in Gene Set Testing. *Bioinforma. Oxf. Engl.* 2019, 35, 1958–1959.
2. Breeze, C.E.; Reynolds, A.P.; van Dongen, J.; Dunham, I.; Lazar, J.; Neph, S.; Vierstra, J.; Bourque, G.; Teschendorff, A.E.; Stamatoyannopoulos, J.A.; et al. EFORGE v2.0: Updated Analysis of Cell Type-Specific Signal in Epigenomic Data. *Bioinforma. Oxf. Engl.* 2019, 35, 4767–4769
3. Hannon, E.; Lunnon, K.; Schalkwyk, L.; Mill, J. Interindividual Methylomic Variation across Blood, Cortex, and Cerebellum: Implications for Epigenetic Studies of Neurological and Neuropsychiatric Phenotypes. *Epigenetics* 2015, 10, 1024–1032
4. Edgar, R.D.; Jones, M.J.; Meaney, M.J.; Turecki, G.; Kobor, M.S. BECon: A Tool for Interpreting DNA Methylation Findings from Blood in the Context of Brain. *Transl. Psychiatry* 2017, 7, e1187

Artículo 2:

Título: Altered methylation pattern in EXOC4 is associated with stroke outcome: an Epigenome-Wide Association Study.

Revista: Clinical Epigenetics.

Objetivo: Identificar variantes epigenéticas asociadas con la evolución neurológica a corto/medio plazo después del ictus.

Resumen:

Antecedentes y objetivo:

El curso neurológico tras un ictus es muy variable y está determinado por factores demográficos, clínicos y genéticos. Sin embargo, otros factores hereditarios, como la metilación del ADN, podrían desempeñar un papel en los cambios neurológicos después del ictus.

Métodos:

Realizamos un estudio de asociación de todo el epigenoma (EWAS) en tres etapas para evaluar la metilación del ADN asociada con la diferencia entre la NIHSS al inicio y al alta (Δ NIHSS) en pacientes con ictus isquémico. Los datos de metilación del ADN en las cohortes de descubrimiento ($n = 643$) y replicación ($n = 62$) se interrogaron con el 450K y el EPIC BeadChip. Las CpG nominales en la cohorte de descubrimiento (valor $p < 10^{-06}$) también se evaluaron en un metanálisis de las cohortes descubrimiento y replicación, utilizando un modelo de efectos fijos. La metilación del tipo celular específico se evaluó utilizando EpiDISH. Realizamos la integración de los datos de metilación con los niveles de expresión de 1305 proteínas plasmáticas medidos por SOMAscan y medimos la expresión de ARN con RT-PCR en un subgrupo de 13 sujetos.

Resultados:

El metanálisis reveló una asociación significativa en todo el epigenoma en EXOC4 (valor $p = 8,4 \times 10^{-08}$) y en MERTK (valor $p = 1,56 \times 10^{-07}$). Solo la metilación en EXOC4 también se asoció en la cohorte descubrimiento y de replicación ($P = 1.14 \times 10^{-06}$ y $P = 1.3 \times 10^{-02}$,

respectivamente). La metilación de EXOC4 se correlacionó negativamente con la evolución (coeficiente = -4,91) y mostró una tendencia hacia una disminución en la expresión de EXOC4 ($\rho = -0,469$, $P = 0,091$). El análisis de enriquecimiento de vías metabólicas de los resultados del metanálisis reveló asociaciones significativas relacionadas con el proceso exocítico. Setenta y nueve proteínas plasmáticas se expresaron diferencialmente en asociación con la metilación de EXOC4. El análisis de las vías metabólicas relacionadas con estas proteínas mostró un enriquecimiento en la activación de células *Natural Killer*.

Conclusiones:

La metilación del ADN de EXOC4 se asocia con una peor evolución neurológica después del ictus. El análisis vías metabólicas sugiere que la vía exocítica podría ser un posible mecanismo biológico subyacente a esta asociación.

RESEARCH

Open Access



Altered methylation pattern in *EXOC4* is associated with stroke outcome: an epigenome-wide association study

Natalia Cullell^{1,2,3†}, Carolina Soriano-Tárraga^{4,5,6†}, Cristina Gallego-Fábrega¹, Jara Cárcel-Márquez¹, Elena Muiño¹, Laia Lluçà-Carol¹, Miquel Lledós¹, Manel Esteller^{7,8,9}, Manuel Castro de Moura⁷, Joan Montaner^{10,11}, Anna Rosell¹⁰, Pilar Delgado¹⁰, Joan Martí-Fàbregas¹², Jerzy Krupinski^{2,13}, Jaume Roquer⁴, Jordi Jiménez-Conde^{4†} and Israel Fernández-Cadenas^{1*†}

Abstract

Background and purpose: The neurological course after stroke is highly variable and is determined by demographic, clinical and genetic factors. However, other heritable factors such as epigenetic DNA methylation could play a role in neurological changes after stroke.

Methods: We performed a three-stage epigenome-wide association study to evaluate DNA methylation associated with the difference between the National Institutes of Health Stroke Scale (NIHSS) at baseline and at discharge (Δ NIHSS) in ischaemic stroke patients. DNA methylation data in the Discovery ($n = 643$) and Replication ($n = 62$) Cohorts were interrogated with the 450 K and EPIC BeadChip. Nominal CpG sites from the Discovery (p value $< 10^{-06}$) were also evaluated in a meta-analysis of the Discovery and Replication cohorts, using a random-fixed effect model. Metabolic pathway enrichment was calculated with methylGSA. We integrated the methylation data with 1305 plasma protein expression levels measured by SOMAscan in 46 subjects and measured RNA expression with RT-PCR in a subgroup of 13 subjects. Specific cell-type methylation was assessed using EpiDISH.

Results: The meta-analysis revealed an epigenome-wide significant association in *EXOC4* (p value $= 8.4 \times 10^{-08}$) and in *MERTK* (p value $= 1.56 \times 10^{-07}$). Only the methylation in *EXOC4* was also associated in the Discovery and in the Replication Cohorts (p value $= 1.14 \times 10^{-06}$ and p value $= 1.3 \times 10^{-02}$, respectively). *EXOC4* methylation negatively correlated with the long-term outcome (coefficient $= -4.91$) and showed a tendency towards a decrease in *EXOC4* expression ($\rho = -0.469$, p value $= 0.091$). Pathway enrichment from the meta-analysis revealed significant associations related to the endocytosis and deubiquitination processes. Seventy-nine plasma proteins were differentially expressed in association with *EXOC4* methylation. Pathway analysis of these proteins showed an enrichment in natural killer (NK) cell activation. The cell-type methylation analysis in blood also revealed a differential methylation in NK cells.

[†]Natalia Cullell, Carolina Soriano-Tárraga, Jordi Jiménez-Conde, Israel Fernández-Cadenas have contributed equally to this article

*Correspondence: israelcadenas@yahoo.es

¹ Stroke Pharmacogenomics and Genetics, IIB-Sant Pau, Institut de Recerca de Sant Pau, Hospital Sant Pau, C/Sant Antoni M^a Claret, 167, 08025 Barcelona, Spain

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusions: DNA methylation of *EXOC4* is associated with a worse neurological course after stroke. The results indicate a potential modulation of pathways involving endocytosis and NK cells regulation.

Introduction

Stroke is a high incidence disease that represents the first cause of death and disability in adults [1, 2]. More than 70% of stroke survivors need help for their daily activity 5 years after an ischaemic stroke [3].

During the acute phase of a stroke, there are dynamic changes in the clinical symptoms that determine the evolution of the lesion and the associated deficits [4]. To measure the outcome of a stroke, two different quantifiable measures are usually considered: the neurological clinical symptoms and the functional independence of patients. There is high variability in the neurological and functional outcomes after stroke which is associated with several factors, including demographic, clinical and genetic factors [5–7]. Different scales are widely used to quantify the neurological deficit and the functional outcome. The National Institutes of Health Stroke Scale (NIHSS) considers 15 different measures of neurological worsening: consciousness, eye movement, vision, coordination, language, sensory function, upper and lower limb strength, facial muscle function, and neglect [8]. The modified Rankin scale (mRS) is used to quantify the functional outcome based on the capacity of patients to be independent in carrying out daily activities [9]. Different variables have been found to be predictors of the early (24 h post-stroke) neurological outcome: baseline NIHSS, tPA treatment, age, stroke subtype, glucose levels, and systolic blood pressure [5–7]. The early neurological outcome has been found to explain up to 30% of the long-term outcome (mRS at 3 months) [5]. The neurological evolution of stroke patients during hospitalization has also been assessed as a good predictor of 30-day and long-term mortality [10]. Reznik et al. compared the predictive value of NIHSS measured at different time-points: baseline, 24 h, and discharge. They concluded that discharge-NIHSS was the best predictor of the 3-month outcome [11]. Thus, the difference between the NIHSS at baseline and the NIHSS at discharge has recently gained importance as a valid outcome variable [12] as it covers the entire period of hospitalization.

Interestingly, genetics seems to play a role in the neurological course. Ibanez L et al. found that 8.7% of the difference between NIHSS at baseline and NIHSS at 24 h was explained by common single nucleotide polymorphisms (SNPs) [13]. However, only three different Genome-Wide Association Studies (GWAS) have been performed in the stroke outcome field [13–15]. Two

genes (*PATJ* and *LOC105372028*) have been associated with long-term functional outcome (at 3 months) [14, 15] and seven loci with the neurological course using the difference between baseline NIHSS and NIHSS at 24 h [13]. However, not all the heritability associated with the neurological course has been completely discovered [13] and other heritable factors, such as epigenetics, could be associated with the post-stroke neurological outcome. It has been demonstrated that epigenetics plays an important role in stroke risk and stroke vascular recurrence [16–21]. Previous Epigenome-Wide Association Studies (EWAS) have identified 22 CpG sites and 21 loci with altered DNA methylation associated with stroke risk [16]. Moreover, biological age calculated with DNA methylation is associated with stroke outcome and mortality [20, 21].

Our aim is to study the epigenetic risk factors and biological mechanisms associated with post-stroke neurological course using the difference in baseline NIHSS and NIHSS at discharge (Δ NIHSS) as outcome variable.

Materials and methods

Data availability

The DNA methylation data analysed in this study are available in GEO. The Discovery data from BASICMAR are available under the GEO accession number “GSE69138”. The Discovery from the GRECOS together with the replication cohort could be identified in GEO under the code “GSE203399”.

Patient selection

We included in the Discovery Cohort 738 Caucasian patients with EWAS data who had suffered a stroke and had had a blood sample taken during the first 24 h following ischaemic stroke. The Discovery consisted in patients from the Mar Hospital, who enrolled 662 patients as part of the BASICMAR register [22–24], and 76 from the GRECOS study [25]. BASICMAR is a prospective register of patients with ischaemic stroke recruited between 2009 and 2012 [22–24]. The GRECOS (Genotyping REcurrence Risk of Stroke) study is a project that enrolled 1,494 Caucasian patients with a first ischaemic stroke and population-based controls between July 2005 and May 2009 from 23 Spanish Hospitals [25]. The patients included from the BASICMAR and the GRECOS study were included in previous EWAS [16, 18, 19]. From the 738 patients, 725 had registered the main variable analysed in this study: the NIHSS at baseline and

the NIHSS at discharge and were included in the EWAS analysis.

In the Replication Cohort, we included 62 Caucasian stroke patients from the EPIGENESIS study. The EPIGENESIS study selected ischaemic stroke patients [26] with a blood sample collected during the first 6 h following onset of symptoms to study epigenetics associated with stroke outcome.

Differences in demographic and clinical variables between the Discovery and Replication cohorts were calculated. Differences between groups were tested with Kruskal–Wallis rank-sum test for nonparametric quantitative variables, while differences for qualitative variables was tested with chi-square test. Statistical significance was defined with p value < 0.05 .

All the projects included in this study have been approved by ethics committees and all the patients have signed informed consent forms. The study was conducted in accordance with Declaration of Helsinki and European guidelines: requirements of the Spanish Law 3/18 on the protection of personal data and the new European Union legislation on personal data, specifically Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016 Data Protection (GDPR).

DNA extraction and bisulphite conversion

Whole blood was obtained in EDTA tubes. DNA from the GRECOS and EPIGENESIS studies was extracted using a Gentra Puregene Blood Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. DNA from BASICMAR was extracted manually using salt precipitation in the National Bank of DNA (Carlos III Institute (ISCIII)).

Bisulphite conversion of DNA was performed before EWAS analysis using the EZ DNA Methylation-Gold™ Kit (Zymo Research, CA, USA). DNA methylation was studied with the 450 K BeadChip (Illumina) in all samples from the Discovery Cohort and EPIC BeadChip (Illumina) in all samples from the Replication Cohort.

Epigenome-wide association study (EWAS)

Methylation raw data were processed using R (<http://www.cran.r-project.org>) and Bioconductor packages (<http://www.bioconductor.org>).

We proceed with quality controls (QCs) using ChAMP package in R [27]. CpG quality controls consisted of the removal of CpG sites with a non-significant detection p value (p value > 0.05), CpG sites from sex chromosomes, CpG sites with affinity for multiple probes, no “CG” probes, probes with bead count < 3 in at least 5% of samples and SNP probes [28]. When performing QCs on samples, we removed samples with more than 1% missing CpG sites and samples with discordance

between genotypic and phenotypic sex (Additional file 1: Figure S1). After this processing, beta values representing methylation of each CpG site were normalized using the Noob function from the minfi package [29]. We also assessed the batch effect by a single value decomposition (SVD) analysis in ChAMP and by performing a multidimensional scaling (MDS) plot where it is shown the distance matrix of each sample depending on the batch (Additional file 1: Figure SII). The proportion of the different blood cell types was estimated for each patient, and beta values were corrected based on these data using the “champ.refbase” function of the ChAMP package [30]. This function has implemented the RefbaseEWAS method, which uses a methylation reference database for each of the major cell types present in blood.

Statistical analysis

We calculated the difference between baseline NIHSS and NIHSS at discharge (Δ NIHSS). First, we evaluated which demographic, cardiovascular and other stroke-related variables were associated with Δ NIHSS. We also assessed whether Δ NIHSS and other variables were associated with the long-term outcome (measured with the Rankin scale at 3 months, mRS) in our cohort using bivariate analysis. Then, we included the significant variables (p value < 0.05) in a backward stepwise regression.

The DNA methylation was considered the dependent variable and the Δ NIHSS the independent variable. We calculated the differential methylation positions using multiple linear regression (lm). First, in the Discovery analysis we considered as covariates the known variables to be associated with the dependent variable (DNA methylation): sex, age, self-reported smoking habits, and the first two principal components (PC) (basic EWAS adjustment). For the PC calculation, we used the function *princomp* in R from stats package which uses a spectral decomposition strategy to study the correlation between the methylation beta values. We selected the first two columns from the loading matrix (eigenvectors) from the output to adjust the results. To ensure that the batch effect was corrected when adjusting by the first two PC, we performed an additional analysis adjusting in the lm by the basic EWAS adjustment + batch. We used the MethylToSNP package in R to evaluate whether any of the CpG sites with a p value < 0.05 had SNP patterns and was not removed during QCs that could cause false positive results [31].

As a secondary analysis, we considered the Δ NIHSS as a dichotomic variable, classifying patients into those with improvement in the outcome (Δ NIHSS ≥ 4) or decline in the outcome (Δ NIHSS < 4), based on previous studies [32, 33]. The purpose was to assess whether nominal CpG sites associated with the continuous Δ NIHSS variable

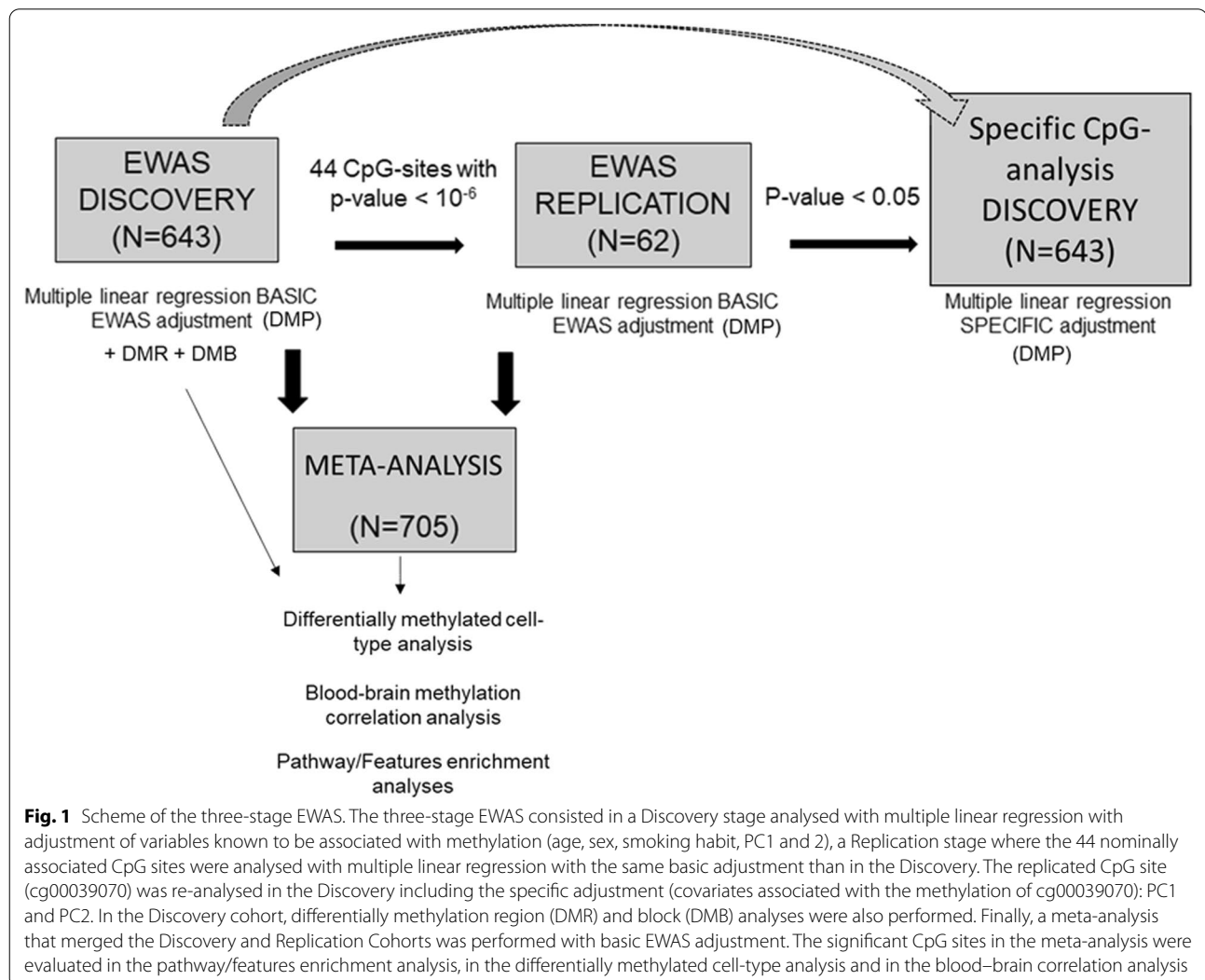
were also associated with neurological improvement or decline.

In the Discovery Analysis, all the significant, with a p value $< 2.4 \times 10^{-07}$ (based on the threshold estimated by Saffari A et al. [34, 35]), and nominal (p value $< 10^{-06}$) differentially methylated positions (DMP) were analysed in the Replication Cohort (Replication Analysis). The CpG methylation sites with p value < 0.05 in the Replication Cohort were considered replicated. The methylation from the replicated CpG sites were analysed in bivariate and backward stepwise regression analyses to identify independent clinical and demographic factors conditioning the methylation pattern of that replicated sites. Finally, the independent variables were used as covariates in a new lm analysis including only the replicated CpG sites in the Discovery Cohort (Fig. 1).

Moreover, we combined the results from the Discovery and Replication Cohorts in a meta-analysis

using METAL [36] based on a random effect model. We selected this model after reviewing the assumptions of random and fixed effect models [37]. Based on Nikolakopoulou A et al., the random effect model could be beneficial on allowing to differ on the true effects and accounting for unexplained heterogeneity [37]. The association of the nominally CpG sites from the Discovery was evaluated in the meta-analysis.

Finally, we considered all the CpG sites that followed the next criteria as validated DMPs: 1) at least a nominal association (p value $< 10^{-06}$) in the Discovery Cohort, 2) significant p value (p value < 0.05) and the same effect direction in the Replication Cohort, and 3) an epigenome-wide significant p value (p value $< 2.4 \times 10^{-07}$) in the meta-analysis.



Feature enrichment and metabolic pathway enrichment analyses

We explored for enrichment of specific features for significant ($n=5$) or nominal ($n=44$) CpG sites from the discovery in comparison with all the CpG sites included in the analysis ($n=423,156$). Features were classified into 1st exon, 3' untranslated region (UTR), 5'UTR, gene body, intergenic region (IGR), transcription start site (TSS) 1500 and TSS200. In this analysis, we classified CpG sites into hypermethylated or hypomethylated in their association with Δ NIHSS. We applied chi-square test to evaluate differences in the features between groups (significant/nominal CpG sites vs. all CpG sites).

We looked for enrichment of metabolic pathways using as input all the CpG sites and the CpG sites that were nominally associated (p value $< 10^{-06}$) with Δ NIHSS in the meta-analysis. The analysis was done with MethylGSA [38], an R package specifically designed for pathway analysis from EWAS results. We applied the three functions (methylglm, methylRRA and methylgometh) from the package (all of them designed to adjust for the number of CpGs in each gene to reduce possible bias). In the main analysis, all the CpG sites from the EWAS meta-analysis were included together with their p value. Using the first and second function, we also restricted the analysis to specific types: CpGs from promoters (TSS1500 or TSS200) or located in gene bodies. Using the second and third function, we were able to perform a secondary analysis filtering out CpG sites based on their p value, selecting the list of nominally associated CpG sites from the meta-analysis. More details for the different functions and the options that we selected are specified in the Supplementary methods. We included in the analyses the three available pathway databases in MethylGSA: Gene Ontology, KEGG and Reactome. We considered significant associations when the p value was < 0.05 and the Q value (false discovery rate (FDR) adjusted p value) was < 0.05 .

Differentially methylated regions (DMRs) and blocks (DMBs)

We evaluated whether differential methylation regions (DMRs) and blocks (DMBs) were associated with Δ NIHSS using the Bumpunther algorithm implemented in the ChAMP package [27, 30]. A DMR was defined in our analysis with the standard parameters: a segment with a minimum of seven CpG sites with a maximum length of 300 nucleotides. DMRs should be separated by a minimum of 1,000 base-pairs to be considered different DMRs. For the DMB calculation, a block was considered a large cluster generated from open sea regions. Each region located in an open sea was collapsed into a unit, calculating the mean methylation and mean position. We

used the standard recommendation of a maximum length of 250,000 nucleotides for a block and to only include blocks with a minimum of 10 regions [27, 30].

Gene expression analysis

We studied the correlation between relative *EXOC4* mRNA levels and *EXOC4* methylation. We analysed by Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) the cDNA isolated from whole blood from 13 healthy controls from the GRECOS study which also have available EWAS data.

Whole blood in EDTA tubes was obtained, and DNA was extracted using a Gentra Puregene Blood Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. We used the 7900 Real-Time PCR system (qRT-PCR) (Applied Biosystems, Foster city, CA, USA) to quantify *EXOC4* expression. We followed a standard TaqMan[®] PCR kit protocol as described previously [15]. Briefly, the *EXOC4* (Hs00253986_m1) probe was used and the results were normalized using endogenous controls: Cyclophilin A (*PPIA*, Hs99999904_m1) and Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*, Hs99999905_m1). We ran triplicates for the 13 samples, and we included an external sample as calibrator. Reactions were analysed with applied Biosystems SDS 7900 system software (Applied Biosystems, Foster city, CA, USA). We calculated the fold change of the average expression (using relative quantification (RQ) values) from *EXOC4* and endogenous controls.

Proteomic and pathway analysis

We used data from SOMAscan[®] Assay (SomaLogic) to find differentially expressed proteins associated with cg00039070 methylation. Briefly, the SOMAscan[®] Assay uses plasma samples to bring data for 1305 proteins using a short single-stranded DNA sequence (SOMAmer reagents) that binds to target proteins and allows their quantification [39].

In this analysis, we included 26 stroke patients and 20 controls for which proteomic (from SOMAscan[®]) and DNA methylation data (from EWAS) were available. For the proteomic assay, blood from patients collected in EDTA tubes was centrifuged at 3,000 g for 10 min to obtain plasma. Plasma samples were frozen at -80°C until they were analysed with SOMAscan[®]. The results were processed as described in [40]. The analysis was adjusted by case/control state.

We investigated differentially expressed proteins in association with cg00039070 methylation. Stroke patients were included from the EPIGENESIS ($N=7$) and GRECOS ($N=19$) cohorts. Controls were included from the ISSYS (Investigating Silent Stroke in hYper-tensives) cohort. It is an observational prospective study in

hypertensive participants to determine the prevalence of silent or magnetic resonance imaging (MRI)-defined brain infarcts and cognitive impairment. This cohort comprises 1000 non-demented individuals, aged 50 to 70 years, and diagnosed with essential hypertension at least one year before inclusion in the ISSYS study [41].

All the proteins associated with cg00039070 methylation (p value < 0.05) were analysed using over-representation analysis (ORA) to find enrichment of metabolic pathways using the Wilcoxon Rank-Sum Test in the Gene Ontology database (Biological Process) in WebGestalt [42, 43].

Differentially methylated cell-type (DMCT) analysis

The cellular component of the blood tissue was estimated and analysed to determine whether the differential methylation was specific to one cell type using EpiDISH [44]. The CellDMC function was used to identify differentially methylated cell types associated with the Δ NIHSS based on the proportion of B cells, CD4+ and CD8+ T cells, NK, neutrophils and monocytes. We studied the nominally associated CpG sites in the meta-analysis to find enrichment of cell-type differential methylation.

Tissue-specific signal detection

We used eFORGE (experimentally derived Functional element Overlap analysis of ReGions from EWAS) [45] to estimate tissue-specific signals from the significant and nominal results from the EWAS meta-analysis. We included analyses with all the different functional elements. A p value < 0.05 was considered a nominal association and a Q value < 0.05 according to FDR adjustment was considered statistically significant. A more complete description of this tool is provided in Supplemental Methods (Additional file 1).

Blood-brain epigenetic correlation

In order to compare the methylation from the significant findings in the meta-analysis between blood and brain, we used the Blood Brain DNA Methylation Comparison Tool [46]. We analysed the correlation of cg00039070 methylation and the four brain regions included in the tool.

Using *Blood-Brain Epigenetic Concordance* (BECon) [47], we also investigated the concordance in the cg00039070 methylation between three brain regions (Brodmann area (BA) 10, BA20 and BA7) and blood using the three metrics available in this tool.

A more detailed explanation for these two methods is included in the Supplemental Methods (Additional file 1).

Results

Discovery stage

A total of 643 patients and 423,156 CpG sites passed QCs in the Discovery analysis (Additional file 1: Figure SI). The median NIHSS at baseline and discharge was 8 and 3, respectively, and the median Δ NIHSS at discharge was positive, indicating an improvement in the neurological status of these patients (Table 1). We wanted to assess which variables were independently associated with the Δ NIHSS in our cohort. With this purpose, we studied the association of different demographic, cardiovascular and stroke-related variables, including the long-term outcome (measured with the Rankin scale at 3 months) with Δ NIHSS. Then, we explored which variables were independently associated with the mRS, to be sure that the main variable in our analysis (Δ NIHSS) was a good predictor of long-term outcome in our cohort. We performed bivariate and backward stepwise regression analyses. NIHSS at baseline (p value $< 2.2 \times 10^{-16}$), mRS (p value $< 2.2 \times 10^{-16}$), treatment with rTPA (p value $= 2.5 \times 10^{-02}$) and atrial fibrillation (p value $= 1.75 \times 10^{-02}$) were independently associated with Δ NIHSS in a multivariate analysis (Additional file 1: Table SI), whereas Δ NIHSS (p value $< 2.2 \times 10^{-16}$), baseline NIHSS (p value $< 2.2 \times 10^{-16}$), age (p value $= 4.72 \times 10^{-10}$), sex (p value $= 1.3 \times 10^{-02}$) and smoking habit (p value $= 4.4 \times 10^{-02}$) were found to be associated with the 3 months of mRS in the stepwise regression analysis (Additional file 1: Table SII). Δ NIHSS at discharge and mRS at 3 months were negatively correlated, indicating that a worsening in the neurological course measured with Δ NIHSS was associated with a worse long-term outcome measured with mRS.

Δ NIHSS was used as the independent variable to calculate differential methylation positions using multiple linear regression (lm) (Fig. 1).

In the Discovery EWAS with basic adjustment (sex, age, smoking habits and the first two PCs), we identified a total of five epigenome-wide (p value $< 2.4 \times 10^{-07}$) CpG sites and 44 nominally CpG sites (p value $< 10^{-06}$) associated with the Δ NIHSS (Fig. 2, Table 2). When the batch variable (defined by the two cohorts included in the discovery) was used as covariate, the results did not change remarkably, and all the 44 CpG sites remained significant with a p value < 0.05 (Additional file 1: Table SIII). Thus, the observed batch effect (Additional file 1: Figure SII) was corrected when adjusting by PCs. Using MethylToSNP, none of the evaluated CpG sites was predicted to have SNP patterns.

Table 1 Demographic and clinical data from the Discovery and Replication Cohorts

	Discovery		Replication	P value
	GRECOS	BASICMAR		
Subjects(n)	59	584	62	
Sex				0.429
Male, n (%)	49 (17%)	322 (55.2%)	32 (51.6%)	
Female, n (%)	10 (83%)	262 (44.8%)	30 (48.4%)	
Age in years, median (IQR)	71 (15)	77 (14)	77 (9.5)	0.22
NIHSS at baseline, median (IQR)	3 (6.5)	5 (8)	15 (11.5)	$< 2.2 \times 10^{-16}$ *
NIHSS at discharge, median (IQR)	1 (2)	3 (6)	4 (12)	2.62×10^{-06} *
Δ NIHSS, median (IQR)	1 (4)	1 (4)	5 (11)	6.6×10^{-04} *
mRS 90 days, median (IQR)	1 (3)	2 (3)	1 (3)	5.4×10^{-04} *
Presence of HTN, n (%)	37 (62.7%)	236 (40.1%)	40 (64.5%)	1.04×10^{-03} *
Smoking, n (%)	13 (22%)	168 (28.8%)	6 (9.67%)	3×10^{-07} *
Presence of AF, n (%)	3 (5%)	201 (34.4%)	25 (40.3%)	2.86×10^{-02} *
Presence of DM, n (%)	17 (28.8%)	421 (72.1%)	8 (12.9%)	9.5×10^{-11} *
Treatment with rtPA, n (%)	9 (15.3%)	91 (15.6%)	53 (85.5%)	$< 2.2 \times 10^{-16}$ *
TOAST				
CES, n (%)	0 (0%)	231 (53.9%)	29 (46.7%)	$< 2.2 \times 10^{-16}$ *
LAS, n (%)	26 (44.1%)	153 (39.6%)	17 (27.4%)	
SVS, n (%)	11 (18.6%)	194 (33.2%)	0 (0%)	
Und, n (%)	21 (35.6%)	4 (0.7%)	13 (2.1%)	
Other, n (%)	1 (1.7%)	2 (0.34%)	0 (0%)	
NA	0 (0%)	0 (0%)	3 (4.8%)	

Main clinical characteristics of patients included in the analysis from the Discovery and Replication Cohorts. P value column indicates differences for each variable between Discovery (both GRECOS and BASICMAR together) and Replication Cohorts

IQR interquartile range; Δ NIHSS = NIHSS at baseline – NIHSS at discharge; HTN: hypertension; AF atrial fibrillation; DM: diabetes mellitus; CES cardioembolic stroke; LAS large artery stroke; SVS small vessel stroke; und undetermined stroke

*Statistically significant p values for differences between cases and controls. NA: not available

We found a differential feature enrichment (p value = 9.23×10^{-03}) when comparing nominal CpG sites ($n=44$) with all the CpG sites included in the analysis ($n=423,156$) (Additional file 1: Table SIV). Specifically, we found that nominal CpG sites that were hypermethylated in association with Δ NIHSS tend to be located in the body of genes rather than in IGR. Hypomethylated nominally CpG sites were found with higher probability in the TSS200 in comparison with all the CpG sites analysed. The differential feature enrichment could suppose different effects in gene expression.

The DMR analysis revealed 50 regions associated (p value < 0.05) (Additional file 1: Table SV) with Δ NIHSS, and the DMB analysis showed a total of 323 blocks associated with Δ NIHSS (Additional file 1: Table SVI).

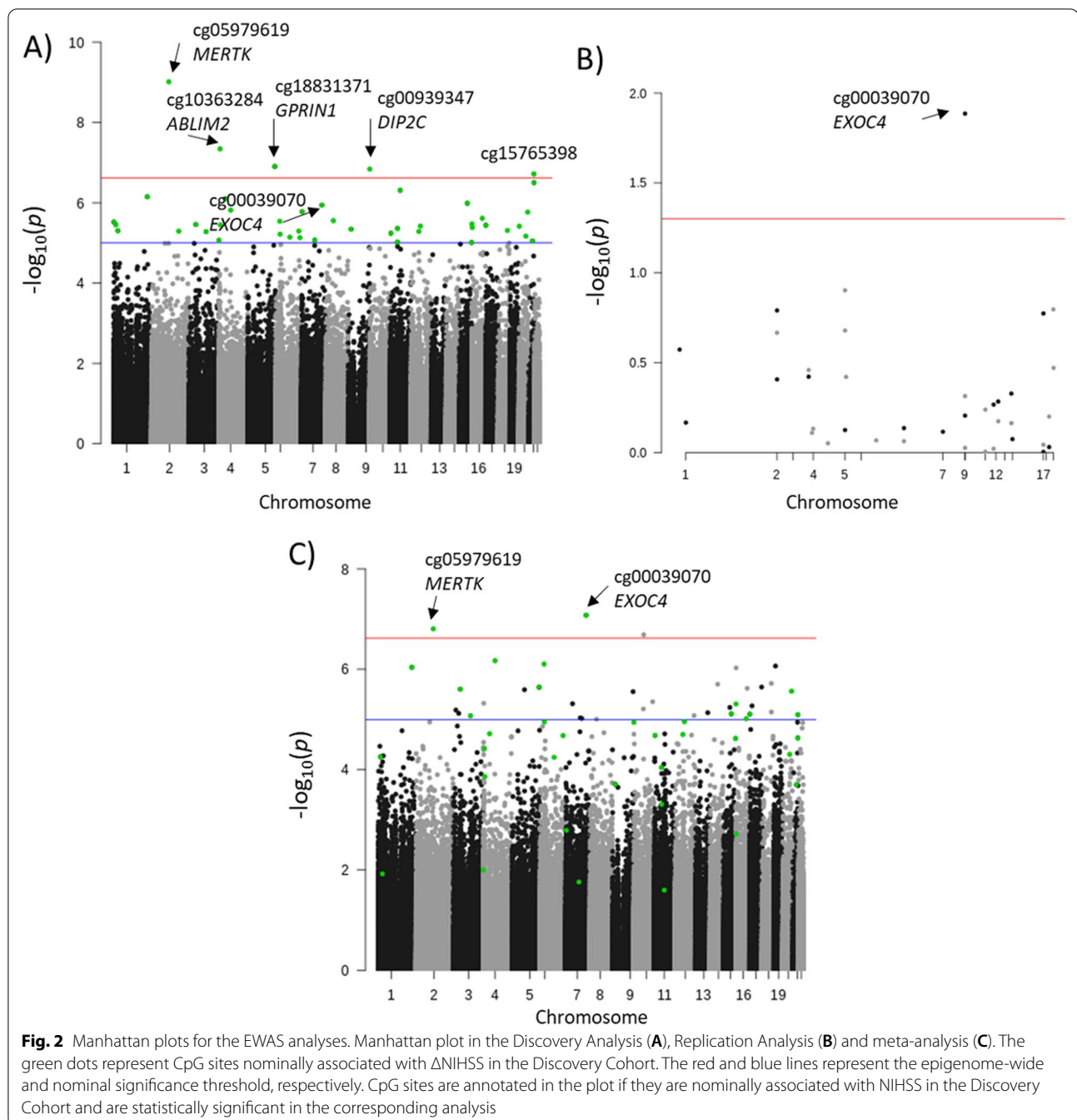
Replication stage

All the CpG sites with a nominal p value ($< 10^{-06}$) in the Discovery Analysis ($n=44$) were evaluated in the Replication Cohort. The median Δ NIHSS also had a positive value in this cohort (Table 1).

Six CpG sites could not be evaluated in the Replication Cohort because they fell to pass QCs (Table 2). From

the 38 CpG sites that could be evaluated in the Replication Cohort, one site was significant (p value < 0.05): cg00039070 (p value = 1.14×10^{-06} , coefficient = -4.91 in the Discovery and p value = 1.14×10^{-02} , coefficient = -2.58 in the Replication Cohort) (Table 2). This CpG site was located in the body of the *EXOC4* gene, a member of the exocyst complex. The effect was the same in both cohorts: higher methylation was associated with stroke worsening measured with Δ NIHSS (Fig. 3A).

We analysed the association of the cg00039070 methylation with demographic and cardiovascular risk variables in the Discovery Cohort. We found sex, batch, PC1, PC2, baseline NIHSS and diabetes mellitus associated with cg00039070 methylation in bivariate analyses. After backward stepwise regression, only PC1 and PC2 remained significant. We re-analysed by lm the association of Δ NIHSS with cg00039070 methylation using only PC1 and PC2 as covariates. Methylation of cg00039070 in *EXOC4* maintained its association with Δ NIHSS with a p value of 2.27×10^{-06} and a coefficient of -4.77 .



EWAS meta-analysis

The meta-analysis combining effect sizes from the Discovery and Replication Cohorts revealed two CpG sites with epigenome-wide significant association, including the CpG site in *EXOC4* (p value = 4.44×10^{-08} , coefficient = -5.47), and 12 CpG sites with nominal association.

(Table 2). From these, all but one presented consistent effect direction in both Discovery and Replication analyses (Table 2).

The meta-analysis considering the Δ NIHSS as a dichotomic variable (Δ NIHSS ≥ 4 vs Δ NIHSS < 4) showed all but two CpG sites from the 38 analysed in the meta-analysis to be significantly associated with Δ NIHSS (p value < 0.05). The CpG site in *EXOC4* was also associated with the dichotomic Δ NIHSS with the same effect

Table 2 EWAS summary statistics in the Discovery, Replication and meta-analyses

CpG	CHR	BP	Gene	Feature	Discovery		P	Replication		Meta-analysis		P	Direction
					COEFF (UCI,LCI)	COEFF (UCI,LCI)		COEFF (UCI,LCI)	COEFF (UCI,LCI)	COEFF (UCI,LCI)	COEFF (UCI,LCI)		
cg00039070	7	133,482,303	EXOC4	Body	− 4.91 (− 6.86, − 2.96)	1.14E-06	1.14E-06	− 2.58 (− 4.62, − 0.54)	1.30E-02	− 5.474 (− 5.4738, − 5.4742)	4.44E-08	—	—
cg05979619	2	111,947,000	MERTK	Body	− 6.21 (− 8.16, − 4.26)	9.72E-10	9.72E-10	− 1.26 (− 3.24, 0.72)	2.15E-01	− 5.250 (− 5.2495, − 5.2501)	1.54E-07	—	—
cg10977834	4	77,075,687	CCNI	5'UTR	4.85 (6.80, 2.90)	1.53E-06	1.53E-06	1.56 (3.55, − 0.43)	1.25E-01	4.875 (4.8749, 4.8747)	1.11E-06	+	+
cg08119231	1	220,747,874	MOSC2	TSS1500	5.01 (6.96, 3.06)	7.07E-07	7.07E-07	0.87 (2.83, − 1.09)	3.91E-01	4.871 (4.8717, 4.8704)	1.11E-06	+	+
cg21574204	6	28,081,180	ZNF165	5'UTR	4.72 (6.67, 2.77)	2.88E-06	2.88E-06	1.27 (3.24, − 0.70)	2.09E-01	4.847 (4.8470, 4.8470)	1.21E-06	+	+
cg01149259	3	46,979,906	NBEAL2	1stExon	4.68 (6.63, 2.73)	3.48E-06	3.48E-06	1.42 (3.41, − 0.57)	1.62E-01	4.768 (4.7684, 4.7682)	1.86E-06	+	+
cg18831371	5	176,610,324	GPRIN1	TSS200	5.35 (7.30, 3.40)	1.25E-07	1.25E-07	0.32 (2.17, − 1.53)	7.48E-01	4.752 (4.7527, 4.7522)	2.02E-06	+	+
cg26050512	20	58,981,011	TH1L	TSS1500	4.83 (6.78, 2.88)	1.71E-06	1.71E-06	− 0.09 (− 1.93, 1.75)	9.30E-01	4.563 (4.5626, 4.5624)	4.83E-06	−	−
cg21404878	15	72,286,264	BRUNOL6	3'UTR	4.94 (6.43, 2.51)	1.03E-06	1.03E-06	0.02 (2.19, − 2.15)	6.69E-01	4.540 (4.5407, 4.5393)	5.61E-06	+	+
cg09548897	16	291,277	AXIN1	Body	− 4.69 (− 6.64, − 2.74)	3.37E-06	3.37E-06	− 0.73 (− 2.67, 1.21)	4.69E-01	− 4.499 (− 4.4986, − 4.4989)	6.80E-06	—	—
cg14659930	3	114,409,454	ZBTB20	5'UTR	4.59 (6.48, 2.58)	5.24E-06	5.24E-06	0.19 (2.08, − 1.70)	3.78E-01	4.453 (4.4544, 4.4511)	8.48E-06	+	+
cg00939347	10	606,320	DIP2C	Body	− 5.32 (− 7.27, − 3.37)	1.45E-07	1.45E-07	− 0.5 (− 2.42, 1.42)	6.22E-01	− 4.450 (− 4.4501, − 4.4506)	8.50E-06	—	—
cg07925823	16	68,264,601	SLC7A6	5'UTR	4.76 (6.71, 2.81)	2.43E-06	2.43E-06	1.4 (3.39, − 0.59)	1.68E-01	4.443 (4.4430, 4.4428)	8.81E-06	+	+
cg20383948	21	45,478,223	COL18A1	Body	5.17 (7.12, 3.22)	3.15E-07	3.15E-07	0.97 (2.93, − 0.99)	3.38E-01	4.431 (4.4307, 4.4305)	9.28E-06	+	+
cg09741713	12	62,078,134	FAM19A2	5'UTR	4.66 (6.63, 2.73)	3.82E-06	3.82E-06	− 1.12 (− 3.08, 0.84)	5.20E-01	4.340 (4.3408, 4.3393)	1.42E-05	+	+
cg12103149	6	30,213,396	TRIM26	TSS200	4.56 (6.58, 2.68)	6.06E-06	6.06E-06	0.08 (1.98, − 1.82)	3.79E-01	4.331 (4.3312, 4.3311)	1.45E-05	+	+
cg10156941	17	1,563,002	PITPNA	TSS200	4.67 (6.62, 2.72)	3.63E-06	3.63E-06	0.12 (1.90, − 1.66)	9.03E-01	4.349 (4.3489, 4.3488)	1.46E-05	+	+
cg15765398	21	44,990,079		IGR	− 5.27 (6.51, 2.61)	1.92E-07	1.92E-07	0.89 (2.85, − 1.07)	1.60E-01	− 4.328 (− 4.3267, − 4.3288)	1.51E-05	+	−
cg12349416	11	4,184,868		IGR	4.58 (6.89, 2.99)	5.73E-06	5.73E-06	0.43 (2.32, − 1.46)	4.84E-01	4.321 (4.3219, 4.3207)	1.55E-05	+	+
cg14482313	12	52,233,105	KRT7	TSS200	4.60 (6.55, 2.65)	5.19E-06	5.19E-06	0.62 (2.55, − 1.31)	5.40E-01	4.304 (4.3047, 4.3037)	1.68E-05	+	+
cg02996131	6	152,637,463	SYNE1	TSS1500	4.60 (6.61, 2.71)	5.08E-06	5.08E-06	0.65 (2.58, − 1.28)	8.63E-01	4.252 (4.2532, 4.2516)	2.11E-05	+	+
cg19935850	4	41,751,422		IGR	4.99 (6.94, 3.04)	7.94E-07	7.94E-07	− 0.14 (− 1.91, 1.63)	8.86E-01	4.182 (4.1825, 4.1808)	2.90E-05	−	−
cg10363284	4	8,005,411	ABLIM2	Body	− 5.54 (− 7.49, − 3.59)	4.55E-08	4.55E-08	0.29 (2.18, − 1.60)	7.76E-01	− 4.179 (− 4.1782, − 4.1790)	2.94E-05	+	−
cg18707780	15	99,733,585	LYSMD4	TSS200	4.46 (6.41, 2.51)	9.75E-06	9.75E-06	− 0.41 (− 2.31, 1.49)	6.85E-01	4.154 (4.1539, 4.1538)	3.42E-05	−	−
cg20648899	6	93,416,423	EPHA7	Body	4.53 (6.60, 2.70)	7.17E-06	7.17E-06	− 0.2 (− 2.03, 1.63)	8.54E-01	4.007 (4.0083, 4.0057)	6.14E-05	+	+
cg07987148	20	46,690,251	TP53RK	TSS1500	4.54 (6.56, 2.66)	6.79E-06	6.79E-06	− 0.41 (− 2.28, 1.46)	9.89E-01	3.980 (3.9813, 3.9790)	6.88E-05	−	−
cg07475390	1	14,114,109		IGR	4.68 (6.53, 2.63)	3.52E-06	3.52E-06	0.7 (2.62, − 1.22)	2.67E-01	3.959 (3.9593, 3.9580)	7.55E-05	−	−
cg00347584	11	47,261,474	NR1H3	Body	4.47 (6.47, 2.57)	9.50E-06	9.50E-06	− 1.26 (− 3.24, 0.72)	9.87E-01	3.909 (3.9098, 3.9081)	9.27E-05	+	+
cg13114315	21	36,174,169	DOPEY2	Body	4.48 (6.43, 2.53)	8.92E-06	8.92E-06	0.49 (2.41, − 1.43)	6.30E-01	3.843 (3.8438, 3.8418)	1.22E-04	+	+
cg14414100	9	19,547,532	SLC24A2	Body	4.63 (6.59, 2.69)	4.55E-06	4.55E-06	− 0.56 (− 2.46, 1.34)	9.40E-01	3.772 (3.7733, 3.7717)	1.62E-04	+	+
cg18795809	4	10,456,907	ZNF518B	5'UTR	4.68 (6.63, 2.73)	3.58E-06	3.58E-06	0.34 (2.22, − 1.54)	7.36E-01	3.771 (3.7708, 3.7703)	1.64E-04	+	+
cg03732020	11	47,261,417	NR1H3	Body	4.64 (6.49, 2.59)	4.35E-06	4.35E-06	− 0.01 (− 1.29, 1.27)	5.76E-01	3.551 (3.5523, 3.5506)	3.83E-04	−	−
cg24978805	7	3,687,557	SDK1	Body	− 4.84 (− 6.79, − 2.89)	1.67E-06	1.67E-06	− 0.35 (− 2.24, 1.54)	7.30E-01	− 3.184 (− 3.1840, − 3.1845)	1.45E-03	—	—

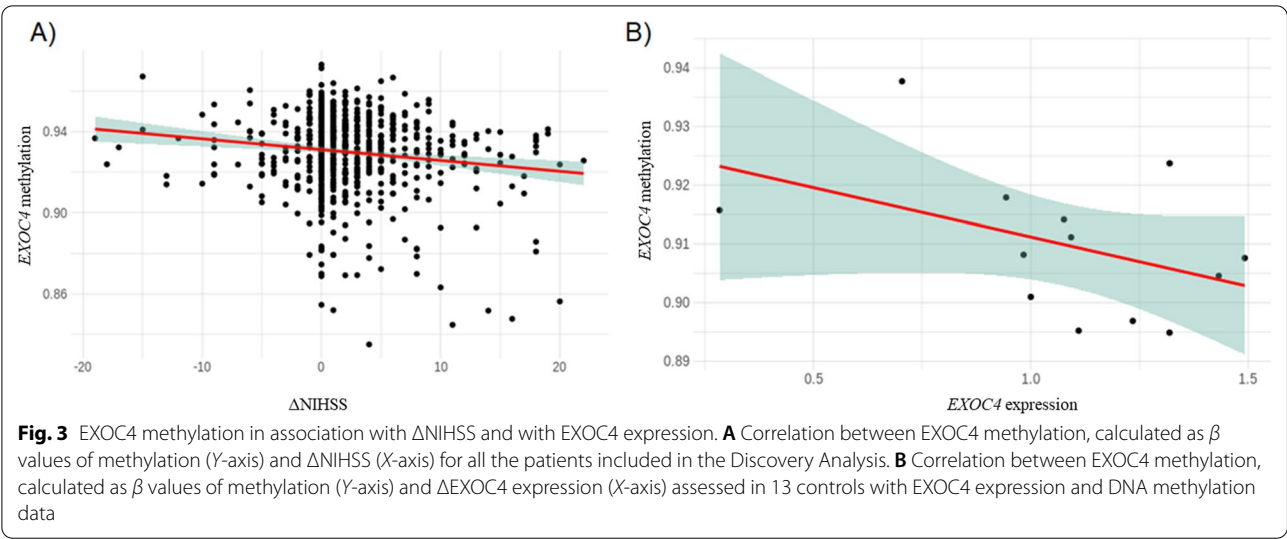
Table 2 (continued)

CpG	CHR	BP	Gene	Feature	Discovery		Replication		Meta-analysis	
					COEFF (UCI,LCI)	P	COEFF (UCI,LCI)	P	COEFF (UCI,LCI)	P
cg08526825	16	2,752,228	<i>SRRM2</i>	TSS200	4.65(− 7.22, − 3.32)	4.13E-06	1.43(3.42, − 0.56)	8.41E-01	3.051(3.0506, 3.0504)	2.28E-03
cg25354926	4	818,823	<i>CPLX1</i>	Body	− 4.49(− 6.44, − 2.54)	8.52E-06	0.95(2.91, − 1.01)	3.47E-01	− 2.784(− 2.7840, − 2.7844)	5.37E-03
cg04886221	1	27,343,238	<i>SYTL1</i>	5'UTR	4.61(6.55, 2.65)	4.97E-06	0.17(1.97, − 1.63)	6.80E-01	2.532(2.5320, 2.5319)	1.14E-02
cg22363670	7	86,643,853	<i>GRM3</i>	TSS200	4.49(6.44, 2.54)	8.55E-06	− 0.3(− 2.16, 1.56)	7.65E-01	2.409(2.4095, 2.4092)	1.59E-02
cg06933752	11	65,122,316	<i>FAU</i>	TSS200	5.08(7.03, 3.13)	4.90E-07	− 0.06(− 1.80, 1.68)	9.51E-01	2.199(2.1995, 2.1994)	2.80E-02
cg04311230	6	159,693,649	<i>SOD2</i>	TSS1500	4.52(6.68, 2.78)	7.40E-06	−	−	−	−
cg04349420	8	48,501,719		IGR	4.73(6.66, 2.76)	2.79E-06	−	−	−	−
cg07797073	1	2,051,352	<i>PRKCZ</i>	Body	4.71(− 6.61, − 2.71)	3.02E-06	−	−	−	−
cg11491381	20	3,082,857	<i>AVP</i>	Body	− 4.66(6.55, 2.65)	3.85E-06	−	−	−	−
cg18862005	2	177,076,135		IGR	4.60(6.56, 2.66)	5.10E-06	−	−	−	−
cg25794823	18	63,949,347	<i>HMSD</i>	TSS200	4.61(− 8.16, − 4.26)	4.90E-06	−	−	−	−

Summary statistics for the 44 CpG sites with nominal association (p value $< 10^{-6}$) in the Discovery Analysis. Significant CpG sites in the meta-analysis are written in bold letters in the CpG site and p value columns. Numbers in *italics* in the p value column for the meta-analysis indicate CpG sites with more significant p values in the meta-analysis compared with the Discovery Analysis

CpG: CpG site ID; CHR: chromosome where the CpG site is located; BP: specific chromosomal position for the CpG site; Gene: Gene annotation (from the Illumina Manifest File); Feature: Genomic location of the CpG site. It could reside in the 5' untranslated region (UTR), between 0 and 200 nucleotides from the transcription start site (TSS), TSS200, or between 200 and 1500 nucleotides from the TSS, TSS1500. It could be also located in the body of the gene, in the 3'UTR or in an intergenic region (IGR); COEFF (UCI,LCI): Effect size for the association of the CpG site with the Δ NIHSS, with information for the upper (UCI) and lower (LCI) 95% confidence intervals; P: p value for the association of the CpG site with the Δ NIHSS

In bold are the significant CpG sites in the meta-analysis



direction (p value = 4.56×10^{-04} , coefficient: -3.63) (Additional file 1: Table SVII).

Pathway enrichment analysis

We identified different significantly enriched pathways after FDR adjustment (Q value < 0.05) using the methylRRA function from methylGSA but only one significant pathway (p value < 0.05) and any significant pathways when using methylglm and methylgometh, respectively

(Table 3). The analyses were performed including as input the list of all the CpG sites (with the information of their p values) or the list of only the nominally associated CpG sites from the meta-analysis. As the feature enrichment analysis indicated an over-representation of CpG sites from the gene body and promoter regions, we decided also to test the pathway enrichment analysis depending on the CpG site location: gene body or gene promoter (TSS1500 and TSS200).

Table 3 Pathway enrichment analysis from meta-analysis EWAS results

Function	Database	Method	Sign CpG cut-off	Type CpGs	ID	Description	Size	P value	Q value
mglm	GO	N/A	No	Promoter	GO:0,030,100	Regulation of endocytosis	242	6.38E-04	5.36E-01
methylRRA	GO	GSA	Nominal	All	GO:0,016,358	Dendrite development	206	3.76E-04	1.62E-02
methylRRA	GO	GSA	Nominal	All	GO:0,048,588	Developmental cell growth	205	1.28E-03	2.75E-02
methylRRA	GO	GSA	No	All	R-HSA-5688426	Deubiquitination	206	1.83E-02	1.83E-02
methylRRA	Reactome	GSA	No	Promoter	R-HSA-5688426	Deubiquitination	206	4.57E-03	1.83E-02
methylRRA	Reactome	GSA	Nominal	Promoter	R-HSA-5688426	Deubiquitination	206	4.72E-03	1.89E-02
methylRRA	Reactome	ORA	Nominal	All	R-HSA-211859	Biological oxidations	203	3.08E-02	3.18E-02
methylRRA	Reactome	ORA	Nominal	All	R-HSA-72203	Processing of Capped Intron-Containing Pre-mRNA	203	3.08E-02	3.18E-02
methylRRA	Reactome	ORA	Nominal	All	R-HSA-162906	HIV infection	205	3.11E-02	3.18E-02
methylRRA	Reactome	ORA	Nominal	All	R-HSA-68882	Mitotic anaphase	209	3.17E-02	3.18E-02
methylRRA	Reactome	ORA	Nominal	All	R-HSA-2555396	Mitotic metaphase and anaphase	210	3.18E-02	3.18E-02
methylRRA	KEGG	ORA	Nominal	All	4510	Focal adhesion	200	3.03E-02	3.05E-02
methylRRA	KEGG	ORA	Nominal	All	4144	Endocytosis	201	3.05E-02	3.05E-02

Description of the significant pathways obtained from MethylGSA analysis using EWAS results from the meta-analysis

Function: Indicates which of the three MethylGSA functions was used; Database: The pathway database used in the analysis; Method: Method used (ORA or GSA) when methylRRA was selected; Sign CpG cut-off: “No” indicates that the analysis was performed including the results from all the CpG sites and “Nominal” indicates that the analysis was performed using the nominally associated CpG sites; type CpGs: “All” indicates that the analysis was performed for all the CpG types, while “promoter” indicates that in the analysis only CpG sites from promoters were analysed; ID: Identifier for the specific pathway from each database; Description: Detail of the pathway; Size: number of genes included in the gene set; and p value: enrichment p value for each gene set; Q value: FDR corrected p value

The endocytosis pathway was found to be enriched using the methylglm function (in the Gene Ontology database) and the methylRRA function (in the KEGG database). This pathway was enriched when all the CpG sites were considered but also when only the suggestive CpG sites were included. When CpG sites were filtered out based on their location, the endocytosis pathway was significant in two situations: when all the CpG sites were included independently of their location and when CpG sites from the promoter regions were selected (Table 3). The deubiquitination pathway was also found to be enriched in different situations with the methylRRA function: 1) using Gene Ontology and Reactome databases; 2) including all the CpG sites and

only the nominal CpG sites; and 3) including only CpG sites in promoters but also with all the CpG site types (Table 3). Different pathways involving cell cycle and development were also overrepresented (Table 3).

Gene expression and proteomic analyses

Only one CpG site (cg00039070 in *EXOC4*) met the pre-established criteria for being considered significant. This CpG site was located in the gene body, 10 kb downstream a predicted enhancer. Given that the effect of gene expression in gene body is variable, we wanted to assess whether this CpG site was affecting *EXOC4* expression. We found a negative correlation between cg00039070 methylation and *EXOC4* mRNA levels (Spearman correlation: -0.469) although it was not significant (p value = 0.091) (Fig. 3B).

To study the effect of *EXOC4* CpG site methylation at the proteomic level, we evaluated the proteins differentially expressed by *EXOC4* methylation using SOMAscan array data from 46 subjects (Additional file 1: Table SVIII). We found 79 differentially expressed proteins (p value < 0.05) in association with cg00039070 methylation. The most significant associations were for IFNA7 (p value = 8.52×10^{-04}) and C8A, C8B, C8G (p value = 1.67×10^{-03}) (Table 4). The pathway enrichment analysis using WebGestalt showed that the most significantly associated pathway related to significant proteins was NK cell activation (p value = 7.13×10^{-04}) (Table 5).

DMCT analysis and tissue specificity analysis

We looked for differentially methylated cell types based on the 14 CpG sites significantly or nominally associated with Δ NIHSS in the meta-analysis. We only identified significant CpG sites in NK (in *NBEAL2* and *SLC7A6* genes) and B cells (in *NBEAL2* gene). However, the

Table 4 SOMAscan results

Protein	t value	P value
IFNA7	-3.59	8.52×10^{-04}
C8A.C8B.C8G	-3.35	1.67×10^{-03}
BMP1	-3.09	3.52×10^{-03}
IGFBP4	-2.85	6.73×10^{-03}
IL17A	-2.60	1.28×10^{-02}
CGA.FSHB	-2.57	1.36×10^{-02}
TGFB1	-2.55	1.43×10^{-02}
IFNL1	-2.51	1.58×10^{-02}
LCN2	-2.49	1.66×10^{-02}
SMAD3	-2.48	1.70×10^{-02}
BMPER	-2.47	1.74×10^{-02}
SERPINE2	-2.46	1.78×10^{-02}
IL23R	-2.44	1.89×10^{-02}
IL18R1	-2.44	1.90×10^{-02}

Summary statistics for the top significant p value < 0.01) association of *EXOC4* methylation and protein levels measured by SOMAscan® Assay

Table 5 Pathway enrichment analysis for significant proteins associated with *EXOC4* methylation

Gene Set	Description	Size	Expect	Ratio	P value
GO:0,030,101	NK cell activation	85	0.21814	27.505	7.13×10^{-04}
GO:0,042,110	T cell activation	452	11.600	8.6207	1.40×10^{-03}
GO:0,050,673	Epithelial cell proliferation	372	0.95468	9.4272	3.10×10^{-03}
GO:0,002,521	Leukocyte differentiation	496	12.729	7.8560	3.30×10^{-03}
GO:0,070,661	Leukocyte proliferation	281	0.72115	11.093	4.54×10^{-03}
GO:0,001,819	Positive regulation of cytokine production	418	10.727	8.3898	8.24×10^{-03}
GO:0,002,250	Adaptive immune response	382	0.98035	8.1604	4.53×10^{-06}
GO:0,002,285	Lymphocyte activation involved in immune response	172	0.44141	13.593	4.57×10^{-06}
GO:0,018,212	Peptidyl-tyrosine modification	389	0.99831	8.0135	5.17×10^{-06}
GO:0,002,449	Lymphocyte mediated immunity	238	0.61079	9.8233	2.9×10^{-05}

Description of the significant pathways from Gene Ontology enriched among proteins significantly associated with Δ NIHSS methylation

Size: number of genes included in the gene set; expect: ratio of enrichment expected by chance in the gene set; Ratio: observed ratio for each specific gene set and p value: enrichment p value for each gene set

results were more significant in NK cells (*NBEAL2*: p value = 5.99×10^{-09} ; *SLC7A6*: p value = 8.65×10^{-08}) than in B cells (*NBEAL2*: p value = 2.59×10^{-07}) (Additional file 1: Figure SIII). There were not any CpG site in *EXOC4* differentially methylated in specific cell types.

We also looked for any tissue-specific regulatory component from the 14 CpG sites significantly and nominally associated with Δ NIHSS in the meta-analysis. Using as functional element the 15 chromatin state marks from the Roadmap project, we identified several brain tissue signals nominally enriched ($p < 0.05$), but not significant after multiple comparison adjustment in our results (Additional file 1: Table SIX).

Blood and brain correlation

We looked for correlation in the cg00039070 methylation between blood and brain. Using the “Blood Brain DNA Methylation Comparison Tool”, we did not find any correlation in cg00039070 between any of the four brain regions included in this tool (prefrontal cortex, entorhinal cortex, superior temporal gyrus and cerebellum) and blood (Additional file 1: Figure SIV).

Using BECon, we found a negative correlation (-0.48) between the cg00039070 methylation in blood and BA20 (in the temporal cortex). This correlation was classified in the highest correlation percentile (90%) (considering all the CpG sites included in the tool) (Additional file 1: Figure SV). The BA7 region (in the parietal cortex) presented an intermediate correlation (in the 50–75% percentile) (Additional file 1: Figure SV). We found a lack of correlation in the prefrontal cortex that was also observed using the “Blood Brain DNA Methylation Comparison Tool”.

Discussion

In this study, our aim was to analyse whether the neurological course could be associated with epigenetic modifications. With this objective, we studied the genome-wide DNA methylation pattern associated with Δ NIHSS at discharge by EWAS. DNA methylation is probably the most studied epigenetic variation, consisting of the addition of a methyl group to a cytosine, mainly in the context of cytosines and guanines (CpG sites). We selected Δ NIHSS at discharge as the main variable in the EWAS because it was independently associated with Rankin at 3 months. The results from our EWAS suggest that the neurological course of stroke patients measured as the difference between NIHSS at baseline and NIHSS at discharge has an impact on DNA methylation in specific CpG sites.

From the 44 candidate CpG sites identified in the Discovery Analysis, two CpG sites, located in genes bodies,

were epigenome-wide significant (p value $< 2.4 \times 10^{-07}$) in the meta-analysis of the Discovery and Replication cohorts, but only one CpG site (cg00039070) located in the body of the *EXOC4* gene accomplished all the pre-established criteria to be considered significant.

EXOC4, also known as *SEC8*, encodes for a subunit in the exocyst complex, a protein complex involved in the tethering of secretory vesicles to the plasma membrane [48]. Different functions are attributed to the exocyst complex, including but not limited to, exocytosis, cell growth cytokinesis and neuronal development [48, 49]. It is highly expressed in the brain and is enriched in axon growth cones and dendritic branches [48]. We found that the methylation pattern identified in association with stroke outcome was enriched for brain specific regulatory signals. We also investigated specifically the correlation of the methylation in cg00039070 between blood and brain using different tools. The results indicated that the effect of the *EXOC4* methylation in brain could be specific for some brain regions and its effect could be exacerbated in patients with stroke.

The pathway enrichment analysis showed that the regulation of DNA methylation in stroke outcome could be mediated by regulation of the endocytosis and the deubiquitination. The endocytosis is the process by which extracellular material is entered to the cell. This process has been showed to be interconnected with the exocytosis in the regulation of different processes such as cell polarity [50]. Gachet et al. [51] demonstrated the affectation of endocytosis in mutated *SEC8* (*EXOC4*) yeasts and a relationship with the cytokinesis process. Later, Jose et al. [52] described the exocyst complex as a key network hub which is regulating and coordinating both endocytosis and exocytosis and the balance among both processes.

The ubiquitin proteasome pathway is involved in the degradation of proteins and is key in the maintenance of the correct neuronal and synaptic function. After stroke, different pathological pathways are activated in response to the neuronal injury such as mitochondrial autophagy, oxidative stress and inflammatory response [53]. All these processes are related to the ubiquitin proteasome system. The specific role of the ubiquitin proteasome system in physiological and pathological processes after stroke is still in investigation but it has been suggested as a potential target for new drugs [53].

Higher methylation in cg00039070 from the *EXOC4* gene seems to be associated with a decrease in the expression of *EXOC4* gene, based on our gene expression results. The higher methylation in cg00039070 identified in patients with a worse stroke outcome could be mediated by the decrease in the *EXOC4* expression. Our proteomic analysis also indicated a decrease in different protein levels linked to higher cg00039070 methylation.

The results from the pathway analysis from proteins differentially expressed linked with *EXOC4* methylation, suggested that the inflammatory pathway, regulated by NK cells, could be involved in the regulation of stroke outcome by methylation. The results from the differential methylation studied by cell type also supported the involvement of this pathway. It showed some associations for the CpG sites identified in our study, especially in NK cells. NK cells are innate immune cells that infiltrate ischaemic stroke lesions in human brains [54]. The function of NK cells is regulated by activation and inhibitory receptors located in the cell surface. Thus, the endocytosis is also important for the NK cell receptors trafficking which is key to modulate the dynamic function of NK cells [55].

The Sec8 protein, encoded by *EXOC4*, has been seen to control the synaptic targeting and the insertion of glutamate receptors in the synapsis, controlling the directional movement of glutamate receptors to the post-synaptic membrane [49]. Another possible hypothesis for the association of *EXOC4* methylation and stroke outcome is the affectation of the glutamate receptors dynamism. In stroke, the release of the glutamate neurotransmitter is associated with ischaemic cell death in a process known as excitotoxicity. Briefly, the glutamate neurotransmitter is increased because of the ischaemic insult [56] and over-activates two kind of glutamate receptors: the N-methyl-D-aspartate receptor (NMDAR) and the α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPA) [56, 57]. Activation of the synaptic NMDAR leads to pro-survival signalling [58], while activation of extra-synaptic NMDAR induces a downstream neurotoxic cascade [59] that finally causes delayed neuronal death. Both NMDAR and AMPAR have been reported to be associated with the Sec8 subunit of the exocyst complex [49, 60], involved in the targeting of these receptors to the post-synaptic membrane [49, 60]. Considering that our results indicated that *EXOC4* methylation is associated with a decrease in the expression of *EXOC4* and with worse neurological course, we hypothesize that Sec8 could be regulating the trafficking of synaptic glutamate receptors related to cell survival in stroke [58].

Moreover, apart from the differential methylation identified in *EXOC4*, we found another gene, *GRM3*, nominally associated with Δ NIHSS in the discovery, that encodes glutamate metabotropic receptor 3, also involved in excitotoxicity processes. This gene was associated with memory impairment in a genetic study in Alzheimer disease patients [61].

A recently published GWAS has identified seven loci associated with stroke outcome measured through the NIHSS scale (calculating the difference between NIHSS

at baseline and at 24 h) [13]. Their functional annotation strongly suggested *GRIA1* and *ADAM23* associated with Δ NIHSS. Both genes are also involved in excitotoxicity processes. Both results support a role of excitotoxicity in processes related to stroke neurological outcomes modulated by genetic and epigenetic variations. Despite clinical trials using drugs to modulate excitotoxicity processes having failed, progress has been made in clarifying the mechanisms that explain this failure [62].

One in vivo study in *EXOC4* mutant's drosophila showed that apart from this gene being involved in glutamate receptor trafficking, it is also required for regulating synaptic microtubule formation and synaptic growth, thus suggesting that *EXOC4* methylation could be altering different processes in the synapsis [63].

The excitotoxicity, neuroinflammatory and the synaptic regulation are pathways that have been suggested to be pathological mediators of ischaemic brain damage [64] and could be potentially regulating the link between the methylation in *EXOC4* and the stroke outcome.

Limitations

The first limitation is the difference in sample size and clinical features between the Discovery and Replication Cohorts. However, we looked for which clinical variables were associated with the methylation of *EXOC4* and we did not find any. Therefore, there is no reason to believe that they would affect *EXOC4* methylation in the Replication Cohort. Despite the differences between both cohorts, we have been able to replicate the results, which reinforces the plausible implication of *EXOC4* methylation in stroke outcome.

Another limitation is the use of whole blood to study DNA methylation in association with stroke outcome. However, other epigenomic, transcriptomic and proteomic studies on stroke [16, 65] have also used blood samples as it is also a relevant tissue in stroke outcome. Additionally, the blood and brain tissues have been found to have a 0.86 correlation in global methylation [66]. For that reason, we performed the analysis and then we correlated the methylation results in brain and blood tissues. Finally, we were not able to find a significant correlation between *EXOC4* mRNA levels and *EXOC4* methylation, despite a trend being observed. Probably, the sample size of for the transcriptomic analysis was not large enough to obtain significant results.

Conclusions

We have expanded the knowledge about biological mechanisms regulating post-stroke outcome and highlighted the relevance of DNA methylation in explaining variability in functional outcome. Despite a small sample size, we had enough statistical power to obtain results

that support the hypothesis of the excitotoxicity, neuroinflammatory and synapsis regulation pathways playing a significant role in stroke and indicate that further research is needed in this field to confirm this pathway as a future therapeutic target.

Abbreviations

AMPA: α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor; DALYs: Disability-adjusted life-years; DMBs: Differential methylation blocks; DMCT: Differentially methylated cell types; DMP: Differentially methylated positions; DMRs: Differential methylation regions; EWAS: Epigenome-wide association study; GSA: Gene set enrichment analysis; GWAS: Genome-wide association studies; MRI: Magnetic resonance imaging; mRS: Modified Rankin Scale; NK: Natural killer; NMDAR: *N*-methyl-D-aspartate receptor; NIHSS: National Institutes of Health Stroke Scale; ORA: Over-representation analysis; PC: Principal components; QCs: Quality controls; TSS: Transcription start site; UTR: Untranslated region.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-022-01340-5>.

Additional file 1. Supplemental Methods.

Acknowledgements

We thank the International Stroke Genetics Consortium, the Spanish Stroke Genetics Consortium, the International Stroke Genetics Consortium, the Global Alliance for Stroke acute and long-term outcome, RETICS Network INVICTUS (RD16/0019/0002, RD16/0019/0010, RD16/0019/0011, RD16/0019/0021) and the RICORDS Stroke network.

Author contributions

I.F.C. and J.J.C. are responsible for the study design and the analytical accuracy of the manuscript. N.C. performed statistical analyses and prepared the manuscript. C.S.T. performed statistical analyses and reviewed the manuscript. C.G.F., J.C.M., L.L.L.C., M.L.L. and E.M. reviewed the manuscript and contributed to data interpretation. M.E. and M.C. performed the genotyping. J.M. and A.R. collected samples included in the study. J.M.F., J.K. and J.R. contributed to data interpretation. All the authors reviewed and approved the manuscript content.

Funding

EPIGENESIS project (Carlos III Institute/Fondo Europeo de Desarrollo Regional (FEDER)- PI17/02089, Marató TV3 and Fundació MútuaTerrassa), MAESTRO project (Carlos III Institute/FEDER—PI18/01338), iBioStroke project (Eranet-Neuron, European research grants), the EPINEXO project- PI20/00678 (Carlos III Institute/FEDER), SEDMAN Study (Boehringer Ingelheim), APHAS Study (Pfizer/Bristol Myers), Fondo Europeo de Desarrollo Regional (FEDER), 2017SGR-1427 (AGAUR), the RETICS Network INVICTUS+ and the RICORDS Stroke network. J. Cárceles-Márquez is supported by an AGAUR Contract (agència de gestió d'ajuts universitaris i de recerca; FI_DGR 2019, grant number 2019_FI_B 00853) co-financed by Fons Social Europeu (FSE). I. Fernandez is the recipient of a research contract from the Miguel Servet Program (CP12/03298) from the Carlos III Institute. M. Lledós is supported by a PFIS Contract (Contratos Predoctorales de Formación en Investigación en Salud): FI19/00309 C. Gallego-Fabrega is supported by a Sara Borrell Contract (CD20/00043) from Instituto de Salud Carlos III and Fondo Europeo de Desarrollo Regional (ISCIII-FEDER). E. Muiño is supported by a Río Hortega Contract (CM18/00198) from the Instituto de Salud Carlos III.

Availability of data and materials

The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the ethics committees of all the participant hospitals (the promoter committees were: Comité Ético de Investigación Clínica del Parc de Salut Mar, Comité Ético d'Investigació Clínica de l'Hospital Universitari Vall d'Hebron and Comité Ético de l'Hospital MútuaTerrassa). All participants provided written informed consent to participate.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Stroke Pharmacogenomics and Genetics, IIB-Sant Pau, Institut de Recerca de Sant Pau, Hospital Sant Pau, C/Sant Antoni M^a Claret, 167, 08025 Barcelona, Spain. ²Neurology, Hospital Universitari MútuaTerrassa/Fundació Docència i Recerca MútuaTerrassa, Terrassa, Spain. ³Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain. ⁴Neurology, Hospital del Mar, Neurovascular Research Group, IMIM, Universitat Autònoma de Barcelona/DCEXS-Universitat Pompeu Fabra, Barcelona, Spain. ⁵Department of Psychiatry, Washington University School of Medicine, Saint Louis, Missouri, USA. ⁶NeuroGenomics and Informatics, Washington University School of Medicine, Saint Louis, Missouri, USA. ⁷Cancer Epigenetics & Biology Program (PEBC), L'Hospitalet, Spain. ⁸Department of Physiological Sciences II, School of Medicine, Universitat de Barcelona, Barcelona, Spain. ⁹Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. ¹⁰Neurovascular Research Laboratory, Vall d'Hebron Institut de Recerca (VHIR), Barcelona, Spain. ¹¹Department of Neurology, Hospital Universitario Virgen Macarena Sevilla, Instituto de Biomedicina de Sevilla, IBIH/Hospital Universitario Virgen del Rocío/CSIC, Universidad de Sevilla, Sevilla, Spain. ¹²Neurology, Hospital de La Santa Creu i Sant Pau, Barcelona, Spain. ¹³Centre for Bioscience, School of HealthCare Science, Manchester Metropolitan University, Manchester, UK.

Received: 20 May 2022 Accepted: 13 September 2022

Published online: 30 September 2022

References

- O'Donnell MJ, Chin SL, Rangarajan S, Xavier D, Liu L, Zhang H, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. *Lancet Lond Engl*. 2016;388(10046):761–75.
- Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*. 2020;396(10258):1204–22.
- Stefan S, Bo N, Jesper P, Teresa U. Long-term survival and function after stroke. *Stroke*. 2019;50(1):53–61.
- Wouters A, Nysten C, Thijs V, Lemmens R. Prediction of outcome in patients with acute ischemic stroke based on initial severity and improvement in the first 24 h. *Front Neurol*. 2018;9:308.
- Heitsch L, Ibanez L, Carrera C, Binkley MM, Strbian D, Tatlisumak T, et al. Early neurological change after ischemic stroke is associated with 90-day outcome. *Stroke*. 2021;52(1):132–41.
- Torres-Aguila NP, Carrera C, Muiño E, Cullell N, Cárceles-Márquez J, Gallego-Fabrega C, et al. Clinical variables and genetic risk factors associated with the acute outcome of ischemic stroke: a systematic review. *J Stroke*. 2019;21(3):276–89.
- Heuschmann PU, Wiedmann S, Wellwood I, Rudd A, Carlo AD, Bejot Y, et al. Three-month stroke outcome: the European Registers of Stroke (EROS) investigators. *Neurology*. 2011;76(2):159–65.
- Taylor-Rowan M, Wilson A, Dawson J, Quinn TJ. Functional Assessment for acute stroke trials: properties, analysis, and application. *Front Neurol*. 2018;9:191. <https://doi.org/10.3389/fneur.2018.00191>.

9. Banks JL, Marotta CA. Outcomes validity and reliability of the modified rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke*. 2007;38:1091–6.
10. Kim YD, Song D, Kim EH, Lee KJ, Lee HS, Nam CM, et al. Long-term mortality according to the characteristics of early neurological deterioration in ischemic stroke patients. *Yonsei Med J*. 2014;55(3):669–75.
11. Reznik ME, Yaghi S, Jayaraman MV, McTaggart RA, Hemendinger M, Mac Grory BC, et al. Baseline NIH stroke scale is an inferior predictor of functional outcome in the era of acute stroke intervention. *Int J Stroke*. 2018;13(8):806–10.
12. Meyer L, Stracke CP, Jungi N, Wallocha M, Broocks G, Sporns PB, et al. Thrombectomy for primary distal posterior cerebral artery occlusion stroke: the TOPMOST study. *JAMA Neurol*. 2021;78(4):434–44.
13. Ibanez L, Heitsch L, Carrera C, Farias FHG, Dhar R, Budde J, et al. Multi-ancestry genetic study in 5876 patients identifies an association between excitotoxic genes and early outcomes after acute ischemic stroke. *MedRxiv*. 2020.
14. Söderholm M, Pedersen A, Lorentzen E, Stanne TM, Bevan S, Olsson M, et al. Genome-wide association meta-analysis of functional outcome after ischemic stroke. *Neurology*. 2019;92(12):e1271–83.
15. Mola-Caminal M, Carrera C, Soriano-Tárraga C, Giralte-Steinhauer E, Díaz-Navarro RM, Tur S, et al. PATJ low frequency variants are associated with worse ischemic stroke functional outcome. *Circ Res*. 2019;124(1):114–20.
16. Soriano-Tárraga C, Lazzcano U, Giralte-Steinhauer E, Avellaneda-Gómez C, Ois Á, Rodríguez-Campello A, et al. Identification of 20 novel loci associated with ischaemic stroke. *Epigenome-wide association study*. *Epigenetics*. 2020. <https://doi.org/10.1080/15592294.2020.1746507>.
17. Davis Armstrong NM, Chen W-M, Brewer MS, Williams SR, Sale MM, Worral BB, et al. Epigenome-wide analyses identify two novel associations with recurrent stroke in the vitamin intervention for stroke prevention clinical trial. *Front Genet*. 2018;9:358.
18. Gallego-Fabrega C, Carrera C, Reny J-L, Fontana P, Slowik A, Pera J, et al. PPM1A methylation is associated with vascular recurrence in aspirin-treated patients. *Stroke J Cereb Circ*. 2016;47(7):1926–9.
19. Gallego-Fabrega C, Carrera C, Reny J-L, Fontana P, Slowik A, Pera J, et al. TRAF3 epigenetic regulation is associated with vascular recurrence in patients with ischemic stroke. *Stroke*. 2016. <https://doi.org/10.1161/STROKEAHA.115.012237>.
20. Soriano-Tárraga C, Mola-Caminal M, Giralte-Steinhauer E, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, et al. Biological age is better than chronological age as predictor of 3-month outcome in ischemic stroke. *Neurology*. 2017;89(8):830–6.
21. Soriano-Tárraga C, Giralte-Steinhauer E, Mola-Caminal M, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, et al. Biological age is a predictor of mortality in ischemic stroke. *Sci Rep*. 2018;8(1):4148.
22. Roquer J, Rodríguez-Campello A, Gomis M, Jiménez-Conde J, Cuadrado-Godia E, Vivanco R, et al. Acute stroke unit care and early neurological deterioration in ischemic stroke. *J Neurol*. 2008;255(7):1012–7.
23. Soriano-Tárraga C, Giralte-Steinhauer E, Mola-Caminal M, Vivanco-Hidalgo RM, Ois A, Rodríguez-Campello A, et al. Ischemic stroke patients are biologically older than their chronological age. *Aging*. 2016;8(11):2655–65.
24. Soriano-Tárraga C, Jiménez-Conde J, Giralte-Steinhauer E, Mola-Caminal M, Vivanco-Hidalgo RM, Ois A, et al. Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum Mol Genet*. 2016;25(3):609–19.
25. Fernández-Cadenas I, Mendióroz M, Giralte D, Nafria C, Garcia E, Carrera C, et al. GRECOS project (genotyping recurrence risk of stroke): the use of genetics to predict the vascular recurrence after stroke. *Stroke*. 2017. <https://doi.org/10.1161/STROKEAHA.116.014322>.
26. Del Río-Espinola A, Fernández-Cadenas I, Giralte D, Quiroga A, Gutiérrez-Aguilló M, Quintana M, et al. A predictive clinical-genetic model of tissue plasminogen activator response in acute ischemic stroke. *Ann Neurol*. 2012;72(5):716–29.
27. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, Teschendorff AE. ChAMP: updated methylation analysis pipeline for illumina beadchips. *Bioinformatics*. 2017;33:3982–4.
28. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation beadchip probes. *Nucleic Acids Res*. 2017. <https://doi.org/10.1093/nar/gkw967>.
29. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30:1363–9.
30. Morris TJ, Butcher LM, Feber A, et al. ChAMP: 450k chip analysis methylation pipeline. *Bioinformatics*. 2014. <https://doi.org/10.1093/bioinformatics/btt684>.
31. LaBarre BA, Goncearenco A, Petrykowska HM, Jaratlerdsiri W, Bornman MSR, Hayes VM, et al. MethyToSNP: identifying SNPs in Illumina DNA methylation array data. *Epigenetics Chromatin*. 2019;12(1):79.
32. Tissue plasminogen activator for acute ischemic stroke. *N. Engl. J. Med*. 1995; 333, 1581–1588.
33. Meyer L, Stracke CP, Jungi N, Wallocha M, Broocks G, Sporns PB, Maegerlein C, Dorn F, Zimmermann H, Naziri W, et al. Thrombectomy for primary distal posterior cerebral artery occlusion stroke: the TOPMOST study. *JAMA Neurol*. 2021;78:434–44.
34. Campagna MP, Xavier A, Lechner-Scott J, Maltby V, Scott RJ, Butzkueven H, Jokubaitis VG, Lea RA. Epigenome-wide association studies: current knowledge. *Strateg Recomm Clin Epigenetics*. 2021;13:214.
35. Saffari A, Silver MJ, Zavattari P, Moi L, Columbano A, Meaburn EL, Dudbridge F. Estimation of a significance threshold for epigenome-wide association studies. *Genet Epidemiol*. 2018;42:20–33.
36. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190–1.
37. Nikolakopoulou A, Mavridis D, Salanti G. Demystifying fixed and random effects meta-analysis. *Evid Based Ment Health*. 2014;17:53–7. <https://doi.org/10.1136/eb-2014-101795>.
38. Ren X, Kuan PF. MethylGSA: a bioconductor package and shiny app for DNA methylation data length bias adjustment in gene set testing. *Bioinform Oxf Engl*. 2019;35:1958–9.
39. Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, et al. Assessment of variability in the SOMAscan assay. *Sci Rep*. 2017;7:14248.
40. Simats A, García-Berrocóso T, Ramiro L, Giralte D, Gill N, Penalba A, et al. Characterization of the rat cerebrospinal fluid proteome following acute cerebral ischemia using an aptamer-based proteomic technology. *Sci Rep*. 2018;8(1):7899.
41. Riba I, Jarca CI, Mundet X, et al. Cognitive assessment protocol design in the ISSYS (investigating silent strokes in hYpertensives: a magnetic resonance imaging study). *J Neurol Sci*. 2012;322:79–88.
42. Nguyen T-M, Shafi A, Nguyen T, Draghici S. Identifying significantly impacted pathways: a comprehensive review and assessment. *Genome Biol*. 2019;20(1):203.
43. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. 2019;47(W1):W199–205.
44. Zheng SC, Breeze CE, Beck S, Teschendorff AE. Identification of differentially methylated cell types in epigenome-wide association studies. *Nat Methods*. 2018;15(12):1059–66.
45. Breeze CE, Reynolds AP, van Dongen J, Dunham I, Lazar J, Neph S, Vierstra J, Bourque G, Teschendorff AE, Stamatoyannopoulos JA, et al. EFORGE v2.0: updated analysis of cell type-specific signal in epigenomic data. *Bioinform Oxf Engl*. 2019;35:4767–9.
46. Hannon E, Lunn K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*. 2015;10:1024–32.
47. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Transl Psychiatry*. 2017;7: e1187.
48. Mei K, Guo W. The exocyst complex. *Curr Biol CB*. 2018;28(17):R922–5.
49. Gerges NZ, Backos DS, Rupasinghe CN, Spaller MR, Esteban JA. Dual role of the exocyst in AMPA receptor targeting and insertion into the postsynaptic membrane. *EMBO J*. 2006;25(8):1623–34.
50. Martin-Urdiroz M, Deeks MJ, Horton CG, Dawe HR, Jourdain I. The exocyst complex in health and disease. *Front Cell Dev Biol*. 2016. <https://doi.org/10.3389/fcell.2016.00024>.
51. Gachet Y, Hyams JS. Endocytosis in fission yeast is spatially associated with the actin cytoskeleton during polarised cell growth and cytokinesis. *J Cell Sci*. 2005;118:4231–42.
52. Jose M, Tollis S, Nair D, Mitteau R, Velours C, Massoni-Laporte A, Royou A, Sibarita J-B, McCusker D. A quantitative imaging-based screen reveals the exocyst as a network hub connecting endocytosis and exocytosis. *Mol Biol Cell*. 2015;26:2519–34.

53. Li Y-C, Wang Y, Zou W. Exploration on the mechanism of Ubiquitin proteasome system in cerebral stroke. *Front Aging Neurosci.* 2022. <https://doi.org/10.3389/fnagi.2022.814463>.
54. Gan Y, Liu Q, Wu W, Yin J-X, Bai X-F, Shen R, Wang Y, Chen J, La Cava A, Poursine-Laurent J, et al. Ischemic neurons recruit natural killer cells that accelerate brain infarction. *Proc Natl Acad Sci USA.* 2014;111:2704–9.
55. Masilamani M, Peruzzi G, Borrego F, Coligan JE. Endocytosis and intracellular trafficking of human natural killer cell receptors. *Traffic Cph Den.* 2009;10:1735–44.
56. Achzet LM, Davison CJ, Shea M, Sturgeon I, Jackson DA. Oxidative stress underlies the ischemia/reperfusion-induced internalization and degradation of AMPA receptors. *Int J Mol Sci.* 2021. <https://doi.org/10.3390/ijms2020717>.
57. Chamorro Á, Dirnagl U, Urra X, Planas AM. Neuroprotection in acute stroke: targeting excitotoxicity, oxidative and nitrosative stress, and inflammation. *Lancet Neurol.* 2016;15(8):869–81.
58. Lai TW, Zhang S, Wang YT. Excitotoxicity and stroke: Identifying novel targets for neuroprotection. *Prog Neurobiol.* 2014;115:157–88.
59. Szydlowska K, Tymianski M. Calcium, ischemia and excitotoxicity. *Cell Calcium.* 2010;47(2):122–9.
60. Sans N, Prybylowski K, Petralia RS, Chang K, Wang Y-X, Racca C, et al. NMDA receptor trafficking through an interaction between PDZ proteins and the exocyst complex. *Nat Cell Biol.* 2003;5(6):520–30.
61. Jablensky A, Morar B, Wiltshire S, Carter K, Dragovic M, Badcock JC, et al. Polymorphisms associated with normal memory variation also affect memory impairment in schizophrenia. *Genes Brain Behav.* 2011;10(4):410–7.
62. Choi DW. Excitotoxicity: still hammering the ischemic brain in 2020. *Front Neurosci.* 2020;14:579953.
63. Liebl FLW, Chen K, Karr J, Sheng Q, Featherstone DE. Increased synaptic microtubules and altered synapse development in *Drosophila* sec8 mutants. *BMC Biol.* 2005. <https://doi.org/10.1186/1741-7007-3-27>.
64. Jayaraj RL, Azimullah S, Beiram R, Jalal FY, Rosenberg GA. Neuroinflammation: friend and foe for ischemic stroke. *J Neuroinflammation.* 2019;16:142.
65. Kamtchum-tatuene J, Jickling GC. Blood biomarkers for stroke diagnosis and management. *Neuromol Med.* 2019;21:344–68.
66. Braun PR, Han S, Hing B, Nagahama Y, Gaul LN, Heinzman JT, Grossbach AJ, Close L, Dlouhy BJ, Howard MA, et al. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Transl Psychiatry.* 2019;9:47.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



SUPPLEMENTARY MATERIAL

DNA methylation is associated with ischemic stroke risk: An Epigenome-Wide association study

Natalia Cullell, MSc^{*1,2,3}, Carolina Soriano-Tárraga, PhD^{*4,5}, Cristina Gallego-Fábrega, PhD^{1,2}, Jara Cárcel, MSc¹, Nuria P Torres-Águila, PhD^{1,6}, Jonathan González-Sánchez, MSc^{1,7}, Elena Muiño, MD, PhD¹, Manel Esteller, PhD^{8,9,10,11}, Manuel Castro de Moura, MSc⁸, Joan Montaner, MD, PhD^{6,12}, Alba Fernández-Sanlés, MSc¹³, Roberto Elosua, MD^{13,14,15}, Pilar Delgado, MD, PhD¹⁶, Joan Martí-Fàbregas, MD, PhD¹⁷, Jerzy Krupinski, MD, PhD^{2,7}, Jaume Roquer, MD⁴, Jordi Jiménez-Conde, MD, PhD^{4**}, Israel Fernández-Cadenas, PhD^{1**}

¹Stroke Pharmacogenomics and Genetics, Biomedical Research Institute Sant Pau, Sant Pau Hospital

²Neurology. Hospital Universitari MútuaTerrassa / Fundacio Docència i Recerca MútuaTerrassa

³Facultat de Medicina. Universitat de Barcelona

⁴Department of Neurology, Hospital del Mar; Neurovascular Research Group, IMIM; Universitat Autònoma de Barcelona/DCEXS-Universitat Pompeu Fabra

⁵Department of Psychiatry. NeuroGenomics and Informatics. Washington University School of Medicine

⁶Evolutionary Developmental Genomics Research group. The Scottish Oceans Institute, University of St Andrews

⁷Centre for bioscience, School of HealthCare Science, Manchester Metropolitan University

⁸Josep Carreras Leukaemia Research Institute (IJC)

⁹Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona

¹⁰Institució Catalana de Recerca i Estudis Avançats (ICREA)

¹¹Centro de Investigación Biomédica en Red Cancer (CIBERONC)

¹²Institute de Biomedicine of Seville / Hospital Universitario Virgen del Rocío/CSIC/University of Seville & Department of Neurology, Hospital Universitario Virgen Macarena

¹³Cardiovascular Epidemiology and Genetics Research Group, IMIM

¹⁴CIBER Cardiovascular diseases (CIBERCV), Instituto Carlos III

¹⁵School of Medicine, University of Vic-Central University of Catalonia

¹⁶Neurovascular Research Laboratory, Vall d'Hebron Institute of Research

¹⁷Neurology, Hospital de la Santa Creu i Sant Pau

Supplementary Data

1-Supplementary Methods

1.1-Cohorts included in the study

2-e-Table legends

3-e-Figure legends

4-e-Tables

5-e-Figures

6-e-References

1-Supplementary Methods

1.1-Cohorts included in the study:

Eighty-seven of these subjects were selected from the GRECOS study¹⁶ and 208 from the EPIGENESIS and Geno-tPA project¹⁷. The replication cohort included 618 stroke patients from the BASICMAR study¹⁸ and 243 controls from the REGICOR project¹⁹.

-GRECOS (the Genotyping RECurrence risk Of Stroke): this project enrolled 1,494 Caucasian patients with a first ischemic stroke and population-based controls between July 2005 and May 2009 from 23 Spanish Hospitals¹⁶. Thus patients were included in previous Epigenome-Wide Association Studies^{9,10}.

-EPIGENESIS (EPIGENetic Study of stroke evolution): this project selected ischemic stroke patients during the first 6 hours after onset symptoms in Vall d'Hebron Hospital to study epigenetics associated with stroke outcome.

-BASICMAR: prospective register of patients with ischemic stroke recruited from 2009 to 2012¹⁸

-REGICOR (REgistre GIroni del COR): population-based cohort recruited in the province of Girona¹⁹. This register includes a randomized representative sample of men and women of the province.

1.2- DMP analysis:

To ensure that the results were not conditioned by the sex status, we run the analysis in the Discovery cohort stratifying by sex. From the 957 significant CpG sites in the Discovery, 953 and 952 were significant ($p < 0.05$) in males and females, respectively, including the top CpGs (CpG in *MAP3K1*: $p_value = 2.06e-10$ in males; $p_value = 1.72e-17$ in females and CpG in *ZFHX3*: $p_value = 2.69e-09$ in males; $p_value = 1.76e-13$ in females).

We also checked whether the methylation of the 867 replicated CpG sites was conditioned by the sex. Our results indicated that methylation of 58 CpG sites was associated with the sex. However, the effect of the sex on the methylation of this site has been eliminated when adjusting results by sex. Moreover, the methylation of the two top CpG sites (*MAP3K* and *ZFHX3*) was not influenced by sex ($p_value = 0.64$ *MAP3K1* and $p_value = 0.63$ *ZFHX3*). Additionally, the correlation between the methylation of the significant CpG sites between males and females was significant ($p < 2.2 \times 10^{-16}$) for all of them.

From the 867 replicated CpG sites, only the methylation profile of 73 CpGs is associated with the age ($p_value < 0.05$) when we analyzed cases and controls together. However, this effect is expected to be corrected when we analyze due to the adjustment of the EWAS by age in cases vs. controls adjusting by age. The methylation of *MAP3K1* nor *ZFHX3* were associated with age ($p_value = 0.95$ and for *ZFHX3* and $p_value = 0.20$ for *MAP3K1* CpG sites).

2- e-Table legends

Table e-1: Significant EWAS results in the discovery.

Summary statistics for all the 957 statistically significant CpG-sites in the discovery cohort.

NA indicates non available results in the replication cohort. *Indicates new CpG-sites not previously associated with stroke

CHR: chromosome; BP: base-pair position; feature: location of the CpG site (1st exon, intergenic region (IGR), body of the gene, promotor region (TSS200 and TSS1500), 5'UTR and 3'UTR); cgi: CpG-site location; logFC (CI): Effect size indicated by logarithm of the fold change with the upper (97.5%) and lower (2.5%) confidence interval ; P: p-value

Table e-2: DMRs in the discovery cohort.

Summary-statistics for the Bonferroni-corrected DMR results in the discovery cohort.

Columns from left to right: CHR: chromosome; Start: Initial base-pair position for a specific DMR; End: Final base-pair position for a specific DMR; no.cpg: number of CpG-sited included in the DMR

Table e-3: Significant EWAS results by stroke subtypes.

Summary statistics of replicated CpG-sites stratified by atherothrombotic (LAS), cardioembolic (CES), small vessel stroke (SVS) and undetermined (UND) stroke subtypes. logFC (CI): Effect size indicated by logarithm of the fold change with the upper (97.5%) and

lower (2.5%) confidence interval ; P: p-value

Table e-4: Significant genes associated with stroke in GWAS analyzed by EWAS

Genes analyzed in our EWAS that has been previously associated with stroke or stroke subtypes (specific subtype stated in the column “subtype”).

Table e-5: Significant EWAS associations in genes associated with stroke in GWAS

Significant CpG sites associates with stroke risk in EWAS from the candidate analysis where CpG sites located in genes previously associated with stroke in GWAS were selected.

CHR: chromosome; BP: base-pair; logFC (CI): Effect size indicated by logarithm of the fold change with the upper (97.5%) and lower (2.5%) confidence interval; P: p-value; Beta: Beta value as effect size measure

*Significant CpG sites aftes Bonferroni adjustment.

Table e-6: Significant EWAS results by stroke subtypes in genes associated with stroke in GWAS.

EWAS results for genes significantly associated with stroke risk in the MEGASTROKE + UK Biobank study.

Table e-7: Summary statistics for significant mQTLs in *MAP3K1* and *ZFHX3*

mQTLs with P-value < 10e-7 for cg10115490 (*MAP3K1*) and cg07786668 (*ZFHX3*)

CHR: chromosome; BP: base-pair; MAF: minor allele frequency; P: p-value; Beta: Beta value as effect size measure

3- e-Figure legends

Figure e-1: Power estimation based on pwrEWAS

Mean power curve for sample sizes between 260 and 300 and a $\Delta\beta$ of 0.05.

Figure e-2: Manhattan plot for results in the discovery cohort

Manhattan plot: red line for statistical significance threshold and blue line for nominal significance.

Figure e-3: CpG-sites features and location

Feature pattern for the replicated DMP in IS compared to controls (Sign) vs all CpG-sites methylation included in the analysis (ALL). A) Percentage of probes from each location in Sign and ALL. B) Percentage of CpG-sites from each feature for Sign and ALL.

Figure e-4: MR results for reverse causality

MR results for A) IS and stroke risk causing *MAP3K1* differential methylation and B) *ZFHX3* differential methylation. Outcome indicates the different GWAs for stroke evaluated: all stroke (AS), all ischemic stroke (AIS), cardioembolic stroke (CES), large-artery stroke (LAS) and small vessel stroke (SVS). The method indicates the MR test used: MR-Egger, weighted-median (WM) and inverse-variance weighted (IVW).

4- e-Tables

Table e-1: Significant EWAS results in the discovery.

CpG	CHR	BP	GENE	feature	cgi	DISCOVERY		REPLICATION	
						logFC(CI)	P	logFC(CI)	P
cg07786668	16	73092391	ZFHX3	5'UTR	island	0.058(0.069-0.048)	2,35E-22	0.017(0.023-0.01)	3,24E-07
*cg10115490	5	56111596	MAP3K1	1stExon	island	0.035(0.041-0.029)	3,22E-22	0.028(0.032-0.023)	9,89E-32
*cg15591678	10	64134155	ZNF365	5'UTR	island	0.036(0.042-0.029)	3,76E-21	0.022(0.026-0.018)	1,12E-21
*cg13432391	19	10828930	DNM2	1stExon	island	0.028(0.034-0.023)	1,74E-19	0.019(0.023-0.015)	4,69E-21
*cg09000178	16	67063319	CBFB	1stExon	island	0.055(0.066-0.044)	2,58E-19	0.023(0.029-0.016)	1,38E-10
*cg10849092	15	40574697		IGR	island	0.029(0.035-0.023)	1,66E-18	0.016(0.019-0.012)	1,49E-18
cg23681311	22	22221878	MAPK1	1stExon	island	0.045(0.054-0.036)	2,21E-18	0.024(0.03-0.019)	1,02E-18
*cg07667295	14	64320118	SYNE2	5'UTR	island	0.046(0.056-0.037)	3,43E-18	0.049(0.057-0.041)	1,10E-30
*cg09238598	5	14871908	ANKH	TSS200	shore	0.057(0.069-0.045)	7,30E-18	0.031(0.038-0.025)	2,73E-19
*cg02516155	17	4046031	CYB5D2	TSS1500	island	0.029(0.035-0.023)	7,44E-18	0.023(0.028-0.018)	1,51E-20
*cg22512322	3	64009096	PSMD6	5'UTR	island	0.038(0.046-0.03)	7,73E-18	0.005(0.009-0.002)	0,00475
*cg06642177	6	134496341	SGK1	Body	island	0.07(0.085-0.055)	1,17E-17	0.013(0.025-0.002)	0,0232
*cg17662034	8	74207518	RDH10	1stExon	island	0.063(0.076-0.049)	1,99E-17	0.029(0.038-0.021)	1,79E-11
*cg14119263	10	5708327	ASB13	Body	island	0.048(0.058-0.038)	2,00E-17	0.027(0.032-0.022)	1,53E-24
*cg14864852	22	39102110	GTPBP1	1stExon	island	0.028(0.034-0.022)	2,99E-17	0.007(0.01-0.004)	1,94E-05
cg25869317	15	101792241	CHSY1	TSS200	island	0.045(0.054-0.035)	5,03E-17	0.025(0.031-0.019)	1,72E-16
*cg08867893	10	64134160	ZNF365	5'UTR	island	0.031(0.038-0.024)	5,68E-17	0.011(0.014-0.008)	4,85E-14
*cg14078059	12	65174660		IGR	island	0.077(0.094-0.06)	1,21E-16	0.034(0.042-0.026)	3,79E-16
*cg10206657	5	56111587	MAP3K1	1stExon	island	0.056(0.069-0.044)	4,11E-16	0.023(0.028-0.018)	1,37E-16
*cg16588137	4	186064512	SLC25A4	1stExon	island	0.024(0.029-0.018)	6,42E-16	0.011(0.014-0.008)	4,63E-13
*cg01668174	2	106810661	UXS1	1stExon	island	0.025(0.031-0.02)	7,56E-16	0.014(0.019-0.01)	1,69E-09
*cg08327423	16	84150746	MBTPS1	TSS1500	island	0.018(0.023-0.014)	7,75E-16	0.011(0.014-0.008)	4,22E-12
*cg18829263	5	129240272	CHSY3	TSS1500	island	0.034(0.042-0.026)	7,91E-16	0.033(0.038-0.029)	1,83E-42
*cg26257177	19	10828928	DNM2	1stExon	island	0.01(0.012-0.008)	1,29E-15	0.006(0.007-0.005)	4,44E-21
*cg09643587	3	107809710	CD47	1stExon	island	0.046(0.057-0.036)	1,68E-15	0.046(0.052-0.039)	1,06E-40
cg04759220	5	78532560	JMY	1stExon	island	0.028(0.035-0.022)	1,69E-15	0.011(0.016-0.007)	2,78E-06
*cg26587051	9	123476683	MEGF9	TSS200	island	0.027(0.034-0.021)	1,94E-15	0.018(0.022-0.013)	5,96E-14
*cg03036064	1	165797085	UCK2	1stExon	island	0.017(0.02-0.013)	2,01E-15	0.01(0.012-0.008)	5,67E-15
*cg17664568	10	98346460	TM9SF3	Body	island	0.022(0.027-0.017)	2,08E-15	0.021(0.025-0.017)	1,63E-22
*cg04794887	4	1858231	LETM1	TSS1500	island	0.036(0.044-0.028)	2,39E-15	0.022(0.027-0.017)	2,48E-16
cg04192862	10	28966472	BAMBI	1stExon	island	0.037(0.046-0.029)	2,85E-15	0.023(0.028-0.017)	2,84E-17
*cg18751306	15	81294292	MESDC1	1stExon	island	0.029(0.035-0.022)	3,00E-15	0.015(0.019-0.01)	6,12E-09
*cg08377314	19	56186477	EPN1	TSS200	island	0.03(0.037-0.023)	3,96E-15	0.012(0.017-0.008)	7,70E-08
*cg03852091	2	242577219	THAP4	TSS1500	island	0.02(0.025-0.016)	5,18E-15	0.014(0.017-0.01)	2,67E-15
*cg05668205	19	16738998	MED26	1stExon	island	0.02(0.024-0.015)	5,63E-15	0.016(0.019-0.012)	6,71E-17
*cg03536846	8	120886254	DEPDC6	Body	island	0.02(0.024-0.015)	6,40E-15	0.011(0.014-0.008)	3,80E-14
*cg05669421	9	111881984	C9orf5	1stExon	island	0.032(0.039-0.024)	7,04E-15	NA	NA
*cg11299854	5	132083184	CCNI2	1stExon	island	0.04(0.05-0.031)	7,16E-15	0.022(0.027-0.017)	4,41E-16
*cg05135536	9	128509002	PBX3	TSS1500	island	0.026(0.032-0.02)	7,27E-15	0.012(0.015-0.008)	8,19E-11
*cg00610228	6	35695934	FKBP5	5'UTR	island	0.023(0.029-0.018)	1,19E-14	0.016(0.02-0.013)	7,93E-18
*cg03732936	1	224371029	DEGS1	5'UTR	island	0.023(0.029-0.018)	1,24E-14	0.022(0.026-0.019)	1,01E-33

*cg27414860	16	4852427	ROGDI	Body	island	0.024(0.03-0.018)	1,58E-14	0.013(0.017-0.01)	3,93E-12
*cg16041611	6	43139680	SRF	1stExon	island	0.025(0.031-0.019)	2,05E-14	0.005(0.009-0)	0,037
*cg10122474	14	88459471	GALC	1stExon	island	0.014(0.017-0.011)	2,16E-14	0.001(0.004--0.001)	0,248
*cg18766912	15	25683909	UBE3A	1stExon	island	0.037(0.046-0.028)	2,34E-14	0.025(0.031-0.018)	7,10E-13
*cg14276025	2	85198562	KCMF1	5'UTR	island	0.016(0.02-0.012)	2,43E-14	0.009(0.013-0.006)	1,49E-07
*cg04084597	11	75479652	DGAT2	TSS200	island	0.033(0.042-0.025)	3,16E-14	0.016(0.021-0.01)	2,44E-08
*cg06793581	16	46865077	C16orf87	TSS200	island	0.027(0.033-0.02)	3,87E-14	0.026(0.03-0.021)	3,39E-31
*cg04971341	1	52607910	ZFYVE9	TSS200	island	0.021(0.026-0.016)	4,12E-14	0.025(0.029-0.021)	1,66E-37
*cg27467996	5	179921670	CNOT6	5'UTR	island	0.022(0.028-0.017)	5,01E-14	0.011(0.015-0.008)	9,06E-10
*cg01281140	7	150065678	REPIN1	TSS200	island	0.023(0.028-0.017)	5,08E-14	0.039(0.043-0.034)	1,12E-60
*cg06825833	13	111566802	ANKRD10	Body	island	0.025(0.032-0.019)	6,47E-14	0.024(0.028-0.02)	2,34E-26
*cg19539824	9	139010884	C9orf69	TSS200	island	0.028(0.035-0.021)	6,85E-14	0.022(0.027-0.017)	7,46E-19
*cg02304070	12	44229711	TMEM117	TSS1500	island	0.015(0.018-0.011)	7,72E-14	0.01(0.013-0.008)	5,22E-14
*cg09977800	17	80477464	FO XK2	TSS200	island	0.028(0.034-0.021)	7,76E-14	0.041(0.046-0.035)	1,52E-44
*cg18455653	20	17662865	RRBP1	1stExon	island	0.048(0.059-0.036)	9,12E-14	0.045(0.052-0.037)	2,94E-28
*cg24906202	15	73345212	NEO1	Body	island	0.03(0.037-0.022)	1,22E-13	0.016(0.021-0.01)	9,81E-09
*cg11424525	21	46494378	ADARB1	TSS200	island	0.023(0.029-0.018)	1,25E-13	0.033(0.038-0.029)	8,19E-45
*cg11637191	1	231557525	EGLN1	1stExon	island	0.018(0.022-0.013)	1,40E-13	0.01(0.013-0.007)	1,67E-11
*cg09572029	13	28194706	POLR1D	TSS1500	island	0.025(0.031-0.019)	1,44E-13	0.029(0.033-0.025)	1,29E-38
*cg23640903	6	13487662	GFOD1	5'UTR	island	0.032(0.039-0.024)	1,61E-13	0.054(0.06-0.049)	3,51E-72
*cg12885832	12	111843885	SH2B3	1stExon	island	0.034(0.043-0.026)	1,82E-13	0.002(0.007--0.003)	0,513
*cg18830052	20	3801106	C20orf29	TSS200	island	0.031(0.038-0.023)	2,01E-13	0.019(0.023-0.015)	2,35E-18
*cg06802374	4	148653449	ARHGAP10	TSS200	island	0.016(0.02-0.012)	2,10E-13	0.012(0.016-0.009)	2,25E-13
*cg11828669	14	64971476	ZBTB1	5'UTR	island	0.02(0.024-0.015)	2,19E-13	0.017(0.02-0.014)	1,27E-25
*cg01725318	4	141348737	CLGN	1stExon	island	0.023(0.029-0.017)	2,44E-13	0.015(0.019-0.011)	4,85E-14
*cg16063115	14	31676811	HECTD1	TSS200	island	0.017(0.022-0.013)	3,06E-13	0.011(0.014-0.008)	2,29E-13
*cg11289281	16	4897921	GLYR1	TSS1500	island	0.023(0.029-0.017)	3,62E-13	0.015(0.019-0.011)	3,27E-13
*cg02198017	14	103801143	EIF5	TSS200	island	0.019(0.024-0.014)	3,84E-13	0.016(0.02-0.013)	3,92E-19
*cg15772223	3	159943466	LOC401097	1stExon	island	0.025(0.031-0.018)	4,09E-13	0.017(0.021-0.013)	4,77E-15
*cg14898822	15	83478541	WHAMM	1stExon	island	0.017(0.021-0.013)	4,51E-13	NA	NA
*cg24794206	6	155316492		IGR	island	0.019(0.024-0.015)	4,56E-13	0.011(0.014-0.007)	5,76E-10
*cg13789548	4	122872838	TRPC3	1stExon	island	0.018(0.023-0.014)	4,64E-13	0.019(0.022-0.015)	2,55E-23
*cg16887862	1	1243669	PUSL1	TSS1500	island	0.029(0.036-0.021)	5,13E-13	0.015(0.019-0.011)	1,66E-13
cg07691609	19	662740	RNF126	Body	island	0.02(0.025-0.015)	5,29E-13	0.012(0.016-0.009)	2,74E-14
*cg08308556	6	157342734	ARID1B	Body	island	0.025(0.031-0.018)	5,86E-13	0.022(0.025-0.018)	3,13E-28
*cg21538902	10	64565003	ADO	1stExon	island	0.025(0.032-0.019)	5,89E-13	0.034(0.039-0.03)	4,52E-43
*cg07331616	11	63933573	MACROD1	1stExon	island	0.024(0.03-0.018)	6,33E-13	0.012(0.015-0.008)	1,40E-10
*cg12639324	10	92617735	HTR7	TSS200	island	0.032(0.04-0.024)	6,60E-13	NA	NA
*cg23706268	13	51484060	RNASEH2B	5'UTR	island	0.011(0.014-0.008)	6,65E-13	0.007(0.009-0.005)	6,06E-11
*cg24511898	2	242576823	ATG4B	TSS1500	island	0.017(0.022-0.013)	7,42E-13	0.016(0.021-0.01)	2,02E-08
*cg04887405	16	84733672	USP10	5'UTR	island	0.02(0.025-0.015)	8,16E-13	0.011(0.014-0.008)	5,17E-11
*cg08195842	13	21278118	IL17D	5'UTR	island	0.02(0.025-0.015)	8,20E-13	0.025(0.028-0.021)	9,79E-35
*cg13470032	9	214612	C9orf66	1stExon	island	0.017(0.021-0.012)	9,05E-13	0.011(0.014-0.009)	3,45E-15
*cg08502703	21	45138838	PDXK	TSS200	island	0.017(0.021-0.012)	9,23E-13	0.011(0.014-0.009)	8,67E-16
*cg03951424	1	61549111	NFIA	Body	island	0.014(0.018-0.011)	9,25E-13	0.012(0.014-0.01)	4,17E-21
*cg05001598	9	114246158	KIAA0368	Body	island	0.02(0.026-0.015)	1,12E-12	0.014(0.017-0.011)	5,35E-22

*cg01628954	18	60190696	ZCCHC2	1stExon	island	0.024(0.03-0.017)	1,18E-12	0.028(0.032-0.023)	1,48E-35
*cg00874357	12	62997129	MIRLET7I	TSS1500	island	0.015(0.019-0.011)	1,23E-12	0.013(0.016-0.009)	4,70E-12
*cg06014792	17	7382760	ZBTB4	5'UTR	shore	0.018(0.023-0.013)	1,29E-12	0.011(0.014-0.008)	4,06E-14
*cg19244874	3	12598623	MKRN2	1stExon	island	0.022(0.027-0.016)	1,41E-12	0.021(0.025-0.017)	4,92E-21
*cg14204791	21	44394797	PKNOX1	5'UTR	island	0.024(0.03-0.017)	2,13E-12	0.014(0.017-0.01)	8,11E-13
*cg04703844	8	74207197	RPL7	TSS1500	island	0.021(0.027-0.016)	2,16E-12	0.024(0.028-0.02)	1,51E-31
*cg02551232	2	99061666	INPP4A	5'UTR	island	0.022(0.028-0.016)	2,52E-12	0.016(0.02-0.012)	2,47E-15
*cg01266790	14	105781312	PACS2	1stExon	island	0.017(0.022-0.013)	2,72E-12	0.007(0.01-0.004)	2,17E-07
*cg23639857	17	79849429	ANAPC11	TSS200	island	0.014(0.018-0.01)	2,94E-12	0.006(0.008-0.004)	2,35E-06
*cg24915717	7	97881573	TECPR1	TSS200	island	0.023(0.029-0.017)	3,19E-12	0.026(0.029-0.022) -0.011(-0.014-- 0.007)	3,21E-41 1,35E-08
*cg25411849	6	89855994	PM20D2	1stExon	island	0.02(0.025-0.014)	3,40E-12	0.014(0.017-0.011)	6,30E-17
*cg23684603	4	185747386	ACSL1	TSS200	island	0.019(0.024-0.014)	4,49E-12	0.027(0.03-0.023)	9,06E-45
*cg11044724	2	7005733	CMPK2	1stExon	island	0.018(0.023-0.013)	4,71E-12	0.016(0.02-0.012)	6,11E-12
*cg20116574	20	44718168	NCOA5	5'UTR	island	0.025(0.031-0.018)	4,78E-12	0.004(0.005-0.003)	2,03E-12
*cg12661145	10	104404054	TRIM8	TSS200	island	0.008(0.01-0.006)	4,88E-12	0.006(0.008-0.004)	5,59E-10
*cg08299976	17	1552302	RILP	Body	island	0.013(0.017-0.01)	4,90E-12	0.017(0.02-0.013)	6,85E-20
*cg01442104	4	159690352	FNIP2	5'UTR	island	0.022(0.028-0.016)	5,05E-12	0.013(0.017-0.01)	1,17E-12
*cg18764107	1	225615669	LBR	1stExon	island	0.022(0.028-0.016)	5,18E-12	0.018(0.021-0.015)	5,06E-28
*cg04396495	5	60626910	ZSWIM6	TSS1500	island	0.023(0.029-0.017)	5,32E-12	0.013(0.017-0.008)	1,25E-07
*cg16344026	5	149109979	PPARGC1B	1stExon	island	0.023(0.029-0.016)	5,49E-12	0.016(0.019-0.013)	2,14E-22
*cg25506900	17	73512855	CASKIN2	TSS1500	island	0.02(0.025-0.015)	5,70E-12	0.009(0.012-0.006)	7,98E-09
*cg19419497	17	74262129	FAM100B	Body	island	0.019(0.024-0.014)	6,12E-12	0.052(0.058-0.047)	5,70E-64
*cg19764539	21	44394723	PKNOX1	5'UTR	island	0.029(0.037-0.021)	6,30E-12	0.008(0.011-0.005)	1,04E-06
cg09915769	19	7660977	KIAA1543	1stExon	island	0.023(0.03-0.017)	6,59E-12	0.001(0.004--0.001)	0,325
*cg24496317	7	24613120	MPP6	1stExon	opensea	0.013(0.017-0.01)	6,66E-12	0.005(0.006-0.003)	6,17E-11
*cg15901997	18	48723744	MEX3C	1stExon	island	0.01(0.012-0.007)	6,86E-12	0.014(0.017-0.01)	1,20E-15
*cg04598683	11	129871924	NCRNA00167	TSS1500	island	0.016(0.021-0.012)	7,06E-12	0.012(0.018-0.005)	0,000323
*cg17255450	4	5710372	EVC2	5'UTR	island	0.05(0.063-0.036)	7,12E-12	0.017(0.021-0.013)	7,19E-20
cg07806715	19	48018254	NAPA	5'UTR	island	0.019(0.025-0.014)	7,32E-12	0.005(0.009-0.002)	0,00294
*cg05425664	17	57184151	TRIM37	1stExon	island	0.021(0.027-0.016)	7,35E-12	0.016(0.019-0.013)	3,47E-24
*cg08625468	17	66287913	SLC16A6	TSS1500	island	0.012(0.016-0.009)	8,32E-12	0.007(0.009-0.004)	1,79E-09
*cg04743947	12	112820215		IGR	island	0.011(0.014-0.008)	9,58E-12	0.01(0.012-0.007)	3,43E-15
*cg27597084	11	124670238	C11orf61	1stExon	island	0.012(0.015-0.008)	9,61E-12	0.027(0.031-0.023)	2,00E-36
*cg20689054	4	153457152	FBXW7	TSS1500	island	0.02(0.026-0.015)	9,69E-12	0.011(0.014-0.008)	4,53E-15
*cg15700006	1	2160297	SKI	1stExon	island	0.017(0.022-0.012)	9,93E-12	0.004(0.007-0)	0,0299
cg16573386	1	1334508	CCNL2	1stExon	island	0.016(0.02-0.012)	9,97E-12	0.001(0.004--0.002) -0.004(-0.006-- 0.002)	0,641 1,12E-05
*cg07813961	3	48956311	C3orf71	1stExon	island	0.015(0.019-0.011)	1,04E-11	0.024(0.03-0.019)	4,36E-16
*cg03024619	3	71803308	GPR27	1stExon	island	0.008(0.01-0.006)	1,17E-11	0.011(0.014-0.008)	2,65E-13
*cg19529645	17	79196743	AZI1	1stExon	island	0.028(0.036-0.02)	1,17E-11	0.004(0.007-0.002)	0,000605
*cg26932552	11	2907116	CDKN1C	TSS200	shore	0.018(0.023-0.013)	1,22E-11	0.036(0.04-0.031)	1,92E-43
*cg06983715	1	86042674	DDAH1	5'UTR	island	0.014(0.018-0.011)	1,23E-11	0.017(0.02-0.013)	1,37E-19
*cg17116739	16	67063591	CBFB	Body	island	0.024(0.03-0.017)	1,34E-11	0.014(0.018-0.01)	5,79E-17
*cg14860494	1	94374966	GCLM	5'UTR	island	0.016(0.021-0.012)	1,45E-11	0.029(0.037-0.021)	1,46E-11
*cg23157089	6	86159962	NT5E	1stExon	island	0.014(0.018-0.01)	1,45E-11		
*cg14810343	5	139028149	CXXC5	TSS200	island	0.029(0.037-0.021)	1,46E-11		

*cg17626178	2	205410273	PAR3B	TSS1500	island	0.021(0.027-0.015)	1,46E-11	0.015(0.018-0.012)	4,00E-18
*cg10423910	5	127420075	SLC12A2	1stExon	island	0.015(0.019-0.011)	1,54E-11	0.013(0.016-0.01)	7,13E-15
*cg08214689	17	16256389	CENPV	1stExon	island	0.016(0.02-0.011)	1,60E-11	0.015(0.018-0.012)	6,21E-23
*cg04729594	2	33172444	LTBP1	1stExon	island	0.022(0.028-0.016)	1,61E-11	0.011(0.015-0.007)	3,13E-08
*cg06699484	5	132166328		IGR	island	0.022(0.028-0.016)	1,63E-11	0.023(0.026-0.02)	3,97E-47
*cg06465196	4	154387603	KIAA0922	1stExon	island	0.013(0.017-0.009)	1,68E-11	0.012(0.014-0.009)	4,23E-21
*cg27250362	16	67597316	CTCF	5'UTR	island	0.019(0.024-0.014)	2,16E-11	0.019(0.023-0.016)	9,95E-29
*cg17813364	7	149321879	ZNF767	Body	island	0.018(0.023-0.013)	2,27E-11	0.01(0.013-0.007)	1,89E-10
*cg19590991	10	11653476	USP6NL	1stExon	island	0.017(0.022-0.012)	2,37E-11	0.015(0.018-0.012)	1,56E-20
*cg04088969	10	11060048	CUGBP2	5'UTR	island	0.022(0.028-0.016)	2,70E-11	0.011(0.015-0.008)	1,21E-10
*cg13800586	2	242626324	DTYMK	5'UTR	island	0.017(0.022-0.012)	2,97E-11	0.009(0.012-0.006)	3,29E-10
*cg02543879	13	53024717	VPS36	1stExon	island	0.018(0.024-0.013)	3,35E-11	0.006(0.01-0.003)	0,000336
cg23281075	1	31538727	PUM1	TSS200	island	0.019(0.025-0.014)	3,40E-11	0.02(0.024-0.017)	1,32E-25
*cg05988158	22	51021418	CHKB-CPT1B	Body	island	0.017(0.022-0.012)	3,57E-11	0.009(0.012-0.006)	1,21E-09
*cg06577463	17	78120714	EIF4A3	1stExon	island	0.02(0.026-0.014)	3,58E-11	0.012(0.015-0.008)	7,78E-10
*cg11052081	2	101767677	TBC1D8	1stExon	island	0.016(0.02-0.011)	3,59E-11	0.019(0.022-0.016)	1,78E-34
*cg05039463	14	91282525	TTC7B	1stExon	island	0.022(0.028-0.016)	3,77E-11	0.016(0.019-0.013)	4,25E-21
*cg09942303	4	166248675	SC4MOL	TSS200	island	0.018(0.024-0.013)	3,88E-11	0.013(0.017-0.009)	7,31E-11
*cg24632597	12	121648093	P2RX4	1stExon	island	0.006(0.007-0.004)	3,89E-11	0.006(0.007-0.004)	2,26E-19
*cg11233236	2	242576818	ATG4B	TSS1500	island	0.014(0.018-0.01)	4,29E-11	0.003(0.005-0)	0,0288
*cg18813733	16	88637115	ZC3H18	5'UTR	island	0.016(0.02-0.011)	4,36E-11	0.009(0.012-0.006)	4,90E-09
*cg18428217	7	6629959	C7orf26	1stExon	island	0.015(0.019-0.011)	4,37E-11	0.016(0.019-0.013)	4,07E-21
*cg12526120	1	193074554	GLRX2	1stExon	island	0.021(0.027-0.015)	4,39E-11	0.014(0.019-0.01)	8,99E-10
*cg09835867	20	31173147		IGR	island	0.016(0.02-0.011)	4,47E-11	0.01(0.014-0.007)	4,48E-11
*cg02831587	11	13690122	FAR1	TSS200	island	0.009(0.012-0.006)	4,59E-11	0.006(0.008-0.004)	1,09E-08
*cg20366603	17	7218821	GPS2	TSS200	island	0.017(0.022-0.012)	5,11E-11	0.016(0.02-0.011)	6,32E-13
*cg26077228	17	80408536	C17orf62	5'UTR	island	0.019(0.024-0.013)	5,40E-11	0.008(0.011-0.004)	7,26E-06
*cg17002091	14	100772763	SLC25A29	1stExon	island	0.025(0.032-0.018)	5,62E-11	0.025(0.029-0.021)	2,03E-30
*cg02543960	18	60383467	PHLPP1	5'UTR	island	0.017(0.022-0.012)	6,00E-11	0.016(0.019-0.013)	2,47E-25
*cg03046332	3	197476906	FYTTD1	5'UTR	island	0.019(0.024-0.013)	6,37E-11	0.014(0.017-0.011)	4,07E-20
*cg05310613	6	139456890	HECA	Body	island	0.025(0.032-0.018)	6,51E-11	0.022(0.027-0.017)	3,47E-19
*cg06202984	9	98079903	FANCC	1stExon	island	0.02(0.026-0.014)	6,63E-11	0.015(0.019-0.011)	1,34E-12
*cg22783262	22	45098346	PRR5-ARHGAP8	TSS200	island	0.016(0.02-0.011)	6,95E-11	0.012(0.016-0.009)	1,32E-10
*cg25410641	2	240322702	HDAC4	TSS200	island	0.019(0.025-0.014)	7,04E-11	0.006(0.009-0.003)	0,000345
*cg11950835	17	25783387		IGR	island	0.02(0.025-0.014)	7,63E-11	0.009(0.012-0.006)	7,39E-09
*cg21135832	4	102268075	PPP3CA	1stExon	island	0.016(0.02-0.011)	7,63E-11	0.015(0.018-0.012)	1,22E-21
*cg19285383	18	5295803	ZFP161	TSS200	island	0.011(0.014-0.008)	7,68E-11	0.012(0.014-0.01)	8,92E-28
*cg27563716	4	184365897	CDKN2AIP	1stExon	island	0.018(0.023-0.013)	7,93E-11	0.015(0.019-0.01)	5,42E-10
*cg20012028	22	39151999	UNC84B	TSS200	island	0.025(0.032-0.018)	8,16E-11	0.014(0.018-0.011)	2,93E-17
*cg21696089	19	33793095	CEBPA	1stExon	island	0.013(0.017-0.01)	8,34E-11	0.004(0.007-0.002)	7,89E-05
*cg21475747	5	177631425	HNRNPAB	TSS200	island	0.018(0.024-0.013)	8,63E-11	0.017(0.021-0.014)	1,08E-19
*cg10196289	8	59572185	NSMAF	1stExon	island	0.015(0.019-0.01)	9,20E-11	0.014(0.017-0.012)	2,98E-23
*cg23702568	10	88281502	WAPAL	5'UTR	island	0.016(0.021-0.011)	9,21E-11	0.01(0.013-0.007)	2,65E-11
*cg09567732	5	177631413	HNRNPAB	TSS200	island	0.036(0.046-0.025)	1,03E-10	0.019(0.024-0.015) -0.008(-0.011-- 0.005)	3,54E-16 2,14E-06
*cg13742648	9	138987085	NACC2	5'UTR	island	0.024(0.031-0.017)	1,04E-10		
*cg20991347	22	44420316	PARVB	Body	island	0.014(0.018-0.01)	1,05E-10	-0.004(-0.008-0)	0,03

*cg21945459	21	47706156	MCM3AP	TSS1500	island	0.023(0.03-0.017)	1,07E-10	0.02(0.024-0.017)	1,70E-24
*cg16340268	1	226925172	ITPKB	5'UTR	island	0.015(0.02-0.011)	1,09E-10	0.014(0.017-0.011)	1,44E-19
*cg01109287	10	65281299	REEP3	5'UTR	island	0.016(0.02-0.011)	1,12E-10	0.015(0.018-0.011)	1,86E-16
*cg07201929	19	33182834	NUDT19	TSS200	island	0.016(0.02-0.011)	1,15E-10	0.015(0.018-0.012)	1,73E-21
*cg01805124	17	30814031	CDK5R1	TSS200	island	0.017(0.022-0.012)	1,16E-10	0.019(0.021-0.016)	9,49E-47
*cg17868942	10	115803819	ADRB1	1stExon	island	0.021(0.027-0.015)	1,16E-10	0.008(0.012-0.004)	4,99E-05
*cg08785922	17	882831	NXN	1stExon	island	0.016(0.02-0.011)	1,19E-10	0.009(0.011-0.006)	1,17E-10
*cg00415642	21	16437351	NRIP1	TSS1500	island	0.016(0.021-0.012)	1,22E-10	0.021(0.024-0.018)	4,15E-38
*cg08522087	5	14871910	ANKH	TSS200	shore	0.025(0.033-0.018)	1,26E-10	0.015(0.019-0.011)	3,70E-11
*cg25433682	16	30407049	ZNF48	TSS200	island	0.022(0.029-0.016)	1,30E-10	0.02(0.024-0.016)	2,07E-24
*cg08158076	1	90287038	LRRC8D	TSS1500	island	0.021(0.027-0.015)	1,35E-10	0.024(0.029-0.02)	2,97E-24
*cg06319359	1	22109595	USP48	5'UTR	island	0.008(0.01-0.006)	1,43E-10	0.002(0.003-0)	0,0224
*cg03214422	6	11538071	TMEM170B	TSS1500	island	0.043(0.056-0.031)	1,46E-10	0.006(0.012-0.001)	0,0281
*cg02385710	11	64884962	ZNHIT2	1stExon	island	0.011(0.014-0.008)	1,65E-10	0.008(0.01-0.006)	3,75E-12
*cg00279406	1	226924739	ITPKB	Body	island	0.017(0.022-0.012)	1,66E-10	0.01(0.013-0.006)	4,32E-08
*cg12160011	6	111804095	REV3L	1stExon	island	0.022(0.028-0.015)	1,76E-10	0.014(0.017-0.011)	1,13E-15
*cg11969191	17	79679066	SLC25A10	TSS1500	island	0.02(0.026-0.014)	1,83E-10	NA	NA
*cg21160852	5	892775	TRIP13	TSS200	island	0.016(0.021-0.011)	1,89E-10	0.025(0.03-0.02)	1,31E-21
*cg01988923	14	77493722	C14orf4	1stExon	island	0.013(0.016-0.009)	1,97E-10	0.01(0.013-0.007)	6,52E-11
*cg02418294	1	14076293	PRDM2	Body	island	0.011(0.014-0.008)	2,03E-10	0.004(0.006-0.002)	0,000128
*cg26975609	7	150777587	FASTK	Body	shore	0.018(0.024-0.013)	2,04E-10	0.008(0.011-0.004)	3,66E-05
*cg19379522	20	35492281	C20orf117	TSS200	island	0.015(0.019-0.01)	2,06E-10	0.008(0.01-0.005)	6,99E-08
*cg05232802	7	156742224	NOM1	TSS200	shore	0.029(0.037-0.02)	2,24E-10	-0.003(-0.008-0.002)	0,285
*cg00850340	10	126432777	FAM53B	5'UTR	island	0.022(0.029-0.016)	2,34E-10	0.031(0.035-0.026)	1,80E-37
*cg18719198	14	55369460	GCH1	5'UTR	island	0.024(0.031-0.017)	2,40E-10	0.026(0.03-0.022)	8,66E-31
*cg16560453	3	101395512	ZBTB11	1stExon	island	0.017(0.022-0.012)	2,43E-10	0.013(0.016-0.01)	1,63E-16
*cg23246706	21	33104009	SFRS15	1stExon	island	0.019(0.025-0.013)	2,46E-10	0.014(0.017-0.011)	1,67E-19
*cg07232033	2	236403031	AGAP1	1stExon	island	0.028(0.036-0.02)	2,56E-10	0.02(0.025-0.016)	4,88E-18
*cg13079099	19	54974383	LENG9	1stExon	island	0.014(0.018-0.01)	2,63E-10	0.011(0.014-0.009)	3,55E-16
*cg03183872	20	3140552	FASTKD5	TSS200	island	0.03(0.039-0.021)	2,65E-10	0.004(0.011--0.002)	0,21
*cg10152756	16	69364566	COG8	3'UTR	island	0.011(0.015-0.008)	2,66E-10	NA	NA
cg13696351	9	140063617	LRRC26	Body	island	0.024(0.032-0.017)	3,04E-10	0.018(0.022-0.014)	1,15E-16
*cg07565236	16	88869998	CDT1	TSS200	island	0.022(0.028-0.015)	3,05E-10	0.012(0.015-0.009)	8,31E-14
*cg19099833	1	201476290	CSRP1	5'UTR	island	0.017(0.023-0.012)	3,06E-10	0.013(0.016-0.01)	2,70E-16
*cg05705366	5	53813915	SNX18	1stExon	island	0.018(0.023-0.012)	3,08E-10	0.032(0.036-0.028)	8,41E-53
*cg06704539	7	102105256	ALKBH4	1stExon	island	0.019(0.024-0.013)	3,30E-10	0.018(0.021-0.015)	1,90E-28
*cg27523523	20	39766360	PLCG1	1stExon	island	0.014(0.018-0.01)	3,33E-10	0.014(0.017-0.011)	2,64E-22
*cg07352798	5	141488591	NDFIP1	1stExon	island	0.012(0.016-0.009)	3,67E-10	-0.059(-0.067--0.051)	1,14E-46
*cg05576451	11	125757753	HYLS1	5'UTR	island	0.023(0.029-0.016)	3,85E-10	0.009(0.013-0.004)	0,000189
*cg16706496	2	97523911	ANKRD39	TSS200	island	0.02(0.026-0.014)	3,87E-10	0.012(0.016-0.008)	1,66E-09
*cg25139493	1	39957400	BMP8A	1stExon	island	0.034(0.045-0.024)	3,95E-10	0.03(0.035-0.025)	6,41E-26
*cg02079880	6	157802436	ZDHHC14	TSS200	island	0.008(0.011-0.006)	4,05E-10	-0.018(-0.021--0.015)	6,63E-30
*cg13895544	3	66550922	LRIG1	TSS200	island	0.012(0.015-0.008)	4,09E-10	0.006(0.008-0.004)	8,02E-08
*cg27313501	16	2034198	GFER	1stExon	island	0.019(0.024-0.013)	4,14E-10	0.022(0.027-0.018)	2,41E-23
*cg14971823	17	73585270	MYO15B	Body	island	0.016(0.021-0.011)	4,35E-10	0.01(0.013-0.007)	1,56E-12

*cg00177013	20	62339100	ARFRP1	5'UTR	island	0.014(0.018-0.009)	4,46E-10	0.017(0.021-0.012)	1,27E-11
*cg08655589	3	14444175	SLC6A6	5'UTR	island	0.034(0.045-0.024)	4,50E-10	0.026(0.034-0.017)	1,96E-09
*cg23656755	1	230203043	GALNT2	1stExon	island	0.032(0.042-0.023)	4,61E-10	0.06(0.068-0.053)	2,91E-52
*cg19513247	17	42297053	UBTF	TSS200	island	0.018(0.023-0.012)	4,62E-10	0.056(0.062-0.05)	3,04E-64
*cg01112643	6	157802216	ZDHC14	TSS1500	island	0.019(0.024-0.013)	4,75E-10	0.027(0.031-0.022)	1,41E-32
*cg12870876	2	203736356	ICA1L	5'UTR	island	0.02(0.027-0.014)	5,00E-10	0.012(0.015-0.009)	1,81E-12
*cg20962543	1	85666741	SYDE2	TSS200	island	0.015(0.019-0.01)	5,01E-10	0.008(0.011-0.005)	3,00E-08
*cg00799842	20	44563271	PCIF1	TSS200	island	0.017(0.022-0.012)	5,10E-10	0.012(0.016-0.008)	1,30E-09
*cg04415045	11	10315754	SBF2	TSS200	island	0.012(0.015-0.008)	5,34E-10	0.01(0.012-0.007)	3,27E-14
*cg02590295	8	66754114	PDE7A	TSS1500	island	0.013(0.017-0.009)	5,37E-10	0.007(0.01-0.004)	5,46E-06
*cg02830438	14	93651416	C14orf109	5'UTR	island	0.019(0.025-0.013)	5,60E-10	0.014(0.018-0.01)	9,02E-12
*cg07428697	7	106301118	FLJ36031	1stExon	island	0.013(0.017-0.009)	5,60E-10	0.007(0.01-0.004)	1,44E-05
*cg10446143	21	44394730	PKNOX1	5'UTR	island	0.024(0.032-0.017)	5,62E-10	0.022(0.028-0.017)	1,04E-13
*cg18644710	2	201172048	SPATS2L	5'UTR	island	0.017(0.022-0.012)	5,80E-10	0.007(0.01-0.005)	7,71E-08
*cg18963390	3	128902676	CNBP	1stExon	island	0.015(0.019-0.01)	5,91E-10	0.003(0.006-0.001)	0,0183
*cg02182114	7	139876578	LOC100134229	TSS1500	island	0.023(0.03-0.016)	6,01E-10	0.051(0.056-0.046)	9,42E-72
*cg12888961	12	96794220	CDK17	5'UTR	island	0.015(0.02-0.011)	6,28E-10	0.015(0.017-0.012)	9,44E-24
*cg13716829	9	140196785	NRARP	TSS200	island	0.019(0.025-0.013)	6,37E-10	0.005(0.009-0.001)	0,0126
*cg14283059	4	186125358	KIAA1430	TSS200	island	0.019(0.024-0.013)	6,42E-10	0.012(0.016-0.009)	3,37E-11
*cg12570942	2	242626270	DTYMK	5'UTR	island	0.018(0.024-0.013)	6,73E-10	0.002(0.005--0.001)	0,112
*cg23453622	17	79885647	LOC92659	TSS200	island	0.018(0.024-0.013)	7,05E-10	0.008(0.011-0.005)	2,32E-07
*cg23722428	10	35931117	FZD8	TSS1500	island	0.026(0.034-0.018)	7,13E-10	0.018(0.022-0.015)	2,62E-20
*cg12970711	22	21984177	YDJC	1stExon	island	0.016(0.02-0.011)	7,49E-10	0.004(0.007-0.001)	0,00471
cg08184047	17	79849980	ANAPC11	5'UTR	island	0.018(0.024-0.013)	7,54E-10	0.027(0.031-0.023)	3,36E-32
*cg05417087	5	139726056	HBEGF	1stExon	island	0.016(0.02-0.011)	7,59E-10	0.015(0.019-0.012)	3,29E-20
*cg02839273	15	85525387	PDE8A	5'UTR	island	0.019(0.025-0.013)	7,70E-10	0.012(0.016-0.008)	8,10E-09
*cg02631957	5	122847966	CSNK1G3	1stExon	island	0.018(0.023-0.012)	8,09E-10	0.014(0.017-0.01)	3,19E-14
*cg11162385	20	25604740	NANP	TSS200	island	0.022(0.028-0.015)	8,21E-10	0.012(0.019-0.005)	0,00125
*cg01419788	14	93581878	ITPK1	5'UTR	island	0.01(0.013-0.007)	8,26E-10	0.003(0.006-0.001)	0,00497
*cg24064055	8	145104176		IGR	island	0.016(0.021-0.011)	8,41E-10	0.005(0.008-0.003)	0,000119
*cg08873779	16	88772658	RNF166	1stExon	island	0.018(0.024-0.013)	8,51E-10	0.01(0.013-0.007)	1,80E-11
*cg26161725	5	36241580	C5orf33	5'UTR	island	0.018(0.024-0.013)	8,59E-10	0.016(0.019-0.012) -0.012(-0.015-- 0.008)	8,97E-21 1,29E-10
*cg17815841	4	39979368	PDS5A	5'UTR	island	0.015(0.02-0.011)	8,95E-10	0.01(0.013-0.007)	1,35E-09
*cg21599869	2	239335445	ASB1	TSS200	island	0.018(0.024-0.013)	9,44E-10	0.018(0.022-0.014)	2,43E-17
*cg21961583	2	170684455	UBR3	1stExon	island	0.017(0.022-0.012)	9,55E-10	0.012(0.015-0.008)	1,85E-12
*cg10314752	14	94640828	PPP4R4	1stExon	island	0.013(0.016-0.009)	9,86E-10	0.004(0.007-0.002)	0,00131
*cg15253587	19	18943093	UPF1	1stExon	island	0.018(0.024-0.012)	1,05E-09	0.016(0.019-0.012)	2,92E-18
*cg00207389	20	18569003	DTD1	Body	island	0.015(0.02-0.01)	1,09E-09	0.019(0.022-0.016)	9,91E-31
*cg07535605	18	158294	USP14	TSS200	island	0.014(0.018-0.009)	1,11E-09	0.005(0.007-0.002)	0,000756
*cg03496713	7	149570987	ATP6V0E2	5'UTR	island	0.015(0.02-0.011)	1,17E-09	0.014(0.018-0.011)	6,28E-14
*cg01622965	7	98476371	TRRAP	5'UTR	island	0.014(0.019-0.01)	1,20E-09	0.017(0.02-0.014)	8,08E-26
*cg25889945	13	77460298	KCTD12	5'UTR	island	0.018(0.023-0.012)	1,25E-09	0.014(0.017-0.01)	3,18E-12
*cg08104527	18	13217953	C18orf1	TSS1500	island	0.02(0.026-0.014)	1,27E-09	0.003(0.006-0)	0,0867
*cg18327423	11	3186549	OSBPL5	1stExon	island	0.011(0.015-0.008)	1,28E-09	0.002(0.005-0)	0,0435
*cg14461650	2	101925089	RNF149	1stExon	island	0.024(0.031-0.016)	1,32E-09	NA	NA
*cg07079421	22	46933261	CELSR1	TSS200	island				

*cg22014600	13	21750851	MRP63	1stExon	island	0.034(0.045-0.024)	1,32E-09	0.007(0.012-0.001)	0,0199
*cg03494795	9	133710827	ABL1	TSS200	island	0.015(0.02-0.011)	1,33E-09	0.012(0.015-0.008)	2,52E-11
*cg12075970	12	122064387	ORAI1	TSS200	island	0.021(0.028-0.015)	1,36E-09	0.028(0.032-0.023)	1,28E-31
*cg22320000	1	245027629	HNRNPU	5'UTR	island	0.017(0.022-0.012)	1,36E-09	0.002(0.006--0.001)	0,226
*cg10792544	9	73704		IGR	island	0.023(0.03-0.016)	1,43E-09	NA	NA
*cg24770850	3	129325130	PLXND1	1stExon	island	0.012(0.016-0.008)	1,45E-09	0.01(0.012-0.008)	3,10E-18
*cg05029189	3	123168386	ADCY5	TSS1500	island	0.021(0.028-0.015)	1,47E-09	0.017(0.021-0.013)	8,50E-19
*cg11496792	9	123477254	MEGF9	TSS1500	island	0.015(0.02-0.01)	1,51E-09	0.018(0.022-0.014)	2,64E-20
*cg02249648	12	1703181	FBXL14	1stExon	island	0.016(0.022-0.011)	1,52E-09	0.006(0.01-0.003)	0,000435
*cg06352072	18	59561651	RNF152	TSS1500	island	0.018(0.023-0.012)	1,52E-09	0.014(0.018-0.011)	4,69E-17
*cg01896926	17	685509	GLOD4	5'UTR	island	0.026(0.034-0.018)	1,62E-09	0.011(0.015-0.006)	3,29E-06
*cg26955196	15	23085768	NIPA1	Body	island	0.018(0.024-0.013)	1,65E-09	0.007(0.011-0.003)	0,00112
*cg05964635	6	15662707	DTNBP1	Body	island	0.013(0.017-0.009)	1,66E-09	0.003(0.005-0.001)	0,0133
*cg13547299	11	1330611	TOLLIP	Body	island	0.016(0.021-0.011)	1,77E-09	0.008(0.011-0.006)	1,78E-08
*cg01470744	9	134406824	UCK1	TSS200	island	0.019(0.025-0.013)	1,82E-09	0.007(0.01-0.004)	2,43E-05
*cg11287521	10	124134015	PLEKHA1	TSS1500	shore	0.02(0.026-0.014)	1,85E-09	0.03(0.035-0.026)	5,15E-34
*cg04830146	17	31203787	MYO1D	Body	island	0.011(0.014-0.007)	1,93E-09	0.007(0.01-0.005)	7,16E-11
*cg10460130	2	242625978	DTYMK	Body	island	0.027(0.036-0.019)	1,93E-09	0.023(0.03-0.016)	4,05E-10
*cg05866713	16	84150400	MBTPS1	1stExon	island	0.009(0.012-0.006)	1,94E-09	0.007(0.009-0.005)	5,77E-14
*cg08497766	10	6019877	IL15RA	TSS1500	island	0.016(0.021-0.011)	1,96E-09	0.01(0.014-0.006)	5,08E-07
*cg03231960	19	49139050	DBP	Body	island	0.021(0.027-0.014)	2,01E-09	0.004(0.007-0.001)	0,0221
*cg09693228	4	2965311	GRK4	TSS200	island	0.012(0.016-0.009)	2,16E-09	0.015(0.017-0.012)	1,21E-32
*cg22534627	12	65004466	RASSF3	1stExon	island	0.02(0.026-0.014)	2,23E-09	0.011(0.015-0.007)	1,45E-08
cg01733795	17	7465439	SEN3	1stExon	island	0.016(0.022-0.011)	2,27E-09	0.011(0.014-0.008)	8,44E-12
*cg02126117	13	26626092	SHISA2	TSS1500	island	0.023(0.03-0.015)	2,42E-09	0.02(0.024-0.017)	3,54E-26
*cg25487903	19	10828950	DNM2	1stExon	island	0.013(0.017-0.009)	2,67E-09	0.013(0.017-0.01)	1,61E-15
*cg16092956	4	185747098	ACSL1	5'UTR	island	0.019(0.025-0.013)	2,70E-09	0.014(0.017-0.01)	8,84E-12
*cg05970992	4	6785456	KIAA0232	5'UTR	shore	0.013(0.017-0.009)	2,73E-09	0.01(0.014-0.007)	6,66E-09
*cg16975959	2	223184510		IGR	island	0.016(0.022-0.011)	2,75E-09	0.016(0.019-0.012)	4,39E-18
*cg11230222	17	40831789	CCR10	Body	island	0.01(0.013-0.007)	2,80E-09	0.006(0.008-0.004)	1,29E-08
*cg03696327	1	101005121	GPR88	Body	island	0.02(0.027-0.014)	2,82E-09	0.017(0.02-0.013)	2,13E-18
*cg25817927	7	158621785	ESYT2	1stExon	island	0.013(0.017-0.009)	2,86E-09	0.009(0.011-0.007)	3,65E-15
*cg15420687	5	159546321	PWWP2A	1stExon	island	0.016(0.021-0.011)	2,95E-09	0.013(0.017-0.01)	2,64E-14
*cg06330324	22	31031670	SLC35E4	TSS200	island	0.009(0.012-0.006)	3,06E-09	0.006(0.008-0.004)	1,53E-07
*cg15798279	7	6523691	KDEL2	1stExon	island	0.017(0.022-0.012)	3,12E-09	0.018(0.022-0.014)	3,62E-18
*cg24640588	4	1283718	MAEA	1stExon	island	0.02(0.026-0.013)	3,26E-09	0.013(0.016-0.009)	2,35E-10
*cg17218495	19	11071743	SMARCA4	5'UTR	island	0.02(0.027-0.014)	3,29E-09	0.009(0.013-0.005)	3,63E-06
cg23962478	22	50354086	PIM3	TSS200	island	0.021(0.027-0.014)	3,34E-09	0.011(0.015-0.008)	2,55E-09
*cg05528568	3	126194162	ZXDC	1stExon	island	0.011(0.014-0.007)	3,35E-09	0.006(0.008-0.004)	6,09E-08
*cg10702645	1	38259565	MANEAL	TSS1500	island	0.015(0.02-0.01)	3,38E-09	0.005(0.008-0.003)	2,44E-05
*cg12869623	17	9548898	USP43	TSS200	island	0.02(0.026-0.013)	3,43E-09	0.019(0.023-0.015)	3,06E-23
*cg02114786	21	46293672	PTTG1IP	5'UTR	island	0.013(0.017-0.009)	3,48E-09	0.011(0.014-0.009)	9,46E-17
*cg19191624	2	32582276	BIRC6	1stExon	island	0.013(0.018-0.009)	3,49E-09	0.005(0.008-0.003)	9,01E-05
*cg16368059	13	50366559	KPNA3	Body	island	0.011(0.015-0.008)	3,50E-09	0.008(0.01-0.006)	2,64E-15
*cg00497630	17	26733052	SLC46A1	1stExon	island	0.014(0.018-0.009)	3,69E-09	0.081(0.089-0.072)	8,84E-66
*cg18001714	5	14871894	ANKH	TSS200	shore	0.02(0.026-0.013)	3,80E-09	0.015(0.018-0.012)	3,25E-20

*cg14702655	14	72398920	RGS6	TSS1500	island	0.013(0.017-0.009)	3,87E-09	0.008(0.011-0.004)	5,29E-05
*cg09009111	18	2847309	EMILIN2	1stExon	island	0.013(0.018-0.009)	3,94E-09	-0.003(-0.006-0.001)	0,114
*cg15739944	9	99212660	HABP4	1stExon	island	0.014(0.018-0.009)	4,02E-09	0.01(0.012-0.007)	4,70E-13
*cg06249499	6	166796752	BRP44L	TSS1500	island	0.017(0.022-0.011)	4,03E-09	0.008(0.011-0.006)	5,21E-09
*cg08723131	20	55043694	C20orf43	1stExon	island	0.01(0.013-0.006)	4,15E-09	0(0.003--0.003)	0,894
*cg04297258	16	88772664	RNF166	1stExon	island	0.021(0.028-0.014)	4,27E-09	0.033(0.038-0.028)	2,81E-37
*cg02431098	1	51425775	FAF1	1stExon	island	0.013(0.017-0.009)	4,34E-09	0.016(0.018-0.013)	4,12E-31
*cg03415497	20	54933931	C20orf108	TSS200	shore	0.02(0.027-0.014)	4,39E-09	0.016(0.019-0.014)	2,57E-33
*cg01263850	1	19283164	IFFO2	TSS1500	island	0.023(0.03-0.015)	4,40E-09	0.033(0.038-0.028)	2,11E-37
*cg06706183	6	53409856	GCLC	TSS200	island	0.009(0.012-0.006)	4,40E-09	0(0.002--0.002)	0,963
*cg00759551	3	133292899	CDV3	1stExon	island	0.014(0.018-0.009)	4,46E-09	0.017(0.02-0.014)	2,22E-25
*cg10299667	17	79679470	SLC25A10	1stExon	island	0.009(0.012-0.006)	4,46E-09	0.001(0.003--0.001)	0,155
cg00614832	16	73092394	ZFHX3	5'UTR	island	0.043(0.056-0.029)	4,49E-09	-0.013(-0.02--0.006)	0,00052
*cg13782561	9	94185786	NFIL3	5'UTR	island	0.011(0.014-0.007)	4,55E-09	0.001(0.002--0.001)	0,409
*cg06618322	15	85923867	AKAP13	TSS200	island	0.021(0.028-0.014)	4,66E-09	0.014(0.017-0.011)	3,32E-16
*cg01637218	1	29063669	YTHDF2	Body	island	0.012(0.016-0.008)	4,69E-09	0.012(0.015-0.009)	6,32E-14
*cg27016262	17	1012835	ABR	Body	island	0.028(0.037-0.019)	4,69E-09	0.024(0.03-0.019)	1,47E-16
*cg18120578	12	77459407	E2F7	TSS200	island	0.02(0.027-0.014)	4,78E-09	0.016(0.019-0.012)	1,17E-15
*cg02680710	6	169310568		IGR	island	0.081(0.107-0.055)	5,13E-09	NA	NA
*cg00666849	2	85839032	C2orf68	1stExon	island	0.015(0.019-0.01)	5,21E-09	-0.001(-0.005-0.003)	0,601
*cg09012337	11	73019505	ARHGEF17	TSS200	island	0.013(0.018-0.009)	5,23E-09	0.01(0.013-0.006)	2,86E-07
*cg02683985	7	559070	PDGFA	5'UTR	island	0.017(0.023-0.012)	5,28E-09	0.021(0.024-0.017)	3,89E-33
*cg06951969	15	66996086	SMAD6	Body	island	0.014(0.018-0.009)	5,39E-09	0.023(0.026-0.02)	6,91E-43
*cg08363925	11	86511378	PRSS23	TSS200	island	0.02(0.026-0.013)	5,44E-09	0.01(0.014-0.007) -0.004(-0.007-- 0.002)	4,55E-08 0,000451
*cg11002404	10	11060038	CUGBP2	5'UTR	island	0.009(0.011-0.006)	5,54E-09	0.035(0.04-0.031)	2,75E-41
*cg01943813	6	24495061	ALDH5A1	TSS200	island	0.021(0.028-0.014)	5,55E-09	0.006(0.009-0.003)	9,41E-05
*cg01025283	3	138327728	FAIM	5'UTR	island	0.014(0.018-0.009)	5,59E-09	0.006(0.009-0.003)	2,24E-05
*cg13865502	9	140196251	NRARP	1stExon	island	0.02(0.026-0.013)	5,65E-09	0.013(0.017-0.01)	1,93E-13
*cg19044674	1	43232628	LEPRE1	1stExon	island	0.015(0.02-0.01)	5,68E-09	0.011(0.015-0.008)	2,47E-09
*cg00787661	2	232329533	NCL	TSS1500	island	0.015(0.02-0.01)	5,72E-09	-0.001(-0.004-0.002)	0,554
*cg20367329	17	17399941	RASD1	TSS1500	island	0.022(0.029-0.015)	5,74E-09	0.008(0.01-0.005)	1,77E-07
cg04355250	12	113796401	PLBD2	1stExon	island	0.014(0.019-0.009)	5,76E-09	0.001(0.003-0)	0,0801
*cg17419935	1	167691120	MPZL1	TSS200	island	0.007(0.01-0.005)	5,91E-09	0.007(0.011-0.004)	1,65E-05
*cg15584531	13	99739316	DOCK9	TSS1500	island	0.017(0.022-0.011)	5,95E-09	0.018(0.021-0.015)	1,09E-28
*cg26732930	15	32162819		IGR	opensea	0.014(0.018-0.009)	5,96E-09	0.015(0.019-0.01)	6,96E-11
*cg07515367	17	45401150	C17orf57	TSS200	island	0.023(0.03-0.015)	5,97E-09	0(0.002--0.002)	0,849
*cg15011622	7	36193062	EEPD1	1stExon	island	0.01(0.013-0.007)	6,23E-09	0.004(0.008-0.001)	0,0259
*cg07753241	3	63849527	ATXN7	TSS1500	island	0.02(0.027-0.014)	6,24E-09	0.006(0.009-0.003) -0.019(-0.022-- 0.016)	0,000302 1,32E-36
*cg20184917	16	89767386	SPATA2L	Body	island	0.016(0.021-0.011)	6,24E-09	0.007(0.01-0.004)	2,32E-05
*cg03457528	8	128806865	MIR1204	TSS1500	island	0.006(0.008-0.004)	6,56E-09	0.008(0.01-0.006)	1,32E-11
*cg21196132	22	37915278	CARD10	TSS200	island	0.02(0.026-0.013)	6,70E-09	0.007(0.01-0.004)	7,69E-06
*cg01190566	19	47363949		IGR	island	0.011(0.014-0.007)	6,80E-09	0.004(0.007-0.001)	0,0201
*cg08005943	1	231004157	C1orf198	1stExon	island	0.014(0.018-0.009)	6,91E-09	0.003(0.006-0.001)	0,0108
*cg04357830	21	36261693	RUNX1	TSS1500	island	0.017(0.022-0.011)	7,26E-09		
*cg08264704	1	228270414	ARF1	5'UTR	island	0.012(0.016-0.008)	7,28E-09		

*cg02994863	1	64059297	PGM1	1stExon	island	0.031(0.041-0.021)	7,40E-09	0.02(0.028-0.012)	1,04E-06
*cg06210969	3	120067732	LRRC58	1stExon	island	0.012(0.015-0.008)	7,45E-09	0.014(0.016-0.011)	1,68E-20
*cg15008401	4	102267974	PPP3CA	1stExon	island	0.013(0.018-0.009)	7,50E-09	0.016(0.019-0.013)	6,13E-21
*cg00816008	16	82203996	MPHOSPH6	TSS200	island	0.01(0.013-0.006)	7,51E-09	0.001(0.003--0.001)	0,501
*cg08903333	21	45209501	RRP1	1stExon	island	0.012(0.016-0.008)	7,70E-09	0.008(0.011-0.006)	4,67E-09
*cg24877195	17	61554604	ACE	1stExon	island	0.016(0.021-0.011)	7,77E-09	0.01(0.013-0.007)	1,51E-11
*cg06066711	1	6662645	KLHL21	1stExon	island	0.015(0.021-0.01)	7,78E-09	0.01(0.013-0.008)	9,15E-13
*cg17906179	3	194991500	C3orf21	1stExon	island	0.016(0.022-0.011)	7,83E-09	0.01(0.013-0.006)	3,97E-08
*cg24252614	5	56247957	MIER3	TSS200	island	0.017(0.023-0.011)	7,91E-09	0.017(0.02-0.014)	3,01E-29
*cg26154342	12	94542977	PLXNC1	1stExon	island	0.016(0.021-0.011)	8,06E-09	0.016(0.019-0.013)	5,59E-22
*cg06920740	11	534896	HRAS	5'UTR	island	0.012(0.016-0.008)	8,15E-09	0.001(0.004--0.002)	0,509
*cg14397361	9	140149997	COBRA1	1stExon	island	0.01(0.013-0.006)	8,15E-09	0.006(0.009-0.004)	2,45E-06
*cg19373170	3	53079105	SFMBT1	1stExon	island	0.013(0.018-0.009)	8,15E-09	0.013(0.016-0.009)	2,15E-14
*cg12423398	17	655593	ELP2P	TSS1500	island	0.009(0.012-0.006)	8,25E-09	0.005(0.008-0.003)	3,07E-05
*cg06584662	3	10362768	SEC13	Body	island	0.005(0.007-0.004)	8,30E-09	-0.034(-0.038--0.03)	1,67E-49
*cg10107929	13	49550206	FNDC3A	5'UTR	island	0.016(0.022-0.011)	8,35E-09	0.009(0.012-0.006)	6,20E-09
*cg20950465	1	53068181	GPX7	1stExon	island	0.021(0.028-0.014)	8,65E-09	0.03(0.036-0.024)	8,27E-19
*cg09349094	1	2323068	MORN1	TSS200	island	0.012(0.016-0.008)	8,79E-09	0.005(0.008-0.002)	0,000376
*cg02750385	13	41885554	NAA16	5'UTR	island	0.018(0.024-0.012)	8,88E-09	0.015(0.018-0.011)	4,39E-13
*cg09055943	21	43431271	ZNF295	TSS1500	island	0.014(0.018-0.009)	8,94E-09	0.013(0.016-0.01)	1,08E-18
*cg07946458	21	38338892	HLCS	5'UTR	island	0.013(0.017-0.009)	9,33E-09	0.008(0.01-0.006)	6,66E-12
*cg27007358	7	21467723	SP4	5'UTR	island	0.015(0.02-0.01)	9,46E-09	0.016(0.02-0.013)	1,07E-19
*cg16869008	8	145550615	DGAT1	TSS200	island	0.014(0.019-0.009)	9,58E-09	0.011(0.014-0.008)	1,33E-14
*cg10239074	2	111880006	BCL2L11	5'UTR	island	0.006(0.008-0.004)	9,67E-09	0.01(0.012-0.008)	4,09E-21
*cg06180363	19	59056346	TRIM28	1stExon	island	0.015(0.02-0.01)	9,74E-09	0.017(0.021-0.013)	1,28E-16
*cg04879696	16	89574810	SPG7	1stExon	island	0.012(0.016-0.008)	9,75E-09	0.005(0.007-0.003)	5,51E-05
*cg25547361	5	76926079	OTP	3'UTR	island	0.014(0.019-0.01)	9,81E-09	0.012(0.015-0.009)	1,83E-14
*cg16697493	16	476419	RAB11FIP3	1stExon	island	0.021(0.028-0.014)	9,98E-09	0.014(0.018-0.011)	4,86E-13
*cg02478956	15	22833681	TUBGCP5	Body	island	0.029(0.038-0.019)	1,00E-08	-0.034(-0.042--0.025)	5,20E-14
*cg25707994	7	157129685	DNAJB6	TSS200	island	0.027(0.035-0.018)	1,00E-08	-0.016(-0.024--0.007)	0,000359
*cg01854842	20	49547693	ADNP	TSS200	shore	0.017(0.022-0.011)	1,05E-08	0.01(0.014-0.007)	6,91E-10
*cg10531748	10	79686265	LOC100128292	TSS1500	island	0.013(0.017-0.008)	1,05E-08	0.009(0.011-0.007)	1,20E-14
*cg01332181	19	47290716	SLC1A5	1stExon	island	0.026(0.034-0.017)	1,08E-08	0.003(0.008--0.002)	0,214
*cg08475379	20	37302989		IGR	island	0.016(0.022-0.011)	1,08E-08	-0.014(-0.019--0.009)	2,87E-08
*cg27297221	19	663093	RNF126	1stExon	island	0.005(0.007-0.004)	1,08E-08	0.007(0.008-0.005)	2,05E-19
*cg06534422	22	24236603	MIF	1stExon	island	0.013(0.017-0.009)	1,09E-08	0.006(0.009-0.003)	0,000215
cg07619799	4	185747409	ACSL1	TSS200	island	0.014(0.018-0.009)	1,10E-08	0.003(0.007-0)	0,0793
*cg17361641	19	18391832	JUND	1stExon	island	0.015(0.02-0.01)	1,10E-08	0.004(0.007-0.002)	0,000827
*cg12094402	4	3767402	ADRA2C	TSS1500	island	0.015(0.02-0.01)	1,11E-08	0.015(0.018-0.012)	7,96E-19
*cg17948846	16	68119548	NFATC3	1stExon	island	0.019(0.026-0.013)	1,11E-08	0.012(0.016-0.009)	1,61E-11
*cg15814736	11	63439115	ATL3	TSS200	island	0.02(0.027-0.014)	1,12E-08	0.009(0.012-0.006)	3,95E-10
*cg23680936	10	64564801	ADO	5'UTR	island	0.009(0.012-0.006)	1,15E-08	0.006(0.008-0.004)	3,47E-07
*cg25390656	10	92980550	PCGF5	5'UTR	island	0.017(0.023-0.011)	1,15E-08	-0.002(-0.006-0.002)	0,383
*cg04889973	12	44229713	TMEM117	TSS1500	island	0.013(0.017-0.009)	1,16E-08	0.005(0.007-0.002)	0,00215

*cg17630771	11	101981004	YAP1	TSS1500	island	0.018(0.025-0.012)	1,18E-08	-0.019(-0.024--0.014)	6,00E-13
*cg01264747	11	43964451	LOC387763	5'UTR	island	0.016(0.021-0.01)	1,20E-08	0.009(0.012-0.006)	2,00E-08
*cg18486630	20	21283939	XRN2	TSS200	island	0.015(0.02-0.01)	1,20E-08	NA	NA
*cg01513081	6	105584780	BVES	TSS1500	island	0.016(0.022-0.011)	1,22E-08	0.009(0.012-0.007)	3,12E-10
*cg02731774	17	45810615	TBX21	1stExon	island	0.01(0.013-0.006)	1,24E-08	0.006(0.008-0.004)	5,99E-09
*cg08514895	8	124286513	ZHX1	5'UTR	island	0.011(0.014-0.007)	1,25E-08	0.008(0.01-0.006)	4,09E-12
*cg13311803	7	101930366	SH2B2	Body	island	0.017(0.022-0.011)	1,25E-08	0.021(0.025-0.018)	1,70E-28
*cg04593445	16	2265125	PGP	TSS1500	island	0.016(0.021-0.01)	1,28E-08	0.006(0.009-0.003)	0,000431
*cg21740452	21	34143691	C21orf66	1stExon	island	0.013(0.017-0.008)	1,28E-08	0.009(0.012-0.006)	9,25E-08
*cg09864858	1	65613311	AK3L1	5'UTR	island	0.017(0.023-0.012)	1,29E-08	0.016(0.02-0.013)	9,18E-18
*cg09795194	17	19770910	ULK2	1stExon	island	0.006(0.008-0.004)	1,34E-08	0.007(0.009-0.006)	3,58E-21
*cg19906131	11	93862516	PANX1	1stExon	island	0.016(0.022-0.011)	1,35E-08	0.01(0.014-0.006)	1,06E-05
*cg24059718	21	45138843	PDXK	TSS200	island	0.005(0.007-0.003)	1,35E-08	0.003(0.005-0.002)	3,40E-09
*cg05324982	17	14204478	HS3ST3B1	TSS200	island	0.017(0.023-0.011)	1,36E-08	0.009(0.013-0.006)	6,22E-07
*cg07706463	1	101005451	GPR88	Body	island	0.013(0.017-0.009)	1,36E-08	0.01(0.013-0.007)	1,15E-12
*cg00784718	1	221916686	DUSP10	TSS1500	island	0.022(0.029-0.015)	1,38E-08	0.015(0.019-0.011)	2,79E-12
*cg02108176	1	22109593	USP48	5'UTR	island	0.015(0.019-0.01)	1,38E-08	0.021(0.024-0.017)	2,65E-27
*cg05914712	19	33864373	CEBPG	TSS1500	island	0.005(0.007-0.004)	1,42E-08	0.005(0.006-0.004) -0.005(-0.008--0.001)	2,01E-17 0,012
*cg11199399	2	33172442	LTBP1	1stExon	island	0.015(0.021-0.01)	1,42E-08	0.005(0.008-0.003)	1,73E-05
*cg04249066	8	144680433	EEF1D	TSS1500	island	0.007(0.009-0.005)	1,45E-08	0.014(0.017-0.011)	1,95E-19
*cg05256242	12	12876945		IGR	island	0.015(0.02-0.01)	1,47E-08	0.011(0.015-0.007)	4,15E-08
*cg23758016	17	73521635	LLGL2	TSS200	island	0.022(0.03-0.015)	1,47E-08	0.021(0.029-0.014)	3,26E-08
*cg11017269	5	95067136	RHOBTB3	1stExon	island	0.021(0.028-0.014)	1,48E-08	-0.002(-0.004-0.001)	0,247
*cg13699657	7	156931944	UBE3C	5'UTR	island	0.01(0.014-0.007)	1,49E-08	0.01(0.015-0.005)	6,14E-05
*cg01398415	1	226309536		IGR	island	0.016(0.021-0.011)	1,51E-08	0.003(0.005-0.001)	0,00148
*cg23278060	6	83777448	DOPEY1	1stExon	island	0.006(0.009-0.004)	1,52E-08	0.018(0.022-0.014)	8,32E-19
*cg25872752	19	920370	KISS1R	Body	island	0.018(0.024-0.012)	1,53E-08	0.011(0.013-0.008)	1,17E-13
*cg05064002	19	17858612	FCHO1	5'UTR	island	0.011(0.015-0.007)	1,56E-08	0.003(0.005-0)	0,0834
*cg10648908	22	43253189	ARFGAP3	1stExon	island	0.015(0.02-0.01)	1,58E-08	0.007(0.009-0.006)	5,41E-21
*cg10101470	1	90286633	LRRC8D	TSS1500	island	0.008(0.011-0.006)	1,63E-08	0.009(0.011-0.006)	4,23E-12
*cg20350269	16	23160019	USP31	1stExon	island	0.011(0.015-0.008)	1,63E-08	0.016(0.019-0.013)	4,59E-21
*cg27425263	15	101792104	CHSY1	1stExon	island	0.016(0.022-0.011)	1,64E-08	0.017(0.021-0.013)	4,92E-16
*cg25820220	8	107282481	OXR1	5'UTR	island	0.016(0.021-0.011)	1,65E-08	0.024(0.027-0.02)	7,82E-35
*cg24594678	15	63796908	USP3	5'UTR	island	0.013(0.017-0.008)	1,69E-08	0.015(0.019-0.011)	2,18E-14
*cg08684830	9	80851119	CEP78	1stExon	island	0.014(0.019-0.009)	1,70E-08	0.009(0.011-0.007)	2,93E-15
*cg24851862	1	228674966	RNF187	TSS200	island	0.01(0.014-0.007)	1,78E-08	0.007(0.009-0.004) -0.003(-0.006--0.001)	1,17E-07 0,00258
*cg07986525	7	26192199	NFE2L3	1stExon	island	0.013(0.017-0.008)	1,79E-08	0.009(0.013-0.006)	2,06E-09
*cg02745632	1	26146759	FAM54B	5'UTR	island	0.008(0.011-0.005)	1,81E-08	0.011(0.015-0.008)	2,91E-12
*cg23572376	6	1610120	FOXC1	TSS1500	island	0.016(0.022-0.011)	1,90E-08	0.01(0.012-0.007) -0.018(-0.023--0.014)	7,70E-11 1,50E-16
*cg24881159	15	31284219	MTMR10	TSS1500	island	0.013(0.017-0.008)	1,91E-08	0.014(0.017-0.011)	1,76E-16
*cg17364044	2	64371202	PELI1	5'UTR	island	0.014(0.019-0.009)	1,93E-08	0.008(0.011-0.006)	1,76E-08
*cg02996314	1	32479425	KHDRBS1	TSS200	island	0.011(0.015-0.007)	1,96E-08		
*cg03648361	17	73893010	TRIM65	1stExon	island	0.014(0.018-0.009)	1,96E-08		
*cg08521010	3	179370821	USP13	TSS200	island	0.012(0.015-0.008)	1,97E-08		

*cg17429424	5	172197911	DUSP1	1stExon	island	0.011(0.014-0.007)	2,01E-08	0.004(0.006-0.002)	0,000642
*cg24817430	15	81294299	MESDC1	1stExon	island	0.023(0.031-0.015)	2,09E-08	0.024(0.029-0.02)	4,11E-22
*cg25179963	6	116422123	NT5DC1	1stExon	island	0.008(0.011-0.006)	2,13E-08	NA	NA
*cg04573959	8	12612631	LONRF1	1stExon	island	0.017(0.023-0.011)	2,14E-08	0.022(0.025-0.018)	1,62E-29
*cg08083016	6	144471643	STX11	TSS200	island	0.018(0.024-0.012)	2,14E-08	0.031(0.034-0.027)	2,13E-54
*cg26986180	12	124246525	DNAH10	TSS1500	island	0.011(0.015-0.007)	2,16E-08	0.006(0.008-0.003)	2,30E-06
*cg07202430	12	111843939	SH2B3	1stExon	island	0.023(0.03-0.015)	2,23E-08	0.008(0.012-0.004)	2,44E-05
*cg17930183	2	36583117	CRIM1	TSS1500	island	0.028(0.038-0.019)	2,26E-08	0.015(0.019-0.011)	1,67E-11
*cg06648971	3	125239030	SNX4	5'UTR	island	0.005(0.007-0.003)	2,28E-08	0.004(0.005-0.003)	2,20E-10
*cg21442773	1	236029828	LYST	5'UTR	island	0.012(0.017-0.008)	2,30E-08	0.005(0.009-0.002)	0,00304
*cg20488697	17	79885644	LOC92659	TSS200	island	0.014(0.018-0.009)	2,35E-08	0.009(0.012-0.005)	7,07E-07
*cg11343870	11	67888858	CHKA	TSS200	island	0.02(0.027-0.013)	2,38E-08	0.03(0.035-0.026)	2,58E-33
*cg13731828	9	139780699	TRAF2	TSS1500	island	0.007(0.01-0.005)	2,44E-08	0.005(0.007-0.004)	1,80E-09
*cg12129103	5	134369903	PITX1	1stExon	island	0.017(0.023-0.011)	2,47E-08	0.019(0.023-0.015)	4,58E-19
*cg16043190	4	78741143	CNOT6L	TSS1500	island	0.015(0.02-0.01)	2,49E-08	0.019(0.022-0.015)	1,23E-25
*cg05052194	1	2160249	SKI	1stExon	island	0.021(0.028-0.014)	2,52E-08	0.013(0.02-0.006)	0,000227
*cg08393828	14	102228221	PPP2R5C	1stExon	island	0.013(0.018-0.009)	2,55E-08	0.026(0.029-0.022)	9,73E-37
*cg19656070	17	3571978	TMEM93	TSS200	island	0.023(0.031-0.015)	2,55E-08	0.01(0.015-0.005)	2,45E-05
*cg26413174	1	55266737	TTC22	1stExon	island	0.018(0.024-0.012)	2,55E-08	0.002(0.005--0.002)	0,333
*cg10729531	20	61492669	TCFL5	1stExon	island	0.014(0.018-0.009)	2,57E-08	0.01(0.013-0.008)	9,50E-14
*cg01698934	16	88729639	MGC23284	TSS200	island	0.005(0.006-0.003)	2,60E-08	0.003(0.004-0.002)	1,91E-09
*cg24657817	4	42153708	BEND4	Body	island	0.009(0.012-0.006)	2,62E-08	0.005(0.006-0.003)	1,58E-07
*cg03183540	8	134931756		IGR	opensea	0.027(0.036-0.018)	2,75E-08	0.017(0.021-0.013)	4,41E-16
*cg01305736	17	40831891	CCR10	Body	island	0.016(0.021-0.01)	2,79E-08	0.029(0.033-0.025)	1,28E-37
*cg09951048	8	21966736	NUDT18	1stExon	island	0.007(0.009-0.004)	2,80E-08	0.008(0.01-0.007)	5,89E-18
*cg07409674	19	18682889	UBA52	5'UTR	island	0.015(0.021-0.01)	2,84E-08	0.007(0.011-0.003)	0,000439
*cg11258043	6	73972820	KHDC1	1stExon	island	0.005(0.007-0.004)	2,84E-08	NA	NA
*cg26087678	15	60884119	RORA	Body	island	0.018(0.025-0.012)	2,89E-08	0.012(0.016-0.009)	1,16E-11
*cg15854067	3	57199318	IL17RD	1stExon	island	0.015(0.02-0.01)	2,92E-08	0.017(0.02-0.014)	1,87E-23
*cg01059126	1	212459029	PPP2R5A	1stExon	island	0.014(0.019-0.009)	2,98E-08	0.014(0.017-0.011)	1,52E-19
*cg10484485	4	154387599	KIAA0922	1stExon	island	0.015(0.021-0.01)	3,01E-08	0.006(0.009-0.003)	0,000376
								-0.011(-0.014--	
*cg22663117	17	79670454	MRPL12	5'UTR	island	0.007(0.01-0.005)	3,02E-08	0.009)	1,33E-16
*cg13309628	20	3801056	C20orf29	TSS200	island	0.013(0.017-0.008)	3,07E-08	0.005(0.008-0.001)	0,00554
*cg01372366	11	48041783	PTPRJ	Body	opensea	0.027(0.036-0.018)	3,14E-08	0.02(0.024-0.016)	4,25E-20
*cg03628682	5	72921796	RGNEF	TSS200	island	0.01(0.014-0.007)	3,15E-08	0.011(0.014-0.009)	2,45E-20
*cg02318426	14	35873683	NFKBIA	1stExon	island	0.011(0.014-0.007)	3,22E-08	0.012(0.015-0.009)	3,65E-15
								-0.019(-0.025--	
*cg05043032	1	95583073	TMEM56	1stExon	island	0.011(0.015-0.008)	3,27E-08	0.012)	6,28E-08
*cg06642941	16	48643813	N4BP1	1stExon	island	0.009(0.012-0.006)	3,36E-08	0.019(0.021-0.016)	1,91E-40
*cg02625641	15	59041175	ADAM10	Body	island	0.023(0.031-0.015)	3,45E-08	0.021(0.029-0.012)	4,00E-06
*cg21544377	12	62997103	MIRLET7I	TSS1500	island	0.011(0.015-0.007)	3,46E-08	-0.001(-0.003-0.001)	0,464
*cg11882432	1	101005509	GPR88	Body	island	0.015(0.02-0.01)	3,48E-08	0.016(0.019-0.012)	4,50E-19
*cg24778909	17	54670875	NOG	TSS200	island	0.018(0.024-0.012)	3,51E-08	0.023(0.027-0.019)	3,97E-29
*cg06959182	5	115910467	SEMA6A	1stExon	island	0.006(0.008-0.004)	3,61E-08	0.008(0.009-0.007)	3,28E-49
								-0.015(-0.018--	
*cg15625219	4	56814934	CEP135	TSS200	island	0.008(0.011-0.005)	3,61E-08	0.012)	7,29E-24
*cg08573313	7	94139369	CASD1	5'UTR	island	0.017(0.023-0.011)	3,77E-08	0.008(0.012-0.005)	1,57E-05

*cg00449899	8	53626436	RB1CC1	5'UTR	island	0.031(0.041-0.02)	3,81E-08	0.005(0.01-0)	0,0558
*cg18770186	8	141521259	CHRA1	TSS200	island	0.019(0.026-0.013)	3,81E-08	0.005(0.009-0.002)	0,00447
*cg01141237	2	198175400		IGR	island	0.018(0.025-0.012)	3,87E-08	0.014(0.018-0.01)	5,45E-12
*cg26283532	4	119757274	SEC24D	1stExon	island	0.011(0.014-0.007)	3,88E-08	0.014(0.017-0.011)	1,36E-17
*cg24721777	17	79885593	LOC92659	TSS200	island	0.015(0.02-0.01)	3,99E-08	0.01(0.014-0.007)	1,76E-10
*cg01631325	18	60191404	ZCCHC2	1stExon	island	0.019(0.025-0.012)	4,20E-08	0.024(0.028-0.02)	4,25E-30
*cg03449631	4	1714019	SLBP	5'UTR	island	0.018(0.024-0.012)	4,20E-08	0.035(0.041-0.03)	7,19E-36
*cg02154956	13	100741410	PCCA	1stExon	island	0.006(0.008-0.004)	4,23E-08	0.005(0.007-0.004)	2,40E-10
*cg04147049	1	33219687	KIAA1522	Body	island	0.011(0.015-0.007)	4,36E-08	0.01(0.013-0.008)	4,32E-14
*cg24662718	1	108507468	VAV3	1stExon	island	0.011(0.014-0.007)	4,37E-08	0.016(0.019-0.013)	1,08E-27
*cg00516597	3	25824837	NGLY1	1stExon	island	0.012(0.016-0.008)	4,39E-08	0.006(0.009-0.003)	0,000246
*cg09597465	12	129308176	SLC15A4	1stExon	island	0.013(0.018-0.009)	4,39E-08	0.009(0.012-0.006)	1,19E-10
*cg14642298	9	89561881	GAS1	5'UTR	island	0.016(0.022-0.011)	4,42E-08	0.014(0.017-0.01)	4,33E-13
*cg21305245	20	32891026	AHCY	5'UTR	island	0.013(0.017-0.008)	4,43E-08	0.011(0.013-0.008)	1,08E-13
*cg06499637	10	105881701	C10orf78	TSS1500	island	0.012(0.016-0.008)	4,53E-08	0.007(0.01-0.005)	1,68E-07
*cg10640371	21	35446012	SLC5A3	1stExon	island	0.012(0.016-0.008)	4,57E-08	0.009(0.011-0.007)	4,89E-14
*cg23064465	4	183838739	DCTD	TSS200	island	0.016(0.022-0.011)	4,59E-08	0.019(0.022-0.015)	1,20E-22
*cg15407517	5	78280882	ARSB	1stExon	island	0.003(0.005-0.002)	4,63E-08	0.006(0.007-0.005)	1,15E-30
*cg13251669	20	60718270	PSMA7	1stExon	island	0.007(0.009-0.004)	4,79E-08	0.013(0.015-0.012)	8,02E-47
*cg07806164	3	23244688	UBE2E2	TSS200	island	0.016(0.022-0.01)	4,80E-08	0.031(0.035-0.027)	3,33E-44
*cg23046629	2	109745841	LOC100287216	Body	island	0.02(0.027-0.013)	4,84E-08	0.002(0.005--0.001)	0,182
*cg17547708	22	51021506	LOC100144603	Body	island	0.015(0.02-0.01)	4,87E-08	0.014(0.017-0.011)	8,72E-18
*cg06849077	5	176245016	UNC5A	Body	island	0.013(0.018-0.009)	4,97E-08	0.01(0.013-0.007)	1,63E-12
*cg22552684	7	42951784	C7orf25	TSS200	island	0.011(0.015-0.007)	4,98E-08	0.009(0.012-0.006) -0.009(-0.014--	3,55E-09
*cg07032309	8	103668104	KLF10	TSS200	island	0.026(0.035-0.017)	5,01E-08	0.004)	0,000254
*cg00636641	9	89561390	GAS1	1stExon	island	0.014(0.019-0.009)	5,08E-08	0.021(0.024-0.018)	8,97E-40
*cg00533923	6	116892543	RWDD1	TSS200	island	0.015(0.021-0.01)	5,14E-08	0.008(0.012-0.004)	9,66E-05
*cg05965902	13	21635727	LATS2	TSS200	island	0.012(0.016-0.008)	5,17E-08	0.02(0.023-0.017)	7,83E-35
*cg01741878	17	79269038	SLC38A10	5'UTR	island	0.013(0.017-0.008)	5,19E-08	0.004(0.007-0.001)	0,0203
*cg14053749	4	122872836	TRPC3	1stExon	island	0.012(0.017-0.008)	5,19E-08	0.012(0.015-0.009)	7,38E-14
*cg03890664	19	10613492	KEAP1	5'UTR	island	0.014(0.018-0.009)	5,23E-08	0.005(0.008-0.002)	0,000271
*cg01490821	2	170684022	UBR3	1stExon	island	0.012(0.017-0.008)	5,26E-08	0.012(0.015-0.008)	4,05E-10
*cg08579070	13	76210054	LMO7	Body	island	0.016(0.022-0.011)	5,29E-08	0.015(0.018-0.011)	2,64E-16
*cg01731839	16	19729756	C16orf88	TSS1500	island	0.017(0.022-0.011)	5,30E-08	0.001(0.006--0.003)	0,576
*cg18507129	3	142839903	CHST2	Body	island	0.011(0.015-0.007)	5,32E-08	0.005(0.007-0.003)	1,81E-06
*cg02756545	1	94884060	ABCD3	1stExon	island	0.013(0.018-0.009)	5,34E-08	0.015(0.019-0.011)	2,52E-13
*cg12996305	17	2295923	MNT	Body	island	0.018(0.025-0.012)	5,42E-08	0.009(0.014-0.004)	0,000176
*cg01134297	16	90014983	DEF8	TSS200	shore	0.012(0.016-0.008)	5,45E-08	0.008(0.011-0.006)	2,88E-09
*cg26864061	3	45730815	SACM1L	1stExon	island	0.014(0.019-0.009)	5,57E-08	0.008(0.011-0.005)	2,95E-08
*cg11769606	5	78281103	ARSB	1stExon	island	0.014(0.018-0.009)	5,58E-08	0.011(0.014-0.008)	4,52E-13
*cg22093503	11	43964098	LOC387763	TSS200	island	0.009(0.013-0.006)	5,60E-08	0.008(0.01-0.006)	5,44E-15
*cg09490387	16	2034188	GFER	1stExon	island	0.018(0.025-0.012)	5,62E-08	-0.001(-0.006-0.004)	0,839
*cg26856583	16	87425887	MAP1LC3B	1stExon	island	0.017(0.023-0.011)	5,64E-08	0.004(0.008-0.001)	0,0126
*cg13493749	9	96214374	FAM120A	1stExon	island	0.008(0.011-0.005)	5,65E-08	0.001(0.002--0.001)	0,359
*cg24341177	12	122710506	DIABLO	Body	island	0.013(0.017-0.008)	5,73E-08	0.003(0.006-0.001)	0,0198
*cg08767710	7	149570989	ATP6V0E2	5'UTR	island	0.017(0.022-0.011)	5,89E-08	0.018(0.021-0.015)	1,18E-25

*cg04986015	1	171810433	DNM3	TSS200	shore	0.012(0.017-0.008)	5,93E-08	0.008(0.011-0.006)	6,09E-09
*cg07944287	19	6110462	RFX2	5'UTR	island	0.013(0.017-0.008)	5,99E-08	0.022(0.025-0.019)	1,58E-38
*cg24463309	12	112036857	ATXN2	1stExon	island	0.015(0.02-0.01)	6,07E-08	0.017(0.02-0.014)	3,93E-26
*cg13134514	1	202780014		IGR	island	0.006(0.009-0.004)	6,17E-08	-0.01(-0.012--0.008)	1,57E-18
*cg09011608	12	110152133	C12orf34	TSS200	island	0.01(0.013-0.006)	6,19E-08	0.004(0.006-0.001)	0,00258
*cg21172814	1	180199001	LHX4	TSS1500	island	0.009(0.012-0.006)	6,25E-08	0.005(0.007-0.002)	0,000146
*cg09874752	10	99531309	SFRP5	1stExon	island	0.018(0.024-0.012)	6,37E-08	0.011(0.014-0.007)	2,83E-08
*cg01249550	3	194353869	TMEM44	1stExon	island	0.014(0.019-0.009)	6,54E-08	0.016(0.019-0.013)	1,03E-22
*cg11785574	1	241803680	OPN3	5'UTR	shore	0.013(0.018-0.009)	6,59E-08	0.01(0.013-0.007)	1,99E-13
*cg20652640	16	88520056	ZFPM1	1stExon	island	0.015(0.02-0.01)	6,61E-08	0.005(0.009-0.002)	0,000719
*cg05884522	6	13615538	NOL7	TSS200	island	0.014(0.019-0.009)	6,70E-08	0.005(0.008-0.002)	0,00209
*cg08300899	9	140064540	LRRC26	TSS200	island	0.026(0.035-0.017)	6,70E-08	0.024(0.03-0.019)	2,21E-16
*cg09780955	10	97321190	SORBS1	TSS200	island	0.01(0.013-0.006)	6,73E-08	0.007(0.009-0.004)	2,23E-08
*cg21316470	15	63340581	TPM1	Body	island	0.025(0.034-0.016)	6,90E-08	0.016(0.02-0.013)	1,61E-20
*cg11156286	12	133338210	ANKLE2	1stExon	island	0.014(0.019-0.009)	6,93E-08	0.01(0.013-0.006)	7,66E-08
*cg26948660	16	14165300	MKL2	5'UTR	island	0.01(0.014-0.007)	6,94E-08	0.011(0.013-0.009)	3,43E-22
*cg16638571	8	81083534	TPD52	Body	island	0.006(0.008-0.004)	7,19E-08	0.007(0.008-0.006)	3,09E-21
*cg09211893	14	100259860	EML1	1stExon	shore	0.027(0.037-0.018)	7,29E-08	0.035(0.039-0.03)	2,07E-48
*cg18261601	3	100053659	NIT2	Body	island	0.011(0.015-0.007)	7,34E-08	0.007(0.01-0.005)	6,49E-08
*cg23183469	4	1686288	FAM53A	TSS1500	island	0.019(0.026-0.012)	7,35E-08	0.05(0.055-0.045)	4,81E-72
*cg12280407	13	77460322	KCTD12	5'UTR	island	0.011(0.015-0.007)	7,36E-08	0.012(0.015-0.01)	3,46E-20
*cg04212568	12	123874243	SETD8	Body	island	0.017(0.023-0.011)	7,50E-08	0.023(0.027-0.02)	3,45E-33
*cg24592364	5	14871736	ANKH	1stExon	shore	0.011(0.015-0.007)	7,54E-08	0.011(0.013-0.009)	1,19E-20
*cg21730067	1	90460918	LOC492303	TSS1500	island	0.008(0.011-0.005)	7,55E-08	0.005(0.007-0.003)	5,34E-07
*cg19262019	19	8510039	HNRNPM	Body	island	0.014(0.019-0.009)	7,64E-08	0.011(0.014-0.007)	9,11E-09
*cg08675654	10	112403955	RBM20	TSS200	island	0.017(0.023-0.011)	7,73E-08	0.02(0.025-0.016)	1,88E-19
*cg01634153	1	39571505	MACF1	Body	island	0.015(0.02-0.009)	7,75E-08	0.018(0.021-0.015)	2,30E-31
*cg01575096	2	238969261	SCLY	TSS1500	island	0.012(0.016-0.008)	7,79E-08	0.007(0.009-0.004)	2,56E-06
*cg01338658	1	182992728	LAMC1	1stExon	island	0.017(0.022-0.011)	7,83E-08	0.023(0.026-0.019)	2,96E-35
*cg22109056	1	9352963	SPSB1	5'UTR	opensea	0.007(0.009-0.005)	7,85E-08	0.005(0.007-0.003)	8,12E-08
*cg03756640	12	102091773	CHPT1	1stExon	island	0.012(0.016-0.008)	7,92E-08	0.011(0.014-0.008)	1,59E-09
*cg21211144	17	16118919	NCOR1	TSS200	island	0.017(0.023-0.011)	7,93E-08	0.015(0.019-0.012)	2,43E-16
*cg06707168	1	231114534	ARV1	TSS1500	island	0.009(0.012-0.006)	8,09E-08	0.01(0.013-0.007)	2,00E-09
*cg14221884	12	63328524	PPM1H	5'UTR	island	0.02(0.027-0.013)	8,09E-08	0.012(0.016-0.008)	5,56E-09
*cg08183916	2	96874535	STARD7	1stExon	island	0.028(0.038-0.018)	8,26E-08	0.008(0.012-0.003)	0,00148
*cg16501572	11	71934660	INPPL1	TSS1500	island	0.009(0.013-0.006)	8,40E-08	0.002(0.004-0)	0,128
*cg22194305	4	2264032	MXD4	TSS1500	island	0.012(0.017-0.008)	8,42E-08	0.013(0.017-0.01)	1,71E-11
*cg24441872	4	53525715	USP46	TSS1500	island	0.016(0.022-0.01)	8,56E-08	0.007(0.011-0.003)	0,000664
*cg11586519	1	1509833	SSU72	Body	island	0.015(0.021-0.01)	8,62E-08	0.013(0.016-0.01)	5,24E-20
*cg06363275	1	86042667	DDAH1	5'UTR	island	0.008(0.011-0.005)	8,68E-08	0.006(0.008-0.004)	3,12E-10
*cg09338033	13	73633700	KLF5	1stExon	island	0.011(0.015-0.007)	8,78E-08	0.013(0.015-0.01)	1,30E-17
*cg16590515	5	79551618	SERINC5	Body	island	0.013(0.018-0.008)	8,78E-08	0.019(0.023-0.016)	1,92E-31
*cg03303757	1	11120145	SRM	TSS200	island	0.003(0.004-0.002)	8,88E-08	0.006(0.008-0.005)	3,24E-17
*cg20247600	2	178129336	NFE2L2	5'UTR	island	0.019(0.025-0.012)	8,89E-08	0.022(0.025-0.018)	8,63E-28
*cg21076271	14	32546633	ARHGAP5	1stExon	island	0.016(0.022-0.01)	8,89E-08	0.011(0.015-0.008)	2,03E-09
*cg01111245	5	36241913	C5orf33	5'UTR	island	0.009(0.013-0.006)	8,94E-08	0.009(0.012-0.007)	8,55E-15

*cg09597785	3	126194459	ZXDC	1stExon	island	0.004(0.005-0.002)	8,95E-08	0.009(0.01-0.007)	4,66E-30
*cg02484274	4	122618147	ANXA5	TSS200	island	0.013(0.018-0.009)	9,10E-08	0.016(0.02-0.013)	9,06E-22
*cg15586420	15	90294517	MESP1	1stExon	island	0.016(0.022-0.01)	9,35E-08	0.028(0.035-0.022)	1,29E-16
*cg15993110	8	17014337	ZDHH2	1stExon	island	0.013(0.017-0.008)	9,43E-08	0.003(0.007-0)	0,0735
*cg06812033	16	50582453	NKD1	5'UTR	island	0.011(0.014-0.007)	9,48E-08	0.01(0.012-0.007)	2,30E-11
*cg06355010	20	13202489	ISM1	1stExon	island	0.017(0.023-0.011)	9,54E-08	0.011(0.015-0.008)	5,57E-11
*cg25589651	6	28891121	TRIM27	1stExon	island	0.021(0.029-0.014)	9,55E-08	0.003(0.007--0.001)	0,17
*cg12108029	8	126010861	SQLE	1stExon	island	0.019(0.025-0.012)	9,61E-08	0.025(0.029-0.021)	3,62E-33
*cg02011918	20	3219539	SLC4A11	TSS1500	island	0.015(0.02-0.009)	9,68E-08	0.014(0.018-0.011)	1,26E-17
*cg07189157	7	149119407		IGR	island	0.011(0.015-0.007)	9,71E-08	0.009(0.011-0.007)	1,09E-13
*cg09586534	2	172290954	DCAF17	5'UTR	island	0.016(0.021-0.01)	9,77E-08	0.028(0.032-0.023)	5,12E-32
*cg13009662	5	127419527	SLC12A2	1stExon	island	0.025(0.034-0.016)	9,79E-08	0.018(0.021-0.015)	4,06E-27
*cg03913271	22	38302504	MICALL1	1stExon	island	0.012(0.016-0.008)	9,80E-08	0.004(0.007-0.001)	0,00588
*cg16466635	7	66093826	KCTD7	TSS200	island	0.015(0.02-0.009)	9,81E-08	0.01(0.013-0.007) -0.004(-0.007-- 0.002)	5,49E-10 0,000818
*cg16370826	16	2040023	SYNGR3	5'UTR	shore	0.008(0.011-0.005)	9,83E-08	0.011(0.015-0.006)	2,01E-06
*cg03327386	18	8609878	RAB12	1stExon	island	0.019(0.026-0.012)	9,84E-08	0.011(0.014-0.009)	2,76E-16
*cg26705583	5	65017946	SGTB	TSS200	island	0.016(0.021-0.01)	9,85E-08	0.006(0.01-0.002)	0,00166
*cg27413118	2	238969595	SCLY	TSS200	island	0.017(0.023-0.011)	9,85E-08	0.026(0.03-0.022)	1,06E-35
*cg09464488	2	175547590	WIPF1	1stExon	island	0.023(0.032-0.015)	9,94E-08	0.014(0.019-0.01)	2,59E-09
*cg05787952	16	734667	JMJD8	5'UTR	opensea	0.021(0.029-0.014)	1,01E-07	0.021(0.025-0.017)	1,25E-23
*cg02600940	2	205410387	PARD3B	TSS200	island	0.012(0.016-0.008)	1,02E-07	0.013(0.016-0.01)	9,60E-22
*cg15359501	1	184006496	GLT25D2	5'UTR	island	0.005(0.007-0.004)	1,02E-07	0.007(0.009-0.005)	3,14E-12
*cg24147485	1	234735701		IGR	island	0.023(0.031-0.015)	1,03E-07	0.006(0.01-0.001)	0,0152
*cg15459695	2	236403016	AGAP1	1stExon	island	0.024(0.032-0.015)	1,04E-07	0.019(0.027-0.011)	2,04E-06
*cg03167369	1	35658023	SFPQ	1stExon	island	0.015(0.02-0.009)	1,05E-07	0.009(0.012-0.006)	1,83E-08
*cg04165030	5	175843293	CLTB	1stExon	island	0.015(0.021-0.01)	1,05E-07	0.026(0.032-0.02)	5,37E-13
*cg00473112	11	507177	RNH1	1stExon	island	0.007(0.01-0.005)	1,07E-07	0.003(0.005-0.002)	0,000492
*cg14045725	17	2240967	SGSM2	Body	island	0.012(0.016-0.008)	1,07E-07	0.006(0.008-0.003)	1,31E-05
*cg07677296	16	1877182	HAGH	5'UTR	island	0.014(0.019-0.009)	1,09E-07	0.003(0.01--0.004)	0,445
*cg11567041	10	12391824	CAMK1D	TSS200	island	0.011(0.015-0.007)	1,11E-07	0.001(0.004--0.001)	0,282
*cg06950567	20	48807318	CEBPB	TSS200	island	0.012(0.017-0.008)	1,11E-07	0.012(0.014-0.009)	2,36E-21
*cg07249227	21	43430503	ZNF295	Body	island	0.016(0.021-0.01)	1,12E-07	0.037(0.042-0.032)	1,29E-44
*cg22827465	19	3094947	GNA11	TSS200	island	0.015(0.02-0.009)	1,13E-07	0.022(0.026-0.019) -0.009(-0.011-- 0.007)	8,11E-30 1,23E-13
*cg17284082	1	92351487	TGFBR3	Body	opensea	0.006(0.008-0.004) -0.017(-0.011-- 0.024)	1,14E-07	-0.016(-0.021-- 0.011)	7,75E-10
*cg17820878	1	27440463	SLC9A1	1stExon	island	0.012(0.016-0.008)	1,15E-07	0.024(0.028-0.021)	2,68E-38
*cg00007226	14	105781366	PACS2	TSS1500	island	0.017(0.023-0.011)	1,16E-07	0.018(0.022-0.014)	7,13E-16
*cg09901735	16	19729769	C16orf88	TSS1500	island	0.013(0.018-0.009)	1,18E-07	0.01(0.014-0.006)	3,02E-06
*cg21503377	1	228871665	RHOU	1stExon	island	0.02(0.027-0.013)	1,19E-07	0.02(0.024-0.016)	4,56E-20
*cg18154422	10	31609200	ZEB1	TSS200	island	0.009(0.013-0.006)	1,20E-07	0.015(0.018-0.013)	7,83E-27
*cg07061867	3	56716962	BMP8A	1stExon	island	0.019(0.026-0.012)	1,20E-07	0.009(0.014-0.004)	0,000642
*cg08748615	1	39957298	PPP1R9B						

*cg12927203	7	56032122	GBAS	TSS200	island	0.016(0.022-0.01)	1,21E-07	0.004(0.007-0.001)	0,0164
*cg19922442	1	8086588	ERRFI1	TSS200	island	0.019(0.026-0.012)	1,22E-07	0.031(0.036-0.026)	6,87E-33
*cg00119117	20	20349621	INSM1	1stExon	island	0.017(0.023-0.011)	1,23E-07	0.015(0.019-0.012)	1,28E-17
*cg12133425	1	247494926	ZNF496	1stExon	island	0.012(0.016-0.008)	1,24E-07	0.01(0.013-0.008)	1,47E-17
*cg12585429	16	49889252		IGR	island	0.019(0.025-0.012)	1,24E-07	0.01(0.014-0.005)	1,77E-05
*cg00373665	1	234735703		IGR	island	0.013(0.017-0.008)	1,26E-07	0.014(0.018-0.01)	5,81E-10
*cg04424385	7	129074328	FAM40B	1stExon	island	0.018(0.025-0.012)	1,26E-07	0.018(0.022-0.014)	1,72E-18
*cg23539933	2	88752119	FOXI3	TSS200	island	0.02(0.027-0.013)	1,26E-07	0.005(0.008-0.002)	0,00322
*cg06590470	9	34523253	ENHO	TSS1500	island	0.012(0.016-0.008)	1,27E-07	0.008(0.011-0.005)	6,56E-07
*cg27594583	7	150716029	ATG9B	Body	island	0.012(0.016-0.008)	1,27E-07	0.009(0.012-0.006)	1,56E-08
*cg04052100	8	30585443	GSR	TSS200	island	0.011(0.015-0.007)	1,28E-07	0.004(0.007-0.002)	0,00163
*cg24608684	6	130686865		IGR	island	0.019(0.026-0.012)	1,28E-07	0.01(0.016-0.005)	0,000608
*cg06023901	6	57037482	BAG2	5'UTR	island	0.011(0.015-0.007)	1,29E-07	-0.008(-0.01--0.006)	5,32E-20
*cg17665931	19	38894173	FAM98C	Body	island	0.014(0.02-0.009)	1,29E-07	-0.005(-0.011-0.001)	0,0782
*cg24255125	11	120856918	GRIK4	Body	island	0.037(0.05-0.023)	1,31E-07	0.01(0.017-0.003)	0,00662
*cg22718540	1	39874874	KIAA0754	TSS1500	island	0.015(0.02-0.009)	1,34E-07	0.008(0.011-0.005)	4,91E-08
*cg23513504	21	45432403	TRAPPC10	1stExon	island	0.014(0.019-0.009)	1,34E-07	0.014(0.018-0.01)	3,24E-12
*cg24497897	11	65658826	CCDC85B	1stExon	island	0.019(0.026-0.012)	1,35E-07	0.011(0.016-0.005)	8,95E-05
*cg05744073	17	1953268	MIR132	Body	island	0.014(0.019-0.009)	1,37E-07	0.018(0.023-0.013)	8,45E-11
*cg11837377	1	230778283	COG2	1stExon	island	0.005(0.007-0.003)	1,39E-07	0.004(0.006-0.003)	3,73E-07
*cg14279702	18	30050859	FAM59A	TSS1500	island	0.012(0.017-0.008)	1,39E-07	0.011(0.014-0.007)	5,07E-09
*cg25096582	14	35183928	CFL2	5'UTR	island	0.009(0.013-0.006)	1,39E-07	-0.069(-0.077--0.06)	1,15E-52
*cg04028315	7	21467710	SP4	5'UTR	island	0.006(0.008-0.004)	1,40E-07	0.003(0.004-0.002)	2,51E-07
*cg16361303	16	9186057	C16orf72	1stExon	island	0.012(0.016-0.007)	1,42E-07	0.009(0.011-0.006)	6,61E-09
*cg02568911	1	6453807	ACOT7	1stExon	island	0.014(0.019-0.009)	1,43E-07	0.006(0.009-0.003)	0,000314
*cg06380157	1	39875049	KIAA0754	TSS200	island	0.018(0.024-0.011)	1,43E-07	0.018(0.021-0.015)	1,82E-27
*cg26363167	8	42010494	AP3M2	1stExon	island	0.007(0.01-0.005)	1,43E-07	0.005(0.007-0.004)	8,95E-17
*cg11719836	1	20209026	OTUD3	1stExon	island	0.005(0.007-0.003)	1,44E-07	0.01(0.012-0.008)	1,04E-25
*cg19080138	7	35734733	HERPUD2	1stExon	island	0.021(0.028-0.013)	1,45E-07	0.009(0.014-0.004)	0,000177
*cg09556952	21	38119946	SIM2	Body	island	0.015(0.021-0.01)	1,48E-07	0.014(0.018-0.011)	1,83E-14
*cg25141720	9	89561826	GAS1	5'UTR	island	0.013(0.018-0.009)	1,48E-07	0.015(0.018-0.012)	2,27E-17
*cg03766306	1	90461091	LOC492303	TSS1500	island	0.011(0.015-0.007)	1,50E-07	0.01(0.013-0.007)	2,41E-09
*cg04994953	6	149638666	MAP3K7IP2	TSS1500	island	0.015(0.021-0.01)	1,50E-07	0.005(0.008-0.002)	0,001
*cg05243095	18	3262534	MYL12B	5'UTR	island	0.022(0.029-0.014)	1,50E-07	0.013(0.017-0.009)	4,10E-10
*cg04220635	1	108743085	SLC25A24	TSS200	island	0.011(0.015-0.007)	1,51E-07	0.021(0.024-0.018)	3,67E-44
*cg05018441	1	32817390		IGR	island	0.011(0.015-0.007)	1,54E-07	0.011(0.015-0.008)	2,61E-10
*cg13169065	2	166650947	GALNT3	TSS200	island	0.015(0.02-0.009)	1,54E-07	0.019(0.023-0.015)	2,77E-19
*cg02835038	21	47743854	PCNT	TSS200	island	0.022(0.03-0.014)	1,55E-07	0.008(0.013-0.003)	0,00119
*cg03848555	9	77567361	C9orf40	1stExon	island	0.005(0.007-0.003)	1,55E-07	0.006(0.008-0.004)	1,82E-09
*cg08889551	1	176175820	RFWD2	1stExon	island	0.011(0.015-0.007)	1,60E-07	0.009(0.012-0.006)	1,68E-07
*cg02695219	18	9136852	ANKRD12	1stExon	island	0.012(0.017-0.008)	1,63E-07	-0.002(-0.005-0.001)	0,185
*cg03007034	22	46068330	ATXN10	Body	island	0.013(0.018-0.008)	1,65E-07	0.002(0.005--0.001)	0,118
*cg25312440	1	76547234	ST6GALNAC3	Body	opensea	0.035(0.048-0.023)	1,65E-07	0.011(0.019-0.003)	0,00476
*cg22402852	4	108853204	CYP2U1	1stExon	island	0.007(0.01-0.004)	1,67E-07	0.005(0.007-0.004)	3,82E-16
*cg21099488	14	56585486	PELI2	Body	island	0.011(0.015-0.007)	1,68E-07	0(-0.004-0.004)	0,978
*cg20526663	16	22385655	CDR2	5'UTR	island	0.011(0.015-0.007)	1,69E-07	0.006(0.008-0.003)	1,11E-05

*cg21160472	1	212782112	ATF3	1stExon	island	0.03(0.041-0.019)	1,69E-07	0.019(0.026-0.011)	1,04E-06
*cg10735361	4	7044595	CCDC96	1stExon	island	0.011(0.016-0.007)	1,70E-07	0.009(0.012-0.006)	2,51E-10
*cg01477633	10	71078714	HK1	Body	island	0.011(0.015-0.007)	1,72E-07	0.005(0.008-0.002)	0,000173
*cg02975259	3	136581033	NCK1	TSS200	island	0.011(0.015-0.007)	1,72E-07	0.008(0.011-0.005)	3,16E-07
*cg15467615	21	47743775	PCNT	TSS1500	island	0.01(0.013-0.006)	1,72E-07	0.006(0.008-0.004)	1,27E-09
*cg08888178	19	344981	MIER2	TSS200	shore	0.021(0.028-0.013)	1,73E-07	0.03(0.036-0.025) -0.006(-0.008-- 0.003)	7,47E-26
*cg09395855	10	120967408	GRK5	5'UTR	island	0.008(0.01-0.005)	1,73E-07		2,53E-06
*cg16268380	15	81294324	MESDC1	1stExon	island	0.019(0.026-0.012)	1,74E-07	0.017(0.025-0.009)	4,75E-05
*cg01297357	12	50451678	ACCN2	5'UTR	island	0.017(0.024-0.011)	1,75E-07	0.014(0.018-0.01)	4,18E-12
*cg02233216	2	30454279	LBH	TSS200	island	0.014(0.019-0.009)	1,75E-07	0.012(0.015-0.01)	1,10E-17
*cg04910183	1	193090988	CDC73	TSS200	island	0.03(0.041-0.019)	1,76E-07	0.015(0.02-0.011)	4,91E-11
*cg18947000	5	179233770	SQSTM1	TSS1500	island	0.01(0.014-0.006)	1,77E-07	0.006(0.009-0.004)	1,57E-06
*cg10235453	16	88497293	ZNF469	Body	island	0.024(0.033-0.015)	1,80E-07	0.069(0.077-0.061)	1,30E-56
*cg01932632	18	12657692	SPIRE1	TSS1500	island	0.01(0.014-0.006)	1,85E-07	0.015(0.017-0.012)	3,50E-32
*cg02308245	16	85833143	COX4NB	1stExon	island	0.011(0.015-0.007)	1,88E-07	-0.02(-0.024--0.016)	1,21E-18
*cg06383048	11	118401878	TMEM25	5'UTR	island	0.01(0.013-0.006)	1,91E-07	0.006(0.008-0.004)	8,04E-09
*cg25660961	1	17764073	RCC2	Body	island	0.011(0.015-0.007)	1,91E-07	-0.003(-0.007-0)	0,0499
*cg27634633	5	133747726	CDKN2AIPNL	TSS200	island	0.011(0.015-0.007)	1,91E-07	0.014(0.017-0.011)	1,71E-20
*cg03437204	19	1401372	GAMT	1stExon	island	0.011(0.015-0.007)	1,93E-07	0.001(0.003--0.001)	0,267
*cg21383720	9	96214543	FAM120AOS	1stExon	island	0.009(0.013-0.006)	1,94E-07	0.005(0.008-0.003)	3,11E-05
*cg23108125	5	176882586	PRR7	Body	island	0.028(0.038-0.018)	1,94E-07	0.007(0.012-0.002)	0,003
*cg08982381	9	80647290	GNAQ	TSS1500	island	0.013(0.018-0.008)	1,95E-07	0.021(0.025-0.017)	4,64E-28
*cg12878260	5	139028156	CXXC5	TSS200	island	0.017(0.023-0.011)	1,96E-07	0.009(0.012-0.006)	1,83E-08
*cg00800773	3	53195164	PRKCD	TSS200	island	0.01(0.014-0.006)	1,97E-07	0.025(0.029-0.02)	6,08E-24
*cg04171308	12	112856542	PTPN11	1stExon	island	0.013(0.017-0.008)	1,97E-07	0.009(0.012-0.006)	5,93E-09
*cg23677570	7	74489617	WBSCR16	5'UTR	island	0.01(0.013-0.006)	1,97E-07	0.003(0.005-0.001)	0,0147
*cg18317494	8	22014496	LGI3	TSS200	island	0.014(0.019-0.009)	1,98E-07	0.008(0.011-0.005)	1,07E-08
*cg22659014	13	30169727	SLC7A1	5'UTR	island	0.013(0.018-0.008)	1,99E-07	0.012(0.015-0.009)	2,03E-14
*cg12588263	13	113862899	PCID2	Body	island	0.014(0.019-0.009)	2,03E-07	0.008(0.012-0.005)	4,50E-06
*cg09424923	12	123635770		IGR	island	0.012(0.016-0.007)	2,06E-07	0.008(0.011-0.005)	3,94E-08
*cg02571907	3	119298511	ADPRH	TSS200	island	0.01(0.013-0.006)	2,11E-07	0.007(0.01-0.005)	2,49E-07
*cg22329233	3	48723293	NCKIPSD	1stExon	island	0.014(0.019-0.009)	2,11E-07	0.024(0.027-0.02)	1,31E-33
*cg01263075	4	89080199	ABCG2	TSS200	island	0.009(0.013-0.006)	2,12E-07	0.005(0.008-0.002)	0,00244
*cg00556087	10	98592104	LCOR	1stExon	island	0.008(0.011-0.005)	2,13E-07	0.007(0.009-0.005)	2,34E-14
*cg06759117	16	53088791	CHD9	TSS200	island	0.009(0.012-0.005)	2,13E-07	0.01(0.013-0.007)	4,19E-10
*cg01014432	7	139478207	TBXAS1	1stExon	island	0.01(0.013-0.006)	2,14E-07	0.001(0.003--0.002)	0,715
*cg23456692	3	5229596	EDEM1	1stExon	island	0.012(0.016-0.008)	2,15E-07	0.008(0.011-0.005)	1,92E-08
*cg07597763	11	59382877	OSBP	1stExon	island	0.013(0.017-0.008)	2,16E-07	0.016(0.019-0.013)	3,28E-22
*cg11725975	4	72052874	SLC4A4	TSS200	island	0.004(0.006-0.003)	2,18E-07	0.009(0.01-0.007)	1,62E-35
*cg23844090	3	44903617	TMEM42	Body	island	0.005(0.007-0.003)	2,18E-07	0.001(0.003--0.001)	0,147
*cg13227851	13	42615145		IGR	island	0.013(0.018-0.008) -0.027(-0.017-- 0.037)	2,19E-07	0.012(0.015-0.009)	1,26E-13
*cg13937905	12	53612551	RARG	Body	shore		2,19E-07	-0.02(-0.026--0.015)	2,92E-12
*cg05639310	4	2420719	ZFYVE28	TSS1500	island	0.013(0.018-0.008)	2,25E-07	0.014(0.017-0.011)	8,82E-23
*cg03351618	2	170551124	KLHL23	1stExon	island	0.01(0.014-0.007)	2,26E-07	0.005(0.008-0.001)	0,0057
*cg10181968	7	139477474	HIPK2	1stExon	island	0.009(0.013-0.006)	2,26E-07	0.015(0.017-0.013)	5,18E-37

*cg19206986	4	42658952	ATP8A1	5'UTR	island	0.019(0.026-0.012)	2,28E-07	0.004(0.007-0)	0,0337
*cg23816049	1	224804081	CNIH3	TSS200	island	0.014(0.02-0.009)	2,28E-07	0.007(0.01-0.004)	2,32E-06
*cg09281539	20	20693126	RALGAPA2	5'UTR	island	0.031(0.043-0.02)	2,33E-07	0.006(0.013-0)	0,056
*cg01190522	6	144606245		IGR	island	0.01(0.014-0.007)	2,35E-07	0.004(0.006-0.001)	0,00417
*cg08270913	1	231114531	ARV1	TSS1500	island	0.009(0.013-0.006)	2,35E-07	0.01(0.013-0.007)	2,52E-12
*cg21226861	10	22605339	COMMD3	1stExon	island	0.01(0.014-0.006)	2,35E-07	0.005(0.007-0.002)	0,000467
*cg14875093	19	38398164	WDR87	TSS1500	island	0.013(0.017-0.008)	2,42E-07	0.01(0.014-0.006)	7,42E-06
*cg10692118	17	43299500	FMNL1	1stExon	island	0.011(0.015-0.007)	2,43E-07	0.009(0.012-0.006)	1,70E-09
*cg08128766	20	5591618	RP5-1022P6.2	1stExon	island	0.004(0.006-0.003)	2,44E-07	0.008(0.009-0.006)	3,33E-27
*cg09218130	12	132413611	PUS1	TSS200	island	0.013(0.018-0.008)	2,44E-07	0.019(0.022-0.016)	6,63E-29
*cg14931122	17	20059106	CYTSB	Body	island	0.012(0.016-0.007)	2,44E-07	0.007(0.01-0.005)	4,42E-07
*cg21991830	8	53626936	RB1CC1	1stExon	island	0.014(0.02-0.009)	2,44E-07	0.015(0.018-0.012)	1,08E-19
*cg05555681	17	36886288	CISD3	TSS1500	island	0.009(0.013-0.006)	2,46E-07	0.009(0.011-0.008)	3,41E-20
*cg11754897	13	28194657	POLR1D	TSS1500	island	0.009(0.012-0.006)	2,48E-07	0.003(0.005-0.001)	0,00302
*cg11819707	17	36507992	SOCS7	TSS200	island	0.012(0.016-0.008)	2,48E-07	0.018(0.021-0.015)	8,34E-30
*cg02811931	7	2598859	IQCE	Body	island	0.015(0.02-0.009)	2,50E-07	0.016(0.019-0.012)	4,43E-19
*cg22253313	4	2011099	WHSC2	TSS200	island	0.016(0.022-0.01)	2,51E-07	0.007(0.01-0.004)	7,13E-05
*cg27583030	4	186064515	SLC25A4	1stExon	island	0.011(0.016-0.007)	2,52E-07	-0.002(-0.006-0.002)	0,295
*cg04738301	12	96184580	NTN4	TSS200	island	0.008(0.011-0.005)	2,53E-07	0.006(0.008-0.005)	3,08E-11
*cg21290042	1	224804053	CNIH3	TSS200	island	0.028(0.038-0.018)	2,55E-07	0.004(0.009--0.001)	0,134
*cg06378494	13	24153531	TNFRSF19	5'UTR	island	0.017(0.023-0.01)	2,58E-07	0.009(0.013-0.006)	1,32E-06
*cg07442657	2	85198334	KCMF1	5'UTR	island	0.014(0.019-0.009)	2,59E-07	0(-0.003-0.003)	0,956
*cg07168160	14	45366564	C14orf28	5'UTR	island	0.008(0.011-0.005)	2,60E-07	0.003(0.005-0.001)	0,00105
*cg01417553	2	9614528	IAH1	TSS200	island	0.014(0.019-0.009)	2,64E-07	0.002(0.004--0.001)	0,197
*cg05422897	2	113403145	SLC20A1	TSS1500	island	0.012(0.016-0.007)	2,65E-07	0.016(0.021-0.011)	1,08E-09
*cg17710545	13	98629067	IPO5	Body	island	0.014(0.019-0.009)	2,66E-07	-0.001(-0.003-0.001)	0,454
*cg20202505	21	47706161	MCM3AP	TSS1500	island	0.011(0.016-0.007)	2,66E-07	0.009(0.012-0.007)	1,08E-14
*cg10496915	6	89855751	PM20D2	TSS200	island	0.011(0.015-0.007)	2,72E-07	-0.003(-0.005-0)	0,0461
*cg02454264	16	4897716	GLYR1	TSS1500	island	0.01(0.014-0.006)	2,74E-07	0.007(0.009-0.005)	1,83E-11
*cg03346041	8	68255952	ARFGEF1	TSS200	island	0.015(0.021-0.01)	2,74E-07	0.023(0.026-0.019)	4,97E-33
*cg22183626	13	45150990	TSC22D1	TSS1500	island	0.014(0.019-0.009)	2,76E-07	0.003(0.006-0)	0,0222
*cg03673687	14	100259804	EML1	1stExon	shore	0.01(0.014-0.006)	2,81E-07	0.011(0.013-0.008)	4,39E-16
*cg06775055	4	18023265	LCORL	1stExon	island	0.003(0.004-0.002)	2,81E-07	0.004(0.005-0.003)	1,97E-15
*cg09174674	1	9884042	CLSTN1	1stExon	island	0.009(0.012-0.006)	2,83E-07	-0.012(-0.017--0.008)	5,81E-07
*cg05266843	8	95961618	TP53INP1	TSS200	island	0.018(0.025-0.011)	2,84E-07	0.019(0.023-0.016)	7,81E-24
*cg14591243	3	136471577	STAG1	TSS1500	island	0.01(0.014-0.006)	2,84E-07	0.006(0.009-0.003)	0,000197
cg02463426	19	16683387	SLC35E1	TSS200	island	0.016(0.022-0.01)	2,89E-07	0.013(0.017-0.009)	1,98E-10
*cg06563451	8	61564681		IGR	island	0.015(0.02-0.009)	2,91E-07	0.003(0.006-0)	0,0772
*cg09344986	7	140774104		IGR	island	0.019(0.026-0.012)	2,92E-07	0.011(0.015-0.007)	1,11E-07
*cg10427040	17	74261306	FAM100B	1stExon	island	0.012(0.017-0.008)	2,92E-07	0.009(0.012-0.006)	1,24E-08
*cg24523758	2	181845005	UBE2E3	TSS200	island	0.011(0.015-0.007)	2,93E-07	0.003(0.005-0)	0,0348
*cg12425673	12	862975	WNK1	1stExon	island	0.01(0.014-0.007)	2,94E-07	0.007(0.01-0.005)	1,80E-09
*cg07485181	2	131851172	FAM168B	TSS200	island	0.012(0.016-0.007)	2,95E-07	0.006(0.009-0.003)	0,000135
*cg06461769	18	9335102	TWSG1	5'UTR	island	0.011(0.015-0.007)	2,96E-07	0.016(0.018-0.014)	3,20E-37
*cg13674535	17	79361026		IGR	island	0.007(0.01-0.005)	2,96E-07	0.005(0.006-0.003)	4,02E-07
*cg22793136	12	94543560	PLXNC1	1stExon	island	0.015(0.021-0.01)	2,97E-07	0.008(0.016-0)	0,039

*cg01751341	16	85045118	ZDHHC7	5'UTR	island	0.007(0.01-0.004)	2,99E-07	0.006(0.007-0.004)	6,39E-10
*cg04672450	1	860185	SAMD11	TSS1500	island	0.013(0.018-0.008)	2,99E-07	0.001(0.004--0.002)	0,444
*cg21140261	1	197871594	C1orf53	TSS200	island	0.011(0.015-0.007)	3,01E-07	0.004(0.008-0.001)	0,0068
*cg13402779	14	50101323	C14orf104	1stExon	island	0.011(0.015-0.007)	3,04E-07	0.002(0.006--0.001)	0,173
*cg04117874	5	149380247	HMGXB3	5'UTR	island	0.003(0.005-0.002)	3,06E-07	0.002(0.003-0.001)	0,000362
*cg26353862	11	10315609	SBF2	1stExon	island	0.014(0.02-0.009)	3,07E-07	0.009(0.013-0.004) -0.015(-0.018-- 0.013)	0,000247
*cg10535933	15	52821418	MYO5A	TSS200	island	0.004(0.006-0.003)	3,10E-07		4,69E-26
*cg11016194	1	42501599		IGR	island	0.015(0.021-0.009)	3,11E-07	0.008(0.011-0.004)	4,62E-05
*cg11689622	1	234744328	IRF2BP2	1stExon	island	0.008(0.011-0.005)	3,11E-07	0.011(0.013-0.009)	1,23E-22
*cg04464465	12	123755775	CDK2AP1	Body	island	0.012(0.017-0.008)	3,15E-07	0.02(0.023-0.017)	4,76E-33
*cg24354901	1	202777590	KDM5B	TSS200	island	0.008(0.011-0.005)	3,17E-07	0.012(0.014-0.01)	4,92E-27
*cg17492326	8	124780746	FAM91A1	TSS200	island	0.022(0.03-0.014)	3,22E-07	-0.001(-0.005-0.004)	0,818
*cg19484886	12	62997435	C12orf61	TSS1500	island	0.015(0.02-0.009)	3,23E-07	0.018(0.021-0.014)	4,98E-19
*cg10391895	11	64781672	ARL2	1stExon	island	0.017(0.024-0.011)	3,26E-07	0.014(0.018-0.009)	4,83E-10
*cg14462015	4	79860507	PAQR3	1stExon	island	0.008(0.011-0.005)	3,26E-07	-0.057(-0.063--0.05)	2,82E-57
*cg12923613	4	83812206	SEC31A	5'UTR	island	0.01(0.014-0.006)	3,27E-07	0.005(0.008-0.001)	0,00416
*cg15140191	6	56819615	BEND6	TSS200	island	0.007(0.01-0.004)	3,32E-07	-0.02(-0.023--0.017)	4,89E-37
*cg09410234	1	111217715	KCNA3	TSS200	island	0.015(0.02-0.009)	3,33E-07	0.023(0.026-0.019)	5,33E-35
*cg00908526	5	11904084	CTNND2	5'UTR	island	0.025(0.035-0.016)	3,39E-07	0.012(0.016-0.008)	1,44E-08
*cg03309786	2	232574968	PTMA	Body	island	0.004(0.006-0.003)	3,41E-07	0.005(0.006-0.004)	8,23E-29
*cg22112144	5	72112333	TNPO1	TSS200	island	0.012(0.016-0.007)	3,42E-07	0.02(0.023-0.018)	1,43E-40
*cg14724283	1	29564502	PTPRU	Body	island	0.016(0.022-0.01)	3,49E-07	0.017(0.02-0.013)	1,90E-20
*cg05307908	4	37687729	RELL1	Body	island	0.01(0.014-0.006)	3,54E-07	0.011(0.013-0.009)	2,66E-18
*cg02331143	1	42128043	HIVEP3	5'UTR	island	0.006(0.009-0.004)	3,55E-07	0.006(0.007-0.005)	9,14E-30
*cg01824804	5	121413478	LOX	1stExon	island	0.019(0.026-0.012)	3,57E-07	0.017(0.021-0.014)	1,75E-19
*cg21848143	1	39269995		IGR	island	0.008(0.01-0.005)	3,58E-07	-0.007(-0.01--0.005)	3,21E-08
*cg04276723	7	38670597	AMPH	Body	island	0.013(0.018-0.008)	3,61E-07	0.004(0.006-0.001)	0,00926
*cg08957360	14	55369030	GCH1	Body	island	0.011(0.015-0.007)	3,63E-07	0.013(0.015-0.01)	7,60E-20
*cg02195676	22	46646564	C22orf40	TSS1500	island	0.013(0.017-0.008)	3,67E-07	0.01(0.012-0.007)	2,94E-12
*cg05151228	1	110752952	KCNC4	TSS1500	island	0.017(0.024-0.011)	3,71E-07	0.006(0.01-0.003)	0,000781
*cg03834947	15	65321822	MTFMT	1stExon	island	0.019(0.026-0.012)	3,72E-07	0.002(0.005--0.001)	0,173
*cg07259398	16	2040017	SYNGR3	5'UTR	shore	0.007(0.009-0.004)	3,73E-07	0(0.002--0.002)	0,751
*cg04543026	1	110453888	CSF1	Body	island	0.012(0.016-0.007)	3,74E-07	0.005(0.009-0.002)	0,00298
*cg24578966	4	84457523	AGPAT9	TSS200	island	0.012(0.017-0.008)	3,79E-07	0.007(0.011-0.004)	3,78E-06
*cg18153326	1	42922229	ZMYND12	TSS1500	island	0.017(0.024-0.011)	3,83E-07	-0.003(-0.006-0)	0,0427
*cg24341498	9	137217390	RXRA	TSS1500	island	0.013(0.018-0.008)	3,83E-07	-0.002(-0.007-0.004)	0,582
*cg06576198	5	159200428		IGR	island	0.016(0.022-0.01)	3,85E-07	0.005(0.007-0.002)	9,38E-05
*cg02547025	2	30454275	LBH	TSS200	island	0.021(0.029-0.013)	3,89E-07	0.02(0.024-0.016)	9,86E-23
*cg25296860	13	20702426		IGR	island	0.014(0.019-0.009)	3,93E-07	0.009(0.013-0.006)	4,58E-07
*cg19375676	1	244615173	ADSS	5'UTR	island	0.01(0.013-0.006)	3,98E-07	0.008(0.011-0.006)	6,76E-11
*cg26911804	13	111806221	ARHGEF7	Body	island	0.005(0.007-0.003)	3,99E-07	0.007(0.008-0.005)	4,24E-18
*cg15752043	10	120789955	NANOS1	1stExon	opensea	0.011(0.015-0.007)	4,04E-07	0.011(0.013-0.008)	1,53E-15
*cg22186533	8	144680410	EEF1D	TSS1500	island	0.009(0.012-0.006)	4,04E-07	0.008(0.01-0.006)	2,31E-12
*cg02201963	18	12702795	PSMG2	TSS1500	island	0.013(0.018-0.008)	4,14E-07	0.01(0.013-0.007)	5,82E-10
*cg26889928	7	138145433	TRIM24	1stExon	island	0.014(0.019-0.008)	4,15E-07	0.013(0.015-0.01)	7,26E-22
*cg22960945	3	9957459	IL17RE	Body	island	0.004(0.005-0.002)	4,22E-07	0.004(0.005-0.003)	3,03E-13

*cg05013695	10	99079798	FRAT1	1stExon	island	0.01(0.014-0.006)	4,27E-07	NA	NA
*cg21800906	11	32914818	QSER1	5'UTR	island	0.01(0.013-0.006)	4,30E-07	0.019(0.023-0.014)	1,74E-14
*cg13970591	19	1491289	REEP6	1stExon	island	0.012(0.017-0.008)	4,31E-07	0.011(0.014-0.007)	1,09E-09
*cg19229702	3	11762011	VGLL4	1stExon	island	0.016(0.022-0.01)	4,37E-07	0.02(0.024-0.016)	3,99E-19
*cg01049678	7	150147536	GIMAP8	TSS1500	opensea	0.015(0.021-0.009)	4,39E-07	-0.013(-0.016--0.01)	1,38E-14
*cg14024356	1	154934173	PYGO2	1stExon	island	0.014(0.02-0.009)	4,39E-07	0.018(0.022-0.015)	1,80E-24
*cg10122932	7	99698990	MCM7	1stExon	island	0.015(0.021-0.009)	4,40E-07	0.01(0.013-0.006)	2,77E-07
*cg03754076	1	232766254		IGR	island	0.012(0.017-0.008)	4,41E-07	0.009(0.012-0.006)	8,37E-10
*cg13464062	7	33148868	RP9	1stExon	island	0.011(0.016-0.007)	4,44E-07	0.018(0.022-0.015)	9,30E-28
*cg18500714	14	100706288	YY1	Body	island	0.029(0.04-0.018)	4,49E-07	0.026(0.035-0.017)	4,97E-08
*cg05554039	11	75379223	MAP6	1stExon	island	0.015(0.021-0.01)	4,51E-07	0.017(0.02-0.013)	1,51E-18
*cg12603531	9	139258402	DNLZ	TSS200	island	0.016(0.022-0.01)	4,59E-07	0.003(0.006--0.001)	0,192
*cg03749777	12	106533046	NUAK1	5'UTR	island	0.014(0.019-0.008)	4,64E-07	0.034(0.038-0.029)	6,61E-38
*cg19144392	22	42666664		IGR	island	0.009(0.012-0.005)	4,65E-07	0.015(0.017-0.012)	2,01E-35
*cg23872935	5	130599536	CDC42SE2	TSS200	island	0.019(0.027-0.012)	4,67E-07	0.016(0.021-0.012)	3,01E-14
*cg08294267	15	72565024	PARP6	TSS1500	island	0.011(0.015-0.007)	4,72E-07	0.016(0.019-0.013)	4,81E-24
*cg05530459	1	225615358	LBR	5'UTR	island	0.009(0.013-0.006)	4,73E-07	0.011(0.013-0.008)	3,68E-13
*cg23857197	1	224033831	TP53BP2	TSS200	island	0.007(0.009-0.004)	4,79E-07	-0.005(-0.007--0.003)	1,47E-05
*cg18858542	15	59664324	MYO1E	Body	island	0.012(0.017-0.008)	4,83E-07	0.003(0.006-0)	0,0355
cg08676905	10	6019609	IL15RA	TSS200	island	0.017(0.023-0.01)	4,86E-07	0.008(0.011-0.004)	4,66E-06
*cg21212389	20	61569206	DIDO1	5'UTR	island	0.006(0.008-0.004)	4,86E-07	0.008(0.009-0.006)	1,13E-21
*cg19612890	12	131356591	RAN	TSS200	island	0.007(0.009-0.004)	4,88E-07	0.011(0.013-0.009)	2,75E-20
*cg27261136	1	202113704	ARL8A	1stExon	island	0.009(0.013-0.006)	4,95E-07	0.009(0.012-0.007)	3,99E-14
*cg11716640	1	113933411	MAGI3	TSS200	island	0.015(0.02-0.009)	4,96E-07	0.008(0.013-0.004)	5,88E-05
*cg02838932	18	12991623	CEP192	5'UTR	island	0.009(0.013-0.006)	4,98E-07	0.016(0.019-0.013)	1,88E-22
*cg15027815	1	202131184	PTPN7	TSS1500	opensea	0.023(0.032-0.014)	5,00E-07	0.013(0.017-0.009)	2,85E-10
*cg05619888	4	122872996	TRPC3	TSS200	island	0.005(0.007-0.003)	5,15E-07	0.006(0.008-0.005)	7,58E-20
*cg13224729	20	306281	SOX12	5'UTR	island	0.005(0.006-0.003)	5,16E-07	0.005(0.006-0.004)	2,55E-14
*cg00375031	8	95835718	INTS8	1stExon	island	0.011(0.015-0.007)	5,20E-07	0.007(0.01-0.004)	2,33E-06
*cg02130453	4	2537789		IGR	island	0.012(0.017-0.007)	5,23E-07	0.025(0.028-0.021)	1,12E-40
*cg18710383	6	111279866	GTF3C6	1stExon	island	0.017(0.023-0.01)	5,28E-07	0.014(0.017-0.01)	7,40E-11
*cg24349555	6	139456526	HECA	1stExon	island	0.013(0.018-0.008)	5,32E-07	0.008(0.011-0.006)	1,46E-08
*cg00753739	10	11653489	USP6NL	1stExon	island	0.016(0.022-0.01)	5,38E-07	0.012(0.016-0.008)	9,72E-09
*cg17760011	3	134205175	CEP63	5'UTR	island	0.008(0.011-0.005)	5,39E-07	0.006(0.009-0.004)	4,02E-08
*cg21754773	3	58477622	KCTD6	TSS200	island	0.011(0.016-0.007)	5,39E-07	0.019(0.022-0.015)	4,46E-22
*cg24512132	1	229478138	C1orf96	Body	island	0.004(0.005-0.002)	5,40E-07	0.002(0.003-0)	0,0119
*cg19907055	16	69345210	VPS4A	TSS200	island	0.012(0.016-0.007)	5,41E-07	0.015(0.017-0.012)	3,51E-27
*cg16377939	16	53088918	CHD9	TSS200	island	0.009(0.013-0.006)	5,45E-07	-0.017(-0.02--0.014)	1,01E-26
*cg25721625	10	44880236	CXCL12	Body	island	0.023(0.032-0.014)	5,46E-07	0.011(0.015-0.007)	2,49E-07
*cg13984040	12	125258948		IGR	opensea	0.044(0.061-0.027)	5,50E-07	0.037(0.044-0.029)	6,53E-22
*cg15627380	17	80656703	RAB40B	TSS200	island	0.012(0.017-0.008)	5,59E-07	-0.011(-0.014--0.007)	1,40E-08
*cg01470283	20	34542109	SCAND1	Body	island	0.011(0.016-0.007)	5,64E-07	0.01(0.013-0.006)	7,89E-09
*cg02737068	2	73340007	RAB11FIP5	5'UTR	island	0.009(0.012-0.005)	5,69E-07	0.012(0.014-0.01)	5,27E-34
*cg09535168	19	19431582	KIAA0892	TSS200	island	0.007(0.01-0.004)	5,69E-07	0.005(0.007-0.003)	4,13E-06
*cg04689407	14	104028889	BAG5	TSS1500	island	0.012(0.017-0.008)	5,70E-07	0(-0.003-0.002)	0,97

*cg18289048	1	225615672	LBR	1stExon	island	0.005(0.007-0.003)	5,71E-07	0.003(0.004-0.002)	4,25E-06
*cg20716688	11	1411018	BRSK2	TSS200	island	0.015(0.02-0.009)	5,71E-07	0.018(0.022-0.015)	1,06E-24
*cg04120766	8	144373364	ZNF696	TSS200	island	0.013(0.017-0.008)	5,75E-07	0.003(0.006-0)	0,0406
*cg16449186	3	186545470		IGR	island	0.008(0.011-0.005)	5,76E-07	0.004(0.007-0.001)	0,00655
*cg18661819	2	131113440	PTPN18	TSS200	island	0.012(0.017-0.008)	5,80E-07	0.019(0.022-0.016)	1,19E-31
*cg22843625	19	50979755	C19orf63	5'UTR	island	0.012(0.017-0.008)	5,84E-07	0.008(0.01-0.006)	1,86E-16
*cg09372755	1	43919584	HYI	5'UTR	island	0.018(0.025-0.011)	5,86E-07	-0.02(-0.027--0.012)	7,01E-07
*cg11382291	4	160024674		IGR	island	0.011(0.016-0.007)	6,00E-07	0.04(0.046-0.035)	5,71E-45
*cg15822346	6	111408761	SLC16A10	TSS200	island	0.013(0.018-0.008)	6,00E-07	0.008(0.011-0.005)	1,88E-07
*cg16551665	17	30814028	CDK5R1	TSS200	island	0.02(0.028-0.013)	6,00E-07	0.02(0.024-0.016)	7,44E-20
*cg00203461	10	70287181	SLC25A16	5'UTR	island	0.011(0.015-0.007)	6,02E-07	0.009(0.012-0.006)	1,12E-09
*cg02898212	8	133787670	PHF20L1	1stExon	island	0.016(0.021-0.01)	6,02E-07	-0.02(-0.024--0.017)	3,98E-27
*cg14915787	3	106959583	LOC344595	Body	island	0.007(0.009-0.004)	6,03E-07	0.008(0.01-0.007)	1,17E-21
*cg14489649	17	40021376	KLHL11	1stExon	island	0.011(0.015-0.007)	6,09E-07	0.01(0.013-0.007)	2,52E-09
*cg05200311	6	18264737	DEK	1stExon	island	0.015(0.021-0.009)	6,10E-07	0.026(0.029-0.023)	4,00E-58
*cg03624415	15	64995499	OAZ2	TSS200	island	0.011(0.016-0.007)	6,15E-07	0.009(0.013-0.004)	4,21E-05
*cg10513276	9	115095793	ROD1	1stExon	island	0.014(0.019-0.009)	6,19E-07	0.017(0.02-0.014)	7,64E-21
*cg08894790	4	177241590	SPCS3	Body	island	0.011(0.015-0.007)	6,20E-07	0.004(0.007-0.001)	0,0184
cg01963056	8	144911482	PUF60	5'UTR	island	0.005(0.006-0.003)	6,21E-07	0.007(0.008-0.006)	1,86E-29
*cg10335659	2	99757356	C2orf15	TSS1500	island	0.012(0.016-0.007)	6,24E-07	0.011(0.016-0.006)	7,56E-05
*cg25962774	5	176831364	F12	Body	island	0.011(0.015-0.007)	6,25E-07	0.002(0.005--0.001)	0,269
*cg19259146	1	26856420	RPS6KA1	1stExon	island	0.011(0.015-0.007)	6,26E-07	0.002(0.005--0.001)	0,146
*cg22935821	12	53738700		IGR	island	0.006(0.008-0.004)	6,26E-07	0.005(0.006-0.003)	1,07E-08
*cg15645079	20	3140569	FASTKD5	TSS200	island	0.009(0.013-0.006)	6,27E-07	0.009(0.012-0.006)	1,29E-07
*cg13497301	2	86564636	REEP1	1stExon	island	0.011(0.015-0.007)	6,28E-07	0.009(0.012-0.006)	2,75E-07
*cg14556391	2	47748407	KCNK12	Body	island	0.01(0.013-0.006)	6,30E-07	0.013(0.018-0.008)	1,34E-06
*cg02008025	2	102314465	MAP4K4	TSS200	island	0.011(0.015-0.007)	6,31E-07	0.018(0.021-0.015)	7,05E-35
*cg10662979	7	87563850	ADAM22	1stExon	island	0.012(0.016-0.007)	6,39E-07	0.014(0.017-0.011)	2,17E-20
*cg02729805	20	30865674	KIF3B	5'UTR	island	0.011(0.015-0.007)	6,47E-07	0(-0.003-0.003)	0,907
*cg24989061	4	185747383	ACSL1	TSS200	island	0.011(0.015-0.007)	6,49E-07	0.011(0.013-0.009)	1,45E-20
*cg06592416	3	57542240	PDE12	1stExon	island	0.008(0.011-0.005)	6,55E-07	0.002(0.004-0)	0,0331
*cg05632623	17	78450055	NPTX1	1stExon	island	0.013(0.018-0.008)	6,64E-07	0.007(0.01-0.003)	9,03E-05
*cg20888386	2	219858260	CRYBA2	TSS200	island	0.013(0.018-0.008)	6,74E-07	0.004(0.007-0.002)	0,000842
*cg20274746	5	65221372	ERBB2IP	TSS1500	island	0.012(0.017-0.008)	6,77E-07	0.018(0.022-0.013)	4,93E-13
*cg22347322	22	43807980	MPPED1	TSS200	island	0.011(0.015-0.007)	6,78E-07	0.005(0.007-0.002)	8,42E-05
*cg24175777	7	149119846		IGR	island	0.011(0.015-0.007)	6,80E-07	0(-0.003-0.002)	0,797
*cg06156768	4	128983279	LARP1B	5'UTR	island	0.013(0.018-0.008)	6,82E-07	0.009(0.012-0.006)	8,33E-09
*cg10573403	14	36003118	INSM2	TSS200	island	0.008(0.012-0.005)	6,83E-07	0.006(0.009-0.003)	5,05E-05
*cg09754948	16	28834200	ATXN2L	TSS1500	island	0.024(0.033-0.015)	6,85E-07	0.003(0.007--0.001)	0,108
*cg10830518	6	111804752	REV3L	TSS1500	island	0.007(0.009-0.004)	6,91E-07	0.003(0.004-0.001)	7,03E-05
*cg22377237	7	100271437	GNB2	1stExon	island	0.013(0.018-0.008)	6,92E-07	0.012(0.014-0.009)	1,52E-14
*cg04307927	17	41132963	AARSD1	TSS1500	island	0.008(0.011-0.005)	6,97E-07	0.013(0.016-0.011)	2,53E-21
*cg26170244	15	65176353		IGR	opensea	0.029(0.04-0.018)	7,01E-07	0.028(0.033-0.023)	2,64E-24
*cg07568751	12	124069168	TMED2	1stExon	island	0.011(0.015-0.007)	7,16E-07	0.004(0.007-0.001)	0,00915
*cg07494420	14	96858044	AK7	TSS1500	island	0.012(0.017-0.008)	7,18E-07	0.028(0.032-0.024)	3,63E-45
*cg11748936	14	75593738	NEK9	1stExon	island	0.009(0.013-0.006)	7,20E-07	0.008(0.01-0.006)	2,95E-13

*cg12818493	21	44527599	U2AF1	5'UTR	island	0.016(0.023-0.01)	7,30E-07	0.054(0.062-0.046)	2,97E-39
*cg03169876	13	21750926	MRP63	5'UTR	island	0.009(0.012-0.005)	7,39E-07	0(-0.002-0.002)	0,946
*cg17083941	11	8285557	LMO1	TSS200	shore	0.021(0.029-0.013)	7,39E-07	0.21(0.233-0.188)	8,37E-69
*cg07696592	3	71114001	FOXP1	Body	island	0.009(0.013-0.006)	7,43E-07	0.013(0.016-0.011)	2,42E-29
*cg02541231	11	2398776	CD81	1stExon	island	0.013(0.019-0.008)	7,55E-07	0.01(0.014-0.006)	5,40E-07
*cg01249735	7	156742241	NOM1	TSS200	shore	0.008(0.011-0.005)	7,60E-07	-0.017(-0.02--0.015)	9,02E-31
*cg17976065	17	15902694	ZSWIM7	Body	island	0.007(0.009-0.004)	7,61E-07	-0.003(-0.005-0)	0,0346
*cg08302713	6	53409850	GCLC	TSS200	island	0.012(0.016-0.007)	7,63E-07	0.019(0.022-0.016)	4,23E-26
*cg09400281	14	100259801	EML1	1stExon	shore	0.004(0.006-0.003)	7,63E-07	0.001(0.004--0.001)	0,372
*cg25974903	1	65211057	RAVER2	1stExon	island	0.011(0.015-0.007)	7,67E-07	0.009(0.011-0.006)	4,43E-09
*cg25560333	10	11060104	CUGBP2	1stExon	island	0.008(0.01-0.005)	7,85E-07	0.004(0.005-0.002)	1,91E-05
*cg02641539	8	105351190	TM7SF4	TSS1500	opensea	0.028(0.038-0.017)	7,86E-07	NA	NA
*cg15515895	17	35766750	ACACA	1stExon	island	0.011(0.016-0.007)	8,05E-07	NA	NA
*cg06974515	5	94956175	GPR150	1stExon	island	0.009(0.013-0.006)	8,06E-07	0.011(0.014-0.008)	3,70E-13
*cg22758970	17	63133468	RGS9	5'UTR	island	0.008(0.011-0.005)	8,08E-07	0.01(0.011-0.008)	4,01E-23
*cg21205504	8	124286560	ZHX1	TSS200	island	0.013(0.019-0.008)	8,10E-07	0.015(0.018-0.011)	3,28E-16
*cg01267522	10	31608548	ZEB1	Body	island	0.01(0.014-0.006)	8,15E-07	0.025(0.028-0.022)	2,09E-50
*cg06529027	2	113033205	ZC3H6	5'UTR	island	0.013(0.018-0.008)	8,16E-07	0.032(0.036-0.028)	1,02E-55
*cg18531954	7	8009369	GLCC1	1stExon	island	0.006(0.008-0.004)	8,18E-07	-0.005(-0.007--0.004)	2,28E-08
*cg08551857	8	82024104	PAG1	1stExon	island	0.01(0.013-0.006)	8,19E-07	0.015(0.017-0.012)	4,31E-30
*cg03614513	1	115632038	TSPAN2	1stExon	island	0.01(0.014-0.006)	8,21E-07	0.002(0.005-0)	0,0405
*cg03461725	15	85525160	PDE8A	TSS200	island	0.007(0.01-0.004)	8,28E-07	-0.049(-0.056--0.043)	2,40E-49
*cg03042149	15	85524778	PDE8A	TSS1500	island	0.007(0.009-0.004)	8,30E-07	0.005(0.006-0.003)	2,60E-07
*cg09006498	2	131850875	FAM168B	5'UTR	island	0.006(0.008-0.004)	8,39E-07	0.015(0.018-0.012)	2,90E-25
*cg00784703	21	45139229	PDXK	1stExon	island	0.005(0.007-0.003)	8,44E-07	0.005(0.006-0.004)	1,06E-16
*cg26963632	16	85558148		IGR	opensea	0.028(0.038-0.017)	8,44E-07	NA	NA
*cg20203327	10	126605322		IGR	island	0.005(0.007-0.003)	8,51E-07	0.002(0.004-0.001)	0,00028
*cg10393931	7	95951432	SLC25A13	1stExon	island	0.018(0.025-0.011)	8,55E-07	-0.015(-0.019--0.011)	1,34E-12
*cg09142020	2	236402329	AGAP1	TSS1500	island	0.024(0.034-0.015)	8,57E-07	0.02(0.025-0.016)	1,63E-17
*cg01902998	14	65346633		IGR	island	0.016(0.023-0.01)	8,59E-07	0.013(0.019-0.007)	1,00E-05
*cg24224803	6	160183104	ACAT2	5'UTR	island	0.011(0.015-0.007)	8,65E-07	0.015(0.018-0.011)	1,26E-15
*cg02624574	17	4900769	INCA1	5'UTR	island	0.003(0.005-0.002)	8,69E-07	0.004(0.006-0.003)	1,43E-09
*cg04615716	16	69364337	COG8	3'UTR	island	0.006(0.008-0.003)	8,77E-07	0.01(0.011-0.008)	1,63E-33
*cg00860370	1	104068121	RNPC3	TSS1500	shore	0.003(0.005-0.002)	8,96E-07	0.004(0.005-0.003)	1,16E-15
*cg09801523	4	128982462	LARP1B	TSS200	island	0.012(0.017-0.008)	9,14E-07	NA	NA
*cg04004861	8	74659081	STAU2	5'UTR	island	0.013(0.018-0.008)	9,15E-07	0.008(0.011-0.005)	4,83E-08
*cg15270851	13	111358291	CARS2	1stExon	island	0.012(0.016-0.007)	9,17E-07	NA	NA
*cg14414971	7	3340979	SDK1	TSS200	island	0.019(0.026-0.011)	9,21E-07	0.015(0.018-0.012)	1,28E-20
*cg02269221	17	27071208	TRAF4	1stExon	island	0.009(0.012-0.005)	9,42E-07	NA	NA
*cg12285409	2	9614722	IAH1	1stExon	island	0.02(0.028-0.012)	9,44E-07	0.023(0.028-0.018)	1,17E-17
*cg20410192	4	185746558	ACSL1	5'UTR	island	0.008(0.012-0.005)	9,49E-07	0.009(0.011-0.006)	7,80E-10
*cg10473508	5	172068142	NEURL1B	TSS200	shore	0.013(0.018-0.008)	9,53E-07	0.004(0.007-0)	0,0367
*cg05914439	17	34958381	MRM1	1stExon	island	0.012(0.017-0.007)	9,55E-07	0.009(0.013-0.005)	1,47E-06
*cg21662506	5	167956425	FBLL1	TSS200	island	0.015(0.02-0.009)	9,55E-07	0.013(0.017-0.009)	5,70E-10
*cg24760647	2	220408209	TMEM198	TSS1500	island	0.013(0.018-0.008)	9,56E-07	0.019(0.023-0.015)	1,65E-20

*cg17083411	1	203274503	BTG2	TSS200	island	0.008(0.012-0.005)	9,66E-07	NA	NA
*cg19509894	19	8070923	ELAVL1	TSS1500	opensea	0.005(0.007-0.003)	9,75E-07	0.007(0.009-0.006)	6,46E-32
*cg05454284	10	32636382	EPC1	TSS1500	island	0.007(0.01-0.004)	9,79E-07	NA	NA
*cg05468351	5	21459598	LOC728411	Body	island	0.015(0.02-0.009)	9,80E-07	0.012(0.019-0.006)	0,000147
*cg10780495	1	8939136	ENO1	TSS1500	island	0.005(0.007-0.003)	9,81E-07	0.019(0.022-0.017)	8,51E-39
*cg11009736	2	119699682	MARCO	TSS200	opensea	0.04(0.055-0.024)	9,83E-07	NA	NA
*cg22503499	6	30294449	TRIM39	TSS200	island	0.01(0.014-0.006)	9,84E-07	NA	NA
*cg05036147	17	79479952	ACTG1	TSS200	island	0.014(0.02-0.009)	9,96E-07	0.016(0.019-0.013)	2,45E-23

Table e-3: DMRs in the discovery cohort

CHR	start	end	no.cpgs	P	Genes
chr5	56110833	56112222	8	4.42E-12	MAP3K1
chr15	101792098	101792838	9	6.20E-09	CHSY1
chr2	242576411	242577535	8	6.70E-09	ATG4B
chr9	123476683	123477254	2	1.77E-08	MEGF9
chr16	73091917	73092823	4	3.66E-08	ZFH3
chr16	67062875	67064157	10	6.48E-08	CBFB
chr8	74206531	74208178	8	7.66E-08	SNORA7
chr17	80477010	80477962	7	5.80E-07	SNORA69
chr17	19770908	19770910	2	8.32E-07	SNORA69
chr5	14871469	14872551	10	8.88E-07	CTB-40H15.4
chr12	44229512	44230058	5	9.76E-07	TMEM117
chr5	78280882	78281103	2	1.02E-06	snoZ6
chr9	139010494	139010884	2	1.21E-06	C9orf69
chr21	44394436	44395084	7	1.54E-06	PKNOX1
chr9	89560881	89561881	5	2.56E-06	GAS1
chr10	5707981	5708674	3	3.30E-06	ASB13
chr16	4852372	4852830	4	3.50E-06	ROGDI
chr3	12598618	12598824	3	3.92E-06	MKRN2
chr1	224370909	224371029	2	4.11E-06	snoU13
chr6	155316380	155316492	2	4.45E-06	TIAM2
chr10	64564801	64565336	3	4.62E-06	ADO
chr1	14075870	14076293	5	5.09E-06	snoU13
chr15	40574544	40574925	3	5.13E-06	PLCB2
chr22	39101979	39102110	3	5.46E-06	GTPBP1
chr16	88772396	88772985	6	6.38E-06	CTU2
chr5	132083105	132083184	2	7.14E-06	CCNI2
chr4	1857817	1858231	2	7.21E-06	LETM1
chr10	64133674	64134403	10	7.64E-06	ZNF365
chr10	92617562	92617735	2	9.06E-06	SNORA17
chr21	33103721	33104301	3	1.08E-05	SCAF4
chr5	78531681	78532944	8	1.25E-05	JMY
chr7	139876520	139876578	2	1.25E-05	KDM7A

chr12	111843567	111843939	5	1.45E-05	SH2B3
chr15	81293687	81294865	6	1.66E-05	MESDC1
chr2	111879711	111880018	4	1.77E-05	SNORA64
chr12	121647936	121648093	3	2.13E-05	P2RX4
chr19	10828146	10829423	19	2.23E-05	DNM2
chr1	52607471	52607933	4	2.34E-05	snoU13
chr2	242625439	242626874	12	2.71E-05	DTYMK
chr19	662740	663107	4	2.82E-05	RNF126
chr10	104403707	104404379	3	3.26E-05	TRIM8
chr2	101767621	101767677	2	3.35E-05	SNORA73
chr21	45138294	45139379	6	4.12E-05	PDXK
chr15	73344867	73345368	3	4.13E-05	NEO1
chr3	71803308	71803558	3	4.21E-05	U7
chr6	43139370	43139870	5	4.42E-05	SNORA38
chr3	159943315	159943466	3	4.69E-05	C3orf80
chr21	46494133	46494823	4	5.36E-05	ADARB1
chr22	43253145	43253208	3	5.97E-05	ARFGAP3
chr14	103801128	103801660	5	6.05E-05	EIF5
chr13	21278118	21278287	2	6.11E-05	IL17D
chr17	30813688	30814031	4	6.71E-05	SNORA69
chr15	83478255	83478541	4	6.99E-05	WHAMM
chr16	88869801	88870304	4	8.22E-05	CDT1
chr14	64319812	64320521	4	8.80E-05	SYNE2
chr9	140063205	140064036	5	9.15E-05	GRIN1
chr2	7005449	7005733	3	9.57E-05	SNORA73
chr4	122872285	122873181	9	1.11E-04	TRPC3
chr5	132165899	132166663	4	1.21E-04	SNORA27
chr1	90286561	90287101	4	1.25E-04	snoU13
chr2	33172442	33172444	2	1.28E-04	SNORA73
chr1	101004934	101005971	6	1.31E-04	snoU13
chr16	4897216	4898306	9	1.34E-04	UBN1
chr14	100772763	100772766	2	1.43E-04	SLC25A29
chr17	66287913	66287915	2	1.53E-04	SNORA69
chr22	45098335	45098367	3	1.62E-04	PRR5

chr17	79885240	79886116	11	1.67E-04	SNORA69
chr20	44717989	44718168	2	1.69E-04	NCOA5
chr10	28966033	28966822	4	1.71E-04	BAMBI
chr4	154387421	154387984	5	1.85E-04	KIAA0922
chr13	51483857	51484248	4	2.08E-04	SNORD36
chr16	67597316	67597374	2	2.33E-04	CTCF
chr4	186064123	186064648	7	2.49E-04	SLC25A4
chr5	127419310	127420219	6	2.82E-04	SLC12A2
chr8	144680115	144680433	5	2.84E-04	TIGD5
chr16	84150069	84151246	9	2.85E-04	RP11-505K9.1
chr5	72111926	72112888	12	2.86E-04	TNPO1
chr7	21467405	21467723	3	3.24E-04	SP4
chr11	67888603	67888858	3	3.33E-04	CHKA
chr14	100259801	100259860	3	3.36E-04	EML1
chr17	16256389	16256568	2	3.64E-04	SNORA69
chr2	85198334	85198562	3	3.76E-04	SNORA73
chr21	36261693	36261702	2	3.81E-04	RUNX1
chr13	28194657	28194884	7	3.97E-04	SNORD36
chr15	25683909	25684205	5	3.98E-04	UBE3A
chr9	140196251	140197048	4	4.04E-04	NRARP
chr17	57184151	57184187	3	4.43E-04	SNORA69
chr19	16738737	16739131	5	5.05E-04	CTC-429P9.2
chr13	77460298	77460322	2	5.07E-04	SNORD36
chr12	62997103	62997631	10	5.34E-04	RP11-631N16.2
chr13	50366190	50366648	4	5.72E-04	SNORD36
chr16	46864715	46865118	5	6.31E-04	SNORD111
chr9	111881453	111882389	6	7.01E-04	TMEM245
chr6	56819432	56819623	4	7.02E-04	SNORA38
chr19	59056208	59056346	2	7.39E-04	TRIM28

Table e-3: Significant EWAS results by stroke subtypes

LAS			CES		SVS		UND	
CpG	logFC(CI)	P	logFC(CI)	P	logFC(CI)	P	logFC(CI)	P
cg07786668**	0.081(0.102-0.06)	3,54E-11	0.091(0.115-0.067)	2,13E-10	0.082(0.127-0.038)	6,88E-04	0.089(0.112-0.066)	8,02E-11
cg10115490	0.045(0.055-0.034)	1,76E-12	0.053(0.064-0.041)	9,85E-14	0.046(0.066-0.026)	4,54E-05	0.054(0.065-0.043)	1,14E-14
cg15591678	0.043(0.052-0.034)	7,53E-14	0.047(0.057-0.038)	1,52E-14	0.032(0.049-0.016)	2,82E-04	0.043(0.053-0.033)	1,54E-11
cg13432391	0.037(0.047-0.028)	1,91E-11	0.044(0.053-0.035)	7,17E-14	0.03(0.046-0.015)	3,31E-04	0.04(0.049-0.031)	1,30E-12
cg09000178	0.09(0.113-0.066)	1,18E-10	0.102(0.129-0.075)	1,27E-10	0.093(0.141-0.045)	4,48E-04	0.091(0.117-0.066)	1,52E-09
cg10849092	0.036(0.045-0.026)	1,26E-10	0.041(0.053-0.03)	4,19E-10	0.035(0.054-0.016)	7,12E-04	0.039(0.049-0.029)	2,00E-10
cg23681311**	0.068(0.088-0.048)	3,79E-09	0.087(0.11-0.065)	9,50E-11	0.074(0.122-0.027)	3,46E-03	0.08(0.102-0.058)	6,99E-10
cg07667295	0.051(0.064-0.038)	2,98E-11	0.055(0.069-0.041)	9,35E-11	0.05(0.074-0.026)	2,09E-04	0.047(0.062-0.032)	5,09E-08
cg09238598	0.059(0.076-0.043)	4,11E-10	0.07(0.088-0.052)	6,62E-11	0.055(0.087-0.024)	1,27E-03	0.061(0.08-0.042)	1,56E-08
cg02516155	0.03(0.038-0.021)	9,25E-10	0.04(0.05-0.029)	7,98E-11	0.031(0.048-0.014)	9,95E-04	0.035(0.045-0.025)	1,42E-09
cg22512322	0.038(0.049-0.026)	1,70E-08	0.042(0.053-0.031)	1,55E-10	0.05(0.07-0.029)	2,34E-05	0.046(0.057-0.036)	3,05E-12
cg06642177	0.085(0.109-0.061)	9,81E-10	0.106(0.135-0.078)	2,20E-10	0.085(0.144-0.025)	6,80E-03	0.094(0.119-0.068)	6,76E-10
cg17662034	0.091(0.119-0.063)	1,07E-08	0.104(0.137-0.071)	2,44E-08	0.089(0.157-0.021)	1,20E-02	0.1(0.132-0.068)	5,38E-08
cg14119263	0.062(0.078-0.045)	8,00E-11	0.063(0.081-0.046)	1,02E-09	0.06(0.093-0.027)	7,57E-04	0.064(0.08-0.048)	8,23E-11
cg14864852	0.037(0.046-0.027)	5,04E-11	0.039(0.051-0.028)	3,41E-09	0.034(0.055-0.012)	3,51E-03	0.04(0.051-0.028)	1,82E-09
cg25869317**	0.064(0.081-0.047)	9,85E-11	0.077(0.098-0.056)	5,56E-10	0.068(0.101-0.035)	2,18E-04	0.075(0.093-0.057)	1,38E-11
cg08867893	0.036(0.045-0.027)	2,96E-11	0.038(0.047-0.028)	4,97E-11	0.029(0.049-0.01)	4,16E-03	0.037(0.048-0.027)	1,05E-09
cg14078059	0.095(0.123-0.067)	2,94E-09	0.102(0.134-0.07)	2,44E-08	0.097(0.156-0.038)	2,25E-03	0.112(0.142-0.082)	3,12E-10
cg10206657	0.065(0.086-0.044)	4,24E-08	0.086(0.111-0.062)	1,30E-09	0.057(0.098-0.015)	9,87E-03	0.08(0.102-0.058)	1,19E-09
cg16588137	0.026(0.035-0.018)	5,59E-08	0.028(0.037-0.019)	1,22E-08	0.029(0.043-0.015)	1,80E-04	0.033(0.04-0.025)	2,27E-12
cg01668174	0.027(0.035-0.018)	9,06E-09	0.032(0.04-0.024)	5,85E-11	0.034(0.052-0.016)	7,38E-04	0.031(0.04-0.022)	2,26E-09
cg08327423	0.02(0.025-0.014)	4,48E-10	0.025(0.032-0.018)	4,27E-10	0.021(0.032-0.009)	8,45E-04	0.02(0.026-0.014)	3,81E-09
cg18829263	0.039(0.05-0.028)	6,27E-10	0.048(0.06-0.035)	1,51E-10	0.032(0.052-0.013)	1,97E-03	0.048(0.06-0.037)	1,29E-11
cg26257177	0.013(0.018-0.009)	2,49E-08	0.018(0.023-0.013)	2,71E-10	0.018(0.029-0.006)	4,52E-03	0.017(0.022-0.012)	2,47E-09
cg09643587	0.054(0.07-0.038)	4,23E-09	0.061(0.08-0.042)	1,80E-08	0.05(0.082-0.017)	4,06E-03	0.059(0.076-0.042)	2,12E-09
cg04759220**	0.029(0.038-0.02)	2,77E-08	0.04(0.051-0.029)	5,05E-10	0.031(0.055-0.007)	1,21E-02	0.036(0.047-0.025)	1,13E-08

cg26587051	0.036(0.045-0.026)	9,18E-11	0.039(0.05-0.029)	9,03E-11	0.029(0.047-0.01)	3,88E-03	0.039(0.049-0.029)	6,28E-11
cg03036064	0.019(0.025-0.014)	2,19E-10	0.022(0.027-0.016)	6,77E-11	0.019(0.027-0.011)	6,52E-05	0.021(0.026-0.016)	4,69E-11
cg17664568	0.022(0.029-0.014)	6,46E-08	0.029(0.037-0.021)	9,59E-10	0.02(0.032-0.007)	2,64E-03	0.026(0.032-0.019)	1,06E-10
cg04794887	0.039(0.05-0.027)	4,92E-09	0.049(0.061-0.036)	2,93E-11	0.046(0.071-0.021)	7,84E-04	0.043(0.055-0.031)	7,15E-10
cg04192862**	0.051(0.066-0.037)	5,59E-10	0.061(0.077-0.044)	4,02E-10	0.052(0.08-0.024)	7,52E-04	0.059(0.074-0.045)	1,82E-11
cg18751306	0.043(0.055-0.03)	5,33E-09	0.048(0.062-0.033)	6,12E-09	0.043(0.072-0.013)	5,91E-03	0.045(0.059-0.031)	1,45E-08
cg08377314	0.033(0.041-0.024)	4,30E-11	0.032(0.041-0.022)	5,80E-09	0.038(0.052-0.024)	5,84E-06	0.034(0.044-0.024)	2,32E-09
cg03852091	0.026(0.034-0.017)	2,78E-08	0.032(0.041-0.023)	9,92E-10	0.02(0.035-0.005)	8,96E-03	0.031(0.04-0.022)	1,61E-09
cg05668205	0.023(0.029-0.016)	2,46E-09	0.03(0.037-0.023)	9,06E-12	0.02(0.031-0.008)	1,36E-03	0.025(0.033-0.018)	9,19E-09
cg03536846	0.019(0.025-0.013)	8,69E-09	0.022(0.029-0.015)	6,77E-09	0.025(0.035-0.014)	3,09E-05	0.022(0.028-0.017)	1,49E-11
cg11299854	0.057(0.075-0.038)	4,23E-08	0.065(0.082-0.048)	2,43E-10	0.059(0.094-0.025)	1,55E-03	0.063(0.08-0.047)	1,84E-10
cg05135536	0.026(0.034-0.018)	8,14E-09	0.028(0.038-0.019)	6,86E-08	0.027(0.04-0.014)	2,72E-04	0.032(0.04-0.025)	5,13E-12
cg00610228	0.029(0.037-0.021)	8,62E-10	0.031(0.039-0.023)	8,30E-11	0.02(0.032-0.007)	3,43E-03	0.03(0.037-0.022)	3,54E-10
cg03732936	0.032(0.042-0.023)	2,46E-09	0.037(0.047-0.026)	1,39E-09	0.03(0.05-0.009)	5,96E-03	0.039(0.049-0.028)	2,89E-10
cg27414860	0.032(0.039-0.024)	8,67E-12	0.034(0.044-0.024)	2,44E-09	0.033(0.048-0.017)	2,27E-04	0.033(0.042-0.025)	5,58E-11
cg16041611	0.044(0.056-0.031)	2,30E-09	0.046(0.06-0.032)	6,29E-09	0.046(0.071-0.02)	9,81E-04	0.048(0.061-0.036)	1,04E-10
cg18766912	0.047(0.062-0.031)	7,02E-08	0.056(0.073-0.039)	7,52E-09	0.052(0.081-0.022)	1,27E-03	0.054(0.07-0.037)	2,17E-08
cg14276025	0.014(0.019-0.009)	5,17E-07	0.019(0.025-0.014)	2,57E-09	0.015(0.026-0.005)	6,20E-03	0.019(0.023-0.014)	9,89E-11
cg04084597	0.036(0.046-0.025)	9,63E-10	0.041(0.053-0.029)	3,73E-09	0.04(0.06-0.019)	4,76E-04	0.036(0.047-0.025)	7,03E-09
cg06793581	0.027(0.035-0.019)	6,87E-09	0.028(0.035-0.02)	1,08E-09	0.027(0.046-0.009)	5,64E-03	0.029(0.038-0.021)	4,59E-09
cg04971341	0.026(0.034-0.019)	4,95E-10	0.03(0.039-0.02)	1,92E-08	0.025(0.038-0.011)	6,22E-04	0.031(0.038-0.023)	5,22E-12
cg27467996	0.026(0.033-0.018)	7,23E-09	0.031(0.04-0.022)	1,63E-09	0.029(0.043-0.015)	1,75E-04	0.027(0.036-0.019)	1,22E-08
cg01281140	0.025(0.031-0.018)	3,43E-10	0.028(0.035-0.021)	3,05E-11	0.025(0.037-0.013)	1,56E-04	0.025(0.032-0.017)	1,08E-08
cg06825833	0.027(0.036-0.018)	5,32E-08	0.034(0.044-0.025)	2,22E-10	0.027(0.041-0.014)	3,02E-04	0.029(0.038-0.021)	2,01E-09
cg19539824	0.038(0.05-0.026)	9,75E-09	0.046(0.059-0.033)	1,42E-09	0.04(0.065-0.014)	3,48E-03	0.045(0.059-0.032)	3,17E-09
cg02304070	0.016(0.022-0.011)	1,68E-07	0.022(0.028-0.015)	6,40E-09	0.015(0.024-0.005)	4,05E-03	0.02(0.025-0.014)	2,93E-10
cg09977800	0.028(0.037-0.019)	7,70E-08	0.034(0.045-0.022)	1,44E-07	0.031(0.049-0.014)	9,55E-04	0.032(0.042-0.023)	7,26E-09
cg18455653	0.062(0.081-0.043)	8,80E-09	0.072(0.093-0.051)	3,62E-09	0.063(0.096-0.03)	5,32E-04	0.067(0.087-0.048)	2,82E-09
cg24906202	0.035(0.046-0.025)	5,28E-09	0.043(0.053-0.032)	1,49E-11	0.04(0.059-0.021)	1,91E-04	0.034(0.045-0.023)	4,21E-08
cg11424525	0.03(0.039-0.021)	5,49E-09	0.034(0.045-0.023)	2,92E-08	0.025(0.043-0.006)	1,24E-02	0.032(0.043-0.022)	4,44E-08
cg11637191	0.016(0.022-0.01)	9,89E-07	0.021(0.028-0.015)	1,89E-08	0.017(0.028-0.006)	4,08E-03	0.023(0.028-0.017)	1,02E-10

cg09572029	0.028(0.036-0.02)	3,04E-09	0.03(0.039-0.021)	1,01E-08	0.034(0.049-0.018)	1,25E-04	0.032(0.041-0.022)	6,91E-09
cg23640903	0.033(0.042-0.023)	1,80E-09	0.038(0.05-0.026)	3,50E-08	0.044(0.062-0.026)	2,23E-05	0.04(0.052-0.029)	2,91E-09
cg18830052	0.04(0.053-0.027)	8,36E-08	0.047(0.061-0.033)	3,21E-09	0.046(0.072-0.02)	1,16E-03	0.049(0.062-0.036)	2,04E-10
cg06802374	0.02(0.026-0.014)	2,37E-09	0.023(0.03-0.016)	5,22E-09	0.019(0.029-0.009)	5,49E-04	0.024(0.03-0.018)	1,21E-11
cg11828669	0.024(0.031-0.016)	2,05E-08	0.026(0.034-0.017)	5,93E-08	0.033(0.047-0.019)	5,48E-05	0.028(0.036-0.021)	3,30E-10
cg01725318	0.028(0.035-0.02)	1,17E-09	0.032(0.041-0.024)	1,98E-10	0.021(0.038-0.005)	1,13E-02	0.027(0.036-0.018)	3,55E-08
cg16063115	0.016(0.021-0.01)	1,35E-07	0.017(0.023-0.012)	1,37E-07	0.018(0.026-0.009)	1,36E-04	0.016(0.021-0.011)	5,53E-09
cg11289281	0.028(0.036-0.02)	2,65E-09	0.038(0.048-0.028)	9,16E-11	0.03(0.05-0.011)	3,56E-03	0.032(0.041-0.022)	5,23E-09
cg02198017	0.021(0.029-0.014)	2,28E-07	0.024(0.031-0.018)	1,05E-09	0.015(0.027-0.004)	1,23E-02	0.026(0.034-0.019)	4,16E-10
cg15772223	0.03(0.04-0.021)	2,58E-08	0.038(0.048-0.028)	4,23E-10	0.025(0.043-0.007)	9,44E-03	0.038(0.047-0.028)	8,82E-11
cg24794206	0.024(0.032-0.017)	1,68E-08	0.027(0.035-0.019)	4,85E-09	0.019(0.03-0.008)	1,08E-03	0.028(0.034-0.021)	1,18E-11
cg13789548	0.02(0.026-0.014)	2,93E-09	0.023(0.03-0.016)	1,72E-08	0.021(0.032-0.009)	1,16E-03	0.022(0.029-0.016)	4,91E-09
cg16887862	0.034(0.046-0.022)	3,49E-07	0.045(0.058-0.032)	3,06E-09	0.029(0.052-0.006)	1,36E-02	0.045(0.058-0.033)	9,99E-10
cg07691609**	0.025(0.033-0.018)	6,38E-09	0.027(0.036-0.018)	1,14E-07	0.018(0.031-0.005)	9,21E-03	0.026(0.034-0.018)	1,53E-08
cg08308556	0.026(0.034-0.017)	1,02E-07	0.029(0.037-0.02)	4,68E-09	0.027(0.042-0.012)	8,62E-04	0.029(0.038-0.02)	2,12E-08
cg21538902	0.028(0.036-0.02)	5,02E-09	0.03(0.039-0.021)	1,39E-08	0.037(0.052-0.022)	2,50E-05	0.032(0.042-0.023)	1,66E-09
cg07331616	0.029(0.039-0.02)	3,37E-08	0.034(0.043-0.024)	1,05E-09	0.032(0.048-0.015)	4,73E-04	0.033(0.042-0.023)	6,07E-09
cg23706268	0.013(0.018-0.009)	1,82E-07	0.015(0.019-0.011)	2,10E-09	0.014(0.024-0.003)	1,62E-02	0.015(0.019-0.01)	5,31E-08
cg24511898	0.018(0.024-0.012)	6,11E-08	0.019(0.026-0.013)	9,60E-08	0.02(0.029-0.01)	2,45E-04	0.019(0.026-0.013)	3,50E-08
cg04887405	0.022(0.03-0.015)	5,19E-08	0.029(0.037-0.022)	9,53E-11	0.026(0.038-0.014)	1,75E-04	0.026(0.033-0.018)	2,71E-09
cg08195842	0.024(0.03-0.017)	2,57E-09	0.025(0.033-0.016)	3,06E-07	0.02(0.032-0.007)	2,51E-03	0.024(0.032-0.016)	4,07E-08
cg13470032	0.019(0.025-0.014)	1,52E-09	0.021(0.028-0.015)	1,09E-08	0.024(0.036-0.013)	1,23E-04	0.02(0.026-0.014)	3,42E-08
cg08502703	0.019(0.026-0.012)	3,52E-07	0.022(0.029-0.015)	8,16E-08	0.019(0.033-0.004)	1,23E-02	0.021(0.028-0.014)	2,23E-07
cg03951424	0.014(0.019-0.01)	4,90E-08	0.017(0.021-0.012)	1,10E-10	0.017(0.026-0.009)	3,09E-04	0.015(0.02-0.01)	1,02E-07
cg05001598	0.027(0.037-0.018)	4,51E-07	0.033(0.043-0.024)	2,22E-09	0.029(0.046-0.011)	1,93E-03	0.031(0.04-0.022)	7,35E-09
cg01628954	0.025(0.032-0.017)	2,98E-09	0.03(0.038-0.021)	3,15E-09	0.025(0.035-0.014)	4,96E-05	0.03(0.037-0.023)	1,17E-11
cg00874357	0.018(0.025-0.012)	2,55E-07	0.022(0.028-0.015)	5,94E-09	0.02(0.031-0.009)	7,08E-04	0.022(0.029-0.016)	7,58E-10
cg06014792	0.023(0.03-0.016)	1,28E-08	0.025(0.033-0.017)	3,87E-08	0.024(0.039-0.008)	4,77E-03	0.025(0.032-0.019)	3,32E-10
cg19244874	0.022(0.03-0.014)	4,33E-07	0.031(0.041-0.022)	7,76E-09	0.025(0.04-0.009)	2,95E-03	0.028(0.037-0.02)	1,03E-08
cg14204791	0.031(0.04-0.021)	1,65E-08	0.031(0.042-0.021)	1,81E-07	0.033(0.051-0.015)	9,28E-04	0.032(0.043-0.022)	6,62E-08
cg04703844	0.023(0.031-0.016)	7,99E-08	0.027(0.036-0.018)	1,59E-07	0.027(0.04-0.014)	2,54E-04	0.028(0.036-0.021)	2,89E-10

cg02551232	0.028(0.038-0.018)	3,76E-07	0.042(0.053-0.031)	3,37E-11	0.028(0.046-0.009)	4,26E-03	0.033(0.043-0.023)	1,92E-08
cg01266790	0.021(0.029-0.013)	3,55E-06	0.03(0.039-0.021)	4,14E-09	0.016(0.032-0.001)	3,96E-02	0.026(0.036-0.017)	3,17E-07
cg23639857	0.018(0.023-0.012)	4,82E-08	0.02(0.026-0.013)	1,86E-07	0.016(0.027-0.006)	3,82E-03	0.021(0.027-0.016)	7,09E-10
cg24915717	0.024(0.031-0.016)	1,52E-08	0.028(0.037-0.018)	2,92E-07	0.026(0.042-0.011)	1,65E-03	0.023(0.031-0.015)	7,30E-07
cg23684603	0.024(0.032-0.015)	3,70E-07	0.029(0.038-0.02)	3,24E-08	0.019(0.036-0.003)	2,46E-02	0.028(0.036-0.02)	8,05E-09
cg11044724	0.019(0.027-0.012)	7,60E-07	0.025(0.032-0.017)	4,14E-08	0.014(0.028-0)	5,43E-02	0.021(0.028-0.013)	2,66E-07
cg20116574	0.031(0.042-0.021)	3,35E-08	0.04(0.052-0.029)	1,54E-09	0.04(0.065-0.015)	2,52E-03	0.035(0.046-0.024)	3,26E-08
cg12661145	0.01(0.014-0.007)	3,26E-07	0.011(0.015-0.008)	5,46E-08	0.011(0.02-0.003)	9,83E-03	0.012(0.016-0.009)	8,17E-09
cg08299976	0.013(0.018-0.008)	2,60E-06	0.017(0.023-0.011)	8,39E-08	0.013(0.025-0.001)	3,01E-02	0.018(0.023-0.012)	1,01E-08
cg01442104	0.021(0.03-0.012)	1,03E-05	0.025(0.035-0.015)	5,44E-06	0.023(0.038-0.007)	5,50E-03	0.027(0.035-0.02)	1,08E-09
cg18764107	0.025(0.033-0.016)	3,84E-07	0.029(0.039-0.02)	4,58E-08	0.031(0.047-0.015)	3,88E-04	0.028(0.038-0.018)	3,30E-07
cg04396495	0.032(0.043-0.021)	1,14E-07	0.039(0.052-0.027)	3,47E-08	0.03(0.051-0.009)	7,17E-03	0.04(0.051-0.028)	3,33E-09
cg16344026	0.022(0.03-0.014)	1,67E-06	0.029(0.038-0.021)	3,43E-09	0.026(0.041-0.01)	2,03E-03	0.027(0.036-0.018)	1,45E-07
cg25506900	0.02(0.028-0.013)	1,86E-06	0.03(0.039-0.021)	2,20E-09	0.025(0.039-0.012)	6,55E-04	0.028(0.036-0.021)	8,37E-11
cg19419497	0.019(0.026-0.013)	7,64E-08	0.022(0.029-0.014)	7,35E-08	0.021(0.032-0.01)	3,68E-04	0.019(0.026-0.012)	2,71E-07
cg19764539	0.026(0.036-0.015)	4,09E-06	0.04(0.049-0.031)	2,69E-12	0.034(0.048-0.02)	2,30E-05	0.032(0.041-0.022)	1,06E-08
cg09915769	0.025(0.035-0.015)	7,24E-06	0.031(0.041-0.02)	2,90E-07	0.027(0.048-0.005)	1,91E-02	0.032(0.044-0.021)	3,45E-07
cg15901997	0.012(0.015-0.008)	2,07E-08	0.011(0.015-0.007)	7,16E-07	0.013(0.021-0.006)	1,20E-03	0.011(0.015-0.007)	3,42E-07
cg04598683	0.019(0.026-0.012)	5,17E-07	0.023(0.029-0.016)	9,73E-10	0.019(0.029-0.008)	1,14E-03	0.021(0.028-0.014)	5,88E-08
cg17255450	0.062(0.084-0.041)	2,14E-07	0.068(0.091-0.046)	8,51E-08	0.058(0.096-0.02)	4,47E-03	0.072(0.095-0.049)	6,42E-08
cg07806715	0.021(0.029-0.014)	2,90E-07	0.028(0.036-0.02)	7,35E-10	0.022(0.035-0.008)	2,92E-03	0.026(0.034-0.018)	9,99E-09
cg05425664	0.022(0.03-0.014)	3,69E-07	0.028(0.035-0.02)	1,39E-09	0.027(0.042-0.012)	8,99E-04	0.026(0.034-0.018)	2,75E-08
cg08625468	0.014(0.019-0.009)	1,29E-07	0.018(0.023-0.013)	3,40E-09	0.014(0.023-0.005)	4,40E-03	0.015(0.02-0.011)	2,78E-08
cg04743947	0.013(0.018-0.008)	9,07E-07	0.015(0.02-0.01)	2,44E-07	0.017(0.026-0.008)	6,90E-04	0.017(0.022-0.013)	1,85E-10
cg27597084	0.014(0.02-0.009)	1,16E-06	0.018(0.025-0.011)	1,08E-06	0.012(0.028-0.004)	1,26E-01	0.018(0.025-0.011)	5,21E-06
cg20689054	0.023(0.031-0.014)	1,68E-06	0.028(0.035-0.02)	3,01E-10	0.023(0.036-0.009)	2,27E-03	0.026(0.035-0.018)	3,30E-08
cg15700006	0.02(0.027-0.013)	3,28E-07	0.026(0.034-0.017)	9,43E-08	0.019(0.033-0.005)	1,04E-02	0.025(0.033-0.018)	5,01E-09
cg16573386**	0.017(0.023-0.011)	6,02E-07	0.022(0.029-0.015)	4,16E-08	0.019(0.032-0.006)	4,65E-03	0.021(0.027-0.014)	2,51E-08
cg19529645	0.036(0.048-0.023)	1,60E-07	0.04(0.052-0.028)	6,32E-09	0.029(0.051-0.007)	1,27E-02	0.038(0.052-0.025)	3,43E-07
cg26932552	0.019(0.026-0.012)	3,04E-07	0.023(0.031-0.015)	2,81E-07	0.024(0.038-0.009)	1,92E-03	0.022(0.029-0.015)	2,32E-08
cg06983715	0.016(0.021-0.011)	4,36E-09	0.018(0.024-0.012)	8,95E-08	0.017(0.028-0.006)	4,90E-03	0.015(0.021-0.01)	4,85E-07

cg17116739	0.028(0.038-0.018)	5,06E-07	0.032(0.043-0.021)	1,83E-07	0.027(0.044-0.01)	2,88E-03	0.032(0.042-0.023)	6,32E-09
cg14860494	0.015(0.021-0.009)	2,15E-06	0.018(0.024-0.012)	2,80E-07	0.023(0.031-0.014)	5,44E-06	0.021(0.027-0.016)	3,27E-10
cg23157089	0.02(0.027-0.013)	3,20E-07	0.022(0.03-0.014)	2,68E-07	0.019(0.034-0.003)	2,04E-02	0.021(0.029-0.013)	1,26E-06
cg14810343	0.039(0.053-0.026)	8,72E-08	0.042(0.058-0.026)	1,18E-06	0.034(0.061-0.006)	1,77E-02	0.047(0.061-0.033)	1,16E-08
cg17626178	0.023(0.033-0.012)	4,73E-05	0.028(0.038-0.018)	7,89E-07	0.019(0.039--0.001)	6,15E-02	0.029(0.039-0.019)	3,52E-07
cg10423910	0.016(0.022-0.01)	1,83E-06	0.022(0.028-0.015)	5,30E-09	0.019(0.03-0.008)	1,63E-03	0.02(0.026-0.014)	2,26E-08
cg08214689	0.018(0.024-0.013)	9,16E-09	0.018(0.024-0.012)	1,26E-07	0.02(0.03-0.01)	4,21E-04	0.02(0.026-0.014)	4,81E-09
cg04729594	0.022(0.031-0.013)	5,69E-06	0.029(0.038-0.02)	5,16E-08	0.018(0.032-0.004)	1,30E-02	0.023(0.032-0.015)	1,41E-06
cg06699484	0.023(0.031-0.014)	6,13E-07	0.025(0.034-0.016)	9,97E-07	0.021(0.036-0.006)	7,06E-03	0.028(0.036-0.021)	6,79E-10
cg06465196	0.015(0.02-0.01)	2,97E-08	0.018(0.024-0.012)	1,29E-07	0.02(0.032-0.007)	2,83E-03	0.019(0.024-0.014)	2,70E-09
cg27250362	0.022(0.029-0.014)	4,43E-07	0.026(0.035-0.017)	3,11E-07	0.023(0.038-0.008)	4,04E-03	0.028(0.036-0.02)	1,44E-09
cg17813364	0.019(0.026-0.013)	7,97E-08	0.023(0.032-0.014)	1,70E-06	0.018(0.03-0.007)	3,56E-03	0.022(0.029-0.015)	2,10E-08
cg19590991	0.018(0.025-0.011)	7,98E-06	0.026(0.034-0.018)	1,21E-08	0.021(0.036-0.006)	6,52E-03	0.029(0.037-0.022)	2,49E-11
cg04088969	0.027(0.036-0.018)	6,02E-08	0.031(0.041-0.021)	3,72E-08	0.023(0.041-0.006)	9,55E-03	0.03(0.04-0.021)	6,62E-09
cg13800586	0.02(0.027-0.012)	1,43E-06	0.026(0.034-0.018)	7,89E-09	0.015(0.028-0.001)	3,39E-02	0.022(0.029-0.014)	3,17E-07
cg02543879	0.019(0.027-0.012)	5,50E-07	0.018(0.026-0.011)	1,03E-05	0.02(0.031-0.008)	1,59E-03	0.022(0.029-0.015)	5,62E-08
cg23281075**	0.019(0.026-0.012)	2,06E-06	0.024(0.032-0.016)	3,14E-08	0.023(0.037-0.009)	1,99E-03	0.021(0.029-0.014)	5,83E-07
cg05988158	0.017(0.023-0.01)	1,09E-06	0.021(0.028-0.014)	4,73E-08	0.019(0.031-0.007)	2,68E-03	0.02(0.027-0.013)	1,60E-07
cg06577463	0.015(0.023-0.008)	8,07E-05	0.025(0.033-0.017)	1,02E-08	0.023(0.038-0.008)	4,50E-03	0.023(0.03-0.016)	3,12E-08
cg11052081	0.017(0.024-0.011)	6,48E-07	0.02(0.026-0.013)	1,45E-07	0.019(0.032-0.007)	3,19E-03	0.022(0.029-0.016)	1,21E-09
cg05039463	0.024(0.032-0.016)	9,84E-08	0.028(0.037-0.018)	2,26E-07	0.012(0.028--0.004)	1,32E-01	0.027(0.035-0.018)	1,28E-07
cg09942303	0.019(0.025-0.012)	5,61E-07	0.021(0.028-0.014)	1,67E-07	0.024(0.037-0.011)	5,40E-04	0.02(0.027-0.013)	1,22E-07
cg24632597	0.008(0.01-0.005)	7,64E-08	0.008(0.011-0.005)	4,58E-07	0.01(0.015-0.004)	1,19E-03	0.009(0.012-0.007)	2,09E-09
cg11233236	0.016(0.023-0.01)	6,20E-06	0.022(0.03-0.015)	9,50E-08	0.016(0.028-0.004)	1,33E-02	0.023(0.03-0.015)	3,22E-08
cg18813733	0.014(0.02-0.009)	2,93E-06	0.019(0.025-0.012)	2,17E-07	0.017(0.027-0.007)	2,01E-03	0.016(0.022-0.009)	3,75E-06
cg18428217	0.016(0.021-0.011)	8,70E-09	0.02(0.027-0.014)	5,01E-08	0.015(0.025-0.005)	5,47E-03	0.021(0.027-0.015)	1,93E-09
cg12526120	0.032(0.044-0.021)	6,02E-07	0.036(0.049-0.024)	2,43E-07	0.027(0.048-0.005)	1,66E-02	0.032(0.045-0.02)	4,87E-06
cg09835867	0.02(0.026-0.014)	1,39E-08	0.024(0.031-0.016)	6,09E-08	0.018(0.031-0.006)	5,28E-03	0.023(0.03-0.016)	4,77E-08
cg02831587	0.011(0.014-0.008)	1,08E-08	0.011(0.015-0.007)	1,56E-06	0.011(0.019-0.004)	4,86E-03	0.012(0.015-0.008)	3,41E-08
cg20366603	0.017(0.024-0.011)	5,77E-07	0.019(0.025-0.013)	6,30E-08	0.023(0.034-0.012)	2,02E-04	0.019(0.025-0.013)	4,64E-08
cg26077228	0.016(0.023-0.009)	9,25E-06	0.025(0.033-0.018)	2,68E-09	0.021(0.031-0.01)	3,39E-04	0.022(0.028-0.016)	4,95E-09

cg17002091	0.029(0.039-0.018)	6,42E-07	0.035(0.046-0.024)	3,18E-08	0.023(0.041-0.005)	1,57E-02	0.03(0.041-0.019)	5,48E-07
cg02543960	0.022(0.03-0.014)	4,61E-07	0.028(0.037-0.019)	2,09E-08	0.017(0.031-0.004)	1,15E-02	0.024(0.032-0.016)	1,34E-07
cg03046332	0.022(0.029-0.015)	1,11E-08	0.027(0.034-0.019)	8,83E-10	0.027(0.04-0.014)	1,34E-04	0.023(0.031-0.016)	7,54E-08
cg05310613	0.035(0.048-0.023)	2,58E-07	0.039(0.052-0.027)	6,33E-08	0.032(0.054-0.01)	5,23E-03	0.043(0.055-0.03)	2,35E-09
cg06202984	0.019(0.026-0.011)	4,84E-06	0.02(0.029-0.012)	1,04E-05	0.026(0.042-0.009)	3,06E-03	0.024(0.032-0.017)	8,38E-09
cg22783262	0.016(0.023-0.01)	3,62E-06	0.021(0.028-0.015)	5,22E-09	0.018(0.027-0.008)	7,63E-04	0.022(0.028-0.016)	9,70E-10
cg25410641	0.02(0.027-0.013)	2,26E-07	0.025(0.033-0.017)	3,97E-08	0.023(0.037-0.009)	1,92E-03	0.023(0.032-0.015)	1,31E-06
cg11950835	0.021(0.029-0.013)	2,10E-06	0.027(0.036-0.018)	2,07E-07	0.021(0.036-0.006)	8,55E-03	0.026(0.034-0.017)	1,49E-07
cg21135832	0.014(0.019-0.008)	2,51E-06	0.02(0.026-0.014)	1,28E-09	0.012(0.02-0.004)	5,81E-03	0.018(0.024-0.011)	6,02E-07
cg19285383	0.011(0.016-0.007)	2,08E-06	0.014(0.019-0.009)	1,07E-07	0.015(0.023-0.007)	8,41E-04	0.014(0.018-0.01)	5,15E-10
cg27563716	0.02(0.027-0.013)	1,03E-07	0.02(0.027-0.012)	4,27E-06	0.021(0.033-0.009)	1,29E-03	0.02(0.026-0.013)	4,26E-08
cg20012028	0.028(0.038-0.017)	2,23E-06	0.03(0.042-0.018)	6,28E-06	0.023(0.044-0.003)	2,68E-02	0.033(0.044-0.023)	5,71E-08
cg21696089	0.016(0.021-0.01)	1,12E-07	0.019(0.026-0.013)	3,22E-07	0.014(0.025-0.003)	1,69E-02	0.017(0.023-0.01)	4,59E-06
cg21475747	0.025(0.034-0.017)	1,25E-07	0.029(0.039-0.02)	7,46E-08	0.023(0.042-0.005)	1,47E-02	0.031(0.04-0.021)	2,85E-08
cg10196289	0.017(0.023-0.01)	1,02E-06	0.016(0.022-0.009)	6,59E-06	0.021(0.032-0.011)	3,52E-04	0.017(0.022-0.012)	5,48E-09
cg23702568	0.017(0.023-0.011)	2,76E-07	0.019(0.026-0.013)	1,37E-07	0.019(0.03-0.007)	2,38E-03	0.018(0.024-0.012)	2,34E-07
cg09567732	0.041(0.056-0.027)	2,72E-07	0.041(0.057-0.024)	7,31E-06	0.041(0.066-0.016)	1,97E-03	0.042(0.057-0.028)	1,91E-07
cg21945459	0.038(0.052-0.025)	1,66E-07	0.045(0.06-0.03)	6,37E-08	0.019(0.049-0.012)	2,23E-01	0.044(0.058-0.029)	9,80E-08
cg16340268	0.016(0.022-0.01)	6,19E-07	0.02(0.026-0.013)	1,04E-07	0.018(0.027-0.009)	2,61E-04	0.019(0.024-0.014)	3,75E-10
cg01109287	0.018(0.024-0.011)	1,36E-06	0.021(0.028-0.014)	1,06E-07	0.022(0.036-0.008)	2,93E-03	0.021(0.028-0.014)	9,65E-08
cg07201929	0.017(0.024-0.009)	1,56E-05	0.018(0.025-0.012)	4,41E-07	0.014(0.025-0.003)	1,55E-02	0.019(0.025-0.013)	6,24E-08
cg01805124	0.021(0.028-0.013)	6,84E-07	0.022(0.032-0.011)	9,44E-05	0.009(0.027-0.009)	3,25E-01	0.024(0.033-0.015)	3,21E-06
cg17868942	0.024(0.032-0.015)	4,95E-07	0.025(0.034-0.015)	1,27E-06	0.026(0.041-0.011)	1,42E-03	0.028(0.037-0.019)	4,66E-08
cg08785922	0.018(0.025-0.011)	2,31E-06	0.025(0.031-0.018)	8,54E-10	0.016(0.029-0.004)	1,34E-02	0.021(0.028-0.014)	1,32E-07
cg00415642	0.018(0.026-0.011)	6,66E-06	0.024(0.031-0.016)	1,48E-08	0.019(0.031-0.007)	3,82E-03	0.025(0.032-0.017)	9,53E-09
cg08522087	0.035(0.049-0.022)	1,88E-06	0.042(0.055-0.029)	2,69E-08	0.028(0.056-0)	4,67E-02	0.037(0.051-0.023)	2,59E-06
cg25433682	0.02(0.028-0.011)	2,42E-05	0.03(0.04-0.021)	3,31E-08	0.031(0.043-0.019)	1,67E-05	0.025(0.032-0.017)	8,56E-09
cg08158076	0.018(0.025-0.012)	1,22E-06	0.023(0.03-0.015)	5,17E-08	0.028(0.041-0.016)	8,24E-05	0.021(0.029-0.014)	7,53E-07
cg06319359	0.009(0.012-0.005)	1,15E-06	0.011(0.015-0.007)	8,75E-07	0.007(0.014-0)	6,18E-02	0.009(0.013-0.006)	3,28E-06
cg03214422	0.052(0.072-0.032)	1,80E-06	0.064(0.085-0.042)	1,02E-07	0.043(0.074-0.011)	9,51E-03	0.064(0.085-0.043)	5,17E-08
cg02385710	0.014(0.019-0.009)	1,27E-06	0.015(0.02-0.01)	1,73E-07	0.014(0.023-0.006)	2,14E-03	0.015(0.02-0.01)	1,14E-07

cg00279406	0.018(0.025-0.011)	1,28E-06	0.02(0.026-0.014)	3,27E-08	0.019(0.031-0.007)	2,86E-03	0.022(0.029-0.015)	8,71E-09
cg12160011	0.023(0.031-0.014)	1,68E-06	0.03(0.039-0.02)	1,28E-08	0.029(0.047-0.011)	2,14E-03	0.025(0.034-0.016)	7,90E-07
cg21160852	0.015(0.021-0.009)	1,27E-06	0.018(0.025-0.012)	5,39E-08	0.017(0.026-0.008)	4,06E-04	0.02(0.026-0.014)	3,47E-09
cg01988923	0.014(0.019-0.008)	1,32E-06	0.015(0.021-0.01)	7,05E-07	0.012(0.021-0.004)	7,31E-03	0.018(0.023-0.013)	9,60E-10
cg02418294	0.016(0.022-0.01)	8,12E-07	0.022(0.029-0.015)	1,03E-08	0.016(0.027-0.005)	6,16E-03	0.02(0.026-0.014)	8,35E-09
cg26975609	0.02(0.028-0.012)	1,72E-06	0.022(0.029-0.015)	6,90E-08	0.024(0.036-0.011)	4,92E-04	0.022(0.03-0.014)	1,32E-06
cg19379522	0.017(0.023-0.01)	1,78E-06	0.019(0.026-0.012)	7,44E-07	0.014(0.024-0.003)	1,31E-02	0.019(0.025-0.012)	7,65E-08
cg00850340	0.024(0.034-0.015)	3,72E-06	0.032(0.043-0.022)	3,80E-08	0.026(0.041-0.011)	1,12E-03	0.029(0.039-0.02)	6,07E-08
cg18719198	0.027(0.037-0.016)	2,13E-06	0.037(0.049-0.025)	1,78E-08	0.036(0.058-0.014)	2,50E-03	0.03(0.042-0.018)	2,34E-06
cg16560453	0.018(0.024-0.011)	8,09E-07	0.019(0.025-0.013)	5,95E-08	0.014(0.024-0.004)	9,06E-03	0.021(0.027-0.015)	3,02E-09
cg23246706	0.026(0.036-0.016)	1,84E-06	0.03(0.041-0.019)	9,69E-07	0.024(0.042-0.006)	9,77E-03	0.028(0.038-0.018)	2,73E-07
cg07232033	0.028(0.039-0.018)	1,40E-06	0.032(0.046-0.019)	1,35E-05	0.019(0.037-0.001)	3,55E-02	0.033(0.043-0.024)	3,57E-09
cg13079099	0.018(0.023-0.012)	2,51E-08	0.018(0.025-0.011)	3,45E-06	0.018(0.029-0.007)	2,85E-03	0.017(0.023-0.011)	2,42E-07
cg13696351**	0.043(0.057-0.028)	7,81E-08	0.048(0.066-0.031)	4,44E-07	0.053(0.086-0.02)	2,58E-03	0.053(0.069-0.037)	1,12E-08
cg07565236	0.041(0.055-0.027)	8,09E-08	0.046(0.062-0.031)	9,23E-08	0.032(0.061-0.003)	3,23E-02	0.044(0.058-0.03)	5,15E-08
cg19099833	0.023(0.032-0.014)	4,27E-06	0.027(0.038-0.016)	8,08E-06	0.018(0.039-0.004)	1,00E-01	0.029(0.039-0.018)	1,52E-06
cg05705366	0.017(0.023-0.011)	7,85E-07	0.021(0.028-0.014)	7,34E-08	0.023(0.036-0.01)	1,07E-03	0.022(0.03-0.014)	5,54E-07
cg06704539	0.018(0.025-0.011)	1,72E-06	0.018(0.025-0.011)	3,88E-06	0.022(0.035-0.009)	2,09E-03	0.018(0.025-0.011)	5,14E-06
cg27523523	0.016(0.021-0.01)	1,88E-07	0.019(0.025-0.013)	1,58E-08	0.014(0.025-0.004)	7,61E-03	0.019(0.025-0.013)	5,03E-08
cg05576451	0.024(0.034-0.015)	1,52E-06	0.027(0.037-0.017)	7,44E-07	0.03(0.048-0.013)	1,36E-03	0.027(0.036-0.017)	3,59E-07
cg16706496	0.031(0.041-0.02)	1,37E-07	0.037(0.048-0.026)	4,21E-09	0.024(0.044-0.003)	2,34E-02	0.026(0.037-0.016)	4,62E-06
cg25139493	0.03(0.042-0.017)	1,29E-05	0.043(0.059-0.027)	1,42E-06	0.043(0.069-0.016)	2,69E-03	0.043(0.058-0.029)	1,28E-07
cg13895544	0.014(0.019-0.009)	4,06E-07	0.018(0.024-0.012)	1,70E-07	0.012(0.022-0.002)	1,58E-02	0.017(0.023-0.011)	3,08E-07
cg27313501	0.019(0.027-0.012)	1,58E-06	0.025(0.033-0.018)	7,08E-09	0.025(0.039-0.012)	7,23E-04	0.024(0.032-0.016)	5,76E-08
cg14971823	0.019(0.026-0.012)	2,54E-06	0.021(0.028-0.013)	1,03E-06	0.02(0.034-0.006)	6,65E-03	0.022(0.03-0.014)	3,93E-07
cg00177013	0.016(0.021-0.011)	3,47E-08	0.017(0.022-0.011)	2,47E-07	0.015(0.022-0.007)	3,62E-04	0.018(0.023-0.013)	1,26E-09
cg08655589	0.037(0.051-0.023)	8,94E-07	0.046(0.06-0.032)	1,41E-08	0.042(0.076-0.009)	1,36E-02	0.036(0.05-0.021)	5,27E-06
cg23656755	0.032(0.045-0.019)	4,20E-06	0.038(0.052-0.023)	1,80E-06	0.04(0.06-0.019)	5,07E-04	0.039(0.051-0.027)	1,06E-08
cg19513247	0.018(0.024-0.011)	9,96E-07	0.021(0.028-0.013)	5,67E-07	0.014(0.024-0.003)	1,38E-02	0.02(0.027-0.013)	2,75E-07
cg01112643	0.018(0.025-0.01)	1,79E-05	0.022(0.03-0.015)	2,75E-07	0.022(0.035-0.01)	1,00E-03	0.022(0.029-0.015)	5,37E-08
cg12870876	0.031(0.042-0.02)	3,49E-07	0.037(0.049-0.026)	1,72E-08	0.014(0.038-0.01)	2,53E-01	0.032(0.044-0.02)	1,21E-06

cg20962543	0.016(0.022-0.009)	1,71E-05	0.02(0.026-0.013)	5,14E-08	0.015(0.025-0.004)	9,41E-03	0.02(0.027-0.013)	1,73E-07
cg00799842	0.018(0.025-0.011)	4,93E-06	0.025(0.033-0.018)	1,41E-08	0.02(0.033-0.007)	4,48E-03	0.02(0.027-0.012)	2,18E-06
cg04415045	0.011(0.017-0.005)	2,92E-04	0.013(0.017-0.008)	1,66E-07	0.012(0.021-0.003)	8,90E-03	0.013(0.018-0.009)	2,53E-07
cg02590295	0.015(0.02-0.009)	1,43E-06	0.017(0.022-0.011)	1,82E-07	0.009(0.019-0.002)	1,15E-01	0.016(0.023-0.01)	3,33E-06
cg02830438	0.018(0.026-0.01)	3,30E-05	0.025(0.033-0.016)	1,39E-07	0.014(0.029-0.001)	6,42E-02	0.023(0.032-0.014)	3,49E-06
cg07428697	0.014(0.019-0.008)	7,10E-06	0.013(0.019-0.007)	1,99E-05	0.015(0.026-0.004)	8,57E-03	0.014(0.019-0.009)	1,06E-06
cg10446143	0.026(0.038-0.015)	2,42E-05	0.035(0.045-0.025)	2,98E-09	0.026(0.045-0.008)	7,04E-03	0.031(0.041-0.021)	4,84E-08
cg18644710	0.025(0.034-0.016)	7,41E-07	0.025(0.034-0.016)	5,94E-07	0.019(0.035-0.003)	2,36E-02	0.027(0.036-0.018)	6,49E-08
cg18963390	0.017(0.024-0.011)	1,49E-06	0.019(0.026-0.012)	7,84E-07	0.019(0.029-0.008)	1,21E-03	0.017(0.024-0.011)	1,49E-06
cg02182114	0.026(0.036-0.016)	3,15E-06	0.028(0.04-0.017)	3,14E-06	0.03(0.05-0.01)	4,85E-03	0.03(0.041-0.02)	2,81E-07
cg12888961	0.015(0.022-0.009)	2,46E-05	0.019(0.026-0.011)	3,63E-06	0.019(0.032-0.006)	5,72E-03	0.019(0.026-0.013)	6,47E-08
cg13716829	0.02(0.028-0.012)	6,23E-06	0.021(0.031-0.012)	1,19E-05	0.023(0.039-0.007)	7,47E-03	0.024(0.032-0.016)	2,00E-07
cg14283059	0.017(0.025-0.009)	4,33E-05	0.023(0.03-0.016)	1,34E-08	0.021(0.032-0.009)	1,09E-03	0.019(0.026-0.012)	5,97E-07
cg23453622	0.019(0.027-0.012)	2,87E-06	0.028(0.036-0.019)	1,10E-08	0.025(0.037-0.013)	1,78E-04	0.023(0.031-0.015)	2,24E-07
cg23722428	0.025(0.037-0.013)	7,04E-05	0.024(0.035-0.012)	1,06E-04	0.027(0.044-0.01)	2,77E-03	0.036(0.046-0.027)	2,06E-10
cg12970711	0.016(0.022-0.009)	1,02E-05	0.019(0.026-0.012)	1,04E-06	0.022(0.033-0.01)	4,95E-04	0.021(0.027-0.015)	4,89E-09
cg08184047**	0.018(0.025-0.011)	6,64E-06	0.02(0.028-0.013)	1,09E-06	0.021(0.034-0.008)	2,28E-03	0.018(0.025-0.011)	3,67E-06
cg05417087	0.018(0.025-0.012)	1,12E-07	0.019(0.027-0.012)	2,46E-06	0.018(0.031-0.004)	1,31E-02	0.023(0.03-0.016)	5,70E-09
cg02839273	0.02(0.028-0.011)	3,32E-05	0.028(0.038-0.019)	1,90E-07	0.012(0.027-0.004)	1,44E-01	0.026(0.034-0.018)	2,60E-08
cg02631957	0.017(0.025-0.01)	3,47E-05	0.023(0.032-0.015)	5,44E-07	0.024(0.038-0.01)	1,63E-03	0.023(0.031-0.015)	7,66E-07
cg11162385	0.022(0.03-0.013)	3,24E-06	0.029(0.037-0.02)	6,59E-09	0.031(0.047-0.015)	4,65E-04	0.023(0.032-0.014)	3,76E-06
cg01419788	0.012(0.016-0.008)	5,61E-08	0.012(0.016-0.007)	9,26E-07	0.01(0.017-0.003)	9,94E-03	0.012(0.016-0.007)	2,40E-06
cg24064055	0.018(0.024-0.012)	1,74E-07	0.017(0.024-0.01)	1,15E-05	0.017(0.028-0.006)	4,29E-03	0.02(0.026-0.013)	1,88E-07
cg08873779	0.026(0.037-0.016)	3,50E-06	0.031(0.042-0.02)	5,68E-07	0.023(0.042-0.004)	2,06E-02	0.031(0.041-0.021)	6,06E-08
cg26161725	0.019(0.027-0.012)	2,72E-06	0.021(0.028-0.014)	4,18E-08	0.012(0.023-0.001)	3,00E-02	0.019(0.027-0.011)	1,49E-05
cg21599869	0.018(0.025-0.011)	4,01E-06	0.018(0.025-0.01)	2,15E-05	0.023(0.037-0.008)	3,35E-03	0.021(0.029-0.014)	5,13E-07
cg21961583	0.019(0.026-0.012)	1,36E-06	0.018(0.025-0.01)	1,10E-05	0.019(0.031-0.006)	4,95E-03	0.021(0.028-0.013)	4,40E-07
cg10314752	0.017(0.024-0.011)	1,78E-06	0.02(0.028-0.012)	1,89E-06	0.019(0.034-0.004)	1,40E-02	0.02(0.028-0.013)	7,33E-07
cg15253587	0.015(0.02-0.01)	1,37E-07	0.016(0.022-0.01)	9,65E-07	0.024(0.033-0.014)	1,27E-05	0.016(0.021-0.011)	7,84E-08
cg00207389	0.018(0.026-0.01)	2,00E-05	0.021(0.028-0.014)	9,64E-08	0.016(0.029-0.003)	1,44E-02	0.019(0.026-0.011)	2,49E-06
cg07535605	0.015(0.021-0.01)	1,33E-06	0.017(0.023-0.011)	5,42E-07	0.023(0.034-0.012)	1,96E-04	0.016(0.022-0.009)	1,23E-05

cg03496713	0.014(0.021-0.008)	6,07E-05	0.021(0.028-0.014)	3,65E-08	0.016(0.026-0.005)	6,02E-03	0.019(0.025-0.013)	3,88E-08
cg01622965	0.013(0.02-0.007)	1,22E-04	0.022(0.028-0.015)	5,08E-09	0.017(0.027-0.007)	1,79E-03	0.018(0.024-0.012)	6,64E-08
cg25889945	0.013(0.019-0.007)	7,92E-05	0.017(0.023-0.01)	2,35E-06	0.017(0.028-0.006)	3,95E-03	0.018(0.022-0.013)	1,03E-10
cg08104527	0.018(0.026-0.011)	1,27E-05	0.023(0.03-0.015)	5,64E-08	0.013(0.024-0.001)	3,33E-02	0.02(0.027-0.012)	1,17E-06
cg14461650	0.011(0.016-0.006)	1,02E-05	0.013(0.018-0.008)	1,81E-06	0.013(0.022-0.004)	5,34E-03	0.013(0.018-0.008)	1,04E-06
cg22014600	0.042(0.056-0.028)	3,19E-08	0.05(0.066-0.034)	1,89E-08	0.054(0.085-0.024)	9,95E-04	0.036(0.052-0.02)	3,11E-05
cg03494795	0.019(0.025-0.013)	6,66E-08	0.019(0.026-0.012)	6,09E-07	0.021(0.031-0.011)	1,91E-04	0.018(0.024-0.012)	1,89E-07
cg12075970	0.023(0.032-0.014)	3,60E-06	0.028(0.039-0.017)	2,03E-06	0.021(0.039-0.002)	2,92E-02	0.024(0.035-0.013)	3,31E-05
cg24770850	0.013(0.018-0.007)	2,14E-05	0.016(0.022-0.01)	6,76E-07	0.014(0.023-0.004)	5,45E-03	0.018(0.023-0.012)	6,97E-09
cg05029189	0.022(0.031-0.012)	1,68E-05	0.032(0.043-0.021)	1,83E-07	0.031(0.047-0.014)	6,52E-04	0.032(0.042-0.023)	5,46E-09
cg11496792	0.014(0.02-0.008)	1,36E-05	0.015(0.021-0.009)	3,18E-06	0.012(0.021-0.002)	1,59E-02	0.015(0.021-0.009)	2,95E-06
cg02249648	0.023(0.032-0.014)	2,65E-06	0.029(0.038-0.02)	3,15E-08	0.028(0.045-0.011)	2,20E-03	0.027(0.035-0.018)	8,46E-08
cg06352072	0.015(0.023-0.008)	8,00E-05	0.018(0.025-0.01)	2,26E-05	0.013(0.023-0.004)	9,36E-03	0.02(0.027-0.013)	1,09E-07
cg01896926	0.03(0.043-0.018)	7,66E-06	0.035(0.048-0.022)	1,90E-06	0.024(0.051--0.004)	8,86E-02	0.032(0.046-0.018)	1,72E-05
cg26955196	0.024(0.033-0.014)	2,93E-06	0.025(0.034-0.016)	3,94E-07	0.036(0.055-0.016)	8,53E-04	0.022(0.031-0.012)	2,79E-05
cg05964635	0.013(0.018-0.008)	2,48E-06	0.015(0.02-0.009)	2,89E-06	0.014(0.023-0.005)	3,82E-03	0.013(0.018-0.008)	2,35E-06
cg13547299	0.014(0.021-0.006)	1,11E-03	0.022(0.029-0.014)	2,65E-07	0.02(0.033-0.007)	3,46E-03	0.021(0.028-0.015)	2,88E-08
cg01470744	0.023(0.033-0.013)	2,71E-05	0.034(0.044-0.024)	1,99E-09	0.021(0.04-0.003)	2,58E-02	0.032(0.042-0.023)	4,43E-09
cg11287521	0.018(0.026-0.01)	2,24E-05	0.022(0.031-0.014)	1,61E-06	0.028(0.041-0.015)	1,07E-04	0.024(0.032-0.015)	1,05E-06
cg04830146	0.013(0.018-0.008)	7,92E-06	0.017(0.023-0.011)	2,44E-07	0.009(0.019-0.001)	7,78E-02	0.016(0.021-0.01)	1,77E-07
cg10460130	0.03(0.042-0.018)	3,42E-06	0.038(0.051-0.024)	1,00E-06	0.032(0.057-0.007)	1,47E-02	0.039(0.052-0.026)	1,30E-07
cg05866713	0.01(0.014-0.006)	1,02E-06	0.011(0.016-0.007)	3,13E-06	0.005(0.015--0.005)	2,84E-01	0.01(0.014-0.005)	3,99E-05
cg08497766	0.015(0.02-0.009)	3,79E-06	0.016(0.023-0.01)	3,02E-06	0.021(0.033-0.01)	7,17E-04	0.018(0.024-0.011)	3,43E-07
cg03231960	0.029(0.041-0.017)	1,35E-05	0.036(0.047-0.024)	6,37E-08	0.022(0.041-0.003)	2,54E-02	0.029(0.04-0.018)	1,45E-06
cg09693228	0.014(0.019-0.008)	4,92E-06	0.015(0.021-0.009)	2,53E-06	0.01(0.021--0.002)	9,55E-02	0.014(0.02-0.008)	1,76E-05
cg22534627	0.02(0.028-0.013)	2,30E-07	0.029(0.037-0.02)	2,31E-09	0.023(0.039-0.008)	4,37E-03	0.024(0.034-0.015)	3,86E-06
cg01733795	0.018(0.025-0.011)	4,97E-06	0.025(0.034-0.016)	3,48E-07	0.021(0.037-0.004)	1,70E-02	0.022(0.03-0.015)	3,03E-07
cg02126117	0.02(0.029-0.01)	7,69E-05	0.028(0.037-0.019)	4,84E-08	0.021(0.033-0.009)	1,57E-03	0.024(0.034-0.015)	4,77E-06
cg25487903	0.013(0.019-0.008)	8,70E-06	0.017(0.023-0.012)	3,66E-08	0.012(0.02-0.004)	6,39E-03	0.016(0.022-0.01)	8,51E-07
cg16092956	0.024(0.032-0.016)	1,72E-07	0.032(0.042-0.022)	2,96E-08	0.014(0.032--0.004)	1,22E-01	0.025(0.034-0.016)	7,42E-07
cg05970992	0.013(0.019-0.008)	4,15E-06	0.021(0.027-0.015)	1,21E-09	0.009(0.019-0)	5,99E-02	0.016(0.021-0.01)	2,76E-07

cg16975959	0.015(0.021-0.009)	7,04E-06	0.019(0.026-0.011)	5,66E-06	0.012(0.022-0.002)	2,02E-02	0.021(0.027-0.014)	1,24E-08
cg11230222	0.01(0.014-0.006)	6,26E-06	0.01(0.015-0.005)	1,03E-04	0.01(0.018-0.003)	1,04E-02	0.012(0.016-0.008)	4,14E-07
cg03696327	0.017(0.025-0.01)	2,08E-05	0.025(0.035-0.015)	3,68E-06	0.019(0.036-0.002)	3,27E-02	0.025(0.034-0.017)	2,84E-07
cg25817927	0.01(0.016-0.005)	5,56E-04	0.015(0.021-0.009)	6,47E-06	0.014(0.026-0.002)	2,08E-02	0.017(0.023-0.012)	5,26E-08
cg15420687	0.019(0.026-0.012)	1,11E-06	0.019(0.027-0.012)	3,10E-06	0.019(0.034-0.004)	1,41E-02	0.02(0.027-0.013)	3,87E-07
cg06330324	0.009(0.013-0.004)	9,85E-05	0.009(0.013-0.005)	2,22E-05	0.012(0.019-0.005)	1,67E-03	0.01(0.014-0.006)	7,80E-06
cg15798279	0.017(0.024-0.01)	8,95E-06	0.018(0.026-0.009)	4,92E-05	0.011(0.023-0.002)	8,61E-02	0.02(0.028-0.012)	1,88E-06
cg24640588	0.022(0.031-0.013)	4,46E-06	0.026(0.035-0.016)	7,83E-07	0.027(0.044-0.009)	3,96E-03	0.025(0.034-0.017)	3,35E-07
cg17218495	0.021(0.03-0.011)	3,05E-05	0.026(0.036-0.017)	4,54E-07	0.035(0.066-0.004)	2,95E-02	0.027(0.038-0.017)	3,49E-06
cg23962478**	0.031(0.042-0.021)	9,78E-08	0.037(0.05-0.025)	1,78E-07	0.031(0.054-0.009)	8,58E-03	0.034(0.045-0.023)	6,95E-08
cg05528568	0.008(0.013-0.004)	2,71E-04	0.013(0.018-0.008)	4,01E-07	0.01(0.017-0.002)	1,71E-02	0.011(0.016-0.007)	7,18E-06
cg10702645	0.023(0.031-0.014)	8,37E-07	0.026(0.036-0.015)	4,51E-06	0.021(0.038-0.004)	1,66E-02	0.025(0.033-0.016)	4,98E-07
cg12869623	0.019(0.028-0.011)	1,89E-05	0.021(0.029-0.012)	6,46E-06	0.023(0.038-0.009)	2,93E-03	0.021(0.029-0.013)	1,52E-06
cg02114786	0.015(0.022-0.008)	2,66E-05	0.02(0.027-0.013)	3,30E-07	0.013(0.025-0.001)	3,91E-02	0.017(0.023-0.011)	2,07E-06
cg19191624	0.015(0.021-0.009)	4,08E-06	0.021(0.027-0.015)	6,35E-10	0.017(0.03-0.005)	7,75E-03	0.017(0.023-0.011)	6,88E-07
cg16368059	0.012(0.017-0.007)	4,22E-06	0.013(0.019-0.008)	2,18E-05	0.006(0.015-0.003)	1,85E-01	0.015(0.021-0.01)	1,47E-07
cg00497630	0.013(0.018-0.007)	1,10E-05	0.012(0.017-0.007)	7,66E-06	0.011(0.02-0.003)	1,03E-02	0.013(0.018-0.009)	2,04E-07
cg18001714	0.018(0.027-0.01)	5,16E-05	0.024(0.032-0.015)	3,81E-07	0.031(0.048-0.013)	1,09E-03	0.024(0.033-0.015)	2,91E-06
cg14702655	0.011(0.017-0.006)	1,05E-04	0.015(0.021-0.009)	3,01E-06	0.016(0.026-0.007)	1,59E-03	0.015(0.021-0.009)	3,64E-06
cg15739944	0.015(0.021-0.009)	7,84E-06	0.016(0.023-0.009)	2,06E-05	0.009(0.02--0.001)	6,30E-02	0.017(0.023-0.01)	1,54E-06
cg06249499	0.018(0.025-0.01)	6,28E-06	0.018(0.026-0.011)	1,79E-06	0.02(0.032-0.007)	2,73E-03	0.022(0.028-0.016)	4,09E-09
cg04297258	0.028(0.038-0.017)	1,16E-06	0.036(0.048-0.024)	7,64E-08	0.019(0.039-0.001)	5,74E-02	0.034(0.045-0.023)	3,56E-08
cg02431098	0.012(0.017-0.007)	2,71E-05	0.013(0.02-0.007)	9,96E-05	0.012(0.022-0.003)	1,25E-02	0.015(0.021-0.008)	1,78E-05
cg03415497	0.023(0.032-0.014)	5,79E-06	0.02(0.03-0.01)	1,29E-04	0.024(0.042-0.006)	1,17E-02	0.024(0.033-0.015)	2,11E-06
cg01263850	0.023(0.032-0.013)	1,26E-05	0.021(0.031-0.011)	9,29E-05	0.025(0.042-0.008)	4,97E-03	0.027(0.036-0.018)	1,55E-07
cg00759551	0.017(0.024-0.01)	5,21E-06	0.018(0.025-0.011)	3,84E-06	0.016(0.029-0.004)	1,38E-02	0.018(0.024-0.012)	3,13E-07
cg06618322	0.029(0.039-0.019)	2,93E-07	0.033(0.046-0.02)	3,29E-06	0.017(0.047-0.014)	2,72E-01	0.029(0.042-0.016)	5,26E-05
cg01637218	0.013(0.018-0.008)	1,87E-06	0.013(0.018-0.008)	8,42E-06	0.013(0.022-0.003)	1,25E-02	0.016(0.02-0.011)	1,79E-08
cg27016262	0.032(0.045-0.019)	3,97E-06	0.037(0.049-0.025)	4,85E-08	0.032(0.052-0.011)	3,93E-03	0.031(0.046-0.017)	6,12E-05
cg18120578	0.028(0.038-0.017)	6,50E-07	0.031(0.042-0.021)	1,59E-07	0.028(0.049-0.008)	8,92E-03	0.031(0.043-0.019)	1,96E-06
cg09012337	0.014(0.019-0.008)	1,02E-05	0.017(0.023-0.011)	1,40E-06	0.014(0.023-0.005)	3,12E-03	0.015(0.02-0.01)	5,76E-07

cg02683985	0.014(0.022-0.007)	2,52E-04	0.022(0.03-0.013)	2,10E-06	0.02(0.032-0.007)	2,63E-03	0.021(0.028-0.013)	6,42E-07
cg06951969	0.013(0.02-0.007)	1,49E-04	0.022(0.029-0.015)	2,45E-08	0.012(0.022-0.003)	9,86E-03	0.02(0.026-0.014)	1,16E-09
cg08363925	0.027(0.037-0.018)	3,34E-07	0.033(0.044-0.022)	1,67E-07	0.023(0.041-0.005)	1,59E-02	0.027(0.037-0.017)	1,35E-06
cg01943813	0.023(0.032-0.013)	1,26E-05	0.024(0.034-0.014)	6,69E-06	0.036(0.056-0.015)	1,19E-03	0.024(0.034-0.015)	2,18E-06
cg01025283	0.016(0.023-0.009)	1,37E-05	0.019(0.026-0.012)	5,86E-07	0.016(0.028-0.003)	1,59E-02	0.017(0.024-0.011)	3,25E-06
cg13865502	0.021(0.031-0.01)	1,54E-04	0.024(0.036-0.012)	1,33E-04	0.021(0.04-0.002)	3,48E-02	0.03(0.039-0.02)	4,51E-08
cg19044674	0.017(0.023-0.01)	1,93E-06	0.019(0.026-0.011)	2,66E-06	0.019(0.033-0.004)	1,62E-02	0.019(0.026-0.012)	2,08E-06
cg00787661	0.016(0.023-0.009)	3,29E-05	0.022(0.03-0.014)	3,68E-07	0.015(0.028-0.001)	3,67E-02	0.017(0.025-0.01)	3,13E-05
cg04355250**	0.015(0.02-0.009)	2,96E-06	0.017(0.023-0.01)	2,48E-06	0.016(0.025-0.008)	7,11E-04	0.014(0.02-0.009)	3,15E-06
cg15584531	0.017(0.023-0.01)	3,63E-06	0.018(0.026-0.01)	2,00E-05	0.027(0.039-0.014)	1,36E-04	0.018(0.025-0.011)	2,64E-06
cg26732930	0.016(0.022-0.009)	4,27E-06	0.022(0.029-0.015)	2,26E-08	0.014(0.024-0.005)	4,29E-03	0.018(0.023-0.012)	3,50E-08
cg07515367	0.022(0.032-0.012)	4,83E-05	0.028(0.038-0.019)	2,18E-07	0.019(0.037-0.001)	3,91E-02	0.026(0.037-0.016)	3,28E-06
cg07753241	0.022(0.032-0.012)	3,40E-05	0.03(0.041-0.02)	3,36E-07	0.015(0.03-0)	5,71E-02	0.029(0.038-0.02)	3,46E-08
cg20184917	0.015(0.021-0.01)	1,03E-06	0.019(0.025-0.012)	3,66E-07	0.021(0.031-0.011)	2,37E-04	0.021(0.028-0.014)	1,53E-07
cg21196132	0.022(0.031-0.013)	3,76E-06	0.022(0.032-0.013)	1,29E-05	0.022(0.039-0.005)	1,42E-02	0.025(0.034-0.016)	1,06E-06
cg01190566	0.013(0.017-0.008)	9,58E-07	0.014(0.02-0.007)	4,11E-05	0.011(0.02-0.002)	1,78E-02	0.014(0.019-0.008)	8,77E-06
cg08005943	0.017(0.023-0.011)	7,47E-07	0.018(0.024-0.011)	1,47E-06	0.017(0.026-0.007)	1,30E-03	0.016(0.022-0.01)	1,31E-06
cg04357830	0.021(0.03-0.013)	5,92E-06	0.028(0.038-0.018)	8,56E-07	0.011(0.028-0.007)	2,15E-01	0.024(0.033-0.014)	5,58E-06
cg08264704	0.013(0.019-0.008)	8,14E-06	0.014(0.02-0.008)	1,39E-05	0.01(0.02-0)	6,15E-02	0.013(0.019-0.007)	2,48E-05
cg02994863	0.039(0.054-0.024)	2,36E-06	0.047(0.063-0.031)	8,28E-08	0.036(0.061-0.01)	7,51E-03	0.04(0.055-0.025)	1,37E-06
cg06210969	0.012(0.017-0.006)	4,81E-05	0.015(0.021-0.01)	2,31E-06	0.011(0.02-0.002)	1,75E-02	0.016(0.022-0.011)	5,00E-08
cg15008401	0.015(0.021-0.009)	5,19E-06	0.019(0.026-0.012)	1,35E-06	0.016(0.027-0.005)	5,93E-03	0.017(0.023-0.011)	4,32E-07
cg08903333	0.01(0.015-0.005)	1,29E-04	0.014(0.02-0.008)	2,78E-05	0.016(0.026-0.005)	4,10E-03	0.014(0.019-0.009)	4,32E-07
cg24877195	0.014(0.021-0.007)	1,08E-04	0.018(0.025-0.01)	1,12E-05	0.014(0.027-0)	5,13E-02	0.022(0.03-0.015)	2,50E-07
cg06066711	0.011(0.017-0.005)	4,28E-04	0.019(0.026-0.012)	1,70E-06	0.013(0.022-0.004)	8,32E-03	0.016(0.023-0.009)	1,45E-05
cg17906179	0.015(0.022-0.008)	8,14E-05	0.026(0.034-0.018)	8,49E-09	0.013(0.024-0.001)	3,83E-02	0.021(0.028-0.013)	7,85E-07
cg24252614	0.021(0.029-0.013)	2,46E-06	0.029(0.037-0.021)	1,19E-09	0.024(0.039-0.009)	2,34E-03	0.024(0.032-0.017)	4,31E-08
cg26154342	0.018(0.024-0.011)	2,13E-06	0.016(0.023-0.009)	1,58E-05	0.013(0.027-0)	5,76E-02	0.022(0.029-0.015)	3,10E-08
cg14397361	0.015(0.021-0.01)	9,02E-07	0.015(0.021-0.009)	1,04E-05	0.018(0.03-0.005)	6,38E-03	0.015(0.021-0.009)	4,02E-06
cg19373170	0.013(0.019-0.008)	8,15E-06	0.015(0.021-0.008)	1,46E-05	0.018(0.028-0.008)	1,09E-03	0.015(0.021-0.009)	4,55E-06
cg12423398	0.009(0.013-0.005)	1,50E-05	0.009(0.013-0.004)	1,24E-04	0.008(0.015-0.002)	1,32E-02	0.011(0.015-0.007)	1,80E-07

cg10107929	0.014(0.02-0.007)	7,67E-05	0.016(0.022-0.009)	1,51E-05	0.015(0.026-0.004)	1,16E-02	0.02(0.027-0.012)	1,40E-06
cg20950465	0.021(0.031-0.012)	3,47E-05	0.02(0.03-0.01)	1,46E-04	0.024(0.039-0.008)	4,56E-03	0.028(0.038-0.018)	6,18E-07
cg09349094	0.013(0.019-0.007)	3,00E-05	0.013(0.019-0.007)	3,50E-05	0.012(0.022-0.003)	1,42E-02	0.014(0.02-0.009)	6,34E-06
cg02750385	0.017(0.025-0.01)	1,60E-05	0.023(0.031-0.014)	7,29E-07	0.023(0.037-0.009)	2,16E-03	0.018(0.027-0.009)	8,01E-05
cg09055943	0.015(0.022-0.009)	1,10E-05	0.019(0.026-0.011)	2,12E-06	0.019(0.031-0.006)	4,35E-03	0.016(0.023-0.01)	5,27E-06
cg07946458	0.016(0.023-0.01)	1,62E-06	0.019(0.027-0.012)	1,22E-06	0.016(0.029-0.003)	2,11E-02	0.019(0.026-0.013)	4,39E-07
cg27007358	0.018(0.025-0.011)	5,57E-06	0.017(0.024-0.01)	1,68E-05	0.021(0.032-0.009)	1,20E-03	0.018(0.026-0.011)	3,38E-06
cg16869008	0.016(0.022-0.01)	1,96E-06	0.016(0.023-0.009)	1,85E-05	0.014(0.024-0.003)	1,53E-02	0.02(0.026-0.013)	3,48E-07
cg10239074	0.009(0.013-0.006)	7,64E-06	0.011(0.016-0.006)	1,10E-05	0.005(0.013--0.004)	2,57E-01	0.011(0.015-0.006)	4,66E-06
cg06180363	0.013(0.02-0.007)	1,10E-04	0.021(0.028-0.014)	1,33E-07	0.017(0.032-0.001)	3,74E-02	0.018(0.025-0.011)	1,35E-06
cg04879696	0.014(0.02-0.009)	1,27E-06	0.016(0.022-0.01)	1,83E-06	0.008(0.02--0.003)	1,59E-01	0.014(0.02-0.009)	2,88E-06
cg25547361	0.014(0.02-0.008)	1,44E-05	0.016(0.022-0.009)	1,33E-05	0.02(0.032-0.008)	1,54E-03	0.014(0.02-0.008)	1,01E-05
cg16697493	0.02(0.03-0.011)	1,04E-04	0.023(0.033-0.013)	1,14E-05	0.024(0.039-0.009)	3,44E-03	0.026(0.036-0.016)	1,54E-06
cg01854842	0.016(0.023-0.008)	6,78E-05	0.016(0.024-0.009)	3,49E-05	0.022(0.033-0.01)	7,62E-04	0.021(0.029-0.014)	3,29E-07
cg10531748	0.014(0.02-0.007)	3,86E-05	0.014(0.02-0.007)	3,98E-05	0.012(0.024-0)	4,25E-02	0.016(0.022-0.01)	1,55E-06
cg27297221	0.006(0.009-0.003)	1,26E-04	0.007(0.011-0.004)	1,07E-04	0.004(0.011--0.002)	1,70E-01	0.009(0.012-0.005)	1,93E-06
cg06534422	0.011(0.016-0.006)	1,00E-04	0.012(0.018-0.007)	6,17E-05	0.013(0.021-0.006)	1,23E-03	0.014(0.02-0.009)	2,34E-06
cg17361641	0.017(0.025-0.01)	9,39E-06	0.025(0.033-0.016)	1,24E-07	0.011(0.025--0.002)	9,07E-02	0.022(0.029-0.014)	3,26E-07
cg12094402	0.018(0.025-0.011)	2,38E-06	0.019(0.026-0.012)	7,99E-07	0.015(0.026-0.004)	7,08E-03	0.017(0.025-0.009)	5,60E-05
cg17948846	0.017(0.025-0.009)	5,70E-05	0.024(0.033-0.015)	9,45E-07	0.017(0.029-0.005)	5,98E-03	0.021(0.029-0.013)	1,25E-06
cg15814736	0.02(0.03-0.01)	1,02E-04	0.022(0.032-0.012)	3,00E-05	0.03(0.047-0.012)	1,60E-03	0.025(0.035-0.015)	4,65E-06
cg23680936	0.008(0.011-0.004)	3,38E-05	0.01(0.013-0.006)	9,74E-08	0.008(0.012-0.003)	1,34E-03	0.01(0.014-0.006)	6,39E-06
cg04889973	0.019(0.026-0.012)	1,18E-06	0.022(0.031-0.013)	9,93E-06	0.015(0.031-0)	4,72E-02	0.022(0.03-0.015)	1,73E-07
cg01264747	0.016(0.023-0.008)	1,45E-04	0.022(0.029-0.014)	6,90E-07	0.013(0.027--0.001)	7,62E-02	0.02(0.027-0.013)	2,06E-07
cg01513081	0.016(0.025-0.008)	2,59E-04	0.021(0.029-0.014)	4,34E-07	0.021(0.04-0.002)	2,84E-02	0.022(0.03-0.015)	2,33E-07
cg02731774	0.01(0.014-0.006)	5,12E-06	0.011(0.015-0.007)	9,30E-07	0.009(0.016-0.003)	5,99E-03	0.01(0.015-0.006)	1,81E-05
cg08514895	0.01(0.015-0.006)	4,64E-06	0.013(0.018-0.008)	1,67E-06	0.01(0.018-0.002)	2,10E-02	0.012(0.016-0.008)	3,20E-07
cg13311803	0.017(0.025-0.01)	1,23E-05	0.023(0.031-0.015)	2,56E-07	0.016(0.029-0.002)	2,55E-02	0.022(0.03-0.014)	7,11E-07
cg04593445	0.012(0.019-0.005)	7,22E-04	0.018(0.025-0.011)	2,80E-06	0.012(0.024--0.001)	6,01E-02	0.016(0.023-0.01)	5,53E-06
cg21740452	0.012(0.018-0.006)	1,74E-04	0.017(0.023-0.011)	1,03E-07	0.015(0.024-0.006)	2,26E-03	0.014(0.02-0.009)	4,16E-06
cg09864858	0.016(0.023-0.009)	1,58E-05	0.02(0.028-0.013)	1,10E-06	0.013(0.026-0.001)	3,29E-02	0.018(0.025-0.011)	6,80E-06

cg09795194	0.009(0.013-0.006)	3,69E-07	0.009(0.013-0.005)	8,48E-06	0.008(0.016-0.001)	3,35E-02	0.01(0.013-0.006)	2,69E-06
cg19906131	0.021(0.029-0.014)	3,08E-07	0.019(0.027-0.012)	4,02E-06	0.03(0.045-0.016)	1,60E-04	0.021(0.029-0.013)	3,03E-06
cg24059718	0.006(0.008-0.003)	4,62E-06	0.005(0.007-0.002)	3,51E-04	0.004(0.008-0)	3,37E-02	0.006(0.008-0.003)	6,08E-06
cg05324982	0.014(0.021-0.006)	5,61E-04	0.021(0.028-0.014)	2,05E-07	0.011(0.023-0)	5,79E-02	0.017(0.023-0.011)	1,29E-06
cg07706463	0.013(0.019-0.007)	2,81E-05	0.013(0.02-0.006)	2,54E-04	0.015(0.025-0.005)	6,44E-03	0.017(0.023-0.012)	7,54E-08
cg00784718	0.023(0.034-0.013)	2,89E-05	0.032(0.043-0.022)	4,72E-08	0.028(0.044-0.011)	1,90E-03	0.026(0.036-0.016)	4,73E-06
cg02108176	0.015(0.022-0.009)	1,42E-05	0.021(0.028-0.014)	8,98E-08	0.012(0.022-0.003)	1,48E-02	0.017(0.023-0.011)	1,11E-06
cg05914712	0.005(0.007-0.003)	3,88E-05	0.006(0.009-0.003)	3,04E-05	0.006(0.011-0.002)	1,18E-02	0.006(0.008-0.003)	8,67E-06
cg04249066	0.008(0.011-0.004)	5,76E-06	0.009(0.013-0.005)	1,03E-05	0.007(0.013-0.001)	1,66E-02	0.008(0.011-0.005)	1,25E-06
cg05256242	0.017(0.024-0.011)	2,82E-06	0.021(0.028-0.013)	2,04E-06	0.018(0.029-0.006)	3,39E-03	0.017(0.025-0.01)	1,85E-05
cg23758016	0.025(0.036-0.014)	2,32E-05	0.034(0.046-0.021)	2,05E-06	0.022(0.043-0.001)	3,66E-02	0.029(0.04-0.017)	5,45E-06
cg11017269	0.023(0.033-0.013)	1,71E-05	0.02(0.031-0.01)	2,36E-04	0.026(0.05-0.002)	3,21E-02	0.024(0.034-0.014)	7,80E-06
cg01398415	0.013(0.02-0.007)	1,63E-04	0.02(0.027-0.012)	3,07E-06	0.018(0.029-0.007)	2,05E-03	0.017(0.023-0.01)	8,76E-06
cg23278060	0.008(0.01-0.005)	2,64E-07	0.005(0.008-0.003)	2,99E-04	0.008(0.013-0.003)	1,58E-03	0.007(0.009-0.004)	1,24E-05
cg25872752	0.016(0.025-0.008)	2,66E-04	0.023(0.031-0.014)	7,08E-07	0.024(0.036-0.012)	2,71E-04	0.022(0.031-0.013)	3,39E-06
cg05064002	0.011(0.016-0.006)	1,85E-05	0.013(0.018-0.007)	9,47E-06	0.011(0.019-0.004)	3,88E-03	0.013(0.018-0.007)	9,06E-06
cg10101470	0.012(0.016-0.007)	1,66E-06	0.013(0.018-0.008)	6,79E-07	0.016(0.026-0.006)	3,39E-03	0.012(0.016-0.007)	9,44E-06
cg20350269	0.012(0.017-0.007)	2,16E-05	0.01(0.016-0.004)	8,67E-04	0.01(0.017-0.003)	8,91E-03	0.012(0.017-0.008)	5,01E-07
cg27425263	0.021(0.029-0.013)	1,68E-06	0.029(0.039-0.019)	1,60E-07	0.018(0.035-0.002)	2,92E-02	0.026(0.035-0.017)	1,63E-07
cg25820220	0.015(0.022-0.008)	1,10E-04	0.014(0.021-0.008)	4,95E-05	0.023(0.033-0.013)	5,15E-05	0.017(0.023-0.011)	8,15E-07
cg24594678	0.012(0.017-0.007)	1,28E-05	0.012(0.018-0.006)	5,69E-05	0.013(0.02-0.006)	6,79E-04	0.011(0.017-0.006)	6,72E-05
cg08684830	0.011(0.017-0.005)	6,76E-04	0.016(0.023-0.01)	5,57E-06	0.014(0.026-0.002)	2,48E-02	0.015(0.022-0.009)	1,40E-05
cg24851862	0.011(0.015-0.006)	1,61E-05	0.014(0.019-0.008)	1,68E-05	0.007(0.015-0.001)	1,03E-01	0.012(0.016-0.007)	8,59E-07
cg07986525	0.015(0.021-0.009)	2,77E-06	0.018(0.024-0.011)	1,22E-06	0.009(0.019-0.001)	8,32E-02	0.016(0.022-0.01)	8,85E-07
cg23572376	0.018(0.025-0.01)	8,06E-06	0.019(0.026-0.011)	2,74E-06	0.023(0.035-0.011)	4,86E-04	0.021(0.028-0.014)	1,43E-07
cg24881159	0.013(0.019-0.008)	5,96E-06	0.014(0.019-0.008)	5,94E-06	0.023(0.033-0.012)	9,38E-05	0.015(0.021-0.009)	2,98E-06
cg17364044	0.013(0.019-0.006)	1,33E-04	0.019(0.025-0.012)	4,10E-07	0.013(0.026-0.001)	3,43E-02	0.02(0.026-0.013)	5,61E-08
cg03648361	0.014(0.02-0.008)	7,28E-06	0.016(0.023-0.009)	1,13E-05	0.018(0.029-0.007)	2,72E-03	0.019(0.026-0.012)	6,77E-07
cg08521010	0.015(0.022-0.009)	1,09E-05	0.02(0.028-0.012)	1,66E-06	0.011(0.023-0.001)	6,87E-02	0.019(0.026-0.012)	5,57E-07
cg17429424	0.015(0.02-0.01)	1,04E-07	0.017(0.023-0.012)	9,07E-08	0.007(0.018-0.004)	2,08E-01	0.015(0.02-0.01)	3,12E-07
cg24817430	0.021(0.03-0.012)	2,12E-05	0.026(0.037-0.015)	1,24E-05	0.023(0.037-0.008)	3,85E-03	0.023(0.033-0.012)	4,38E-05

cg04573959	0.019(0.028-0.01)	4,77E-05	0.02(0.029-0.011)	1,89E-05	0.021(0.033-0.009)	1,10E-03	0.02(0.027-0.013)	5,20E-07
cg08083016	0.016(0.023-0.009)	1,77E-05	0.014(0.023-0.005)	2,73E-03	0.019(0.032-0.007)	3,92E-03	0.018(0.025-0.01)	2,46E-05
cg26986180	0.01(0.015-0.005)	1,92E-04	0.014(0.019-0.008)	5,58E-06	0.012(0.022-0.003)	1,48E-02	0.014(0.019-0.008)	3,59E-06
cg07202430	0.038(0.052-0.024)	1,08E-06	0.04(0.056-0.024)	5,91E-06	0.043(0.07-0.017)	2,40E-03	0.042(0.057-0.027)	8,27E-07
cg17930183	0.025(0.037-0.013)	1,56E-04	0.031(0.045-0.018)	1,59E-05	0.029(0.054-0.003)	2,72E-02	0.035(0.048-0.022)	1,42E-06
cg06648971	0.005(0.008-0.002)	1,82E-03	0.007(0.01-0.004)	9,19E-05	0.004(0.01--0.001)	1,05E-01	0.006(0.009-0.004)	7,63E-06
cg21442773	0.012(0.018-0.007)	2,15E-05	0.014(0.02-0.008)	6,03E-06	0.015(0.025-0.005)	4,02E-03	0.013(0.019-0.008)	1,16E-05
cg20488697	0.016(0.022-0.01)	4,00E-07	0.017(0.025-0.01)	1,11E-05	0.008(0.02--0.004)	1,72E-01	0.017(0.023-0.01)	2,23E-06
cg11343870	0.025(0.035-0.015)	2,47E-06	0.031(0.041-0.02)	1,81E-07	0.014(0.029--0.001)	7,22E-02	0.029(0.04-0.019)	5,13E-07
cg13731828	0.007(0.01-0.003)	6,20E-04	0.008(0.012-0.005)	1,65E-05	0.006(0.011-0)	3,99E-02	0.009(0.012-0.006)	2,07E-07
cg12129103	0.017(0.024-0.009)	1,58E-05	0.023(0.032-0.013)	7,30E-06	0.014(0.025-0.002)	1,98E-02	0.02(0.027-0.013)	4,61E-07
cg16043190	0.016(0.024-0.008)	2,32E-04	0.022(0.03-0.014)	1,37E-06	0.016(0.031-0.002)	2,91E-02	0.02(0.029-0.012)	5,86E-06
cg05052194	0.018(0.025-0.01)	1,73E-05	0.026(0.036-0.016)	2,09E-06	0.009(0.022--0.003)	1,36E-01	0.024(0.033-0.016)	4,82E-07
cg08393828	0.014(0.021-0.007)	6,42E-05	0.013(0.02-0.007)	1,34E-04	0.009(0.02--0.001)	8,10E-02	0.017(0.023-0.011)	2,89E-07
cg19656070	0.024(0.035-0.013)	2,12E-05	0.029(0.041-0.017)	7,86E-06	0.024(0.049--0.001)	5,83E-02	0.027(0.039-0.015)	2,14E-05
cg10729531	0.013(0.019-0.006)	3,28E-04	0.017(0.025-0.01)	8,38E-06	0.008(0.018--0.003)	1,47E-01	0.017(0.024-0.01)	1,62E-05
cg01698934	0.005(0.007-0.003)	2,94E-05	0.005(0.007-0.002)	1,83E-04	0.007(0.011-0.003)	5,15E-04	0.005(0.007-0.003)	2,52E-05
cg24657817	0.01(0.015-0.005)	2,38E-04	0.013(0.018-0.008)	4,38E-06	0.01(0.019-0.001)	2,64E-02	0.012(0.017-0.007)	6,25E-06
cg03183540	0.036(0.048-0.023)	2,04E-07	0.025(0.038-0.012)	3,43E-04	0.031(0.05-0.012)	2,50E-03	0.03(0.042-0.018)	4,27E-06
cg01305736	0.017(0.025-0.01)	2,04E-05	0.018(0.025-0.01)	1,12E-05	0.019(0.033-0.006)	5,82E-03	0.02(0.027-0.013)	1,68E-07
cg09951048	0.007(0.01-0.004)	5,36E-05	0.009(0.012-0.005)	1,36E-05	0.005(0.012--0.003)	1,88E-01	0.008(0.011-0.005)	2,20E-05
cg07409674	0.014(0.02-0.007)	1,11E-04	0.016(0.023-0.01)	6,81E-06	0.018(0.029-0.008)	1,31E-03	0.018(0.024-0.012)	1,02E-07
cg26087678	0.021(0.029-0.013)	3,34E-06	0.025(0.034-0.016)	7,21E-07	0.018(0.033-0.002)	2,83E-02	0.021(0.031-0.011)	5,67E-05
cg15854067	0.015(0.022-0.008)	1,19E-04	0.023(0.031-0.015)	1,87E-07	0.017(0.03-0.003)	2,04E-02	0.021(0.028-0.014)	3,97E-07
cg01059126	0.014(0.02-0.007)	1,29E-04	0.017(0.025-0.01)	7,77E-06	0.01(0.021-0)	6,06E-02	0.017(0.024-0.01)	1,42E-05
cg10484485	0.015(0.022-0.009)	1,94E-05	0.02(0.029-0.012)	8,31E-06	0.018(0.033-0.002)	2,83E-02	0.018(0.026-0.01)	1,68E-05
cg13309628	0.012(0.018-0.006)	1,21E-04	0.017(0.023-0.011)	5,33E-07	0.012(0.022-0.002)	1,84E-02	0.017(0.023-0.011)	3,96E-07
cg01372366	0.03(0.042-0.019)	8,66E-07	0.026(0.038-0.014)	7,64E-05	0.025(0.043-0.007)	7,68E-03	0.028(0.04-0.016)	2,53E-05
cg03628682	0.013(0.019-0.008)	1,48E-05	0.019(0.025-0.012)	1,36E-07	0.014(0.025-0.003)	1,34E-02	0.017(0.023-0.011)	4,73E-07
cg02318426	0.01(0.014-0.006)	2,46E-05	0.014(0.02-0.008)	2,73E-05	0.013(0.024-0.003)	1,23E-02	0.01(0.016-0.005)	1,38E-04
cg06642941	0.009(0.013-0.005)	1,14E-05	0.007(0.012-0.003)	1,62E-03	0.011(0.016-0.006)	1,47E-04	0.011(0.015-0.007)	7,12E-07

cg02625641	0.022(0.032-0.012)	4,26E-05	0.028(0.039-0.018)	1,19E-06	0.031(0.054-0.007)	1,18E-02	0.025(0.036-0.014)	2,55E-05
cg11882432	0.012(0.018-0.006)	2,93E-04	0.018(0.025-0.011)	4,64E-06	0.025(0.038-0.012)	3,73E-04	0.019(0.026-0.012)	8,44E-07
cg24778909	0.016(0.024-0.008)	7,39E-05	0.023(0.032-0.014)	2,74E-06	0.022(0.035-0.01)	8,99E-04	0.02(0.028-0.012)	3,34E-06
cg06959182	0.005(0.008-0.003)	1,01E-04	0.005(0.009-0.002)	1,47E-03	0.007(0.011-0.003)	2,71E-03	0.006(0.008-0.003)	8,92E-06
cg08573313	0.014(0.022-0.007)	1,70E-04	0.022(0.03-0.013)	2,39E-06	0.018(0.032-0.005)	1,05E-02	0.022(0.03-0.014)	3,10E-07
cg18770186	0.016(0.025-0.007)	4,96E-04	0.022(0.031-0.012)	1,46E-05	0.024(0.041-0.008)	5,08E-03	0.022(0.032-0.013)	8,20E-06
cg01141237	0.02(0.028-0.012)	5,00E-06	0.022(0.032-0.012)	6,58E-05	0.029(0.046-0.011)	2,04E-03	0.02(0.03-0.01)	1,74E-04
cg26283532	0.012(0.016-0.007)	7,16E-07	0.009(0.015-0.004)	5,11E-04	0.009(0.016-0.003)	7,48E-03	0.012(0.016-0.007)	3,20E-06
cg24721777	0.017(0.024-0.01)	5,83E-06	0.017(0.026-0.008)	3,48E-04	0.019(0.033-0.006)	6,02E-03	0.023(0.03-0.016)	2,41E-08
cg01631325	0.025(0.034-0.016)	5,34E-07	0.025(0.035-0.014)	1,23E-05	0.024(0.041-0.007)	6,47E-03	0.021(0.031-0.012)	3,58E-05
cg03449631	0.018(0.027-0.01)	5,72E-05	0.024(0.034-0.015)	4,36E-06	0.026(0.041-0.011)	1,43E-03	0.025(0.033-0.017)	7,01E-08
cg02154956	0.007(0.01-0.004)	1,59E-05	0.008(0.011-0.005)	9,50E-06	0.002(0.008-0.003)	4,01E-01	0.006(0.01-0.003)	9,25E-04
cg04147049	0.009(0.014-0.005)	1,89E-04	0.013(0.018-0.007)	1,07E-05	0.008(0.016-0)	5,73E-02	0.013(0.017-0.008)	8,32E-07
cg24662718	0.01(0.014-0.005)	3,46E-05	0.012(0.017-0.007)	1,55E-05	0.011(0.017-0.004)	1,59E-03	0.014(0.019-0.009)	1,48E-06
cg00516597	0.012(0.017-0.007)	3,50E-05	0.017(0.023-0.011)	3,47E-07	0.013(0.023-0.004)	7,40E-03	0.012(0.018-0.006)	7,53E-05
cg09597465	0.011(0.018-0.005)	4,80E-04	0.02(0.027-0.012)	1,43E-06	0.011(0.022-0.001)	6,58E-02	0.021(0.028-0.015)	3,16E-08
cg14642298	0.019(0.028-0.011)	4,43E-05	0.027(0.037-0.017)	7,69E-07	0.029(0.049-0.01)	4,42E-03	0.023(0.033-0.014)	1,01E-05
cg21305245	0.012(0.018-0.007)	2,46E-05	0.014(0.02-0.008)	1,14E-05	0.015(0.023-0.008)	4,27E-04	0.014(0.018-0.009)	2,56E-07
cg06499637	0.012(0.018-0.006)	1,10E-04	0.016(0.022-0.01)	2,04E-06	0.013(0.023-0.003)	1,33E-02	0.014(0.02-0.007)	4,98E-05
cg10640371	0.013(0.019-0.006)	3,50E-04	0.016(0.022-0.01)	3,33E-06	0.015(0.025-0.004)	8,11E-03	0.016(0.023-0.01)	3,20E-06
cg23064465	0.018(0.025-0.011)	5,22E-06	0.023(0.032-0.015)	7,71E-07	0.016(0.031-0.001)	3,70E-02	0.019(0.027-0.011)	1,08E-05
cg15407517	0.004(0.006-0.003)	3,55E-06	0.004(0.006-0.002)	4,59E-05	0.004(0.007-0.001)	2,57E-02	0.005(0.007-0.003)	1,34E-06
cg13251669	0.006(0.009-0.003)	3,00E-04	0.008(0.011-0.004)	1,84E-05	0.008(0.012-0.003)	1,69E-03	0.006(0.009-0.003)	2,73E-05
cg07806164	0.017(0.025-0.008)	1,19E-04	0.02(0.028-0.012)	3,11E-06	0.025(0.039-0.011)	1,33E-03	0.022(0.029-0.016)	5,46E-09
cg17547708	0.012(0.019-0.006)	4,19E-04	0.014(0.021-0.008)	5,33E-05	0.016(0.027-0.006)	3,05E-03	0.017(0.023-0.011)	1,19E-06
cg06849077	0.014(0.02-0.008)	2,58E-05	0.015(0.021-0.009)	7,55E-06	0.013(0.025-0.001)	3,20E-02	0.018(0.025-0.012)	2,32E-07
cg22552684	0.011(0.016-0.006)	4,63E-05	0.011(0.016-0.005)	1,34E-04	0.008(0.017-0.001)	6,94E-02	0.012(0.016-0.007)	4,55E-06
cg00636641	0.016(0.023-0.009)	2,35E-05	0.018(0.026-0.009)	8,27E-05	0.016(0.028-0.004)	1,01E-02	0.019(0.027-0.012)	2,98E-06
cg00533923	0.016(0.023-0.009)	1,24E-05	0.014(0.02-0.007)	1,01E-04	0.022(0.033-0.012)	1,85E-04	0.014(0.022-0.006)	8,35E-04
cg05965902	0.015(0.021-0.009)	3,68E-06	0.011(0.017-0.004)	1,30E-03	0.015(0.026-0.004)	7,08E-03	0.016(0.021-0.01)	1,90E-07
cg01741878	0.013(0.019-0.008)	5,17E-06	0.015(0.022-0.009)	1,57E-05	0.012(0.022-0.002)	2,07E-02	0.014(0.02-0.009)	1,62E-06

cg14053749	0.013(0.018-0.007)	1,19E-05	0.015(0.021-0.009)	1,44E-05	0.012(0.025-0)	4,67E-02	0.016(0.022-0.009)	3,82E-06
cg03890664	0.014(0.02-0.007)	7,14E-05	0.018(0.025-0.011)	1,48E-06	0.009(0.022--0.004)	1,78E-01	0.015(0.022-0.009)	1,54E-05
cg01490821	0.013(0.018-0.008)	6,60E-06	0.012(0.019-0.006)	3,68E-04	0.015(0.026-0.004)	8,14E-03	0.013(0.019-0.008)	1,10E-05
cg08579070	0.016(0.023-0.008)	5,61E-05	0.02(0.027-0.013)	7,80E-07	0.019(0.029-0.009)	5,01E-04	0.018(0.026-0.01)	1,71E-05
cg18507129	0.011(0.016-0.007)	1,51E-06	0.01(0.017-0.004)	1,37E-03	0.011(0.021-0.001)	3,35E-02	0.013(0.019-0.008)	1,58E-05
cg02756545	0.015(0.022-0.008)	6,48E-05	0.019(0.027-0.011)	6,16E-06	0.013(0.027--0.001)	7,56E-02	0.019(0.026-0.013)	2,30E-07
cg12996305	0.019(0.027-0.011)	1,60E-05	0.018(0.027-0.009)	1,34E-04	0.026(0.04-0.011)	9,14E-04	0.022(0.03-0.013)	5,03E-06
cg01134297	0.012(0.017-0.007)	4,07E-05	0.014(0.019-0.008)	2,67E-06	0.019(0.029-0.009)	4,43E-04	0.013(0.018-0.007)	2,61E-05
cg26864061	0.014(0.02-0.007)	5,38E-05	0.016(0.024-0.009)	5,24E-05	0.016(0.029-0.003)	1,50E-02	0.014(0.021-0.008)	7,54E-05
cg11769606	0.018(0.025-0.01)	7,27E-06	0.022(0.03-0.013)	2,20E-06	0.016(0.03-0.003)	2,04E-02	0.018(0.026-0.011)	8,31E-06
cg22093503	0.011(0.016-0.006)	1,04E-04	0.014(0.019-0.009)	6,42E-07	0.01(0.017-0.002)	1,11E-02	0.014(0.019-0.009)	1,05E-07
cg26856583	0.016(0.023-0.008)	7,94E-05	0.018(0.027-0.008)	4,65E-04	0.018(0.032-0.004)	1,54E-02	0.018(0.027-0.01)	5,88E-05
cg24341177	0.014(0.019-0.009)	2,44E-07	0.018(0.024-0.012)	3,16E-07	0.014(0.024-0.004)	1,02E-02	0.015(0.021-0.009)	5,00E-06
cg08767710	0.015(0.023-0.007)	2,81E-04	0.022(0.031-0.013)	3,33E-06	0.018(0.033-0.003)	2,33E-02	0.02(0.029-0.012)	1,38E-05
cg04986015	0.013(0.018-0.007)	6,77E-05	0.015(0.02-0.009)	1,63E-06	0.012(0.022-0.003)	1,41E-02	0.016(0.021-0.011)	8,88E-09
cg07944287	0.014(0.02-0.009)	4,15E-06	0.018(0.025-0.012)	5,07E-07	0.021(0.032-0.01)	4,09E-04	0.017(0.024-0.011)	1,86E-06
cg24463309	0.014(0.021-0.007)	1,02E-04	0.019(0.027-0.01)	7,33E-05	0.018(0.031-0.006)	5,84E-03	0.017(0.024-0.011)	3,43E-06
cg09011608	0.01(0.015-0.006)	3,49E-05	0.013(0.019-0.007)	1,61E-05	0.01(0.019-0.002)	1,77E-02	0.014(0.019-0.01)	4,81E-08
cg21172814	0.01(0.014-0.005)	4,93E-05	0.013(0.018-0.008)	3,01E-06	0.011(0.018-0.004)	4,10E-03	0.011(0.015-0.007)	4,14E-06
cg09874752	0.019(0.027-0.011)	1,17E-05	0.021(0.031-0.011)	7,81E-05	0.018(0.031-0.005)	8,94E-03	0.02(0.028-0.011)	1,53E-05
cg01249550	0.012(0.018-0.006)	1,48E-04	0.015(0.022-0.009)	2,13E-05	0.014(0.024-0.003)	1,37E-02	0.017(0.023-0.01)	1,33E-06
cg11785574	0.016(0.022-0.011)	1,03E-07	0.016(0.022-0.009)	3,51E-05	0.014(0.022-0.005)	2,44E-03	0.016(0.022-0.01)	3,37E-06
cg20652640	0.021(0.03-0.012)	8,90E-06	0.031(0.041-0.021)	2,01E-08	0.015(0.032--0.002)	7,54E-02	0.022(0.031-0.013)	7,11E-06
cg05884522	0.015(0.021-0.008)	3,13E-05	0.013(0.019-0.006)	3,30E-04	0.019(0.029-0.008)	1,13E-03	0.014(0.02-0.008)	4,56E-05
cg08300899	0.029(0.042-0.016)	2,58E-05	0.034(0.047-0.02)	4,52E-06	0.033(0.055-0.012)	3,61E-03	0.035(0.047-0.024)	7,75E-08
cg09780955	0.009(0.013-0.004)	1,27E-04	0.01(0.015-0.005)	2,24E-04	0.008(0.016-0)	4,58E-02	0.012(0.017-0.007)	9,28E-06
cg21316470	0.024(0.036-0.011)	2,69E-04	0.028(0.044-0.013)	4,94E-04	0.014(0.044--0.016)	3,46E-01	0.033(0.045-0.021)	4,93E-07
cg11156286	0.014(0.02-0.008)	1,59E-05	0.017(0.024-0.01)	5,23E-06	0.022(0.034-0.01)	6,99E-04	0.015(0.022-0.008)	4,53E-05
cg26948660	0.011(0.016-0.006)	4,53E-05	0.014(0.019-0.008)	1,63E-06	0.015(0.024-0.006)	2,02E-03	0.013(0.018-0.007)	2,62E-05
cg16638571	0.006(0.009-0.003)	3,97E-04	0.01(0.014-0.005)	3,44E-05	0.007(0.015-0)	5,98E-02	0.008(0.013-0.004)	5,39E-04
cg09211893	0.03(0.042-0.018)	3,48E-06	0.028(0.043-0.014)	2,29E-04	0.028(0.048-0.008)	8,16E-03	0.033(0.045-0.02)	1,38E-06

cg18261601	0.011(0.016-0.006)	1,19E-04	0.015(0.021-0.008)	3,35E-05	0.012(0.025--0.001)	6,62E-02	0.017(0.023-0.01)	2,40E-06
cg23183469	0.016(0.025-0.007)	8,39E-04	0.02(0.028-0.011)	1,15E-05	0.017(0.03-0.004)	1,47E-02	0.022(0.031-0.013)	1,46E-05
cg12280407	0.013(0.018-0.007)	2,90E-05	0.018(0.024-0.012)	6,39E-08	0.014(0.025-0.003)	1,75E-02	0.016(0.021-0.01)	1,13E-06
cg04212568	0.015(0.023-0.008)	1,51E-04	0.021(0.028-0.013)	9,83E-07	0.021(0.036-0.005)	9,45E-03	0.021(0.029-0.013)	4,25E-06
cg24592364	0.015(0.022-0.009)	9,13E-06	0.022(0.029-0.015)	5,04E-08	0.009(0.021--0.003)	1,47E-01	0.017(0.024-0.01)	1,04E-05
cg21730067	0.007(0.011-0.004)	2,65E-04	0.011(0.015-0.006)	1,52E-05	0.008(0.016-0)	3,85E-02	0.01(0.014-0.005)	7,03E-05
cg19262019	0.015(0.022-0.008)	9,21E-05	0.016(0.023-0.009)	1,39E-05	0.014(0.026-0.003)	1,65E-02	0.015(0.022-0.008)	6,16E-05
cg08675654	0.019(0.027-0.011)	6,06E-06	0.018(0.027-0.009)	2,28E-04	0.017(0.032-0.002)	2,56E-02	0.023(0.032-0.014)	1,84E-06
cg01634153	0.016(0.023-0.009)	3,39E-05	0.017(0.024-0.009)	5,60E-05	0.019(0.033-0.006)	6,24E-03	0.019(0.026-0.012)	7,82E-07
cg01575096	0.011(0.016-0.005)	1,27E-04	0.013(0.019-0.008)	7,05E-06	0.014(0.021-0.006)	1,14E-03	0.012(0.017-0.006)	4,36E-05
cg01338658	0.02(0.027-0.012)	9,89E-07	0.018(0.027-0.01)	3,43E-05	0.015(0.03--0.001)	5,87E-02	0.022(0.03-0.015)	2,06E-07
cg22109056	0.007(0.011-0.004)	2,40E-05	0.009(0.013-0.005)	1,50E-05	0.01(0.018-0.003)	1,01E-02	0.008(0.011-0.005)	1,37E-05
cg03756640	0.014(0.02-0.009)	7,26E-07	0.013(0.019-0.007)	1,05E-04	0.01(0.019-0.001)	3,05E-02	0.017(0.023-0.012)	6,47E-08
cg21211144	0.013(0.021-0.005)	1,46E-03	0.017(0.025-0.01)	1,80E-05	0.019(0.035-0.004)	1,58E-02	0.019(0.027-0.011)	2,69E-05
cg06707168	0.01(0.015-0.006)	2,13E-05	0.012(0.017-0.007)	1,39E-05	0.01(0.018-0.002)	1,42E-02	0.012(0.017-0.008)	3,02E-07
cg14221884	0.023(0.034-0.012)	9,13E-05	0.024(0.036-0.013)	9,20E-05	0.029(0.047-0.012)	1,63E-03	0.031(0.041-0.021)	5,01E-08
cg08183916	0.038(0.054-0.021)	1,79E-05	0.051(0.067-0.035)	2,43E-08	0.028(0.059--0.002)	6,76E-02	0.041(0.058-0.024)	6,38E-06
cg22194305	0.014(0.021-0.008)	2,90E-05	0.017(0.023-0.01)	2,63E-06	0.015(0.025-0.005)	5,85E-03	0.017(0.023-0.011)	1,99E-07
cg24441872	0.014(0.022-0.007)	2,96E-04	0.017(0.024-0.01)	1,02E-05	0.026(0.041-0.01)	1,93E-03	0.018(0.026-0.01)	3,56E-05
cg11586519	0.015(0.021-0.01)	5,81E-07	0.018(0.025-0.01)	1,64E-05	0.018(0.03-0.005)	6,44E-03	0.014(0.021-0.006)	8,24E-04
cg06363275	0.009(0.014-0.005)	1,33E-04	0.013(0.017-0.008)	1,95E-06	0.008(0.017-0)	4,31E-02	0.011(0.016-0.007)	3,92E-06
cg09338033	0.015(0.021-0.009)	1,11E-05	0.018(0.024-0.011)	1,96E-06	0.011(0.022-0)	4,30E-02	0.02(0.026-0.013)	4,33E-08
cg16590515	0.014(0.02-0.009)	1,18E-06	0.018(0.025-0.011)	1,03E-06	0.02(0.029-0.01)	1,77E-04	0.015(0.023-0.008)	1,39E-04
cg03303757	0.003(0.005-0.001)	4,98E-04	0.004(0.006-0.002)	3,06E-04	0.003(0.005-0)	5,45E-02	0.004(0.005-0.002)	2,41E-05
cg20247600	0.018(0.026-0.01)	2,95E-05	0.02(0.029-0.01)	8,97E-05	0.03(0.048-0.013)	1,16E-03	0.023(0.032-0.015)	7,73E-07
cg21076271	0.017(0.025-0.009)	6,09E-05	0.02(0.028-0.012)	2,08E-06	0.02(0.037-0.003)	2,02E-02	0.021(0.03-0.012)	7,85E-06
cg01111245	0.009(0.013-0.005)	1,56E-05	0.009(0.013-0.005)	1,07E-04	0.005(0.012--0.003)	2,18E-01	0.013(0.017-0.008)	2,28E-07
cg09597785	0.004(0.006-0.002)	1,77E-05	0.005(0.007-0.002)	1,34E-04	0.004(0.009-0)	6,87E-02	0.004(0.006-0.002)	2,40E-05
cg02484274	0.017(0.024-0.01)	8,41E-06	0.019(0.025-0.012)	3,30E-07	0.016(0.03-0.002)	2,47E-02	0.016(0.024-0.009)	7,37E-05
cg15586420	0.017(0.024-0.009)	4,36E-05	0.02(0.028-0.012)	3,13E-06	0.021(0.036-0.006)	6,57E-03	0.018(0.026-0.01)	5,24E-05
cg06812033	0.014(0.02-0.008)	1,16E-05	0.017(0.023-0.011)	3,73E-07	0.008(0.017-0)	6,21E-02	0.016(0.022-0.011)	5,90E-08

cg06355010	0.018(0.027-0.009)	1,04E-04	0.019(0.027-0.011)	1,49E-05	0.016(0.03-0.001)	3,77E-02	0.019(0.027-0.012)	2,65E-06
cg12108029	0.02(0.029-0.011)	3,95E-05	0.022(0.032-0.012)	2,52E-05	0.022(0.038-0.005)	1,07E-02	0.021(0.031-0.012)	3,25E-05
cg02011918	0.014(0.022-0.006)	1,03E-03	0.022(0.029-0.014)	1,69E-07	0.021(0.033-0.009)	1,27E-03	0.02(0.027-0.013)	3,13E-07
cg07189157	0.01(0.015-0.005)	2,40E-04	0.01(0.015-0.005)	4,27E-04	0.011(0.017-0.004)	2,00E-03	0.012(0.017-0.008)	4,67E-07
cg09586534	0.017(0.024-0.01)	1,28E-05	0.019(0.026-0.012)	1,23E-06	0.017(0.03-0.004)	1,06E-02	0.019(0.027-0.011)	1,12E-05
cg13009662	0.025(0.038-0.012)	3,01E-04	0.032(0.048-0.017)	6,78E-05	0.031(0.061-0.002)	3,79E-02	0.03(0.043-0.017)	1,59E-05
cg03913271	0.014(0.02-0.007)	6,89E-05	0.02(0.028-0.012)	3,32E-06	0.008(0.022--0.005)	2,11E-01	0.019(0.026-0.013)	2,08E-07
cg16466635	0.013(0.019-0.007)	1,06E-04	0.016(0.022-0.01)	3,10E-06	0.012(0.021-0.003)	9,40E-03	0.015(0.022-0.008)	6,04E-05
cg03327386	0.023(0.035-0.011)	3,11E-04	0.028(0.039-0.017)	2,38E-06	0.024(0.047-0.001)	3,77E-02	0.029(0.04-0.017)	4,10E-06
cg26705583	0.014(0.021-0.008)	1,77E-05	0.016(0.024-0.009)	3,74E-05	0.004(0.018--0.01)	5,70E-01	0.013(0.02-0.005)	1,10E-03
cg27413118	0.015(0.023-0.007)	4,20E-04	0.022(0.03-0.014)	5,20E-07	0.013(0.026-0.001)	3,17E-02	0.018(0.025-0.011)	1,84E-06
cg09464488	0.019(0.028-0.01)	8,91E-05	0.026(0.036-0.016)	2,50E-06	0.021(0.038-0.003)	2,23E-02	0.023(0.032-0.014)	6,59E-06
cg05787952	0.027(0.038-0.016)	4,90E-06	0.027(0.038-0.017)	3,91E-06	0.033(0.05-0.017)	3,11E-04	0.028(0.038-0.018)	3,68E-07
cg22274117	0.029(0.038-0.019)	1,09E-07	0.02(0.031-0.009)	4,55E-04	0.011(0.03--0.007)	2,27E-01	0.023(0.034-0.013)	2,44E-05
cg02600940	0.01(0.016-0.005)	2,85E-04	0.012(0.018-0.007)	6,39E-05	0.013(0.023-0.004)	9,34E-03	0.013(0.019-0.008)	2,13E-06
cg15359501	0.006(0.009-0.003)	2,32E-04	0.006(0.009-0.003)	2,47E-04	0.004(0.011--0.002)	1,41E-01	0.006(0.009-0.003)	8,57E-05
cg24147485	0.025(0.036-0.015)	1,41E-05	0.023(0.034-0.013)	4,84E-05	0.028(0.046-0.011)	2,94E-03	0.024(0.034-0.014)	7,02E-06
cg15459695	0.025(0.034-0.016)	7,61E-07	0.017(0.028-0.005)	4,28E-03	0.021(0.038-0.004)	1,95E-02	0.027(0.037-0.017)	5,32E-07
cg03167369	0.014(0.022-0.007)	2,29E-04	0.016(0.022-0.009)	1,21E-05	0.02(0.033-0.007)	3,01E-03	0.017(0.024-0.01)	1,98E-05
cg04165030	0.016(0.023-0.009)	2,79E-05	0.017(0.024-0.01)	1,56E-05	0.024(0.037-0.01)	1,03E-03	0.015(0.023-0.008)	9,40E-05
cg00473112	0.015(0.021-0.008)	3,29E-05	0.018(0.025-0.011)	4,69E-06	0.014(0.026-0.003)	1,66E-02	0.014(0.021-0.006)	9,72E-04
cg14045725	0.006(0.009-0.003)	1,85E-04	0.009(0.013-0.005)	4,94E-05	0.006(0.011-0.001)	2,39E-02	0.008(0.011-0.004)	1,39E-05
cg19009714	0.011(0.016-0.006)	8,03E-05	0.014(0.019-0.008)	1,06E-05	0.006(0.014--0.001)	1,03E-01	0.013(0.019-0.008)	9,72E-06
cg07677296	0.012(0.017-0.007)	2,03E-05	0.013(0.02-0.007)	1,62E-04	0.015(0.024-0.007)	1,19E-03	0.013(0.02-0.007)	1,51E-04
cg07249227	0.013(0.019-0.007)	6,31E-05	0.015(0.02-0.009)	4,44E-07	0.011(0.02-0.002)	1,82E-02	0.013(0.02-0.007)	1,21E-04
cg22827465	0.018(0.026-0.009)	5,61E-05	0.019(0.027-0.012)	4,05E-06	0.013(0.024-0.002)	1,88E-02	0.016(0.021-0.01)	1,55E-06
cg00294419	0.017(0.024-0.01)	3,96E-06	0.018(0.026-0.011)	6,34E-06	0.023(0.034-0.011)	4,39E-04	0.018(0.025-0.012)	9,46E-07
cg17820878	-0.023(-0.015--0.031)	6,74E-07	-0.018(-0.009--0.026)	4,85E-05	-0.021(-0.006--0.037)	6,89E-03	-0.02(-0.012--0.029)	1,29E-05
cg00007226	0.011(0.016-0.006)	7,07E-05	0.012(0.018-0.006)	9,28E-05	0.013(0.021-0.005)	2,97E-03	0.013(0.018-0.008)	3,14E-06
cg09901735	0.015(0.021-0.009)	8,10E-06	0.02(0.027-0.013)	1,57E-07	0.015(0.028-0.002)	2,40E-02	0.017(0.024-0.01)	1,74E-05

cg21503377	0.015(0.023-0.007)	2,94E-04	0.015(0.022-0.008)	4,18E-05	0.022(0.033-0.011)	3,60E-04	0.017(0.026-0.009)	8,86E-05
cg18154422	0.015(0.022-0.009)	2,22E-05	0.018(0.024-0.011)	1,25E-06	0.017(0.029-0.006)	4,00E-03	0.014(0.021-0.007)	1,79E-04
cg07061867	0.022(0.032-0.012)	2,63E-05	0.025(0.036-0.014)	3,61E-05	0.023(0.043-0.003)	2,35E-02	0.025(0.034-0.015)	5,55E-06
cg08748615	0.016(0.023-0.009)	1,08E-05	0.022(0.029-0.014)	2,42E-07	0.01(0.025--0.005)	1,91E-01	0.018(0.025-0.01)	1,02E-05
cg13129662	0.021(0.03-0.012)	2,45E-05	0.026(0.037-0.015)	1,62E-05	0.022(0.04-0.005)	1,40E-02	0.022(0.032-0.012)	7,41E-05
cg12927203	0.017(0.025-0.009)	2,96E-05	0.018(0.027-0.009)	1,31E-04	0.019(0.033-0.005)	1,14E-02	0.019(0.027-0.011)	1,98E-05
cg19922442	0.016(0.025-0.006)	1,51E-03	0.025(0.033-0.016)	3,25E-07	0.013(0.027--0.001)	7,52E-02	0.023(0.032-0.013)	8,08E-06
cg00119117	0.019(0.027-0.01)	2,84E-05	0.025(0.034-0.015)	3,72E-06	0.018(0.031-0.004)	1,18E-02	0.022(0.031-0.013)	5,04E-06
cg12133425	0.012(0.018-0.006)	1,31E-04	0.013(0.02-0.006)	3,63E-04	0.016(0.028-0.004)	1,05E-02	0.014(0.02-0.008)	2,09E-05
cg12585429	0.021(0.032-0.009)	5,99E-04	0.027(0.038-0.015)	2,03E-05	0.02(0.04-0)	5,49E-02	0.027(0.038-0.016)	1,23E-05
cg00373665	0.013(0.019-0.007)	8,79E-05	0.014(0.02-0.008)	1,20E-05	0.015(0.023-0.007)	7,22E-04	0.012(0.018-0.007)	2,59E-05
cg04424385	0.016(0.025-0.008)	4,41E-04	0.021(0.03-0.011)	5,40E-05	0.018(0.034-0.002)	2,62E-02	0.019(0.027-0.011)	1,38E-05
cg23539933	0.02(0.029-0.01)	1,03E-04	0.026(0.038-0.014)	3,80E-05	0.028(0.046-0.01)	3,69E-03	0.025(0.033-0.016)	4,96E-07
cg06590470	0.013(0.019-0.008)	7,78E-06	0.015(0.022-0.008)	3,01E-05	0.014(0.023-0.005)	4,64E-03	0.015(0.021-0.009)	8,60E-06
cg27594583	0.015(0.021-0.008)	2,19E-05	0.02(0.026-0.014)	3,31E-09	0.014(0.023-0.006)	2,22E-03	0.021(0.027-0.014)	1,19E-08
cg04052100	0.01(0.015-0.005)	9,15E-05	0.014(0.019-0.008)	2,85E-06	0.016(0.027-0.005)	6,29E-03	0.014(0.019-0.008)	5,38E-06
cg24608684	0.018(0.027-0.009)	2,43E-04	0.024(0.033-0.014)	3,97E-06	0.025(0.043-0.007)	6,92E-03	0.019(0.03-0.009)	5,30E-04
cg24255125	0.025(0.045-0.006)	1,10E-02	0.048(0.064-0.032)	1,36E-07	0.048(0.077-0.02)	1,63E-03	0.044(0.061-0.028)	1,34E-06
cg22718540	0.016(0.022-0.009)	5,07E-06	0.02(0.028-0.013)	1,58E-06	0.013(0.028--0.002)	9,29E-02	0.017(0.024-0.01)	4,53E-06
cg23513504	0.013(0.021-0.006)	7,27E-04	0.019(0.026-0.011)	3,68E-06	0.018(0.031-0.006)	5,60E-03	0.017(0.024-0.01)	1,38E-05
cg24497897	0.02(0.029-0.01)	1,88E-04	0.021(0.03-0.012)	2,44E-05	0.023(0.039-0.008)	4,08E-03	0.02(0.029-0.011)	6,17E-05
cg05744073	0.012(0.018-0.006)	8,51E-05	0.016(0.023-0.009)	1,42E-05	0.011(0.021-0.001)	2,59E-02	0.015(0.02-0.009)	1,22E-06
cg11837377	0.005(0.007-0.003)	4,33E-05	0.006(0.009-0.004)	3,67E-06	0.006(0.01-0.002)	6,18E-03	0.006(0.008-0.003)	9,68E-05
cg14279702	0.015(0.02-0.009)	3,15E-06	0.014(0.02-0.008)	1,43E-05	0.014(0.024-0.004)	9,34E-03	0.014(0.019-0.008)	6,30E-06
cg04028315	0.007(0.011-0.003)	4,11E-04	0.008(0.012-0.004)	3,58E-04	0.009(0.016-0.002)	2,00E-02	0.008(0.012-0.004)	9,61E-05
cg16361303	0.014(0.02-0.009)	1,20E-06	0.013(0.019-0.008)	2,00E-05	0.008(0.019--0.002)	1,05E-01	0.015(0.02-0.009)	8,95E-07
cg02568911	0.018(0.026-0.009)	9,11E-05	0.022(0.031-0.012)	4,87E-05	0.015(0.032--0.002)	7,32E-02	0.026(0.035-0.017)	1,95E-07
cg06380157	0.018(0.026-0.011)	8,05E-06	0.018(0.027-0.009)	1,09E-04	0.022(0.035-0.01)	1,04E-03	0.017(0.026-0.009)	5,02E-05
cg26363167	0.007(0.011-0.003)	7,13E-04	0.009(0.013-0.005)	1,42E-05	0.011(0.019-0.002)	1,54E-02	0.009(0.013-0.005)	5,80E-05
cg11719836	0.005(0.007-0.003)	9,01E-05	0.006(0.008-0.003)	8,97E-06	0.003(0.007-0)	8,68E-02	0.006(0.008-0.003)	2,00E-05
cg19080138	0.023(0.033-0.013)	1,71E-05	0.022(0.032-0.012)	2,58E-05	0.023(0.041-0.004)	1,74E-02	0.021(0.031-0.011)	6,81E-05

cg09556952	0.015(0.022-0.007)	4,10E-04	0.018(0.026-0.01)	3,25E-05	0.022(0.036-0.009)	1,91E-03	0.02(0.028-0.012)	3,83E-06
cg25141720	0.015(0.022-0.009)	2,37E-05	0.012(0.019-0.006)	2,59E-04	0.015(0.025-0.005)	3,50E-03	0.016(0.023-0.009)	4,54E-05
cg03766306	0.009(0.015-0.004)	1,17E-03	0.012(0.017-0.007)	1,77E-05	0.017(0.026-0.007)	8,42E-04	0.01(0.015-0.005)	2,01E-04
cg04994953	0.015(0.023-0.008)	1,99E-04	0.018(0.025-0.011)	5,23E-06	0.013(0.023-0.003)	1,23E-02	0.018(0.025-0.012)	7,83E-07
cg05243095	0.021(0.031-0.011)	7,20E-05	0.019(0.031-0.007)	2,29E-03	0.032(0.051-0.013)	1,82E-03	0.023(0.034-0.012)	4,77E-05
cg04220635	0.014(0.02-0.009)	6,28E-06	0.016(0.023-0.01)	3,94E-06	0.02(0.032-0.008)	2,26E-03	0.016(0.022-0.009)	1,42E-05
cg05018441	0.013(0.018-0.008)	5,83E-06	0.012(0.018-0.006)	7,76E-05	0.011(0.02-0.002)	1,59E-02	0.014(0.019-0.008)	1,63E-06
cg13169065	0.014(0.021-0.007)	1,27E-04	0.016(0.022-0.009)	1,24E-05	0.017(0.028-0.005)	7,06E-03	0.015(0.021-0.009)	9,75E-06
cg02835038	0.026(0.037-0.016)	4,94E-06	0.025(0.038-0.013)	9,23E-05	0.039(0.066-0.012)	6,22E-03	0.028(0.038-0.017)	1,43E-06
cg03848555	0.005(0.008-0.002)	1,03E-03	0.007(0.01-0.003)	1,21E-04	0.003(0.011--0.004)	3,45E-01	0.006(0.009-0.003)	2,68E-04
cg08889551	0.011(0.016-0.006)	4,52E-05	0.013(0.018-0.008)	1,45E-06	0.012(0.02-0.004)	5,37E-03	0.011(0.017-0.005)	2,62E-04
cg25312440	0.044(0.061-0.028)	8,84E-07	0.039(0.057-0.021)	5,58E-05	0.053(0.076-0.03)	5,00E-05	0.038(0.055-0.02)	5,82E-05
cg22402852	0.008(0.012-0.004)	6,95E-05	0.007(0.01-0.003)	4,91E-04	0.009(0.016-0.002)	1,22E-02	0.009(0.013-0.005)	1,80E-05
cg20526663	0.01(0.015-0.004)	6,61E-04	0.013(0.018-0.008)	7,94E-06	0.017(0.027-0.008)	1,13E-03	0.012(0.017-0.007)	2,22E-05
cg21160472	0.038(0.056-0.02)	6,36E-05	0.052(0.072-0.032)	2,49E-06	0.041(0.081-0.001)	4,66E-02	0.05(0.071-0.029)	9,12E-06
cg10735361	0.01(0.015-0.005)	1,63E-04	0.013(0.019-0.008)	1,02E-05	0.013(0.023-0.003)	1,52E-02	0.012(0.018-0.007)	7,21E-05
cg01477633	0.012(0.017-0.007)	3,19E-05	0.015(0.02-0.01)	3,34E-07	0.016(0.026-0.007)	2,01E-03	0.014(0.019-0.008)	1,23E-05
cg02975259	0.009(0.014-0.005)	8,85E-05	0.008(0.013-0.003)	9,53E-04	0.01(0.015-0.005)	5,10E-04	0.011(0.016-0.006)	3,08E-05
cg15467615	0.01(0.015-0.005)	2,97E-04	0.013(0.019-0.008)	3,89E-06	0.008(0.019--0.003)	1,66E-01	0.014(0.019-0.008)	4,17E-06
cg08888178	0.015(0.025-0.005)	3,70E-03	0.022(0.031-0.013)	1,27E-05	0.023(0.039-0.008)	5,03E-03	0.023(0.032-0.014)	7,92E-06
cg16268380	0.019(0.029-0.009)	4,99E-04	0.019(0.032-0.006)	4,65E-03	0.013(0.029--0.003)	9,61E-02	0.019(0.028-0.01)	3,91E-05
cg01297357	0.016(0.024-0.008)	1,31E-04	0.02(0.029-0.011)	5,41E-05	0.025(0.039-0.012)	4,24E-04	0.02(0.029-0.011)	1,92E-05
cg02233216	0.013(0.019-0.006)	3,31E-04	0.019(0.025-0.012)	4,14E-07	0.011(0.021-0.001)	2,67E-02	0.015(0.021-0.009)	1,03E-05
cg04910183	0.038(0.058-0.018)	3,02E-04	0.044(0.063-0.025)	2,61E-05	0.036(0.069-0.002)	3,67E-02	0.045(0.063-0.026)	1,01E-05
cg18947000	0.009(0.014-0.004)	3,58E-04	0.012(0.017-0.007)	1,29E-05	0.009(0.017-0)	4,20E-02	0.01(0.015-0.005)	7,42E-05
cg10235453	0.017(0.029-0.005)	5,20E-03	0.029(0.041-0.017)	6,32E-06	0.027(0.045-0.01)	3,00E-03	0.023(0.033-0.013)	2,03E-05
cg01932632	0.008(0.013-0.004)	9,78E-04	0.011(0.016-0.007)	1,18E-05	0.013(0.02-0.005)	1,23E-03	0.009(0.013-0.004)	1,58E-04
cg06383048	0.01(0.014-0.005)	1,74E-04	0.012(0.016-0.007)	2,33E-06	0.008(0.015-0.001)	1,97E-02	0.009(0.014-0.005)	4,33E-05
cg27634633	0.01(0.016-0.005)	1,29E-04	0.009(0.014-0.004)	5,37E-04	0.015(0.023-0.007)	4,07E-04	0.011(0.015-0.006)	1,06E-05
cg21383720	0.01(0.016-0.005)	4,17E-04	0.015(0.021-0.009)	2,46E-06	0.01(0.02-0)	4,39E-02	0.013(0.018-0.007)	2,20E-05
cg23108125	0.029(0.044-0.014)	2,18E-04	0.044(0.061-0.027)	3,16E-06	0.033(0.069--0.002)	6,02E-02	0.043(0.06-0.026)	3,18E-06

cg08982381	0.01(0.016-0.004)	1,17E-03	0.015(0.022-0.008)	4,43E-05	0.009(0.019-0.001)	6,40E-02	0.016(0.022-0.009)	7,73E-06
cg12878260	0.016(0.024-0.008)	2,16E-04	0.021(0.031-0.011)	8,38E-05	0.013(0.028-0.002)	7,84E-02	0.02(0.028-0.012)	6,82E-06
cg00800773	0.01(0.015-0.005)	1,26E-04	0.011(0.015-0.006)	1,60E-05	0.007(0.014-0)	4,27E-02	0.008(0.013-0.004)	6,33E-04
cg04171308	0.012(0.019-0.005)	6,66E-04	0.016(0.023-0.009)	3,04E-05	0.014(0.025-0.004)	9,51E-03	0.015(0.021-0.01)	1,16E-06
cg23677570	0.01(0.015-0.005)	2,06E-04	0.012(0.018-0.007)	3,39E-05	0.008(0.017-0.002)	1,02E-01	0.011(0.016-0.005)	2,07E-04
cg18317494	0.012(0.019-0.006)	6,12E-04	0.014(0.02-0.008)	3,61E-06	0.018(0.026-0.009)	2,67E-04	0.013(0.02-0.006)	5,99E-04
cg22659014	0.012(0.018-0.006)	2,53E-04	0.017(0.023-0.01)	4,62E-06	0.01(0.019-0.001)	3,91E-02	0.018(0.025-0.012)	3,62E-07
cg12588263	0.017(0.024-0.01)	6,06E-06	0.019(0.026-0.013)	3,36E-07	0.019(0.032-0.006)	7,17E-03	0.013(0.021-0.005)	1,27E-03
cg09424923	0.011(0.018-0.005)	4,71E-04	0.018(0.026-0.01)	2,11E-05	0.002(0.015-0.011)	7,69E-01	0.017(0.024-0.01)	5,44E-06
cg02571907	0.012(0.016-0.007)	2,14E-06	0.014(0.019-0.009)	3,17E-07	0.01(0.019-0.001)	2,57E-02	0.01(0.014-0.005)	1,29E-04
cg22329233	0.011(0.017-0.005)	3,37E-04	0.019(0.026-0.011)	6,78E-06	0.021(0.032-0.009)	1,18E-03	0.017(0.024-0.01)	7,92E-06
cg01263075	0.008(0.012-0.003)	5,92E-04	0.011(0.016-0.007)	7,15E-06	0.011(0.017-0.005)	5,57E-04	0.009(0.013-0.006)	6,09E-06
cg00556087	0.008(0.012-0.003)	9,01E-04	0.01(0.015-0.005)	8,39E-05	0.007(0.014-0.001)	8,95E-02	0.013(0.017-0.009)	2,50E-08
cg06759117	0.008(0.011-0.004)	6,64E-05	0.01(0.015-0.006)	2,39E-05	0.012(0.017-0.006)	3,19E-04	0.009(0.013-0.005)	4,32E-05
cg23456692	0.011(0.016-0.005)	2,69E-04	0.012(0.017-0.007)	1,85E-05	0.019(0.027-0.011)	5,35E-05	0.013(0.019-0.008)	2,83E-06
cg07597763	0.013(0.018-0.008)	1,54E-06	0.008(0.014-0.001)	1,68E-02	0.012(0.02-0.004)	4,37E-03	0.014(0.019-0.008)	8,68E-06
cg11725975	0.004(0.006-0.001)	1,05E-03	0.004(0.007-0.002)	9,01E-04	0.003(0.006-0.001)	1,50E-01	0.004(0.006-0.002)	1,34E-04
cg13227851	0.014(0.02-0.008)	6,81E-06	0.013(0.019-0.007)	3,20E-05	0.01(0.02-0)	4,83E-02	0.011(0.019-0.004)	4,24E-03
cg13937905	-0.036(-0.022--0.051)	3,66E-06	-0.03(-0.015--0.046)	1,75E-04	-0.044(-0.017--0.071)	2,07E-03	-0.031(-0.017--0.045)	5,18E-05
cg05639310	0.015(0.022-0.009)	1,43E-05	0.019(0.027-0.011)	8,54E-06	0.017(0.032-0.003)	1,93E-02	0.017(0.024-0.01)	1,31E-05
cg03351618	0.01(0.015-0.005)	4,76E-04	0.011(0.016-0.006)	3,67E-05	0.012(0.019-0.005)	1,94E-03	0.011(0.016-0.006)	6,31E-05
cg10181968	0.009(0.014-0.005)	1,52E-04	0.01(0.016-0.004)	9,29E-04	0.01(0.018-0.002)	1,78E-02	0.011(0.016-0.007)	2,72E-06
cg19206986	0.019(0.029-0.009)	2,22E-04	0.024(0.034-0.014)	1,60E-05	0.009(0.027-0.008)	2,82E-01	0.025(0.036-0.015)	1,09E-05
cg23816049	0.012(0.018-0.005)	4,97E-04	0.017(0.025-0.009)	8,84E-05	0.017(0.029-0.005)	8,24E-03	0.017(0.023-0.011)	1,65E-07
cg01190522	0.008(0.013-0.003)	1,94E-03	0.014(0.019-0.008)	1,05E-05	0.012(0.02-0.004)	3,31E-03	0.011(0.015-0.006)	1,93E-05
cg08270913	0.009(0.013-0.004)	4,41E-04	0.012(0.017-0.007)	2,62E-05	0.005(0.013-0.004)	2,63E-01	0.013(0.018-0.008)	1,91E-06
cg21226861	0.011(0.016-0.006)	1,03E-04	0.014(0.019-0.008)	3,79E-06	0.009(0.019-0.001)	6,25E-02	0.014(0.019-0.009)	7,17E-07
cg14875093	0.014(0.021-0.007)	9,00E-05	0.019(0.026-0.012)	6,98E-07	0.019(0.031-0.008)	2,14E-03	0.02(0.027-0.013)	8,86E-07
cg10692118	0.012(0.017-0.006)	5,25E-05	0.016(0.022-0.011)	1,45E-07	0.012(0.021-0.003)	1,40E-02	0.014(0.02-0.009)	2,05E-06
cg08128766	0.005(0.007-0.002)	5,64E-05	0.005(0.007-0.002)	2,97E-04	0.005(0.009-0.001)	1,87E-02	0.005(0.007-0.003)	8,44E-06

cg09218130	0.013(0.02-0.007)	1,09E-04	0.014(0.02-0.008)	1,19E-05	0.011(0.021-0.001)	3,67E-02	0.012(0.018-0.005)	6,98E-04
cg14931122	0.013(0.018-0.007)	7,47E-05	0.012(0.018-0.006)	1,04E-04	0.014(0.023-0.004)	6,87E-03	0.013(0.019-0.007)	2,55E-05
cg21991830	0.015(0.022-0.008)	4,59E-05	0.017(0.024-0.01)	4,09E-06	0.013(0.026-0)	4,77E-02	0.015(0.021-0.008)	1,10E-05
cg05555681	0.008(0.013-0.004)	3,75E-04	0.012(0.018-0.007)	3,16E-05	0.006(0.014-0.002)	1,52E-01	0.011(0.015-0.006)	1,21E-05
cg11754897	0.008(0.013-0.003)	1,26E-03	0.013(0.018-0.007)	5,83E-06	0.009(0.019-0)	4,77E-02	0.011(0.016-0.007)	4,69E-06
cg11819707	0.012(0.018-0.005)	4,69E-04	0.017(0.023-0.011)	1,95E-07	0.017(0.029-0.006)	4,47E-03	0.015(0.021-0.009)	1,31E-05
cg02811931	0.015(0.022-0.007)	2,73E-04	0.017(0.025-0.008)	4,04E-04	0.014(0.027-0.001)	3,82E-02	0.017(0.023-0.01)	1,99E-06
cg22253313	0.019(0.028-0.01)	4,97E-05	0.019(0.028-0.01)	7,97E-05	0.012(0.028-0.004)	1,46E-01	0.022(0.031-0.013)	1,34E-05
cg04738301	0.011(0.016-0.006)	2,05E-05	0.011(0.017-0.006)	3,61E-05	0.009(0.019-0.001)	7,88E-02	0.012(0.017-0.007)	1,82E-05
cg06378494	0.015(0.023-0.008)	4,61E-05	0.018(0.027-0.009)	2,25E-04	0.023(0.035-0.012)	2,91E-04	0.017(0.025-0.008)	2,14E-04
cg07168160	0.008(0.012-0.004)	1,05E-04	0.008(0.012-0.004)	2,33E-04	0.008(0.016-0)	4,68E-02	0.008(0.012-0.004)	4,25E-04
cg05422897	0.012(0.017-0.006)	6,67E-05	0.013(0.018-0.007)	3,99E-05	0.016(0.024-0.008)	2,64E-04	0.012(0.018-0.007)	3,53E-05
cg20202505	0.012(0.019-0.006)	3,71E-04	0.014(0.021-0.007)	2,14E-04	0.009(0.024-0.006)	2,13E-01	0.013(0.019-0.007)	7,53E-05
cg02454264	0.014(0.021-0.006)	3,65E-04	0.017(0.024-0.009)	2,37E-05	0.006(0.025-0.012)	4,86E-01	0.015(0.022-0.009)	2,80E-05
cg03346041	0.015(0.022-0.007)	2,24E-04	0.02(0.029-0.011)	2,15E-05	0.015(0.029-0.001)	3,83E-02	0.021(0.028-0.013)	1,66E-06
cg22183626	0.012(0.019-0.005)	7,37E-04	0.015(0.022-0.008)	4,26E-05	0.013(0.022-0.004)	6,48E-03	0.016(0.022-0.009)	1,59E-05
cg03673687	0.012(0.017-0.006)	4,65E-05	0.016(0.022-0.009)	5,49E-06	0.013(0.023-0.004)	8,80E-03	0.015(0.021-0.01)	6,74E-07
cg06775055	0.003(0.005-0.002)	1,18E-04	0.003(0.005-0.001)	4,77E-04	0.003(0.007-0)	3,05E-02	0.003(0.005-0.002)	2,98E-04
cg05266843	0.021(0.032-0.01)	2,32E-04	0.023(0.035-0.012)	1,83E-04	0.023(0.044-0.002)	3,27E-02	0.03(0.04-0.019)	3,89E-07
cg14591243	0.01(0.014-0.005)	8,65E-05	0.012(0.017-0.007)	1,73E-05	0.01(0.019-0.002)	1,33E-02	0.011(0.015-0.006)	7,26E-06
cg02463426	0.02(0.028-0.012)	6,26E-06	0.022(0.031-0.014)	1,56E-06	0.021(0.037-0.006)	7,15E-03	0.021(0.028-0.013)	1,12E-06
cg09344986	0.02(0.029-0.012)	9,45E-06	0.017(0.026-0.008)	6,07E-04	0.022(0.036-0.007)	4,81E-03	0.019(0.027-0.011)	2,25E-05
cg10427040	0.017(0.026-0.008)	3,46E-04	0.022(0.03-0.013)	2,45E-06	0.023(0.04-0.006)	1,14E-02	0.019(0.028-0.011)	4,41E-05
cg24523758	0.013(0.019-0.007)	5,17E-05	0.014(0.02-0.007)	7,55E-05	0.012(0.023-0.002)	2,37E-02	0.015(0.021-0.009)	8,28E-06
cg12425673	0.011(0.016-0.006)	7,47E-05	0.01(0.016-0.005)	2,92E-04	0.014(0.022-0.006)	1,46E-03	0.011(0.016-0.007)	5,87E-06
cg07485181	0.01(0.016-0.004)	1,55E-03	0.016(0.022-0.01)	2,88E-06	0.012(0.021-0.002)	1,50E-02	0.013(0.019-0.008)	2,24E-05
cg06461769	0.011(0.016-0.005)	2,18E-04	0.013(0.018-0.008)	7,53E-06	0.009(0.019-0.002)	9,90E-02	0.014(0.019-0.009)	8,29E-07
cg13674535	0.009(0.012-0.005)	4,28E-05	0.007(0.011-0.003)	7,17E-04	0.011(0.016-0.006)	1,68E-04	0.01(0.014-0.007)	1,89E-07
cg22793136	0.019(0.025-0.012)	4,15E-07	0.018(0.025-0.01)	2,98E-05	0.014(0.025-0.002)	1,99E-02	0.023(0.03-0.016)	4,20E-09
cg01751341	0.007(0.011-0.004)	1,10E-04	0.007(0.012-0.003)	1,51E-03	0.006(0.013-0.001)	9,04E-02	0.01(0.014-0.006)	1,30E-06
cg21140261	0.012(0.017-0.007)	1,50E-05	0.013(0.019-0.008)	6,36E-06	0.011(0.019-0.003)	1,08E-02	0.012(0.017-0.007)	1,25E-05

cg04117874	0.003(0.005-0.002)	3,07E-04	0.004(0.006-0.002)	3,32E-05	0.003(0.007-0)	3,18E-02	0.004(0.006-0.002)	8,09E-05
cg26353862	0.013(0.021-0.006)	5,83E-04	0.015(0.023-0.008)	9,58E-05	0.017(0.028-0.006)	4,39E-03	0.019(0.025-0.012)	9,81E-07
cg11016194	0.013(0.02-0.006)	5,52E-04	0.015(0.023-0.008)	8,19E-05	0.019(0.031-0.007)	2,33E-03	0.016(0.023-0.009)	3,42E-05
cg11689622	0.008(0.011-0.004)	8,52E-05	0.009(0.014-0.005)	4,86E-05	0.009(0.014-0.003)	6,20E-03	0.007(0.011-0.004)	1,68E-04
cg04464465	0.012(0.019-0.006)	3,27E-04	0.016(0.023-0.008)	1,43E-04	0.014(0.025-0.004)	1,09E-02	0.017(0.023-0.011)	9,98E-08
cg24354901	0.009(0.013-0.004)	1,46E-04	0.007(0.011-0.002)	3,44E-03	0.008(0.014-0.002)	9,60E-03	0.007(0.011-0.004)	1,68E-04
cg19484886	0.014(0.022-0.006)	5,51E-04	0.017(0.024-0.01)	8,91E-06	0.022(0.033-0.011)	4,41E-04	0.014(0.021-0.007)	1,18E-04
cg10391895	0.019(0.027-0.012)	2,58E-06	0.016(0.024-0.007)	5,29E-04	0.017(0.03-0.004)	1,07E-02	0.017(0.027-0.008)	6,38E-04
cg12923613	0.011(0.016-0.005)	1,43E-04	0.014(0.019-0.008)	2,88E-06	0.015(0.025-0.004)	8,32E-03	0.014(0.019-0.008)	1,76E-06
cg09410234	0.014(0.021-0.007)	2,71E-04	0.015(0.022-0.008)	8,90E-05	0.016(0.027-0.005)	5,68E-03	0.015(0.021-0.008)	3,57E-05
cg00908526	0.027(0.039-0.014)	5,39E-05	0.023(0.035-0.011)	3,05E-04	0.023(0.04-0.006)	1,05E-02	0.027(0.04-0.015)	5,70E-05
cg03309786	0.005(0.008-0.003)	1,99E-04	0.006(0.008-0.003)	7,56E-05	0.001(0.005-0.004)	7,26E-01	0.006(0.008-0.003)	1,62E-05
cg22112144	0.015(0.022-0.008)	3,23E-05	0.016(0.024-0.009)	5,17E-05	0.013(0.025-0)	4,38E-02	0.02(0.027-0.013)	3,14E-07
cg14724283	0.015(0.022-0.007)	3,50E-04	0.021(0.029-0.012)	4,58E-06	0.022(0.04-0.004)	1,90E-02	0.021(0.029-0.012)	8,36E-06
cg05307908	0.009(0.014-0.004)	4,32E-04	0.013(0.018-0.007)	2,06E-05	0.011(0.019-0.003)	6,23E-03	0.012(0.018-0.007)	4,97E-05
cg02331143	0.007(0.01-0.004)	1,67E-05	0.005(0.009-0.002)	1,76E-03	0.005(0.011-0)	4,53E-02	0.007(0.01-0.004)	5,90E-05
cg01824804	0.02(0.028-0.012)	6,81E-06	0.018(0.028-0.007)	1,14E-03	0.016(0.028-0.003)	1,57E-02	0.019(0.028-0.009)	4,01E-04
cg04276723	0.012(0.018-0.006)	2,07E-04	0.012(0.018-0.005)	6,41E-04	0.012(0.021-0.003)	9,08E-03	0.015(0.021-0.008)	1,44E-05
cg08957360	0.01(0.015-0.005)	3,10E-04	0.012(0.018-0.007)	6,26E-06	0.013(0.021-0.004)	3,99E-03	0.014(0.02-0.009)	2,81E-06
cg02195676	0.012(0.017-0.006)	1,10E-04	0.014(0.02-0.007)	5,58E-05	0.016(0.027-0.006)	3,72E-03	0.014(0.021-0.008)	4,34E-05
cg05151228	0.02(0.029-0.012)	1,08E-05	0.02(0.03-0.01)	1,57E-04	0.02(0.035-0.005)	1,03E-02	0.019(0.028-0.01)	9,38E-05
cg04543026	0.012(0.017-0.006)	5,08E-05	0.013(0.019-0.007)	8,18E-05	0.015(0.025-0.004)	8,24E-03	0.016(0.022-0.01)	2,64E-06
cg24578966	0.01(0.016-0.004)	1,45E-03	0.014(0.02-0.007)	3,17E-05	0.02(0.032-0.009)	1,14E-03	0.015(0.021-0.01)	7,19E-07
cg06576198	0.019(0.026-0.011)	4,33E-06	0.02(0.029-0.011)	5,20E-05	0.021(0.036-0.007)	5,80E-03	0.017(0.026-0.009)	1,76E-04
cg02547025	0.026(0.04-0.012)	3,00E-04	0.033(0.048-0.019)	2,64E-05	0.023(0.056-0.01)	1,57E-01	0.035(0.05-0.021)	1,13E-05
cg25296860	0.015(0.023-0.007)	4,13E-04	0.017(0.025-0.009)	5,18E-05	0.02(0.035-0.006)	7,12E-03	0.016(0.023-0.009)	5,25E-05
cg19375676	0.011(0.017-0.006)	1,09E-04	0.013(0.019-0.008)	1,24E-05	0.013(0.022-0.003)	1,03E-02	0.012(0.018-0.007)	1,32E-05
cg26911804	0.006(0.008-0.003)	8,56E-05	0.005(0.008-0.002)	2,92E-03	0.007(0.014-0.001)	2,81E-02	0.007(0.01-0.004)	3,31E-05
cg15752043	0.009(0.015-0.003)	2,70E-03	0.013(0.019-0.007)	6,75E-05	0.013(0.022-0.004)	7,28E-03	0.015(0.019-0.01)	4,17E-08
cg22186533	0.011(0.017-0.006)	5,47E-05	0.017(0.022-0.011)	2,69E-07	0.01(0.019-0.001)	2,83E-02	0.015(0.02-0.009)	4,16E-07
cg02201963	0.012(0.019-0.005)	7,87E-04	0.015(0.022-0.009)	1,21E-05	0.015(0.026-0.004)	1,04E-02	0.015(0.021-0.009)	6,31E-06

cg26889928	0.013(0.02-0.006)	7,79E-04	0.017(0.025-0.01)	3,04E-05	0.016(0.027-0.005)	4,77E-03	0.016(0.023-0.01)	6,57E-06
cg22960945	0.005(0.007-0.003)	4,55E-06	0.004(0.006-0.002)	1,90E-04	0.005(0.009-0.002)	5,33E-03	0.004(0.006-0.002)	2,90E-04
cg21800906	0.008(0.013-0.004)	2,32E-04	0.008(0.012-0.003)	1,26E-03	0.009(0.016-0.002)	1,11E-02	0.011(0.015-0.006)	6,08E-06
cg13970591	0.013(0.019-0.007)	4,78E-05	0.014(0.02-0.008)	1,77E-05	0.019(0.03-0.008)	1,36E-03	0.016(0.023-0.01)	2,16E-06
cg19229702	0.015(0.022-0.008)	9,60E-05	0.018(0.026-0.01)	4,52E-05	0.018(0.033-0.002)	2,56E-02	0.019(0.027-0.011)	1,64E-05
cg14024356	0.015(0.022-0.008)	2,75E-05	0.013(0.021-0.005)	1,51E-03	0.014(0.026-0.002)	2,07E-02	0.015(0.021-0.008)	5,09E-05
cg10122932	0.016(0.023-0.009)	3,91E-05	0.017(0.026-0.009)	1,49E-04	0.019(0.036-0.003)	2,44E-02	0.017(0.025-0.008)	2,06E-04
cg03754076	0.015(0.022-0.007)	1,41E-04	0.019(0.026-0.011)	1,51E-05	0.019(0.034-0.004)	1,44E-02	0.017(0.025-0.01)	2,56E-05
cg13464062	0.011(0.016-0.006)	5,65E-05	0.012(0.019-0.005)	8,52E-04	0.016(0.024-0.007)	5,77E-04	0.016(0.021-0.011)	7,85E-09
cg18500714	0.034(0.05-0.017)	9,05E-05	0.039(0.056-0.021)	3,38E-05	0.042(0.077-0.008)	1,84E-02	0.038(0.055-0.021)	2,93E-05
cg05554039	0.018(0.025-0.011)	2,25E-06	0.017(0.026-0.008)	4,45E-04	0.021(0.033-0.008)	2,72E-03	0.019(0.027-0.012)	5,81E-06
cg03749777	0.013(0.019-0.006)	3,00E-04	0.015(0.022-0.009)	2,83E-05	0.012(0.022-0.002)	1,79E-02	0.015(0.022-0.008)	5,47E-05
cg19144392	0.006(0.011-0.002)	8,49E-03	0.009(0.014-0.005)	2,85E-05	0.009(0.015-0.002)	1,23E-02	0.009(0.013-0.005)	1,76E-05
cg23872935	0.015(0.025-0.005)	3,40E-03	0.024(0.033-0.015)	9,19E-07	0.023(0.04-0.007)	6,84E-03	0.018(0.027-0.009)	2,12E-04
cg08294267	0.013(0.019-0.007)	1,01E-04	0.016(0.022-0.01)	3,71E-06	0.011(0.021-0.001)	3,47E-02	0.015(0.022-0.009)	2,57E-05
cg05530459	0.009(0.013-0.004)	3,19E-04	0.012(0.017-0.006)	3,40E-05	0.007(0.014-0.001)	7,23E-02	0.012(0.016-0.007)	2,22E-06
cg18858542	0.013(0.02-0.006)	2,84E-04	0.011(0.018-0.005)	6,68E-04	0.009(0.017-0.001)	2,87E-02	0.014(0.02-0.008)	2,08E-05
cg08676905**	0.021(0.029-0.012)	1,01E-05	0.017(0.027-0.008)	4,33E-04	0.024(0.038-0.009)	2,87E-03	0.021(0.029-0.012)	5,81E-06
cg21212389	0.007(0.01-0.004)	2,18E-05	0.007(0.01-0.003)	5,63E-04	0.008(0.014-0.001)	1,86E-02	0.008(0.011-0.005)	1,12E-05
cg19612890	0.007(0.011-0.004)	5,32E-05	0.009(0.013-0.005)	1,02E-05	0.01(0.017-0.004)	2,62E-03	0.008(0.012-0.005)	2,16E-05
cg27261136	0.009(0.014-0.005)	1,64E-04	0.011(0.016-0.007)	7,46E-06	0.008(0.016-0.001)	3,47E-02	0.011(0.016-0.006)	3,80E-05
cg11716640	0.014(0.021-0.006)	6,11E-04	0.013(0.02-0.006)	2,29E-04	0.019(0.028-0.01)	1,26E-04	0.013(0.02-0.006)	3,50E-04
cg02838932	0.01(0.014-0.005)	3,88E-05	0.009(0.013-0.005)	1,02E-04	0.011(0.019-0.004)	5,07E-03	0.01(0.014-0.006)	2,13E-05
cg15027815	0.031(0.042-0.021)	1,27E-07	0.023(0.034-0.012)	9,04E-05	0.025(0.043-0.007)	8,17E-03	0.028(0.04-0.016)	1,46E-05
cg05619888	0.006(0.01-0.003)	9,73E-04	0.007(0.01-0.004)	9,66E-05	0.003(0.01--0.003)	2,81E-01	0.006(0.01-0.002)	1,32E-03
cg13224729	0.004(0.007-0.001)	3,18E-03	0.005(0.008-0.003)	3,37E-05	0.004(0.008-0.001)	8,47E-02	0.006(0.009-0.004)	1,91E-06
cg00375031	0.009(0.015-0.004)	1,24E-03	0.015(0.021-0.009)	5,79E-06	0.014(0.022-0.005)	2,67E-03	0.01(0.015-0.006)	2,05E-05
cg02130453	0.014(0.021-0.008)	5,07E-05	0.017(0.024-0.009)	2,17E-05	0.01(0.022--0.001)	8,54E-02	0.014(0.021-0.007)	1,30E-04
cg18710383	0.016(0.025-0.007)	7,07E-04	0.023(0.032-0.014)	2,14E-06	0.007(0.019-0.004)	1,99E-01	0.017(0.027-0.007)	6,95E-04
cg24349555	0.014(0.021-0.008)	5,11E-05	0.014(0.021-0.008)	3,23E-05	0.011(0.02-0.001)	3,01E-02	0.016(0.022-0.009)	9,55E-06
cg00753739	0.011(0.019-0.003)	9,42E-03	0.021(0.03-0.013)	5,34E-06	0.019(0.031-0.007)	2,23E-03	0.02(0.028-0.013)	9,39E-07

cg17760011	0.009(0.012-0.005)	1,11E-05	0.009(0.014-0.005)	7,26E-05	0.008(0.016-0.001)	2,20E-02	0.009(0.014-0.004)	2,18E-04
cg21754773	0.015(0.022-0.008)	4,73E-05	0.02(0.028-0.013)	1,39E-06	0.01(0.022--0.003)	1,40E-01	0.017(0.023-0.01)	1,51E-06
cg24512132	0.004(0.006-0.002)	5,12E-05	0.003(0.005-0.001)	6,27E-03	0.003(0.007--0.001)	1,61E-01	0.003(0.005-0.002)	5,64E-04
cg19907055	0.012(0.018-0.007)	2,78E-05	0.011(0.017-0.006)	1,85E-04	0.005(0.016--0.006)	3,62E-01	0.013(0.02-0.007)	1,21E-04
cg25721625	0.025(0.036-0.013)	5,49E-05	0.025(0.035-0.015)	1,05E-05	0.02(0.038-0.003)	2,24E-02	0.027(0.037-0.016)	3,26E-06
cg13984040	0.066(0.087-0.046)	1,23E-08	0.038(0.062-0.014)	2,50E-03	0.061(0.097-0.024)	2,08E-03	0.053(0.073-0.032)	4,19E-06
cg01470283	0.013(0.018-0.008)	8,32E-06	0.016(0.023-0.01)	3,40E-06	0.016(0.028-0.003)	1,62E-02	0.012(0.018-0.005)	4,39E-04
cg02737068	0.011(0.016-0.006)	5,43E-05	0.013(0.02-0.006)	1,85E-04	0.01(0.02--0.001)	7,99E-02	0.013(0.019-0.007)	5,27E-05
cg09535168	0.007(0.01-0.004)	7,18E-05	0.007(0.011-0.004)	2,38E-04	0.012(0.017-0.006)	8,76E-05	0.007(0.011-0.004)	1,23E-04
cg18289048	0.005(0.008-0.002)	6,61E-04	0.005(0.008-0.002)	2,66E-04	0.006(0.011-0.002)	8,96E-03	0.006(0.009-0.003)	2,96E-05
cg20716688	0.015(0.023-0.008)	1,82E-04	0.021(0.029-0.013)	3,24E-06	0.013(0.026-0.001)	4,13E-02	0.02(0.028-0.012)	2,00E-06
cg04120766	0.011(0.017-0.005)	6,46E-04	0.015(0.022-0.008)	4,36E-05	0.006(0.018--0.007)	3,39E-01	0.013(0.02-0.007)	2,00E-04
cg16449186	0.009(0.013-0.004)	2,84E-04	0.009(0.014-0.005)	1,49E-04	0.009(0.015-0.002)	1,04E-02	0.01(0.014-0.006)	2,34E-06
cg18661819	0.012(0.018-0.006)	1,88E-04	0.016(0.023-0.008)	6,87E-05	0.009(0.021--0.003)	1,27E-01	0.017(0.024-0.01)	5,36E-06
cg22843625	0.015(0.023-0.008)	1,48E-04	0.012(0.019-0.004)	3,51E-03	0.012(0.025-0)	5,86E-02	0.016(0.023-0.01)	6,11E-06
cg11382291	0.012(0.018-0.007)	6,00E-05	0.015(0.021-0.009)	2,51E-06	0.009(0.021--0.002)	9,47E-02	0.013(0.02-0.006)	2,13E-04
cg15822346	0.012(0.018-0.005)	5,68E-04	0.02(0.028-0.013)	1,08E-06	0.013(0.024-0.002)	2,32E-02	0.017(0.024-0.009)	2,79E-05
cg16551665	0.018(0.028-0.009)	2,21E-04	0.025(0.038-0.011)	6,34E-04	0.011(0.029--0.006)	2,02E-01	0.026(0.037-0.016)	8,28E-06
cg00203461	0.012(0.017-0.006)	5,65E-05	0.011(0.016-0.005)	3,15E-04	0.013(0.02-0.006)	7,10E-04	0.012(0.017-0.006)	3,87E-05
cg14915787	0.006(0.009-0.002)	1,52E-03	0.006(0.01-0.003)	1,03E-03	0.007(0.012-0.002)	7,36E-03	0.007(0.01-0.004)	1,53E-05
cg14489649	0.012(0.017-0.007)	3,07E-05	0.012(0.018-0.006)	2,57E-04	0.011(0.019-0.003)	1,16E-02	0.011(0.015-0.006)	5,56E-05
cg05200311	0.014(0.021-0.008)	7,68E-05	0.018(0.025-0.01)	2,75E-05	0.017(0.028-0.005)	5,60E-03	0.019(0.026-0.011)	1,15E-05
cg03624415	0.013(0.019-0.008)	7,34E-06	0.012(0.018-0.006)	1,06E-04	0.01(0.019-0.002)	1,85E-02	0.01(0.015-0.004)	1,14E-03
cg10513276	0.015(0.023-0.007)	3,87E-04	0.018(0.026-0.009)	6,61E-05	0.011(0.028--0.007)	2,21E-01	0.019(0.028-0.01)	4,64E-05
cg08894790	0.013(0.019-0.008)	7,44E-06	0.013(0.02-0.007)	1,25E-04	0.015(0.026-0.004)	7,71E-03	0.013(0.019-0.006)	2,16E-04
cg01963056**	0.004(0.007-0.002)	8,09E-04	0.003(0.006-0.001)	4,21E-03	0.004(0.008-0)	4,49E-02	0.005(0.008-0.003)	6,46E-05
cg10335659	0.009(0.015-0.004)	1,23E-03	0.014(0.02-0.008)	1,12E-05	0.016(0.024-0.007)	5,76E-04	0.011(0.017-0.006)	1,52E-04
cg22935821	0.005(0.008-0.002)	1,22E-03	0.007(0.01-0.004)	1,41E-05	0.007(0.011-0.003)	2,79E-03	0.007(0.01-0.003)	1,20E-04
cg15645079	0.01(0.015-0.005)	1,26E-04	0.008(0.012-0.004)	3,86E-04	0.007(0.014-0)	4,73E-02	0.009(0.013-0.004)	1,74E-04
cg13497301	0.01(0.015-0.006)	4,16E-05	0.01(0.016-0.004)	1,81E-03	0.011(0.019-0.004)	2,37E-03	0.012(0.017-0.006)	5,89E-05
cg14556391	0.009(0.014-0.005)	7,00E-05	0.009(0.014-0.004)	8,61E-04	0.008(0.015-0.001)	2,18E-02	0.009(0.014-0.004)	2,59E-04

cg02008025	0.009(0.015-0.004)	1,49E-03	0.011(0.018-0.005)	1,06E-03	0.011(0.021-0.001)	3,02E-02	0.014(0.02-0.008)	8,44E-06
cg10662979	0.013(0.02-0.007)	1,85E-04	0.017(0.023-0.01)	2,46E-06	0.018(0.029-0.006)	4,25E-03	0.017(0.023-0.01)	2,16E-06
cg24989061	0.011(0.016-0.005)	1,29E-04	0.013(0.019-0.006)	2,15E-04	0.009(0.02--0.001)	6,73E-02	0.011(0.017-0.005)	5,71E-04
cg06592416	0.009(0.013-0.005)	3,57E-05	0.01(0.014-0.005)	5,26E-05	0.011(0.019-0.003)	7,14E-03	0.01(0.014-0.006)	4,15E-06
cg05632623	0.015(0.021-0.009)	1,30E-05	0.018(0.026-0.011)	5,45E-06	0.014(0.025-0.004)	1,01E-02	0.016(0.024-0.008)	1,48E-04
cg20888386	0.016(0.022-0.01)	1,49E-06	0.019(0.027-0.011)	8,99E-06	0.013(0.024-0.003)	1,49E-02	0.016(0.022-0.009)	2,12E-05
cg20274746	0.012(0.018-0.006)	2,41E-04	0.012(0.019-0.006)	1,90E-04	0.013(0.02-0.005)	3,00E-03	0.012(0.018-0.007)	5,32E-05
cg22347322	0.011(0.016-0.005)	1,06E-04	0.015(0.021-0.009)	2,30E-06	0.01(0.018-0.002)	1,24E-02	0.012(0.017-0.007)	1,29E-05
cg06156768	0.014(0.021-0.006)	4,27E-04	0.019(0.028-0.011)	2,10E-05	0.013(0.025-0.001)	3,12E-02	0.02(0.028-0.013)	8,18E-07
cg10573403	0.008(0.013-0.004)	3,18E-04	0.009(0.014-0.004)	2,70E-04	0.009(0.016-0.002)	1,31E-02	0.011(0.015-0.007)	1,28E-06
cg10830518	0.008(0.011-0.004)	3,60E-05	0.007(0.011-0.003)	6,27E-04	0.008(0.015-0.001)	2,40E-02	0.008(0.012-0.004)	4,93E-05
cg22377237	0.015(0.021-0.008)	1,33E-05	0.011(0.017-0.004)	1,83E-03	0.015(0.024-0.006)	1,65E-03	0.013(0.019-0.006)	1,68E-04
cg04307927	0.01(0.015-0.005)	4,09E-04	0.012(0.019-0.006)	3,28E-04	0.011(0.021-0)	4,30E-02	0.013(0.019-0.007)	2,61E-05
cg26170244	0.037(0.051-0.024)	5,55E-07	0.024(0.038-0.01)	1,44E-03	0.033(0.057-0.009)	7,81E-03	0.029(0.045-0.014)	3,47E-04
cg07568751	0.013(0.018-0.007)	3,85E-05	0.011(0.017-0.005)	5,99E-04	0.012(0.021-0.002)	1,81E-02	0.012(0.017-0.007)	3,66E-05
cg07494420	0.012(0.017-0.006)	5,30E-05	0.013(0.019-0.006)	1,42E-04	0.018(0.027-0.008)	5,31E-04	0.012(0.018-0.005)	5,66E-04
cg11748936	0.009(0.014-0.003)	2,30E-03	0.008(0.013-0.004)	4,10E-04	0.01(0.017-0.003)	4,63E-03	0.011(0.015-0.007)	4,70E-06
cg12818493	0.014(0.023-0.005)	1,80E-03	0.017(0.025-0.009)	4,48E-05	0.018(0.029-0.008)	1,35E-03	0.019(0.027-0.012)	3,35E-06
cg17083941	0.024(0.033-0.014)	3,15E-06	0.02(0.031-0.009)	6,66E-04	0.027(0.04-0.013)	3,68E-04	0.023(0.033-0.014)	6,64E-06
cg07696592	0.01(0.015-0.005)	6,55E-05	0.008(0.013-0.003)	1,29E-03	0.015(0.022-0.007)	2,65E-04	0.011(0.016-0.007)	6,18E-06
cg02541231	0.011(0.017-0.004)	3,18E-03	0.017(0.024-0.01)	4,09E-06	0.017(0.028-0.007)	2,30E-03	0.014(0.021-0.008)	4,73E-05
cg08302713	0.012(0.017-0.006)	9,63E-05	0.011(0.017-0.005)	2,25E-04	0.004(0.011--0.003)	2,14E-01	0.013(0.018-0.008)	6,78E-07
cg25974903	0.011(0.017-0.005)	2,18E-04	0.016(0.022-0.01)	4,08E-06	0.015(0.026-0.005)	5,80E-03	0.014(0.021-0.007)	8,64E-05
cg25560333	0.009(0.013-0.006)	3,01E-06	0.006(0.01-0.002)	4,08E-03	0.011(0.017-0.004)	1,84E-03	0.008(0.012-0.004)	4,25E-04
cg06974515	0.012(0.016-0.007)	3,13E-07	0.01(0.015-0.006)	5,76E-05	0.011(0.018-0.005)	1,58E-03	0.011(0.015-0.006)	2,70E-05
cg22758970	0.008(0.012-0.004)	3,23E-04	0.009(0.013-0.004)	2,39E-04	0.011(0.017-0.004)	1,64E-03	0.009(0.013-0.005)	3,91E-05
cg21205504	0.013(0.019-0.006)	1,66E-04	0.013(0.02-0.007)	2,07E-04	0.014(0.025-0.003)	1,42E-02	0.014(0.021-0.006)	5,82E-04
cg01267522	0.012(0.017-0.007)	1,49E-05	0.01(0.015-0.005)	1,39E-04	0.013(0.022-0.005)	3,95E-03	0.012(0.017-0.007)	7,07E-06
cg06529027	0.013(0.019-0.007)	3,13E-05	0.011(0.017-0.004)	1,27E-03	0.013(0.024-0.003)	1,32E-02	0.013(0.019-0.007)	5,69E-05
cg08551857	0.008(0.013-0.003)	1,66E-03	0.013(0.018-0.008)	1,55E-06	0.011(0.018-0.004)	1,97E-03	0.011(0.016-0.006)	8,93E-05
cg03614513	0.01(0.015-0.005)	1,01E-04	0.012(0.017-0.007)	8,24E-06	0.006(0.014--0.002)	1,14E-01	0.01(0.015-0.005)	3,53E-04

cg03042149	0.007(0.011-0.003)	1,15E-03	0.009(0.013-0.005)	1,33E-05	0.008(0.016-0.001)	2,58E-02	0.009(0.013-0.006)	6,27E-07
cg09006498	0.007(0.01-0.004)	7,85E-05	0.007(0.011-0.003)	2,68E-04	0.006(0.012-0)	4,68E-02	0.006(0.009-0.003)	4,36E-05
cg00784703	0.007(0.011-0.004)	6,50E-05	0.008(0.012-0.005)	1,85E-05	0.007(0.014-0)	5,52E-02	0.008(0.011-0.004)	1,35E-04
cg20203327	0.005(0.007-0.002)	1,38E-03	0.007(0.01-0.003)	1,13E-04	0.005(0.01-0)	7,15E-02	0.006(0.009-0.004)	1,98E-05
cg09142020	0.029(0.041-0.016)	2,20E-05	0.027(0.041-0.012)	5,83E-04	0.032(0.055-0.009)	7,79E-03	0.03(0.042-0.017)	1,48E-05
cg01902998	0.016(0.024-0.008)	2,13E-04	0.017(0.025-0.009)	1,59E-04	0.016(0.03-0.001)	3,68E-02	0.014(0.023-0.005)	2,34E-03
cg24224803	0.011(0.017-0.006)	1,48E-04	0.015(0.021-0.01)	1,93E-06	0.012(0.022-0.002)	2,56E-02	0.013(0.019-0.007)	2,72E-05
cg02624574	0.004(0.006-0.002)	8,09E-05	0.005(0.007-0.003)	4,95E-05	0.004(0.008-0)	4,35E-02	0.006(0.008-0.004)	7,34E-07
cg04615716	0.007(0.011-0.003)	5,93E-04	0.01(0.015-0.006)	1,01E-05	0.005(0.012--0.002)	1,25E-01	0.01(0.014-0.005)	2,79E-05
cg00860370	0.004(0.006-0.001)	1,41E-03	0.004(0.006-0.002)	1,37E-04	0.002(0.006--0.001)	1,60E-01	0.004(0.005-0.002)	1,83E-04
cg04004861	0.012(0.018-0.005)	3,99E-04	0.015(0.022-0.009)	3,26E-05	0.01(0.02--0.001)	7,15E-02	0.012(0.018-0.006)	2,37E-04
cg14414971	0.021(0.032-0.01)	2,97E-04	0.022(0.034-0.01)	3,72E-04	0.016(0.033--0.001)	6,62E-02	0.024(0.033-0.014)	6,62E-06
cg12285409	0.019(0.029-0.008)	7,15E-04	0.021(0.031-0.011)	9,43E-05	0.028(0.043-0.013)	5,47E-04	0.023(0.033-0.014)	6,16E-06
cg20410192	0.008(0.012-0.004)	7,92E-05	0.01(0.014-0.005)	1,12E-04	0.008(0.015-0.002)	1,76E-02	0.008(0.013-0.004)	2,23E-04
cg10473508	0.018(0.027-0.01)	6,37E-05	0.025(0.033-0.017)	1,28E-07	0.014(0.029--0.001)	6,71E-02	0.021(0.029-0.013)	1,06E-06
cg05914439	0.013(0.02-0.006)	3,34E-04	0.014(0.021-0.007)	1,56E-04	0.018(0.029-0.007)	2,70E-03	0.015(0.022-0.009)	2,44E-05
cg21662506	0.015(0.022-0.007)	1,85E-04	0.014(0.022-0.005)	2,82E-03	0.013(0.026-0.001)	3,31E-02	0.02(0.027-0.012)	4,15E-06
cg24760647	0.013(0.019-0.007)	2,73E-05	0.013(0.02-0.007)	1,31E-04	0.016(0.024-0.008)	5,16E-04	0.013(0.02-0.006)	3,80E-04
cg19509894	0.005(0.008-0.003)	1,67E-05	0.003(0.006-0.001)	1,23E-02	0.007(0.011-0.003)	1,48E-03	0.005(0.008-0.003)	4,58E-05
cg05468351	0.015(0.022-0.007)	1,73E-04	0.017(0.025-0.009)	6,97E-05	0.029(0.041-0.016)	8,19E-05	0.017(0.025-0.009)	8,02E-05
cg10780495	0.005(0.007-0.002)	2,12E-04	0.006(0.008-0.003)	9,82E-05	0.006(0.009-0.002)	5,52E-03	0.005(0.008-0.002)	2,47E-04
cg05036147	0.015(0.023-0.007)	2,07E-04	0.017(0.026-0.009)	1,76E-04	0.012(0.025--0.001)	7,81E-02	0.015(0.023-0.007)	3,26E-04

Table e-4: Significant genes associated with stroke in GWAS analyzed by EWAS

Genes	Phenotype
CASZ1	AS
WNT2B	AS
RGS7	CES
KCNK3	AS
TM4SF4	LAS
TM4SF1	LAS
ANK2	AIS
EDNRA	LAS
FGA	AIS
LOC100505841	AIS
NKX2-5	CES
SLC22A7	AS
ZNF318	AS
CDK6	AIS
Chr9p21	AS
LINC01492	LAS
SH3PXD2A	AS
OBFC1	AS
PDE3A	AIS
TBX3	AS
LRCH1	AS
FURIN	AIS
FES	AIS
PRPF8	AIS
ILF3	AIS
SLC44A2	AIS
SMARCA4	AS
LDLR	AS
TSPAN2	LAS
PMF1	AS
SEMA4A	AS
PITX2	CES
FOXF2	AS
HDAC9	LAS
ABO	AIS
MMP12	AIS
ZFHX3	CES
SH2B3	AIS
ZCCHC14	AS
TWIST1	LAS
NOS3	AS
COL4A1	AS
DYRK1A	AS

Table e-5: Significant EWAS associations in genes associated with stroke in GWAS

cpg	CHR	BP	Gene	feat.cgi	logFC(CI)	P
cg07786668*	16	73092391	ZFHX3	5'UTR-island	0,058(0,069-0,048)	2,35E-22
cg12885832*	12	111843885	SH2B3	1stExon-island	0,034(0,043-0,026)	1,82E-13
cg17218495*	19	11071743	SMARCA4	5'UTR-island	0,020(0,027-0,014)	3,29E-09
cg00614832*	16	73092394	ZFHX3	5'UTR-island	0,043(0,056-0,029)	4,49E-09
cg07202430*	12	111843939	SH2B3	1stExon-island	0,023(0,030-0,015)	2,23E-08
cg03614513*	1	115632038	TSPAN2	1stExon-island	0,010(0,014-0,006)	8,21E-07
cg00373279*	19	10765277	ILF3	5'UTR-island	0,013(0,018-0,008)	2,80E-06
cg07266431*	7	92460113	CDK6	Body-shelf	0,026(0,037-0,015)	4,83E-06
cg18346784	1	10754162	CASZ1	5'UTR-island	0,003(0,005-0,002)	5,85E-05
cg21303386	1	241520311	RGS7	5'UTR-island	0,028(0,042-0,014)	8,76E-05
cg17318642	12	111842773	SH2B3	TSS1500-shore	0,006(0,008-0,003)	1,09E-04
cg09811183	1	115631515	TSPAN2	Body-shore	0,026(0,040-0,012)	3,81E-04
cg05101437	7	92262970	CDK6	Body-opensea	0,006(0,009-0,003)	4,76E-04
cg23249336	13	47127304	LRCH1	5'UTR-island	0,007(0,011-0,003)	5,56E-04
cg23231666	1	10699975	CASZ1	Body-island	-0,002(-0,001--0,003)	5,68E-04
cg04443357	13	47128300	LRCH1	Body-island	0,006(0,010-0,003)	5,93E-04
cg20013286	1	113008897	WNT2B	TSS1500-island	0,002(0,004-0,001)	7,47E-04
cg02469210	12	111843919	SH2B3	1stExon-island	0,002(0,003-0,001)	7,61E-04
cg05767489	1	10725218	CASZ1	Body-island	-0,011(-0,004--0,017)	8,29E-04
cg09089230	17	1588195	PRPF8	TSS200-island	0,008(0,012-0,003)	9,52E-04
cg23892580	19	11071746	SMARCA4	5'UTR-island	0,007(0,011-0,003)	1,06E-03
cg25242556	15	91423375	FURIN	Body-island	-0,002(-0,001--0,003)	1,77E-03
cg01446895	12	111856416	SH2B3	Body-island	-0,002(-0,001--0,004)	1,92E-03
cg18885392	7	18832559	HDAC9	Body-opensea	-0,037(-0,014--0,060)	1,96E-03
cg26699757	7	92462202	CDK6	Body-island	0,004(0,007-0,001)	2,57E-03
cg22159483	17	1588198	PRPF8	TSS200-island	0,008(0,013-0,003)	2,57E-03
cg19348484	15	91413236	FURIN	5'UTR-shore	0,023(0,039-0,008)	3,09E-03
cg21295222	1	10712421	CASZ1	Body-shore	-0,010(-0,003--0,017)	4,33E-03
cg12035827	1	10855226	CASZ1	5'UTR-island	0,009(0,016-0,003)	4,34E-03
cg10219512	6	43337449	ZNF318	TSS1500-island	0,002(0,003-0,000)	4,34E-03
cg00907204	7	92461971	CDK6	Body-shore	0,008(0,013-0,002)	4,62E-03
cg20250874	13	47127354	LRCH1	5'UTR-island	0,005(0,009-0,002)	4,72E-03
cg05309964	1	156182615	PMF1	TSS200-island	0,002(0,003-0,001)	4,76E-03
cg00277384	10	105454768	SH3PXD2A	Body-shore	0,014(0,023-0,004)	4,91E-03
cg19128723	10	105668237	OBFC1	Body-opensea	0,003(0,005-0,001)	5,28E-03
cg02722808	10	105595043	SH3PXD2A	Body-opensea	0,017(0,029-0,005)	5,46E-03
cg17480466	1	240986748	RGS7	Body-island	-0,006(-0,002--0,011)	5,74E-03
cg08080174	16	72992580	ZFHX3	Body-shore	0,028(0,047-0,008)	6,46E-03
cg02337568	12	111843385	SH2B3	TSS1500-island	0,003(0,005-0,001)	6,81E-03
cg05278074	7	92442568	CDK6	Body-opensea	-0,027(-0,007--0,047)	7,82E-03

cg06751603	19	11152063	SMARCA4	Body-island	-0,005(-0,001--0,009)	8,18E-03
cg10376337	16	87446508	ZCCHC14	Body-shore	0,007(0,013-0,002)	8,26E-03
cg26573797	16	72991926	ZFHX3	Body-shore	-0,004(-0,001--0,006)	8,38E-03
cg20668662	10	105452885	SH3PXD2A	Body-island	0,002(0,003-0,000)	8,90E-03
cg04910662	16	73083014	ZFHX3	5'UTR-shore	0,013(0,024-0,003)	9,30E-03
cg04121771	3	149191624	TM4SF4	TSS1500- opensea	0,027(0,048-0,007)	9,94E-03
cg11037064	6	43336719	ZNF318	1stExon-island	0,002(0,003-0,000)	1,05E-02
cg10779184	7	92462706	CDK6	5'UTR-island	0,003(0,006-0,001)	1,07E-02
cg18522266	19	11072136	SMARCA4	5'UTR-island	0,002(0,003-0,000)	1,08E-02
cg11084266	1	156187819	PMF1	Body-shore	0,006(0,011-0,001)	1,11E-02
cg11006812	1	113044437	WNT2B	Body-island	0,008(0,014-0,002)	1,12E-02
cg04902826	10	105677453	OBFC1	5'UTR-shore	0,012(0,021-0,003)	1,12E-02
cg20318235	1	10707917	CASZ1	Body-island	-0,009(-0,002--0,017)	1,16E-02
cg25181684	10	105456466	SH3PXD2A	Body-shelf	0,003(0,005-0,001)	1,16E-02
cg07996255	16	73093309	ZFHX3	TSS1500-shore	0,008(0,013-0,002)	1,18E-02
cg13351962	19	11200983	LDLR	Body-island	0,002(0,004-0,001)	1,21E-02
cg16516270	1	156195431	PMF1	Body-opensea	0,004(0,008-0,001)	1,25E-02
cg23246821	3	149095573	TM4SF1	TSS200- opensea	0,009(0,015-0,002)	1,28E-02
cg16935061	6	1395278	FOXF2	3'UTR-shore	-0,045(-0,010--0,080)	1,28E-02
cg04804052	19	11094658	SMARCA4	TSS200-shelf	0,019(0,034-0,004)	1,29E-02
cg23979520	11	102745854	MMP12	TSS200- opensea	-0,015(-0,003--0,027)	1,29E-02
cg09397246	15	91427361	FES	TSS1500-shore	0,027(0,048-0,006)	1,35E-02
cg05972185	16	72972644	ZFHX3	Body-opensea	0,032(0,058-0,007)	1,38E-02
cg10126205	7	19158664	TWIST1	TSS1500-shore	0,008(0,014-0,002)	1,48E-02
cg06330618	15	91428456	FES	1stExon-shore	0,019(0,035-0,004)	1,53E-02
cg11582017	15	91414499	FURIN	5'UTR-island	0,004(0,007-0,001)	1,60E-02
cg20705938	4	111558988	PITX2	TSS1500-shore	0,007(0,013-0,001)	1,63E-02
cg10481072	7	92459517	CDK6	Body-shelf	-0,021(-0,004--0,038)	1,66E-02
cg25647583	15	91427184	FES	TSS1500-shore	0,021(0,039-0,004)	1,72E-02
cg10689147	6	43336734	ZNF318	1stExon-island	0,002(0,004-0,000)	1,86E-02
cg26847100	19	10748778	SLC44A2	Body-shore	0,023(0,042-0,004)	1,94E-02
cg25320499	1	113052563	WNT2B	Body-shore	0,021(0,038-0,003)	1,95E-02
cg07434438	16	72961899	ZFHX3	Body-opensea	0,020(0,037-0,003)	1,95E-02
cg24304658	2	26950770	KCNK3	Body-island	-0,002(0,000--0,003)	1,98E-02
cg02854164	16	87474977	ZCCHC14	Body-opensea	0,007(0,013-0,001)	2,00E-02
cg09754110	1	10714160	CASZ1	Body-island	-0,002(0,000--0,005)	2,11E-02
cg24075935	1	10772782	CASZ1	5'UTR- opensea	0,022(0,041-0,003)	2,12E-02
cg24182603	12	111843578	SH2B3	TSS200-island	0,003(0,005-0,000)	2,26E-02
cg16469740	7	19018055	HDAC9	Body-opensea	-0,019(-0,003--0,035)	2,36E-02
cg26405020	15	91427363	FES	TSS1500-shore	0,020(0,038-0,003)	2,39E-02
cg22719037	12	111842896	SH2B3	TSS1500-shore	0,005(0,009-0,001)	2,41E-02
cg27442412	1	115632570	TSPAN2	TSS1500- island	-0,007(-0,001--0,012)	2,49E-02

cg02818775	10	105362057	SH3PXD2A	Body-island	-0,006(-0,001--0,011)	2,50E-02
cg05102394	4	148402899	EDNRA	Body-shore	0,023(0,042-0,003)	2,55E-02
cg12477923	10	105365600	SH3PXD2A	Body-shelf	0,019(0,036-0,002)	2,62E-02
cg15001636	19	11144178	SMARCA4	Body-island	-0,015(-0,002--0,028)	2,68E-02
cg27455918	7	92370191	CDK6	Body-opensea	-0,017(-0,002--0,031)	2,70E-02
cg05250953	7	92463343	CDK6	5'UTR-island	0,001(0,002-0,000)	2,80E-02
cg09223396	16	72893924	ZFHX3	Body-opensea	-0,004(0,000--0,007)	2,82E-02
cg05180856	15	91428056	FES	5'UTR-shore	0,014(0,027-0,001)	2,94E-02
cg12548629	7	92463334	CDK6	5'UTR-island	0,001(0,003-0,000)	2,96E-02
cg24113439	19	11216129	LDLR	Body-island	0,006(0,011-0,001)	2,97E-02
cg13797285	1	115631930	TSPAN2	Body-island	0,004(0,008-0,000)	2,97E-02
cg04892643	7	18906433	HDAC9	Body-opensea	-0,011(-0,001--0,021)	3,03E-02
cg13539203	2	26950545	KCNK3	Body-island	-0,012(-0,001--0,023)	3,04E-02
cg21644703	19	10735482	SLC44A2	TSS1500-shore	-0,016(-0,002--0,030)	3,05E-02
cg18461436	3	149096828	TM4SF1	TSS1500-opensea	0,008(0,015-0,001)	3,18E-02
cg12081743	7	18535499	HDAC9	TSS1500-opensea	0,004(0,007-0,000)	3,27E-02
cg07974958	1	10707454	CASZ1	3'UTR-shore	-0,006(0,000--0,011)	3,27E-02
cg10483525	16	73080909	ZFHX3	5'UTR-shore	0,012(0,023-0,001)	3,28E-02
cg05483388	3	149192609	TM4SF4	1stExon-opensea	-0,010(-0,001--0,020)	3,42E-02
cg11998200	7	92274858	CDK6	Body-opensea	0,004(0,009-0,000)	3,59E-02
cg24504936	1	10844072	CASZ1	5'UTR-opensea	-0,003(0,000--0,007)	3,69E-02
cg06254768	4	114064176	ANK2	Body-opensea	0,022(0,043-0,001)	3,72E-02
cg26312150	7	19156550	TWIST1	1stExon-island	0,010(0,020-0,001)	3,74E-02
cg10582871	1	10725332	CASZ1	Body-island	-0,004(0,000--0,009)	3,78E-02
cg15635503	19	11101976	SMARCA4	Body-shelf	-0,001(0,000--0,003)	3,81E-02
cg04101363	16	72830284	ZFHX3	Body-opensea	-0,003(0,000--0,006)	3,84E-02
cg13348905	19	10735578	SLC44A2	TSS1500-shore	-0,020(-0,001--0,040)	3,86E-02
cg03738979	19	11145704	SMARCA4	Body-island	-0,001(0,000--0,003)	3,87E-02
cg19590227	1	10755322	CASZ1	5'UTR-island	0,003(0,006-0,000)	3,91E-02
cg09306186	16	73081722	ZFHX3	1stExon-island	0,007(0,013-0,000)	3,94E-02
cg17167425	10	105480849	SH3PXD2A	Body-opensea	0,008(0,016-0,000)	4,01E-02
cg01515508	1	10728270	CASZ1	Body-shelf	-0,001(0,000--0,003)	4,19E-02
cg24102726	1	10855553	CASZ1	5'UTR-island	0,002(0,003-0,000)	4,22E-02
cg12796481	1	10796411	CASZ1	5'UTR-opensea	0,007(0,014-0,000)	4,28E-02
cg17944110	1	10856657	CASZ1	1stExon-island	0,006(0,011-0,000)	4,29E-02
cg13458645	4	111544271	PITX2	Body-island	0,004(0,008-0,000)	4,56E-02
cg22508928	1	10710134	CASZ1	Body-shore	-0,009(0,000--0,018)	4,65E-02
cg09442403	3	149089298	TM4SF1	Body-opensea	0,006(0,011-0,000)	4,73E-02
cg09761690	16	72820681	ZFHX3	3'UTR-shore	-0,009(0,000--0,019)	4,74E-02
cg23136742	2	26950956	KCNK3	Body-island	-0,032(0,000--0,064)	4,76E-02
cg08108188	1	10853918	CASZ1	5'UTR-island	0,004(0,007-0,000)	4,76E-02
cg01517364	16	87446080	ZCCHC14	Body-shore	-0,006(0,000--0,012)	4,77E-02

cg26796245	4	113970820	ANK2	5'UTR- opensea	0,018(0,036-0,000)	4,77E-02
cg05265940	16	87446987	ZCCHC14	Body-island	-0,004(0,000--0,009)	4,82E-02
cg21602363	17	1588247	PRPF8	TSS200-island	0,004(0,008-0,000)	4,86E-02
cg24259688	15	91414844	FURIN	5'UTR-island	-0,006(0,000--0,013)	4,90E-02
cg16278728	12	111885550	SH2B3	Body-opensea	-0,001(0,000--0,002)	4,90E-02

Table e-6: Significant EWAS results by stroke subtypes in genes associated with stroke in GWAS.

					AIS		LAS		CES		SVS		UND	
CpG	CHR	BP	Gene	feat.cgi	logFC(CI)	P	logFC(CI)	P	logFC(CI)	P	logFC(CI)	p.value	logFC(CI)	p.value
cg07786668	16	73092391	ZFH3	5'UTR-island	0.058(0.069-0.048)	2,35E-22	0.081(0.102-0.06)	3,54E-11	0.091(0.115-0.067)	2,13E-10	0.082(0.127-0.038)	6,88E-04	0.089(0.112-0.066)	8,02E-11
cg12885832	12	111843885	SH2B3	1stExon-island	0.034(0.043-0.026)	1,82E-13	0.039(0.051-0.028)	4,28E-09	0.048(0.06-0.035)	9,98E-11	0.042(0.063-0.02)	4,12E-04	0.039(0.051-0.027)	1,30E-08
cg17218495	19	11071743	SMARCA4	5'UTR-island	0.02(0.027-0.014)	3,29E-09	0.021(0.03-0.011)	3,05E-05	0.026(0.036-0.017)	4,54E-07	0.035(0.066-0.004)	2,95E-02	0.027(0.038-0.017)	3,49E-06
cg00614832	16	73092394	ZFH3	5'UTR-island	0.043(0.056-0.029)	4,49E-09	0.048(0.068-0.028)	8,38E-06	0.065(0.087-0.042)	2,06E-07	0.047(0.085-0.01)	1,44E-02	0.065(0.087-0.042)	3,39E-07
cg07202430	12	111843939	SH2B3	1stExon-island	0.023(0.03-0.015)	2,23E-08	0.038(0.052-0.024)	1,08E-06	0.04(0.056-0.024)	5,91E-06	0.043(0.07-0.017)	2,40E-03	0.042(0.057-0.027)	8,27E-07
cg03614513	1	115632038	TSPAN2	1stExon-island	0.01(0.014-0.006)	8,21E-07	0.01(0.015-0.005)	1,01E-04	0.012(0.017-0.007)	8,24E-06	0.006(0.014--0.002)	1,14E-01	0.01(0.015-0.005)	3,53E-04
cg00373279	19	10765277	ILF3	5'UTR-island	0.013(0.018-0.008)	2,80E-06	0.011(0.018-0.004)	1,84E-03	0.013(0.02-0.007)	7,68E-05	0.016(0.026-0.006)	2,83E-03	0.013(0.02-0.006)	4,30E-04
cg07266431	7	92460113	CDK6	Body-shelf	0.026(0.037-0.015)	4,83E-06	0.032(0.046-0.019)	7,19E-06	0.021(0.037-0.006)	8,37E-03	0.025(0.046-0.005)	1,50E-02	0.03(0.045-0.016)	7,18E-05

Table e-7: Summary statistics for mQTLs in *MAP3K1* and *ZFHX3*

SNP	CHR	BP	alleleA	alleleB	MAF	P	Beta
<i>MAP3K1</i>							
rs4801075	18	54723754	C	T	0.0176	4.65E-10	2.423
rs150220406	19	46394268	A	G	0.0146	6.53E-10	-2.222
rs75888346	19	14912618	C	T	0.0058	3.75E-09	4.003
rs59754912	1	91209842	G	T	0.0147	1.14E-08	-2.457
rs67018530	1	91138501	A	G	0.0145	1.27E-08	-2.448
rs78027485	18	54669321	G	T	0.0116	1.94E-08	2.687
rs78504775	18	54674304	A	G	0.0116	1.94E-08	-2.687
rs112831652	5	115819727	C	T	0.0058	1.97E-08	-3.777
rs75916710	5	115821451	C	G	0.0058	1.97E-08	3.777
rs76283481	5	115822218	A	G	0.0058	1.97E-08	-3.777
rs80080318	7	54069673	A	T	0.0117	3.36E-08	2.639
rs55932208	17	16838875	C	G	0.0148	6.10E-08	2.325
rs11200507	10	124016195	C	T	0.0116	6.64E-08	-2.585
rs11200508	10	124016409	A	G	0.0116	6.64E-08	-2.585
rs11200509	10	124016442	C	G	0.0116	6.64E-08	-2.585
rs4751882	10	124017887	A	G	0.0116	6.64E-08	-2.585
rs75926138	10	124018378	A	G	0.0116	6.64E-08	-2.585
rs4751884	10	124019956	A	G	0.0116	6.64E-08	2.585
rs4752677	10	124021233	C	T	0.0116	6.64E-08	2.585
rs79671903	10	124024365	C	T	0.0116	6.64E-08	-2.585
rs10887115	10	124029062	C	T	0.0116	6.64E-08	-2.585
rs11200522	10	124031218	A	G	0.0116	6.64E-08	-2.585
rs75189955	10	124031335	A	T	0.0116	6.64E-08	2.585
rs10887117	10	124032655	A	G	0.0116	6.64E-08	2.585
rs10887118	10	124042621	C	G	0.0116	6.64E-08	-2.585
rs11200528	10	124042994	A	G	0.0116	6.64E-08	2.585
rs10887119	10	124043596	A	G	0.0116	6.64E-08	-2.585
rs11200535	10	124050439	A	G	0.0116	6.64E-08	-2.585
rs11200537	10	124053547	A	C	0.0116	6.64E-08	2.585
rs11200538	10	124055736	A	G	0.0116	6.64E-08	2.585
rs10887121	10	124055922	A	G	0.0116	6.64E-08	2.585
rs78017737	10	124058935	C	T	0.0116	6.64E-08	-2.585
rs11200541	10	124061121	C	T	0.0116	6.64E-08	2.585
rs74949254	10	124054794	C	G	0.0117	6.79E-08	-2.590
rs148168465	8	122066679	G	T	0.0353	8.53E-08	-1.399
rs139061303	8	122067069	C	T	0.0353	8.53E-08	-1.399
rs61748971	13	100381591	C	G	0.0175	1.33E-07	2.088
rs11029139	11	26103163	A	G	0.0145	1.48E-07	2.278
rs2349610	11	26108605	C	T	0.0145	1.48E-07	-2.278
rs113074219	9	11010827	C	T	0.0059	1.86E-07	-3.564
rs17695152	1	77759856	G	T	0.0237	2.33E-07	-1.792
rs142726777	8	122065839	C	T	0.0325	2.36E-07	1.400
rs144099668	8	122067155	G	T	0.0325	2.36E-07	1.400
rs12302973	12	69369309	A	T	0.0266	2.40E-07	1.749
rs8058442	16	81401142	C	T	0.0176	2.52E-07	-2.076
rs58655554	16	81402276	A	G	0.0176	2.52E-07	-2.076

rs150186906	8	122066502	C	T	0.0296	2.54E-07	-1.450
rs141670571	8	122066969	C	G	0.0296	2.54E-07	1.450
rs140625613	8	122067157	C	T	0.0296	2.54E-07	1.450
rs12217181	10	124058904	C	T	0.0145	2.88E-07	2.222
rs78175645	5	52939032	A	G	0.1287	3.80E-07	-0.817
rs1347231	8	122100181	C	T	0.0473	4.67E-07	-1.180
rs1864950	5	148409765	C	T	0.0179	5.41E-07	-2.016
rs138570613	8	122066886	A	G	0.0325	5.60E-07	-1.345
rs74030547	16	81174695	A	G	0.0518	5.84E-07	1.157
rs74030549	16	81174785	C	T	0.0518	5.84E-07	-1.157
rs117006360	16	81174978	A	G	0.0518	5.84E-07	1.157
rs117414860	16	81174992	G	T	0.0518	5.84E-07	1.157
rs4871159	8	122099756	A	C	0.0471	6.03E-07	1.172
rs11200544	10	124068989	A	G	0.0149	6.05E-07	2.162
rs79437924	8	122088028	C	G	0.0468	6.41E-07	-1.167
rs73317195	8	122088533	G	T	0.0468	6.41E-07	-1.167
rs16894647	8	122093191	C	G	0.0468	6.41E-07	-1.167
rs73318906	8	122095342	A	G	0.0468	6.41E-07	-1.167
rs9935161	16	11524858	C	G	0.0268	7.85E-07	-1.608
rs4237539	10	124065660	C	G	0.0147	8.07E-07	-2.148
rs4317918	10	124065870	C	T	0.0147	8.07E-07	2.148
rs1022780	10	124068439	G	T	0.0148	8.80E-07	2.147
rs149703719	8	122054525	C	T	0.0497	9.30E-07	1.125
rs73315279	8	122055019	A	G	0.0497	9.30E-07	1.125
rs73317155	8	122059785	C	G	0.0497	9.30E-07	-1.125
rs16894592	8	122060129	A	G	0.0497	9.30E-07	-1.125
rs75299560	8	122060233	A	G	0.0497	9.30E-07	-1.125
rs4483161	8	122060882	A	G	0.0497	9.30E-07	1.125
rs4310214	8	122060930	C	T	0.0497	9.30E-07	-1.125
rs4300011	8	122061094	A	G	0.0497	9.30E-07	1.125
rs5010752	8	122061168	C	T	0.0497	9.30E-07	-1.125
rs73317162	8	122061193	A	G	0.0497	9.30E-07	1.125
rs7835412	8	122061527	C	T	0.0497	9.30E-07	-1.125
rs7817321	8	122061749	C	T	0.0497	9.30E-07	1.125
rs7813342	8	122061907	C	T	0.0497	9.30E-07	-1.125
rs7813643	8	122062068	C	T	0.0497	9.30E-07	-1.125
rs10505389	8	122062152	C	T	0.0497	9.30E-07	1.125
rs74476105	8	122062344	A	T	0.0497	9.30E-07	-1.125
rs75313161	8	122062350	A	G	0.0497	9.30E-07	-1.125
rs16894623	8	122062457	A	C	0.0497	9.30E-07	-1.125
rs73317172	8	122062488	G	T	0.0497	9.30E-07	-1.125
rs140422925	8	122067824	C	T	0.0497	9.30E-07	-1.125
rs9969654	8	122069760	C	T	0.0497	9.30E-07	1.125
rs2117064	8	122072128	C	G	0.0497	9.30E-07	1.125
rs6469984	8	122078427	C	G	0.0497	9.30E-07	-1.125
rs74438538	5	52851864	G	T	0.1316	9.73E-07	0.786
rs75644904	5	52852055	A	G	0.1316	9.73E-07	-0.786
rs79379207	5	52861699	A	G	0.1316	9.73E-07	0.786
rs56291043	5	52861800	A	G	0.1316	9.73E-07	-0.786
ZFHX3							
rs113569939	12	96033048	C	G	0.0211	6.85E-11	2.102
rs111761836	12	96037517	A	G	0.0211	6.85E-11	2.102

rs113035257	12	96031764	A	G	0.0210	7.36E-11	-2.099
rs112227669	12	96036797	C	T	0.0210	7.36E-11	-2.099
rs484416	1	10491209	A	G	0.0265	1.18E-09	-2.010
rs2847336	1	10495110	C	T	0.0294	2.48E-08	1.761
rs117839871	20	48820986	A	G	0.0366	2.74E-08	-1.648
rs666103	1	10511589	C	T	0.0292	3.02E-08	1.750
rs1307497	1	10512056	C	T	0.0292	3.02E-08	1.750
rs944973	1	10516393	C	G	0.0294	3.35E-08	-1.749
rs74621069	20	41823592	C	T	0.0407	5.05E-08	1.379
rs79608766	20	41823630	C	T	0.0407	5.05E-08	1.379
rs12047540	1	111834864	A	G	0.0325	1.08E-07	1.455
rs68144419	1	111836930	C	T	0.0325	1.08E-07	-1.455
rs12022369	1	111837751	A	T	0.0325	1.08E-07	1.455
rs12023362	1	111838942	A	T	0.0325	1.08E-07	1.455
rs56101238	14	102244476	A	G	0.0240	1.41E-07	-1.904
rs12184718	13	20885789	A	T	0.0647	1.86E-07	-1.105
rs77287810	1	174981454	A	T	0.0325	2.65E-07	1.576
rs80154038	7	86076026	A	G	0.0233	2.79E-07	1.815
rs7785342	7	86095053	A	C	0.0233	2.79E-07	-1.815
rs6950753	7	86100128	A	G	0.0233	2.79E-07	-1.815
rs73907527	20	41834758	A	G	0.0322	2.84E-07	-1.437
rs113776847	20	41835211	A	T	0.0322	2.84E-07	-1.437
rs676995	1	10498453	C	T	0.0322	3.06E-07	1.553
rs1307489	1	10507236	A	G	0.0322	3.06E-07	-1.553
rs1308520	1	10510682	A	G	0.0322	3.06E-07	-1.553
rs6540931	1	10512509	C	T	0.0322	3.06E-07	1.553
rs2483686	1	10515414	C	T	0.0322	3.06E-07	1.553
rs8084639	18	4456714	C	T	0.0237	3.09E-07	1.814
rs7235642	18	4454027	C	G	0.0238	3.19E-07	1.816
rs7236361	18	4454161	A	T	0.0238	3.19E-07	1.816
rs7537164	1	10519090	C	T	0.0324	3.32E-07	-1.553
rs117184875	20	48824804	A	C	0.0357	3.43E-07	1.494
rs3927586	1	10518596	C	T	0.0325	3.44E-07	-1.555
rs10898428	11	85688893	C	T	0.0500	3.58E-07	1.192
rs78278955	11	85726956	A	C	0.0500	3.58E-07	1.192
rs11234514	11	85730012	A	G	0.0500	3.58E-07	1.192
rs10170617	2	39640893	A	G	0.1105	5.39E-07	-0.850
rs16987892	20	41797018	G	T	0.0237	5.97E-07	-1.572
rs73280321	20	41800274	A	G	0.0237	5.97E-07	-1.572
rs73280324	20	41801407	C	T	0.0237	5.97E-07	1.572
rs74620441	20	41802141	C	T	0.0237	5.97E-07	-1.572
rs112463332	16	9278010	A	C	0.0234	6.74E-07	1.763
rs7864025	9	6957604	G	T	0.0238	7.01E-07	-1.846
rs956229	7	86092160	A	G	0.0292	8.23E-07	1.581
rs10231314	7	86092334	C	T	0.0292	8.23E-07	1.581
rs10250585	7	86092926	A	T	0.0292	8.23E-07	-1.581
rs10235600	7	86093414	C	T	0.0292	8.23E-07	1.581
rs10265248	7	86093744	A	G	0.0292	8.23E-07	1.581
rs59251779	7	86094206	C	T	0.0292	8.23E-07	1.581
rs61052994	7	86094401	C	T	0.0292	8.23E-07	-1.581
rs183100113	7	86154356	A	G	0.0205	8.45E-07	-1.847
rs28709459	7	86105944	C	T	0.0235	8.55E-07	1.742

rs17160821	7	86107405	C	T	0.0235	8.55E-07	-1.742
rs10281570	7	86120164	C	T	0.0235	8.55E-07	1.742

5- Figures

Figure e-1

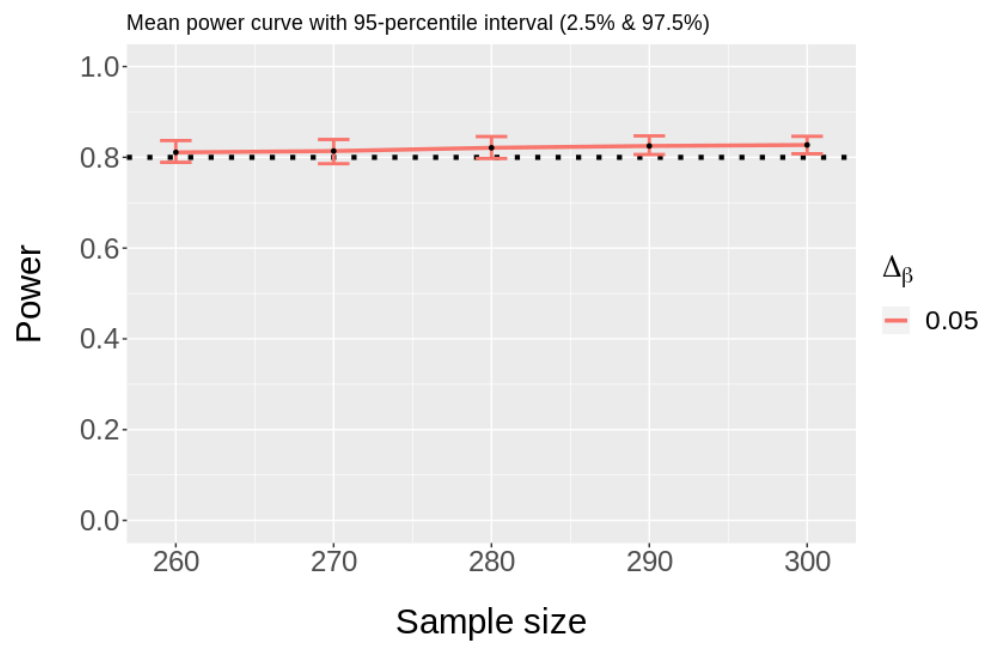


Figure e-2

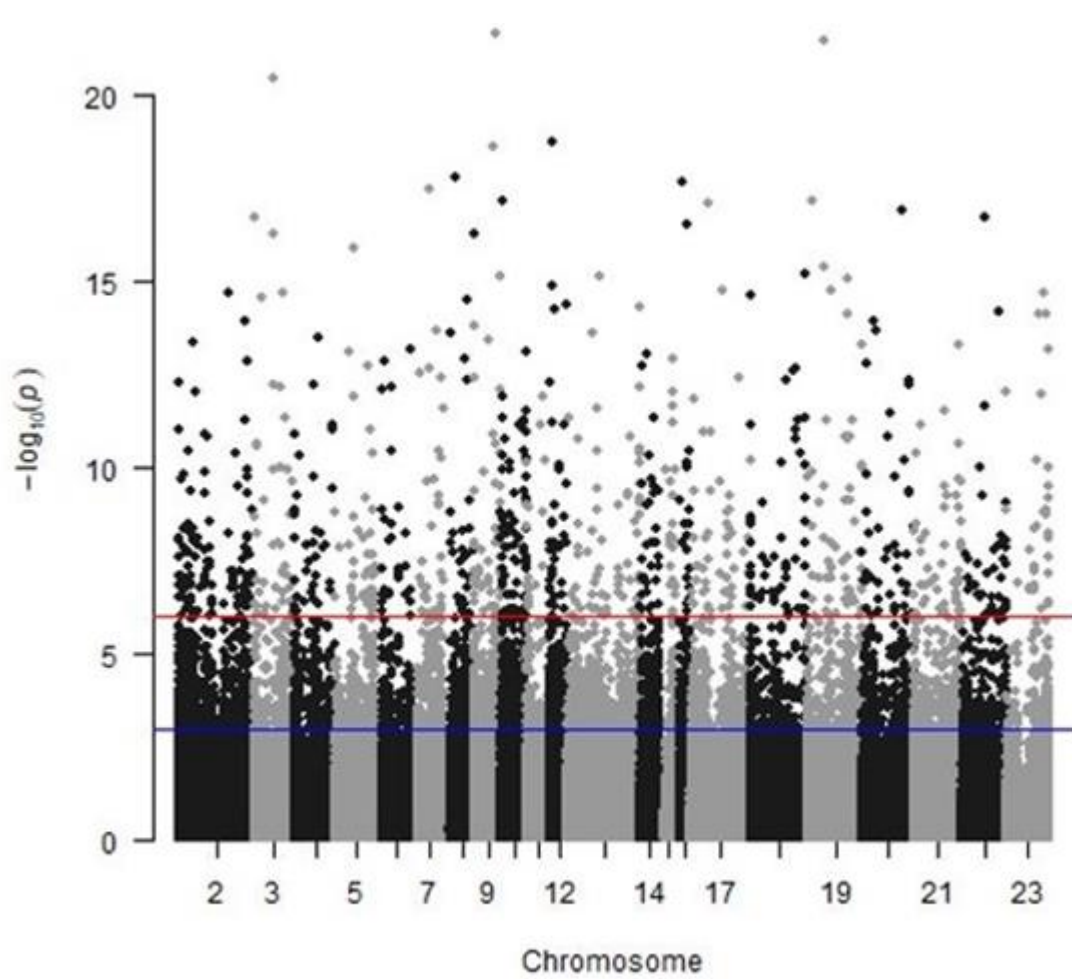
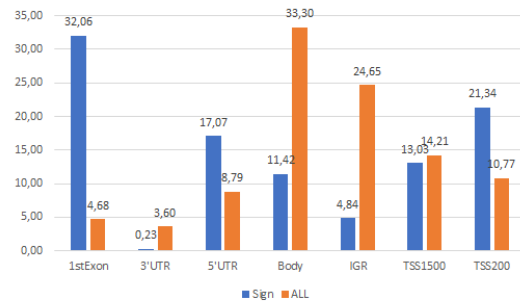


Figure e-3

a)



b)

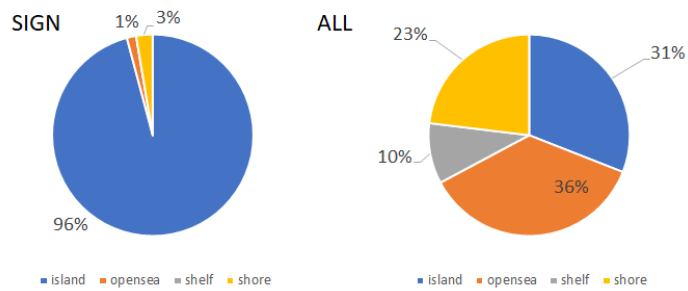
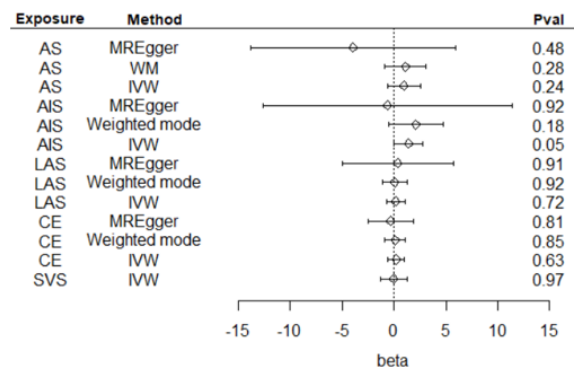
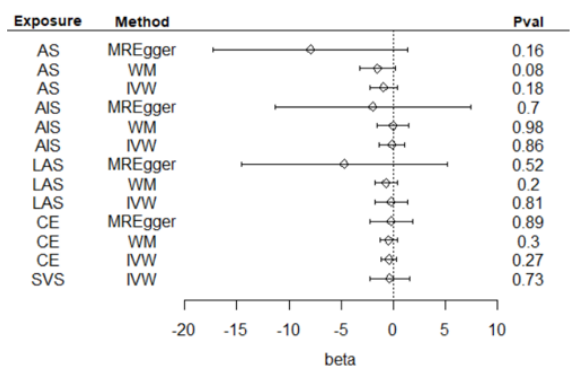


Figure e-4

A)



B)



6- e-References

9. Gallego-Fabrega C, Carrera C, Reny J-L, Fontana P, Slowik A, Pera J, Pezzini A, Serrano-Heras G, Segura T, Martí-Fàbregas J, et al. TRAF3 Epigenetic Regulation Is Associated With Vascular Recurrence in Patients With Ischemic Stroke. *Stroke J. Cereb. Circ.* 2016;
10. Gallego-Fabrega C, Carrera C, Reny J-L, Fontana P, Slowik A, Pera J, Pezzini A, Serrano-Heras G, Segura T, Martí-Fàbregas J, et al. PPM1A Methylation Is Associated With Vascular Recurrence in Aspirin-Treated Patients. *Stroke J. Cereb. Circ.* 2016;47:1926–9.
15. Min JL, Hemani G, Davey Smith G, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinforma. Oxf. Engl.* 2018;34:3983–3989.
16. Fernández-Cadenas I, Mendióroz M, Giralt D, Nafria C, Garcia E, Carrera C, Gallego-Fabrega C, Domingues-Montanari S, Delgado P, Ribó M, et al. GRECOS Project (Genotyping Recurrence Risk of Stroke): The Use of Genetics to Predict the Vascular Recurrence after Stroke. *Stroke.* 2017;
17. Carrera C, Cullell N, Torres-Águila N, Muiño E, Bustamante A, Dávalos A, López-Cancio E, Ribó M, Molina CA, Giralt-Steinhauer E, et al. Validation of a clinical-genetics score to predict hemorrhagic transformations after rtPA. *Neurology.* 2019;93:e851–e863.
18. Roquer J, Rodríguez-Campello A, Gomis M, Jiménez-Conde J, Cuadrado-Godia E, Vivanco R, Giralt E, Sepúlveda M, Pont-Sunyer C, Cucurella G, et al. Acute stroke unit care and early neurological deterioration in ischemic stroke. *J. Neurol.* 2008;255:1012–1017.
19. Sayols-Baixeras S, Lluís-Ganella C, Subirana I, Salas LA, Vilahur N, Corella D, Muñoz D, Segura A, Jimenez-Conde J, Moran S, et al. Identification of a new locus and validation of previously reported loci showing differential methylation associated with smoking. The REGICOR study. *Epigenetics.* 2015;10:1156–1165.
20. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, Beck S. ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics.* 2014;
21. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res.* 2017;45:e22.

Artículo 3:

Título: ICA1L Is Associated with Small Vessel Disease: A Proteome-Wide Association Study in Small Vessel Stroke and Intracerebral Haemorrhage.

Revista: International Journal of Molecular Sciences.

Objetivo: Identificar proteínas reguladas genéticamente asociadas al riesgo de ictus lacunar y compararlo con las hemorragias intracerebrales mediante integración de datos genómicos con datos proteómicos utilizando la estrategia de PWAS.

Resumen:

Antecedentes y objetivos:

Los ictus lacunares (SVS) y las hemorragias intracerebrales (HIC) se producen de forma aguda en la enfermedad cerebral de pequeño vaso (SVD). Estudios genéticos previos que han combinado ambos fenotipos han identificado 3 loci asociados con ambos rasgos. Sin embargo, no se ha estudiado la regulación genética *en cis* a nivel de proteína asociada con la SVD.

Métodos:

Realizamos un estudio de asociación de todo el proteoma (PWAS) usando FUSION para integrar datos genómicos (GWAS) y datos proteómicos cerebrales para identificar mecanismos comunes que regulan tanto SVS como HIC. Los proteomas cerebrales los obtuvimos del estudio ROS/MAP (N=376 sujetos y 1443 proteínas) de la región de la corteza prefrontal dorsolateral (dPFC) y los datos GWAS de ictus lacunar a partir del estudio MEGASTROKE (N=237 511) y MTAG-ICH-SVS de Chung et al. (N=240,269). Realizamos PWAS y luego análisis de colocación con COLOC. Los resultados significativos y nominales se validaron utilizando datos proteómicos de replicación de la misma región cerebral (N=152). Los resultados replicados (valor $q < 0,05$) se investigaron más a fondo para determinar la relación de causalidad utilizando randomización mendeliana con SMR.

Resultados:

Una proteína (ICA1L) se asoció significativamente con el riesgo de ictus lacunar ($z = -4,42$ y $P = 9,6 \times 10^{-06}$) e HIC no lobar ($z = -4,8$ y $p = 1,58 \times 10^{-06}$) en el PWAS, con una alta probabilidad posterior 4 de colocación. En el PWAS de validación, ICA1L permaneció asociado significativamente con ambos fenotipos. Los resultados de SMR para ICA1L indicaron una asociación causal de los niveles de expresión de proteína en el cerebro con SVS (valor $p = 3,66 \times 10^{-05}$) e HIC no lobar ($p = 1,81 \times 10^{-05}$).

Conclusiones:

Nuestros resultados muestran que la asociación de ICA1L con SVS e HIC no lobar está condicionada por la regulación en cis de sus niveles de proteína en el cerebro.



Article

ICA1L Is Associated with Small Vessel Disease: A Proteome-Wide Association Study in Small Vessel Stroke and Intracerebral Haemorrhage

Natalia Cullell ^{1,2,3,4}, Cristina Gallego-Fábrega ¹ , Jara Cárcel-Márquez ^{1,5} , Elena Muíño ¹, Laia Lluçà-Carol ¹ , Miquel Lledós ¹ , Jesús M. Martín-Campos ¹ , Jessica Molina ², Laura Casas ², Marta Almeria ², Israel Fernández-Cadenas ^{1,2,*} and Jerzy Krupinski ^{2,*}

- ¹ Stroke Pharmacogenomics and Genetics Group, Biomedical Research Institute Sant Pau, 08041 Barcelona, Spain; ncullell@mutuaterrassa.cat (N.C.); cristina.gallego.fabrega@gmail.com (C.G.-F.); jara.carcel@gmail.com (J.C.-M.); elena.muinho@gmail.com (E.M.); laialluciacarol@gmail.com (L.L.-C.); miquel.lledos@gmail.com (M.L.); jmartinca@santpau.cat (J.M.M.-C.)
- ² Neurology Department, Hospital Universitari MútuaTerrassa, 08221 Terrassa, Spain; jessicamolina@mutuaterrassa.cat (J.M.); lcasas@mutuaterrassa.cat (L.C.); malmeria@mutuaterrassa.cat (M.A.)
- ³ Fundació Docència i Recerca MútuaTerrassa, 08221 Terrassa, Spain
- ⁴ Department of Medicine, Universitat de Barcelona, 08036 Barcelona, Spain
- ⁵ Department of Medicine, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain
- * Correspondence: israelcadenas@yahoo.es (I.F.-C.); jkrupinski@mutuaterrassa.es (J.K.)



Citation: Cullell, N.; Gallego-Fábrega, C.; Cárcel-Márquez, J.; Muíño, E.; Lluçà-Carol, L.; Lledós, M.; Martín-Campos, J.M.; Molina, J.; Casas, L.; Almeria, M.; et al. ICA1L Is Associated with Small Vessel Disease: A Proteome-Wide Association Study in Small Vessel Stroke and Intracerebral Haemorrhage. *Int. J. Mol. Sci.* **2022**, *23*, 3161. <https://doi.org/10.3390/ijms23063161>

Academic Editor: Teresa Gasull

Received: 10 February 2022

Accepted: 10 March 2022

Published: 15 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Small vessel strokes (SVS) and intracerebral haemorrhages (ICH) are acute outcomes of cerebral small vessel disease (SVD). Genetic studies combining both phenotypes have identified three loci associated with both traits. However, the genetic cis-regulation at the protein level associated with SVD has not been studied before. We performed a proteome-wide association study (PWAS) using FUSION to integrate a genome-wide association study (GWAS) and brain proteomic data to discover the common mechanisms regulating both SVS and ICH. Dorsolateral prefrontal cortex (dPFC) brain proteomes from the ROS/MAP study (N = 376 subjects and 1443 proteins) and the summary statistics for the SVS GWAS from the MEGASTROKE study (N = 237,511) and multi-trait analysis of GWAS (MTAG)-ICH-SVS from Chung et al. (N = 240,269) were selected. We performed PWAS and then a co-localization analysis with COLOC. The significant and nominal results were validated using a replication dPFC proteome (N = 152). The replicated results (q -value < 0.05) were further investigated for the causality relationship using summary data-based Mendelian randomization (SMR). One protein (ICA1L) was significantly associated with SVS (z -score = -4.42 and p -value = 9.6×10^{-6}) and non-lobar ICH (z -score = -4.8 and p -value = 1.58×10^{-6}) in the discovery PWAS, with a high co-localization posterior probability of 4. In the validation PWAS, ICA1L remained significantly associated with both traits. The SMR results for ICA1L indicated a causal association of protein expression levels in the brain with SVS (p -value = 3.66×10^{-5}) and non-lobar ICH (p -value = 1.81×10^{-5}). Our results show that the association of ICA1L with SVS and non-lobar ICH is conditioned by the cis-regulation of its protein levels in the brain.

Keywords: small vessel disease; small vessel stroke; intracerebral haemorrhage; PWAS; proteome-wide association study; genome-wide association study (GWAS)

1. Introduction

Cerebral small vessel disease (SVD) is a chronic and progressive heterogeneous disorder affecting small arteries, arterioles, venules, and capillaries in the brain [1,2]. Among the different outcomes of SVD, small vessel stroke (SVS) and spontaneous intracerebral haemorrhage (ICH) are the most acute and devastating manifestations [3–5]. One-quarter

of all ischemic strokes (IS) are SVS [6], and ICH is the most common cause of haemorrhagic stroke [7]. The incidence of SVD increases with age, and it is, thus, expected that with the increase in life expectancy, the frequency of SVD manifestations will also increase in following years [1,3].

SVD causes parenchyma lesions, such as white matter hyperintensities (WMHs), enlarged perivascular spaces, lacunes, small subcortical infarcts, cerebral microbleeds, and brain atrophy, among others [1,3]. These parenchyma lesions are used as surrogate markers of SVD and can be quantified using brain imaging techniques. The extent of these lesions predicts poor outcomes after ICH [8] and SVS [9]. The two major pathophysiological mechanisms beyond SVD are arteriosclerosis and cerebral amyloid angiopathy (CAA) [1]. Cardiovascular risk factors, mainly hypertension, are associated with arteriosclerotic SVD. It can cause the rupture or occlusion of small arteries and arterioles, producing non-lobar (or deep) ICH and SVS, respectively [10].

There is no effective treatment for SVD, apart from controlling cardiovascular risk factors [4]. Genome-wide association studies (GWAS) have been helpful in identifying genes involved in SVD, analysing magnetic resonance image (MRI) SVD surrogate markers [11–21], and, moreover, specifically studying SVS and ICH.

Two GWAS have been performed with ICH patients. The first study identified a risk locus for non-lobar ICH in 1q22, overlapping with *PMF1* and *SLC25A44* genes [5,11,12]. The second performed a meta-analysis that allowed for the identification of *APOE* associated with lobar ICH [4]. In order to improve the statistical power of these studies, Chung et al. applied a novel strategy, using a multi-trait analysis of GWAS (MTAG) and combining genotype data from two related phenotypes, ICH and SVS. This approach permitted the identification of two novel associations in 2q33 (*ICA1L*) and 13q34 (*COL4A2* and *COL4A1*) associated with non-lobar ICH [4]. Interestingly, these loci presented a nominal association ($p\text{-value} < 5 \times 10^{-3}$) with non-lobar ICH and SVS in the corresponding GWAS but not at a genome-wide level [4]. Three different GWAS in SVS have identified significant associations with this phenotype in 16q24.2 (*ZCCHC14*) [6], 2q33.2 (*ICA1L* and *WDR12*) [22], *ULK4*; *SPI1-SLC39A13-PSMC3-RAPSN*, and *ZBTB14-EPB41L3* [23]. Recently, an MTAG combining SVS and WMH allowed for the identification of seven loci associated with SVS for the first time but some of them were previously associated with ICH (*SLC25A44-PMF1-BGLAP*, *LOX-ZNF474-LOC100505841*, *FOXF2-FOXQ1*, *SH3PXD2A*, *VTA1-GPR126*, *HTRA1-ARMS2*, and *COL4A2*) [23].

Proteome-wide association studies (PWAS) have been proposed as a novel strategy to recover causal protein-coding genes and to identify new associations for a disease. Since proteins are the focus of PWAS [24], it is a potentially beneficial tool for the identification of drug targets for disease [25]. This method, combined with Bayesian or Mendelian randomization, allows for the identification of gene-phenotype associations causally mediated by protein levels [24–26]. We performed a proteome-wide association study (PWAS) to integrate the genomic and proteomic data to discover common mechanisms regulating the two main acute manifestations of SVD: SVS and ICH.

2. Results

2.1. Stringent Criterion PWAS Analysis

From the 1475 proteins evaluated in the discovery cohort, only *ICA1L* reached a significant q -value in association with SVS ($z\text{-score} = -4.93$; $q\text{-value} = 1.42 \times 10^{-2}$) and with non-lobar ICH ($z\text{-score} = -4.8$; $q\text{-value} = 2.33 \times 10^{-3}$) (Table 1). The results from the Bayesian colocalization analysis showed a high posterior probability for the rs7582720 to be the causal variant for both the *ICA1L* levels and the significant GWAS traits (SVS and non-lobar ICH) (posterior probability 4 (PP4) = 0.9 for SVS and non-lobar ICH) (Table 1). The association of *ICA1L* with SVS and non-lobar ICH was the only one to reach the stringent criterion, and it was considered for the replication stage. Both associations resulted as significant in the replication analysis (the $z\text{-score} = -3.55$; $q\text{-value} = 5.08 \times 10^{-3}$, and

PP4 = 0.99 for SVS and the z-score = -4.07 , q -value = 6.19×10^{-4} , and PP4 = 0.99 for non-lobar ICH) (Table 1).

Table 1. PWAS results for the stringent criteria. Significant PWAS associations using the stringent criteria in the discovery and replication. P0–P1: initial-final position for the analysed protein; pQTL ID: ID for the pQTL strongly associated with ICA1L; pQTL z-score: z-score for the pQTL strongly associated with ICA1L; pQTL z-score: z-score in the correspondent GWAS for the best pQTL; NSNP: number of SNPs in the locus; TWAS z-score: z-score for the PWAS association; PWAS p -value: p -value for the PWAS association; PP4: posterior probability 4 in the COLOC analysis; q -value: FDR-adjusted PWAS p -value.

Trait	Protein	Chr	P0–P1	pQTL ID	pQTL z-Score	pQTL GWAS z-Score	Discovery				Replication				
							NSNP	PWAS z-Score	PWAS p -Value	PP4	q -Value	PWAS z-Score	PWAS p -Value	PP4	q -Value
SVS	ICA1L	2	203640690	rs7582720	7.21	−4.426	81	−4.43	9.60×10^{-6}	0.90	1.42×10^{-2}	−3.55	3.91×10^{-4}	0.99	5.08×10^{-3}
Non-lobar ICH			203736708			−4.801		−4.80	1.58×10^{-6}	0.93	2.33×10^{-3}	−4.07	4.76×10^{-5}	0.99	6.19×10^{-4}

We also tested, specifically, those proteins from loci associated previously with SVD. The PWAS aimed to interrogate 26 proteins corresponding to 13 different loci previously associated with ICH (4 loci) and/or SVS (12 loci). From these, only ICA1L was present among the proteins analysed in the PWAS.

2.2. SMR and Conditional Analysis

In order to ensure that the genetic association of ICA1L with SVS and non-lobar ICH was mediated by the cis-regulation of the ICA1L protein levels, SMR and a conditional analysis were performed. The SMR results indicated that the pQTL was associated with SVS (p -value = 3.66×10^{-5}) and non-lobar ICH (p -value = 1.81×10^{-5}) through the ICA1L protein expression regulation (Table 2, Figure 1). Similarly, the conditional analysis plot showed a conditional effect by the expression levels of ICA1L on the GWAS results (Figure 2). Table 2 shows the Summary statistics for the SMR results with the beta (b_{SMR}) and p -values (p_{SMR}) for all the significant and replicated associations from PWAS.

Table 2. SMR results.

Trait	Protein	b_{SMR}	p_{SMR}
SVS	ICA1L	−2.78	3.66×10^{-5}
SVS	SPATA20	−0.24	3.69×10^{-1}
ICH	ICA1L	−1.55	1.81×10^{-5}
Non-lobar ICH	ICA1L	−1.55	1.81×10^{-5}
Lobar ICH	SPATA20	−0.05	7.98×10^{-1}

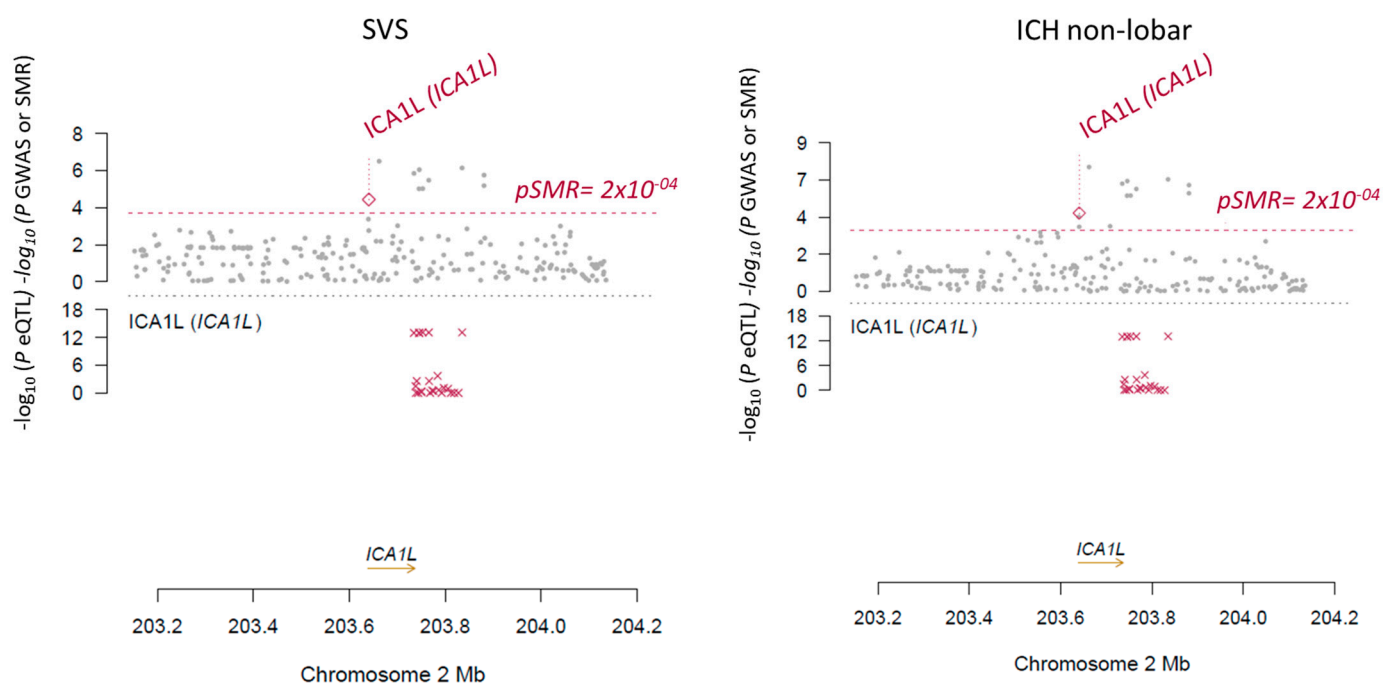


Figure 1. Significant SMR results for the significant and replicated PWAS results. Each “x” in figure represents a pQTL for ICA1L.

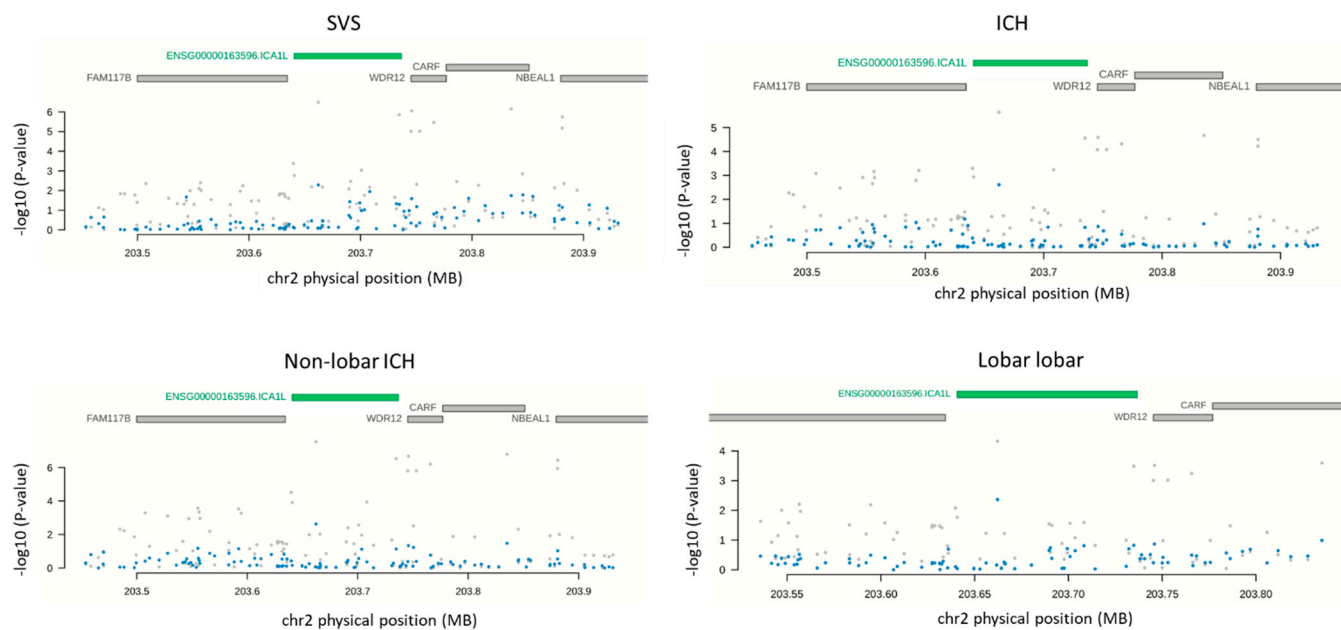


Figure 2. GWAS results before conditioning by ICA1L levels (grey) or after conditioning by protein levels (blue).

2.3. Open Criterion PWAS Analysis

In considering the open criterion (p -value $< 1 \times 10^{-5}$ in the discovery PWAS), five proteins reached a nominal association with all ICH, four proteins with SVS (excluding ICA1L) and lobar ICH, and two proteins with non-lobar ICH (excluding ICA1L) (Table 3). In the replication of the open analysis, only SPATA20 and ICA1L were significant (p -value < 0.05), with a PP4 higher than 0.5 in the replication proteome panel. SPATA20 was significantly associated with SVS and lobar ICH (q -value = 1.95×10^{-3} and PP4 = 0.97 for SVS; 8.82×10^{-3} and PP4 = 0.55 for lobar ICH). ICA1L was significantly associated with all ICH (q -value = 1.25×10^{-2} and PP4 = 0.99 for all ICH). Two proteins (NRBF2 and OSBPL11) could not be evaluated in the replication because they were not present in the panel (Table 3). Results from the SMR indicated that an ICA1L abundance was mediating the association through the pQTL and all ICH (p -value = 3.66×10^{-5}). However, the SMR results for the SPATA20 were not significant (p -value = 3.69×10^{-1} for SVS and 7.98×10^{-1} for lobar ICH) (Table 2). The conditional analyses for both proteins showed that the genetic results were conditionally affected by a protein abundance (Figures 2 and 3).

Table 3. PWAS results for the open criteria. The results for the discovery and replication of PWAS associations for proteins with a nominal p -value in the discovery (open criteria).

Trait	Protein	Chr	pQTL ID	Discovery						Replication				
				pQTL Z-Score	pQTL GWAS Z-Score	NSNP	PWAS Z-Score	PWAS p-Value	PP4	q-Value	PWAS Z-Score	PWAS p-Value	PP4	q-Value
SVS	SPATA20	17	rs878619	−10.89	−2.917	102	3.71	2.10×10^{-4}	0.17	3.10×10^{-1}	3.79	1.50×10^{-4}	0.97 *	1.95×10^{-3}
SVS	ALDH2	12	rs4648328	−7.25	−3.437	61	3.58	3.38×10^{-4}	0.29	4.99×10^{-1}	3.38	7.28×10^{-4}	0.27	9.46×10^{-3}
SVS	EXOC6	10	rs980204	4.98	−2.469	210	−3.58	3.41×10^{-4}	0.19	5.03×10^{-1}	3.37	7.36×10^{-1}	-	1
SVS	OSBPL11	3	rs2922170	4.11	3577	111	3.58	3.48×10^{-4}	0.35	5.14×10^{-1}	-	-	-	-
ICH	ICA1L	2	rs7582720	7.21	−3.935	81	−3.94	8.32×10^{-5}	0.68	1.23×10^{-1}	−3.30	9.66×10^{-4}	0.99 *	1.25×10^{-2}
ICH	NRBF2	10	rs4379723	−4.19	−3.96	77	3.85	1.19×10^{-4}	0.18	1.76×10^{-1}	-	-	-	-
ICH	DLGAP2	8	rs7842425	4.32	−3.742	225	−3.74	1.83×10^{-4}	0.12	2.70×10^{-1}	1.14	2.54×10^{-1}	-	1
ICH	SPATA20	17	rs878619	−10.89	−2.89	102	3.58	3.47×10^{-4}	0.12	5.12×10^{-1}	3.27	1.09×10^{-3}	0.33	1.41×10^{-2}
ICH	MADD	11	rs11570115	−9.09	−2.565	99	3.54	4.06×10^{-4}	0.03	5.99×10^{-1}	1.88	6.04×10^{-2}	-	0.78
Non-lobar ICH	SPATA20	17	rs878619	−10.89	−2.672	102	3.69	2.19×10^{-4}	0.42	3.23×10^{-1}	3.51	4.44×10^{-4}	0.27	5.77×10^{-3}
Non-lobar ICH	NRBF2	10	rs4379723	−4.19	−3.569	77	3.43	6.00×10^{-4}	0.29	8.90×10^{-1}	-	-	-	-
Lobar ICH	MRVI1	11	rs753002	−5.36	2.452	211	−3.69	2.20×10^{-4}	0.02	3.25×10^{-1}	−1.87	6.22×10^{-2}	-	0.80
Lobar ICH	NRBF2	10	rs4379723	−4.19	−3.804	77	3.68	2.33×10^{-4}	0.11	3.44×10^{-1}	-	-	-	-
Lobar ICH	SPATA20	17	rs878619	−10.89	−3.186	102	3.51	4.44×10^{-4}	0.28	6.55×10^{-1}	3.40	6.79×10^{-4}	0.55 *	8.82×10^{-3}
Lobar ICH	ICA1L	2	rs7582720	7.21		81	−3.30	9.6×10^{-4}	0.2	1	−2.5692	1.02×10^{-2}	0.9	0.13

* Significant PP4 in the replication. pQTL ID: ID for the pQTL strongly associated with proteins; pQTL z-score: z-score for the pQTL strongly associated with proteins; pQTL z-score: z-score in the correspondent GWAS for the best pQTL; NSNP: number of SNPs in the locus; TWAS z-score: z-score for the PWAS association; PWAS p-value: p-value for the PWAS association; PP4: posterior probability 4 in the COLOC analysis; q-value: FDR-adjusted PWAS p-value.

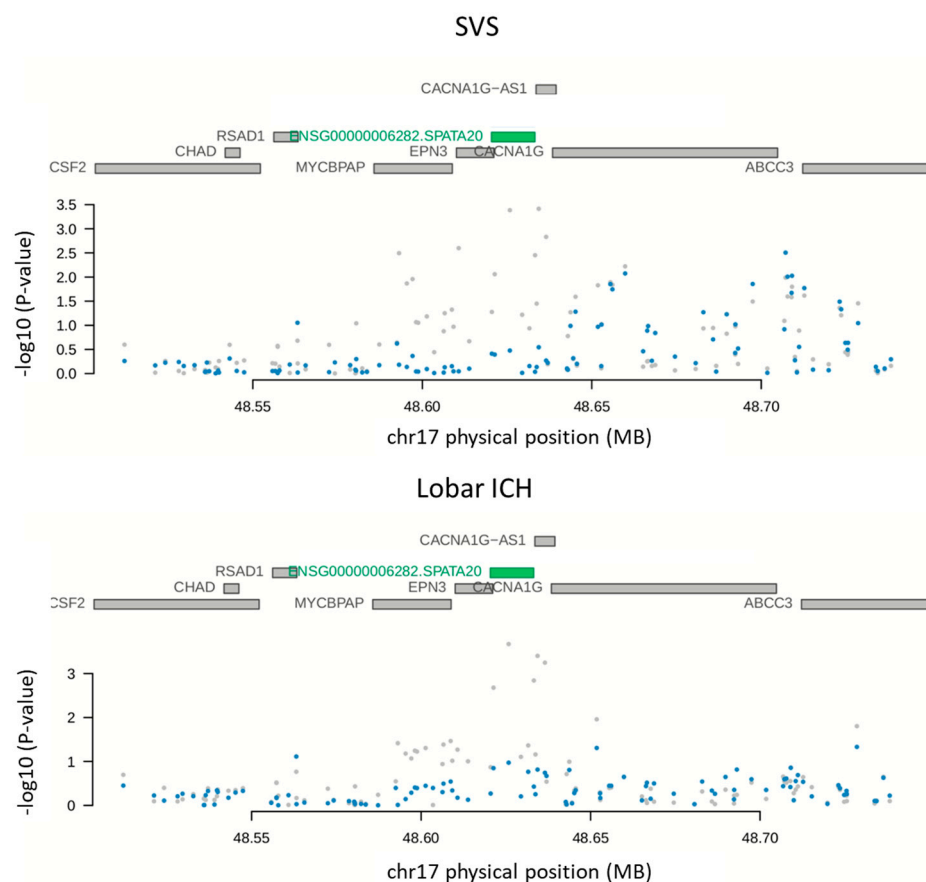


Figure 3. GWAS results before conditioning by the SPATA20 levels (grey) or after conditioning by protein levels (blue).

3. Discussion

In the present study, we performed a PWAS for the two main acute manifestations of SVD: SVS and ICH. We identified an association at the proteomic level of ICA1L with SVS and non-lobar ICH. Moreover, ICA1L was also associated with lobar ICH in our study. However, it was not a statistically significant association after a multivariable correction.

Although different loci have previously been found to be associated with ICH and SVS, how these genes contribute to their risk is not fully understood. One approach used to prioritize the likely causal gene within an associated locus with multiple genes is the use of a transcriptome-wide association study (TWAS) [27]. The TWAS from the SVS GWAS results found that the expression of *SLC25A44* in the arteries and the 2q33.2 locus (*CARF*, *FAM117B*, *ICA1L*, and *NBEAL1*) in the arteries and brain was causally associated with an SVS risk [23]. Very similar results regarding the 2q33.2 locus were previously obtained by Persyn et al. and Sargurupremraj et al. in their TWAS of WMH [16,17]. In contrast, the *ULK4* expression was negatively associated with an SVS risk in the arterial tissues, whole blood, and in the brain [23].

Our study implemented a similar strategy, focused on proteins, that allowed for a better characterization of a previously known association of ICA1L with non-lobar ICH, SVS [4,23], and related SVD traits, such as WMH [16,17] and fractional anisotropy (FA) [16]. Based on our results, genetics are responsible for the cis-regulation of an ICA1L abundance in brain tissue. Specifically, the top pQTL for ICA1L determines a higher abundance in ICA1L, which, in turn, reduces the risk for non-lobar ICH and SVS.

Although no previous PWAS has investigated the association of a protein abundance with SVS nor ICH, an ICA1L abundance was previously found to be associated with Alzheimer's disease (AD) risk using a PWAS approach [25,28]. These studies used the same proteomic data as in our work and found that a higher expression of the ICA1L protein

was linked with a decrease in AD risk [25,28]. Thus, both results for AD and SVD showed that elevated ICA1L levels are a protective factor. These AD studies, using single-cell RNA-sequencing, also identified enrichment in the ICA1L expression in glutamatergic excitatory neurons [25,28]. Excitatory processes involving glutamate are related to the cell death features in stroke and ICH [29,30]. Furthermore, there is a clear relationship between AD and SVD. Both have common features that lead to dementia, such as a blood–brain barrier breakdown [31]. Moreover, cerebral amyloid angiopathy and arteriolosclerosis, common SVD mechanisms, worsen the threshold for AD dementia. Different SVD features have been found in post-mortem brains from AD patients, such as small cortical and subcortical infarcts, microbleeds, perivascular spacing, and WMH [32]. Additionally, *ApoE* is the major genetic factor associated both with lobar ICH and AD [33]. The *Ica1l* results suggest that this protein could also be commonly regulated in AD and SVD.

SPATA20 was also identified as a potential protein affecting SVS and lobar ICH. Previous genetic studies using a GWAS found polymorphisms in this gene associated with platelet distribution width [34], reticulocyte count [34], and subcortical volume [35]. However, pleiotropy was not found in the SMR analysis, indicating that different genetic regulations could be mediating the protein abundance and the trait risk.

The main limitation of our study is the number of proteins that we were able to evaluate. In this sense, only ICA1L was present in the proteomic panel from all proteins from loci previously associated with SVS or ICH. Moreover, we only had access to proteomic data from dPFC, and other brain regions could be relevant in SVD.

The results from this PWAS, combined with the previous results at the genomic and transcriptomic levels, highlight the relevance of ICA1L in the risk of SVD and specific SVD outcomes. As all these outcomes (SVS, non-lobar ICH, and lobar ICH) are common manifestations of SVD, the ICA1L regulation could be a common mechanism in SVD. Given that this regulation is also produced at the protein level and in brain tissue, future studies could focus on the investigation of ICA1L as a potential drug target.

4. Materials and Methods

4.1. Proteome-Wide Associations Study (PWAS)

We used FUSION software (Boston, MA, USA) [36] to estimate the association of brain protein levels with SVS and ICH. FUSION is implemented to identify associations between a functional phenotype (i.e., protein expression levels) and a GWAS phenotype. FUSION takes two inputs: 1) Precomputed functional weights, and 2) GWAS summary statistics unified to a reference single-nucleotide polymorphism (SNP) panel. In the PWAS, the functional weights consist of estimations of the strength of association between proteomic expression levels and the SNPs from the cis-locus of the protein (only considered in the heritable proteins). The SNPs associated with cis-locus protein expression levels are protein quantitative-traits (or pQTLs).

In our PWAS analysis, we selected, as functional weights, the freely available pre-computed weights for the proteome expression analysis published by Wingo et al. [28]. As mentioned, in this case, the functional weights consisted of a summary-based file with data for the association between protein levels and SNPs (for the same locus of the protein). In the study from Wingo et al., from which we selected the functional weights, post-mortem dorsolateral prefrontal cortex (dPFC) human brain proteomes from 400 participants of European ancestry from the ROS/MAP cohorts were generated [37]. Proteomic sequencing was performed using isobaric tandem mass tag (TMT) peptide labelling combined with liquid chromatography coupled to mass spectrometry. They carried out different quality controls (QCs), which are described in the original article [28]. First, they randomized the 400 samples and 100 global internal standards into 50 batches before TMT labeling with the TMT10plex kit (Thermo Fisher Scientific, Waltham, MA, USA). Each batch contained two exact global internal standards. After running mass spectrometry, they assigned a peptide spectral using the Proteome Discoverer suite v.2.3 (Thermo Fisher Scientific) and a search in the canonical UniProtKB human proteome database. Then, the peptide spectral

matches were collated into proteins. They did not include, in the analysis, the peptides whose protein quantification was outside the 95% confidence interval of the global internal standard, nor proteins with missing values in more than 50% of the samples. After the QCs, 8356 proteins remained for analysis. The abundance of these proteins was log2 transformed. Then, QC for the samples was also performed, removing samples outstanding more than four standard deviations from the first two principal components. The effect of the batch, sex, age of death, study, mass spectrometry reporter quantification mode, postmortem interval, and clinical diagnosis was detected and removed [28].

They coupled this proteomic data with genetic data for weight calculation purposes, and the genotype data from 1,190,321 HapMap SNPs present in 489 individuals from the same cohort was processed. The genotypes were obtained from whole-genome sequencing or genome-wide genotyping. Classical QCs were applied to the genotype data: removal of variants deviating from the HardyWeinberg equilibrium, with a missing genotype rate of >5%, and rare variants with a minor allele frequency of <1%. Individuals with >5% genotyping missingness and related individuals were also excluded from the analysis. For genome-wide genotyped data, imputation was performed, and only HapMap SNPs selected from the 1000 Genomes Project were included. After considering all the QCs for proteomic and genotyping data, the functional weight included dPFC proteome abundance from heritable proteins (a total of 1475 heritable proteins) in 376 subjects, combined with 1,190,321 SNPs [28].

For the PWAS, we selected the GWAS summary statistics from the traits of interest considering the largest sample size published. For SVS, we selected summary statistics for the European ancestry MEGASTROKE SVS GWAS (N = 5386 SVS patients and 406,111 controls) [22]. For ICH, we selected the three different ICH subtypes (all ICH, ICH lobar, and ICH non-lobar) GWAS summary statistics available from the MTAG published by Chung et al. (N = 6255 ICH cases and 233,058 control subjects) [4].

Using FUSION, we first unified the GWAS SNPs (total SNPs SVS = 8,177,651; total SNPs all ICH = 7,136,230; total SNPs non-lobar ICH = 7,136,333; and total SNPs lobar ICH = 7,495,490) with the reference panel from the 1000 Genomes Project (Linkage disequilibrium (LD) reference panel). Secondly, SNPs from the GWAS not present in the LD reference panel were imputed. Finally, the association between the functional weights (1475 proteins) and the GWAS summary statistics was estimated. We corrected the results, considering multiple testing by 1475 proteins with a False Discovery Rate (FDR) multivariable test correction (q -values).

4.2. Colocalization Analysis

We performed a Bayesian colocalization analysis for the PWAS results with a p -value of <0.05. The objective of this analysis was to assess whether the association of protein levels with the GWAS phenotype was caused by the same variant (or pQTL). We used COLOC, currently implemented in FUSION, for this purpose. Five different COLOC hypotheses were tested: no association (posterior probability (PP0), functional association only (PP1), GWAS association only (PP2), independent functional/GWAS associations (PP3), or colocalized functional/GWAS associations (PP4)). We considered plausible causal relationships for those with a posterior PP4 > 0.5.

We considered two criteria to assign significant terms to the PWAS results from the discovery analysis. First, using a stringent criterion, we considered significant PWAS results, those proteins with a significant q -value (<0.05) in the discovery PWAS, and a PP4 > 0.5 in the colocalization analysis [28]. Proteins following this criterion were considered in the replication analysis. Second, using an open criterion, we selected for the replication of all the proteins with a nominal p -value in the discovery PWAS (p -value < 10^{-5}) (Figure 4).

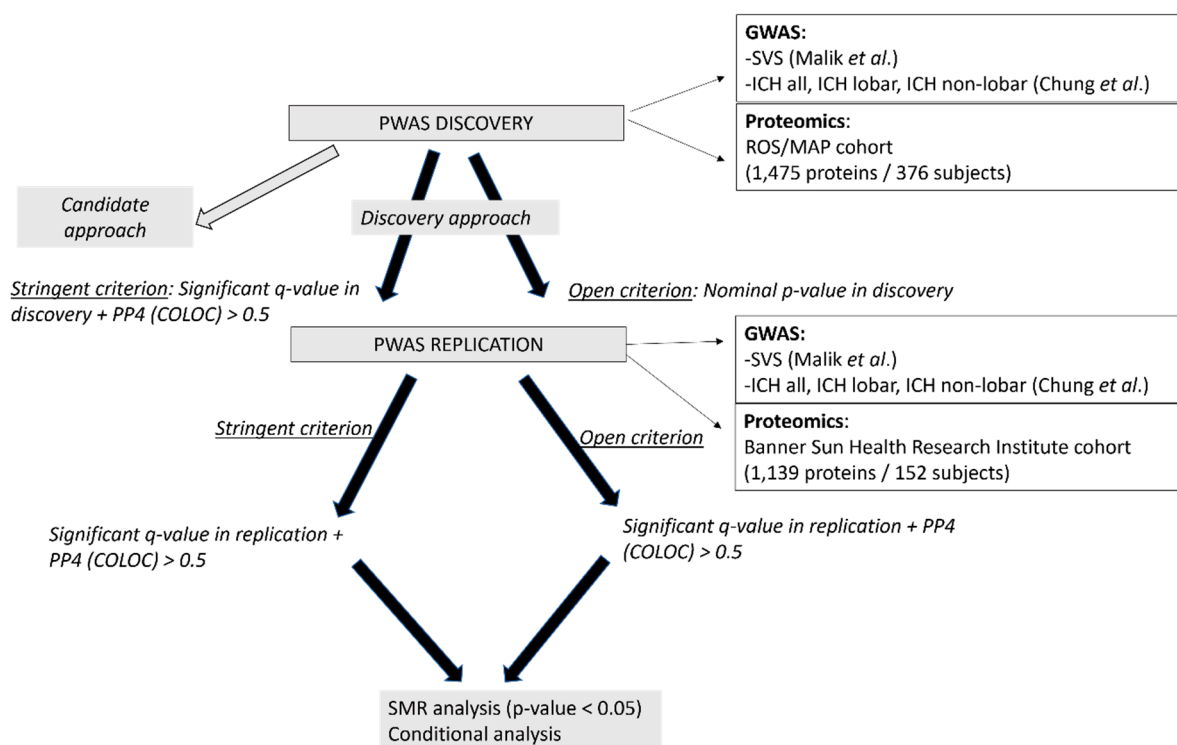


Figure 4. Schematic study workflow considering the discovery and the replication PWAS but also the sensitivity analyses.

4.3. Replication Analysis

For replication purposes, we selected a confirmation dPFC brain proteome panel obtained in 198 participants of European ancestry from the Banner Sun Health Research Institute [38]. Proteomic profiles were obtained using the same methodology used in the discovery panel [28]. After applying the QCs previously described, a total of 152 patients with proteomic and genetic data were used in the functional weight construction. Finally, proteomic weights included 1139 heritable proteins [28] that were combined with the above-mentioned GWAS summary statistics using the same protocol we used for the discovery panel.

All the proteins evaluated were considered significant in the replication study when the q -value was <0.05 (adjusting for all the proteins analysed in the replication) and the PP4 was >0.5 (Figure 1).

4.4. Summary-Based Mendelian Randomization (SMR)

We applied SMR to test whether the cis-regulated protein abundance was pleiotropically regulating the relationship between the SNPs and the trait of interest [39]. To perform the analysis for proteins achieving the stringent PWAS criterion, we followed the predetermined criteria of at least one pQTL at a p -value $< 5 \times 10^{-8}$ and included only SNPs with pQTL p -values $< 1.5654 \times 10^{-3}$. SNPs with LD r^2 between the top SNP > 0.9 or < 0.05 were excluded [39]. For the results from the open criterion, we included SNPs with pQTL p -values < 0.05 .

We considered that protein levels were mediating the effect of an SNP on the phenotype when they had a p -value < 0.05 in SMR.

4.5. Conditional Analysis

For the significant PWAS proteins, we identified SNPs from the GWAS (p -value < 0.05) in the regions cis-overlapped with the selected proteins. In order to assess whether the significant associations between these polymorphisms and the trait were mediated by the

cis-regulated protein abundance, we applied conditional tests using FUSION [27]. We plotted the results to observe that mediation visually, and we calculated conditional statistics.

Author Contributions: I.F.-C. and J.K. were responsible for the study design and the analytical accuracy of the manuscript. N.C. performed the statistical analyses and prepared the manuscript. C.G.-F., J.C.-M., L.L.-C., M.L., E.M., J.M.M.-C., J.M., L.C. and M.A. reviewed the manuscript and contributed to the data interpretation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The EPIGENESIS project (Carlos III Institute-PI17/02089, Marató TV3 and Fundació MútuaTerrassa), the MAESTRO project- PI18/01338 (Carlos III Institute), the iBioStroke project (Eranet-Neuron, European research grants), the EPINEXO project-PI20/00678 (Carlos III Institute/Fondo Europeo de Desarrollo Regional (FEDER), the SEDMAN Study (Boehringer Ingelheim), the APHAS project (Pfizer/BMS), and the European Regional Development Fund (ERDF)/Fondo FEDER and 2017SGR-1427 (AGAUR). I. Fernández is the recipient of a research contract from the Miguel Servet Program (CP12/03298) from the Carlos III Institute. J. Cárcel-Márquez is supported by an AGAUR Contract (Agència de Gestió d'Ajuts Universitaris i de Recerca; FI_DGR 2019, grant number 2019_FI_B 00853), co-financed by Fons Social Europeu (FSE). M. Lledós is supported by a PFIS Contract (Contratos Predoctorales de Formación en Investigación en Salud): FI19/00309. C. Gallego-Fabrega is supported by a Sara Borrell Contract (CD20/00043) from the Carlos III Institute and Fondo Europeo de Desarrollo Regional (ISCIII-FEDER). E. Muiño is supported by a Río Hortega Contract (CM18/00198) from the Carlos III Institute.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki guidelines. Each Institutional Review Board from previously published studies was responsible for approving the various studies included in this manuscript.

Informed Consent Statement: The informed consent of the patients was obtained for all the studies included in this article.

Data Availability Statement: The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: We would like to thank the International Stroke Genetics Consortium, the Spanish Stroke Genetics Consortium, the International Stroke Genetics Consortium, and the Global Alliance for Acute and Long-term Outcome Studies in stroke and the RETICS-INVICTUS Network (RD16/0019/0011).

Conflicts of Interest: The authors declare that they have no competing interests.

References

1. Pantoni, L. Cerebral Small Vessel Disease: From Pathogenesis and Clinical Characteristics to Therapeutic Challenges. *Lancet Neurol.* **2010**, *9*, 689–701. [\[CrossRef\]](#)
2. Uniken Venema, S.M.; Marini, S.; Lena, U.K.; Morotti, A.; Jessel, M.; Moomaw, C.J.; Kourkoulis, C.; Testai, F.D.; Kittner, S.J.; Brouwers, H.B.; et al. Impact of Cerebral Small Vessel Disease on Functional Recovery After Intracerebral Hemorrhage. *Stroke* **2019**, *50*, 2722–2728. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Ter Telgte, A.; van Leijsen, E.M.C.; Wiegertjes, K.; Klijn, C.J.M.; Tuladhar, A.M.; de Leeuw, F.-E. Cerebral Small Vessel Disease: From a Focal to a Global Perspective. *Nat. Rev. Neurol.* **2018**, *14*, 387–398. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Chung, J.; Marini, S.; Pera, J.; Norrving, B.; Jimenez-Conde, J.; Roquer, J.; Fernandez-Cadenas, I.; Tirschwell, D.L.; Selim, M.; Brown, D.L.; et al. Genome-Wide Association Study of Cerebral Small Vessel Disease Reveals Established and Novel Loci. *Brain* **2019**, *142*, 3176–3189. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Woo, D.; Falcone, G.J.; Devan, W.J.; Brown, W.M.; Biffi, A.; Howard, T.D.; Anderson, C.D.; Brouwers, H.B.; Valant, V.; Battey, T.W.K.; et al. Meta-Analysis of Genome-Wide Association Studies Identifies 1q22 as a Susceptibility Locus for Intracerebral Hemorrhage. *Am. J. Hum. Genet.* **2014**, *94*, 511–521. [\[CrossRef\]](#)
6. Traylor, M.; Malik, R.; Nalls, M.A.; Cotlarciuc, I.; Radmanesh, F.; Thorleifsson, G.; Hanscombe, K.B.; Langeveld, C.; Saleheen, D.; Rost, N.S.; et al. Genetic Variation at 16q24.2 Is Associated with Small Vessel Stroke. *Ann. Neurol.* **2017**, *81*, 383–394. [\[CrossRef\]](#)
7. Qureshi, A.I.; Mendelow, A.D.; Hanley, D.F. Intracerebral Haemorrhage. *Lancet* **2009**, *373*, 1632–1644. [\[CrossRef\]](#)
8. Sato, S.; Delcourt, C.; Heeley, E.; Arima, H.; Zhang, S.; Al-Shahi Salman, R.; Stapf, C.; Woo, D.; Flaherty, M.L.; Vagal, A.; et al. Significance of Cerebral Small-Vessel Disease in Acute Intracerebral Hemorrhage. *Stroke* **2016**, *47*, 701–707. [\[CrossRef\]](#)
9. Kim, B.J.; Lee, S.-H. Prognostic Impact of Cerebral Small Vessel Disease on Stroke Outcome. *J. Stroke* **2015**, *17*, 101–110. [\[CrossRef\]](#)

10. Regenhardt, R.W.; Das, A.S.; Lo, E.H.; Caplan, L.R. Advances in Lacunar Stroke Pathophysiology: A Review. *JAMA Neurol.* **2018**, *75*, 1273–1281. [[CrossRef](#)]
11. Fornage, M.; Debette, S.; Bis, J.C.; Schmidt, H.; Ikram, M.A.; Dufouil, C.; Sigurdsson, S.; Lumley, T.; DeStefano, A.L.; Fazekas, F.; et al. Genome-Wide Association Studies of Cerebral White Matter Lesion Burden: The CHARGE Consortium. *Ann. Neurol.* **2011**, *69*, 928–939. [[CrossRef](#)] [[PubMed](#)]
12. Haffner, C.; Malik, R.; Dichgans, M. Genetic Factors in Cerebral Small Vessel Disease and Their Impact on Stroke and Dementia. *J. Cereb. Blood Flow Metab.* **2016**, *36*, 158–171. [[CrossRef](#)] [[PubMed](#)]
13. Verhaaren, B.F.J.; Debette, S.; Bis, J.C.; Smith, J.A.; Ikram, M.K.; Adams, H.H.; Beecham, A.H.; Rajan, K.B.; Lopez, L.M.; Barral, S.; et al. Multiethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI. *Circ. Cardiovasc. Genet.* **2015**, *8*, 398–409. [[CrossRef](#)] [[PubMed](#)]
14. Traylor, M.; Zhang, C.R.; Adib-Samii, P.; Devan, W.J.; Parsons, O.E.; Lanfranconi, S.; Gregory, S.; Cloonan, L.; Falcone, G.J.; Radmanesh, F.; et al. Genome-Wide Meta-Analysis of Cerebral White Matter Hyperintensities in Patients with Stroke. *Neurology* **2016**, *86*, 146–153. [[CrossRef](#)] [[PubMed](#)]
15. Traylor, M.; Tozer, D.J.; Croall, I.D.; Lisiecka-Ford, D.M.; Olorunda, A.O.; Boncoraglio, G.; Dichgans, M.; Lemmens, R.; Rosand, J.; Rost, N.S.; et al. Genetic Variation in PLEKHG1 Is Associated with White Matter Hyperintensities (n = 11,226). *Neurology* **2019**, *92*, e749–e757. [[CrossRef](#)]
16. Persyn, E.; Hanscombe, K.B.; Howson, J.M.M.; Lewis, C.M.; Traylor, M.; Markus, H.S. Genome-Wide Association Study of MRI Markers of Cerebral Small Vessel Disease in 42,310 Participants. *Nat. Commun.* **2020**, *11*, 2175. [[CrossRef](#)]
17. Sargurupremraj, M.; Suzuki, H.; Jian, X.; Sarnowski, C.; Evans, T.E.; Bis, J.C.; Eiriksdottir, G.; Sakaue, S.; Terzikhan, N.; Habes, M.; et al. Cerebral Small Vessel Disease Genomics and Its Implications across the Lifespan. *Nat. Commun.* **2020**, *11*, 6285. [[CrossRef](#)]
18. Rutten-Jacobs, L.C.A.; Tozer, D.J.; Duering, M.; Malik, R.; Dichgans, M.; Markus, H.S.; Traylor, M. Genetic Study of White Matter Integrity in UK Biobank (N = 8448) and the Overlap with Stroke, Depression, and Dementia. *Stroke* **2018**, *49*, 1340–1347. [[CrossRef](#)]
19. Armstrong, N.J.; Mather, K.A.; Sargurupremraj, M.; Knol, M.J.; Malik, R.; Satizabal, C.L.; Yanek, L.R.; Wen, W.; Gudnason, V.G.; Dueker, N.D.; et al. Common Genetic Variation Indicates Separate Etiologies for Periventricular and Deep White Matter Hyperintensities. *Stroke* **2020**, *51*, 2111–2121. [[CrossRef](#)]
20. Knol, M.J.; Lu, D.; Traylor, M.; Adams, H.H.H.; Romero, J.R.J.; Smith, A.V.; Fornage, M.; Hofer, E.; Liu, J.; Hostettler, I.C.; et al. Association of Common Genetic Variants with Brain Microbleeds: A Genome-Wide Association Study. *Neurology* **2020**, *95*, e3331–e3343. [[CrossRef](#)]
21. Li, H.-Q.; Cai, W.-J.; Hou, X.-H.; Cui, M.; Tan, L.; Yu, J.-T.; Dong, Q. Alzheimer’s Disease Neuroimaging Initiative Genome-Wide Association Study of Cerebral Microbleeds on MRI. *Neurotox. Res.* **2020**, *37*, 146–155. [[CrossRef](#)] [[PubMed](#)]
22. Malik, R.; Chauhan, G.; Traylor, M.; Sargurupremraj, M.; Okada, Y.; Mishra, A.; Rutten-Jacobs, L.; Giese, A.-K.; van der Laan, S.W.; Gretarsdottir, S.; et al. Multiancestry Genome-Wide Association Study of 520,000 Subjects Identifies 32 Loci Associated with Stroke and Stroke Subtypes. *Nat. Genet.* **2018**, *50*, 524–537. [[CrossRef](#)] [[PubMed](#)]
23. Traylor, M.; Persyn, E.; Tomppo, L.; Klasson, S.; Abedi, V.; Bakker, M.K.; Torres, N.; Li, L.; Bell, S.; Rutten-Jacobs, L.; et al. Genetic Basis of Lacunar Stroke: A Pooled Analysis of Individual Patient Data and Genome-Wide Association Studies. *Lancet Neurol.* **2021**, *20*, 351–361. [[CrossRef](#)]
24. Brandes, N.; Linial, N.; Linial, M. PWAS: Proteome-Wide Association Study—Linking Genes and Phenotypes by Functional Variation in Proteins. *Genome Biol.* **2020**, *21*, 173. [[CrossRef](#)]
25. Ou, Y.-N.; Yang, Y.-X.; Deng, Y.-T.; Zhang, C.; Hu, H.; Wu, B.-S.; Liu, Y.; Wang, Y.-J.; Zhu, Y.; Suckling, J.; et al. Identification of Novel Drug Targets for Alzheimer’s Disease by Integrating Genetics and Proteomes from Brain and Blood. *Mol. Psychiatry* **2021**, *1*–9. [[CrossRef](#)]
26. Wingo, T.S.; Liu, Y.; Gerasimov, E.S.; Gockley, J.; Logsdon, B.A.; Duong, D.M.; Dammer, E.B.; Lori, A.; Kim, P.J.; Ressler, K.J.; et al. Brain Proteome-Wide Association Study Implicates Novel Proteins in Depression Pathogenesis. *Nat. Neurosci.* **2021**, *24*, 810–817. [[CrossRef](#)] [[PubMed](#)]
27. Malik, R.; Dichgans, M. Challenges and Opportunities in Stroke Genetics. *Cardiovasc. Res.* **2018**, *114*, 1226–1240. [[CrossRef](#)]
28. Wingo, A.P.; Liu, Y.; Gerasimov, E.S.; Gockley, J.; Logsdon, B.A.; Duong, D.M.; Dammer, E.B.; Robins, C.; Beach, T.G.; Reiman, E.M.; et al. Integrating Human Brain Proteomes with Genome-Wide Association Data Implicates New Proteins in Alzheimer’s Disease Pathogenesis. *Nat. Genet.* **2021**, *53*, 143–146. [[CrossRef](#)]
29. Liu, D.-Z.; Sharp, F.R. Excitatory and Mitogenic Signaling in Cell Death, Blood-Brain Barrier Breakdown, and BBB Repair after Intracerebral Hemorrhage. *Transl. Stroke Res.* **2012**, *3*, 62–69. [[CrossRef](#)]
30. Lai, T.W.; Zhang, S.; Wang, Y.T. Excitotoxicity and Stroke: Identifying Novel Targets for Neuroprotection. *Prog. Neurobiol.* **2014**, *115*, 157–188. [[CrossRef](#)]
31. Chagnot, A.; Barnes, S.R.; Montagne, A. Magnetic Resonance Imaging of Blood-Brain Barrier Permeability in Dementia. *Neuroscience* **2021**, *474*, 14–29. [[CrossRef](#)] [[PubMed](#)]
32. Kalaria, R.N.; Sepulveda-Falla, D. Cerebral Small Vessel Disease in Sporadic and Familial Alzheimer Disease. *Am. J. Pathol.* **2021**, *191*, 1888–1905. [[CrossRef](#)] [[PubMed](#)]
33. Bordes, C.; Sargurupremraj, M.; Mishra, A.; Debette, S. Genetics of Common Cerebral Small Vessel Disease. *Nat. Rev. Neurol.* **2022**, *1*–18. [[CrossRef](#)]

34. Vuckovic, D.; Bao, E.L.; Akbari, P.; Lareau, C.A.; Mousas, A.; Jiang, T.; Chen, M.-H.; Raffield, L.M.; Tardaguila, M.; Huffman, J.E.; et al. The Polygenic and Monogenic Basis of Blood Traits and Diseases. *Cell* **2020**, *182*, 1214–1231.e11. [[CrossRef](#)] [[PubMed](#)]
35. Van der Meer, D.; Frei, O.; Kaufmann, T.; Shadrin, A.A.; Devor, A.; Smeland, O.B.; Thompson, W.K.; Fan, C.C.; Holland, D.; Westlye, L.T.; et al. Understanding the Genetic Determinants of the Brain with MOSTest. *Nat. Commun.* **2020**, *11*, 3512. [[CrossRef](#)]
36. Gusev, A.; Ko, A.; Shi, H.; Bhatia, G.; Chung, W.; Penninx, B.W.J.H.; Jansen, R.; de Geus, E.J.C.; Boomsma, D.I.; Wright, F.A.; et al. Integrative Approaches for Large-Scale Transcriptome-Wide Association Studies. *Nat. Genet.* **2016**, *48*, 245–252. [[CrossRef](#)]
37. Bennett, D.A.; Buchman, A.S.; Boyle, P.A.; Barnes, L.L.; Wilson, R.S.; Schneider, J.A. Religious Orders Study and Rush Memory and Aging Project. *J. Alzheimer's Dis.* **2018**, *64*, S161–S189. [[CrossRef](#)]
38. Beach, T.G.; Adler, C.H.; Sue, L.I.; Serrano, G.; Shill, H.A.; Walker, D.G.; Lue, L.; Roher, A.E.; Dugger, B.N.; Maarouf, C.; et al. Arizona Study of Aging and Neurodegenerative Disorders and Brain and Body Donation Program. *Neuropathology* **2015**, *35*, 354–389. [[CrossRef](#)]
39. Zhu, Z.; Zhang, F.; Hu, H.; Bakshi, A.; Robinson, M.R.; Powell, J.E.; Montgomery, G.W.; Goddard, M.E.; Wray, N.R.; Visscher, P.M.; et al. Integration of Summary Data from GWAS and EQTL Studies Predicts Complex Trait Gene Targets. *Nat. Genet.* **2016**, *48*, 481–487. [[CrossRef](#)]

Artículo 4:

Título: Identification of genetic regulation at the proteomic level in the long-term stroke outcome: A Proteome-Wide Association study.

Revista: No publicado aún.

Objetivo: Identificar proteínas reguladas genéticamente asociadas al pronóstico funcional a largo plazo (mRS a los 3 meses) mediante un estudio de integración de datos genómicos y proteómicos utilizando la estrategia de PWAS.

Resumen:

Antecedentes y objetivos:

Antecedentes: La evolución del ictus es muy variable y depende de factores demográficos, clínicos y genéticos. El estudio de la modulación de la variabilidad funcional después de un ictus es fundamental para identificar nuevas terapias. Un nuevo enfoque es el estudio de cómo las variantes genéticas afectan al pronóstico funcional del ictus a través de la modulación a nivel proteómico.

Objetivos: Identificar nuevos biomarcadores proteicos para la evolución a largo plazo (escala de Rankin modificada a los tres meses [mRS3]) combinando datos ómicos procedentes de la genómica y la proteómica.

Métodos: Se realizó un estudio de asociación de *Proteome-Wide Association study* (PWAS) utilizando FUSION para integrar los datos proteómicos del córtex prefrontal dorsolateral (dPFC) de la cohorte ROS/MAP (1,475 proteínas heredables en 376 sujetos) y el GWAS de la evolución funcional post-ictus de la cohorte GODS (N = 1,791 y 8,895,027 variantes de nucleótido único (SNV)). Aplicamos un análisis bayesiano de colocación con COLOC y randomización mendeliana (SMR) para validar la causalidad de los resultados del PWAS. Replicamos los resultados significativos en un nuevo conjunto de datos proteómicos del dPFC del *Banner Sun Health Research Institute* (1.139 proteínas heredables en 152 sujetos) (Replicación 1). En una replicación 2, combinamos el conjunto de datos de ROS/MAP con un nuevo GWAS de mRS3 (N = 688 y 7.502.471 SNV).

Metaanalizamos los resultados del descubrimiento y ambas cohortes de replicación utilizando el método AW-Fisher y estudiamos si algunas vías estaban enriquecidas en el metaanálisis. También realizamos un estudio de *Transcriptome-Wide association study* (TWAS) utilizando datos transcriptómicos del Consorcio CommonMind (5.420 genes en 452 muestras humanas de la región dPFC).

Resultados: Tres proteínas se asociaron significativamente con la evolución del ictus ($P < 0,05$) en el PWAS de descubrimiento con una probabilidad posterior 4 (PP4) en el análisis de COLOC $> 0,5$. Las tres proteínas fueron significativas en la Replicación1 pero sólo THEM4 se asoció significativamente con mRS3 en la Replicación2. Dos de las tres proteínas (THEM4 y APOL2) fueron significativas tras el ajuste de pruebas múltiples en el metaanálisis. Los resultados del SMR confirmaron la asociación causal de los niveles de THEM4 y APOL2 en la región dPFC con el pronóstico funcional del ictus. Las vías de señalización relacionadas con la regulación de los ácidos grasos y de AMPK se encontraron enriquecidas en los resultados del metaanálisis.

Conclusiones: Nuestros resultados indican que las variantes genéticas asociadas con el pronóstico del ictus a través de la regulación cis a nivel proteómico están modulando vías clave en el ictus isquémico.

Identification of genetic regulation at the proteomic level in the long-term stroke outcome: A Proteome-Wide Association study.

Natalia Cullell, MSc^{*1,2,3}, Cristina Gallego-Fábrega, PhD¹, Jara Cárcel, MSc¹, Elena Muiño, MD, PhD¹, Laia Llucià-Carol^{1,8}, Miquel Lledós, MSc¹, Jesús M Martín-Campos¹, Israel Fernández-Cadenas, PhD^{1,2**}, Jerzy Krupinski, MD, PhD^{2 **}

1- Stroke Pharmacogenomics and Genetics group, Biomedical Research Institute Sant Pau (IIB Sant Pau). 08041. Barcelona

2- Neurology. Hospital Universitari MútuaTerrassa/ Fundacio Docència i Recerca MutuaTerrassa, 08221. Terrassa, Spain.

3- Facultat de Medicina. Universitat de Barcelona, 08036. Barcelona, Spain

****Correspondence:**

Israel Fernández Cadenas: israelcadenas@yahoo.es

Jerzy Krupinski: jkrupinski@mutuaterrassa.es

Abstract

Background: The stroke outcome is highly variable and depends on demographic, clinical and genetic factors. To study how the stroke outcome is modulated is critical to identify new therapies. A new approach would be to study how genetic variants are affecting the stroke outcome through the modulation at the proteomic level.

Objectives: To identify new protein biomarkers for the long-term outcome (modified Rankin scale at three months (mRS3)) combining -omic data coming from genomics and proteomics.

Methods: We performed a discovery proteome-wide association study (PWAS) using FUSION to integrate the dorsolateral prefrontal cortex (dPFC) proteomic data from the ROS/MAP cohort (1,475 heritable proteins in 376 subjects) and the stroke outcome GWAS from the GODS cohort (N = 1,791 and 8,895,027 single nucleotide variants (SNV)). We applied a Bayesian colocalization analysis with COLOC and a Summary-based-data Mendelian Randomization (SMR) approach to validate the causality from the PWAS results. We replicated the significant results in a new dPFC proteomic dataset from the Banner Sun Health Research Institute (1,139 heritable proteins in 152 subjects) (Replication 1). In a replication 2, we combined the ROS/MAP dataset with a new mRS3 GWAS (N = 688 and 7,502,471 SNV). We meta-analyzed the results from the discovery and both replication cohorts using the adaptively weighted Fisher's method (AW-Fisher) and we studied whether some pathways were enriched in the meta-analysis. We also performed a transcriptome wide association study (TWAS) using transcriptomic data from the CommonMind Consortium (5,420 genes in 452 human dPFC samples).

Results: Three proteins were significantly associated with the stroke outcome ($P < 0.05$) in the discovery PWAS with a posterior probability 4 (PP4) in the COLOC > 0.5 . The three proteins were significant in the Replication1 but only THEM4 was significantly associated with mRS3 in the Replication2. Two out of the three proteins (THEM4 and APOL2) were significant after multiple testing adjustment in the meta-analysis. The SMR results confirmed the causal association of THEM4 and APOL2 levels in the dPFC with the stroke outcome. Fatty acids regulation and AMPK signaling pathways were enriched from the meta-analysis results.

Conclusions: Our results indicated that genetic variants associated with the stroke outcome through the cis-regulation at the proteomic level are modulating key pathways in stroke.

Introduction

Stroke is a leading cause of death and disability worldwide (1). Genetically, it is classified as a complex disease for which multiple genes have been recognized to be involved in its risk (2). Getting insights into the genetics of stroke outcome is also relevant to develop new strategies that could modify the outcome depending on the patient's genetic background. Previous studies demonstrated that stroke outcome is heritable. The 8.7% of the variability in the National Institutes of Health Stroke Scale (NIHSS) at 24h is explained by genetics (3).

Genome-wide Association Studies (GWAS) have revealed distinct genetic architecture involved in stroke outcome compared with stroke risk (3–5). Different stroke outcome measures have been used in these GWAS. Two studies focused on the long-term outcome, measured with the 3-months modified Rankin Scale (mRS3) (4,5) while the latest study published in this field (3) analysed the early neurological instability as a measurable variable of outcome. They calculated the difference in the National Institutes of Health Stroke Scale (NIHSS) between the baseline (<6h) and 24h (Δ NIHSS) and demonstrated that this variable is a good predictor of the long-term outcome (6). Using the mRS3, two different genes have been found. First, Mola-Caminal, identified variants in *PATJ* associated with worse functional outcome after stroke (4). Then, Södelholm et al., found a variant in *LOC105372028*, a trans-expression quantitative trait locus for *PPP1R21* (5). Using the Δ NIHSS variable, Ibanez et al., found eight genes from the excitotoxicity pathway that explained up to 1.8% of the variability in the stroke outcome (3).

Although GWAS have been successful in the identification of genetic risk factors for diseases, the causal genes affected by these variants are difficult to identify. Recently, different studies have used a new strategy based on the integration of genomic findings from GWAS with tissue-specific transcriptomic (transcriptome-wide association studies; TWAS) or proteomic data (proteome-wide association studies; PWAS). TWAS and PWAS allow the identification of genes associated with a disease through its regulation at the transcriptomic or proteomic level. Comparing with GWAS, both TWAS and PWAS enable a more functional approach to understand a disease (7). If Mendelian Randomization and Colocalization analyses are coupled to TWAS/PWAS, it allows the determination of causality to the findings (8).

Few studies have applied this approach in stroke, and they have focused only on stroke risk. The cis-regulation at the transcriptomic level of 19 genes in blood and adipose tissue were identified by Yang et al. (9) in association with stroke risk. Traylor et al. (10), found six genes using TWAS whose expression was regulating small vessel stroke (SVS) risk. Using PWAS, we published in a previous study (11), the association of *ICA1L* at the proteomic level with the risk of SVS and non-lobar intracerebral haemorrhage, two conditions sharing pathophysiological mechanisms. This result was further validated by two additional studies (12,13). Apart from the *ICA1L* finding, the PWAS study published by Zhang et al., also found *CAND2*, *ALDH2* and *MADD* associated with SVS risk (13).

As far as we know, there is no publication using PWAS nor TWAS in stroke outcome. The objective of our study is to find causal functional associations of genes regulating stroke outcome mediated by the cis-regulation at the proteomic level. Given that protein levels are dynamic and druggable, we first focused on the PWAS results. We then aimed to validate the causal effect of proteins identified by PWAS using Summary-based-data Mendelian Randomization (SMR) and Colocalization analysis. To identify whether significant results from PWAS had also an effect at the transcriptomic level, we performed a TWAS analysis.

Methods

Study design

We performed a three step PWAS which included a discovery analysis and two independent replications. In the discovery analysis, we combined one proteomic dataset obtained from the dorsolateral prefrontal cortex (dPFC) region of human brains with a long-term stroke outcome (mRS3) GWAS. The replication1 consisted of an independent proteomic dataset from the same brain region combined with the same GWAS than in the discovery. The replication2 consisted of the same proteomic dataset than in the discovery but combined with a different stroke outcome GWAS, analysing also de mRS3 as variable. Then, the results from the discovery and both replications were meta-analyzed. The results from the meta-analysis were further explored with the analysis of pathway enrichment.

The causality and validity of the results was assessed with colocalization and summary-based-data Mendelian Randomization analyses.

Finally, we also performed TWAS focusing on the transcriptomic data from the same validated proteins from the PWAS.

Proteomic dataset selection

We selected two different proteomic datasets, the first from the ROS/MAP cohort (14), used in the discovery stage and the second, for replication, from the Banner Sun Health Research Institute (15) (Replication1). Both contained proteomic data from the dPFC region from post-mortem human brains. For both cohorts, the proteomic data was generated using isobaric tandem mass tag (TMT) peptide labelling and liquid chromatography combined with mass spectrometry. Specific quality controls were applied to proteins and samples (described in the original article (16))

For PWAS analysis purpose, the proteomic data was converted into functional weights precomputed by Wingo et al. (16). Briefly, functional weights are computed for heritable proteins and consist of summary statistics of the association strength between protein quantitative-traits (pQTLs) and the expression of the corresponding cis-protein.

For the computation of the functional weights from the ROS/MAP cohort, the proteomic data from 400 participants was coupled with genomic data from the same cohort obtained from whole-genome sequencing or genome-wide genotyping. Genomic data consisted in 1,190,321 HapMap SNPs, obtained after proceeding with classical quality controls, previously described in the manuscript published by Wingo et al (16). Functional weights from the ROS/MAP cohort consisted in proteome abundance for 1,475 heritable proteins and 1,190,321 single-nucleotide polymorphisms in 376 subjects. After the same quality controls than for the ROS/MAP cohort, a total of 152 subjects had genomic and proteomic data in the Banner cohort (included in the Replication1). In this case, the functional weights consisted of 1,147 proteins.

GWAS selection

We selected the long-term outcome GWAS published by Mola-Caminal et al. (GODS GWAS) to be analysed in the discovery PWAS. Specifically, we studied the summary-statistics for the stringent analysis (N subjects = 1,791 and 8,895,027 SNV) because it was considered to have the most powerful results in the original manuscript.

We also selected the results for a GWAS (not previously published) that we performed in 688 stroke patients that accomplished the GODS GWAS selection criteria. This GWAS was considered as part of the replication (Replication2). Briefly, patients with > 18 years-old, from European descent, with a confirmed stroke by neuroimaging, an initial NIHSS > 4, previous functional independence (mRS at baseline < 3) and information of the mRS at 3 months were considered in the analysis. We applied the standard quality controls in GWAS for SNV and subjects. For SNV: we removed those with a call rate < 0.05, non-biallelic, ambiguous and duplicated SNV, and those that were not in Hardy-Weinberg equilibrium (HWE) with a threshold of $<1 \times 10^{-10}$. For samples, we removed all patients with > 3% of missingness, with differences in phenotypic and genotyping sex, duplicated samples and those with a π -hat > 0.2 in the identity by state (IBD). Genotyped samples were imputed using the Michigan Imputation Server Pipeline (<https://imputationserver.sph.umich.edu/index.html>) with the HRC r1.1 2016 (GRCh37/hg19) panel (European population). The association analysis was performed using fastGWA (17) in 7,502,471 SNV and considering as covariates the NIHSS at discharge, the stroke subtype using the The Trial of Org 101072 in Acute Stroke Treatment (TOAST) classification, age, sex and the first two principal components.

Transcriptomic data selection

We selected the transcriptomic data available with the functional weight format for 5,420 genes in 452 human dPFC brain samples from the CommonMind Consortium (CMC) Single tissue gene expression assay obtained from RNA-seq analysis, available in the FUSION website (<http://gusevlab.org/projects/fusion/>).

PWAS and TWAS analysis

We performed the proteome or transcriptome-wide association analysis with FUSION. This software, in a first step, unify the GWAS and reference SNV from the 1000 Genomes project. Then, all the missed SNV from the reference panel in the GWAS are imputed using the IMPG algorithm. Finally, the association between proteomics/transcriptomics with GWAS is estimated.

We considered a significant association in the PWAS or TWAS when the p-value was < 0.05.

Colocalization

To assess whether the same SNV was associated with protein / transcriptomic levels and with the stroke outcome, we performed Bayesian colocalization with COLOC, implemented in FUSION. COLOC follows a Bayesian approach to establish the posterior probability 4 (PP4) for the same SNV to be pQTL and associated with the trait.

Summary-based-data Mendelian Randomization (SMR)

To establish causality between the SNV and the trait caused by the proteomic or transcriptomic levels we tested pleiotropy using SMR. SNV in linkage disequilibrium (LD) ($r^2 > 0.9$) with the top SNP > 0.9 were excluded (18). We considered causal associations those with a p-value < 0.05 in the SMR analysis.

Meta-analysis

To meta-analyze the results from the Discovery and the Replication1 and 2, we applied the adaptively weighted Fisher's method (AW-Fisher)(19), specifically developed for meta-analysis

in -omic studies, as recently has been used by Zhenyao et al.(20) in TWAS. We performed the analysis in R using the AWFisher package (version 1.10.0) (21).

Pathway enrichment analysis

We selected the significant proteins after Bonferroni correction from the meta-analysis to study for some pathway enrichment. We used the web-based integrated data mining system WebGestalt (22) with the KEGG database for this purpose.

Conditional analysis

To assess whether the results from the GWAS with a p-value < 0.05 in the cis regions from the significant and validated PWAS results were conditioned by the proteomic levels, we performed conditional tests with FUSION. Regional Manhattan plots showed the effects of the conditioning by proteomic levels.

Multi tissue Mendelian Randomization

We performed a two-sample Mendelian Randomization (MR) with the R package TwoSampleMR version 3.6.2. (23) using as exposure the GTEX data about THEM4 eQTLs for 20 different tissues (<https://gtexportal.org/home/>). We analysed as outcome the GODS GWAS. We considered eQTLs with a p-value < 10^{-8} , we harmonize the data and clumped the SNV with a linkage disequilibrium (LD) $R^2 < 0.01$. We tested the causal association between the exposure and the outcome with the Wald test because only one SNV remained after filtering out the exposures. We considered significant causal association those with a p-value < 0.05 in the test.

RESULTS

PWAS discovery

The PWAS combining the proteomic data from the ROS/MAP cohort with the GODS GWAS resulted in a total of 85 proteins significantly associated (p-value < 0.05) with the long-term outcome after stroke (Table S1). The results from the Bayesian colocalization analysis showed that for three proteins (THEM4, GSTP1 and APOL2), their pQTL had a probability (PP4) > 50% to be the causal variant for both the protein levels and the stroke outcome (Table 1). Higher THEM4 levels were associated with a higher mRS score at three months (PWAS Z-score = 3.04), while higher GSTP1 and APOL2 levels were linked with lower mRS score, thus with a better stroke outcome (PWAS Z-score = -2.56 and -2.37 for GSTP1 and APOL2, respectively).

Table 1: Significant PWAS results in the Discovery, Replication 1, Replication 2 and meta-analysis

ID	DISCOVERY					REPLICATION1				REPLICATION2				META-ANALYSIS	
	pQTL	pQTL Z	PWAS Z	PWAS P	COLOC PP4	pQTL	pQTL Z	PWAS Z	PWAS P	pQTL	pQTL Z	PWAS Z	PWAS P	P	Q-value
THEM4	rs16833668	-10.7	3.04	2.35E-03	0.5	rs16833668	-8.69	2.84	4.51E-03	rs16833668	-10.8	2.3	2.35E-02	1.51E-04	6.93E-03
GSTP1	rs7941648	-8.99	-2.56	1.04E-02	0.7	rs1695	-9.13	-2.06	3.93E-02	rs7941648	-9	-0.06	9.55E-01	1.40E-02	6.42E-01
APOL2	rs8136528	-6.19	-2.37	1.77E-02	0.8	rs9619597	-6.5	-3.31	9.32E-04	rs8136528	-6.2	0.6	5.28E-01	8.66E-04	3.98E-02

Significant results (PWAS p-value < 0.05) from the Discovery PWAS with a colocalization PP4 > 0.5.

PP4: posterior probability 4; Q-value: P-value with Bonferroni correction

PWAS replication

49 out of the 85 significant proteins from the discovery were available in the proteomic panel from Banner cohort (replication1). From these, a total of 21 proteins were significantly associated with mRS3 in the replication1 (Table S2). THEM4, GSTP1 and APOL2, the three proteins significantly associated with the stroke outcome in the discovery with a $PP4 > 0.5$, were validated in the replication1 (p -value = 4.51×10^{-3} , 3.93×10^{-2} and 9.32×10^{-4} , respectively) with the same effect regardless the stroke outcome. Only THEM4 shared the same pQTL in the discovery and the replication1 (Table 1).

In the replication2, using a second GWAS for stroke outcome, three out of the 85 significant proteins from the discovery were significant (Table S3). From THEM4, GSTP1 and APOL2 proteins, only THEM4 was associated with the mRS3 (p -value = 2.35×10^{-2}) (Table 1), being also rs16833668 the most significant pQTL. In the replication2, the results also indicated that higher THEM4 levels were associated with worse outcome.

The meta-analysis of the results from the discovery, replication1 and replication2 showed 30 proteins with association with the stroke outcome (p -value < 0.05) (Table S4). Five of these proteins reached a significant Bonferroni corrected p -value (Q -value < 0.05), including THEM4 and APOL2 (Q -value = 6.93×10^{-3} and 3.98×10^{-2} , respectively) (Table 1).

SMR results

From the three proteins that were significant and had a $PP4 > 0.5$ in the discovery, the levels of THEM4 and APOL2 were identified to be causally associated with stroke outcome in the SMR analysis (p -value = 2.8×10^{-3} THEM4 and p -value = 9.4×10^{-3} APOL2) (Table 2). GSTP1 levels were not demonstrated to be causally associated with stroke outcome in the SMR analysis.

Table 2: Significant SMR results

Gene	b_SMR	se_SMR	p_SMR
THEM4	0.62	0.21	2.85E-03
APOL2	-0.84	0.32	9.40E-03
CCDC91	1.52	0.61	1.21E-02
LACTB2	1.54	0.62	1.32E-02
PTGR2	-0.69	0.29	1.86E-02
RASA2	-1.70	0.74	2.23E-02
NDUFAF1	-0.61	0.27	2.43E-02
SOD3	0.69	0.31	2.52E-02
ZADH2	1.02	0.47	2.92E-02
PFKM	1.56	0.72	3.12E-02
GPRIN3	1.03	0.51	4.17E-02
TTC38	0.32	0.16	4.22E-02
ECI1	0.82	0.40	4.26E-02
C14orf159;DGLUCY	-0.08	0.04	4.29E-02
LIN7A	-2.42	1.20	4.41E-02
CA4	0.57	0.29	4.96E-02

Significant results from the SMR analysis. *b*_SMR: beta value in the SMR analysis, *se*: standard error in the SMR analysis and *p*_SMR is the pvalue in the SMR analysis.

TWAS results

Considering the transcriptomic data from the dPFC, the TWAS analysis revealed a significant association of the mRNA levels for THEM4 and GSTP1 with the mRS3. Only THEM4 presented a high probability in the colocalization analysis for the eQTL variant to be shared between the mRNA levels and the phenotype (Table 3)

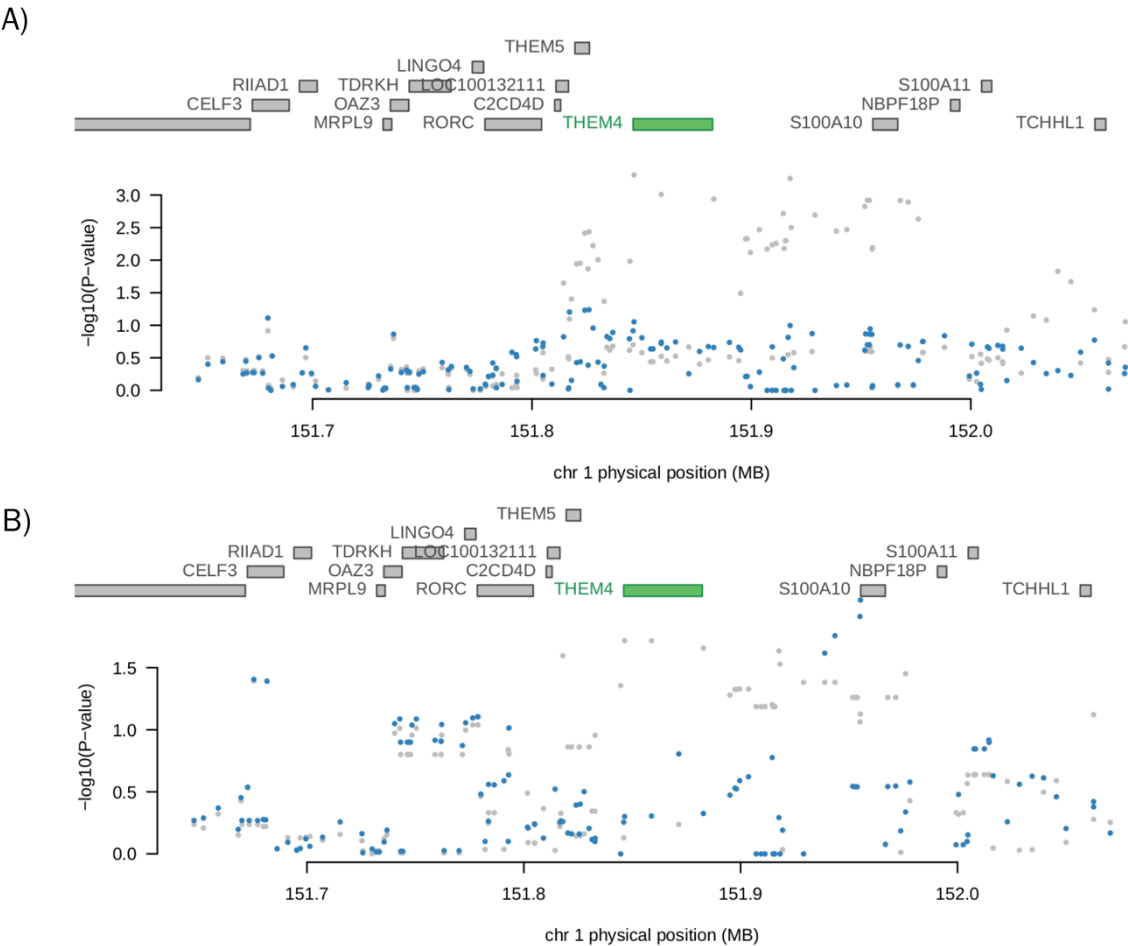
Table 3: TWAS results using transcriptomic data from the DLPC

ID	EQTL.ID	EQTL.Z	TWAS.Z	TWAS.P	COLOC.PP4
THEM4	rs13320	-15.62	3.36	7.71E-04	0.80
APOL2	rs129607	-4.79	1.459	1.45E-01	NA
GSTP1	rs614080	9.94	-2.41	1.60E-02	0.21

Conditional analysis

We conditioned the GWAS results for THEM4 considered in the discovery and the replication2 by their cis-protein levels. We found a significant loss of statistical significance in both GWAS when considering THEM4 levels, demonstrating the relevance of the cis-protein regulation in the association of THEM4 genetic variants with the stroke outcome (Figure 1).

Figure 1: Conditional analysis



Conditional analysis for the GODS GWAS used in the discovery (A) and for the new GWAS with GODS criteria used in the Replication 2 (B).

Pathway enrichment analysis

Using the KEGG database for the pathway enrichment analysis, we found the fatty acid elongation and the AMPK signaling pathways to be enriched among the five significant proteins associated with the stroke outcome from the meta-analysis.

Table 3: Pathway enrichment analysis results

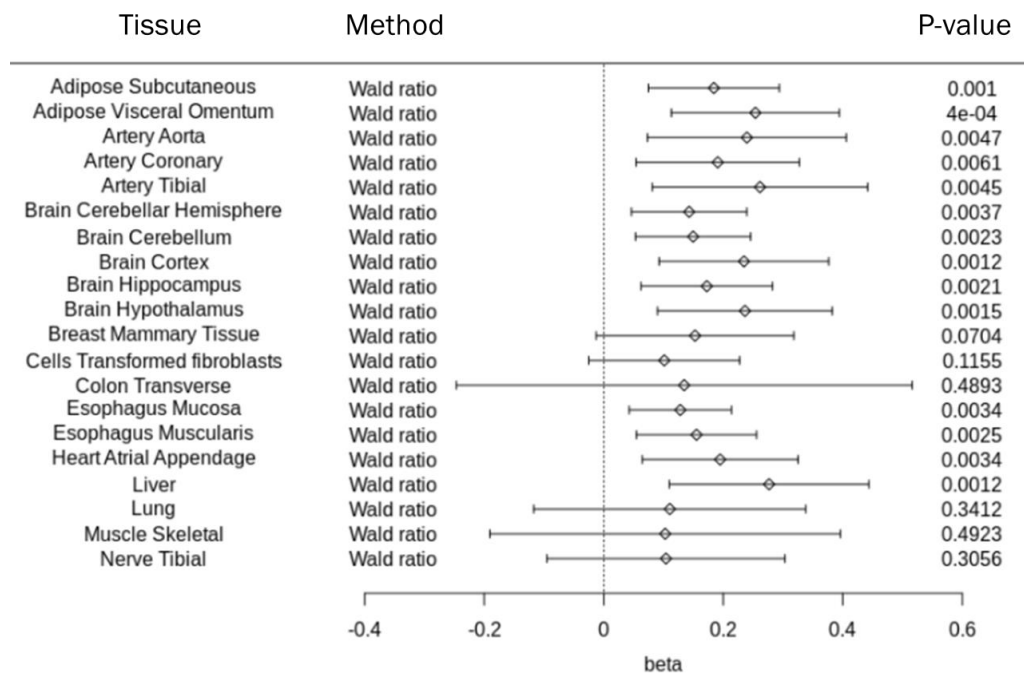
Gene Set	Description	Size	Ratio	P-Value
hsa00062	Fatty acid elongation	30	121.18	8.24E-03
hsa04152	AMPK signaling pathway	120	30.296	3.27E-02

Significant pathways enriched from the meta-analysis results. The gene set is the name of the enriched pathway. The size is the number of genes belonging to the specific pathway. The ratio corresponds to the expected ratio.

Multi tissue MR

To evaluate the tissue specificity for the causal association of THEM4 levels on the stroke outcome, we performed multi tissue MR analyses. We selected 20 tissue specific transcriptomic data from the GTEX to study the causal association between the THEM4 levels and the stroke outcome. Globally, we found that THEM4 levels were associated with the mRS3 in many different tissues, including arteries, different brain regions, heart and liver, among others (Figure 2).

Figure 2: Mendelian Randomization results



Forest plot for the MR results for 20 different tissues included in the GTEx. Each row indicates the causal association for the THEM4 levels in the indicated tissue in the stroke outcome.

DISCUSSION

The results for the stroke outcome PWAS indicate a cis-regulation at the proteomic level in the determination of the mRS3. As far as we know, this is the first study integrating genomic and proteomic data in the study of the stroke outcome.

THEM4 protein levels were found to be negatively associated with the long-term stroke outcome. These results were validated in two independent cohorts (one for proteomic data and another for genomic data). Moreover, our results indicated that THEM4 levels were causally mediating the genetic association at this locus with the mRS at three months. THEM4 or Carboxy-Terminal Modulator Protein (CTMP) is a mitochondrial protein which is translocated into the cytosol after damage signals such as ischemia. In the cytosol, THEM4 inhibits the protein kinase B (PKB). The inhibition of PKB induces apoptotic processes, which are hallmarks in postischemic neurons (24). In cell models of oxygen/glucose deprivation (OGD) have been proved the involvement of THEM4 in apoptotic processes after hypoxia (25). In animal models of rats, the implication of CTMP after ischemia has been also demonstrated. Rats after the induction of an ischemic injury, increase CTMP levels and it leads to a decrease in the brain ischemic tolerance (26). The increase in CTMP has been also being linked with a worsening of neurological outcomes (26). Interestingly, THEM4 is modulated by sevoflurane, a volatile anaesthetics drug. In a rat model of cerebral ischemia, the preconditioning with sevoflurane was induced neuroprotection via the inhibition of CTMP (27,28).

The effect of volatile anaesthetics on stroke has been also studied in human, rising some controversial results. Raub et al. (29), included patients which undergone a surgery and were administrated with volatile anaesthetics. The anaesthetics concentration was considered the primary exposure variable and the stroke occurrence the primary outcome of the study. The results of the study demonstrated a protective effect of volatile anaesthetics, including sevoflurane, which was dose-dependent in patients undergoing a noncardiac surgery (29). An observational analysis in patients from the MR-CLEAN trial also demonstrated that the use of volatile anaesthetics supposed a better neurological outcome in stroke patients receiving endovascular treatment compared with general anaesthesia (30). The AnStroke trial was designed to compare the effect of general anaesthesia versus conscious sedation during the endovascular treatment. Non-significant results were found in terms of acute nor long-term outcome between groups (31). However, these studies were not designed to test the effect of sevoflurane in the stroke outcome and different volatile anaesthetics were used. Further studies would need to be designed to demonstrate in humans the effect of sevoflurane as a neuroprotective drug.

APOL2 levels were also significantly associated with stroke outcome in the PWAS meta-analysis. APOL2 is an apolipoprotein almost exclusively expressed in the brain. Genetic variants in APOL2 locus has been previously associated with ischemic stroke risk and also with the risk for atherothrombotic stroke (32).

Apart from THEM4 and APOL2, three additional proteins were found significantly associated with stroke outcome in the meta-analysis. We included them in a pathway enrichment analysis, and we found an enrichment of the fatty acid elongation and AMPK regulation pathways. The synthesis of fatty acids is linked with the brain injury in murine stroke models (33). Moreover, the free fatty acid levels in cerebrospinal fluid have been studied as potential prognostic value in ischemic stroke (34). In the other hand, the AMPK pathway is cue in the regulation of molecular mechanisms occurring after stroke (35). AMPK is a kinase protein with a key action

after deleterious stimulus such as ischemia. The increase in AMPK levels has been found to be related with a reduction in neuroinflammation, oxidative stress, apoptosis, mitochondrial dysfunction and glutamate excitotoxicity (36). Thus, it has been suggested as a good target in stroke therapy.

The main limitation of this study is that we were not able to study by PWAS other proteomic datasets obtained from other brain tissues. Despite the TWAS results seems to show that the regulation of THEM4 is not specific for the dPFC, the regulation of THEM4 protein in association with stroke outcome would need to be checked. However, our study has supported the power of -omic studies in humans to reach results that can be validated in animal models. In this case, the relevance of THEM4 in ischemic processes was tested and confirmed in previous animal models. Given our results in human and the previous results *in vitro* and *in vivo*, it would be interesting to study the regulation of THEM4 as a modulator of the neurological outcome.

Funding

EPIGENESIS project (Carlos III Institute/ Fondo Europeo de Desarrollo Regional (FEDER)-PI17/02089, Marató TV3 and Fundació MútuaTerrassa), MAESTRO project (Carlos III Institute/ FEDER - PI18/01338), iBioStroke project (Eranet-Neuron, European research grants), the EPINEXO project- PI20/00678 (Carlos III Institute / FEDER), SEDMAN Study (Boehringer Ingelheim), APHAS Study (Pfizer/Bristol Myers), Fondo Europeo de Desarrollo Regional (FEDER), 2017SGR-1427 (AGAUR), the RETICS Network INVICTUS + and the RICORDS Stroke network

J. Cárcel-Márquez is supported by an AGAUR Contract (agència de gestió d'ajuts universitaris i de recerca; FI_DGR 2019, grant number 2019_FI_B 00853) co-financed by Fons Social Europeu (FSE).

I. Fernandez is the recipient of a research contract from the Miguel Servet Program (CP12/03298) from the Carlos III Institute.

M. Lledós is supported by a PFIS Contract (Contratos Predoctorales de Formación en Investigación en Salud): FI19/00309

C. Gallego-Fabrega is supported by a Sara Borrell Contract (CD20/00043) from Instituto de Salud Carlos III and Fondo Europeo de Desarrollo Regional (ISCIII-FEDER).

E. Muiño is supported by a Río Hortega Contract (CM18/00198) from the Instituto de Salud Carlos III.

Acknowledgments

We thank the International Stroke Genetics Consortium, the Spanish Stroke Genetics Consortium, the International Stroke Genetics Consortium, and the Global Alliance for Stroke acute and long-term outcome genetics and the RETICS Network INVICTUS (RD16/0019/0011).

Author Contributions

I.FC. and J.JC. are responsible for the study design and the analytical accuracy of the manuscript. N.C. performed statistical analyses and prepared the manuscript. C.GF., J.CM., L.LLC, M.LL., E.M. and J.MC reviewed the manuscript and contributed to the data interpretation. All the authors reviewed and approved the manuscript content.

Institutional Review Board

The study was conducted according to the guidelines of the Declaration of Helsinki. The different studies included in this manuscript were responsible for the aprovement by the Institutional Review Board.

Informed Consent

All the studies included as part of this article were responsible for the obtention of the informed consent of the patients.

Data Availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no competing interests.

REFERENCES

1. GBD 2016 Stroke Collaborators. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* mayo de 2019;18(5):439-58.
2. Debette S, Mishra A, Malik R, Hachiya T, Jürgenson T, Namba S, et al. Stroke genetics informs drug discovery and risk prediction across ancestries [Internet]. 2022 [citado 17 de agosto de 2022]. Disponible en: <https://doi.org/10.21203/rs.3.rs-1175817/v1>
3. Ibanez L, Heitsch L, Carrera C, Farias FHG, Del Aguila JL, Dhar R, et al. Multi-ancestry GWAS reveals excitotoxicity associated with outcome after ischaemic stroke. *Brain J Neurol.* 29 de julio de 2022;145(7):2394-406.
4. Mola-Caminal M, Carrera C, Soriano-Tárraga C, Giralte-Steinhauer E, Díaz-Navarro RM, Tur S, et al. PATJ Low Frequency Variants Are Associated With Worse Ischemic Stroke Functional Outcome. *Circ Res.* 4 de enero de 2019;124(1):114-20.
5. Söderholm M, Pedersen A, Lorentzen E, Stanne TM, Bevan S, Olsson M, et al. Genome-wide association meta-analysis of functional outcome after ischemic stroke. *Neurology.* 19 de marzo de 2019;92(12):e1271-83.
6. Heitsch L, Ibanez L, Carrera C, Binkley MM, Strbian D, Tatlisumak T, et al. Early Neurological Change After Ischemic Stroke Is Associated With 90-Day Outcome. *Stroke.* enero de 2021;52(1):132-41.
7. Li B, Ritchie MD. From GWAS to Gene: Transcriptome-Wide Association Studies and Other Methods to Functionally Understand GWAS Discoveries. *Front Genet.* 30 de septiembre de 2021;12:713230.
8. Brandes N, Linial N, Linial M. PWAS: proteome-wide association study—linking genes and phenotypes by functional variation in proteins. *Genome Biol.* 14 de julio de 2020;21(1):173.
9. Yang J, Yan B, Fan Y, Yang L, Zhao B, He X, et al. Integrative analysis of transcriptome-wide association study and gene expression profiling identifies candidate genes associated with stroke. *PeerJ.* 2019;7:e7435.
10. Traylor M, Persyn E, Tomppo L, Klasson S, Abedi V, Bakker MK, et al. Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies. *Lancet Neurol.* mayo de 2021;20(5):351-61.
11. Cullell N, Gallego-Fábrega C, Cárcel-Márquez J, Muiño E, Llucià-Carol L, Lledós M, et al. ICA1L Is Associated with Small Vessel Disease: A Proteome-Wide Association Study in Small Vessel Stroke and Intracerebral Haemorrhage. *Int J Mol Sci.* 15 de marzo de 2022;23(6):3161.
12. Wu BS, Chen SF, Huang SY, Ou YN, Deng YT, Chen SD, et al. Identifying causal genes for stroke via integrating the proteome and transcriptome from brain and blood. *J Transl Med.* 21 de abril de 2022;20(1):181.

13. Zhang C, Qin F, Li X, Du X, Li T. Identification of novel proteins for lacunar stroke by integrating genome-wide association data and human brain proteomes. *BMC Med.* 23 de junio de 2022;20(1):211.
14. Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA. Religious Orders Study and Rush Memory and Aging Project. *J Alzheimers Dis JAD.* 2018;64(s1):S161-89.
15. Beach TG, Adler CH, Sue LI, Serrano G, Shill HA, Walker DG, et al. Arizona Study of Aging and Neurodegenerative Disorders and Brain and Body Donation Program. *Neuropathol Off J Jpn Soc Neuropathol.* agosto de 2015;35(4):354-89.
16. Wingo AP, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, et al. Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis. *Nat Genet.* febrero de 2021;53(2):143-6.
17. Jiang L, Zheng Z, Qi T, Kemper KE, Wray NR, Visscher PM, et al. A resource-efficient tool for mixed model association analysis of large-scale data. *Nat Genet.* diciembre de 2019;51(12):1749-55.
18. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet.* mayo de 2016;48(5):481-7.
19. Li J, Tseng GC. An adaptively weighted statistic for detecting differential gene expression when combining multiple transcriptomic studies. *Ann Appl Stat.* junio de 2011;5(2A):994-1019.
20. Ye Z, Mo C, Ke H, Yan Q, Chen C, Kochunov P, et al. Meta-Analysis of Transcriptome-Wide Association Studies across 13 Brain Tissues Identified Novel Clusters of Genes Associated with Nicotine Addiction. *Genes.* 23 de diciembre de 2021;13(1):37.
21. Zhiguang Huo. AWFisher: An R package for fast computing for adaptively weighted fisher's method. 2022;
22. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* 1 de julio de 2005;33(Web Server issue):W741-748.
23. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Loos R, editor. *eLife.* 30 de mayo de 2018;7:e34408.
24. Parcellier A, Tintignac LA, Zhuravleva E, Cron P, Schenk S, Bozulic L, et al. Carboxy-Terminal Modulator Protein (CTMP) is a mitochondrial protein that sensitizes cells to apoptosis. *Cell Signal.* abril de 2009;21(4):639-50.
25. Huang CY, Chen JJ, Wu JS, Tsai HD, Lin H, Yan YT, et al. Novel link of anti-apoptotic ATF3 with pro-apoptotic CTMP in the ischemic brain. *Mol Neurobiol.* abril de 2015;51(2):543-57.
26. Li J, Shan W, Zuo Z. Age-Related Upregulation of Carboxyl Terminal Modulator Protein Contributes to the Decreased Brain Ischemic Tolerance in Older Rats. *Mol Neurobiol.* julio de 2018;55(7):6145-54.

27. Chen Y, Nie H, Tian L, Tong L, Deng J, Zhang Y, et al. Sevoflurane preconditioning-induced neuroprotection is associated with Akt activation via carboxy-terminal modulator protein inhibition. *Br J Anaesth*. febrero de 2015;114(2):327-35.
28. Neag MA, Mitre AO, Catinean A, Mitre CI. An Overview on the Mechanisms of Neuroprotection and Neurotoxicity of Isoflurane and Sevoflurane in Experimental Studies. *Brain Res Bull*. diciembre de 2020;165:281-9.
29. Raub D, Platzbecker K, Grabitz SD, Xu X, Wongtangman K, Pham SB, et al. Effects of Volatile Anesthetics on Postoperative Ischemic Stroke Incidence. *J Am Heart Assoc*. 2 de marzo de 2021;10(5):e018952.
30. Berkhemer OA, van den Berg LA, Fransen PSS, Beumer D, Yoo AJ, Lingsma HF, et al. The effect of anesthetic management during intra-arterial therapy for acute stroke in MR CLEAN. *Neurology*. 16 de agosto de 2016;87(7):656-64.
31. Löwhagen Hendén P, Rentzos A, Karlsson JE, Rosengren L, Leiram B, Sundeman H, et al. General Anesthesia Versus Conscious Sedation for Endovascular Treatment of Acute Ischemic Stroke: The AnStroke Trial (Anesthesia During Stroke). *Stroke*. junio de 2017;48(6):1601-7.
32. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, et al. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat Genet*. octubre de 2012;44(10):10.1038/ng.2397.
33. Janssen L, Ai X, Zheng X, Wei W, Caglayan AB, Kilic E, et al. Inhibition of Fatty Acid Synthesis Aggravates Brain Injury, Reduces Blood-Brain Barrier Integrity and Impairs Neurological Recovery in a Murine Stroke Model. *Front Cell Neurosci* [Internet]. 2021 [citado 15 de diciembre de 2022];15. Disponible en: <https://www.frontiersin.org/articles/10.3389/fncel.2021.733973>
34. Wei XJ, Han M, Wei GC, Duan CH. Prognostic value of cerebrospinal fluid free fatty acid levels in patients with acute ischemic stroke. *Front Hum Neurosci* [Internet]. 2015 [citado 15 de diciembre de 2022];9. Disponible en: <https://www.frontiersin.org/articles/10.3389/fnhum.2015.00402>
35. Qin C, Yang S, Chu YH, Zhang H, Pang XW, Chen L, et al. Signaling pathways involved in ischemic stroke: molecular mechanisms and therapeutic interventions. *Signal Transduct Target Ther*. 6 de julio de 2022;7(1):1-29.
36. Jiang S, Li T, Ji T, Yi W, Yang Z, Wang S, et al. AMPK: Potential Therapeutic Target for Ischemic Stroke. *Theranostics*. 10 de agosto de 2018;8(16):4535-51.

Supplementary

Identification of genetic regulation at the proteomic level in the long-term stroke outcome: A Proteome-Wide Association study.

Natalia Cullell, MSc^{*1,2,3}, Cristina Gallego-Fábrega, PhD¹, Jara Cárcel, MSc¹, Elena Muiño, MD, PhD¹, Laia Llucià-Carol^{1,8}, Miquel Lledós, MSc¹, Jesús M Martín-Campos¹, Israel Fernández-Cadenas, PhD^{1,2**}, Jerzy Krupinski, MD, PhD^{2 **}

1- Stroke Pharmacogenomics and Genetics group, Biomedical Research Institute Sant Pau (IIB Sant Pau). 08041. Barcelona

2- Neurology. Hospital Universitari MútuaTerrassa/ Fundacio Docència i Recerca MutuaTerrassa, 08221. Terrassa, Spain.

3- Facultat de Medicina. Universitat de Barcelona, 08036. Barcelona, Spain

Table S1: Discovery PWAS results

ID	pQTL	pQTL Z	PWAS Z	PWAS P	COLOCPP4
FXN	rs2498432	5.73	-4.07	4.77E-05	0.4
TBC1D1	rs12511756	5.67	-3.21	1.33E-03	0.3
GFM1	rs17699324	-5.24	-3.20	1.38E-03	0.3
LACTB2	rs12681420	6.6	3.18	1.49E-03	0.4
ADCY9	rs2230739	-5.23	3.11	1.88E-03	0.3
THEM4	rs16833668	-10.76	3.04	2.35E-03	0.5
TTC38	rs6008495	7.72	2.88	3.92E-03	0.2
NDUFAF1	rs17730589	-8.41	-2.80	5.08E-03	0.2
GPNMB	rs199357	-8.67	2.70	6.95E-03	0.1
KIAA1549	rs10259252	4.6	-2.68	7.43E-03	0.1
CYB5R3	rs137087	-6.84	-2.59	9.48E-03	0.3
PRKCB	rs196002	-5.16	2.58	9.82E-03	0.2
SLC22A23	rs9503598	-4.52	2.58	9.88E-03	0.1
RALGPS1	rs10987552	3.26	2.57	1.01E-02	0.02
GSTP1	rs7941648	-8.99	-2.56	1.04E-02	0.7
GNPDA1	rs164080	-4.46	-2.55	1.08E-02	0.1
PAG1	rs10504731	4.46	2.54	1.11E-02	0.1
GUCY1A3	rs3796582	3.46	2.53	1.13E-02	0.003
RPAP1	rs1200353	-3.8	-2.52	1.17E-02	0.03
HDHD3	rs1043836	-7.3	-2.52	1.19E-02	0.1
MECR	rs547976	5.91	2.50	1.24E-02	0.02
TMEM245	rs489283	6.75	2.50	1.24E-02	0.2
CARS2	rs7337089	-8.25	-2.49	1.30E-02	0.1
PARK2	rs9346868	-3.72	-2.47	1.34E-02	0.02
GC	rs222047	-3.89	2.46	1.40E-02	0.01
DUS2	rs7193701	3.55	2.45	1.44E-02	0.01
SNPH	rs11905792	-4.55	2.44	1.46E-02	0.5
GRID1	rs7075405	-5.87	2.42	1.54E-02	0.1
ALDH9A1	rs12408101	-7.27	-2.40	1.63E-02	0.2
DCXR	rs4625783	-3.61	2.40	1.64E-02	0.01
SOD3	rs699473	6.24	2.39	1.67E-02	0.2
PFKFB3	rs4310524	-3.2	2.39	1.69E-02	0.02
PRICKLE1	rs3747562	-3.97	2.38	1.72E-02	0.1
APOL2	rs8136528	-6.19	-2.37	1.77E-02	0.8
JMJD7	rs8025614	-6.51	2.37	1.77E-02	0.1
PPA1	rs7905657	-3.64	2.31	2.07E-02	0.01
SRSF9	rs11065126	-7.06	-2.30	2.13E-02	0.2
EGFR	rs759159	-4.44	2.30	2.14E-02	0.3
PSEN1	rs17781801	-5.09	2.29	2.23E-02	0.1
GDE1	rs2187659	-4.3	2.28	2.25E-02	0.02
CSAD	rs822682	-4.38	2.28	2.28E-02	0.03
LSP1	rs11041476	5.9	2.27	2.31E-02	0.2

LXN	rs3867391	8.34	2.27	2.33E-02	0.3
DCAKD	rs2239919	-3.7	2.25	2.47E-02	0.02
PREP	rs717324	-4.49	-2.24	2.49E-02	0.04
PNP	rs1049564	-4.97	-2.24	2.50E-02	0.1
FAM98C	rs2074981	-4.71	2.24	2.53E-02	0.03
PYGM	rs624975	-4.5	2.22	2.62E-02	0.02
ZADH2	rs7230037	-9.33	2.20	2.81E-02	0.1
WVOX	rs7184686	6.28	-2.19	2.84E-02	0.03
DGUOK	rs1653259	-3.76	2.19	2.85E-02	0.00
LRPPRC	rs7568481	5.72	-2.17	3.00E-02	0.1
GPRIN3	rs2005104	6.02	2.16	3.06E-02	0.2
LYSMD2	rs16964505	4.4	-2.16	3.12E-02	0.02
LIN7A	rs7977901	-5.71	-2.15	3.18E-02	0.1
HMGCLL1	rs10498804	5.22	2.14	3.20E-02	0.04
UAP1L1	rs7467117	-7.49	-2.13	3.34E-02	0.1
WIPF1	rs1010027	-3.57	2.12	3.41E-02	0.003
ESYT2	rs1061733	-4.87	2.12	3.41E-02	0.1
RDH11	rs718212	4.65	2.09	3.65E-02	0.1
PYCRL	rs11549789	-5.96	2.08	3.71E-02	0.03
PIK3CD	rs4240895	4.08	-2.08	3.76E-02	0.01
CRAT	rs7849270	-3.42	2.08	3.79E-02	0.01
TBCA	rs13185613	-4.19	-2.07	3.83E-02	0.01
CARKD	rs179356	10	-2.07	3.87E-02	0.1
PFKM	rs2269935	-7.48	2.06	3.90E-02	0.1
N6AMT2	rs17318109	-4.7	2.06	3.93E-02	0.05
PLCL1	rs1595823	5.79	2.04	4.10E-02	0.1
C14orf159	rs4900072	-13.62	-2.04	4.12E-02	0.1
TAPT1	rs16893162	4	2.04	4.17E-02	0.00
PCYT1A	rs884534	-4.94	-2.02	4.30E-02	0.1
SNUPN	rs8025681	-3.91	2.02	4.31E-02	0.003
ALKBH4	rs2410988	-4.65	-2.02	4.31E-02	0.1
DNAJB1	rs8104440	4.41	2.02	4.37E-02	0.03
PAK4	rs692191	4.18	2.01	4.41E-02	0.01
TRIM65	rs8066711	5.96	-2.01	4.44E-02	0.1
ESD	rs8192892	-6.69	-2.00	4.52E-02	0.03
MTHFD1L	rs6922269	3.42	-2.00	4.53E-02	0.004
CCDC91	rs1133028	-6.38	2.00	4.55E-02	0.2
ABHD10	rs17429033	-7.13	-2.00	4.58E-02	0.1
RCSD1	rs926521	-5.94	-1.99	4.62E-02	0.1
SLC2A11	rs875643	-7.97	1.99	4.62E-02	0.03
CMC2	rs2602424	-10.21	1.99	4.67E-02	0.1
ITGA6	rs10166912	3.9	-1.97	4.84E-02	0.003
TRAP1	rs2791	-7.91	1.97	4.88E-02	0.05

Significant proteins (P WAS p-value < 0.05) in the P WAS from the Discovery cohort. COLOC PP4 indicates the posterior probability 4 for the proteins, genetic variant and trait to colocalize.

Table S2: Replication1 PWAS results

ID	pQTL	pQTL.Z	PWAS.Z	PWAS.P	COLOCPP4
TTC38	rs9627287	5.41	4.01	6.14E-05	0.033
APOL2	rs9619597	-6.5	-3.31	9.32E-04	0.759
PREP	rs9500087	-4.4	-3.16	1.56E-03	0.147
TRAP1	rs11646287	-6.87	3.11	1.85E-03	0.451
TBC1D1	rs12511756	3.65	-2.97	2.95E-03	0.015
THEM4	rs16833668	-8.69	2.84	4.51E-03	0.484
LACTB2	rs2035196	4.28	2.76	5.80E-03	0.033
GPNMB	rs7808488	-5.22	2.64	8.39E-03	0.226
LXN	rs17699478	5.27	2.36	1.83E-02	0.109
LSP1	rs907611	3.99	2.33	1.96E-02	0.012
PFKM	rs10875746	-7.7	2.30	2.13E-02	0.149
PSEN1	rs10498532	-4.62	2.28	2.25E-02	0.03
ZADH2	rs7230037	-4.52	2.23	2.55E-02	0.02
MECR	rs547976	5.97	2.15	3.14E-02	0.021
CARKD	rs2765341	7.97	-2.15	3.18E-02	0.073
N6AMT2	rs9579937	-4.05	2.09	3.67E-02	0.012
GSTP1	rs1695	-9.13	-2.06	3.93E-02	0.315
NDUFAF1	rs7168431	-4.96	-1.99	4.66E-02	0.181
C14orf159	rs4900071	-10.59	-1.99	4.71E-02	0.096
PYCRL	rs12546272	5.37	1.98	4.75E-02	0.047
JMJD7	rs1197690	-4.92	1.96	4.96E-02	0.074

Significant proteins (PWAS p-value < 0.05) in the PWAS from the Replication cohort1 from the significant proteins in the discovery. COLOC PP4 indicates the posterior probability 4 for the proteins, genetic variant and trait to colocalize.

Table S3: Replication2 PWAS results

ID	PQTL	EQTL Z	PWAS Z	PWAS P	COLOC PP4
THEM4	rs16833668	-10.76	2.27	2.35E-02	0.159
SLC22A23	rs9503598	-4.52	-2.23	2.59E-02	0.054
RDH11	rs718212	4.65	-2.15	3.14E-02	0.05

Significant proteins (PWAS p-value < 0.05) in the PWAS from the Replication cohort2 from the significant proteins in the discovery. COLOC PP4 indicates the posterior probability 4 for the proteins, genetic variant and trait to colocalize.

Table S4: Meta-analysis of the discovery, replication1 and replication2 results.

ID	P_value	Q_value
TTC38	1.89E-05	8.68E-04
THEM4	1.51E-04	6.93E-03
TBC1D1	2.39E-04	1.10E-02
LACTB2	4.86E-04	2.24E-02
APOL2	8.66E-04	3.98E-02
GPNMB	2.63E-03	1.21E-01
TRAP1	3.85E-03	1.77E-01
GFM1	5.61E-03	2.58E-01
NDUFAF1	8.81E-03	4.05E-01
MECR	1.34E-02	6.15E-01
GSTP1	1.40E-02	6.42E-01
LXN	1.45E-02	6.65E-01
PSEN1	1.65E-02	7.61E-01
CYB5R3	1.70E-02	7.80E-01
JMJD7	2.05E-02	9.44E-01
KIAA1549	2.09E-02	9.61E-01
CARS2	2.44E-02	1
PFKM	2.51E-02	1
ALDH9A1	3.15E-02	1
RDH11	3.26E-02	1
PRKCB	3.42E-02	1
CARKD	3.45E-02	1
DCAKD	3.88E-02	1
GNPDA1	3.90E-02	1
N6AMT2	3.92E-02	1
HDHD3	4.13E-02	1
RPAP1	4.20E-02	1
TMEM245	4.43E-02	1
PYCRL	4.61E-02	1
C14orf159	4.98E-02	1

Meta-analysis of the PWAS results from the discovery, replication1 and replication2.

Q-value indicates the Bonferroni corrected P-value.

Table S5: Significant TWAS results

ID	EQTL.ID	EQTL.Z	TWAS.Z	TWAS.P	COLOC.PP4
EYA1	rs1908036	-3.77	3.95	7.86E-05	0.12
COL23A1	rs11249560	-4.2	3.82	1.33E-04	0.04
C1orf159	rs540049	7.54	3.68	2.37E-04	0.77
PEBP1	rs904661	5.18	3.45	5.56E-04	0.80
SCHIP1	rs13064339	-7.44	-3.45	5.66E-04	0.89
SCAI	rs10986743	-3.96	3.44	5.74E-04	0.27
ABCA4	rs553608	6.94	3.43	6.01E-04	0.84
RPL23AP82	rs131734	-3.68	-3.41	6.40E-04	0.01
SCAP	rs878659	9.01	-3.38	7.27E-04	0.69
THEM4	rs13320	-15.62	3.36	7.71E-04	0.80
CASQ1	rs3747623	-15.75	-3.36	7.91E-04	0.77
UTP23	rs7823271	-7.42	-3.25	1.15E-03	0.72
HCG4	rs385492	9.71	-3.23	1.25E-03	0.79
B3GALT1	rs1448839	3.81	-3.22	1.30E-03	0.01
DDX17	rs12172202	4.41	3.13	1.73E-03	0.36
DNMT3B	rs4911259	5.38	3.13	1.78E-03	0.63
ADAR	rs1127309	-7.99	3.12	1.80E-03	0.58
SH3TC1	rs11735274	3.22	3.08	2.10E-03	0.00
GUCY1A3	rs10517621	6.61	3.04	2.35E-03	0.04
FAM71D	rs8014841	-4.8	-2.99	2.83E-03	0.51
OR14I1	rs28498595	-4.39	-2.98	2.92E-03	0.05
MTHFR	rs1476413	6.13	2.97	2.97E-03	0.52
LOC100506207	rs1335639	10.77	2.96	3.11E-03	0.34
ZCRB1	rs1376931	7.77	-2.96	3.12E-03	0.30
LAMC2	rs2276543	13.84	-2.94	3.25E-03	0.56
NDNL2	rs17681991	-4.09	-2.93	3.40E-03	0.03
AARS2	rs325001	-6.31	-2.92	3.45E-03	0.16
COMMD7	rs927239	4.97	2.92	3.54E-03	0.58
TC2N	rs1884840	14.87	-2.91	3.61E-03	0.61
C2orf40	rs3820893	-7.05	2.90	3.68E-03	0.02
ACOX1	rs11868566	-3.71	2.90	3.75E-03	0.00
STEAP3	rs838090	-5.34	-2.90	3.75E-03	0.14
SMARCAL1	rs6708185	4.77	2.90	3.75E-03	0.25
NOP56	rs6132905	10.33	-2.88	3.92E-03	0.37
SYT10	rs861525	-8.68	2.88	3.99E-03	0.18
TG	rs853321	6.07	2.88	4.03E-03	0.34
ZFYVE27	rs10786370	6.38	-2.87	4.05E-03	0.46
BTBD2	rs12461737	-7.73	-2.87	4.08E-03	0.46
ERAP1	rs27038	10.55	-2.87	4.12E-03	0.77
GBP4	rs12063900	4.33	2.86	4.18E-03	0.13
KIF9	rs4683297	-3.96	2.86	4.25E-03	0.03
SLC12A9	rs2396755	3.46	-2.86	4.25E-03	0.02
NUDT9P1	rs7069106	-3.25	-2.85	4.34E-03	0.00
KHNYN	rs4982912	-9.69	2.83	4.59E-03	0.46

MRPS10	rs7750260	8.21	-2.82	4.80E-03	0.36
PLSCR1	rs1997604	-5.39	-2.81	4.96E-03	0.43
REEP1	rs17510512	-3.92	-2.81	5.02E-03	0.01
AMDHD1	rs7486703	-6.77	-2.80	5.10E-03	0.48
C9orf102	rs7023522	-6.58	2.79	5.35E-03	0.39
PLA2G5	rs660414	-8.28	2.78	5.51E-03	0.41
LACTB2	rs13280922	-6.56	2.77	5.62E-03	0.38
BICD2	rs10992469	5.08	2.76	5.70E-03	0.10
NCLN	rs566581	-9	2.75	5.90E-03	0.34
MEPE	rs10516794	-6.29	2.75	6.04E-03	0.12
USP40	rs838705	8.12	2.73	6.25E-03	0.30
SNRPN	rs752874	-7.62	-2.73	6.36E-03	0.05
KIAA0040	rs2272785	-6.93	-2.72	6.46E-03	0.36
TMEM106C	rs12721420	6.14	-2.71	6.80E-03	0.51
NMT1	rs9898793	5.7	-2.71	6.83E-03	0.35
NDUFAF1	rs3743038	-8.84	-2.70	6.93E-03	0.14
L1TD1	rs2457820	-4.71	-2.69	7.11E-03	0.03
NDUFV3	rs4148974	-8.99	-2.69	7.14E-03	0.24
JOSD1	rs5750630	-6.19	-2.69	7.22E-03	0.38
TTC38	rs6008495	8.93	2.68	7.34E-03	0.15
AEBP1	rs3087370	-8.32	2.67	7.64E-03	0.07
ITGA1	rs2938793	-3.66	-2.66	7.93E-03	0.01
XKR9	rs6991329	15.34	-2.65	8.07E-03	0.31
ESYT2	rs11761517	-5.5	-2.65	8.14E-03	0.06
ALAD	rs818702	-7.37	-2.64	8.23E-03	0.12
NUPL2	rs7808488	-11.72	2.64	8.39E-03	0.31
EIF5	rs2281824	-5.09	-2.63	8.44E-03	0.27
CCDC111	rs12108497	13.37	-2.63	8.46E-03	0.63
CATSPERB	rs10134181	8.41	-2.63	8.54E-03	0.30
ZNF567	rs3108545	-5.36	2.63	8.56E-03	0.29
SERPINB1	rs12525298	6.59	2.63	8.59E-03	0.19
OCEL1	rs3745164	-4.54	-2.63	8.65E-03	0.09
RPL13	rs3751691	-5.08	2.62	8.67E-03	0.13
GXYLT1	rs2215454	6.29	-2.61	9.05E-03	0.29
SLC29A3	rs12412201	-4.96	-2.60	9.30E-03	0.03
MRPL34	rs6903	8.5	2.60	9.43E-03	0.29
WISP1	rs2013146	-4.75	2.59	9.66E-03	0.03
CHRNA6	rs11984732	-3.55	2.58	9.92E-03	0.01
LOC100272217	rs4740350	10.33	-2.58	9.93E-03	0.50
ITGA4	rs10205700	-4.29	2.57	1.00E-02	0.03
PTK7	rs3793017	-3.85	-2.57	1.01E-02	0.01
APOA1BP	rs4661188	-13.73	-2.57	1.03E-02	0.28
GFM1	rs2307064	-4.79	-2.57	1.03E-02	0.08
PPHLN1	rs4768401	-7.02	2.57	1.03E-02	0.27
GLDN	rs10519317	-4.42	-2.56	1.03E-02	0.04
PAFAH2	rs3008244	6.98	2.56	1.03E-02	0.28

UGCG	rs667463	-3.8	-2.56	1.04E-02	0.10
ECHDC3	rs718641	-5.93	-2.56	1.05E-02	0.04
SERPINE1	rs2227631	4.04	2.55	1.07E-02	0.01
HDDC3	rs1052532	-5.69	-2.55	1.09E-02	0.26
CBY1	rs2072794	-4.77	-2.54	1.11E-02	0.45
MGC2752	rs11666790	-5.07	-2.54	1.11E-02	0.08
RSRC1	rs7613824	-3.59	-2.53	1.13E-02	0.01
FAM149A	rs11725687	-8.55	-2.53	1.13E-02	0.28
NUDT7	rs11863999	5.64	2.53	1.13E-02	0.24
SLMO1	rs529538	-10.29	2.53	1.14E-02	0.28
RPAP1	rs1200345	6.67	-2.53	1.15E-02	0.25
GLCE	rs4777128	-6.66	-2.52	1.16E-02	0.25
LOC401321	rs9648389	-5.14	-2.51	1.21E-02	0.21
REV3L	rs173287	7.45	-2.51	1.21E-02	0.14
SCNN1D	rs11260570	7.18	2.51	1.22E-02	0.36
H6PD	rs2268175	-9.32	-2.51	1.22E-02	0.25
LOC643837	rs3094315	-6.39	-2.51	1.22E-02	0.25
ZNF169	rs11793128	6.02	2.50	1.24E-02	0.24
FAM101B	rs28445967	6.05	2.50	1.25E-02	0.25
IL1RAP	rs9825567	-2.9	2.50	1.25E-02	0.00
ZFAT	rs17781964	4.74	2.49	1.26E-02	0.10
LRP2BP	rs6815168	-8.72	2.49	1.27E-02	0.10
CENPO	rs17799872	-4.77	2.49	1.29E-02	0.13
ABHD3	rs7343008	-4.49	2.48	1.30E-02	0.14
CTNNA3	rs10762129	-3.32	-2.48	1.31E-02	0.01
TMEM229A	rs10487808	4.77	2.47	1.34E-02	0.02
ZXDC	rs3230	-5.12	-2.47	1.34E-02	0.28
SERHL2	rs137094	5.28	2.47	1.36E-02	0.23
RASA2	rs2640017	4.8	-2.46	1.40E-02	0.13
PLA2G12A	rs9995011	-7.03	-2.45	1.43E-02	0.06
SKOR1	rs11631030	-3.4	-2.45	1.44E-02	0.00
RALGPS1	rs11795097	6.96	2.45	1.45E-02	0.20
TPRG1L	rs4648537	6.64	2.44	1.45E-02	0.04
KIF15	rs17631716	-5.89	-2.44	1.45E-02	0.04
CABYR	rs4800550	-11.19	-2.44	1.46E-02	0.24
SNPH	rs11905792	-5.89	2.44	1.47E-02	0.70
IKBKAP	rs838825	-6.92	-2.44	1.48E-02	0.23
CBLN3	rs4982912	-10.24	2.42	1.54E-02	0.46
RFWD3	rs3815826	8.33	-2.42	1.55E-02	0.10
ZADH2	rs7230037	-7.65	2.41	1.58E-02	0.13
TOMM22	rs1056610	-11.18	-2.41	1.59E-02	0.15
GSTP1	rs614080	9.94	-2.41	1.60E-02	0.21
RAB36	rs4820541	6.82	2.41	1.62E-02	0.20
AACSP1	rs11249583	-5.81	2.40	1.63E-02	0.20
LOC100130673	rs10273165	11.22	2.40	1.64E-02	0.19
PEAK1	rs11632999	8.71	-2.40	1.65E-02	0.42

PGBD1	rs13408	-6.43	-2.40	1.66E-02	0.11
NSUN4	rs17357621	-12.5	2.39	1.66E-02	0.18
HES6	rs11694250	8.33	-2.39	1.70E-02	0.19
CCDC157	rs5763928	5.02	2.39	1.70E-02	0.15
ATP6V0A2	rs882563	7.16	-2.39	1.70E-02	0.10
NEAT1	rs2073800	4.94	2.38	1.73E-02	0.30
ZCCHC7	rs1536876	-8.08	-2.38	1.73E-02	0.19
TTLL3	rs7652460	-5.48	2.38	1.75E-02	0.02
F2R	rs464494	4.77	-2.37	1.77E-02	0.22
IL18R1	rs7568913	-13.09	2.37	1.78E-02	0.08
LAMB3	rs2076220	4.1	-2.37	1.78E-02	0.04
CCT6P1	rs7796045	-3.78	-2.36	1.81E-02	0.04
LOC100131067	rs2270823	13.33	2.36	1.82E-02	0.04
MXRA8	rs1739855	-5.26	-2.36	1.85E-02	0.21
TMEM132E	rs11651584	-6.88	2.36	1.85E-02	0.19
PQLC1	rs12606223	-4.63	2.35	1.88E-02	0.03
PTX3	rs9289983	-6.35	2.35	1.89E-02	0.18
RNH1	rs12420868	-6.53	-2.34	1.93E-02	0.06
FAM131C	rs9442235	-8.53	2.33	1.96E-02	0.40
MPP4	rs16838095	6.1	2.33	1.96E-02	0.12
SEC16B	rs7556610	-3.96	-2.33	2.00E-02	0.01
LOC100506795	rs271764	-7.6	-2.33	2.00E-02	0.18
PNMAL1	rs8107491	7.43	2.32	2.01E-02	0.43
FAM184B	rs4334752	-10.68	2.32	2.02E-02	0.18
POLR1B	rs13412193	7.17	-2.32	2.06E-02	0.08
WDR60	rs710421	4.55	-2.31	2.07E-02	0.05
HS3ST5	rs9372385	3.42	2.31	2.07E-02	0.00
RRP15	rs10863387	14.86	2.31	2.09E-02	0.23
RSPH1	rs2839536	5.66	-2.30	2.13E-02	0.16
FADD	rs1131715	5.56	-2.30	2.16E-02	0.16
SPG7	rs7184655	-6.99	2.29	2.19E-02	0.12
A1BG	rs1265960	-6.34	2.29	2.20E-02	0.16
ZNF195	rs6578341	-8.54	-2.29	2.20E-02	0.17
TRPC6	rs7126833	3.55	-2.29	2.22E-02	0.01
PCBD1	rs923177	7.99	2.29	2.23E-02	0.15
SGCE	rs10953114	2.98	2.28	2.23E-02	0.02
MED24	rs9916158	-11.35	-2.28	2.24E-02	0.08
LOC100132111	rs6587622	-6.11	2.28	2.25E-02	0.16
SSR1	rs1050226	5.05	-2.28	2.28E-02	0.04
LOC646903	rs2130319	5.06	-2.28	2.29E-02	0.11
NAGLU	rs1032070	-3.78	-2.27	2.32E-02	0.00
SETD1A	rs11150599	3.98	2.27	2.32E-02	0.01
HIST2H2BA	rs6600662	4.79	-2.27	2.33E-02	0.08
LOC154761	rs7811904	5	-2.27	2.35E-02	0.12
INTS10	rs10788700	-5.54	-2.26	2.37E-02	0.23
TRIM66	rs10840106	5.84	2.26	2.38E-02	0.09

SULT1C4	rs10173529	5.75	-2.26	2.39E-02	0.03
MSH3	rs1677703	-13.95	-2.26	2.41E-02	0.45
ACTR10	rs7150530	5.81	2.26	2.41E-02	0.57
SLC27A6	rs2577472	-11.04	-2.25	2.42E-02	0.19
NUSAP1	rs17730589	10.41	2.25	2.44E-02	0.17
SPECC1L	rs5996675	4.48	-2.25	2.44E-02	0.27
HIBADH	rs1404282	6.51	-2.25	2.45E-02	0.06
PKD2L1	rs7084075	7.87	2.25	2.47E-02	0.12
SPTBN5	rs1113333	5.26	2.25	2.47E-02	0.05
NSUN6	rs7895709	-7.07	2.24	2.49E-02	0.35
CDKL1	rs2153553	-8.81	-2.24	2.49E-02	0.05
PREP	rs717324	-5.2	-2.24	2.49E-02	0.13
TCTN2	rs3704	-9.44	-2.24	2.52E-02	0.07
PITPNM3	rs11653722	-6.98	2.24	2.53E-02	0.13
N6AMT2	rs9579937	4.77	-2.24	2.53E-02	0.08
LOXL2	rs4327870	3.94	2.24	2.53E-02	0.01
RIBC2	rs2272804	-14.14	2.23	2.56E-02	0.10
PRPSAP2	rs8067882	-10.83	-2.23	2.57E-02	0.15
RPL13A	rs4802612	5.57	-2.23	2.57E-02	0.15
HTR6	rs10159394	5.34	2.23	2.58E-02	0.14
PLCH1	rs13100169	8.06	-2.22	2.62E-02	0.06
PRODH	rs9604911	9.26	-2.22	2.65E-02	0.10
TMEM63A	rs360103	-7.12	2.21	2.70E-02	0.14
RPAIN	rs8072363	8.21	-2.21	2.70E-02	0.14
MRPL53	rs1859734	-8	2.21	2.71E-02	0.14
LDHAL6A	rs7938977	9	2.21	2.72E-02	0.11
ENAH	rs6702452	4.82	-2.20	2.75E-02	0.06
CTSC	rs217109	-5.69	2.20	2.78E-02	0.05
MXRA7	rs9303003	9.75	2.20	2.81E-02	0.10
NUDT6	rs17408757	4.82	2.20	2.82E-02	0.08
TGS1	rs765158	3.63	-2.20	2.82E-02	0.00
MTAP	rs871024	-4.05	-2.19	2.84E-02	0.03
DZIP1	rs6492817	5.73	2.19	2.85E-02	0.13
CLCN3	rs10520163	4.81	2.19	2.85E-02	0.08
CCDC122	rs12853286	8.26	2.18	2.89E-02	0.09
RMND1	rs7774446	5.77	-2.18	2.89E-02	0.06
ARHGAP29	rs1541098	5.36	2.18	2.89E-02	0.03
PARP14	rs1995993	-4.92	2.18	2.90E-02	0.09
ZNF25	rs2505202	7.14	2.18	2.91E-02	0.19
CRAT	rs3124510	-6.33	2.18	2.91E-02	0.13
SUGT1P3	rs2180960	3.49	2.18	2.93E-02	0.00
DHFR	rs2560424	-13.86	-2.18	2.93E-02	0.46
MIA3	rs3748662	-4.56	2.18	2.95E-02	0.04
ELOVL1	rs1199039	-4.4	-2.17	3.00E-02	0.12
PGAP3	rs9972882	3.53	-2.17	3.01E-02	0.16
PXMP2	rs4883545	-6.62	-2.17	3.01E-02	0.12

CENPV	rs3112518	11.43	2.17	3.02E-02	0.13
GBP3	rs17433780	-17.61	2.17	3.03E-02	0.14
C5orf34	rs6896693	4.03	-2.16	3.04E-02	0.01
CNIH4	rs6686449	-4.82	-2.16	3.05E-02	0.06
ACAT1	rs7943669	4.43	2.16	3.05E-02	0.03
ESCO1	rs2847135	9.88	-2.16	3.05E-02	0.13
PLD1	rs4586847	5.28	-2.16	3.06E-02	0.12
PLCL1	rs2054125	-3.59	2.16	3.07E-02	0.02
LRIG2	rs946220	-3.84	-2.16	3.07E-02	0.07
PCMTD1	rs6999321	-4.62	2.15	3.12E-02	0.10
CTSK	rs11576175	-6.79	2.15	3.15E-02	0.09
MLC1	rs6010138	-4.15	2.15	3.16E-02	0.02
MTFP1	rs4820013	6.14	2.15	3.16E-02	0.06
GDI2	rs2797496	-11.36	-2.15	3.16E-02	0.15
SGCB	rs7440594	3.68	2.15	3.18E-02	0.00
RERG	rs7131767	5.9	2.15	3.18E-02	0.08
LPIN3	rs6129815	4.89	-2.14	3.24E-02	0.07
IKBKE	rs1953090	-8.52	2.14	3.25E-02	0.13
CCDC62	rs1696320	-5.65	-2.14	3.28E-02	0.06
TOP2B	rs2037860	-3.92	2.13	3.29E-02	0.02
TIE1	rs3768046	-5.84	-2.13	3.33E-02	0.12
MDM1	rs11177160	7.09	-2.12	3.37E-02	0.12
DPY19L2	rs2429406	9.42	-2.12	3.39E-02	0.05
RAPGEFL1	rs4566234	6.5	2.12	3.42E-02	0.04
IARS	rs17677599	-4.09	2.11	3.47E-02	0.01
HTR7P1	rs3825272	-11.1	2.10	3.53E-02	0.06
KIAA0226L	rs259734	-6.49	-2.10	3.54E-02	0.05
ANKRD29	rs1788758	-6.25	2.10	3.60E-02	0.12
SUDS3	rs7974630	6.19	-2.10	3.60E-02	0.25
ZC3H12D	rs421453	5.36	-2.10	3.61E-02	0.16
UST	rs9322158	5.97	2.10	3.61E-02	0.29
LAMTOR3	rs11944405	-11.82	-2.09	3.64E-02	0.11
AIFM2	rs15801	-11.02	-2.09	3.67E-02	0.11
LOC220729	rs660325	3.6	2.09	3.70E-02	0.01
ITGA2	rs27915	6.74	2.08	3.71E-02	0.07
ADORA1	rs3766558	4.19	-2.08	3.75E-02	0.03
VOPP1	rs4948036	-5.48	-2.08	3.75E-02	0.04
ZDHHC14	rs9355674	-4.78	2.07	3.83E-02	0.03
RNF215	rs2072157	3.79	-2.07	3.84E-02	0.01
SLC40A1	rs1123110	9.57	-2.06	3.90E-02	0.29
RBBP8	rs10853510	-3.18	-2.06	3.93E-02	0.00
ABCA11P	rs6810900	-6.9	2.06	3.95E-02	0.05
C19orf54	rs2369302	4.51	-2.06	3.96E-02	0.03
D2HGDH	rs7563672	-8.05	-2.06	3.97E-02	0.11
DHRS13	rs17794574	-3.95	2.06	3.97E-02	0.03
FBXO10	rs12237351	-4.39	-2.06	3.97E-02	0.04

BCL2L2	rs8020117	4.52	-2.05	4.01E-02	0.04
FHOD3	rs1893434	-3.93	-2.05	4.04E-02	0.01
C1orf52	rs11161570	7.66	2.05	4.05E-02	0.10
TRIM47	rs9908862	-8.21	2.05	4.05E-02	0.09
RRP7A	rs4820470	-7.12	2.05	4.08E-02	0.04
UBE2Z	rs15563	-6.24	2.05	4.09E-02	0.11
DCAKD	rs7225329	-9.65	2.04	4.10E-02	0.10
HMGA1P7	rs11970500	-7.61	-2.04	4.10E-02	0.11
AMACR	rs10941112	-16.29	2.04	4.12E-02	0.13
LARP7	rs1486865	-5.96	2.04	4.13E-02	0.11
C10orf90	rs11245016	7.03	-2.04	4.16E-02	0.10
COQ7	rs4630530	5.35	-2.04	4.18E-02	0.02
DCXR	rs4969481	6.72	2.03	4.21E-02	0.18
ADCY10	rs7516279	6.43	2.03	4.25E-02	0.04
PDE1B	rs10783631	-3.81	-2.03	4.25E-02	0.01
SCP2	rs11206043	-5.65	-2.03	4.28E-02	0.10
EPS8L2	rs4078520	6.99	2.03	4.28E-02	0.02
TRIM5	rs3802981	-4.77	-2.03	4.28E-02	0.21
RNF40	rs7187359	8.47	2.02	4.31E-02	0.10
MAPKBP1	rs1918310	-6.6	-2.02	4.32E-02	0.07
ZNF543	rs7250371	4.3	2.02	4.36E-02	0.01
ANKS1A	rs820089	7.67	2.02	4.36E-02	0.10
ABAT	rs3760024	-4.78	2.02	4.37E-02	0.04
EHD3	rs617163	6.46	2.02	4.37E-02	0.37
ATAD3C	rs1695824	-9.35	-2.01	4.41E-02	0.05
KLC3	rs12981382	9.05	2.01	4.41E-02	0.05
NODAL	rs4746035	4.49	-2.01	4.46E-02	0.05
SLC10A5	rs1559950	-3.81	2.01	4.49E-02	0.00
HMG20A	rs11632999	5.28	-2.00	4.51E-02	0.22
STAT6	rs12307379	6.29	-2.00	4.52E-02	0.14
ZNF646	rs749671	-5.13	2.00	4.53E-02	0.08
ZNF534	rs2290747	8.14	2.00	4.54E-02	0.07
FAAH	rs6662982	-5.19	2.00	4.56E-02	0.13
FDX1	rs17111255	6.22	2.00	4.59E-02	0.10
IGFBP2	rs284526	6.19	-2.00	4.60E-02	0.13
MYSM1	rs232851	15.19	-1.99	4.61E-02	0.06
TRMU	rs11090872	3.42	-1.99	4.62E-02	0.01
AP1M1	rs12461400	-6.34	1.99	4.67E-02	0.10
ANKRD18A	rs10758513	-8.52	1.99	4.69E-02	0.06
PRKD3	rs6730503	4.64	-1.99	4.70E-02	0.04
TNKS2	rs3802650	-9.27	1.98	4.73E-02	0.08
PGM1	rs3790861	-5.06	1.98	4.73E-02	0.03
POLE	rs10870483	16.19	1.98	4.74E-02	0.09
ZFAND1	rs1844358	-9.14	1.98	4.74E-02	0.05
CDH10	rs4290981	-4.7	-1.98	4.78E-02	0.04
MLF1	rs11924696	3.98	-1.98	4.79E-02	0.03

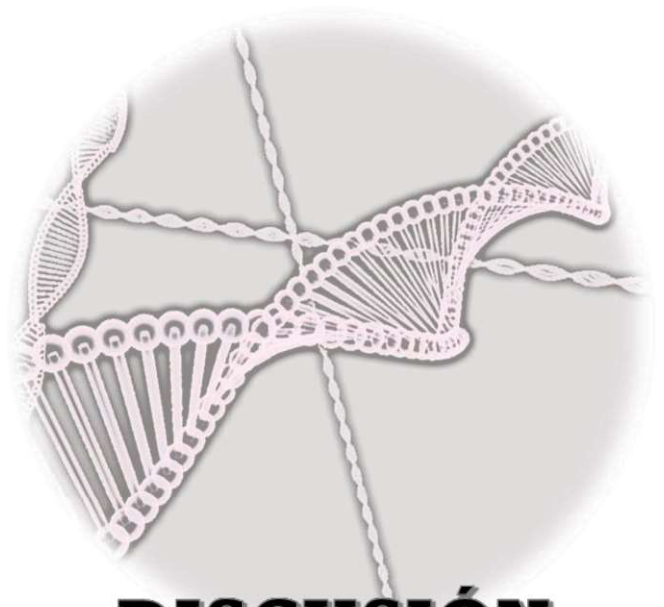
LOC100506710	rs2988011	5.08	-1.98	4.82E-02	0.08
RPGRIP1L	rs8050354	10.9	-1.97	4.86E-02	0.15
PNMAL2	rs4803967	-6.32	-1.97	4.90E-02	0.15
C17orf97	rs11150882	13.68	-1.97	4.91E-02	0.13
CHD1L	rs12726206	6.32	-1.97	4.94E-02	0.47
LRRC38	rs3795750	-4.84	1.96	4.99E-02	0.04

Table S6: GWAS results in THEM4

MarkerName	CHR	BP	A1	A2	Effect	P	A1	A2	Effect	P
rs2932586	1	151922808	t	c	0.10	1.96E-04	t	c	0.34	2.32E-02
rs3007691	1	151927901	t	c	0.09	3.22E-04	t	c	0.31	3.16E-02
rs12116923	1	151919326	a	g	0.09	3.91E-04	a	g	0.34	2.32E-02
rs13320	1	151846517	t	c	0.09	4.90E-04	t	c	0.33	1.91E-02
rs2279502	1	151917652	a	g	0.09	5.57E-04	a	g	0.34	2.32E-02
rs2999508	1	151934447	a	c	-0.09	5.96E-04	a	c	-0.31	3.16E-02
rs1038747	1	151942345	t	g	0.09	6.05E-04	t	g	0.31	3.16E-02
rs2999512	1	151946092	a	g	0.09	7.50E-04	a	g	0.31	3.16E-02
rs6685187	1	151858888	a	g	-0.08	9.78E-04	a	g	-0.33	1.91E-02
rs9943281	1	151878223	t	g	0.08	1.09E-03	t	g	0.33	2.20E-02
rs12067594	1	151865780	a	g	0.08	1.10E-03	a	g	0.33	1.91E-02
rs11204915	1	151882826	t	c	0.08	1.15E-03	t	c	0.33	2.20E-02
rs6698740	1	151859050	a	g	0.08	1.19E-03	a	g	0.33	1.91E-02
rs1060870	1	151862661	a	g	0.08	1.29E-03	a	g	0.33	1.91E-02
rs11204914	1	151881093	a	g	0.08	1.33E-03	a	g	0.33	2.20E-02
kgp12473722	1	151574295	a	g	0.55	1.57E-03	a	g	-0.87	3.82E-03
rs7542137	1	151868892	t	c	0.08	1.71E-03	t	c	0.33	1.91E-02
rs1602251	1	151903671	a	g	-0.08	1.72E-03	a	g	-0.29	4.69E-02
rs6693396	1	151930673	t	g	0.09	1.83E-03	t	g	0.32	4.15E-02
rs2337685	1	151923919	t	c	-0.08	1.98E-03	t	c	-0.35	2.95E-02
rs6682605	1	151926457	t	c	-0.08	2.01E-03	t	c	-0.32	4.15E-02
rs6693388	1	151930645	t	c	0.08	2.02E-03	t	c	0.32	4.15E-02
rs2337686	1	151925941	a	c	0.08	2.03E-03	a	c	0.32	4.15E-02
rs6587634	1	151929022	t	g	-0.08	2.03E-03	t	g	-0.32	4.15E-02
rs6684114	1	151928715	a	g	0.08	2.03E-03	a	g	0.32	4.15E-02
rs57785325	1	151928471	a	g	0.08	2.03E-03	a	g	0.32	4.15E-02
rs6693901	1	151925192	t	c	0.08	2.05E-03	t	c	0.32	4.15E-02
rs6675930	1	151924753	t	c	-0.08	2.05E-03	t	c	-0.32	4.15E-02
rs16833670	1	151922513	a	g	0.08	2.07E-03	a	g	0.35	2.95E-02
rs61817724	1	151921111	a	c	0.08	2.09E-03	a	c	0.35	2.95E-02
rs7538652	1	151920040	t	g	-0.08	2.16E-03	t	g	-0.35	2.95E-02
rs2999534	1	151976070	t	c	0.08	2.33E-03	t	c	0.30	3.53E-02
rs10788818	1	151892199	t	c	-0.08	2.36E-03	t	c	-0.28	4.73E-02

rs9943251	1	151917115	a	g	-0.08	2.42E-03	a	g	-0.35	2.95E-02
rs6587638	1	151938138	a	g	-0.08	3.05E-03	a	g	-0.32	4.15E-02
rs16833668	1	151918164	a	c	-0.08	3.15E-03	a	c	-0.35	2.95E-02
rs2130525	1	151902101	t	c	-0.08	3.19E-03	t	c	-0.33	2.20E-02
rs6703465	1	151939567	a	g	0.08	3.31E-03	a	g	0.32	4.15E-02
rs2338209	1	151903823	a	g	-0.08	3.33E-03	a	g	-0.29	4.69E-02
rs6662611	1	151936485	a	g	0.08	3.35E-03	a	g	0.32	4.15E-02
rs6587635	1	151938055	t	c	0.08	3.36E-03	t	c	0.32	4.15E-02
rs6587637	1	151938131	a	g	0.08	3.37E-03	a	g	0.32	4.15E-02
rs6587636	1	151938080	a	g	0.08	3.37E-03	a	g	0.32	4.15E-02
rs971887	1	151943497	a	g	-0.08	3.38E-03	a	g	-0.32	4.15E-02
rs2999544	1	151903634	a	g	0.08	3.40E-03	a	g	0.29	4.69E-02
rs11204919	1	151933799	t	g	0.08	3.43E-03	t	g	0.32	4.15E-02
rs7541099	1	151938763	t	c	-0.08	3.58E-03	t	c	-0.32	4.15E-02
rs1907040	1	151901940	a	g	0.08	3.75E-03	a	g	0.29	4.69E-02
rs6662728	1	151936578	a	g	0.08	3.76E-03	a	g	0.32	4.15E-02
rs28415528	1	151879761	a	g	-0.08	4.08E-03	a	g	-0.36	2.15E-02
rs10788820	1	151896233	t	c	0.07	4.40E-03	t	c	0.28	4.73E-02
rs3007685	1	151898037	t	g	-0.07	4.67E-03	t	g	-0.28	4.73E-02
rs3007670	1	151897616	t	g	-0.07	4.69E-03	t	g	-0.28	4.73E-02
rs3007711	1	151897377	t	g	0.07	4.70E-03	t	g	0.28	4.73E-02
rs4845713	1	151894887	a	g	0.07	4.72E-03	a	g	0.28	4.73E-02
rs4845409	1	151894379	t	c	0.07	5.19E-03	t	c	0.28	4.73E-02
rs9435964	1	151901342	a	g	0.07	5.46E-03	a	g	0.29	4.69E-02
rs11204920	1	151933962	t	c	0.08	5.49E-03	t	c	0.32	4.15E-02
rs6672064	1	151884499	a	g	0.08	6.08E-03	a	g	0.31	4.88E-02
rs7541453	1	151838092	t	c	0.09	6.24E-03	t	c	0.32	2.08E-02
rs56313205	1	151904109	t	c	-0.08	6.93E-03	t	c	-0.31	4.88E-02
rs6701374	1	151903858	a	g	0.08	7.14E-03	a	g	0.31	4.88E-02
rs58087484	1	151879939	t	c	-0.07	7.55E-03	t	c	-0.31	4.88E-02
rs2999542	1	151899559	t	c	-0.07	7.62E-03	t	c	-0.29	4.69E-02
rs6657658	1	151939864	t	c	0.07	7.65E-03	t	c	0.32	4.15E-02
rs2338206	1	151892226	a	g	0.07	8.48E-03	a	g	0.31	4.88E-02
rs56170235	1	151872754	a	c	-0.07	9.40E-03	a	c	-0.31	4.88E-02

rs57674198	1	151876073	t	c	0.07	9.59E-03	t	c	0.31	4.88E-02
rs61817695	1	151879676	t	g	-0.07	9.74E-03	t	g	-0.31	4.88E-02
rs6684312	1	151844650	a	g	-0.07	1.04E-02	a	g	-0.31	4.40E-02
rs6690449	1	151859166	t	c	0.07	1.23E-02	t	c	0.31	4.40E-02
rs56000812	1	151899104	a	g	0.07	1.26E-02	a	g	0.31	4.88E-02
rs56078223	1	151898843	t	g	0.07	1.26E-02	t	g	0.31	4.88E-02
rs61817697	1	151890958	a	c	0.07	1.31E-02	a	c	0.31	4.88E-02
rs6703772	1	151939856	a	g	0.07	2.19E-02	a	g	0.32	4.15E-02
rs12566232	1	151839838	a	c	0.07	3.93E-02	a	c	0.32	4.33E-02
rs998454	1	151817998	t	c	-0.07	3.96E-02	t	c	-0.41	2.53E-02



DISCUSIÓN

DISCUSIÓN

Los estudios que componen esta tesis se plantearon con el objetivo de dar respuestas a aspectos de salud que tienen un gran impacto en nuestra sociedad como son el riesgo de sufrir un ictus y la posterior evolución neurológica, ya que esta se asocia al estado funcional a largo plazo. Debido al envejecimiento de la sociedad, el número de pacientes con ictus es cada vez mayor (2,6). Tanto por su alta incidencia, como por la gran carga personal, familiar y económica que representa a medio-largo plazo en nuestro sistema de salud, es importante encontrar los mecanismos y los factores que puedan ayudar a predecir el riesgo de sufrir un ictus, así como mejorar el pronóstico, minimizar la discapacidad de los pacientes y reducir su impacto social.

El ictus es una enfermedad compleja en la que su riesgo se explica, en parte, por factores demográficos, clínicos y genéticos. Sin embargo, estos factores no definen toda su heredabilidad. Es importante explorar otros factores heredables, como la epigenética, que explican parte del componente hereditario de las enfermedades complejas.

Por todo ello, con estos objetivos, los estudios de esta tesis fueron diseñados para dar respuesta a estos problemas a través de las tecnologías *ómicas*. Principalmente, de la epigenómica, aunque también haciendo uso de datos genómicos, proteómicos y análisis bioinformáticos, para intentar encontrar nuevos mecanismos asociados al riesgo y al pronóstico tras un ictus.

Nuestros estudios han demostrado que las modificaciones epigenéticas, a través de la metilación del ADN, se asocian tanto al riesgo como a la progresión después de un ictus. Las principales asociaciones de la metilación con el riesgo de ictus identificadas son de tipo causal, por lo que primero se encuentran las modificaciones en la metilación que posteriormente modifican el riesgo de tener un ictus.

Las asociaciones identificadas parecen mostrar una implicación de la metilación en la regulación de genes que es más relevante en células NK, importantes en procesos inflamatorios como los que se producen después de una isquemia cerebral.

También hemos identificado biomarcadores proteicos cuya regulación genética se asocia al riesgo de tener un ictus de tipo lacunar así como de padecer una hemorragia

intracerebral. Además, también hemos identificado una regulación genética a nivel de proteína asociada con el pronóstico funcional a largo plazo. Debido al carácter modificable de las proteínas, el hecho de encontrar una regulación a este nivel puede conllevar su estudio como dianas para futuros fármacos.

Nuestros resultados tanto de epigenética como de integración genómica con proteómica indican que algunas vías previamente conocidas asociadas al ictus, como la inflamación, la apoptosis o la excitotoxicidad, se encuentran reguladas a través de modificaciones epigenéticas o genéticas.

El hallazgo de factores asociados al riesgo y al pronóstico del ictus aumenta el conocimiento acerca de los mecanismos implicados pero también puede permitir el desarrollo de terapias o fármacos para la prevención precoz de la enfermedad y la mejora de los síntomas neurológicos y alteraciones funcionales de las personas que han sufrido un ictus.

Con el objetivo de analizar si la metilación puede ser un factor causal asociado al riesgo de ictus, realizamos un estudio EWAS en dos fases, el cual nos permitió identificar y replicar un total de 821 CpGs asociadas al riesgo de ictus. Se sabía, gracias a estudios previos, que la metilación se asocia al riesgo de ictus (164–166). Sin embargo, estos nuevos resultados han permitido expandir el conocimiento que se tenía acerca del papel de la metilación en esta enfermedad. Se ha aumentado el número de loci significativos asociados al riesgo de ictus de 25 (164–166) hasta 836. De los cinco loci (siete genes) identificados muy recientemente en población del norte de China por Sun et al., (166) ninguno fue significativo en nuestra cohorte. Estas diferencias pueden ser debidas a la diferente población analizada y al uso de una metodología alternativa. En cambio, la metilación diferencial encontrada por Soriano-Tárraga et al., en 20 loci (164), también fue identificada en nuestros análisis. El hecho de haber identificado en nuestro estudio muchos más loci asociados al riesgo del ictus puede ser, en parte, gracias a un mayor tamaño muestral (considerando la cohorte de descubrimiento y de replicación), al poder estadístico superior al 0.8 (calculado específicamente para estudios EWAS) y al uso de una metodología reciente diseñada para realizar análisis EWAS. El paquete de R *Meffil*, utilizado para realizar los análisis, efectúa una estimación automática de los parámetros

de la normalización que permite reducir la variación técnica en los niveles de metilación y así disminuir el número de falsos positivos, incrementando el poder estadístico (225).

Aunque la metodología de los EWAS es una buena aproximación que permite encontrar puntos de metilación (CpGs) distribuidos a lo largo del genoma y regiones asociadas a un fenotipo, presenta ciertos problemas relacionados con las características intrínsecas de la epigenética. Por un lado, la metilación del ADN es específica de tejido. Al utilizar un tejido como el sanguíneo para realizar el estudio EWAS, como en nuestro caso, existe una mezcla de diferentes tipos celulares. Por lo tanto, las asociaciones encontradas pueden producirse por la diferente metilación en un tipo celular concreto. Para evitar que este problema afectara a nuestros resultados, por un lado, se realizó un ajuste por conteo celular estimado en nuestras muestras. Por otro lado, aplicamos la metodología publicada por Zheng et al., (111) para identificar si existía un predominio de metilación diferencial en algún tipo celular. Los resultados, aplicados por primera vez en un estudio EWAS de ictus, indicaron que las células NK estaban mediando parte de la asociación que encontramos en el EWAS. Más de dos tercios de las asociaciones epigenéticas identificadas en nuestro estudio se asociaron al riesgo de ictus en este tipo celular concreto. El papel de estas células NK en el ictus se ha descrito previamente y está relacionado con procesos de inflamación, inmunodepresión e infecciones post-ictus (226,227).

Además, la epigenética es variable en el tiempo, pudiéndose producir antes o después del ictus. Por lo tanto, para poder estudiar con más profundidad qué tipo de relación puede tener la metilación en el ictus, realizamos un análisis de causalidad mediante RM de los dos genes que encontramos más significativos en el EWAS. La RM es una metodología con un uso muy extendido actualmente ya que permite establecer relaciones de causalidad entre dos factores sin la necesidad de llevar a cabo un ensayo clínico (20). Además, puede ser una herramienta útil a la hora de encontrar fármacos que puedan ser eficaces en el tratamiento de una enfermedad (20). Mediante el uso de información genética disponible, se aplicó la RM por primera vez en un estudio de EWAS en ictus y permitió identificar una relación de causalidad entre la metilación en *MAP3K1* y el riesgo de ictus aterotrombótico y cardioembólico y la metilación en *ZFX3* y el riesgo de ictus aterotrombótico, cardioembólico y lacunar.

MAP3K1 y *ZFHX3* son los dos genes cuya metilación se asoció de forma más significativa con el riesgo de ictus en nuestros análisis. El gen *MAP3K1* pertenece a la vía de señalización de las MAPK. Este gen y otros miembros de la misma vía metabólica se identificaron como significativos en nuestros análisis. Los miembros de la vía de señalización de las MAPK responden a diferentes estímulos como pueden ser la isquemia o la inflamación. Además, algunos miembros significativos en nuestro estudio pertenecientes a esta vía, como *MAPK1*, son diana de fármacos aprobados o en estudio para el ictus, como es el caso del ácido acetilsalicílico o la minociclina. El estudio en más profundidad del papel de la metilación en estos genes puede ser importante para plantear un análisis de reposicionamiento farmacológico.

El gen *ZFHX3* codifica un factor de transcripción implicado en la diferenciación neuronal. En nuestro estudio, a parte de la asociación de la metilación en este gen con el cómputo de todos los subtipos de ictus, también se asoció específicamente al riesgo de ictus de tipo cardioembólico. El gen *ZFHX3* ya se había asociado previamente al riesgo de este subtipo de ictus en estudios genéticos mediante GWAS (72), y a ictus isquémico en estudios epigenéticos con EWAS (164). Estos resultados, junto con los obtenidos de la RM, están en concordancia con los resultados de genética publicados (72) en los que se ha visto una asociación de *ZFHX3* con el riesgo de ictus cardioembólico.

Al encontrar que podía existir una correlación entre la genética y epigenética en ciertos genes en asociación al riesgo de ictus, también analizamos si la metilación de otros genes asociados al riesgo de ictus en el GWAS del estudio MEGASTROKE (72) (en el que se incluía un mayor tamaño muestral hasta el momento de la publicación de nuestro artículo) también tenían una asociación epigenética con el riesgo de ictus. Un total de cinco genes (*SH2B3*, *SMARCA4*, *TSPAN2*, *ILF3* y *CDK6*), aparte de *ZFHX3*, asociados previamente a ictus mediante GWAS, también se asociaron al riesgo de ictus en nuestro estudio epigenético. Posteriormente, hemos analizado los nuevos genes asociados al riesgo de ictus en el estudio GWAS publicado recientemente, después de la publicación del artículo de esta tesis, por parte del consorcio GIGASTROKE (75). Hemos encontrado cuatro nuevos genes (*GRK5*, *PTPN11*, *LAMC1* y *MESDC1*) cuya metilación se asocia también al riesgo de ictus de forma estadísticamente significativa en nuestro estudio. Además, en el caso del gen *MESDC1*, un total de 3 CpGs están asociadas ($p\text{-value} < 10^{-6}$).

⁰⁷), haciendo que a nivel de región también se encuentre una asociación. Nuestros resultados permiten reforzar el papel que estos genes tienen en el ictus, cuya asociación viene determinada de forma tanto genética como epigenética.

Aunque el estudio en detalle de los genes identificados como más significativos es importante, el hecho de haber obtenido un número tan grande de resultados significativos hace que sea muy difícil el estudio de cada uno de los genes de forma independiente. Por ello, es importante el uso de estrategias que permitan unificar los diferentes resultados para encontrar características comunes. En nuestro caso, realizamos un análisis de enriquecimiento de localización de la metilación dentro de los genes. La localización de la metilación dentro del genoma es importante para poder predecir cómo pueden estar afectando estas modificaciones a la expresión génica. En nuestro caso, identificamos un enriquecimiento de la regulación de la metilación en el primer exón de los genes asociado con el riesgo de ictus. En general, la metilación en el primer exón se ha asociado a una disminución de la expresión génica, lo que nos permitiría pensar en nuestro caso que la metilación disminuida en casos comparado con los controles podría estar relacionado con un aumento de la expresión de genes en el ictus (228).

Por otro lado, realizamos un estudio de vías metabólicas con el mismo objetivo de unificar los resultados. Encontramos un enriquecimiento de genes pertenecientes a la vía de la fagocitosis. La regulación de este proceso es importante para dar respuesta a estímulos como la isquemia y poder activar procesos de reparación del tejido dañado (229). Además, otras vías estrechamente relacionadas con procesos isquémicos en el ictus, como es el caso de la respuesta al estrés oxidativo (230) o la regulación del *epidermal growth factor* (231), también fueron significativos en nuestro análisis.

La progresión de los déficits neurológicos y funcionales en el ictus presentan una variación interindividual muy importante. Es relevante identificar los factores que se asocian a la progresión para poder administrar las mejores terapias de forma individualizada. A pesar de que se han descrito diferentes factores, tanto demográficos,

clínicos y genéticos que están asociados (29–31), sigue faltando conocimiento en esta área para poder aplicarse de forma rutinaria en la práctica clínica habitual.

En el manuscrito 2 se realizó un estudio de metilación del ADN de genoma completo (EWAS) en tres fases (descubrimiento, replicación y metaanálisis). Para cuantificar la evolución neurológica utilizamos como variable la diferencia entre la NIHSS basal y la NIHSS al alta (Δ NIHSS). La NIHSS al alta es una buena predictora de la progresión de los pacientes a los tres meses después del ictus (232), pero se ha visto que durante la estancia hospitalaria se producen variaciones en la NIHSS. Por lo tanto, la diferencia entre la NIHSS basal y la NIHSS al alta, la Δ NIHSS, es una variable más dinámica que permite capturar los diferentes procesos relacionados con la recuperación de la isquemia que se producen en el ictus (233). En nuestro estudio, la Δ NIHSS se asoció de forma inversa e independiente al mRS a los 90 días después de un ictus. De esta manera, los pacientes con mejoría durante la estancia hospitalaria (Δ NIHSS > 0) presentaron un mRS menor (más independencia funcional) a los tres meses.

La aproximación realizada en este estudio permitió identificar la metilación de un gen (*EXOC4*) asociada al pronóstico neurológico después de un ictus. El gen *EXOC4* se expresa en cerebro, especialmente en conos de crecimiento de axones y ramas dendríticas, y pertenece al complejo de exocitosis, implicado en la secreción de vesículas hacia la membrana plasmática (234,235). En las neuronas es importante el transporte hacia la membrana plasmática de proteínas que participan en la sinapsis. Para ello, la célula utiliza el complejo de exocitosis para poder formar las vesículas cuyo contenido será liberado en la membrana plasmática. En el caso del gen *EXOC4*, que codifica por la subunidad Sec8 del complejo de exocitosis, se ha visto que está implicado en el transporte de receptores de glutamato (NMDAR y AMPAR) a la membrana post-sináptica (236,237). El papel del glutamato en el ictus está ampliamente descrito ya que participa en el proceso conocido como excitotoxicidad. Después de un ictus, el proceso isquémico favorece la liberación de glutamato, el cual se une a sus receptores post-sinápticos (238). Según el receptor al que se une se genera una respuesta de señalización de supervivencia o de muerte celular (239,240). Según indican nuestros resultados, una mayor metilación del gen parece asociarse con una disminución de su expresión y a su vez con una mejoría neurológica de los pacientes. Además, los resultados de este

estudio también indican que la metilación diferencial de *EXOC4* promueve cambios en los niveles de expresión de diferentes proteínas. Todos estos cambios, junto con la función de *EXOC4* en las neuronas, sugieren que su disminución en la expresión debido a una mayor metilación podría estar relacionada con cambios en la movilización de los receptores de glutamato, conllevando a cambios en la supervivencia neuronal.

Para poder comprobar si a parte de los resultados obtenidos en *EXOC4* había otros genes asociados de forma nominal a la Δ NIHSS en nuestro EWAS que estuvieran también relacionados con la vía de exocitosis, hicimos un análisis de enriquecimiento de vías metabólicas. Esta aproximación nos permitió ver un enriquecimiento de las vías de “procesos de exocitosis” y “transporte de vesículas a la membrana celular”. Como se ha comentado, los procesos de exocitosis y transporte a la membrana están estrechamente relacionados con la excitotoxicidad. Además, otro gen perteneciente a la vía de excitotoxicidad, el *GRM3*, previamente asociado a deterioro de la memoria en la enfermedad de Alzheimer (241), se asoció de forma nominal a la Δ NIHSS.

Los diferentes resultados obtenidos en este análisis parecen apoyar el papel de la metilación en la regulación del pronóstico post-ictus mediante la regulación de procesos de excitotoxicidad. Además, uno de los pocos estudios GWAS publicados en relación al pronóstico del ictus, identificó también este proceso como clave. En concreto, se identificaron ocho loci asociados a la diferencia entre la NIHSS basal y la NIHSS a las 24h. La anotación funcional de estos loci identificó a dos genes, *GRIA1* y *ADAM23*, ambos pertenecientes a la vía de la excitotoxicidad (82).

La principal limitación de este estudio es que se realizó el análisis de metilación en el tejido que teníamos disponible para los pacientes, que fue el tejido sanguíneo. Sin embargo, los resultados obtenidos en tejido sanguíneo pueden ser más útiles como biomarcadores. Además, para evitar sesgos, realizamos un estudio de metilación diferencial de tipos celulares. Las células NK fueron las que mostraron una tendencia más significativa a ser las responsables de las asociaciones observadas en el EWAS. Además, el estudio de vías metabólicas de las proteínas diferencialmente expresadas en relación a la metilación de *EXOC4* también mostró la vía de activación de las NK como diferencialmente regulada por metilación en el pronóstico post-ictus.

Los datos *ómicos* permiten obtener una gran cantidad de información que ha ayudado en muchas patologías, como en el caso del ictus, a aumentar el conocimiento sobre los mecanismos implicados en la enfermedad. Sin embargo, debido a la gran cantidad de datos obtenidos, la integración de los diferentes datos *ómicos*, aunque es compleja, ayuda a priorizar los genes causales de las asociaciones y dilucidar aquellos con una implicación en la enfermedad. En el caso de los PWAS, estos permiten la integración de datos genómicos (obtenidos mediante GWAS) con datos proteómicos. De esta manera, se puede saber si existe una regulación de tipo cis entre el gen y los niveles de proteína, que a su vez se asocie con un fenotipo. En este tipo de regulación, los polimorfismos de un gen modifican la expresión de la proteína codificada por el mismo gen.

En nuestro caso, realizamos un PWAS con datos genómicos de ictus de tipo lacunar y datos proteómicos de tejido cerebral. Además, también se realizó el análisis PWAS con datos de HIC, y comparamos ambos análisis. Los dos tipos de ictus comparten mecanismos fisiopatológicos, principalmente la arteriosclerosis y la angiopatía cerebral amiloide, y son manifestaciones de la enfermedad de pequeño vaso. Por eso, en estudios previos de GWAS, los dos fenotipos ya habían sido comparados e incluso analizados conjuntamente utilizando la aproximación de MTAG (77). Estos estudios habían identificado algunos genes comunes como factores de riesgo en ambas patologías, e incluso una regulación de polimorfismos en estos genes como eQTLs a nivel de expresión en cis. Sin embargo, no se había analizado previamente si estos genes podían estar regulando también los niveles de proteínas y de esta manera estar contribuyendo a aumentar o disminuir el riesgo tanto de ictus lacunar como de HIC.

Por este motivo, primero se realizó el PWAS en estos dos subtipos de ictus utilizando datos proteómicos de cerebro (del córtex prefrontal dorsolateral). Los resultados mostraron que el gen *ICA1L* se asociaba a ictus lacunar e HIC de tipo no lobar a través de la regulación de los niveles proteicos. La asociación con el ictus lacunar e HIC a nivel genético se había encontrado previamente. También se había visto, mediante TWAS, que la expresión de *ICA1L* en arterias y cerebro se asociaba de forma causal con el riesgo de ictus lacunar y con la presencia de hiperintensidades de sustancia blanca, un fenotipo común en el ictus lacunar.

El gen *ICA1L* (*Islet Cell Autoantigen 1 Like*) se encuentra regulando actividades de unión de proteínas y regula procesos de transporte en la célula. Además, el cerebro es el tejido donde la proteína codificada por este gen tiene mayor expresión (242). Específicamente, se ha visto una expresión de *ICA1L* en neuronas excitatorias glutamatérgicas. Como se ha comentado anteriormente, estas neuronas participan en los procesos de excitotoxicidad en respuesta al estímulo isquémico provocado por un ictus (243–245).

En un estudio PWAS realizado en la enfermedad de Alzheimer, el gen *ICA1L* se había encontrado aumentando el riesgo de padecer la enfermedad, mediante una disminución en la abundancia de la proteína (243,244). En nuestro estudio, los resultados obtenidos son muy similares. Los polimorfismos en *ICA1L* parecen estar determinando una mayor abundancia de la proteína y esto reduce el riesgo de HIC no lobar e ictus lacunar. La fisiopatología en la enfermedad de Alzheimer está estrechamente relacionada con la de HIC e ictus lacunar, ya que tanto la angiopatía amiloide cerebral como la arteriolosclerosis son comunes en ambas y, de hecho, empeoran la demencia en el Alzheimer. Además, en ambos casos, procesos como la rotura de la barrera hematoencefálica favorecen el desarrollo de la demencia (246). Genes como *APOE* son el mayor componente genético asociado tanto a HIC como a Alzheimer (247). Por lo tanto, estos resultados parecen indicar que la regulación de los niveles de la proteína *ICA1L* a través de la genética pueden ser un mecanismo común en el ictus lacunar, HIC y en la enfermedad de Alzheimer.

Posteriormente a la publicación del artículo 3 de esta tesis, se publicaron dos nuevos estudios que han permitido corroborar nuestros resultados. Por un lado, Wu et al. (248), publicó un análisis PWAS realizado con los mismos datos analizados en nuestro estudio. Por un lado, incluyó los datos de transcriptómica del ROS/MAP y, por otro, los datos de GWAS del estudio MEGASTROKE. En su publicación se incluyeron datos de todos los tipos de ictus (isquémico y hemorrágico), de ictus isquémicos y de los diferentes subtipos etiológicos de ictus. Sin embargo, no se analizaron los ictus hemorrágicos por separado. Además, como en nuestro caso, realizaron análisis posteriores de RM y colocación para comprobar la causalidad de sus resultados. Sin embargo, su estudio no incluyó un análisis de replicación en una nueva cohorte con datos de proteómica. Entre los diferentes resultados obtenidos en su PWAS, solo pudo ser validado tanto por

RM como por colocación los resultados de ICA1L, obteniendo los mismos resultados que en nuestro estudio: la disminución de los niveles de ICA1L se asoció con un mayor riesgo de padecer ictus de tipo lacunar.

En el segundo de los artículos, Zhang et al. (249), realizó un análisis PWAS solo para el subtipo de ictus lacunar. En este caso tampoco se estudiaron los ictus hemorrágicos. En este estudio también realizaron análisis de RM y colocación e incluyeron la replicación con la cohorte de *Banner Sun Health Research Institute* con datos de proteómica tal como habíamos realizado en nuestro estudio. La diferencia en este estudio fue que incluyeron los datos del GWAS de ictus lacunares publicado por Traylor et al., (71) en lugar de los del estudio MEGASTROKE. Este GWAS se realizó exclusivamente en pacientes con este subtipo de ictus. Sin embargo, el número de muestras totales analizadas (considerando casos y controles), es inferior al del estudio MEGASTROKE (262,136 en el manuscrito de Traylor et al., vs 411,497 en el estudio MEGASTROKE). En el estudio de Zhang et al., se identificaron y validaron cuatro proteínas asociadas al riesgo de ictus lacunar, entre ellas ICA1L, con resultados similares a los obtenidos en nuestro estudio. Además, identificaron una asociación significativa para CAND2, ALDH2 y MADD. En nuestro estudio, estas tres proteínas también se asociaron con un valor de $p < 0.05$ al riesgo de ictus lacunar. Sin embargo, al realizar el ajuste por False-Discovery Rate (FDR), la asociación solo se mantuvo para ALDH2, pero con una probabilidad posterior 4 en el análisis de colocación < 0.75 , por lo que no la consideramos validada. De forma similar, en el estudio de Wu et al., también se identificó ALDH2 pero tampoco se validó por colocación. Probablemente, las diferencias en los resultados obtenidos entre el estudio de Zhang et al., comparado con el de Wu et al., y el nuestro vienen determinadas por el uso de estudios GWAS diferentes. La confirmación en un tercer GWAS sería necesaria para corroborar los nuevos resultados de proteínas asociadas a ictus lacunar. Sin embargo, es importante que dos estudios diferentes han replicado nuestros resultados obtenidos en relación a la asociación de ICA1L con el riesgo de ictus lacunar. Además, nuestro estudio permitió añadir la relevancia que ICA1L también presenta en un fenotipo correlacionado, las HIC.

Los resultados de este análisis, junto con los resultados previos a nivel genético y de transcriptómica, pueden tener una implicación futura a nivel de desarrollo de un fármaco que tenga ICA1L como diana para la enfermedad de pequeño vaso.

En el manuscrito 4 se analizó mediante PWAS el pronóstico funcional a largo plazo. En este análisis se identificaron las proteínas THEM4 y APOL2 cuyos niveles en el córtex prefrontal dorsolateral estaban regulados genéticamente en asociación con el pronóstico funcional a largo plazo. En el caso de THEM4, que fue significativo en todos los análisis (descubrimiento, replicación 1 y 2, y metaanálisis), una mayor expresión de la proteína se asoció causalmente a un peor estado funcional del paciente a los tres meses. Esta asociación de causalidad identificada en la región del córtex prefrontal dorsolateral, también la identificamos a nivel de expresión génica (mediante RM), en otras regiones como en tejido adiposo, arterial, cerebral (hemisferio cerebelar, cerebelo, hipocampo, hipotálamo y córtex), esófago, corazón e hígado.

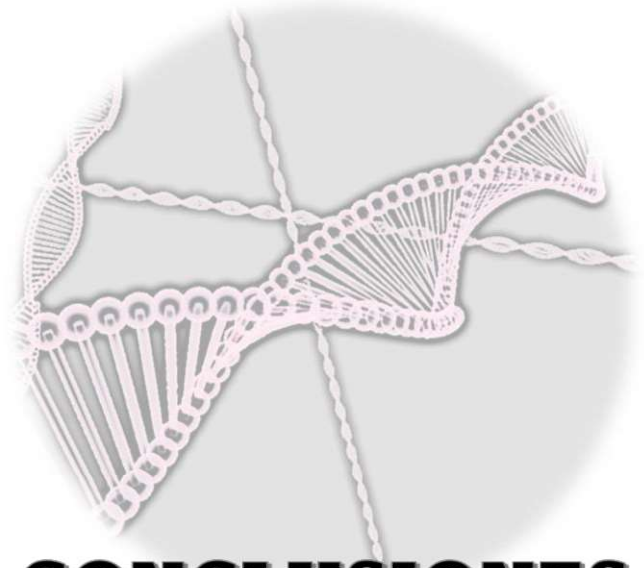
THEM4, también conocido como *Carboxy-Terminal Modulator Protein* (CTMP), es una proteína mitocondrial que se trasloca al citosol en respuesta a un estímulo isquémico como el ictus (250). En estudios *in vitro* e *in vivo* se ha encontrado que THEM4 es capaz de inhibir la función inhibidora de la apoptosis de la proteína kinasa AKT. De esta manera, ante un estímulo como el ictus, THEM4 favorece los procesos de apoptosis (251). THEM4 es diana de un anestésico volátil, el sevoflurano, utilizado en procedimientos como la trombectomía mecánica en el tratamiento del ictus. En ratas, la exposición a sevoflurano previa a la isquemia cerebral disminuye el tamaño del infarto y mejora la puntuación en las pruebas neurológicas realizadas a los animales. Estos resultados sugieren un papel neuroprotector de sevoflurano en los estudios *in vivo* (252). En humanos también se ha estudiado el efecto de los anestésicos volátiles durante diferentes procesos quirúrgicos, incluida la trombectomía mecánica. El principal objetivo de estos estudios ha sido el de comparar el uso de la anestesia general con el de la sedación consciente en relación a diferentes efectos adversos. Los resultados son controvertidos, encontrándose en algunos casos un efecto protector de los anestésicos volátiles (como el sevoflurano) en la ocurrencia de ictus como efecto adverso en diferentes procedimientos quirúrgicos (253). También se observó en pacientes participantes en el ensayo MR-CLEAN, como el uso de anestésicos volátiles se asociaba

con una mejor evolución neurológica en pacientes con ictus (254). Sin embargo, otros ensayos han identificado un efecto neutro en el pronóstico después del ictus según el uso de anestesia vs sedación (255). Sin embargo, el objetivo de estos estudios no ha tenido como enfoque estudiar el posible efecto neuroprotector de fármacos como el sevoflurano. Considerando nuestros resultados del PWAS y los resultados en modelos animales, sería interesante desarrollar estudios en humanos para poder testar el efecto de sevoflurano como fármaco neuroprotector.

La asociación de APOL2 con el pronóstico funcional de los pacientes se replicó en la primera replicación y también se identificó en el metaanálisis. El incremento en los niveles de APOL2 se encontraron asociados causalmente con un mejor pronóstico funcional a los tres meses. APOL2 pertenece a la familia de las apolipoproteínas y variantes genéticas en APOL2 se habían visto previamente asociadas al riesgo de ictus (256).

Globalmente, los resultados del metaanálisis de este análisis PWAS indican un enriquecimiento de proteínas pertenecientes a la vía de elongación de los ácidos grasos y a la regulación de AMPK. Ambas vías metabólicas tienen una estrecha relación con el ictus. Por un lado, los ácidos grasos se han visto asociado con daño cerebral en modelos animales de ictus (257) y también se ha estudiado su relación con el pronóstico en el ictus (258). La vía de AMPK, una proteína kinasa, se activa tras un ictus y se asocia con una reducción de procesos dañinos en el ictus como la neuroinflamación, el estrés oxidativo, la apoptosis o la excitotoxicidad y se ha postulado como potencial diana terapéutica (259).

La identificación de proteínas previamente estudiadas por su relación con el ictus mediante la aproximación PWAS pone de manifiesto la relevancia de la integración de diferentes tecnologías *ómicas* para identificar dianas potencialmente relevantes en enfermedades complejas como el ictus.



CONCLUSIONES

CONCLUSIONES

1-El estudio EWAS ha permitido la identificación de 821 CpG (situadas en 836 loci) asociadas al riesgo de ictus, siendo la metilación en *MAP3K1* y *ZFHX3* las más significativas. Los resultados han mostrado que la metilación identificada en este estudio no es específica de un subtipo de ictus concreto.

1.1-La regulación de la fagocitosis es la vía metabólica que se ha encontrado más enriquecida entre los genes diferencialmente metilados en relación al riesgo de ictus.

1.2-Las células NK se han identificado como las principales mediadoras de la asociación de la metilación en el riesgo de ictus en el tejido sanguíneo.

1.3-La metilación en *MAP3K1* y *ZFHX3* se relaciona causalmente con el riesgo de ictus cardioembólico y aterotrombótico y en el gen *ZFHX3* también con el ictus lacunar, según los resultados del análisis de RM.

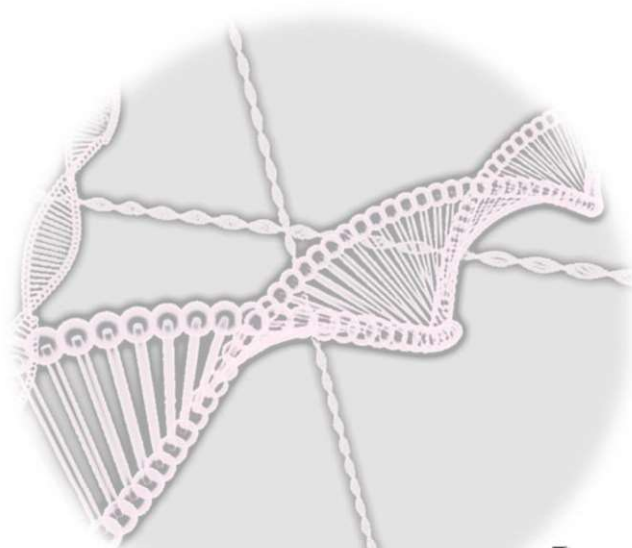
2-En el estudio EWAS, en el que se ha analizado la metilación asociada a la Δ NIHSS alta, se ha identificado una mayor metilación en la CpG cg00039070, en el gen *EXOC4*, asociada con un peor pronóstico neurológico post-ictus.

2.1-Las células NK son las mediadoras de la metilación diferencial asociada al pronóstico neurológico a medio plazo obtenido en el EWAS.

2.3-La exocitosis se identificó como la vía metabólica más enriquecida entre los genes cuya metilación se asoció a la Δ NIHSS en el metaanálisis realizado. Además, la vía de la regulación de la activación de las células NK se encontró enriquecida entre las proteínas moduladas por la metilación diferencial en *EXOC4*.

3-La integración de datos genómicos y proteómicos de cerebro en un análisis PWAS ha permitido la identificación de una regulación genética en cis de la proteína ICA1L en asociación causal con el riesgo de ictus lacunar e HIC de tipo no lobar. Se ha observado una menor abundancia de la proteína ICA1L en cerebro en asociación con un aumento del riesgo de ictus lacunar e ICH.

4-Se ha identificado una regulación genética de dos proteínas, THEM4 y APOL2, en asociación causal con el pronóstico funcional a largo plazo de pacientes que han sufrido un ictus. El aumento de la expresión de THEM4 se asocia a un peor pronóstico, mientras que el incremento de APOL2 se asocia a un mejor pronóstico funcional a los tres meses.



BIBLIOGRAFÍA

BIBLIOGRAFÍA

1. Sacco RL, Kasner SE, Broderick JP, Caplan LR, Connors JJ (Buddy), Culebras A, et al. An Updated Definition of Stroke for the 21st Century. *Stroke*. 1 de julio de 2013;44(7):2064-89.
2. Johnson CO, Nguyen M, Roth GA, Nichols E, Alam T, Abate D, et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 1 de mayo de 2019;18(5):439-58.
3. Heart Disease and Stroke Statistics—2021 Update | *Circulation* [Internet]. [citado 19 de agosto de 2021]. Disponible en: <https://www.ahajournals.org/doi/10.1161/CIR.0000000000000950>
4. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*. octubre de 2020;396(10258):1204-22.
5. Wafa HA, Wolfe CDA, Emmett E, Roth GA, Johnson CO, Wang Y. Burden of Stroke in Europe. *Stroke*. agosto de 2020;51(8):2418-27.
6. SEN. El Atlas del Ictus. España 2019. En 2019;
7. Amarenco P, Bogousslavsky J, Caplan LR, Donnan GA, Hennerici MG. Classification of stroke subtypes. *Cerebrovasc Dis Basel Switz*. 2009;27(5):493-501.
8. Morais Filho AB de, Rego TL de H, Mendonça L de L, Almeida SS de, Nóbrega ML da, Palmieri T de O, et al. The physiopathology of spontaneous hemorrhagic stroke: a systematic review. *Rev Neurosci*. 26 de agosto de 2021;32(6):631-58.
9. Boehme AK, Esenwa C, Elkind MSV. Stroke Risk Factors, Genetics, and Prevention. *Circ Res*. 3 de febrero de 2017;120(3):472-95.
10. Musuka TD, Wilton SB, Traboulsi M, Hill MD. Diagnosis and management of acute ischemic stroke: speed is critical. *CMAJ Can Med Assoc J*. 8 de septiembre de 2015;187(12):887-93.
11. Donnan GA, Fisher M, Macleod M, Davis SM. Stroke. *Lancet Lond Engl*. 10 de mayo de 2008;371(9624):1612-23.
12. Adams HP, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. enero de 1993;24(1):35-41.
13. Kleindorfer DO, Towfighi A, Chaturvedi S, Cockcroft KM, Gutierrez J, Lombardi-Hill D, et al. 2021 Guideline for the Prevention of Stroke in Patients With Stroke and Transient Ischemic Attack: A Guideline From the American Heart Association/American Stroke Association. *Stroke* [Internet]. julio de 2021 [citado 25 de abril de 2023];52(7):e364-467. Disponible en: <https://www.ahajournals.org/doi/10.1161/STR.0000000000000375>

14. Qureshi AI, Mendelow AD, Hanley DF. Intracerebral haemorrhage. *Lancet*. 9 de mayo de 2009;373(9675):1632-44.
15. Yew KS, Cheng E. Acute Stroke Diagnosis. *Am Fam Physician*. 1 de julio de 2009;80(1):33-40.
16. Kolominsky-Rabas PL, Weber M, Gefeller O, Neundoerfer B, Heuschmann PU. Epidemiology of Ischemic Stroke Subtypes According to TOAST Criteria. *Stroke* [Internet]. diciembre de 2001 [citado 25 de abril de 2023];32(12):2735-40. Disponible en: <https://www.ahajournals.org/doi/10.1161/hs1201.100209>
17. Gardener H, Sacco RL, Rundek T, Battistella V, Cheung YK, Elkind MSV. Race and Ethnic Disparities in Stroke Incidence in the Northern Manhattan Study. *Stroke* [Internet]. abril de 2020 [citado 25 de abril de 2023];51(4):1064-9. Disponible en: <https://www.ahajournals.org/doi/10.1161/STROKEAHA.119.028806>
18. Dufouil C, Beiser A, McLure LA, Wolf PA, Tzourio C, Howard VJ, et al. A Revised Framingham Stroke Risk Profile to Reflect Temporal Trends. *Circulation*. 21 de marzo de 2017;135(12):1145-59.
19. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet Lond Engl*. 10 de julio de 2010;376(9735):112-23.
20. Georgakis MK, Gill D. Mendelian Randomization Studies in Stroke: Exploration of Risk Factors and Drug Targets With Human Genetic Data. *Stroke*. agosto de 2021;52(9):2992-3003.
21. Frerich S, Malik R, Georgakis MK, Sinner MF, Kittner SJ, Mitchell BD, et al. Cardiac Risk Factors for Stroke: A Comprehensive Mendelian Randomization Study. *Stroke* [Internet]. abril de 2022 [citado 27 de julio de 2022];53(4). Disponible en: <https://www.ahajournals.org/doi/10.1161/STROKEAHA.121.036306>
22. Biddinger KJ, Emdin CA, Haas ME, Wang M, Hindy G, Ellinor PT, et al. Association of Habitual Alcohol Intake With Risk of Cardiovascular Disease. *JAMA Netw Open* [Internet]. 25 de marzo de 2022 [citado 27 de julio de 2022];5(3):e223849. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8956974/>
23. Millwood IY, Walters RG, Mei XW, Guo Y, Yang L, Bian Z, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *The Lancet* [Internet]. 4 de mayo de 2019 [citado 27 de julio de 2022];393(10183):1831-42. Disponible en: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(18\)31772-0/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)31772-0/fulltext)
24. Elkind MSV, Carty CL, O'Meara ES, Lumley T, Lefkowitz D, Kronmal RA, et al. Hospitalization for infection and risk of acute ischemic stroke: the Cardiovascular Health Study. *Stroke*. julio de 2011;42(7):1851-6.
25. Wouters A, Nysten C, Thijs V, Lemmens R. Prediction of Outcome in Patients With Acute Ischemic Stroke Based on Initial Severity and Improvement in the First 24 h. *Front Neurol*. 7 de mayo de 2018;9:308.

26. Saver JL, Altman H. RELATION BETWEEN NEUROLOGIC DEFICIT SEVERITY AND FINAL FUNCTIONAL OUTCOME SHIFTS AND STRENGTHENS DURING FIRST HOURS AFTER ONSET. *Stroke J Cereb Circ.* junio de 2012;43(6):1537-41.
27. Taylor-Rowan M, Wilson A, Dawson J, Quinn TJ. Functional Assessment for Acute Stroke Trials: Properties, Analysis, and Application. *Front Neurol.* 2018;9:191.
28. Banks JL, Marotta CA. Outcomes Validity and Reliability of the Modified Rankin Scale: Implications for Stroke Clinical Trials. *Stroke.* 1 de marzo de 2007;38(3):1091-6.
29. Heitsch Laura, Ibanez Laura, Carrera Caty, Binkley Michael M., Strbian Daniel, Tatlisumak Turgut, et al. Early Neurological Change After Ischemic Stroke Is Associated With 90-Day Outcome. *Stroke.* 1 de enero de 2021;52(1):132-41.
30. Torres-Aguila NP, Carrera C, Muiño E, Cullell N, Cárcel-Márquez J, Gallego-Fabrega C, et al. Clinical Variables and Genetic Risk Factors Associated with the Acute Outcome of Ischemic Stroke: A Systematic Review. *J Stroke.* septiembre de 2019;21(3):276-89.
31. A B, T GB, N R, V L, M R, C M, et al. Ischemic stroke outcome: A review of the influence of post-stroke complications within the different scenarios of stroke care. *Eur J Intern Med* [Internet]. abril de 2016 [citado 28 de julio de 2022];29. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/26723523/>
32. Spronk E, Sykes G, Falcione S, Munsterman D, Joy T, Kamtchum-Tatuene J, et al. Hemorrhagic Transformation in Ischemic Stroke and the Role of Inflammation. *Front Neurol* [Internet]. 14 de mayo de 2021 [citado 20 de marzo de 2023];12:661955. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8160112/>
33. Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, Michel P. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score. *Neurology.* 12 de junio de 2012;78(24):1916-22.
34. Kwok CS, Potter JF, Dalton G, George A, Metcalf AK, Ngeh J, et al. The SOAR Stroke Score Predicts Inpatient and 7-Day Mortality in Acute Stroke. *Stroke* [Internet]. julio de 2013 [citado 28 de julio de 2022];44(7):2010-2. Disponible en: <https://www.ahajournals.org/doi/10.1161/STROKEAHA.113.001148>
35. Griffiths A. *Genética*. 7.^a ed. McGRAW-HILL interamericana; 2002. 849 p.
36. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet.* abril de 2019;20(4):207-20.
37. Thompson & Thompson. *Genética en medicina*. 7.^a ed. Elsevier Masson;
38. Green ED, Watson JD, Collins FS. Human Genome Project: Twenty-five years of big biology. *Nature.* 1 de octubre de 2015;526(7571):29-31.
39. National Human Genome Research Institute. Human Genome Project Timeline [Internet]. Disponible en: <https://www.genome.gov/human-genome-project/timeline>
40. Schloss JA, Gibbs RA, Makhijani VB, Marziali A. Cultivating DNA Sequencing Technology After the Human Genome Project. *Annu Rev Genomics Hum Genet.* 31 de agosto de 2020;21:117-38.

41. Moraes F, Góes A. A decade of human genome project conclusion: Scientific diffusion about our genome knowledge. *Biochem Mol Biol Educ Bimon Publ Int Union Biochem Mol Biol*. 6 de mayo de 2016;44(3):215-23.
42. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 13 de septiembre de 1996;273(5281):1516-7.
43. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. agosto de 2019;20(8):467-84.
44. Shendure J, Findlay GM, Snyder MW. Genomic medicine -- progress, pitfalls, and promise. *Cell*. 21 de marzo de 2019;177(1):45-57.
45. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. *Nat Rev Methods Primer*. 26 de agosto de 2021;1(1):1-21.
46. Manolio TA, Collins FS. The HapMap and Genome-Wide Association Studies in Diagnosis and Therapy. *Annu Rev Med*. 2009;60:443-56.
47. Gibbs RA, Belmont JW, Hardenbol P, Willis TD, Yu F, Yang H, et al. The International HapMap Project. *Nature*. diciembre de 2003;426(6968):789-96.
48. Slatkin M. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*. junio de 2008;9(6):477-85.
49. Kruglyak L. The road to genome-wide association studies. *Nat Rev Genet*. abril de 2008;9(4):314-8.
50. Nannya Y, Taura K, Kurokawa M, Chiba S, Ogawa S. Evaluation of genome-wide power of genetic association studies based on empirical data from the HapMap project. *Hum Mol Genet*. 15 de octubre de 2007;16(20):2494-505.
51. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol*. noviembre de 2012;11(11):951-62.
52. Markus HS. Genetics studies in ischaemic stroke. *Transl Stroke Res*. diciembre de 2010;1(4):238-45.
53. International Stroke Genetics Consortium (ISGC), Wellcome Trust Case Control Consortium 2 (WTCCC2), Bellenguez C, Bevan S, Gschwendtner A, Spencer CCA, et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet*. 5 de febrero de 2012;44(3):328-33.
54. Meschia JF, Arnett DK, Ay H, Brown RD, Benavente O, Cole JW, et al. Stroke Genetics Network (SiGN) Study: Design and rationale for a genome-wide association study of ischemic stroke subtypes. *Stroke J Cereb Circ*. octubre de 2013;44(10):2694-702.
55. Matarín M, Brown WM, Scholz S, Simón-Sánchez J, Fung HC, Hernandez D, et al. A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release. *Lancet Neurol*. mayo de 2007;6(5):414-20.

56. Matarin M, Singleton A, Hardy J, Meschia J. The genetics of ischaemic stroke. *J Intern Med.* febrero de 2010;267(2):139-55.
57. Kubo M, Hata J, Ninomiya T, Matsuda K, Yonemoto K, Nakano T, et al. A nonsynonymous SNP in PRKCH (protein kinase C eta) increases the risk of cerebral infarction. *Nat Genet.* febrero de 2007;39(2):212-7.
58. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med.* 23 de abril de 2009;360(17):1718-28.
59. International Stroke Genetics Consortium, Wellcome Trust Case-Control Consortium 2. Failure to validate association between 12p13 variants and ischemic stroke. *N Engl J Med.* 22 de abril de 2010;362(16):1547-50.
60. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol.* octubre de 2008;64(4):402-9.
61. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, et al. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet.* agosto de 2009;41(8):876-8.
62. Schunkert H, Götz A, Braund P, McGinnis R, Tregouet DA, Mangino M, et al. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation.* 1 de abril de 2008;117(13):1675-84.
63. Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, et al. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. *Ann Neurol.* mayo de 2009;65(5):531-9.
64. Anderson CD, Biffi A, Rost NS, Cortellini L, Furie KL, Rosand J. Chromosome 9p21 in ischemic stroke: population structure and meta-analysis. *Stroke.* junio de 2010;41(6):1123-31.
65. Kilarski LL, Achterberg S, Devan WJ, Traylor M, Malik R, Lindgren A, et al. Meta-analysis in more than 17,900 cases of ischemic stroke reveals a novel association at 12q24.12. *Neurology.* 19 de agosto de 2014;83(8):678-85.
66. Malik R, Traylor M, Pulit SL, Bevan S, Hopewell JC, Holliday EG, et al. Low-frequency and common genetic variation in ischemic stroke. *Neurology.* 29 de marzo de 2016;86(13):1217-26.
67. Traylor M, Mäkelä KM, Kilarski LL, Holliday EG, Devan WJ, Nalls MA, et al. A novel MMP12 locus is associated with large artery atherosclerotic stroke using a genome-wide age-at-onset informed approach. *PLoS Genet.* julio de 2014;10(7):e1004469.
68. Pulit SL, McArdle PF, Wong Q, Malik R, Gwinn K, Achterberg S, et al. The NINDS Stroke Genetics Network: a genome-wide association study of ischemic stroke and its subtypes. *Lancet Neurol.* febrero de 2016;15(2):174-84.
69. Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, the Stroke Genetics Network (SiGN), and the International Stroke Genetics Consortium (ISGC). Identification of additional risk loci for

- stroke and small vessel disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* junio de 2016;15(7):695-707.
70. Traylor M, Malik R, Nalls MA, Cotlarciuc I, Radmanesh F, Thorleifsson G, et al. Genetic variation at 16q24.2 is associated with small vessel stroke. *Ann Neurol.* marzo de 2017;81(3):383-94.
 71. M T, E P, L T, S K, V A, Mk B, et al. Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies. *Lancet Neurol.* 25 de marzo de 2021;20(5):351-61.
 72. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet.* 2018;50(4):524-37.
 73. Malik R, Rannikmäe K, Traylor M, Georgakis MK, Sargurupremraj M, Markus HS, et al. Genome-wide meta-analysis identifies 3 novel loci associated with stroke. *Ann Neurol.* 2018;84(6):934-9.
 74. Keene KL, Hyacinth HI, Bis JC, Kittner SJ, Mitchell BD, Cheng YC, et al. A GWAS Meta-Analysis of Stroke in 22,000 individuals of African descent identifies novel associations with stroke. *Stroke.* agosto de 2020;51(8):2454-63.
 75. Debette S, Mishra A, Malik R, Hachiya T, Jürgenson T, Namba S, et al. Stroke genetics informs drug discovery and risk prediction across ancestries [Internet]. 2022 [citado 1 de marzo de 2022]. Disponible en: <https://doi.org/10.21203/rs.3.rs-1175817/v1>
 76. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am J Hum Genet.* 3 de abril de 2014;94(4):511-21.
 77. Chung J, Marini S, Pera J, Norrving B, Jimenez-Conde J, Roquer J, et al. Genome-wide association study of cerebral small vessel disease reveals established and novel loci. *Brain.* octubre de 2019;142(10):3176-89.
 78. Jian X, Fornage M. Imaging Endophenotypes of Stroke as a Target for Genetic Studies. *Stroke.* 1 de junio de 2018;49(6):1557-62.
 79. ISGC. Acute Endophenotypes Working Group. Disponible en: <https://www.strokegenetics.org/node/327>
 80. Carrera C, Cárcel-Márquez J, Cullell N, Torres-Águila N, Muiño E, Castillo J, et al. Single nucleotide variations in ZBTB46 are associated with post-thrombolytic parenchymal haematoma. *Brain J Neurol.* 4 de septiembre de 2021;144(8):2416-26.
 81. Muiño E, Cárcel-Márquez J, Carrera C, Llucà-Carol L, Gallego-Fabrega C, Cullell N, et al. RP11-362K2.2:RP11-767I20.1 Genetic Variation Is Associated with Post-Reperfusion Therapy Parenchymal Hematoma. A GWAS Meta-Analysis. *J Clin Med.* 16 de julio de 2021;10(14):3137.
 82. Ibanez L, Heitsch L, Carrera C, Farias FHG, Del Aguila JL, Dhar R, et al. Multi-ancestry GWAS reveals excitotoxicity associated with outcome after ischaemic stroke. *Brain J Neurol.* 25 de febrero de 2022;awac080.

83. Torres-Aguila NP, Carrera C, Giese AK, Cullell N, Muiño E, Cárcel-Marquez J, et al. GENOME-WIDE ASSOCIATION STUDY OF WHITE BLOOD CELL COUNTS IN ISCHEMIC STROKE PATIENTS. *Stroke*. diciembre de 2019;50(12):3618-21.
84. Mola-Caminal M, Carrera C, Soriano-Tárraga C, Giralt-Steinhauer E, Díaz-Navarro RM, Tur S, et al. PATJ Low Frequency Variants Are Associated with Worse Ischemic Stroke Functional Outcome: A Genome-Wide Meta-Analysis. *Circ Res*. 4 de enero de 2019;124(1):114-20.
85. Söderholm M, Pedersen A, Lorentzen E, Stanne TM, Bevan S, Olsson M, et al. Genome-wide association meta-analysis of functional outcome after ischemic stroke. *Neurology*. 19 de 2019;92(12):e1271-83.
86. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet*. agosto de 2016;17(8):487-500.
87. Lacal I, Ventura R. Epigenetic Inheritance: Concepts, Mechanisms and Perspectives. *Front Mol Neurosci* [Internet]. 2018 [citado 8 de febrero de 2022];11. Disponible en: <https://www.frontiersin.org/article/10.3389/fnmol.2018.00292>
88. Berdasco M, Esteller M. Clinical epigenetics: seizing opportunities for translation. *Nat Rev Genet*. febrero de 2019;20(2):109-27.
89. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res*. marzo de 2011;21(3):381-95.
90. Stanzione R, Cotugno M, Bianchi F, Marchitti S, Forte M, Volpe M, et al. Pathogenesis of Ischemic Stroke: Role of Epigenetic Mechanisms. *Genes*. 13 de enero de 2020;11(1):89.
91. Ng GYQ, Yun-An L, Sobey CG, Dheen T, Fann DYW, Arumugam TV. Epigenetic regulation of inflammation in stroke. *Ther Adv Neurol Disord*. 26 de abril de 2018;11:1756286418771815.
92. Bao MH, Szeto V, Yang BB, Zhu S zhen, Sun HS, Feng ZP. Long non-coding RNAs in ischemic stroke. *Cell Death Dis*. 15 de febrero de 2018;9(3):281.
93. Kaikkonen MU, Lam MTY, Glass CK. Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc Res*. 1 de junio de 2011;90(3):430-40.
94. Ghafouri-Fard S, Shirvani-Farsani Z, Hussien BM, Taheri M, Arefian N. Emerging Impact of Non-coding RNAs in the Pathology of Stroke. *Front Aging Neurosci* [Internet]. 2021 [citado 14 de febrero de 2022];13. Disponible en: <https://www.frontiersin.org/article/10.3389/fnagi.2021.780489>
95. Rakyen VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet*. agosto de 2011;12(8):529-41.
96. Janssen SM, Lorincz MC. Interplay between chromatin marks in development and disease. *Nat Rev Genet*. 4 de octubre de 2021;1-17.
97. Krupinski J, Carrera C, Muiño E, Torres N, Al-Baradie R, Cullell N, et al. DNA Methylation in Stroke. Update of Latest Advances. *Comput Struct Biotechnol J*. 9 de diciembre de 2017;16:1-5.

98. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet.* julio de 2012;13(7):484-92.
99. Laird PW. Principles and challenges of genome-wide DNA methylation analysis. *Nat Rev Genet.* marzo de 2010;11(3):191-203.
100. Campagna MP, Xavier A, Lechner-Scott J, Maltby V, Scott RJ, Butzkueven H, et al. Epigenome-wide association studies: current knowledge, strategies and recommendations. *Clin Epigenetics.* 4 de diciembre de 2021;13(1):214.
101. Mao R, Chou LS. Methylation Analysis by Restriction Endonuclease Digestion and Real-Time PCR. *Clin Chem.* 1 de julio de 2010;56(7):1050-2.
102. Zaina S, Heyn H, Carmona FJ, Varol N, Sayols S, Condom E, et al. DNA methylation map of human atherosclerosis. *Circ Cardiovasc Genet.* octubre de 2014;7(5):692-700.
103. Soriano-Tárraga C, Jiménez-Conde J, Giralt-Steinhauer E, Mola M, Ois Á, Rodríguez-Campello A, et al. Global DNA Methylation of Ischemic Stroke Subtypes. *PLoS ONE* [Internet]. 30 de abril de 2014 [citado 19 de junio de 2020];9(4). Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4005764/>
104. Liu Y, Wu H, Zhang W, Zhang H, Chen H, Zhou G, et al. Endonuclease-assisted hydrogel bead array for digital analysis of circulating tumor DNA methylation. *Sens Actuators B Chem.* 1 de febrero de 2020;304:127381.
105. Bibikova M, Fan JB. GoldenGate assay for DNA methylation profiling. *Methods Mol Biol Clifton NJ.* 2009;507:149-63.
106. Dedeurwaerder S, Defrance M, Bizet M, Calonne E, Bontempi G, Fuks F. A comprehensive overview of Infinium HumanMethylation450 data processing. *Brief Bioinform.* noviembre de 2014;15(6):929-41.
107. Jelinek J, Lee JT, Cesaroni M, Madzo J, Liang S, Lu Y, et al. Digital Restriction Enzyme Analysis of Methylation (DREAM). *Methods Mol Biol Clifton NJ.* 2018;1708:247-65.
108. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature.* noviembre de 2009;462(7271):315-22.
109. Visone R, Bacalini MG, Franco SD, Ferracin M, Colorito ML, Pagotto S, et al. DNA methylation of shelf, shore and open sea CpG positions distinguish high microsatellite instability from low or stable microsatellite status colon cancer stem cells. *Epigenomics.* mayo de 2019;11(6):587-604.
110. CNCB; NGDC. EWAS Atlas@ EWAS Open Platform. Disponible en: <https://ngdc.cncb.ac.cn/ewas/atlas>
111. Zheng SC, Breeze CE, Beck S, Teschendorff AE. Identification of differentially methylated cell types in epigenome-wide association studies. *Nat Methods.* 2018;15(12):1059-66.
112. Zhang X, Hamblin MH, Yin KJ. Non-coding RNAs and Stroke. *Neurosci Rev J Bringing Neurobiol Neurol Psychiatry.* febrero de 2019;25(1):22-6.

113. Demyanenko S, Dzreyan V, Sharifulina S. Histone Deacetylases and Their Isoform-Specific Inhibitors in Ischemic Stroke. *Biomedicines*. 11 de octubre de 2021;9(10):1445.
114. Mirzaei H, Momeni F, Saadatpour L, Sahebkar A, Goodarzi M, Masoudifar A, et al. MicroRNA: Relevance to stroke diagnosis, prognosis, and therapy. *J Cell Physiol*. 2018;233(2):856-65.
115. Dykstra-Aiello C, Jickling GC, Ander BP, Shroff N, Zhan X, Liu D, et al. Altered Expression of Long Noncoding RNAs in Blood After Ischemic Stroke and Proximity to Putative Stroke Risk Loci. *Stroke*. diciembre de 2016;47(12):2896-903.
116. Guo X, Yang J, Liang B, Shen T, Yan Y, Huang S, et al. Identification of Novel LncRNA Biomarkers and Construction of LncRNA-Related Networks in Han Chinese Patients with Ischemic Stroke. *Cell Physiol Biochem Int J Exp Cell Physiol Biochem Pharmacol*. 2018;50(6):2157-75.
117. Xu X, Zhuang C, Chen L. Exosomal Long Non-Coding RNA Expression from Serum of Patients with Acute Minor Stroke. *Neuropsychiatr Dis Treat*. 2020;16:153-60.
118. New epigenetic players in stroke pathogenesis: From non-coding RNAs to exosomal non-coding RNAs. *Biomed Pharmacother*. 1 de agosto de 2021;140:111753.
119. Zhu W, Tian L, Yue X, Liu J, Fu Y, Yan Y. LncRNA Expression Profiling of Ischemic Stroke During the Transition From the Acute to Subacute Stage. *Front Neurol [Internet]*. 2019 [citado 17 de febrero de 2022];10. Disponible en: <https://www.frontiersin.org/article/10.3389/fneur.2019.00036>
120. Deng QW, Li S, Wang H, Sun HL, Zuo L, Gu ZT, et al. Differential long noncoding RNA expressions in peripheral blood mononuclear cells for detection of acute ischemic stroke. *Clin Sci Lond Engl* 1979. 31 de julio de 2018;132(14):1597-614.
121. Wang M, Wang L, Pu L, Li K, Feng T, Zheng P, et al. LncRNAs related key pathways and genes in ischemic stroke by weighted gene co-expression network analysis (WGCNA). *Genomics*. 1 de mayo de 2020;112(3):2302-8.
122. Tan KS, Armugam A, Sepramaniam S, Lim KY, Setyowati KD, Wang CW, et al. Expression profile of MicroRNAs in young stroke patients. *PloS One*. 2 de noviembre de 2009;4(11):e7689.
123. Tan JR, Tan KS, Koo YX, Yong FL, Wang CW, Armugam A, et al. Blood microRNAs in Low or No Risk Ischemic Stroke Patients. *Int J Mol Sci*. 22 de enero de 2013;14(1):2072-84.
124. Jickling GC, Ander BP, Zhan X, Noblett D, Stamova B, Liu D. microRNA expression in peripheral blood cells following acute ischemic stroke and their predicted gene targets. *PloS One*. 2014;9(6):e99283.
125. Wang W, Sun G, Zhang L, Shi L, Zeng Y. Circulating microRNAs as novel potential biomarkers for early diagnosis of acute stroke in humans. *J Stroke Cerebrovasc Dis Off J Natl Stroke Assoc*. diciembre de 2014;23(10):2607-13.
126. Li P, Teng F, Gao F, Zhang M, Wu J, Zhang C. Identification of Circulating MicroRNAs as Potential Biomarkers for Detecting Acute Ischemic Stroke. *Cell Mol Neurobiol*. 1 de abril de 2015;35(3):433-47.

127. Huang S, Lv Z, Guo Y, Li L, Zhang Y, Zhou L, et al. Identification of Blood Let-7e-5p as a Biomarker for Ischemic Stroke. *PLoS ONE*. 24 de octubre de 2016;11(10):e0163951.
128. Sepramaniam S, Tan JR, Tan KS, DeSilva DA, Tavintharan S, Woon FP, et al. Circulating MicroRNAs as Biomarkers of Acute Stroke. *Int J Mol Sci*. enero de 2014;15(1):1418-32.
129. Sørensen SS, Nygaard AB, Nielsen MY, Jensen K, Christensen T. miRNA expression profiles in cerebrospinal fluid and blood of patients with acute ischemic stroke. *Transl Stroke Res*. diciembre de 2014;5(6):711-8.
130. Bam M, Yang X, Sen S, Zumbun EE, Dennis L, Zhang J, et al. Characterization of Dysregulated miRNA in Peripheral Blood Mononuclear Cells from Ischemic Stroke Patients. *Mol Neurobiol*. febrero de 2018;55(2):1419-29.
131. Sonoda T, Matsuzaki J, Yamamoto Y, Sakurai T, Aoki Y, Takizawa S, et al. Serum MicroRNA-Based Risk Prediction for Stroke. *Stroke*. junio de 2019;50(6):1510-8.
132. Su Y, Zhang L, Zhou Y, Ding L, Li L, Wang Z. The progress of research on histone methylation in ischemic stroke pathogenesis. *J Physiol Biochem* [Internet]. 2 de septiembre de 2021 [citado 17 de febrero de 2022]; Disponible en: <https://doi.org/10.1007/s13105-021-00841-w>
133. Hanjin C, Tao L, Pengfei L, Ali Y, Huajun Z, Jiekun L, et al. Altered Long Noncoding RNA and Messenger RNA Expression in Experimental Intracerebral Hemorrhage - a Preliminary Study. *Cell Physiol Biochem*. 2018;45(3):1284-301.
134. Wang J, Zhu Y, Jin F, Tang L, He Z, He Z. Differential expression of circulating microRNAs in blood and haematoma samples from patients with intracerebral haemorrhage. *J Int Med Res*. junio de 2016;44(3):419-32.
135. Ouyang Y, Li D, Wang H, Wan Z, Luo Q, Zhong Y, et al. MiR-21-5p/dual-specificity phosphatase 8 signalling mediates the anti-inflammatory effect of haem oxygenase-1 in aged intracerebral haemorrhage rats. *Aging Cell*. diciembre de 2019;18(6):e13022.
136. Guo D, Liu J, Wang W, Hao F, Sun X, Wu X, et al. Alteration in abundance and compartmentalization of inflammation-related miRNAs in plasma after intracerebral hemorrhage. *Stroke*. junio de 2013;44(6):1739-42.
137. Leung LY, Chan CPY, Leung YK, Jiang HL, Abrigo JM, Wang DF, et al. Comparison of miR-124-3p and miR-16 for early diagnosis of hemorrhagic and ischemic stroke. *Clin Chim Acta Int J Clin Chem*. 10 de junio de 2014;433:139-44.
138. Sultan W, Machado LGDD, Ali MG, Tramontana A, Bayoumy AE, Baxter SG, et al. MicroRNAs as biomarkers in spontaneous intracerebral hemorrhage: A systematic review of recent clinical evidence. *Clin Neurol Neurosurg*. 1 de febrero de 2022;213:107130.
139. Mens MMJ, Heshmatollah A, Fani L, Ikram MA, Ikram MK, Ghanbari M. Circulatory MicroRNAs as Potential Biomarkers for Stroke Risk. *Stroke*. marzo de 2021;52(3):945-53.
140. Chandran R, Mehta SL, Vemuganti R. Non-coding RNAs and Neuroprotection after Acute CNS Injuries. *Neurochem Int*. diciembre de 2017;111:12-22.

141. Wang L, Zheng S, Zhang L, Xiao H, Gan H, Chen H, et al. Histone Deacetylation 10 Alleviates Inflammation After Intracerebral Hemorrhage via the PTPN22/NLRP3 Pathway in Rats. *Neuroscience*. 15 de abril de 2020;432:247-59.
142. Wang J, Cao B, Zhao H, Gao Y, Luo Y, Chen Y, et al. Long noncoding RNA H19 prevents neurogenesis in ischemic stroke through p53/Notch1 pathway. *Brain Res Bull*. agosto de 2019;150:111-7.
143. Zeng W, Jin J. The correlation of serum long non-coding RNA ANRIL with risk factors, functional outcome, and prognosis in atrial fibrillation patients with ischemic stroke. *J Clin Lab Anal*. agosto de 2020;34(8):e23352.
144. Zhu M, Li N, Luo P, Jing W, Wen X, Liang C, et al. Peripheral Blood Leukocyte Expression of lncRNA MIAT and Its Diagnostic and Prognostic Value in Ischemic Stroke. *J Stroke Cerebrovasc Dis Off J Natl Stroke Assoc*. febrero de 2018;27(2):326-37.
145. Edwardson MA, Zhong X, Fiandaca MS, Federoff HJ, Cheema AK, Dromerick AW. Plasma microRNA markers of upper limb recovery following human stroke. *Sci Rep*. 22 de agosto de 2018;8(1):12558.
146. Zeng L, Liu J, Wang Y, Wang L, Weng S, Tang Y, et al. MicroRNA-210 as a novel blood biomarker in acute cerebral ischemia. *Front Biosci Elite Ed*. 1 de junio de 2011;3:1265-72.
147. Huang S, Zhao J, Huang D, Zhuo L, Liao S, Jiang Z. Serum miR-132 is a risk marker of post-stroke cognitive impairment. *Neurosci Lett*. 26 de febrero de 2016;615:102-6.
148. Zhang Y, Cheng L, Chen Y, Yang GY, Liu J, Zeng L. Clinical predictor and circulating microRNA profile expression in patients with early onset post-stroke depression. *J Affect Disord*. 15 de marzo de 2016;193:51-8.
149. Liang HB, He JR, Tu XQ, Ding KQ, Yang GY, Zhang Y, et al. MicroRNA-140-5p: A novel circulating biomarker for early warning of late-onset post-stroke depression. *J Psychiatr Res*. agosto de 2019;115:129-41.
150. Zheng HW, Wang YL, Lin JX, Li N, Zhao XQ, Liu GF, et al. Circulating MicroRNAs as potential risk biomarkers for hematoma enlargement after intracerebral hemorrhage. *CNS Neurosci Ther*. diciembre de 2012;18(12):1003-11.
151. Wang MD, Wang Y, Xia YP, Dai JW, Gao L, Wang SQ, et al. High Serum MiR-130a Levels Are Associated with Severe Perihematoma Edema and Predict Adverse Outcome in Acute ICH. *Mol Neurobiol*. marzo de 2016;53(2):1310-21.
152. Hu YL, Wang H, Huang Q, Wang G, Zhang HB. MicroRNA-23a-3p promotes the perihematoma edema formation after intracerebral hemorrhage via ZO-1. *Eur Rev Med Pharmacol Sci*. mayo de 2018;22(9):2809-16.
153. Zhu Y, Wang JL, He ZY, Jin F, Tang L. Association of Altered Serum MicroRNAs with Perihematoma Edema after Acute Intracerebral Hemorrhage. *PLOS ONE*. 24 de julio de 2015;10(7):e0133783.
154. Kashif H, Shah D, Sukumari-Ramesh S. Dysregulation of microRNA and Intracerebral Hemorrhage: Roles in Neuroinflammation. *Int J Mol Sci*. 29 de julio de 2021;22(15):8115.

155. Uzdensky. Histone acetylation and deacetylation in ischemic stroke [Internet]. [citado 17 de febrero de 2022]. Disponible en: <https://www.nrronline.org/article.asp?issn=1673-5374;year=2021;volume=16;issue=8;spage=1529;epage=1530;aulast=Uzdensky>
156. Hurtado O, Hernández-Jiménez M, Zarruk JG, Cuartero MI, Ballesteros I, Camarero G, et al. Citicoline (CDP-choline) increases Sirtuin1 expression concomitant to neuroprotection in experimental stroke. *J Neurochem*. septiembre de 2013;126(6):819-26.
157. Sukumari-Ramesh S, Alleyne CH, Dhandapani KM. The Histone Deacetylase Inhibitor Suberoylanilide Hydroxamic Acid (SAHA) Confers Acute Neuroprotection After Intracerebral Hemorrhage in Mice. *Transl Stroke Res*. 1 de abril de 2016;7(2):141-8.
158. Yang H, Ni W, Jiang H, Lei Y, Su J, Gu Y, et al. Histone Deacetylase Inhibitor Scriptaid Alleviated Neurological Dysfunction after Experimental Intracerebral Hemorrhage in Mice. *Behav Neurol*. 12 de agosto de 2018;2018:6583267.
159. Yang H, Ni W, Wei P, Li S, Gao X, Su J, et al. HDAC inhibition reduces white matter injury after intracerebral hemorrhage. *J Cereb Blood Flow Metab Off J Int Soc Cereb Blood Flow Metab*. mayo de 2021;41(5):958-74.
160. Bonsack F, Sukumari-Ramesh S. Entinostat improves acute neurological outcomes and attenuates hematoma volume after Intracerebral Hemorrhage. *Brain Res*. 1 de febrero de 2021;1752:147222.
161. Baccarelli A, Wright R, Bollati V, Litonjua A, Zanobetti A, Tarantini L, et al. Ischemic Heart Disease and Stroke in Relation to Blood DNA Methylation. *Epidemiol Camb Mass*. noviembre de 2010;21(6):819-28.
162. Gómez-Úriz AM, Milagro FI, Mansego ML, Cordero P, Abete I, De Arce A, et al. Obesity and ischemic stroke modulate the methylation levels of KCNQ1 in white blood cells. *Hum Mol Genet*. 1 de marzo de 2015;24(5):1432-40.
163. Rask-Andersen M, Martinsson D, Ahsan M, Enroth S, Ek WE, Gyllensten U, et al. Epigenome-wide association study reveals differential DNA methylation in individuals with a history of myocardial infarction. *Hum Mol Genet*. 1 de noviembre de 2016;25(21):4739-48.
164. Soriano-Tárraga C, Lazcano U, Giralte-Steinhauer E, Avellaneda-Gómez C, Ois Á, Rodríguez-Campello A, et al. Identification of 20 novel loci associated with ischaemic stroke. Epigenome-wide association study. *Epigenetics*. 6 de abril de 2020;1-10.
165. Shen Y, Peng C, Bai Q, Ding Y, Yi X, Du H, et al. Epigenome-Wide Association Study Indicates Hypomethylation of MTRNR2L8 in Large-Artery Atherosclerosis Stroke. *Stroke*. junio de 2019;50(6):1330-8.
166. Sun H, Xu J, Hu B, Liu Y, Zhai Y, Sun Y, et al. Association of DNA Methylation Patterns in 7 Novel Genes With Ischemic Stroke in the Northern Chinese Population. *Front Genet*. 2022;13:844141.
167. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 24 de enero de 2013;49(2):359-67.

168. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 10 de diciembre de 2013;14(10):3156.
169. Soriano-Tárraga C, Giralt-Steinhauer E, Mola-Caminal M, Vivanco-Hidalgo RM, Ois A, Rodríguez-Campello A, et al. Ischemic stroke patients are biologically older than their chronological age. *Aging.* 25 de 2016;8(11):2655-66.
170. Wang C, Ni W, Yao Y, Just A, Heiss J, Wei Y, et al. DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: The NAS, and KORA F4. *EBioMedicine.* enero de 2021;63:103151.
171. Gallego-Fabrega C, Muiño E, Cullell N, Cárcel-Márquez J, Lazcano U, Soriano-Tárraga C, et al. Biological Age Acceleration Is Lower in Women With Ischemic Stroke Compared to Men. *Stroke.* 25 de febrero de 2022;STROKEAHA121037419.
172. Zhang Y, Long H, Wang S, Xiao W, Xiong M, Liu J, et al. Genome-Wide DNA Methylation Pattern in Whole Blood Associated With Primary Intracerebral Hemorrhage. *Front Immunol.* 2021;12:702244.
173. Davis Armstrong NM, Chen WM, Brewer MS, Williams SR, Sale MM, Worrall BB, et al. Epigenome-Wide Analyses Identify Two Novel Associations With Recurrent Stroke in the Vitamin Intervention for Stroke Prevention Clinical Trial. *Front Genet* [Internet]. 6 de septiembre de 2018 [citado 16 de marzo de 2021];9. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6135883/>
174. Davis Armstrong NM, Chen WM, Hsu FC, Brewer MS, Cullell N, Fernández-Cadenas I, et al. DNA methylation analyses identify an intronic ZDHHC6 locus associated with time to recurrent stroke in the Vitamin Intervention for Stroke Prevention (VISP) clinical trial. *PLoS ONE.* 12 de julio de 2021;16(7):e0254562.
175. Gallego-Fabrega C, Carrera C, Reny JL, Fontana P, Slowik A, Pera J, et al. TRAF3 Epigenetic Regulation Is Associated With Vascular Recurrence in Patients With Ischemic Stroke. *Stroke.* 2016;47(5):1180-6.
176. Gallego-Fabrega C, Carrera C, Reny JL, Fontana P, Slowik A, Pera J, et al. PPM1A Methylation Is Associated With Vascular Recurrence in Aspirin-Treated Patients. *Stroke.* 2016;47(7):1926-9.
177. Soriano-Tárraga C, Mola-Caminal M, Giralt-Steinhauer E, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, et al. Biological age is better than chronological as predictor of 3-month outcome in ischemic stroke. *Neurology.* 22 de agosto de 2017;89(8):830-6.
178. Soriano-Tárraga C, Giralt-Steinhauer E, Mola-Caminal M, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, et al. Biological Age is a predictor of mortality in Ischemic Stroke. *Sci Rep.* 7 de marzo de 2018;8(1):4148.
179. Soriano-Tárraga C, Lazcano U, Jiménez-Conde J, Ois A, Cuadrado-Godia E, Giralt-Steinhauer E, et al. Biological age is a novel biomarker to predict stroke recurrence. *J Neurol.* enero de 2021;268(1):285-92.
180. Kim BJ, Kim Y, Youn DH, Park JJ, Rhim JK, Kim HC, et al. Genome-wide blood DNA methylation analysis in patients with delayed cerebral ischemia after subarachnoid hemorrhage. *Sci Rep.* 10 de julio de 2020;10(1):11419.

181. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* 4 de enero de 2016;44(Database issue):D877-81.
182. Pereira S, Gibbs RA, McGuire AL. Open Access Data Sharing in Genomic Research. *Genes.* 29 de agosto de 2014;5(3):739-47.
183. Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* agosto de 1999;9(8):677-9.
184. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, et al. Ensembl 2021. *Nucleic Acids Res.* 8 de enero de 2021;49(D1):D884-91.
185. Safran M, Rosen N, Twik M, BarShir R, Stein TI, Dahary D, et al. The GeneCards Suite. En: Abugessaisa I, Kasukawa T, editores. *Practical Guide to Life Science Databases [Internet]*. Singapore: Springer; 2021 [citado 21 de febrero de 2022]. p. 27-56. Disponible en: https://doi.org/10.1007/978-981-16-5812-9_2
186. Murphy M, Brown G, Wallin C, Tatusova T, Pruitt K, Murphy T, et al. Gene Help: Integrated Access to Genes of Genomes in the Reference Sequence Collection [Internet]. Gene Help [Internet]. National Center for Biotechnology Information (US); 2021 [citado 21 de febrero de 2022]. Disponible en: <http://www.ncbi.nlm.nih.gov/books/NBK3841/>
187. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 8 de enero de 2021;49(D1):D480-9.
188. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 1 de enero de 2006;34(Database issue):D668-672.
189. Chen X, Ji ZL, Chen YZ. TTD: Therapeutic Target Database. *Nucleic Acids Res.* 1 de enero de 2002;30(1):412-5.
190. Whirl-Carrillo M, Huddart R, Gong L, Sangkuhl K, Thorn CF, Whaley R, et al. An Evidence-Based Framework for Evaluating Pharmacogenomics Knowledge for Personalized Medicine. *Clin Pharmacol Ther.* septiembre de 2021;110(3):563-72.
191. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 8 de enero de 2019;47(D1):D1005-12.
192. Xiong Z, Li M, Yang F, Ma Y, Sang J, Li R, et al. EWAS Data Hub: a resource of DNA methylation array data and metadata. *Nucleic Acids Res.* 8 de enero de 2020;48(D1):D890-5.
193. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: a database of human genotype–phenotype associations. *Bioinformatics.* 15 de octubre de 2016;32(20):3207-9.
194. Mountjoy E, Schmidt EM, Carmona M, Schwartzentruber J, Peat G, Miranda A, et al. An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nat Genet.* noviembre de 2021;53(11):1527-33.

195. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* junio de 2013;45(6):580-5.
196. Zheng Z, Huang D, Wang J, Zhao K, Zhou Y, Guo Z, et al. QTLbase: an integrative resource for quantitative trait loci across multiple human molecular phenotypes. *Nucleic Acids Res.* 8 de enero de 2020;48(D1):D983-91.
197. Vösa U, Claringbould A, Westra HJ, Bonder MJ, Deelen P, Zeng B, et al. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis [Internet]. *bioRxiv*; 2018 [citado 21 de febrero de 2022]. p. 447367. Disponible en: <https://www.biorxiv.org/content/10.1101/447367v1>
198. Liang L, Morar N, Dixon AL, Lathrop GM, Abecasis GR, Moffatt MF, et al. A cross-platform analysis of 14,177 expression quantitative trait loci derived from lymphoblastoid cell lines. *Genome Res.* abril de 2013;23(4):716-26.
199. Gaunt TR, Shihab HA, Hemani G, Min JL, Woodward G, Lyttleton O, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biol.* 31 de marzo de 2016;17(1):61.
200. Allman PH, Aban IB, Tiwari HK, Cutter GR. An introduction to Mendelian randomization with applications in neurology. *Mult Scler Relat Disord.* agosto de 2018;24:72-8.
201. Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. *JAMA.* 21 de noviembre de 2017;318(19):1925-6.
202. Koellinger PD, de Vlaming R. Mendelian randomization: the challenge of unobserved environmental confounds. *Int J Epidemiol.* 1 de junio de 2019;48(3):665-71.
203. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiol Camb Mass.* enero de 2017;28(1):30-42.
204. Zhou X, Wang L, Xiao J, Sun J, Yu L, Zhang H, et al. Alcohol consumption, DNA methylation and colorectal cancer risk: Results from pooled cohort studies and Mendelian randomization analysis. *Int J Cancer.* 25 de enero de 2022;
205. Portilla-Fernández E, Hwang SJ, Wilson R, Maddock J, Hill WD, Teumer A, et al. Meta-analysis of epigenome-wide association studies of carotid intima-media thickness. *Eur J Epidemiol.* 2021;36(11):1143-55.
206. Georgakis MK, Gill D, Webb AJS, Evangelou E, Elliott P, Sudlow CLM, et al. Genetically determined blood pressure, antihypertensive drug classes, and risk of stroke subtypes. *Neurology.* 28 de julio de 2020;95(4):e353-61.
207. Ni G, Moser G, Wray NR, Lee SH. Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood. *Am J Hum Genet.* 7 de junio de 2018;102(6):1185-94.
208. Lu Q, Li B, Ou D, Erlendsdottir M, Powles RL, Jiang T, et al. A Powerful Approach to Estimating Annotation-Stratified Genetic Covariance via GWAS Summary Statistics. *Am J Hum Genet.* 7 de diciembre de 2017;101(6):939-64.

209. Zhang Y, Lu Q, Ye Y, Huang K, Liu W, Wu Y, et al. SUPERGNOVA: local genetic correlation analysis reveals heterogeneous etiologic sharing of complex traits. *Genome Biol.* 7 de septiembre de 2021;22(1):262.
210. Guo H, Li JJ, Lu Q, Hou L. Detecting local genetic correlations with scan statistics. *Nat Commun.* 1 de abril de 2021;12(1):2033.
211. Cai H, Cai B, Liu Z, Wu W, Chen D, Fang L, et al. Genetic correlations and causal inferences in ischemic stroke. *J Neurol.* julio de 2020;267(7):1980-90.
212. White MJ, Yaspan BL, Veatch OJ, Goddard P, Risse-Adams OS, Contreras MG. Strategies for Pathway Analysis using GWAS and WGS Data. *Curr Protoc Hum Genet.* enero de 2019;100(1):e79.
213. Ren X, Kuan PF. methylGSA: a Bioconductor package and Shiny app for DNA methylation data length bias adjustment in gene set testing. *Bioinforma Oxf Engl.* 1 de junio de 2019;35(11):1958-9.
214. Maleki F, Ovens K, Hogan DJ, Kusalik AJ. Gene Set Analysis: Challenges, Opportunities, and Future Research. *Front Genet [Internet].* 2020 [citado 23 de febrero de 2022];11. Disponible en: <https://www.frontiersin.org/article/10.3389/fgene.2020.00654>
215. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 2 de julio de 2019;47(W1):W199-205.
216. Li B, Ritchie MD. From GWAS to Gene: Transcriptome-Wide Association Studies and Other Methods to Functionally Understand GWAS Discoveries. *Front Genet [Internet].* 30 de septiembre de 2021 [citado 28 de julio de 2022];12:713230. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8515949/>
217. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet [Internet].* marzo de 2016 [citado 29 de julio de 2022];48(3):245-52. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4767558/>
218. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet [Internet].* mayo de 2016 [citado 29 de julio de 2022];48(5):481-7. Disponible en: <https://www.nature.com/articles/ng.3538>
219. Yang J, Yan B, Fan Y, Yang L, Zhao B, He X, et al. Integrative analysis of transcriptome-wide association study and gene expression profiling identifies candidate genes associated with stroke. *PeerJ.* 2019;7:e7435.
220. Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov.* enero de 2019;18(1):41-58.
221. Konuma T, Ogawa K, Okada Y. Integration of genetically regulated gene expression and pharmacological library provides therapeutic drug candidates. *Hum Mol Genet.* 1 de febrero de 2021;30(3-4):294-304.

222. Lau A, So HC. Turning genome-wide association study findings into opportunities for drug repositioning. *Comput Struct Biotechnol J*. 1 de enero de 2020;18:1639-50.
223. Sanseau P, Agarwal P, Barnes MR, Pastinen T, Richards JB, Cardon LR, et al. Use of genome-wide association studies for drug repositioning. *Nat Biotechnol*. 10 de abril de 2012;30(4):317-20.
224. Wu Z, Wang Y, Chen L. Network-based drug repositioning. *Mol Biosyst*. 7 de mayo de 2013;9(6):1268-81.
225. Min JL, Hemani G, Davey Smith G, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics* [Internet]. 1 de diciembre de 2018 [citado 2 de agosto de 2022];34(23):3983-9. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6247925/>
226. Chen C, Ai QD, Chu SF, Zhang Z, Chen NH. NK cells in cerebral ischemia. *Biomed Pharmacother Biomedecine Pharmacother*. enero de 2019;109:547-54.
227. Gan Y, Liu Q, Wu W, Yin JX, Bai XF, Shen R, et al. Ischemic neurons recruit natural killer cells that accelerate brain infarction. *Proc Natl Acad Sci U S A*. 18 de febrero de 2014;111(7):2704-9.
228. Brenet F, Moh M, Funk P, Feierstein E, Viale AJ, Socci ND, et al. DNA methylation of the first exon is tightly linked to transcriptional silencing. *PloS One*. 18 de enero de 2011;6(1):e14524.
229. Ju H, Park KW, Kim ID, Cave JW, Cho S. Phagocytosis converts infiltrated monocytes to microglia-like phenotype in experimental brain ischemia. *J Neuroinflammation*. 18 de julio de 2022;19(1):190.
230. Orellana-Urzúa S, Rojas I, Líbano L, Rodrigo R. Pathophysiology of Ischemic Stroke: Role of Oxidative Stress. *Curr Pharm Des*. 2020;26(34):4246-60.
231. Peng DH, Liu YY, Chen W, Hu HN, Luo Y. Epidermal growth factor alleviates cerebral ischemia-induced brain injury by regulating expression of neutrophil gelatinase-associated lipocalin. *Biochem Biophys Res Commun*. 16 de abril de 2020;524(4):963-9.
232. Reznik ME, Yaghi S, Jayaraman MV, McTaggart RA, Hemendinger M, Mac Grory BC, et al. Baseline NIH Stroke Scale is an inferior predictor of functional outcome in the era of acute stroke intervention. *Int J Stroke Off J Int Stroke Soc*. octubre de 2018;13(8):806-10.
233. Meyer L, Stracke CP, Jungi N, Wallocha M, Broocks G, Sporns PB, et al. Thrombectomy for Primary Distal Posterior Cerebral Artery Occlusion Stroke: The TOPMOST Study. *JAMA Neurol*. 1 de abril de 2021;78(4):434-44.
234. Simats A, García-Berrocó T, Ramiro L, Giralt D, Gill N, Penalba A, et al. Characterization of the rat cerebrospinal fluid proteome following acute cerebral ischemia using an aptamer-based proteomic technology. *Sci Rep*. 21 de mayo de 2018;8(1):7899.
235. Mei K, Guo W. The exocyst complex. *Curr Biol CB*. 10 de septiembre de 2018;28(17):R922-5.

236. Gerges NZ, Backos DS, Rupasinghe CN, Spaller MR, Esteban JA. Dual role of the exocyst in AMPA receptor targeting and insertion into the postsynaptic membrane. *EMBO J.* 19 de abril de 2006;25(8):1623-34.
237. Sans N, Prybylowski K, Petralia RS, Chang K, Wang YX, Racca C, et al. NMDA receptor trafficking through an interaction between PDZ proteins and the exocyst complex. *Nat Cell Biol.* junio de 2003;5(6):520-30.
238. Achzet LM, Davison CJ, Shea M, Sturgeon I, Jackson DA. Oxidative Stress Underlies the Ischemia/Reperfusion-Induced Internalization and Degradation of AMPA Receptors. *Int J Mol Sci.* 13 de enero de 2021;22(2):E717.
239. Lai TW, Zhang S, Wang YT. Excitotoxicity and stroke: identifying novel targets for neuroprotection. *Prog Neurobiol.* abril de 2014;115:157-88.
240. Szydłowska K, Tymianski M. Calcium, ischemia and excitotoxicity. *Cell Calcium.* febrero de 2010;47(2):122-9.
241. Jablensky A, Morar B, Wiltshire S, Carter K, Dragovic M, Badcock JC, et al. Polymorphisms associated with normal memory variation also affect memory impairment in schizophrenia. *Genes Brain Behav.* junio de 2011;10(4):410-7.
242. Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics MCP.* febrero de 2014;13(2):397-406.
243. Ou YN, Yang YX, Deng YT, Zhang C, Hu H, Wu BS, et al. Identification of novel drug targets for Alzheimer's disease by integrating genetics and proteomes from brain and blood. *Mol Psychiatry [Internet].* octubre de 2021 [citado 3 de agosto de 2022];26(10):6065-73. Disponible en: <https://www.nature.com/articles/s41380-021-01251-6>
244. Wingo AP, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, et al. Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis. *Nat Genet.* febrero de 2021;53(2):143-6.
245. Liu DZ, Sharp FR. Excitatory and Mitogenic Signaling in Cell Death, Blood-brain Barrier Breakdown, and BBB Repair after Intracerebral Hemorrhage. *Transl Stroke Res.* julio de 2012;3(Suppl 1):62-9.
246. Chagnot A, Barnes SR, Montagne A. Magnetic Resonance Imaging of Blood-Brain Barrier permeability in Dementia. *Neuroscience.* 15 de octubre de 2021;474:14-29.
247. Bordes C, Sargurupremraj M, Mishra A, Debette S. Genetics of common cerebral small vessel disease. *Nat Rev Neurol.* febrero de 2022;18(2):84-101.
248. Wu BS, Chen SF, Huang SY, Ou YN, Deng YT, Chen SD, et al. Identifying causal genes for stroke via integrating the proteome and transcriptome from brain and blood. *J Transl Med [Internet].* 21 de abril de 2022 [citado 3 de agosto de 2022];20:181. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9022281/>

249. Zhang C, Qin F, Li X, Du X, Li T. Identification of novel proteins for lacunar stroke by integrating genome-wide association data and human brain proteomes. *BMC Med.* 23 de junio de 2022;20(1):211.
250. Fayard E, Xue G, Parcellier A, Bozulich L, Hemmings BA. Protein kinase B (PKB/Akt), a key mediator of the PI3K signaling pathway. *Curr Top Microbiol Immunol.* 2010;346:31-56.
251. Li J, Shan W, Zuo Z. Age-Related Upregulation of Carboxyl Terminal Modulator Protein Contributes to the Decreased Brain Ischemic Tolerance in Older Rats. *Mol Neurobiol.* julio de 2018;55(7):6145-54.
252. Chen Y, Nie H, Tian L, Tong L, Deng J, Zhang Y, et al. Sevoflurane preconditioning-induced neuroprotection is associated with Akt activation via carboxy-terminal modulator protein inhibition. *Br J Anaesth.* febrero de 2015;114(2):327-35.
253. Raub D, Platzbecker K, Grabitz SD, Xu X, Wongtangman K, Pham SB, et al. Effects of Volatile Anesthetics on Postoperative Ischemic Stroke Incidence. *J Am Heart Assoc* [Internet]. 2 de marzo de 2021 [citado 16 de diciembre de 2022];10(5):e018952. Disponible en: <https://www.ahajournals.org/doi/10.1161/JAHA.120.018952>
254. Berkhemer OA, van den Berg LA, Fransen PSS, Beumer D, Yoo AJ, Lingsma HF, et al. The effect of anesthetic management during intra-arterial therapy for acute stroke in MR CLEAN. *Neurology.* 16 de agosto de 2016;87(7):656-64.
255. Löwhagen Hendén P, Rentzos A, Karlsson JE, Rosengren L, Leiram B, Sundeman H, et al. General Anesthesia Versus Conscious Sedation for Endovascular Treatment of Acute Ischemic Stroke: The AnStroke Trial (Anesthesia During Stroke). *Stroke.* junio de 2017;48(6):1601-7.
256. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, et al. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat Genet* [Internet]. octubre de 2012 [citado 16 de diciembre de 2022];44(10):10.1038/ng.2397. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3651583/>
257. Janssen L, Ai X, Zheng X, Wei W, Caglayan AB, Kilic E, et al. Inhibition of Fatty Acid Synthesis Aggravates Brain Injury, Reduces Blood-Brain Barrier Integrity and Impairs Neurological Recovery in a Murine Stroke Model. *Front Cell Neurosci* [Internet]. 2021 [citado 15 de diciembre de 2022];15. Disponible en: <https://www.frontiersin.org/articles/10.3389/fncel.2021.733973>
258. Wei XJ, Han M, Wei GC, Duan CH. Prognostic value of cerebrospinal fluid free fatty acid levels in patients with acute ischemic stroke. *Front Hum Neurosci* [Internet]. 2015 [citado 15 de diciembre de 2022];9. Disponible en: <https://www.frontiersin.org/articles/10.3389/fnhum.2015.00402>
259. Jiang S, Li T, Ji T, Yi W, Yang Z, Wang S, et al. AMPK: Potential Therapeutic Target for Ischemic Stroke. *Theranostics* [Internet]. 10 de agosto de 2018 [citado 15 de diciembre de 2022];8(16):4535-51. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6134933/>