UNIVERSITAT DE BARCELONA

Estudios genéticos masivos en el ictus isquémico: factores de riesgo y pronóstico

Nuria Paz Torres Águila

Estudios genéticos masivos en el ictus isquémico: factores de riesgo y pronóstico

Nuria P Torres-Águila

Tesis doctoral

UNIVERSITAT DE BARCELONA

ESTUDIOS GENÉTICOS MASIVOS EN EL ICTUS ISQUÉMICO: FACTORES DE RIESGO Y PRONÓSTICO

Memoria presentada por

Nuria Paz Torres Águila

Para optar al grado de

Doctora por la Universidad de Barcelona

Programa de Genética

Realizado en "Fundació Doncència i Recerca Mútua Terrassa"

en colaboración con "Vall d'Heborn Institut de Recerca"

Fundació Docència i Recerca MútuaTerrassa

Vall d'Hebron Institut de Recerca

Barcelona, 27 de Septiembre de 2019

Doctoranda	Director	Tutor
Nuria Paz Torres Águila	Israel Fernández Cadenas	Ricard Albalat Rodríguez

AGRADECIMIENTOS

Quisiera agradecer a todas las personas que me han apoyado durante estos últimos años en los que he desarrollado la presente Tesis. En primer lugar, a mi Director de Tesis Israel Fernández, que además de proponer y diseñar estudios, ha sido un jefe de grupo excelente, siempre disponible para la resolución de dudas de última hora. Agradecer especialmente a mi familia, mis amigos y mi pareja que me han dado apoyo moral y ánimos cuando más lo he necesitado. También agradecer a los miembros del 'Oikolab', que me han ayudado a llegar hasta aquí, puesto que apostaron por mí cuando estaba empezando mi carrera investigadora. Y finalmente, agradecer a los miembros del laboratorio de 'Farmacogenómica y genética Neurovascular', por aportar tanto ideas para los proyectos como risas, haciendo ameno el trabajo día a día; agradecer en especial a Caty que me mostró las bases de los análisis GWAS y tuvo paciencia en mis inicios.

¡Muchísimas gracias!

SINOPSIS

Stroke is a complex disease that affects the brain and vascular system. This disease has a high mortality, being the second cause of death worldwide and the fist cause of disability in adults of developed countries. Currently, more than 30 loci have been found associated with ischemic stroke risk; however, these genetic risk factors did not explain all heritability associated with this disease. Additionally, little is known about the genetics behind stroke outcome. Recently, two Genome-Wide Association Studies had reported two new candidate genes for ischemic stroke outcome. However, even a large sample size has been used in these studies, the results found are fewer that the expected. This could be due to the high heterogeneity of ischemic stroke outcome.

The principal aims of this thesis were, on one hand, to explore whether new stroke genetic risk factors can be found in specific homogeneous populations, and, on the other hand, to find new genetic risk factors of ischemic stroke outcome by using a new approach focus on acute endophenotypes associated with stroke outcome. In order to reach the first aim, it has been performed a case-control genome-wide association analysis using a Spanish cohort as a discovery and an international cohort as replication, with an additional validation in a second independent Spanish cohort. For second aim, it has been studied the clinical variables associated with the acute phase outcome of ischemic stroke. The variable leukocyte count was selected as an interesting endophenotype to find new genetic risk factors associated with stroke outcome.

As a results of this research, a new locus associated with lacunar stroke risk has been described, giving new information about the physiopathology of this disease. Additionally, the enzyme codified in the candidate gen *MAN2B1* is a potential biomarker for lacunar stroke diagnosis. Moreover, it has been described in this thesis the first locus associated with leukocyte counts during the acute phase of ischemic stroke that additionally was associated with stroke outcome. Thus, the genetic analysis of acute endophenotypes has been confirmed as a useful approach to find new genetic risk factors associated with stroke outcome.

ÍNDICE

ABREVIATURAS

3'-UTR Región no traducida 3'

5'-UTR Región no traducida 5'

ADN Ácido desoxirribonucleico

ARN Ácido ribonucleico

AVM Malformación arteriovenosa

BBB Barrera hematoencefálica ('Blood-brain barrier')

CARASAL Arteriopatía relacionada con catepsina A con derrames cerebrales y leucoencefalopatía ('Cathepsin A-related arteriopathy with strokes and leukoencephalopathy')

CE Ictus cardioembólico

CpG Dímero citosina-fosfato-guanina

DADA2 Deficiencia de adenosina desaminasa 2

eQTL loci característico de expresión cuantitativo (expression quantitative trait loci)

EWAS análisis epigenético masivo de asociación ('Epigenome-wide association study')

GISCOME 'Genetics of Ischemic Stroke Functional Outcome'

GODS 'Genetic contribution to functional Outcome and Disability after Stroke'

GWAS análisis genómico masivo de asociación ('Genome-wide association study')

HDL lipoproteína de elevada densidad

LAA ictus aterotrombótico

LD desequilibrio de ligamiento

MR aleatorización Mendeliana

mRS escala de rankin modificada

NGS secuenciación de nueva generación

NIHSS 'National Institute of Health Stroke Scale'

NINDS-SiGN 'National Institute of Neurological Disorders Stroke Genetics Network'

PCR reacción en cadena de la polimerasa

PH-2 hematoma parenquimatoso 2

rtPA activador de plasminógeno tisular recombinante

SNP polimorfismo de nucleótido único

SVS ictus lacunar

TOAST 'Trial of ORG 10172 in Acute Stroke Tratment'

WMH hiperintensidades de sustancia blanca

INTRODUCCIÓN

1. Introducción

1.1. El Ictus

Definición

El ictus es una enfermedad cardiovascular de elevada incidencia mundial, siendo actualmente la segunda mayor causa de mortalidad. Esta enfermedad afecta a los vasos sanguíneos que irrigan el cerebro y ocurre cuando alguno de estos vasos se ocluye o se rompe, causando la interrupción del riego sanguíneo en una parte del cerebro. Si esta ausencia de riego perdura en el tiempo, provoca la muerte de las células neuronales por la falta de oxígeno y de nutrientes.

La sintomatología del ictus es diversa, comprendiendo desde dificultad en el habla a la pérdida de conciencia, pasando por la parálisis de extremidades y la pérdida de visón. Asimismo, las guías médicas dictan una sintomatología más concreta para ayudar al paciente a determinar si está sufriendo un ictus; estas son:

- Súbito dolor de cabeza, de gran intensidad y sin causa aparente.
- Pérdida de la visión, total o parcial o incluso visión doble con sensación de mareo.
- Dificultad en el habla.
- Pérdida de fuerza y sensibilidad en un lado del cuerpo, ya se cara, brazos o piernas.
- Pérdida de equilibrio, inestabilidad o sensación de vértigo y confusión.

Clasificación

Se pueden distinguir dos tipos diferentes de ictus dependiendo de la causa de éste. Si se produce una obstrucción del vaso sanguíneo se clasifica como ictus isquémico, por lo contrario, si el ictus ocurre como consecuencia de la rotura de un vaso sanguíneo se conoce como ictus hemorrágico (Figura 1).

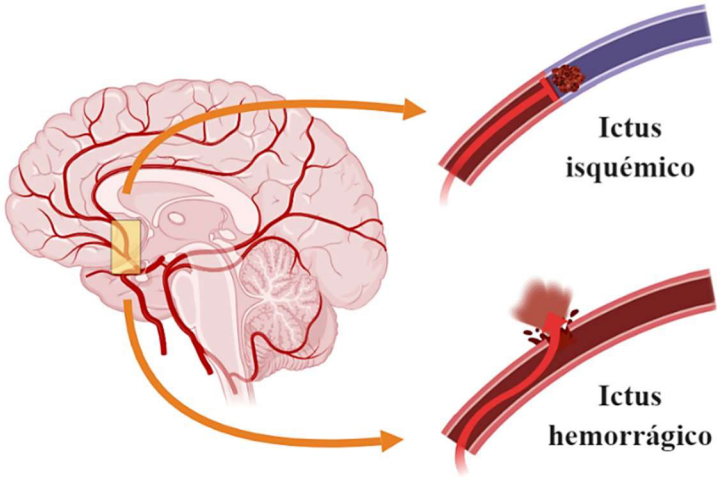

Figura 1. Representación gráfica de los principales tipos de ictus. Realizada en BioRender[1].

Ictus isquémico

El ictus isquémico es el tipo de infarto cerebral más frecuente a nivel mundial, supone el 85% de los casos reportados[2] y actualmente se puede tratar con trombolíticos y/o trombectomía mecánica. Asimismo, ambos tratamientos deben ser aplicados dentro de las primeras 24 horas desde el inicio de los síntomas. El suministro de trombolíticos al paciente se suele hacer por vía endovenosa, llamándose "tratamiento endovascular". El trombolítico más ampliamente utilizado actualmente es el activador de plasminógeno tisular recombinante o rtPA (de sus siglas en inglés 'recombinant tissue plasminogen activator'). Este fármaco deshace el trombo que ocluye el vaso permitiendo recuperar el flujo sanguíneo; concretamente, cataliza el paso de plasminogeno a plasmina, la cual hidroliza las cadenas de fibrina que forman el coágulo provocando su degradación. Por otra parte, la trombectomía mecánica, tal y como su nombre indica, consiste en retirar de manera mecánica el trombo que ocluye el vaso. Existen diversos instrumentos médicos que se pueden utilizar con este fin, los más utilizados son los 'stent' de segunda generación y los catéteres de aspiración[3].

Es posible distinguir diversos subtipos de ictus isquémico. La clasificación etiológica más extendida es el 'Trial of ORG 10172 in Acute Stroke Tratment' (TOAST)[4], el cual distingue cinco subgrupos diferentes de ictus isquémico:

- Ictus aterotrombótico o LAA ('large artery aterosclerosis'): presentan una estenosis u oclusión, mayor de un 50% en una arteria cerebral principal o en una rama cortical.
- Ictus cardioembólico o CE ('cardioembolic stroke'): oclusión de arteria cerebral causada por un trombo o coágulo producido o despendido de una cavidad del corazón.
- Ictus lacunar o SVS ('small vessel stroke'): oclusión de un vaso sanguíneo pequeño que provoca un infarto cerebral menor.
- Ictus indeterminado o criptogenético: presentan más de una causa de ictus, o el estudio completo del paciente no ha desvelado ninguna causa, o el estudio ha sido incompleto.
- Otro: el ictus se da debido a una causa infrecuente, como vasculopatías, enfermedades hematológicas, etc.

Dependiendo la etiología del ictus isquémico es posible recetar al paciente fármacos antiplaquetarios y/o anticoagulantes como estrategia de prevención secundaria. Ambos fármacos evitan la formación de futuros trombos, pero utilizan estrategias moleculares diferentes. Los antiplaquetarios evitan la señalización molecular de agregación plaquetaria, y se recetan a pacientes con ictus isquémico de etiología no cardioembólica. Por otra parte, los anticoagulantes ralentizan la formación de coágulos sanguíneos. Concretamente afectan a la síntesis de fibrina, la cual es la proteína principal en los coágulos y es esencial para su

formación. Estos últimos se administran a pacientes de ictus isquémico de etiología cardioembólica.

Ictus hemorrágico

El ictus hemorrágico se da por la rotura de un vaso cerebral (normalmente una pequeña arteria) causando la formación de un hematoma en el tejido cerebral pudiéndose extender a zonas más profundas del cerebro. La masa de sangre coagulada provoca la interrupción física del tejido cerebral y presión sobre el tejido circundante. Una vez se detiene la hemorragia, la sangre se desintegra lentamente y se absorbe durante un período que puede durar de semanas a meses. Actualmente no existe un tratamiento específico para este tipo de ictus por lo que suelen presentar una elevada mortalidad; sin embargo, el ictus hemorrágico es menos frecuente, llegando a ser el 15% del total de casos de ictus[2].

Es posible distinguir dos grandes subtipos de ictus hemorrágico dependiendo el lugar donde se produce la hemorragia:

- Intracraneal: es el tipo más común de ictus hemorrágico. Se produce por la rotura de un vaso en el interior del cerebro. Las causas más comunes de este tipo de ictus es la hipertensión y la edad avanzada, pero también puede ser ocasionado como consecuencia de la enfermedad genética "Malformación arteriovenosa" (AVM, del inglés 'arteriovenous malformation'). Se sub-clasifica en hemorragia intracraneal lobar y hemorragia intracraneal profunda.
- Subaracnoidea: la hemorragia se produce en el territorio entre el cerebro y el tejido colindante (el espacio subaracnoideo). Normalmente se produce por la rotura de la pared del vaso debido a un ensanchamiento de la misma, comúnmente llamado aneurisma.

Como tratamiento preventivo secundario para el ictus hemorrágico únicamente existe la posibilidad de retirar los fármacos antiplaquetarios o anticoagulantes, en el caso de que el paciente los estuviese tomando con anterioridad. Asimismo, para algunos anticoagulantes, como pueden ser los anti-trombina (dabigatrán) y los anti-Factor X de coagulación (rivaroxaban, apixaban, y edoxaban), existe también la posibilidad de administrar al paciente un antídoto para contrarrestar de manera inmediata el efecto de estos anticoagulantes.

Epidemiología

Actualmente el ictus tiene una gran incidencia mundial[5], llegando a los 15 millones de casos anuales. Asimismo, es la segunda causa de mortalidad a nivel mundial (con 5,5 millones de muertes anuales) y la primera causa de discapacidad en adultos en los países desarrollados.

Concretamente en España, las enfermedades del sistema circulatorio fueron la mayor causa de muerte en 2017, con 122.466 defunciones (28,8% del total). Asimismo, dentro de este grupo, las enfermedades cerebrovasculares fueron la segunda mayor causa de muerte, con un total de 26.937 defunciones[6]. En el caso del ictus, esta enfermedad presenta una elevada incidencia en la población española, con 276.200 casos reportados en los últimos 12 meses[6]. Asimismo, según datos del Instituto Nacional de Estadística[6], en 2017 se produjeron 7.643 defunciones debidas a infarto cerebral (ictus isquémico), 1.190 debidas a hemorragia intraencefálica no traumática (ictus hemorrágico intracraneal) y 1.153 debidas a hemorragia subaracnoidea (ictus hemorrágico subaracnoideo). Así pues, el ictus es la tercera causa principal de muerte en territorio español[7] y se encuentra entre las 10 primeras causas de discapacidad y mortalidad[7].

La elevada incidencia del ictus, tanto a nivel nacional como mundial, justifica la necesidad de investigar las causas y los factores de riesgo de esta enfermedad, así como los posibles factores implicados en el pronóstico del paciente.

Factores de riesgo

En la actualidad hay descritos diversos factores de riesgo clínicos para el ictus, tanto isquémico como hemorrágico, comunes con otras enfermedades cardiovasculares[8,9]. Estos factores se pueden clasificar en modificables y no modificables (Tabla 1). De entre los factores no modificables destacan la edad (a mayor edad mayor riesgo) y el sexo (el ictus es más frecuente en hombres, pero con mayor mortalidad en mujeres). Respecto a los factores de riesgo modificables, el más común es la hipertensión, así como la diabetes, la fibrilación auricular, y el tabaquismo[5,9].

Estos factores de riesgo se han utilizado para crear escalas de riesgo cardiovascular, comúnmente llamados escalas o 'scores' de predicción de riesgo. Estos 'scores' se utilizan actualmente en la clínica como una herramienta de apoyo para discernir entre pacientes que necesitan un seguimiento más constante de los que no. Uno de los 'scores' más conocido y ampliamente utilizado en clínica es el "Framingham Risk Score"[10]. Este 'score' está estratificado por sexo, existiendo un 'score' para hombres y otro para mujeres, e incluye varios de los factores de riesgo clínicos. En concreto, las variables incluidas son edad, colesterol total, colesterol HDL (de las siglas en inglés 'High Density Lipoproteins'), presión sanguínea sistólica, diabetes, y tabaquismo. El "Framingham Risk score" puede calcular la probabilidad de que un paciente tenga un evento cardiovascular en los próximos 10 años. En el caso de España, el 'score' de riesgo más utilizado es el "REGICOR"[11], el cual es una adaptación del "Framingham Risk score" calibrado para población catalana. No obstante, estos 'scores' están diseñados para enfermedades cardiovasculares por lo que no son específicos para ictus.

Tabla 1. Factores de riesgo descritos para enfermedades cerebrovasculares.

Factores de Riesgo no modificables	Factores de riesgo modificables	
	Bien documentados	Menos documentados
- Edad - Bajo peso al nacer - Etnia/raza - Factores genéticos	- Inactividad física - Dislipemia - Dieta y nutrición - Hipertensión - Obesidad y distribución de la grasa corporal - Diabetes mellitus - Tabaquismo - Fibrilación auricular - Otras cardiopatías - Estenosis asintomática de la arteria carótida - Anemia falciforme	- Migraña - Síndrome metabólico - Consumo de alcohol - Abuso de drogas - Apnea del sueño - Hiperhomocisteinemia - Lipoproteína(a) elevada - Hípercoagulabilidad - Inflamación e infecciones

Tabla basada en los factores descritos en Meschia et al. (2014)[8].

Progresión del ictus isquémico

El ictus isquémico es una enfermedad de aparición súbita en la cual se pueden distinguir diferentes fases durante su progresión. Estas fases se dividen según el tiempo trascurrido desde que se produce la obstrucción del vaso en: "fase aguda", "fase sub-aguda", y "fase tardía".

Fase aguda

La fase transcurrida durante las primeras 24-48h desde la aparición de los primeros síntomas del ictus se define como "fase aguda". Es durante esta fase cuando se puede intervenir para tratar el ictus y prevenir un daño neurológico mayor. El daño neurológico se puede medir mediante diferentes escalas neurológicas, la más ampliamente utilizada es la escala NIHSS (de sus siglas en inglés 'National Institute of Helath Stroke Scale')[12,13]. Esta escala evalúa diferentes aspectos neurológicos del paciente y puntúa cada uno de ellos creando un 'score' que determina la gravedad del daño neurológico (Tabla 2). La puntuación de esta escala varía desde 0 a 42, siendo 0 un estado neurológico bueno y 42 el más grave. Cuando se observa un incremento de 4 puntos en la escala NIHSS, se considera que el paciente tiene un deterioro neurológico sintomático.

Tabla 2. Parámetros evaluados en la escala de ictus NIHSS.

Parámetro neurológico a evaluar	Rango de Puntuación
1a. Nivel de conciencia	0 – 3
1b. Preguntas de nivel de conciencia	0 – 2
1c. Comandos de nivel de conciencia	0 – 2
2. Movilidad Ocular	0 – 2
3. Visual	0 – 3
4. Parálisis facial	0 – 3
5. Motricidad Brazos (5a. Izquierdo; 5b. Derecho)	0 – 4 (UN, por amputación)
6. Motricidad Piernas (6a. Izquierda; 6b. Derecha)	0 – 4 (UN, por amputación)
7. Ataxia de extremidades	0 – 2 (UN, por amputación)
8. Sensorial	0 – 2
9. Lenguaje (afasia)	0 – 3
10. Disartria	0 – 2 (UN, por intubación)
11. Extinción y falta de atención (antes de negligencia)	0 – 2

UN = no evaluable.

Fase sub-aguda

Tras la fase aguda prosigue la "fase sub-aguda". En esta etapa se puede observar el estado neurológico del paciente y determinar si el tratamiento ha sido exitoso, suele comprender a partir de 24 horas hasta 10 días después de sufrir el ictus. Cabe destacar que este periodo de tiempo no está estandarizado, por lo que la definición de fase sub-aguda es relativa. En la fase sub-aguda también se pueden evaluar y tratar complicaciones médicas asociadas al ictus como pueden ser la trasformación hemorrágica, el edema, y/o las infecciones, entre otras.

Fase tardía

La última etapa tras el ictus es la que se define como "fase tardía", en la cual se estudia la recuperación final del paciente normalmente en una visita de control, a los 3 meses o al año. En esta etapa se evalúa tanto el estado neurológico como el estado funcional del paciente. Para medir el estado funcional existen diferentes escalas, las más ampliamente utilizada es la Escala de Rankin modificada (mRS; de sus siglas en inglés 'modified Rankin Scale')[14]. La mRS puntúa el estado del paciente de 0 a 6, siendo el estado peor cuanto mayor es la puntuación (Tabla 3).

Asimismo, esta escala también suele utilizarse para dicotomizar los pacientes entre dependientes e independiente. Normalmente, en investigación se utilizan los valores de mRS

entre 0 y 2 como independientes y entre 3 y 6 para dependientes. Asimismo, esta definición es relativa, en algunos casos se utilizan únicamente los valores más extremos e incluso se pueden excluir los individuos con valor de mRS igual a 6, puesto que este valor indica mortalidad.

Tabla 3. Descripción de los niveles de la escala de Rankin modificada (mRS).

Puntuación	Estado funcional
0	**Sin síntomas**
1	**Sin incapacidad importante:** Capaz de realizar sus actividades y obligaciones habituales.
2	**Incapacidad leve:** Incapaz de realizar algunas de sus actividades previas, pero capaz de velar por sus intereses y asuntos sin ayuda.
3	**Incapacidad moderada:** Síntomas que restringen significativamente su estilo de vida o impiden su subsistencia autónoma.
4	**Incapacidad moderadamente severa:** Síntomas que impiden claramente su subsistencia independiente aunque sin necesidad de atención continua.
5	**Incapacidad severa:** Totalmente dependiente, necesitando asistencia constante día y noche.
6	**Muerte**

Factores asociados al Pronóstico

Existe un gran interés por definir los factores que determinan el pronóstico del ictus, ya que estos pueden ser de gran utilidad para mejorar el estado clínico de los pacientes, tanto a nivel funcional como neurológico.

Actualmente, igual que en el caso del riesgo de ictus, se han descrito diversos factores clínicos asociados con el estado final del paciente[15]. De entre estos destacan: la edad, el sexo, la severidad del ictus, la etiología del ictus, la presencia de factores de riesgo cardiovasculares (fibrilación auricular, enfermedad coronaria, insuficiencia cardíaca, diabetes mellitus, infarto de miocardio previo, tabaquismo), tener alguna comorbilidad (cáncer, demencia, disfunción renal), tener discapacidad previa al ictus y los niveles de glucosa en la admisión a urgencias[15].

Biomarcadores en el ictus isquémico

Un biomarcador es una indicación objetiva del estado médico del paciente que puede medirse de manera precisa y con reproducibilidad[16]. Algunos de los factores asociados al pronóstico pueden ser utilizados como biomarcadores para inferir el estado final del paciente tras padecer un ictus. La severidad inicial del ictus, medida con la escala NIHSS en el

momento de la admisión a urgencias, es uno de los factores más claramente relacionados con el pronóstico debido a la gran cantidad de estudios que soportan esta asociación[17–22]. Por ello, la severidad inicial del ictus es posible categorizarla como un biomarcador útil y suele ser incluida en los 'scores' de predicción.

Sin embargo, de cara a determinar las bases genéticas y moleculares que dirigen el pronóstico del paciente, los biomarcadores puede tener una capacidad limitada debido a que no siempre son causales. No obstante, los biomarcadores clasificados como endofenotipos sí poseen esta característica. Un endofenotipo es un fenotipo que puede ser descrito dentro de otro fenotipo mayor y que posee un componente genético heredable, el cual proporciona causalidad al endofenotipo.

En el caso del ictus (fenotipo mayor), algunos factores que pueden ser clasificados como endofenotipos son la presión sangínea[23], el número de células inmunes[24], los niveles de glucosa en sangre[25], o la concentración de hemoglobina[26], entre otros. Diversos estudios han podido relacionar algunos de estos endofenotipos con el estado final del paciente después de padecer un ictus[27–29].

Células del sistema inmunitario

De cara a modificar endofenotipos asociados al pronóstico para mejorar el estado final del paciente, uno de los endofenotipos más interesantes es el número de células inmunes, así como el porcentaje de cada tipo celular. En la literatura está descrito que durante la fase aguda del ictus hay una respuesta inflamatoria muy elevada por parte del sistema inmunitario[30] (Figura 2). Además, a nivel fisiológico se ha demostrado que los neutrófilos son las primeras células del sistema inmune que llegan al lugar del daño cerebral, promoviendo el reclutamiento de más células del sistema inmunitario y desencadenando la respuesta inflamatoria[31] (Figura 3).

Por todo ello se cree que las células del sistema inmune pueden jugar un papel importante en el pronóstico del ictus. Existen varios estudios que soportan esta idea. Un estudio a destacar es el de Nardi et al.[32], en el cual se describe la asociación del número de células inmunológicas (leucocitos) con el estado neurológico del paciente a las 72h post-ictus, es decir tras la fase aguda, y con el estado funcional del paciente a los 3 meses. De igual manera, en un estudio posteriro[33], se reportó la asociación de los niveles de neutrófilos antes del tratamiento trombolítico con el pronóstico del paciente a los 3 meses, hallando una asociación inversa: a mayor número de neutrófilos peor pronóstico.

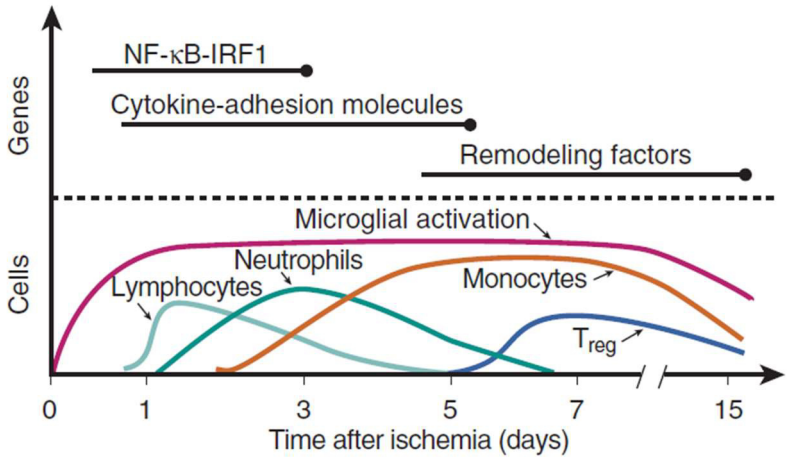

Figura 2. Representación gráfica de la respuesta inflamatoria en ratón tras sufrir una isquemia cerebral. El eje horizontal representa el tiempo tras la isquemia, y el eje vertical el número de células inmunes, así como los genes expresados. Figura extraída de Anrather et al. (2016)[30].

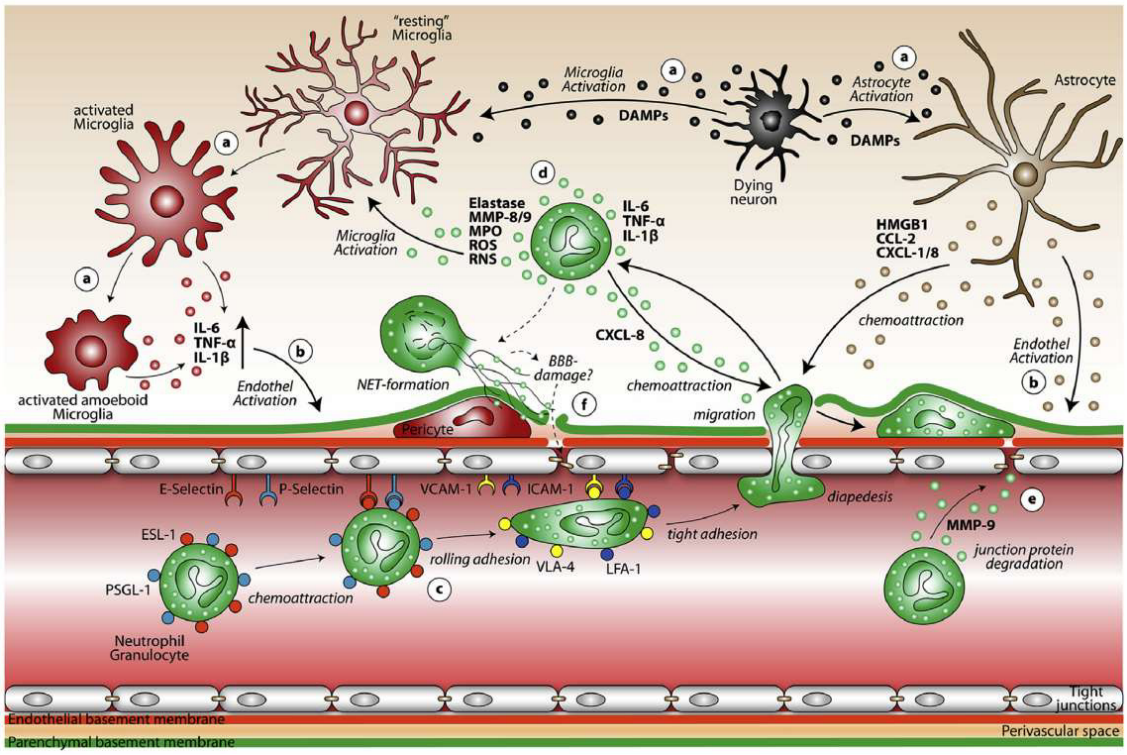

Figura 3. Esquematización de la señalización molecular que promueve el reclutamiento de neutrófilos en el ictus isquémico. Figura extraída de Strecker et al. (2016)[31].

1.2. Genética

Introducción

La genética es la ciencia que estudia la herencia biológica y la variabilidad entre organismos. Pese que la herencia biológica se ha estudiado incluso en la época clásica, se considera que la genética moderna se inició en el siglo XIX con el trabajo de Gregor Mendel, el cual estableció las bases de la herencia genética.

Posteriormente, el descubrimiento del ADN, de los genes y de los cromosomas permitió el avance de la genética molecular, siendo en el siglo XX cuando se realizaron los mayores avances. Entre estos destacan: el descubrimiento del código genético y el establecimiento de las bases de la regulación de la expresión génica. Asimismo, durante las últimas décadas del siglo XX se iniciaron los proyectos de secuenciación genética a gran escala, como la secuenciación de genomas completos, los cuales son esenciales para realizar estudios de genética molecular en la actualidad.

Proyectos genómicos humanos

De entre el gran número de proyectos de secuenciación realizados en las últimas décadas, existen tres grandes proyectos que han revolucionado el campo de la genómica humana: el Proyecto Genoma Humano, el Proyecto HapMap, y el Proyecto 1000 Genomas.

Proyecto Genoma Humano

El Proyecto Genoma Humano[34,35] se inició oficialmente en el año 1990 con el objetivo de obtener la secuenciación completa del genoma humano, para poder describir y estudiar el funcionamiento de todos los genes del *Homo sapiens*. El proyecto se finalizó en el año 2003, estimando el total de genes del ser humano entre 20.000 y 25.000 genes. Pese que actualmente diversos de estos genes aún no están descritos a nivel funcional, gracias a los resultados obtenidos en el Proyecto Genoma Humano, ha sido posible describir nuevos elementos moduladores de la expresión génica como micro-ARNs o islas CpGs, entre otros.

Asimismo, en este proyecto también se recalcó la presencia de los cambios genéticos o polimorfismos. Gracias a la secuenciación del proyecto Genoma Humano se observó que existían gran cantidad de pequeños cambios en la secuencia genética a lo largo del genoma al comparar diferentes individuos de la población humana.

Proyecto HapMap

El Proyecto HapMap tenía como objetivo determinar los cambios genéticos presentes en la población que eran heredados conjuntamente (haplotipo), con el fin de crear un mapa de referencia de los mismos. Por definición, un haplotipo únicamente está formado por polimorfismos de nucleótido único o SNP (de sus siglas en inglés 'Single Nucleotide Polymorphism'). Como su propio nombre indica, un SNP es un cambio único de nucleótido

por otra de las bases nucleotídicas del ADN (adenina, timina, guanina o citosina) en una posición concreta del genoma (Figura 4). La segregación entre SNPs se puede calcular mediante el desequilibrio de ligamiento (LD, del inglés 'linkage disequilibrium'). El desequilibrio de ligamiento se define como la diferencia entre la frecuencia observada de una combinación particular de alelos en dos SNPs y la frecuencia esperada para la asociación aleatoria. Así pues, esta escala de asociación va desde 1 a 0, siendo 1 totalmente asociados y 0 totalmente independientes. Gracias a esta posición única en el genoma es posible crear un mapa con su localización, y al conocer que SNPs están en LD es posible determinar un 'tag-SNP' que represente la distribución alélica del haplotipo al completo (Figura 4).

Los resultados del proyecto HapMap permitieron, junto con los arrays de genotipado de polimorfismos, inferir los polimorfismos no genotipados para generar los primeros resultados de estudios de genoma completo[36,37].

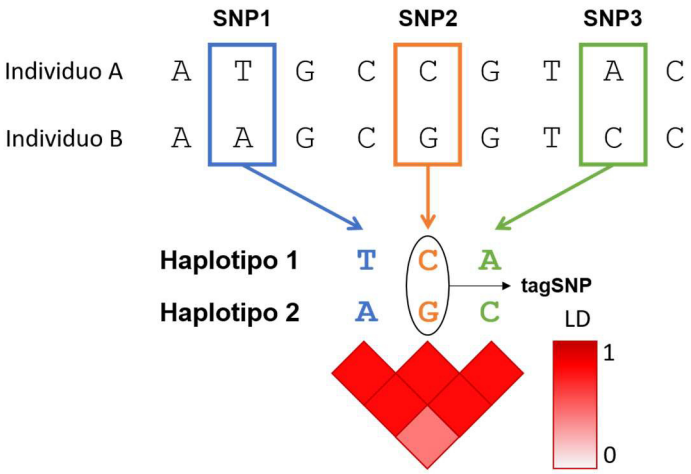

Figura 4. Esquematización gráfica de SNP y haplotipo. LD = desequilibrio de ligamiento (del inglés, 'linkage disequilibrium').

Proyecto 1000 Genomas

El "Proyecto 1000 Genomas"[38] tenía como objetivo encontrar el mayor número posible de cambios genéticos presentes entre la población con una frecuencia de al menos un 1%. Para ello, se secuenciaron un total de 1000 genomas de diferentes personas y poblaciones con tal de tener una distribución alélica representativa de la población humana mundial. Los resultados de este proyecto, iniciado el 2008 y finalizado en 2015, se utilizaron para establecer un genoma de referencia para el ser humano, pudiéndose apreciar diferencias en la distribución alélica de diversos polimorfismos al compararlas entre las poblaciones participantes.

Actualmente se utilizan los resultados del proyecto 1000 Genomas como referencia en los estudios de genoma completo para estimar o imputar los polimorfismos no genotipados.

Bases de la genética molecular moderna

La genética molecular humana actual se basa principalmente en los 'cambios genéticos' o 'polimorfos' presentes a lo largo de toda la secuencia genómica. Existen diferentes tipos de cambios genéticos, sin embargo, los SNPs representan el 90% de los cambios genéticos y, por ello, son los más frecuentes y ampliamente estudiados. Se pueden distinguir diferentes tipos de SNPs según su localización en el genoma, sub-clasificados a su vez según la implicación biológica que pueden llegar a tener.

SNPs codificantes

Un SNP codificante (o 'coding' SNP) es aquel polimorfismo de cambio único que se localiza en la región codificante de un gen, en los exones. Así pues, los SNPs exónicos suponen una alteración en el codón del gen y dependiendo del código genético (Figura 5) puede suponer un cambio en la estructura de la proteína o no. Se pueden distinguir tres tipos diferentes de SNPs codificantes:

- Cambios sinónimos: el cambio de nucleótido no supone ningún cambió en el aminoácido codificado en la proteína.
- Cambios no-sinónimos de cambio de sentido o 'missense': el cambio de nucleótido supone un cambio en el aminoácido codificado en la proteína.
- Cambios no-sinónimos de parada o 'nonsense': el cambio de nucleótido cambia el codón codificante de aminoácido por un codón de parada de la traducción.

Primera posición del codón	Segunda posición del codón: A		C		G		U		Tercera posición del codón
A	AAA	Lys	ACA	Thr	AGA	Arg	AUA	Ile	A
	AAG	Lys	ACG	Thr	AGG	Arg	AUG	**Met**	G
	AAC	Asn	ACC	Thr	AGC	Ser	AUC	Ile	C
	AAU	Asn	ACU	Thr	AGU	Ser	AUU	Ile	U
C	CAA	Gln	CCA	Pro	CGA	Arg	CUA	Leu	A
	CAG	Gln	CCG	Pro	CGG	Arg	CUG	Leu	G
	CAC	His	CCC	Pro	CGC	Arg	CUC	Leu	C
	CAU	His	CCU	Pro	CGU	Arg	CUU	Leu	U
G	GAA	Glu	GCA	Ala	GGA	Gly	GUA	Val	A
	GAG	Glu	GCG	Ala	GGG	Gly	GUG	Val	G
	GAC	Asp	GCC	Ala	GGC	Gly	GUC	Val	C
	GAU	Asp	GCU	Ala	GGU	Gly	GUU	Val	U
U	UAA	**STOP**	UCA	Ser	UGA	**STOP**	UUA	Leu	A
	UAG	**STOP**	UCG	Ser	UGG	Trp	UUG	Leu	G
	UAC	Tyr	UCC	Ser	UGC	Cys	UUC	Phe	C
	UAU	Tyr	UCU	Ser	UGU	Cys	UUU	Phe	U

Aminoácidos polares
Aminoácidos no polares
Codón de inicio
Codón de finalización

Figura 5. Representación del código genético que determina la traducción de ADN a aminoácidos. Ala = alanina; Arg = arginina; Asn = asparraguina; Asp = ácido aspártico; Cys = Cisteína; Gln = glutamina; Glu = ácido glutámico; Gly = glicina; His = histidina; Ile = isoleucina; Leu = leucina; Lys = lisina; Met = metionina; Phe = fenilalanina; Pro = prolina; Ser = serina; Thr = treonina; Trp = triptófano; Tyr = tirosina; Val = valina; STOP = codón de finalización.

SNPs no codificantes

Cuando el SNP está localizado en una región genética no codificante para proteína (intrones, regiones 5'-UTR o 3'-UTR, pormotores, etc.) se conoce como SNP no-codificante ('non-coding' SNP). Los SNPs no-codificantes, pese a no tener una implicación directa en la codificación del gen, pueden tener implicaciones a nivel de transcripción de genes, llegando incluso a afectar a los niveles de proteína de los mismos, al afectar otras estructuras genéticas como son promotores, separadores, regiones de unión a factores de trascripción, etc. Por ejemplo, un SNP intrónico puede tener implicaciones a nivel estructural de la proteína, mediante la creación de isoformas de la misma como resultado de cambiar el patrón de procesamiento o 'splicing'[39]. Asimismo, también es posible que los SNPs no-codificantes no supongan ningún cambio a nivel biológico.

Genotipado de polimorfismos

El genotipado de los SNPs se puede hacer mediante diversos métodos bioquímicos incluidos en las técnicas de secuenciación de nueva generación (NGS, del inglés 'new generation sequencing'). Estos métodos se basan en la hibridación de sondas marcadas, normalmente mediante fluorescencia, con el ADN genómico del individuo a genotipar. Midiendo el marcaje de cada sonda se puede determinar la carga alélica de cada SNP en el individuo.

En la actualidad, existen 'arrays' o chips comerciales con los que se pueden identificar desde cientos de miles a millones de SNPs. Los SNPs incluidos en estos chips suelen ser 'tagSNPs', los cuales representan la distribución alélica de un haplotipo (Figura 4). Utilizando un único tagSNP por haplotipo es posible obtener una representación de la variabilidad genética del genoma al completo sin tener que incluir en el chip los millones de SNPs descritos en la totalidad del genoma humano. Esta técnica de genotipado es de gran utilidad para los estudios genómicos de asociación.

Tipos de estudios genéticos

El objetivo principal de todo estudio genético es describir la influencia de la genética, si es posible a nivel funcional, sobre el fenotipo de interés estudiado. Así pues, una de las aplicaciones que tienen los estudios genéticos es determinar los mecanismos moleculares causantes de enfermedades humanas, excluyendo aquellas que no tienen características heredables.

Se pueden distinguir dos tipos de enfermedades con base genética: las enfermedades monogénicas y las enfermedades poligénicas, también llamadas enfermedades complejas. Una enfermedad monogénica es aquella causada por la presencia de una o varias mutaciones en un único gen. Normalmente estas enfermedades suelen tener una frecuencia

baja en la población (menor al 5%) por lo que se clasifican como enfermedades raras o infrecuentes. Por otra parte, las enfermedades complejas tienen una causa multifactorial, es decir, es la suma de más de un factor lo que finalmente acaba causando la enfermedad. Se pueden distinguir diferentes tipos de factores de riesgo: factores clínicos y demográficos, factores ambientales y factores genéticos. Cada uno de estos factores aporta una pequeña parte del riesgo a padecer la enfermedad y suelen implicar diversos genes. Entre las enfermedades complejas se encuentran varias de las complicaciones médicas más frecuentes en los países desarrollados, como son la diabetes tipo II, las enfermedades cardiovasculares, o la obesidad.

Los estudios genéticos se pueden dividir en dos grades categorías: los estudios de gen candidato, y los estudios de genoma completo.

Estudios de gen candidato

Los estudios de gen candidato se basan en seleccionar uno o varios genes interesantes en relación al fenotipo estudiado, normalmente en base a literatura previa, y estudiarlos mediante técnicas de genética molecular (PCR, hibridación *in situ*, etc.) y/o ingeniería genética (knock-out, knock-in, etc.). Normalmente, en estos estudios se seleccionan uno o varios SNPs pertenecientes al gen de interés y se analizan las diferencias en la frecuencia alélica entre casos y controles. Asimismo, también se pueden realizar análisis funcionales adicionales en modelos animales o cultivos celulares para validar la asociación encontrada entre gen y fenotipo, aunque éstos pueden presentar dificultades en su translacionalidad a seres humanos, debido a las diferencias genéticas entre especies o a la falta de cambios sistémicos, respectivamente.

No obstante, pese a que estos estudios pueden aportar mucha información de la funcionalidad e implicación del gen en la enfermedad de interés, tienen la limitación de depender del conocimiento previo existente. Asimismo, antes de que se extendieran las nuevas técnicas de secuenciación y se iniciaran los grandes proyectos de genoma humano, únicamente se podían realizar estudios de gen candidato, los cuales resultaron poco eficientes de cara a estudiar las enfermedades complejas e imposibles de aplicar en enfermedades monogénicas si no se conocía el gen responsable.

Estudios de genoma completo

Este tipo de estudio, a diferencia de los estudios de gen candidato, no priorizan ningún gen en concreto y analizan el genoma en su totalidad. De entre este tipo de estudios, los más utilizados en genética humana son los estudios de asociación de genoma completo, también conocidos como GWAS (de sus siglas en inglés 'Genome-Wide Association Study').

Los estudios GWAS se basan en analizar miles de SNPs representativos de todo el genoma para buscar polimorfismos asociados al fenotipo de interés. En este tipo de estudios es posible analizar hasta 15 millones de SNPs a partir de los SNPs genotipados en un chip comercial gracias a los métodos informáticos de imputación. Sin embargo, es necesario realizar diversos controles de calidad bioinformáticos[40,41] para evitar los falsos positivos en el análisis de asociación, por lo que también requieren personal especializado. Entre esto controles de calidad se incluyen:

- Calidad de genotipado: se excluyen aquellos SNPs que no están genotipados en un 10% o más de la población de estudio. Para garantizar un número mínimo de pacientes incluidos en el análisis de asociación posterior.
- Calidad de imputación: se excluyen aquellos SNPs que la probabilidad de que la imputación sea correcta es inferior a un 30%, esto se traduce en un valor de r^2 menor a 0.3.
- Frecuencia de los polimorfismos: se excluyen aquellos polimorfismos que tiene una frecuencia inferior al 1% en la población de estudio. Estos polimorfismos se categorizan como variantes raras o infrecuentes y se analizan con algoritmos diferentes a los de las variables comunes.
- Sexo discordante: se excluyen los individuos que presentan una discordancia de sexo al comparar los datos fenotípicos descritos en la base de datos con los observados al genotipar los cromosomas sexuales. Esto es indicativo de una incorrecta genotipación o de un error en la anotación de los datos o en el procesamiento de las muestras.
- Calidad de genotipado por individuo: se excluyen aquellos individuos que presentan más de un 10% de los SNPs no genotipados. En este caso se considera que el proceso de genotipación no se ha realizado correctamente y los datos no son fiables.
- Distribución étnica: se realizan análisis informáticos para determinar la distribución étnica de la población de estudio, si está no es uniforme se seleccionan únicamente los individuos agrupados en la etnia de interés a estudiar.

Los resultados de los estudios de GWAS se suelen representa gráficamente en los llamados “Manhattan plot” (Figura 6). En estos gráficos los SNPs se representan como puntos donde su valor en el eje horizontal es la posición en el cromosoma y el valor en el eje vertical es el valor de asociación de cada SNP con el fenotipo, representado por el valor estadístico “*p*” en una escala logarítmica invertida (cuanta más asociación menor valor de *p* y mayor valor en el eje Y).

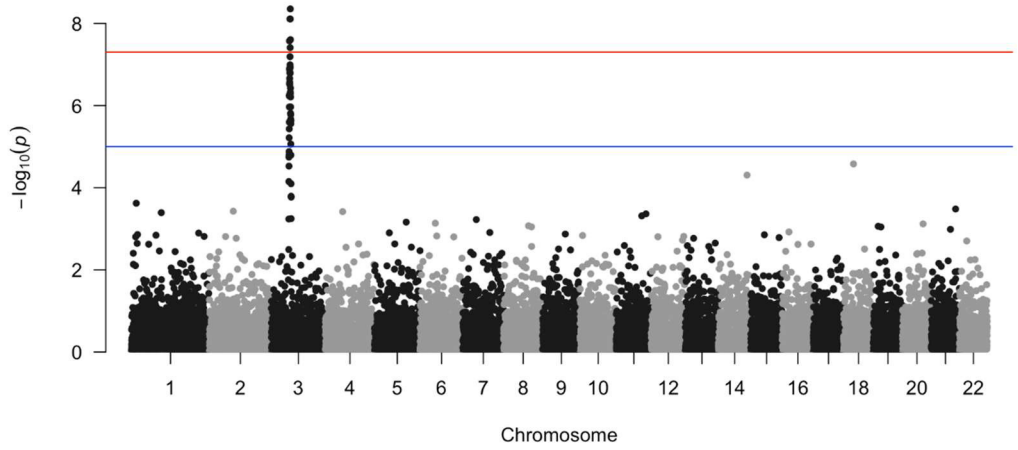

Figura 6. Gráfico "Manhattan". Realizado en el programa informático R con el repositorio de datos "gwasResutls". La línea roja representa el umbral de significancia de los estudios GWAS ($p < 5x10^{-8}$), y la línea azul representa el umbral sugestivo ($p < 1x10^{-5}$).

De cara a desvelar los factores genéticos que influyen en las enfermedades complejas, los estudios GWAS presentan diversas ventajas. Por una parte, no hay problemas de translacionalidad puesto que la observación se realiza directamente en población humana. También presentan la ventaja de ser estudios no sesgados, ya que se analiza al completo el genoma. Finalmente, al ser la genética una variable no modificable y obtener los resultados de ensayos no sesgados, en los GWAS se reduce la posibilidad de obtener falsos positivos, aumentando la reproducibilidad de este tipo de estudios.

No obstante, debido a la gran variabilidad genética entre poblaciones y al gran número de polimorfismos estudiados, los estudios GWAS suelen presentar la desventaja de necesitar un gran número de participantes para obtener resultados estadísticamente significativos; además de requerir personal cualificado para el manejo de los datos.

1.3. La genética en el ictus

El ictus es una enfermedad compleja en la cual intervienen diversos factores en su patología, tanto ambientales como genéticos. Asimismo, esta enfermedad posee cierto grado de heredabilidad. En el estudio realizado por Bevan et al. (2012)[42] se estimó que la heredabilidad del ictus era de entre un 16% y un 40%. Debido a ello, durante los últimos años se han realizado diversos estudios genéticos masivos con el objetivo de determinar los factores genéticos responsables tanto del riesgo de ictus como del pronóstico del mismo.

Factores de riesgo genéticos en el ictus

En los últimos años se han realizado varios estudios genéticos masivos internacionales, los cuales han dado como resultado más de 30 loci relacionados con el riesgo de ictus. De entre estos, los más importantes a destacar son los estudios SiGN[43], METASTROKE[44], MEGASTROKE[45] y MEGASTROKE-UKBiobank[46].

METASTROKE

En marzo de 2012 se publicó el estudio "METASTROKE". Este estudio se caracterizó por la realización de un meta-análisis de 12 GWAS a modo de cohorte de descubrimiento, con lo que se incluían 10.307 casos de ictus y 19.326 controles, y posteriormente realizar una replicación de los resultados en dos cohortes externas, una de etnia caucásica (con 13.435 casos y 29.269 controles) y otra de etnia sur-asiática (con 2.385 casos y 5.193 controles).

En este estudio se describieron cuatro loci significativos: *ABO* (para ictus, tanto hemorrágico como isquémico), *HDAC9* (para ictus isquémico aterotrombótico), *PITX2* y *ZFHX3* (para ictus isquémico cardioembólico). Asimismo, este estudio también se reportó que en estos loci había gran cantidad de SNPs poco frecuentes (frecuencia menor al 5%) los cuales podían justificar la baja heredabilidad de algunos subtipos de ictus. Así pues, el estudio METASTROKE puso de manifiesto la necesidad de utilizar grandes cohortes para encontrar los loci asociados al ictus.

SiGN

El estudio SiGN, fue realizado por el "National Institute of Neurological Disorders Stroke Genetics Network" (NINDS-SiGN) en febrero de 2016. En este proyecto, separado en dos fases, participaron: 16.851 pacientes y 32.473 controles en la primera etapa y, 20.941 casos y 364.736 controles en la segunda etapa.

Gracias a este estudio se identificó un nuevo locus de riesgo para el ictus aterotombótico (el locus 1p13.2, próximo a *TSPAN2*) y se confirmó la asociación genética de cuatro loci previamente descritos (*PITX2* y *ZFHX3*, para ictus cardioembólico; *HDAC9* para ictus aterotrobótico; y *ALDH2*, para ictus lacunar).

MEGASTROKE

Uno de los estudios más extensos realizados hasta la fecha es el estudio "MEGASTROKE" realizado por Malik et al. (2018)[45]. En este estudio se incluyeron 521.612 participantes (67.162 casos y 454.450 controles), provenientes de diferentes etnias (caucásicos, latinos, africanos, y asiáticos) con el objetivo de tener una representación de varias etnias no caucásicas.

Este estudio realizó un análisis por separado de cada una de las etnias para posteriormente realizar un estudio multiétnico, tanto para el riesgo de ictus en general como para el riesgo de cada uno de los subtipos de ictus. Mediante esta estrategia, se encontraron un total de 32 loci asociados al riesgo de esta enfermedad, de los cuales 22 eran nuevas asociaciones genéticas (Tabla 4).

MEGASTROKE-UKBiobank

El último estudio a gran escala realizado es el estudio surgido de la colaboración entre el consorcio MEGASTROKE y el UK Biobank, publicado en julio de 2018. En este estudio se incluyeron un total de 823.869 controles (775.530 europeos) y 72.147 casos de ictus (45.570 europeos), de los cuales 63.969 eran ictus isquémicos (37.845 europeos).

Puesto que los individuos del UK Biobank eran europeos, se realizó primero un análisis incluyendo únicamente esta población (tanto de UK Biobank como de MEGASTROKE) y posteriormente se realizó un análisis multiétnico utilizando las diferentes cohortes incluidas en el MEGASTROKE. Gracias a esta estrategia se encontraron tres nuevos loci asociados al riesgo de ictus (hemorrágico e isquémico): rs1799983, localizado en el gen *NOS3*; rs9521634, localizado en el gen *COL4A1*; y rs720470, cercano al gen *DYRK1A*.

Asimismo, pese a todos los estudios realizados hasta la fecha, la heredabilidad explicada por los factores genéticos encontrados es menor a la esperada, indicando que aún pueden existir más factores genéticos no descritos. La dificultad para encontrar todos estos factores de riesgo puede ser debida a las diferencias genéticas entre poblaciones.

Tabla 4. Resumen de los loci asociados con el ictus encontrados en el estudio MEGASTROKE.

SNP	Locus genómico	Gen más próximo	Localización relativa al gen	Fenotipo asociado
rs880315	1p36	*CASZ1*	Intrónico	Ictus (todos)
rs12037987	1p13	*WNT2B*	Intrónico	Ictus (todos)
rs146390073	1q43	*RGS7*	Intrónico	Cardioembólico
rs12476527	2p23	*KCNK3*	5′-UTR	Ictus (todos)
rs7610618	3q25	*TM4SF4–TM4SF1*	Intergénico	Aterotrombótico
rs34311906	4q25	*ANK2*	Intergénico	Ictus isquémico
rs17612742	4q31	*EDNRA*	Intrónico	Aterotrombótico
rs6825454	4q31	*FGA*	Intergénico	Ictus isquémico
rs11957829	5q23	*LOC100505841*	Intrónico	Ictus isquémico
rs6891174	5q35	*NKX2-5*	Intergénico	Cardioembólico
rs16896398	6p21	*SLC22A7–ZNF318*	Intergénico	Ictus (todos)
rs42039	7q21	*CDK6*	3′-UTR	Ictus isquémico
rs7859727	9p21	*Chr9p21*	ncRNA intrónico	Ictus (todos)
rs10820405	9q31	*LINC01492*	ncRNA intrónico	Aterotrombótico
rs2295786	10q24	*SH3PXD2A*	Intergénico	Icuts (todos)
rs7304841	12p12	*PDE3A*	Intrónico	Ictus isquémico
rs35436	12q24	*TBX3*	Intergénico	Ictus (todos)
rs9526212	13q14	*LRCH1*	Intrónico	Ictus (todos)
rs4932370	15q26	*FURIN–FES*	Intergénico	Ictus isquémico
rs11867415	17p13	*PRPF8*	Intrónico	Ictus isquémico
rs2229383	19p13	*ILF3–SLC44A2*	Exónico sinónimo	Ictus isquémico
rs8103309	19p13	*SMARCA4–LDLR*	Intergénico	Ictus (todos)
rs12124533	1p13	*TSPAN2*	Intergénico	Aterotrombótico
rs1052053	1q22	*PMF1–SEMA4A*	Exónico no sinónimo	Ictus (todos)
rs13143308	4q25	*PITX2*	Intergénico	Cardioembólico
rs4959130	6p25	*FOXF2*	Intergénico	Ictus (todos)
rs2107595	7p21	*HDAC9–TWIST1*	Intergénico	Aterotrombótico
rs635634	9q34	*ABO*	Intergénico	Ictus isquémico
rs2005108	11q22	*MMP12*	Intergénico	Ictus isquémico
rs3184504	12q24	*SH2B3*	Exónico no sinónimo	Ictus isquémico
rs12932445	16q22	*ZFHX3*	Intrónico	Cardioembólico
rs12445022	16q24	*ZCCHC14*	Intergénico	Ictus (todos)

Es posible que un factor de riesgo genético mucho más presente en una población específica no sea detectable mediante estudios GWAS multiétnicos, pero si cuando se realiza el estudio en dicha población, aumentando la frecuencia de este factor y su posible detección. De hecho, este escenario ha sido descrito anteriormente para población australiana en ictus isquémico aterotrombótico. En el estudio de Holliday et al. (2012)[47], los autores realizaron un GWAS en población australiana, incluyendo 1.244 controles y 1.162 casos de ictus isquémico. Sin embargo, cuando se analizó el riesgo de ictus por subtipos los autores encontraron asociaciones significativas, pese a disminuir el número de casos (421 ictus aterotrombóticos), y pudieron replicar sus resultados en una cohorte internacional. Esto puede ser debido a que al analizar un único subtipo de ictus se disminuye la heterogeneidad de la muestra y se aumenta el poder estadístico.

En base a este estudio se plantea una nueva estrategia para la búsqueda de nuevos factores de riesgo genéticos: analizar en primer lugar una cohorte geográficamente homogénea y posteriormente validar los resultados en una población internacional. Puesto que no existe hasta la fecha ningún análisis genético de riesgo de ictus en población hispano-ibérica, la población española resulta un buen candidato para encontrar nuevos factores genéticos asociados con el ictus.

Factores genéticos en el pronóstico del ictus

Existen diversos estudios de gen candidato, la mayoría con estudios funcionales en ratón como modelo animal, que han señalado a varios genes como posibles responsables del empeoramiento del pronóstico de los pacientes tras un ictus (Tabla 5). Sin embargo, pese a presentar resultados muy prometedores la translacionalidad a humanos no ha sido posible.

Recientemente, se han publicado estudios GWAS[48,49] enfocados en el pronóstico funcional de los pacientes de ictus isquémico a los 3 meses, medido con la escala mRS. Estos estudios han encontrado dos loci asociados al estado funcional del paciente al tercer mes, demostrando la influencia de la genética en la evolución de los pacientes y en su grado de discapacidad.

Tabla 5. Tabla resumen de los genes candidatos estudiados par a el pronóstico del ictus.
Tabla adaptada de Lindgren & Maguire (2016)[50].

Gen	Fenotipo asociado	Referencia
COX-1	Resultado vascular, mortalidad	Cao et al.[51]
NINJ2	Recurrencia en ictus aterotrombótico	Zhang et al.[52]
TLR4	Resultado neurológico	Weinstein et al.[53]
COL3A1	Recurrencia, pronóstico, mortalidad	Lv et al.[54]
GPIIIa	Ictus, infarto de miocardio, muerte ; Índice Barthel	Yeh et al.[55] ; Maguire et al.[56]
PAI-1	Ictus, infarto de miocardio, muerte	Yeh et al.[55]
Factor VII	Ictus, infarto de miocardio, muerte ; mortalidad post-ictus	Yeh et al.[55] ; Heywood et al.[57]
MTHFR	Ictus, infarto de miocardio, muerte	Yeh et al.[55]
eNOS	Ictus, infarto de miocardio, muerte	Yeh et al.[55]
APOE	Muerte temprana por ictus ; impedimento neurológico al año, discapacidad funcional severa, y dependencia ; pronóstico al año ; Recuperación al mes post-ictus isquémico y escala de Rankin modificada al 3r mes post-ictus isquémico	Gromadzka et al.[58] ; Gromadzka et al.[59] ; Sarzynska-Dlugosz et al.[60] ; Cramer et al.[61]
GPIIb	Mortalidad post-ictus	Carter et al.[62]
IGF-1	Ocurrencia, severidad, y estado funcional	Aberg et al.[63]
MPO	Riesgo de infarto cerebral, estado funcional a corto plazo	Hoy et al.[64]
SIGMAR1	Estado al 3r mes	Lovkvist et al.[65]
APOD	Estado al 3r mes	Lovkvist et al.[65]
COX-2	Escala Glasgow ; escala de Rankin modificada	Maguire et al.[56]
BDNF	Recuperación motora post-ictus ; plasticidad del córtex motor en el ictus agudo ; parámetros clínicos y resultado funcional en pacientes con ictus (isquémico y hemorragia intracraneal) ; recuperación al 1r mes post-ictus y 3r mes post-ictus ; 2-semanas post-ictus y 1-año post-ictus	Chang et al.[66] ; Di Lazzaro et al.[67] ; Mirowska-Guzel et al.[68] ; Cramer et al.[61] ; Kim et al.[69]
CYPC19	Estado clínico en pacientes de ictus isquémico tratados con copidogrel, escala de Rankin modificada	Qiu et al.[70]
CRP	Capacidad funcional	Guo et al.[71]
COMT	Índice Barthel y evaluación motora de Rivermead en admisión, tras 4 semanas y a los 6 meses	Liepert et al.[72]
SERT	Depresión post-ictus	Kohen et al.[73]

GISCOME

El estudio GISCOME[49] ('Genetics of Ischemic Stroke Functional Outcome') utilizó los datos de 12 proyectos internacionales que disponían de datos genéticos para realizar un análisis de asociación en relación al pronóstico del paciente. Así pues, en este estudio se incluyeron un total de 6.165 individuos y se analizó la asociación genética con el pronóstico funcional a los 60-190 días medido con la escala mRS dicotomizada en 0-2 y 3-6. Este estudio identificó un único locus asociado al pronóstico funcional del paciente (rs1842681), el cual se encuentra en el gen *LOC105372028* y está descrito como un eQTL de *PPP1R21* ('Protein Phosphatase 1 Regulatory Subunit 21'). No obstante, en este estudio no ser realizó una replicación de los resultados y presentaba una considerable heterogeneidad en relación a las cohortes utilizadas y a los datos clínicos disponibles.

GODS

En el estudio GODS[48] ('Genetic controbution to functional Outcome and Disability after Stroke'), el total de muestras analizadas (n=2.482) era inferior al estudio GISCOME, pero los criterios de inclusión utilizados eran más estrictos con el objetivo de homogeneizar mejor la cohorte utilizada: se excluyeron pacientes con un mRS previo superior a 3 y pacientes que presentaron una recurrencia vascular en los primeros 3 meses tras sufrir el ictus; ambos factores podían suponer un sesgo importante para el estudio. En este proyecto se describió un nuevo locus asociado al pronóstico del paciente a los 90 días post-ictus isquémico, validado mediante una replicación de los resultados en una cohorte internacional de 1.257 pacientes. Además, a nivel funcional, se describió que el locus candidato estaba situado en el gen *PATJ y* presentaba una asociación con la expresión de dicho gen en tejido humano. Esta primera aproximación al estudio funcional del locus, abre las puertas para la comprensión a nivel molecular del pronóstico del paciente tras un ictus isquémico.

Pese a que los GWAS realizados han utilizado cohortes con un gran número de pacientes, el número de loci encontrados es muy reducido. Esto puede ser debido al origen multifactorial del pronóstico del ictus, puesto que puede dificultar el conseguir una cohorte homogénea donde realizar el análisis genético de asociación y, por lo tanto, reducir el poder estadístico del análisis. Una posible estrategia para solventar este problema es analizar de manera individual cada uno de los factores asociados al pronóstico del ictus (endofenotipos), disminuyendo así la heterogeneidad y facilitando el descubrimiento de nuevos factores de riesgo genéticos.

Perspectivas de futuro

El objetivo principal de estudiar la genética en las enfermedades humanas es poder aproximarnos, a largo plazo, a una medicina personalizada, en la cual se tenga en cuenta la predisposición genética única del paciente en el riesgo y la progresión de las enfermedades. Así pues, en el caso del ictus, se podría utilizar la genética como un biomarcador para detectar la etiología del ictus o posibles complicaciones médicas, con el fin de mejorar el tratamiento de la enfermedad y el pronóstico del paciente.

En un principio, la genética se considera un factor no modificable, puesto que es invariable en el individuo desde su nacimiento. Sin embargo, gracias a la farmacología, existen fármacos capaces de alterar la función de las proteínas codificadas por la genética y así alterar este factor no modificable. Además, en los últimos años también se han desarrollado técnicas de edición genética que permite realizar modificaciones en el código genético de los individuos, la llamada terapia génica.

Una primera aproximación a la medicina personalizada es la utilización de los polimorfismos de riesgo en los scores de predicción. En el caso del ictus, recientemente se ha publicado un estudio para determinar el riesgo de sangrado de los pacientes tratados con rtPA tras sufrir un ictus isquémico[74].

Otra herramienta útil para la medicina es el uso de la genética en el reposicionamiento de fármacos. Al detectar un posible gen o proteína responsable de la enfermedad, es posible realizar una búsqueda *in silico* de fármacos previamente aprobados que puedan tener efecto sobre dicho gen o proteína y, por lo tanto, tratar la enfermedad. En referencia al ictus isquémico, el consorcio del MEGASTROKE realizó un estudio de reposicionamiento de fármacos para detectar posibles medicamentos que afectaran a los genes candidatos encontrados y así poder utilizarlos como fármacos preventivos de ictus isquémico. Mediante esta estrategia se encontraron 2 genes cuyas proteínas codificadas, potencialmente, podrían ser moduladas por fármacos utilizados en la práctica clínica habitual.

Por otra parte, la implementación de otros análisis masivos, además de los GWAS, hacen posible aportar más información sobre las proteínas y rutas metabólicas implicadas en el ictus y con ello mejorar el tratamiento y pronóstico de esta enfermedad.

Un ejemplo es el análisis de Aleatorización Mendeliana (o MR, del inglés 'Mendelian Randomization'). Este análisis utiliza la información genética para determinar si la asociación encontrada entre dos fenotipos es de causa o efecto. Por ejemplo, la MR puede determinar si la asociación encontrada entre niveles elevados de una proteína y la ocurrencia del ictus es causal o no. Para ello se observa si los factores genéticos que determinan los niveles de dicha proteína también se asocian a la incidencia de ictus y viceversa, si los factores

genéticos de riesgo de ictus se asocian a los niveles de proteína. Si se da el primer caso, la asociación sería causal (niveles elevados de la proteína causarían ictus) mientras que, si se da el segundo caso, la asociación sería de efecto (el hecho de tener un ictus elevaría los niveles de proteína). Así pues, el análisis MR es de gran utilidad para determinar la causalidad de determinados factores. Actualmente, mediante este tipo de análisis se ha demostrado la causalidad en referencia al riesgo de ictus del tabaco[75], la diabetes tipo II[76] y el colesterol en sangre[77].

Por otra parte, estudios recientes han demostrado la implicación de la epigenética en diversas enfermedades, incluyendo la recurrencia vascular después de un ictus[78], y la edad biológica en el riesgo de ictus[79]. El análisis más ampliamente utilizado para estudiar la epigenética son los análisis epigenéticos masivos o EWAS (del inglés 'Epigenome-Wide Association Study'). En los EWAS se miden los niveles de metilación de las islas CpGs presentes por todo el genoma y se determina si estos niveles se asocian o no al fenotipo de interés, por lo que presentan las mismas ventajas que los GWAS. Así pues, los EWAS serían de utilidad para determinar otros componentes implicados en el ictus[80], no detectables por GWAS.

Finalmente, es importante destacar la necesidad de realizar análisis de integrómica de datos. En los análisis de integrómica, se utilizan los diferentes resultados obtenidos en estrategias "omicas" (genómica, exómica, proteómica, etc.) para determinar genes, proteínas y/o rutas metabólicas comunes entre los diferentes análisis. Así pues, este tipo de análisis es útil para para discernir los falsos positivos de asociaciones reales y aporta mayor información de las posibles rutas metabólicas implicadas en el fenotipo estudiado. Hasta la fecha no se ha realizado ningún análisis de integrómica en ictus, puesto que aún es necesario realizar más estudios "ómicos".

OBJETIVOS

2. Objetivos

Teniendo en cuenta el escenario descrito, se plantearon las siguientes hipótesis: 1) Existen factores genéticos de riesgo de ictus específicos de población española. Un análisis mediante herramientas genéticas masivas nos permitirá descubrir estos factores de riesgo genéticos; y 2) El estudio genético de endofenotipos asociados con el deterioro neurológico durante la fase aguda del ictus nos permitirá encontrar factores de riesgo genéticos asociados con la evolución neurológica y funcional de los pacientes. De igual manera, en base a estas hipótesis se plantearon los siguientes objetivos para la presente Tesis:

1. Encontrar nuevos factores de riesgo genéticos asociados al ictus isquémico en población española y validar estos polimorfismos de riesgo en población internacional.
2. Realizar estudios funcionales para determinar el efecto de los factores genéticos asociados a las variables de estudio.
3. Realizar una revisión bibliográfica de los factores clínicos, demográficos y genéticos asociados con la evolución neurológica y funcional del paciente durante la fase aguda del ictus isquémico para identificar potenciales endofenotipos.
4. Encontrar factores de riesgo genéticos asociados con el pronóstico del ictus isquémico mediante una estrategia de análisis de endofenotipos, centrado en las células del sistema inmune.

INFORME DEL DIRECTOR

3. Informe del Director de Tesis

La doctoranda Nuria Paz Torres Águila presenta la memoria de la Tesis Doctoral titulada "Estudios genéticos masivos en el ictus isquémico: factores de riesgo y pronóstico". Su director de Tesis, el Dr. Israel Fernández Cadenas, informa que la tesis doctoral está compuesta por dos artículos aceptados para su publicación y un manuscrito original en fase de revisión por parte de la revista. En todos estos artículos la doctoranda figura como primera autora.

Hago constar que ambos artículos de la presente Tesis han sido aceptados en revistas internacionales de prestigio para las correspondientes áreas de conocimiento, en las cuales se incluyen Neurología clínica y Enfermedades vasculares periféricas. Todas las revistas donde la doctoranda ha publicado sus trabajos se encuentran incluidas en 'InCites Journal Citation Reports' y, en todos los casos, se trata de publicaciones que han pasado el filtro de evaluadores anónimos designados por los editores de las revistas.

A continuación, se detalla la contribución científica de la doctoranda en cada uno de los artículos, así como el factor de impacto de las revistas ('2018 Journal Impact Factor') en las que han sido publicados.

Artículo 1: "A genome-wide association study reveals a new locus in *MAN2B1* gene for lacunar stroke risk" Nuria P Torres-Aguila, Caty Carrera, Elena Muiño, Natalia Cullell, Jara Cárcel-Márquez, Jonathan González-Sánchez, Cristina Gallego-Fabrega, Anna Penalba-Morenilla. Alejandro Bustamante, Jesús Pizarro, Marimar Freijo, Juan F Arenillas, Victor Obach, José Álbarez-Sabín, Carlos Molina, Marc Ribó, Jordi Jiménez-Conde, Jaume Roquer, Tomás Sobrino, Francisco Campos, José Castillo, Lucia Muñoz-Narbona, Elena Lopez-Cancio, Antoni Dávalos, Rosa Diaz-Navarro, Silvia Tur, Cristòfil Vives-Bauza, Gemma Serrrano-Heras, Tomás Segura, Laura Ibañez, Laura Heitshc, Jerzy Krupinski, Joan Martí-Fàbregas, Raquel Delgado-Mederos, Stephanie Debette, Martin Dichgans, Rainer Malik, Rafael de Cid, Lauro Sumoy, Victor Moreno, Carlos Cruchaga, Jin Moo-Lee, Joan Montaner, Israel Fernandez-Cadenas. On behalf of the MEGASTROKE consortium, the GeneStroke consortium, the International Stroke Genetics Consortium (ISGC) and the Acute Endophenotypes group of the ISGC. ***Neurology*** (Manuscrito en fase de Revisión)

- Factor de impacto: 8.689
- Posición en el área: 10/199, primer decil y primer cuartil (Q1) en el área de Neurología Clínica.

En este artículo la doctoranda fue la principal autora del trabajo, siendo responsable de la realización de todos los análisis genéticos incluidos en el artículo y de la realización y

posterior análisis de los ensayos funcionales. Concretamente, la doctoranda realizó los análisis GWAS de descubrimiento y replicación realizados en las cohortes españolas para ictus isquémico y para cada uno de los subtipos (aterotrombóticos, cardioembólicos y lacunares). También, validó/replicó el locus encontrado en los datos administrados por el consorcio MEGASTROKE y realizó análisis de asociación a eQTLs para determinar posibles genes candidatos. Asimismo, tras identificar a *MAN2B1* como principal candidato, realizó los ensayos enzimáticos de α-manosidasa en muestras de plasma de pacientes y controles. Finalmente, la doctoranda tuvo un papel principal en la interpretación y discusión de los resultados, así como en la elaboración de las figuras y tablas. La redacción del manuscrito fue realizada por la doctoranda, así como la revisión del manuscrito tras las sugerencias de los otros autores.

Artículo 2: "Clinical variables and genetic risk factors associated with the acute outcome of ischemic stroke. A systematic review." Nuria P Torres-Aguila, Caty Carrera, Elena Muiño, Natalia Cullell, Jara Cárcel-Márquez, Cristina Gallego-Fabrega, Jonathan González-Sánchez, Alejandro Bustamante, Pilar Delgado, Laura Ibañez, Laura Heitsch, Jerzy Krupinski, Joan Montaner, Joan Martí-Fàbregas, Carlos Cruchaga, Jin-Moo Lee, Israel Fernandez-Cadenas. Acute endophenotypes group of the International Stroke Genetics Consortium (ISGC). ***Journal of Stroke***. (Aceptado)

- Factor de impacto: 5.571
- Posición en el área: 24/199, primer cuartil (Q1) en el área de Neurología Clínica; 7/65, primer decil y primer cuartil (Q1) en el área de Enfermedades periféricas vasculares.

La doctoranda fue la principal autora de este trabajo. Se encargó de la revisión bibliográfica, de establecer los criterios de inclusión para la revisión, de la redacción del manuscrito y de la elaboración de las figuras y tablas incluidas. También, tras recibir los comentarios de la revisión por parte de la revista, se encargó de realizar las modificaciones correspondientes y de redactar tanto la carta al editor como la respuesta a los revisores.

Artículo 3: "Genome-wide association study of white blood cell counts in ischemic stroke patients" Nuria P Torres-Aguila, Caty Carrera, Anne-Katrine Gise, Natalia Cullell, Elena Muiño, Jara Cárcel-Márquez, Cristina Gallego-Fabrega, Jonathan González-Sánchez, Marimar Freijo, José Álvarez-Sabín, Carlos Molina, Marc Ribó, Jordi Jimenez.Conde, Jaume Roquer, Tomás Sobrino, Francisco Campos, José Castillo, Lucia Muñoz-Narbona, Elena Lopez-Cancio, Antoni Dàvalos, Rosa Diaz-Navarro, Silvia Tur, Cristófil Vives-Bauza, Gemma Serrano-Heras, Tomás Segura, Jerzy Krupinski, Raquel Delgado-Mederos, Joan Martí-Fàbergas, Laura Heitsch, Laura Ibañez, Carlos Cruchaga, Natalia S Rost, Joan Montaner, Jin-Moo Lee, Israel Fernandez-Cadenas. ***Stroke***. (Aceptado)

- Factor de impacto: 6.046
- Posición en el área: 20/199, primer cuartil (Q1) en el área de Neurología Clínica; 5/65, primer decil y primer cuartil (Q1) en el área de Enfermedades periféricas vasculares.

En trabajo se realizó con la colaboración del grupo de investigación de la Dr. Natalia S Rost del Massachusetts General Hospital (USA). Asimismo, la doctoranda fue la principal autora de este trabajo puesto que estableció la estrategia de análisis del estudio y realizó el análisis GWAS de descubrimiento. También, se encargó de validar la replicación en los datos proporcionados por el grupo de la Dr. Natalia S Rost y de realizar los posteriores análisis *in silico* del locus encontrado. Asimismo, la doctoranda fue responsable de la redacción del manuscrito y de la elaboración de las tablas y figuras incluidas; además, realizó los cambios y modificaciones recomendados por los revisores de la revista y redactó la correspondiente carta al editor y la respuesta a los revisores.

Firma,

Dr. Israel Fernández Cadenas

PUBLICACIONES

4. Publicaciones

4.1. Artículo 1

A genome-wide association study reveals a new locus in *MAN2B1* gene for lacunar stroke risk

RESUMEN

Objetivo: Nuestro objetivo es explorar si existen nuevos factores de riesgo genético de ictus en poblaciones geográficamente homogéneas.

Métodos: Realizamos un estudio caso-control de asociación de genoma completo en tres etapas en una cohorte española de 1.752 casos de ictus isquémico y 5.227 controles. Para la replicación utilizamos una cohorte internacional (MEGASTROKE, n = 521.612) y una cohorte española independiente (n = 1.720). La clasificación etiológica se realizó siguiendo los criterios de TOAST. También realizamos test SMR-HEIDI, análisis de qRT-PCR y ensayos de actividad de α-manosidasa.

Resultados: Encontramos un locus asociado con riesgo de ictus isquémico, ubicado en el gen *MAN2B1* con 20 polimorfismos sugestivos (p <1x10-5; polimorfismo de menor p: rs34324185; p = 4.63x10-8; odds ratio = 1.38 [1.28-1.68]). En la replicación, rs34324185 y rs8107196, los polimorfismos significativos del análisis GWAS, se replicaron en el subanálisis de ictus lacunar (rs34324185: MEGASTROKE: p = 5.90x10-3, odds ratio = 1.06 [1.02-1.12]; Cohorte de replicación 2: p = 0.04, odds ratio = 1.59 [1.01-2.49]). Los *análisis in silico* revelaron que los alelos de riesgo del locus se asociaron con una mayor expresión de *MAN2B1* (p = 1.40x10-5), confirmada por experimentos de qRT-PCR (p = 0.04). Los ensayos enzimáticos mostraron que la actividad de α-manosidasa era mayor en pacientes con ictus lacunar en comparación con los controles (p = 0.02) y otros subtipos de ictus (p = 0.04).

Conclusiones: Hemos encontrado un nuevo locus asociado con el riesgo de ictus lacunar, que proporciona nueva información sobre los factores de riesgo y procesos biológicos asociados con este tipo de ictus. El análisis de poblaciones homogéneas de un mismo territorio podría ser una estrategia interesante para encontrar nuevos factores de riesgo genético asociados con el ictus.

A genome-wide association study reveals a new locus in *MAN2B1* gene for lacunar stroke risk

Nuria P Torres-Aguila, MSc[1,2]; Caty Carrera, MSc[1,2]; Elena Muiño, MD[1]; Natalia Cullell, MSc[1,3]; Jara Cárcel-Márquez, MSc[1]; Jonathan González-Sánchez, MSc[1,3,4]; Cristina Gallego-Fabrega, PhD[1,3]; Anna Penalba-Morenilla, BSc[2]; Alejandro Bustamante, MD, PhD[2]; Jesús Pizarro, BSc[2]; Marimar Freijo, MD[5]; Juan F Arenillas, MD,PhD[6]; Victor Obach, MD[7]; José Álvarez-Sabín, MD, PhD[8]; Carlos Molina, MD, PhD[8]; Marc Ribó, MD, PhD[8]; Jordi Jiménez-Conde, MD, PhD[9]; Jaume Roquer, MD, PhD[9]; Tomás Sobrino, PhD[10]; Francisco Campos, MD[10]; José Castillo, MD[10]; Lucia Muñoz-Narbona, PhD[11]; Elena Lopez-Cancio, MD, PhD[12]; Antoni Dávalos, MD, PhD[11]; Rosa Diaz-Navarro, MD[13]; Silvia Tur, MD[13]; Cristòfol Vives-Bauza, PhD[13]; Gemma Serrano-Heras, PhD[14]; Tomás Segura, MD[14]; Laura Ibañez, PhD[15]; Laura Heitsch, MD[16,17]; Jerzy Krupinski, MD, PhD[4,18]; Joan Martí-Fàbregas, MD, PhD[19]; Raquel Delgado-Mederos, MD, PhD[19]; Stephanie Debette, MD, PhD[20]; Martin Dichgans, MD PhD[21]; Rainer Malik, PhD[21]; Rafael de Cid, PhD[22]; Lauro Sumoy, PhD[23]; Victor Moreno, MD, PhD[24]; Carlos Cruchaga, PhD[15]; Jin Moo-Lee, MD, PhD[17]; Joan Montaner, MD, PhD[2,25]; Israel Fernandez-Cadenas, PhD[1]†. On behalf of the MEGASTROKE consortium, the GeneStroke consortium, the International Stroke Genetics Consortium (ISGC) and the Acute Endophenotypes group of the ISGC.

1. Stroke Pharmacogenomics and Genetics Laboratory, Fundació Insititut de Recerca Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
2. Neurovascular Research Laboratory, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain
3. Stroke Pharmacogenomics and Genetics Laboratory, Fundació Docència I Recerca Mútua Terrassa, Hospital Mútua de Terrassa, Terrassa (Barcelona), Spain
4. School of Healthcare Science, Manchester Metropolitan University, Manchester, United Kingdom
5. Department of Neurology, Hospital de Basurto, Bilbao, Spain
6. Stroke Unit, Department of Neurology, Hospital Clínico Universitario, University of Valladolid, Valladolid, Spain
7. Department of Neurology, Hospital Clínic i Provincial de Barcelona, Barcelona, Spain
8. Stroke Unit, Department of Neurology, Hospital Universitari Vall d'Hebron, Barcelona, Spain
9. Department of Neurology, IMIM-Hospital del Mar; Neurovascular Research Group, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques); Universitat Autònoma de Barcelona/DCEXS-Universitat Pompeu Fabra, Barcelona, Spain

10. Clinical Neurosciences Research Laboratory, Health Research Institute of Santiago de Compostela (IDIS), Santiago de Compostela, Spain
11. Department of Neurosciences, Hospital Germans Trias I Pujol, Universitat Autònoma de Barcelona, Barcelona, Spain
12. Stroke Unit, Hospital Universitario Central de Asturias, Oviedo, Spain
13. Department of Neurology, Hospital Universitari Son Espases, Institut d'Investigació Sanitaria de les Illes Balears (IdISBa), Palma de Mallorca, Spain
14. Department of Neurology, University Hospital of Albacete, Albacete, Spain
15. Department of Psychiatry, Washington University School of Medicine, Saint Louis, MO, USA
16. Division of Emergency Medicine, Washington University School of Medicine, Saint Louis, MO, USA
17. Department of Neurology, Washington University School of Medicine, Saint Louis, MO, USA
18. Neurology Service, Hospital Universitari Mútua Terrassa, Terrassa (Barcelona), Spain
19. Stroke Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
20. INSERM U1219 Bordeaux Population Health Research Center, University of Bordeaux, France
21. Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany
22. GenomesForLife-GCAT Lab Group, Program of Predictive and Personalized Medicine of Cancer (PMPPC), Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain
23. High Content Genomics and Bioinformatics Unit, Program of Predictive and Personalized Medicine of Cancer (PMPPC), Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain
24. Oncology Data Analytics Program, Catalan Institute of Oncology (ICO), IDIBELL, CIBERESP and Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain
25. Stroke Research Program, Institute of Biomedicine of Seville, IBiS/Hospital Universitario Virgen del Rocío/CSIC/University of Seville & Department of Neurology, Hospital Universitario Virgen Macarena, Seville, Spain

†Corresponding author:

Israel Fernandez-Cadenas

Stroke Pharmacogenomics and Genetics Laboratory,
Hospital de la Santa Creu I Sant Pau Research Institute,
C/ Sant Antoni Maria Claret, 167, 08025 Barcelona, Spain
Phone: +34 932 74 60 00; e-mail: israelcadenas@yahoo.es

ABSTRACT

Objective: We aim to explore whether new stroke genetic risk factors can be found in geographically homogeneous populations. **Methods:** We performed a three-stage case-control genome-wide association study in a Spanish cohort of 1,752 ischemic stroke cases and 5,227 controls. For replication we used an international cohort (MEGASTROKE, n = 521,612) and an independent Spanish cohort (n = 1,720). Etiological classification was performed following TOAST criteria. We also performed Summary-based Mendelian Randomization and HEterogeneity In Dependent Instruments tests, Real-Time Quantitative Reverse Transcription PCR analysis, and α-mannosidase activity assay. **Results:** We found a locus associated with ischemic stroke risk, located within the *MAN2B1* gene with 20 suggestive polymorphisms ($p<1x10^{-5}$; top polymorphism: rs34324185; $p = 4.63x10^{-8}$; odds ratio = 1.38[1.28-1.68]). In replication, rs34324185 and rs8107196, the genome-wide significant polymorphisms of the discovery, were replicated in lacunar stroke sub-analyses (rs34324185: MEGASTROKE: $p = 5.90x10^{-3}$, odds ratio = 1.06[1.02-1.12]; Replication cohort 2: $p = 0.04$, odds ratio = 1.59[1.01-2.49]). *In silico* analyses revealed that risk alleles of the locus were associated with higher *MAN2B1* expression ($p = 1.40x10^{-5}$), confirmed by Real-Time Quantitative Reverse Transcription PCR experiments ($p = 0.04$). Enzyme experiments showed that α-mannosidase activity was higher in lacunar stroke patients compared to controls ($p = 0.02$) and other stroke subtypes ($p = 0.04$). **Conclusions:** We have found a new locus associated with lacunar stroke risk, providing new information about the genetics behind lacunar stroke physiopathology. The analysis of specific populations could be an interesting strategy to find new genetic risk factors associated with stroke.

INTRODUCTION

Ischemic stroke (IS) is a complex disease with a substantial genetic component, the heritability of which ranges from 16% to 40%[1]. Although clinical risk factors for IS are well established, the genetic risk variants have not been completely identified. Several genome-wide association studies (GWAS) have found genes associated with the risk of stroke, which have been confirmed in independent studies[2–4]. In the MEGASTROKE[4] study, the largest and most recent genetic study on stroke with 521,612 individuals, authors found 32 loci associated with stroke risk, 22 of which were novel. Those genetic risk factors can be combined with clinical data to develop scores for predicting stroke risk or etiology, with higher accuracy compared to clinical data alone models[5–8]. Moreover, genomic approaches, as GWAS or exome sequencing, are unbiased strategies that are useful for drug discovery [4,9,10].

However, loci associated with IS only explain approximately 5-10% of the heritable risk[1]. In addition, due to

differences in the genetic background among worldwide populations, private genetic risk loci are described for geographically-specific populations[11,12]. A previous GWAS[11] found an association for large artery atherosclerosis (LAA) in an Australian cohort of 421 LAA cases and 1,244 controls. Interestingly, authors replicated the finding in an international cohort composed of different ethnicities; showing that the analysis of geographically homogenous populations and etiological sub-analysis could be useful for finding new genetic associations. We aimed to find new genetic risk factors for IS following the strategy of analyze homogeneous populations and perform the replication in large international cohorts.

METHODS

Study population

In our Discovery cohort we included 6,979 Spanish participants (5,227 controls and 1,752 IS cases). For replication analysis, we studied two different cohorts: Replication cohort 1, from MEGASTROKE consortium, which included 521,612 individuals (454,450 controls and 67,162 IS cases); and, Replication cohort 2, which included 1,720 Spanish participants (189 controls and 1,531 cases).

From the 8,699 Spanish participants included in this study, 3,783 individuals were recruited in 14 different hospitals throughout Spanish territory, and 4,916 were provided by the GCAT[13,14] project. The number of participants included from each center is detailed in the Supplement (table e-1).

The IS patients were recruited if they had a measurable neurologic deficit on the NIHSS within 6 hours of last known normal, had a stroke diagnosis performed by an experienced neurologist at each center and confirmed by neuroimaging, were older than 18 years of age, and were recruited at one of the 14 hospitals included in the study. Etiologic subgroups were classified following TOAST criteria[15]. These patients were recruited as part of the GENISIS[16], GODS[17], and CONIC[18] projects (appendix e-1).

Controls were subjects without a history of ischemic stroke, older than 18, who declared they were free of neurovascular diseases by direct interview before recruitment. The control cohort was collected in primary care centers from Barcelona city and in hospitals throughout Spanish territory as a part of the GCAT[13,14], CONIC[18], GRECOS[19], and ISSYS[20] projects (appendix e-1).All participants were genotyped with the Illumina® Human Core Exome chip with the exception of the GCAT cohort, genotyped with the Illumina® Infimum Multi-Ethnic Global

consortium array. The demographic characteristics, the prevalence of cardiovascular risk factors, and the detailed number of samples of the cohorts included on this study are detailed in Table 1.

Written informed consent was obtained from all subjects (cases and controls) with approval from the ethics committee of all participating institutions.

Genetic analyses

The same quality control (QC) pipeline was applied to the Discovery cohort and Replication cohort 2. Samples were tested for missingness (≤5%), relatedness (pi-hat>0.18), heterozygosity, and sex discrepancies. Single nucleotide polymorphisms (SNPs) with genotype call rate <95% and/or a *p* value deviating from the Hardy-Weinberg equilibrium at $p < 1\times10^{-6}$ were excluded, and to avoid strand discrepancies A/T and C/G SNPs were excluded. Mitochondrial and sexual chromosomes X and Y were also excluded, after the sex discrepancy test was performed. After the first round of QCs, cohorts were aligned with 1000G Phase 3 Panel reference and frequency alleles were compared; then cohorts were merged using PLINK[21] software. Afterwards, we performed a second round of QCs and also performed multidimensional scaling plots with principal components (PCs). PCs were used to identify and remove ethnic outliers and to adjust for population stratification in the downstream analyses. Genotype imputation was performed on Michigan Imputation Server Portal[22] using 1000G Phase 3v5 panel. We removed SNPs with r^2 < 30%, SNPs with minor allele frequency (MAF) < 1%, and insertions and deletions.

We performed the SNP association analyses using SNPTEST[23–25] software and age, sex, and principal components were used as covariates. For Linkage Disequilibrium analysis, we used SNP-clip tool from the LDlink[26] web portal. We selected the European population as reference and applied an R^2 cut-off of 0.9. The QCs of Replication cohort 1 are described on Malik et al.[4]. Briefly, SNPs were removed if they had extreme effect values ($\beta > 5$ or $\beta < -5$), MAF < 0.01, or effective allele count < 10 (calculated as the product of imputation accuracy, by number of cases, and by twice MAF).

Experimental assays

Functional annotations

We performed an *in silico* analysis using data available on GTEx portal, from The Genotype-Tissue Expression (GTEx) Project[27], to study possible eQTLs associated with our candidate SNPs. Secondly, we performed the SMR (Summary-based Mendelian Randomization) and HEIDI (HEterogeneity In Dependent Instruments)

tests[28,29] with eQTLs of whole-blood tissue (V7 release of GTEx eQTL data[27]) and brain tissue (Brain-eMeta eQTL data[30]).

Transcriptomic analysis

We performed a Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) with cDNA from 66 blood samples of healthy individuals from GRECOS[19] study, to confirm the results of *in silico* analysis. RNA extraction and cDNA synthesis were performed following the manufacture's protocol using RiboPure™-Blood kit (Ambion™) and MultiScribe™ Reverse Transcriptase (Thermo Fisher Scientific), respectively.

Relative mRNA levels were measured on a 7900 Real-Time PCR System (Applied Biosystems, Foster city, CA, USA) using a TaqMan fluorogenic probes for Mannosidase Alpha Class 2B Member 1 (*MAN2B1*, Hs01051311_m1) and normalizing with Cyclophilin A (*PPIA*, Hs99999904_m1) and Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*, Hs99999905_m1). qRT-PCR was performed using a standard TaqMan® PCR kit protocol as described previously[17]. Reactions were performed in 384-well plates; all reactions were run in triplicate and analyzed using the Applied Biosystems SDS 7900 system software (Applied Biosystems, Foster city, CA, USA). The data was analyzed as relative quantification (RQ) values and calculated as fold change expression values vs the average expression of the endogenous controls. We used an external sample from a healthy individual as a calibrator; the same calibrator was included in each experiment performed.

Enzyme activity experiments

We tested enzyme activity of α-mannosidase on 74 plasma samples with a commercial kit assay (Sigma Aldrich®) that was specifically designed to measure the activity of lysosomal α-mannosidase. We used 96-wells plates and followed the manufacture's protocol. Patients' plasma samples were collected during the acute phase of stroke (first 24h after stroke onset). Plasma samples from controls were collected as part of the GRECOS[19] study. Cases and controls were selected based on the rs34324185 genotype, in order to have an equal distribution of alleles in each group studied. For each sample analyzed, 10 µL of plasma were added to 90 µL of Reaction Substrate Buffer. After 25 min of incubation at room temperature (25°C), the reaction was stopped by adding 100 µL of Stopping Reagent. Optical Density (OD) was measured at 405 nm using a multi-plaque spectrophotometer. A commercial plasma sample (Sigma Aldrich®) was used as a positive control in the colorimetric assay and for the blank setting. Results

were expressed as U/L, where one unit is the amount of enzyme that will convert 1.0 µmole of 4-Nitrophenyl-α-D-manno-pyranoside to 4-Nitrophenol and α-D-Mannose per minute at 25 °C and pH 4.5 (the optimum pH of lysosomal α-mannosidase).

Statistical analyses of experimental assays

Statistical analyses were performed using the SPSS statistical package, version 17.0 (IBM, Chicago, US). Univariate analysis for cases-controls was evaluated by $\chi 2$ for categorical variables. A T-test was used for continuous variables, which normal distribution was assessed by Kolmogorov-Smirnov and Shapiro-Wilk tests. Outliers were determined in SPSS based on Z-scores and Inter Quartile range (IQ), calculated as Quartile3(Q3) minus Quartile1(Q1). Samples were considered as outliers and were excluded when Z-score >3 or Z-score<-3, or value>Q3+(1'5xIQ), or value<Q1-(1'5xIQ).

Sample size calculation was performed with RStudio software, version 1.0.1533, using the *epi.ccsize* function of the epiR[31] package, version 0.9-99.

Data Availability

The entire dataset, including all data used in this study, and completely anonymized, is located in a Dropbox folder and will be shared following request by qualified investigators.

RESULTS

Genetics analysis

In our Discovery cohort, we included 6,974 participants (1,752 IS cases and 5,222 controls), and from the 263,277 genotyped SNPs a total of 7,963,764 were included after imputation (during QCs 124,872 SNPs were excluded prior imputation and 39,109,351 were excluded after imputation).

The GWAS analysis revealed 20 suggestive SNPs at locus 19p13.13 ($p < 1x10^{-5}$; Table 2), located in the intronic region of the *MAN2B1* gene associated with IS (Figure 1). The analyses by stroke subtypes did not reveal any prominent association with the locus in the Discovery cohort (Table 3). Both top SNPs, rs8107196 and rs34324185, were on linkage disequilibrium ($R^2>0.8$). We performed a sample size calculation for rs34324185, based on the odds ratio obtained on the Discovery analysis. The minimum sample size required for replication was 637 cases and 637 controls. The replication analysis in the MEGASTROKE cohort (Replication cohort 1) showed an association of both linkage top SNPs with the lacunar stroke subtype on Trans-ethnical and only-Europeans analyses, which both were

statistically significant after Bonferroni correction ($p < 0.025$; Table 4). We performed a second replication using a Spanish cohort (Replication cohort 2) composed by 1,720 participants (1,531 cases (93 lacunar stroke cases) and 189 controls). In this second replication we obtained similar results, rs8107196 and rs34324185 were associated with the risk of the lacunar stroke subtype ($p < 0.05$; Table 4).

Experimental assays

On one hand, rs34324184 and rs8107196, the top SNPs of 19p13.13 locus, were searched in the GTEx web portal and showed association with eQTLs of different genes, including eQTLs of *MAN2B1* from whole blood and other different tissues (table e-2 and e-3). On the other hand, for SMR-HEIDI analysis, we performed a multi-SNP-based SMR test with all the SNPs of 19p13.13 locus that reached suggestive significance in the Discovery analysis (Table 2) and targeted the top SNP rs34324185. Results confirmed that our locus was only associated with *MAN2B1* expression, for whole blood and brain tissues (Table 5). Moreover, results showed that risk alleles for lacunar stroke were associated with higher *MAN2B1* expression. To confirm this data, we performed qRT-PCR analysis with cDNA from the blood of 66 healthy controls with rs34324185 genotyping data. We found that the relative RNA expression levels of *MAN2B1* from homozygous risk allele group (TT: n = 10, RQ mean = 0.75±0.06) were significantly higher compared to the other rs34324185 genotype groups (GG: n = 25, RQ mean = 0.65±0.03; TG: n = 30, RQ mean = 0.65±0.02; $p = 0.04$).

In order to study whether α-mannosidase activity (AMA) could be also associated with IS we first analyzed 24 controls and 14 IS cases selected by rs34324185 genotype to avoid false positive differences due to rs34324185 allele distribution (Figure 2A-B). IS patients showed significantly higher AMA than controls (AMA IS patients, 14.25 U/L; AMA controls, 10.72 U/L; $p = 0.03$; Figure 2A). Secondly, to test differences between IS etiologies, we performed a second experiment including 35 new samples: 12 controls, 6 cardioembolic strokes, 8 atherothrombotic strokes, and 9 lacunar strokes with a balanced proportion of rs34324185 alleles (Figure 2C-D). We found that AMA was significantly higher in lacunar strokes compared to controls (AMA lacunar, 12.58 U/L; AMA controls, 9.05 U/L; $p = 0.02$), and compared to atherothrombotic and cardioembolic strokes (AMA of lacunar strokes, 12.58 U/L; AMA atherothrombotic

and carioembolic strokes, 9.99 U/L; p = 0.04).

DISCUSSION

In this study we have found a locus in the intronic region of the *MAN2B1* gene associated with the lacunar stroke risk. We were able to replicate this result in two different cohorts with the same effect directionality as the Discovery cohort.

Using *in silico* analysis, we have found that the locus 19p13.13 was an eQTL of *MAN2B1* in blood and brain tissues. In all the cases, the risk allele for lacunar stroke was associated with higher expression of *MAN2B1*.We confirmed this association by performing qRT-PCR on blood samples.

As *MAN2B1* gene encodes for the lysosomal α-mannosidase enzyme, we performed α-mannosidase enzyme activity assays on plasma samples from controls and IS patients in order to clarify the association of this enzyme with the disease. We found that IS patients had significantly higher levels of activity compared to controls. Moreover, lacunar stroke patients showed the highest levels of activity compared to the other stroke subtypes in a group of samples with the same rs34324185 genotype distribution. As a result, our findings suggest that *MAN2B1* locus may have a role in the pathophysiology and in the risk of lacunar stroke. Additionally, our enzyme experiments suggest that α-mannosidase activity could potentially be a biomarker for the diagnosis of lacunar stroke during the acute phase of stroke. However, further research is needed in order to confirm this potential role as biomarker.

Besides, we detected different MAFs in each worldwide population for our top SNP rs34324185 (lower in East Asian (EAS), African (AFR) and European-non-Iberian (EUR-noIBS) compared to Iberians (IBS); MAF-EAS, 0.06; MAF-AFR, 0.08; MAF-EUR-noIBS, 0.37; and MAF-IBS, 0.44, in the Ensemble Portal). In fact, it is known that genetic risk factors could have different allele frequencies in specific populations, hence the use of homogeneous populations in GWAS analyses allow us to detect new genetic associations in stroke, as we observed in this study.

The biological reasons for the association between *MAN2B1* polymorphisms and the risk of lacunar stroke are unknown. The *MAN2B1* gene encodes for an enzyme that hydrolyzes terminal non-reducing α-D-mannose residues in α-D-mannosides. It is able to cleave all known types of α-mannosidic linkages, thereby releasing mannoses into the medium. This enzyme is a member of glycosyl-hydrolases family 38 and its activity is necessary for the catabolism of N-linked carbohydrates during the glycoprotein turnover. The

MAN2B1 protein is localized in the lysosome lumen, and also in the extracellular space due to neutrophil degranulation as part of the azurophilic granule content.

Interestingly, two genes associated with monogenic small vessel disease and lacunar stroke: *CTSA* (cathepsin A; described for CARASAL, cathepsinA-related arteriopathy with strokes and leukoencephalopathy), and *ADA2* (adenosine deaminase 2; described for DADA2, deficiency of adenosine deaminase 2) had its encoded proteins located in the lysosome lumen, in the extracellular space and are members of the azurophilic granule luminal proteins[32], as also observed for MAN2B1. Those facts suggested a potential role of the lysosomes or the proteins located in lysosomes in the progression of small vessel diseases and/or lacunar stroke.

Additionally, overexpression of another member of glycosyl-hydrolases family 38, *MAN2C1* (Mannosidase Alpha Class 2C Member 1, located in the cytoplasm) leads to protein underglycosylation and upregulation of the endoplasmic reticulum-associated protein degradation pathway[33]. Protein underglycosilation could have consequences on extracellular matrix composition, since glycoproteins are one of its structural components, and that could lead to disruption of the blood brain barrier and lacunar stroke[34].

Moreover, glycoproteins and glycosylation levels play an important role in the inflammatory process[35–37]. One study[36] describes how high-mannose hypoglycosilated N-glycans on the endothelial surface could increase the monocyte recruitment and promote the atherosclerosis plaque progression. This process in small vessels can lead to microinfarcts or lacunar stroke[34].

All these evidences together suggest a connection between *MAN2B1* and lacunar stroke. However, more studies are required to clarify the molecular mechanisms behind the association of *MAN2B1* with lacunar stroke risk.

Our study has some limitations. The sample size of Replication cohort 2 had a reduce number of lacunar strokes (n, 93) and controls (n, 189). However, despite this limited sample size we have replicated for a second time the association with lacunar stroke for rs34324185 and rs8107196. Besides, we have not studied differences between hemorrhagic stroke and ischemic stroke. It would be interesting since lacunar stroke and hemorrhagic stroke share some subjacent pathology in the context of small vessel disease.

In summary, we have found a new locus in the intronic region of the *MAN2B1* gene associated with the risk of lacunar stroke, which we have successfully replicated in two different cohorts, and we have demonstrated an implication of the encoded protein within the disease. Therefore, our study provides new information about genetic and molecular mechanisms behind the physiopathology of lacunar strokes. Moreover, α-mannosidase activity could potentially be a biomarker for the diagnosis of lacunar stroke. However, further research is needed to clarify the reasons why polymorphisms of *MAN2B1* are associated with lacunar stroke risk.

ACKNOWLEDGMENTS

This study was performed with the support of: GENERATION study, funded by the Instituto de Salud Carlos III (Carlos III Health Institute) and the European Regional Development Fund (ERDF); GENISIS study, funded by the National Institute of Health (K23 NS099487, and R01NIH NS085419); the MAESTRO study, funded by the Carlos III Health Institute and the ERDF; the EPIGENESIS study, funded by the Carlos III Institute and the ERDF; the Epigenesis study, funded by Marató de TV3; the INVICTUS Network RD16/0019, funded by RETICS, the Carlos III Health Institute and the ERDF.

I. Fernandez (CPII17/00021), F. Campos (CP14/00154) and T. Sobrino (CPII17/00027) are recipients of a research contract from Miguel Servet Program from the Carlos III Health Institute. R. de Cid was supported by the "Ramón y Cajal" researcher program (RYC-2011-07822). A. Bustamante is supported by a Juan Rodes research contract from the Carlos III Health Institute (JR16/00008).

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 10/22/2018.

This study makes use of data generated by the "GCAT |Genomes for Life Project. GCAT Projct, funded by MINECO "Acción de Dinamización" ADE 10/00026 and Generalitat de Catalunya (SGR 2014-1589). Cohort study of the Genomes of Catalonia", PMPPC-IGTP. IGTP is part of the CERCA Program / Generalitat de Catalunya. A full list of the investigators who contributed to the generation of the data is available from www.genomesforlife.com. The authors would like to thank all the GCAT volunteers for generously helping with this research.

The authors would like to thank all contributors and investigators of the MEGASTROKE project. The MEGASTROKE project received funding from sources specified at http://www.megastroke.org/acknowledgments.html.

CONFLICTS OF INTEREST

Nothing to repot.

REFERENCES

1. 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 [online serial]. 2012;43:3161–3167. Accessed at: https://www.ahajournals.org/doi/10.1161/STROKEAHA.112.665760.

2. Traylor M, Farrall M, Holliday EG, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol [online serial]. 2012;11:951–962. Accessed at: http://linkinghub.elsevier.com/retrieve/pii/S147444221270234X%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/23041239.

3. 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 [online serial]. 2016;15:174–184. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/26708676.

4. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet [online serial]. 2018;50:524–537. Accessed at: http://www.nature.com/articles/s41588-018-0058-3.

5. del Río-Espínola A, Fernández-Cadenas I, Giralt D, et al. A predictive clinical-genetic model of tissue plasminogen activator response in acute ischemic stroke. Ann Neurol [online serial]. 2012;72:716–729. Accessed at: http://doi.wiley.com/10.1002/ana.23664.

6. Malik R, Bevan S, Nalls MA, et al. Multilocus Genetic Risk Score Associates With Ischemic Stroke in Case–Control and Prospective Cohort Studies. Stroke [online serial]. 2014;45:394–402. Accessed at: https://www.ahajournals.org/doi/10.1161/STROKEAHA.113.002938.

7. Muiño E, Krupinski J, Carrera C, Gallego-Fabrega C, Montaner J, Fernández-Cadenas I. An Inflammatory Polymorphisms Risk Scoring System for the Differentiation of Ischemic Stroke Subtypes. Mediators Inflamm [online serial]. 2015;2015:1–7. Accessed at: http://www.hindawi.com/journals/mi/2015/569714.

8. Pulit SL, Weng L-C, McArdle PF, et al. Atrial fibrillation genetic risk differentiates cardioembolic stroke from other stroke subtypes. Neurol Genet [online serial]. 2018;4:e293. Accessed at: http://ng.neurology.org/lookup/doi/10.1212/NXG.0000000000000293.

9. Roth EM. Alirocumab for hyperlipidemia: ODYSSEY Phase III clinical trial results and US FDA approval indications. Future Cardiol [online serial]. 2016;12:115–128. Accessed at: https://www.futuremedicine.com/doi/10.2217/fca.15.78.

10. Sabatine MS, De Ferrari GM, Giugliano RP, et al. Clinical Benefit of Evolocumab by Severity and Extent of Coronary Artery Disease. Circulation [online serial]. 2018;138:756–766. Accessed at: https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.118.034309.

11. Holliday EG, Maguire JM, Evans T-J, et al. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. Nat Genet [online serial]. 2012;44:1147–1151. Accessed at: http://www.nature.com/articles/ng.2397.

12. Chi C, Shao X, Rhead B, et al. Admixture mapping reveals evidence of differential multiple sclerosis risk by genetic ancestry. Bush W, editor. PLOS Genet [online serial]. 2019;15:e1007808.

Accessed at: http://dx.plos.org/10.1371/journal.pgen.1007808.

13. Obón-Santacana M, Vilardell M, Carreras A, et al. GCAT|Genomes for life: a prospective cohort study of the genomes of Catalonia. BMJ Open [online serial]. 2018;8:e018324. Accessed at: http://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2017-018324.

14. Galván-Femenía I, Obón-Santacana M, Piñeyro D, et al. Multitrait genome association analysis identifies new susceptibility genes for human anthropometric variation in the GCAT cohort. J Med Genet [online serial]. 2018;55:765–778. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/30166351.

15. Chung J-W, Park SH, Kim N, et al. Trial of ORG 10172 in Acute Stroke Treatment (TOAST) classification and vascular territory of ischemic stroke lesions diagnosed by diffusion-weighted imaging. J Am Heart Assoc [online serial]. 2014;3. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/25112556.

16. Heitsch L, Ibanez L, Carrera C, et al. Meta-analysis of Transethnic Association (MANTRA) Reveals Loci Associated With Neurological Instability After Acute Ischemic Stroke. Int Stroke Conf. 2017.

17. Mola-Caminal M, Carrera C, Soriano-Tárraga C, et al. PATJ Low Frequency Variants Are Associated With Worse Ischemic Stroke Functional Outcome. Circ Res [online serial]. 2019;124:114–120. Accessed at: https://www.ahajournals.org/doi/10.1161/CIRCRESAHA.118.313533.

18. Domingues-Montanari S, Fernández-Cadenas I, Del Río-Espinola A, et al. KCNK17 genetic variants in ischemic stroke. Atherosclerosis [online serial]. 2010;208:203–209. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/19647252.

19. Fernández-Cadenas I, Mendióroz M, Giralt D, et al. GRECOS Project (Genotyping Recurrence Risk of Stroke). Stroke [online serial]. 2017;48:1147–1153. Accessed at: https://www.ahajournals.org/doi/10.1161/STROKEAHA.116.014322.

20. 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 [online serial]. 2012;322:79–81. Accessed at: https://linkinghub.elsevier.com/retrieve/pii/S0022510X12003231.

21. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet [online serial]. 2007;81:559–575. Accessed at: https://linkinghub.elsevier.com/retrieve/pii/S0002929707613524.

22. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. Nat Genet [online serial]. 2016;48:1284–1287. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/27571263.

23. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet [online serial]. 2007;39:906–913. Accessed at: http://www.nature.com/articles/ng2088.

24. Burton PR, Clayton DG, Cardon LR, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature [online serial]. 2007;447:661–678. Accessed at: http://www.nature.com/doifinder/10.1038/nature05911.

25. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet [online serial]. Nature Publishing

Group; 2010;11:499. Accessed at: https://doi.org/10.1038/nrg2796.

26. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics [online serial]. 2015;31:3555–3557. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/26139635.

27. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet [online serial]. 2013;45:580–585. Accessed at: http://www.nature.com/articles/ng.2653.

28. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet [online serial]. 2016;48:481–487. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/27019110.

29. Wu Y, Zeng J, Zhang F, et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun [online serial]. 2018;9:918. Accessed at: https://doi.org/10.1038/s41467-018-03371-0.

30. Qi T, Wu Y, Zeng J, et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat Commun [online serial]. 2018;9:2282. Accessed at: https://doi.org/10.1038/s41467-018-04558-1.

31. Stevenson M, Nunes T, Sanchez J, et al. EpiR: An R package for the analysis of epidemiological data. 2013. p. 9–43.

32. Lübke T, Lobel P, Sleat DE. Proteomics of the lysosome. Biochim Biophys Acta [online serial]. 2009;1793:625–635. Accessed at: http://www.ncbi.nlm.nih.gov/pubmed/18977398.

33. Bernon C, Carré Y, Kuokkanen E, et al. Overexpression of Man2C1 leads to protein underglycosylation and upregulation of endoplasmic reticulum-associated degradation pathway. Glycobiology [online serial]. 2011;21:363–375. Accessed at: https://academic.oup.com/glycob/article-lookup/doi/10.1093/glycob/cwq169.

34. Regenhardt RW, Das AS, Lo EH, Caplan LR. Advances in Understanding the Pathophysiology of Lacunar Stroke. JAMA Neurol [online serial]. 2018;75:1273. Accessed at: http://archneur.jamanetwork.com/article.aspx?doi=10.1001/jamaneurol.2018.1073.

35. Chacko BK, Scott DW, Chandler RT, Patel RP. Endothelial Surface N -Glycans Mediate Monocyte Adhesion and Are Targets for Anti-inflammatory Effects of Peroxisome Proliferator-activated Receptor γ Ligands. J Biol Chem [online serial]. 2011;286:38738–38747. Accessed at: http://www.jbc.org/lookup/doi/10.1074/jbc.M111.247981.

36. Scott DW, Chen J, Chacko BK, Traylor JG, Orr AW, Patel RP. Role of Endothelial N-Glycan Mannose Residues in Monocyte Recruitment During Atherogenesis. Arterioscler Thromb Vasc Biol [online serial]. 2012;32. Accessed at: https://www.ahajournals.org/doi/10.1161/ATVBAHA.112.253203.

37. Scott DW, Vallejo MO, Patel RP. Heterogenic Endothelial Responses to Inflammation: Role for Differential N -Glycosylation and Vascular Bed of Origin. J Am Heart Assoc [online serial]. 2013;2. Accessed at: https://www.ahajournals.org/doi/10.1161/JAHA.113.000263.

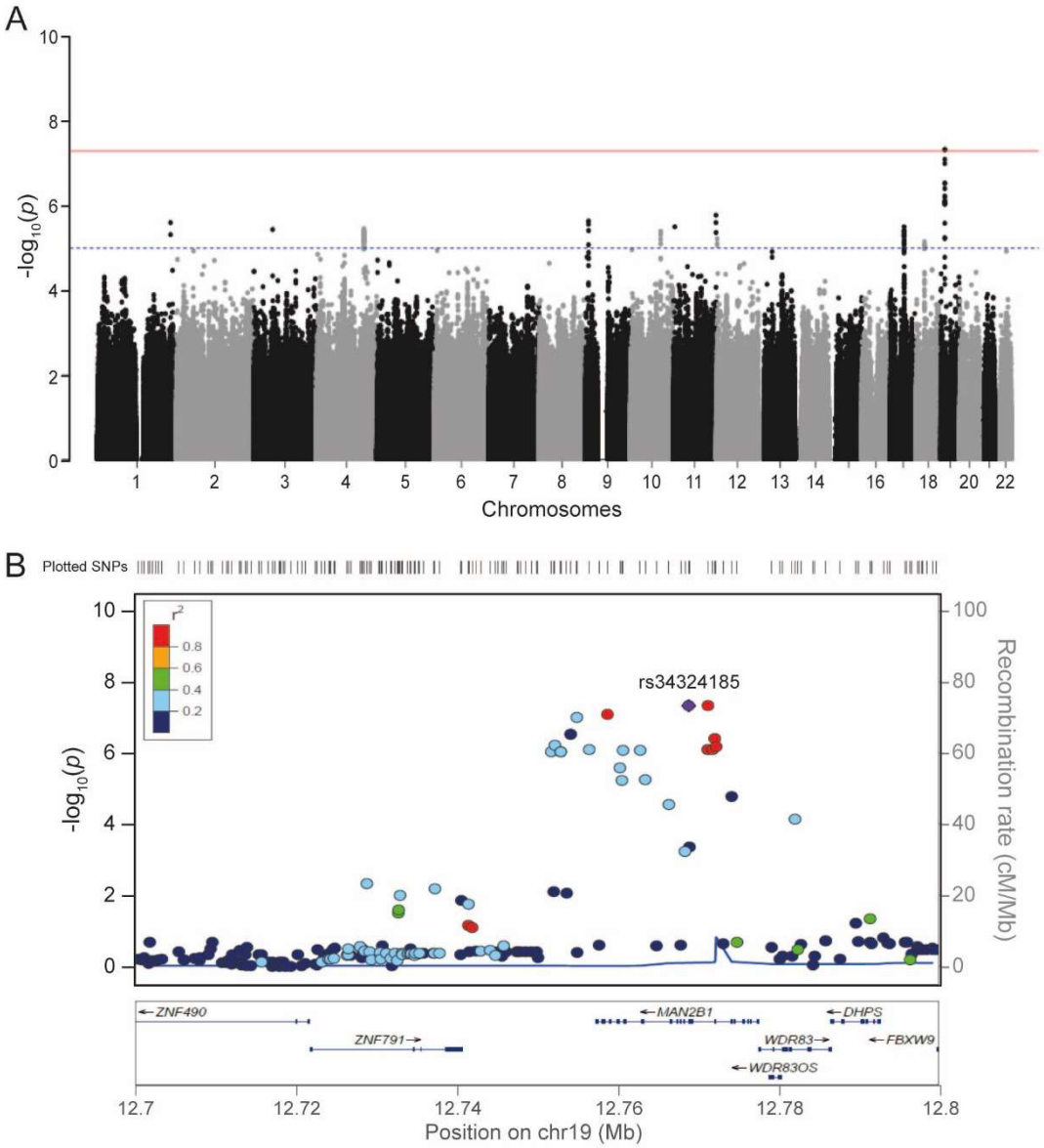

Figure 1. Results from the Discovery case-control GWAS analysis of Ischemic stroke. A) Manhattan plot of the Discovery case-control GWAS analysis of Ischemic stroke. The continuous line represents genome-wide significance ($5.0x10^{-08}$). The discontinuous line represents suggestive threshold ($1.0x10^{-05}$). B) Zoom Plot of the Discovery case-control GWAS analysis of Ischemic stroke performed on the LocusZoom portal. rs34324185 are the top SNP of the locus. The X axis shows chromosome location, and the Y axis shows the negative logarithm to base 10 of the *p* value ($-\log_{10}(p)$).

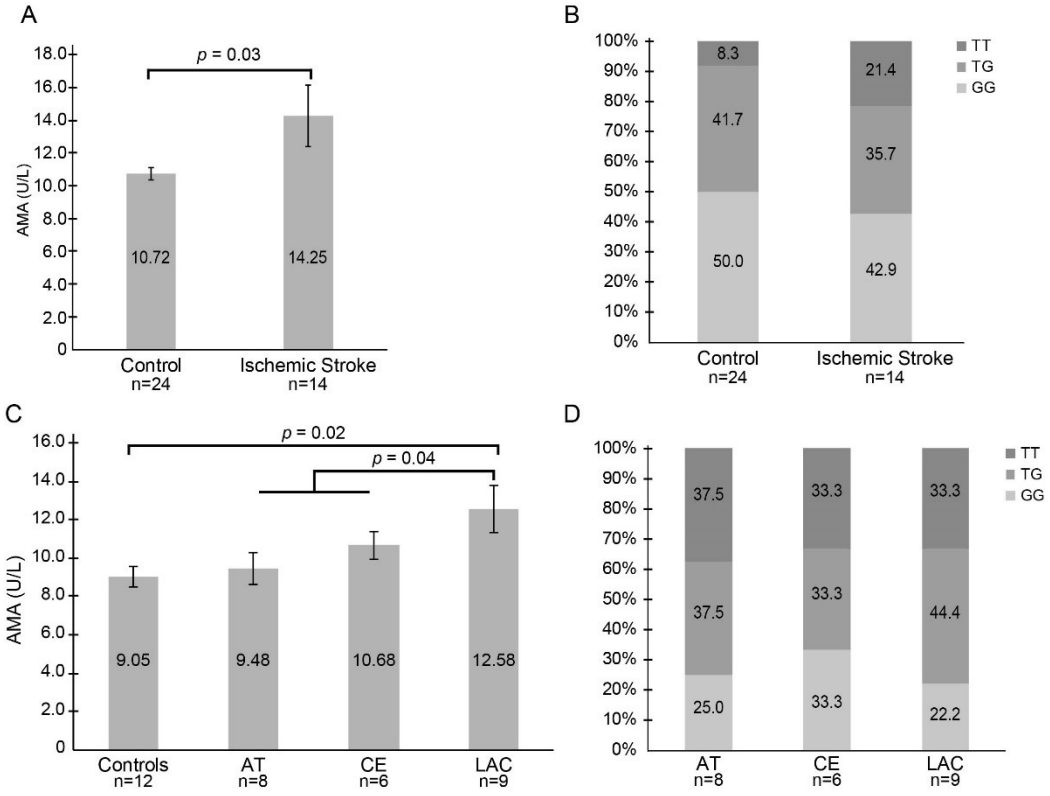

Figure 2. Results of α-mannosidase activity assays. A) Case-control α-mannosidase activity (AMA) assay. B) Allelic distribution of rs34324185 (%) for each group of case-control AMA experiment. The Chi-Square Test between phenotype (case/control) and genotype (GG/TG/TT) was not significant ($p = 0.52$). C) Results of second AMA assay, for control group and each stroke etiology group. D) Allelic distribution of rs34324185 (%) for each stroke etiology group of second AMA assay. The Chi-Square Test between etiology (AT/CE/LAC) and genotype (GG/TG/TT) was not significant ($p = 0.99$). For all panels, phenotype groups are shown on the X axis and the number of samples included (n) is detailed under the name category. In panels A and C, it is represented the mean (histogram) and standard error (error bars) of AMA for each group and the AMA is shown on the Y axis in units per liter (U/L). In panels B and D, the percentage (%) is shown on the Y axis. GG = reference allele homozygote, TG = heterozygote, TT = minor allele homozygote; AT = Atherothrombotic stroke, CE = Cardioembolic stroke, LAC = Lacunar stroke.

Table 1. Detailed characteristics of the Discovery cohort, the Replication cohort 1 and the Replication cohort 2.

Descriptive Variables	Discovery cohort		Replication cohort 1				Replication cohort 2	
	Controls	IS cases	Controls	IS cases	Controls	IS cases	Controls	IS cases
Ethnicity	IBS	IBS	EUR	EUR	TRANS	TRANS	IBS	IBS
Sample size (n)	5,227	1,752	406,111	34,217	454,450	60,341	189	1,531
Mean Age ±SD	52.3 ±7.9	73.2 ±12.6	55.1 ±5.0	68.8 ±10.2	55.3 ±5.8	68.6 ±9.9	67.8 ±7.1	72.9 ±12.6
Woman (%)	2,864 (54.8)	783 (44.6)	148,614 (39.5)	18,211 (47.1)	166,698 (50.8)	25,226 (43.0)	102 (54.0)	575 (37.6)
HT (%)	1,018 (19.5)	1,241 (70.7)	*NA*	*NA*	*NA*	*NA*	134 (70.9)	1008 (65.8)
DM (%)	269 (5.1)	504 (28.7)	*NA*	*NA*	*NA*	*NA*	28 (14.8)	369 (24.1)
AF (%)	11 (0.2)	568 (32.3)	*NA*	*NA*	*NA*	*NA*	*NA*	486 (31.7)
DL (%)	211 (4.0)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
TOAST AT (% IS cases)		292 (16.8)		4,373 (12.8)		6,688 (11.1)		244 (15.9)
TOAST CE (% IS cases)		757 (43.2)		7,193 (21.0)		9,006 (14.9)		690 (45.1)
TOAST LAC (% IS case)		123 (7.0)		5,386 (15.7)		11,710 (19.4)		93 (6.1)
Other TOAST categories*		580 (33.0)		17,265 (50.5)		32,937 (54.6)		504 (32.9)

For the Discovery cohort, the Replication cohort 1 (European and trans-ethnical) and the Replication cohort 2, it is detailed for cases and controls: the ethnicity, the sample size, the mean age (±SD), the percentage of women, the percentage of patients with hypertension (HT), diabetes mellitus(DM), atrial fibrillation (AF) and dyslipidemia (DL), and the proportion of ischemic stroke patients for each TOAST category. NA was considered when more than 80% of the data was not available. IBS = Iberians; EUR= Europeans; TRANS = Trans-ethnical; IS = ischemic stroke; AT = Atherothrombotic stroke; CE = cardioembolic stroke; LAC = Lacunar stroke; SD = Standard Deviation; n = number of samples; NA = non-available. *Including Undetermined category of the TOAST classification criteria.

Table 2. List of SNPs located in the loci associated with ischemic stroke in the Discovery cohort that reached suggestive significance.

SNP	Locus	A1	A2	MAF	P value	OR [95% CI]
rs34324185	**MAN2B1**	**T**	**G**	**0.401**	**$4.87x10^{-8}$**	**1.37 [1.22-1.54]**
rs8107196	**MAN2B1**	**G**	**C**	**0.401**	**$4.87x10^{-8}$**	**1.37 [1.22-1.54]**
rs74181667	MAN2B1	G	A	0.400	$8.30x10^{-8}$	1.37 [1.22-1.54]
rs8110545	MAN2B1	C	T	0.317	$8.73x10^{-8}$	0.72 [0.63-0.81]
rs8110264	MAN2B1	T	C	0.259	$3.43x10^{-7}$	0.71 [0.62-0.81]
rs8107354	MAN2B1	A	G	0.259	$3.43x10^{-7}$	0.71 [0.62-0.81]
rs1133330	MAN2B1	T	C	0.427	$5.68x10^{-7}$	1.34 [1.20-1.51]
rs8112964	ZNF490, MAN2B1	C	T	0.293	$7.15x10^{-7}$	0.73 [0.64-0.83]
rs2303731	MAN2B1	C	T	0.429	$9.08x10^{-7}$	1.34 [1.19-1.50]
rs8104226	MAN2B1	A	G	0.287	$9.81x10^{-7}$	0.73 [0.64-0.83]
rs3815914	MAN2B1	T	C	0.287	$9.96x10^{-7}$	0.73 [0.64-0.83]
rs17476839	MAN2B1	T	C	0.287	$9.96x10^{-7}$	0.73 [0.64-0.83]
rs4804727	ZNF490, MAN2B1	C	T	0.287	$1.09x10^{-6}$	0.73 [0.64-0.83]
rs4804728	ZNF490, MAN2B1	G	A	0.287	$1.09x10^{-6}$	0.73 [0.64-0.83]
rs11670251	ZNF490, MAN2B1	T	G	0.288	$1.11x10^{-6}$	0.73 [0.64-0.83]
rs10415457	MAN2B1	G	A	0.429	$1.13x10^{-6}$	1.33 [1.19-1.50]
rs8108316	MAN2B1	G	A	0.429	$1.13x10^{-6}$	1.33 [1.19-1.50]
rs12984441	MAN2B1	A	C	0.320	$2.30x10^{-6}$	0.74 [0.66-0.84]
rs4804205	MAN2B1	A	C	0.321	$7.08x10^{-6}$	0.75 [0.67-0.85]
rs8102193	MAN2B1	T	C	0.294	$7.13x10^{-6}$	0.75 [0.66-0.85]

Genome-wide significant SNPs are in bold. A1 = allele 1 (minor allele), A2 = allele 2, MAF = minor allele frequency, OR = odds ratio in reference to minor allele, 95% CI = 95 % confidence interval of 95%. N total = number of participants included in the analysis, N cases = number of patients included in the analysis.

Table 3. Genome-wide significant SNPs from the Discovery case-control GWAS of Ischemic stroke, in each TOAST category.

Analysis	N (n cases)	SNP	A1	A2	MAF	P value	OR [95% CI]
AT	5520 (293)	rs34324185	T	G	0.382	$9.21x10^{-3}$	1.33 [1.07-1.64]
		rs8107196	G	C	0.382	$9.21x10^{-3}$	1.33 [1.07-1.64]
CE	5984 (757)	rs34324185	T	G	0.389	$1.33x10^{-4}$	1.39 [1.17-1.64]
		rs8107196	G	C	0.389	$1.33x10^{-4}$	1.39 [1.17-1.64]
LAC	5350 (123)	rs34324185	T	G	0.381	$4.34x10^{-4}$	1.68 [1.26-2.26]
		rs8107196	G	C	0.381	$4.34x10^{-4}$	1.68 [1.26-2.26]
UNDET	5699 (472)	rs34324185	T	G	0.388	$5.01x10^{-4}$	1.40 [1.15-1.69]
		rs8107196	G	C	0.388	$5.01x10^{-4}$	1.40 [1.15-1.69]

IS = ischemic stroke; AT = Atherothrombotic stroke; CE = Cardioembolic stroke; LAC = Lacunar stroke; UNDET = Undetermined stroke. A1 = allele 1 (minor allele), A2 = allele 2, MAF = minor allele frequency, OR = odds ratio, 95% CI = 95 % confidence interval of 95%. N total = number of participants included in the analysis, N cases = number of patients included in the analysis.

Table 4. Results of the Replication analyses for the genome-wide significant SNPs found in the Discovery analysis.

	Analysis	N (n cases)	SNP	A 1	A 2	MAF	P	OR [95% CI]
Replication cohort 1	IS (EUR)	440,328 (34,217)	rs34324185	T	G	0.363	0.44	1.01 [0.99 - 1.03]
			rs8107196	G	C	0.364	0.50	1.01 [0.97 - 1.01]
	LAC (EUR)	411,497 (5,386)	**rs34324185**	**T**	**G**	**0.365**	**9.70 x10^{-3}**	**1.06 [1.01 - 1.12]**
			rs8107196	**G**	**C**	**0.365**	**0.01**	**1.06 [1.01 - 1.12]**
	IS (TRANS)	514,791 (60,341)	rs34324185	T	G	0.334	0.31	1.01 [0.99 - 1.03]
			rs8107196	G	C	0.335	0.33	1.01 [0.99 - 1.03]
	LAC (TRANS)	466,160 (11,710)	**rs34324185**	**T**	**G**	**0.315**	**5.90 x10^{-3}**	**1.06 [1.02 - 1.11]**
			rs8107196	**G**	**C**	**0.314**	**7.85 x10^{-3}**	**1.06 [1.01 - 1.12]**
Replication	LAC (IBS)	282 (93)	**rs34324185**	**T**	**G**	**0.358**	**0.04**	**1.59 [1.01 - 2.49]**
			rs8107196	**G**	**C**	**0.353**	**0.03**	**1.64 [1.04 - 2.58]**

Detailed *p* value and odds ratio for rs34324185 and rs8107196. Significant results are in bold. IS = Ischemic stroke; LAC = lacunar stroke; EUR= European; TRANS=Trans-ethnical; IBS = Iberians (Spanish); N = sample size; n cases = number of cases included in the sample size; A1 = allele 1 (minor allele); A2 = allele 2; MAF = minor allele frequency; P = *p* value; OR=Odds ratio; 95% CI = 95% confidence interval.

Table 5. Results of multi-SNP-based SMR and HEIDI tests for whole blood and brain tissues.

Gene	Tissue	Top SNP	A 1	A 2	GWAS		eQTL		SMR		multi-SMR	HEIDI	
					B ±SE	P	B ±SE	P	B ±SE	P	P	P	N
MAN 2B1	Whole Blood	rs343 24185	T	G	0.32 ±0.06	$4.63x 10^{-8}$	0.16 ±0.03	$1.57x 10^{-8}$	2.00 ±0.51	$8.52x 10^{-5}$	$6.42x 10^{-6}$	0.83	4
MAN 2B1	Brain	rs343 24185	T	G	0.32 ±0.06	$4.63x 10^{-8}$	0.24 ±0.05	$6.04x 10^{-7}$	1.28 ±0.35	$2.29x 10^{-4}$	$1.18x 10^{-5}$	0.30	5

A1 = allele 1 (minor allele); A2= allele 2; P = *p* value; B = beta coefficient; SE = Standard Error; multi-SMR = multi-SNP-based SMR; N = number of SNPs included in HEIDI test.

A genome-wide association study reveals a new locus in *MAN2B1* gene for lacunar stroke risk

SUPPLEMENTAL MATERIAL

appendix e-1

GENISIS cohort: Genetics of Early Neurological Instability after Ischemic Stroke (GENISIS)[e1] is an international study currently recruiting patients from four different locations: United States, Finland, Poland, and Spain. The inclusion criteria for the GENISIS study are IS patients (age ≥ 18 years) collected from 2003 to 2017 with a measurable neurologic deficit on the NIHSS within 6 hours of last known normal. Patients who received endovascular thrombectomy, or for whom consent and/or a blood sample could not be obtained were excluded. For our study we only included Spanish patients. Genotyping was performed with Human Core Exome chip (Illumina®).

GODS cohort: The Genetic contribution to functional Outcome and Disability after Stroke (GODS)[e2] project is a study that aimed to find genetic factors associated with stroke outcome. All participants met the following criteria: (1) European descent, aged >18 years, diagnosis of IS in the anterior vascular territory; (2) assessed by a neurologist during the acute phase of stroke; (3) initial stroke severity >4, according to the National Institutes of Health Stroke Scale (NIHSS); (4) information on post-stroke functional status at 3 months (or alternatively between 3 and 6 months); (5) evidence of acute IS in a neuroimaging study; (6) lack of concomitant disease. Individuals with stroke recurrence during the follow-up period were excluded, in addition to posterior vascular territory and lacunar strokes. Samples were genotyped at the Genetic and Molecular Epidemiology Laboratory of McMaster University (David Braley Research Institute) in Ontario, Canada, with Human Core Exome chip (Illumina®).

CONIC cohort: The CONtrol ICtus (CONIC) study [e3] is a national study focus on find new genetic risk factors for ischemic stroke, it is a case-control matched study. Control participants were recruited in Vall d'Hebron Hospital between 2007 and 2008. All controls were older than 65 years of age and declared free of dementia, neurovascular and/or cardiovascular disease, as evaluated by self-description during a direct interview before recruitment. Subjects with a history of first and/or second-degree neurovascular disorder were also excluded from the study.

The IS cases were admitted to the emergency department of a university with a documented middle cerebral artery (MCA) occlusion on transcranial Doppler ultrasonography (TCD) and received tPA in a standard 0.9-mg/kg dose (10% bolus, 90% continuous infusion over 1 hour) within 4.5 hours of symptom onset following National Institute of Neurological Disorders and Stroke (NINDS) recommendations. Cases and controls were genotyped with Human Core Exome chip (Illumina®).

GRECOS cohort: The Genotyping RECurrence Risk Of Stroke (GRECOS)[e4] project is a national study that aimed to find genetic factors associated with recurrence after stroke. Control participants were relatives of patients (wife or husband, without any consanguinity between cases and controls) and healthy volunteers visiting the same hospital for routine testing. They were >65 years of age and classified as free of neurovascular and cardiovascular history and family history by direct interview before recruitment. All samples were genotyped with Human Core Exome chip (Illumina®).

ISSYS cohort: The Investigating Silent Stroke in hYpertensives: A magnetic resonance imaging Study (ISSYS)[e5] 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 old, and diagnosed of essential hypertension at least one year before inclusion in the ISSYS study. Those individuals were genotyped with Human Core Exome chip (Illumina®).

GCAT cohort: GCAT|Genomes for Life Study [e6,e7] is a long-term project that was set up to integrate and assess the role of epidemiological, environmental and omic factors (ie, genomic, metabolomic, proteomic, epigenomic) in the development of chronic diseases. GCAT aims to assess the prevalence of risk factors and their association with disease incidence over time. The GCAT cohort is a prospective collection recruited from the general population of the north-east region of Spain, Catalonia. The GCAT Study have recruited 20 000 participants aged 40–65 years. Participants complete a self-administered computer-based questionnaire that collects data on a large number of lifestyle and health factors that are of interest in epidemiological and genetic studies. Participants who agreed to take part in the study completed a self-administered computer-driven questionnaire, and underwent blood pressure, cardiac frequency and anthropometry measurements. Participants will be followed for 20 years after recruitment. Genome-wide genotypes have been generated using Illumina Infinium SNP-bead array technology (Multi-Ethnic Global

(MEGAEX, V.2) consortium array), a multipurpose, multiethnic genotyping array with two million selected markers (including previously described germline mutations, insertions-deletions (InDels) and SNPs). Genotyping was performed at the Genomics and Bioinformatics Unit of the PMPPC Institute for Health Science Research Germans Trias i Pujol, in Badalona, Spain. Data is deposited in the public repository EGA accessible upon demand (EGAD00010001665). All genotyped individuals from Caucasian ancestry born in Spain were included in the study, and those with myocardial infarction or heart diseases were excluded.

table e-1. Detailed number of individuals included from each participating Hospital for the Discovery cohort and the Replication cohort 2.

Hospitals (City)	Discovery cohort			Replication cohort 2		
	Controls	Cases	**Total**	Controls	Cases	**Total**
Hospital del Mar (Barcelona)	0	711	**711**	0	58	**58**
Hospital Universitari Germans Tries i Pujol (Badalona)	0	330	**330**	0	10	**10**
Hospital Universitari Vall d'Hebron (Barcelona)	310	306	**616**	189	357	**546**
Hospital Universitari Son Espases (Palma de Mallorca)	0	227	**227**	0	70	**70**
Hospital Universitario Río Hortega (Valladolid)	0	23	**23**	0	2	**2**
Hospital Universitari Mútua de Terrassa (Terrassa)	1	22	**23**	0	53	**53**
Hospital Universitario Virgen del Rocío Y Virgen Macarena (Sevilla)	0	16	**16**	0	35	**35**
Hospital Clinic de Barcelona (Barcelona)	0	52	**52**	0	112	**112**
Hospital Universitario Doctor Josep Trueta (Girona)	0	22	**22**	0	0	**0**
Complejo Hospitalario Universitrio de Albacete (Albacete)	0	36	**36**	0	174	**174**
Hospital Universitario de Basurto (Bilbao)	0	5	**5**	0	9	**9**
Hospital Clínico Universitario de Santiago (Galicia)	0	0	**0**	0	607	**607**
Hospital de la Santa Creu i Sant Pau (Barcelona)	0	2	**2**	0	39	**39**
Hospital Arnau de Vilanova (Lleida)	0	0	**0**	0	5	**5**
GCAT Project	4,916	0	**4,916**	0	0	**0**
TOTAL	**5,227**	**1,752**	**6,979**	**189**	**1,531**	**1,720**

table e-2. eQTLs from GTEx data associated with rs34324185.

Gencode Id	Gene Symbol	P	NES	Tissue
ENSG00000104774.8	**MAN2B1**	8.20E-13	0.24	Nerve - Tibial
ENSG00000104774.8	**MAN2B1**	2.00E-10	0.21	Thyroid
ENSG00000104774.8	**MAN2B1**	8.50E-10	0.18	Cells - Transformed fibroblasts
ENSG00000104774.8	**MAN2B1**	**1.60E-08**	**0.16**	**Whole Blood**
ENSG00000105576.11	TNPO2	2.10E-08	0.14	Artery - Tibial
ENSG00000104774.8	**MAN2B1**	5.70E-08	0.24	Esophagus - Muscularis
ENSG00000105576.11	TNPO2	1.10E-07	0.17	Adipose - Subcutaneous
ENSG00000104774.8	**MAN2B1**	1.70E-07	0.27	Heart - Left Ventricle
ENSG00000104774.8	**MAN2B1**	2.30E-07	0.14	Lung
ENSG00000104774.8	**MAN2B1**	2.80E-07	0.15	Adipose - Subcutaneous
ENSG00000104774.8	**MAN2B1**	3.10E-07	0.28	Colon - Sigmoid
ENSG00000104774.8	**MAN2B1**	3.80E-07	0.16	Skin - Not Sun Exposed (Suprapubic)
ENSG00000104774.8	**MAN2B1**	4.40E-07	0.35	Testis
ENSG00000104774.8	**MAN2B1**	5.90E-07	0.15	Muscle - Skeletal
ENSG00000104774.8	**MAN2B1**	6.60E-07	0.19	Esophagus - Mucosa
ENSG00000104774.8	**MAN2B1**	6.70E-07	0.16	Skin - Sun Exposed (Lower leg)
ENSG00000104774.8	**MAN2B1**	1.10E-06	0.15	Artery - Tibial
ENSG00000104774.8	**MAN2B1**	1.70E-05	0.30	Adrenal Gland
ENSG00000105576.11	TNPO2	2.00E-05	0.17	Artery - Aorta
ENSG00000095066.7	HOOK2	2.30E-05	0.19	Heart - Left Ventricle

P = P value. NES = Normalized Effect Size.

table e-3. eQTLs from GTEx data associated with rs8107196.

Gencode Id	Gene Symbol	P	NES	Tissue
ENSG00000104774.8	**MAN2B1**	8.3e-13	0.24	Nerve - Tibial
ENSG00000104774.8	**MAN2B1**	2.3e-10	0.21	Thyroid
ENSG00000104774.8	**MAN2B1**	**2.6e-9**	**0.17**	**Whole Blood**
ENSG00000104774.8	**MAN2B1**	4.6e-9	0.17	Cells - Transformed fibroblasts
ENSG00000104774.8	**MAN2B1**	4.1e-8	0.28	Heart - Left Ventricle
ENSG00000105576.11	TNPO2	7.2e-8	0.17	Adipose - Subcutaneous
ENSG00000105576.11	TNPO2	8.7e-8	0.14	Artery - Tibial
ENSG00000104774.8	**MAN2B1**	1.5e-7	0.23	Esophagus - Muscularis
ENSG00000104774.8	**MAN2B1**	6.0e-7	0.28	Colon - Sigmoid
ENSG00000104774.8	**MAN2B1**	6.4e-7	0.14	Lung
ENSG00000104774.8	**MAN2B1**	7.3e-7	0.14	Adipose - Subcutaneous
ENSG00000104774.8	**MAN2B1**	8.9e-7	0.19	Esophagus - Mucosa
ENSG00000104774.8	**MAN2B1**	1.3e-6	0.16	Skin - Not Sun Exposed (Suprapubic)
ENSG00000104774.8	**MAN2B1**	1.3e-6	0.15	Skin - Sun Exposed (Lower leg)
ENSG00000104774.8	**MAN2B1**	1.9e-6	0.33	Testis
ENSG00000104774.8	**MAN2B1**	2.1e-6	0.14	Artery - Tibial
ENSG00000104774.8	**MAN2B1**	2.5e-6	0.14	Muscle - Skeletal
ENSG00000104774.8	**MAN2B1**	1.3e-5	0.31	Adrenal Gland
ENSG00000105576.11	TNPO2	3.0e-5	0.16	Artery - Aorta

P = P value; NES = Normalized Effect Size.

e-References

e1. Heitsch L, Ibanez L, Carrera C, et al. Meta-analysis of Transethnic Association (MANTRA) Reveals Loci Associated With Neurological Instability After Acute Ischemic Stroke. In: *International Stroke Conference*. ; 2017.

e2. Mola-Caminal M, Carrera C, Soriano-Tárraga C, et al. PATJ Low Frequency Variants Are Associated With Worse Ischemic Stroke Functional Outcome. *Circ Res*. 2019;124(1):114-120.

e3. Domingues-Montanari S, Fernández-Cadenas I, Del Río-Espinola A, et al. KCNK17 genetic variants in ischemic stroke. *Atherosclerosis*. 2010;208(1):203-209.

e4. Fernández-Cadenas I, Mendióroz M, Giralt D, et al. GRECOS Project (Genotyping Recurrence Risk of Stroke). *Stroke*. 2017;48(5):1147-1153. doi:10.1161/STROKEAHA.116.014322

e5. 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(1-2):79-81.

e6. Obón-Santacana M, Vilardell M, Carreras A, et al. GCAT|Genomes for life: a prospective cohort study of the genomes of Catalonia. *BMJ Open*. 2018;8(3):e018324.

e7. Galván-Femenía I, Obón-Santacana M, Piñeyro D, et al. Multitrait genome association analysis identifies new susceptibility genes for human anthropometric variation in the GCAT cohort. *J Med Genet*. 2018;55(11):765-778.

4.2. Artículo 2

Clinical Variables and Genetic Risk Factors Associated with the Acute Outcome of Ischemic Stroke: A systematic review

RESUMEN

El ictus es una enfermedad compleja y una de las principales causas de morbilidad y mortalidad entre la población adulta. Se sabe que una gran variedad de factores influye en el pronóstico del paciente, incluidas las variables demográficas, las comorbilidades o la genética. En esta revisión, exponemos lo que se sabe sobre la influencia de las variables clínicas y los factores de riesgo genéticos relacionados en el pronóstico del ictus isquémico, centrándonos en el pronóstico agudo y subagudo (dentro de las 24h-48h después del ictus y hasta el día 10, respectivamente), ya que son los primeros indicadores de daño tras un ictus.

Se realizaron búsquedas en la base de datos de PubMed para encontrar artículos que investigaron la interacción entre variables clínicas o factores genéticos y el pronóstico agudo o subagudo del ictus. Finalmente se incluyeron un total de 61 estudios en esta revisión.

Con respecto a los datos recopilados, las variables asociadas constantemente con el pronóstico agudo del ictus son: niveles de glucosa, presión arterial, presencia de fibrilación auricular, tratamiento previo con estatinas, gravedad del ictus, tipo de tratamiento agudo realizado, complicaciones neurológicas graves, niveles de leucocitos y factores de riesgo genéticos. Se requieren más investigaciones y esfuerzos internacionales en este campo, que deben incluir estudios de asociación de todo el genoma.

JoS
JOURNAL OF STROKE

Journal of Stroke 2019;21(3):?-?
https://doi.org/10.5853/jos.2019.01522

Review

Clinical Variables and Genetic Risk Factors Associated with the Acute Outcome of Ischemic Stroke: A Systematic Review

Nuria P Torres-Aguila,[a,b] Caty Carrera,[a,b] Elena Muiño,[a] Natalia Cullell,[c] Jara Cárcel-Márquez,[a] Cristina Gallego-Fabrega,[a,c] Jonathan González-Sánchez,[a,c,d] Alejandro Bustamante,[b] Pilar Delgado,[b] Laura Ibañez,[e] Laura Heitsch,[f,g] Jerzy Krupinski,[c,d] Joan Montaner,[h] Joan Martí-Fàbregas,[i] Carlos Cruchaga,[e] Jin-Moo Lee,[g] Israel Fernandez-Cadenas,[a] Acute Endophenotypes Group of the International Stroke Genetics Consortium (ISGC)

[a]Stroke Pharmacogenomics and Genetics Laboratory, Sant Pau Research Institute, Barcelona, Spain
[b]Neurovascular Research Laboratory, Vall d'Hebron Research Institute (VHIR), Autonomous University of Barcelona, Barcelona, Spain
[c]Stroke Pharmacogenomics and Genetics Laboratory, Mutua Terrasa Foundation of Teaching and Research, Mutua Terrassa Hospital, Terrassa, Spain
[d]Health Care Science Department, The Manchester Metropolitan University of All Saints, Manchester, UK
[e]Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA
[f]Division of Emergency Medicine, Washington University School of Medicine, St. Louis, MO, USA
[g]Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA
[h]Department of Neurology, Virgin Rocío and Macarena Hospitals, Institute of Biomedicine of Seville (IBiS), Seville, Spain
[i]Stroke Unit, Department of Neurology, Saint Cross and Saint Pau Hospital, Barcelona, Spain

Stroke is a complex disease and one of the main causes of morbidity and mortality among the adult population. A huge variety of factors is known to influence patient outcome, including demographic variables, comorbidities or genetics. In this review, we expound what is known about the influence of clinical variables and related genetic risk factors on ischemic stroke outcome, focusing on acute and subacute outcome (within 24 to 48 hours after stroke and until day 10, respectively), as they are the first indicators of stroke damage. We searched the PubMed data base for articles that investigated the interaction between clinical variables or genetic factors and acute or subacute stroke outcome. A total of 61 studies were finally included in this review. Regarding the data collected, the variables consistently associated with acute stroke outcome are: glucose levels, blood pressure, presence of atrial fibrillation, prior statin treatment, stroke severity, type of acute treatment performed, severe neurological complications, leukocyte levels, and genetic risk factors. Further research and international efforts are required in this field, which should include genome-wide association studies.

Keywords Stroke; Outcome; Clinical variables; Genetics

Correspondence: Israel Fernandez-Cadenas
Stroke Pharmacogenomics and Genetics Laboratory, Sant Pau Research Institute, C/Sant Antoni Maria Claret, 167, Barcelona 08025, Spain
Tel: +34-932746000
Fax: +34-935537864
E-mail: israelcadenas@yahoo.es

Co-correspondence: Jin-Moo Lee
Department of Neurology, Washington University School of Medicine, 660 S Euclid Ave, St. Louis, MO 63110, USA
Tel: +1-314-362-7382
Fax: +1-314-747-2244
E-mail: leejm@wustl.edu

Received: June 12, 2019
Revised: August 7, 2019
Accepted: August 28, 2019

Introduction

Stroke is one of the main causes of morbidity and mortality worldwide. In addition, as stroke is a main cause of disability in adults, there is a huge interest in improving the recovery of patients post-stroke.

A wide variety of factors are known to influence the outcome[1,2] after stroke, most of which are clinical variables related with the disease (stroke severity, etiology, etc.), cardiovascular risk factors (hypertension, heart failure, etc.) and other demographic variables (age, sex, etc.). However, there are studies that present contradictory results, making the relationship between clinical variables and stroke outcome not so clear.

In addition, ischemic stroke is a complex disease with a substantial genetic component, the heritability of which ranges from 16% to 40%.[3] Several genome-wide association studies (GWAS) have found genes associated with stroke risk and have been confirmed in independent studies.[4-6] However, with the exception of two recent GWAS,[7,8] the studies performed to find genetic variables associated with stroke outcome are candidate gene studies that have not been consistently replicated.[9]

Fast fibrinolysis or thrombectomy treatments are related with better recovery.[10,11] This suggests that outcome-related molecular mechanisms are taking place in the first 24 to 48 hours, the period defined as the acute phase of stroke. Acute outcome is defined as the outcome during the acute phase, and it is the first indicator of the impact of stroke on patient health. Acute outcome commonly reports the neurological status of the patient, usually measured by the National Institute of Health Stroke Scale (NIHSS). This scale is a systematic assessment tool that provides a quantitative measure of neurological deficit, evaluating different neurological aspects (consciousness, language, neglect, etc.), and can be used to predict long term outcome.[12,13] Consequently, acute outcome is associated with long-term outcome.

In this review, we detailed what is known about the influence of clinical variables and related genetic risk factors on the acute and subacute outcome of patients after an ischemic stroke (within 24 to 48 hours after stroke and until day 10, respectively). The aim of this review is to summarize all the knowledge acquired in recent years that could be useful for clinical practice and to perform studies in the field.

Methods

We used the National Center for Biotechnology Information (NCBI) website to search in the PubMed database. The keywords used were: "ischemic stroke," "neurological" or "neurologic," "associated" or "predictor," and "outcome." We included articles that searched for a relationship of acute and subacute outcome with other clinical variables or genetic factors, and which were written in English or Spanish. We excluded animal trials, childhood trials and articles with less than 100 patients analyzed. Using these criteria, we found 1,321 different articles by May 2019, plus six specific articles that were searched for specific clinical variables. A total of 61 were finally included, excluding process is detailed in a flow diagram performed following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[14] statements (Figure 1).

Variables associated with acute outcome

We classified the variables into three fields: (1) baseline variables, (2) early outcome variables, and (3) genetic factors, which summarized a total of 38, 20, and three articles, respectively (briefly detailed in Table 1).

Baseline variables

We defined as baseline variables those clinical factors present at the time of stroke onset and which are non-modifiable. These variables include demographics, comorbidities and pharmacological treatments prior to stroke.

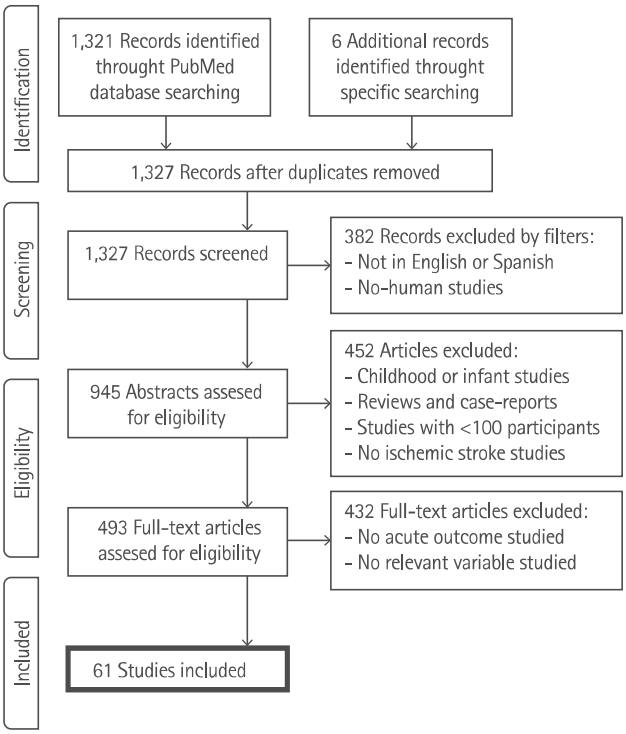

Figure 1. Flow diagram of the systematic review.

https://doi.org/10.5853/jos.2019.01522

Table 1. Detailed summary of each article included in this review

Study	Outcome studied (definition)	Cohort size (n)	Variable studied	Influence
Adams et al. (1999)[12]	7-day and 3-mo outcome (measured by Barthel Index and the Glasgow Outcome Scale)	1,281	Stroke severity	Association
Kugler et al. (2003)[16]	Early recovery at 24 hr and 1 wk (Barthel Index)	2,219	Age	Week influence (only at 1 wk)
Siegler et al. (2013)[18]	END (increase in NIHSS score of ≥2 points within 24 hr)	366	Age	Independent association
			Sex	No association
			Stroke severity	Independent association
Yeo et al. (2013)[19]	ENI (reduction of ≥10 points on NIHSS score, or score of 4 or less, at 2 hr); CNI (reduction in NIHSS score of ≥8 points between 2 and 24 hr, or an NIHSS score of ≤4 at 24 hr)	263	Age	Non-independent association
			Sex	Female gender associated with CNI
			Stroke severity	Independent predictor of CNI
Naess et al. (2014)[20]	7-day NIHSS, neurological worsening, mortality	1,867	Age	>80 yr associated with worse outcome
Boehm et al. (2014)[21]	END (increase of ≥2 points on NIHSS score during first 24 hr after hospitalization)	4,925	Age	Covariate
			Sex	Non-independent association
			Ethnicity	Non-independent association
Geng et al. (2017)[22]	END (increase of ≥2 points on NIHSS score during 1st wk after stroke)	1,064	Age	No association
			Sex	No association
			Diabetes mellitus	Association with END
			Hyperlipidemia	LDL and total cholesterol were associated with END, but not triglycerides
			Body mass index	No association with END
Hassaballa et al. (2001)[25]	7-day and 3-mo outcome (measured by Glasgow Outcome Scale)	1,093	Ethnicity	No association
Machumpurath et al. (2011)[26]	ENR (improvement at least 50% on NIHSS score within 24 hr)	161	Diabetes mellitus	Association (hyperglycemia patients were less likely to have ENR)
Roquer et al. (2014)[27]	END (increase of ≥4 points on NIHSS score during first 72 hr after stroke)		Diabetes mellitus	Association with END
Tang et al. (2016)[28]	Favorable neurological outcome (decrease of ≥4 points on NIHSS score or score of 0 at 24 hr, decrease of ≥8 points on NIHSS score or an score of 0 at 7 days; good functional outcome (mRS 0–1) at 3 mo	419	Diabetes mellitus	Predictor of unfavorable outcome
Yi et al. (2016)[29]	END (increase of ≥2 points on NIHSS score within 10 days after admission)	426	Diabetes mellitus	Association with END
Hui et al. (2018)[30]	END (increase of ≥2 points on NIHSS score within 5 days after stroke)	336	Diabetes mellitus	Association with END
Forlivesi et al. (2018)[31]	No neurological improvement (NIHSS score at 24 hr ≥NIHSS score at baseline)	200	Diabetes mellitus	Association with END
Vlcek et al. (2003)[32]	5-day outcome (Rankin Scale score >2 was defined as poor outcome)	372	Blood pressure	Independent association with poor outcome (high diastolic BP)
Castillo et al. (2004)[33]	END (diminution on Canadian Stroke Scale of ≥1 points within first 48 hr); neurological outcome and mortality at 3 mo	304	Blood pressure	Extreme values of BP were associated with poor outcome
Pezzini et al. (2011)[34]	END (increase of ≥4 points on NIHSS score at 48 hr); 90-day functional status (measured by mRS)	264	Blood pressure	Association, but dependent on stroke etiology
Geeganage et al. (2011)[35]	Death or neurological deterioration at 10 days	1,479	Blood pressure	Association (high systolic BP)
Kvistad et al. (2013)[36]	CNR (no ischemic stroke symptoms at 24 hr); favorable short-term outcome (7-day mRS score of 0–1)	749	Blood pressure	No association
Chung et al. (2015)[37]	END within 72 hr (increase of NIHSS score of ≥2 points)	1,116	Blood pressure	Independent association with END (high systolic BP)

Table 1. Continued

Study	Outcome studied (definition)	Cohort size (n)	Variable studied	Influence
Gill et al. (2016)[38]	Early neurological outcome (improvement of NIHSS score at 24 hr)	327	Blood pressure	Independent association with ENR (low diastolic BP)
Kellert et al. (2017)[39]	ENI (improvement of ≥20% on NIHSS score, or improvement of ≥8 points on NIHSS score); long-term functional outcome (mRS at 90 days)	28,976	Blood pressure	No association
Kang et al. (2017)[40]	END (worsening by 2 points on NIHSS score) at 1, 2 and 3 days	2,545	Blood pressure	Independent association (systolic BP)
Keezer et al. (2008)[41]	Poor outcome at 10 days (Rankin Scale score >3)	364	Blood pressure	Independent association with poor outcome (high and low BP values)
Sare et al. (2009)[42]	Neurological impairment (high 7-day NIHSS score than median NIHSS score); 90-day functional outcome (measured by mRS)	1,722	Blood pressure	Association with neurological impairment and poor outcome (high systolic BP)
Zhang et al. (2018)[43]	END (increase in NIHSS score ≥4 or increase in Ia of NIHSS ≥1 within 72 hr after recanalization treatment)	278	Blood pressure	Independent association (high systolic BP)
			Stroke etiology	Independent association in intravenous treated patients (large artery occlusion)
Sanák et al. (2010)[45]	24 hr and 7-day NIHSS score; 7-day mortality	157	Atrial fibrillation	Association with 7-day mortality
Yaghi et al. (2016)[46]	ENR (decrease of ≥8 points in NIHSS score, or score of 0–1 at 24 hr)	306	Atrial fibrillation	Significantly more present on non-ENR group; independent negative association with ENR
Restrepo et al. (2009)[47]	7-day NIHSS score	142	Hyperlipidemia	Association with hyperlipidemia history
Choi et al. (2012)[48]	END (increase in NIHSS score of ≥4 at 24 hr) or ENR (reduction of NIHSS score of ≥4) within a week after stroke onset	736	Hyperlipidemia	Extreme triglyceride levels associated with poor outcome
Branscheidt et al. (2016)[51]	ENR (improve >40% on NIHSS score at 24 hr); good outcome (mRS 0–1), favorable outcome (mRS 0–2) and mortality at 3 mo	896	Body mass index	No association
Power et al. (2013)[53]	NIHSS score at baseline and 24 hr	229	Renal dysfunction	Association
Lo et al. (2015)[54]	NIHSS improvement at 24 hr post-thrombolysis; 3-mo functional independence; 30-day mortality	199	Renal dysfunction	No association
Yu et al. (2009)[56]	10-day functional outcome (mRS)	339	Prior statin treatment	Association
			Prior antithrombotic treatment	No association
Ní Chróinín et al. (2011)[58]	7- and 28-day functional outcome (mRS); 7-, 28-, 90-day, and 1-yr mortality	448	Prior statin treatment	Associated with good outcome
Tsivgoulis et al. (2015)[59]	ECR (reduction of ≥10 points NIHSS score at 24 hr); good functional outcome (mRS 0–1) and mortality at 3 mo	1,660	Prior statin treatment	Association with ECR
Yi et al. (2017)[60]	Neurological deterioration (increase of 2 points of NIHSS during 10 days after admission)	1,124	Prior statin treatment	Concomitant use of antiplatelet and statins was associated with a favorable outcome
			Prior antithrombotic treatment	Concomitant use of antiplatelet and statins was associated with a favorable outcome
Cappellari et al. (2011)[61]	Neurological improvement (reduction of ≥4 points in NIHSS score between 24 and 72 hr)	250	Prior statin treatment	Prior and continued use of statins after stroke was associated with worse outcome
McAlpine et al. (2014)[63]	ENR (diminution on NIHSS score during first 24 hr after stroke)	158	Leukoaraiosis	No association

https://doi.org/10.5853/jos.2019.01522

Table 1. Continued

Study	Outcome studied (definition)	Cohort size (n)	Variable studied	Influence
Saposnik et al. (2008)[64]	7-, 30-day, and 1-yr mortality; neurological deterioration (measured by Canadian Neurological Scale, worsening neurological deficit or deterioration in the level of consciousness)	3,631	Stroke severity	Independent association
Kim et al. (2017)[65]	Early dramatic recovery (reduction of ≥8 points in NIHSS score or NIHSS score of 0–1 at 24 hr)	102	Stroke severity	Independent association
Schmitz et al. (2017)[66]	ENR (NIHSS score improvement of ≥4 points at 24 hr)	557	Stroke etiology	Cardioembolic stroke patients more likely to have ENR
Forlivesi et al. (2017)[67]	Neurological improvement (NIHSS score improvement of ≥4 points or NIHSS score of 0) at 7 days	122	Stroke etiology	Large artery strokes had lower odds ratio than cardioembolic strokes
Ciccone et al. (2013)[68]	Neurologic deficit (NIHSS score ≥6) at 7 days; functional outcome (mRS) and mortality at 90 days	362	Acute treatment	No association
Saver et al. (2015)[69]	NIHSS score changes at 27 hr; 3-mo functional outcome (mRS)	196	Acute treatment	Mechanical thrombectomy after IVT treatment had higher NIHSS score decrease
Jovin et al. (2015)[70]	ENR (decrease of 4 points in NIHSS at 24 hr); functional (Barthel Index) and neurological (NIHSS score) outcome at 90 days	206	Acute treatment	Mechanical thrombectomy had better outcome
Fiorelli et al. (1999)[71]	END (increase of NIHSS score of ≥4 at 24 hr post-stroke onset); 3-mo disability (mRS score ≥1) and 3-mo death	609	Hemorrhagic transformation	Independent association (server HT)
Kablau et al. (2011)[72]	ENR (decrease of >4 on NIHSS score) and END (increase of >4 on NIHSS score) at 5 days	122	Hemorrhagic transformation	No association with END; non-severe HT more common on ENR
Dharmasaroja et al. (2011)[73]	ENR (NIHSS of 0 to 4 at 24 hr)	203	Hemorrhagic transformation	Inversely association with ENR
Gill et al. (2016)[74]	Reduction in NIHSS score after 24 hr	339	Hemorrhagic transformation	Inversely associated (server HT)
Boehme et al. (2013)[77]	END (NIHSS score increase of ≥2 at 24 hr)	334	Infections	Non-independent association
Nardi et al. (2012)[80]	NIHSS score at baseline and at 72 hr; functional outcome (mRS) at discharge	811	Leukocyte counts	Independent association
Kumar et al. (2013)[81]	Neurological deterioration (NIHSS score increase of ≥2 within 24 hr)	292	Leukocyte counts	Association
Tian et al. (2018)[82]	ENI (decrease NIHSS score of ≥4 points or complete recovery after 24 hr of intravenous treatment)	240	Leukocyte counts	Independent association
Furlan et al. (2016)[84]	7-, 30-, and 90-day mortality rate	9,230	Blood platelet counts	Non-independent association for 7-day mortality rate; associated with 30- and 90-day mortality
Turcato et al. (2017)[85]	Lack of neurological improvement at 7 days (no NIHSS score of 0, nor NIHSS score ≤4 from baseline)	316	Red blood cell counts	Association with worse outcome
Pinho et al. (2018)[86]	NIHSS score at baseline and NIHSS score changes at 24 hr	602	Red blood cell counts	No association
Furlan et al. (2016)[87]	7-, 30-, and 90-day mortality rate	9,230	Red blood cell counts	High hemoglobin associated with high 7-day mortality
Yi et al. (2017)[88]	10-day END (NIHSS score increase of ≥2 points)	396	Genetic factors	CYP polymorphism associated with CYP plasma metabolites levels in END patients
Yi et al. (2017)[89]	10-day END (NIHSS score increase of ≥2 points)	297	Genetic factors	3 SNPs independent risk predictors for END
Yi et al. (2017)[90]	10-day END (NIHSS score increase of ≥2 points)	850	Genetic factors	High-risk interactive genotypes were associated with END

END, early neurological deterioration; NIHSS, National Institute of Health Stroke Scale; ENI, early neurological improvement; CNI, continuous neurological improvement; LDL, low density lipoprotein; ENR, early neurological recovery; mRS, modified Rankin Scale; BP, blood pressure; CNR, complete neurological recovery; ECR, early clinical recovery; HT, hemorrhagic transformation; CYP, cytochrome P450.

Demographics

In the literature, age, sex and race are the demographic variables that have been most often related to the acute outcome of stroke. We found seven articles that reported association of age, sex and/or race to acute outcome.

The relation of stroke outcome with age is well established.[15] Some prior studies reported advancing age as a major negative factor in morbidity, mortality, and long-term stroke outcome. Regarding acute outcome, there are also several studies reporting its association with age.[16-21] Kugler et al.[16] analyzed a cohort of 2,219 patients to study the association between age and early recovery after ischemic stroke. The authors studied the functional status with the Barthel Index score at 24 hours after admission, at 1 week and at discharge. Linear multiple regression showed significant independent negative influence of age on functional status at 1 week and at discharge, although this influence was weak. Other authors[20] analyzed age as a dichotomic variable, dividing the cohort used into ≥80 or <80 years old (592 and 1,275 patients were included in each group, respectively). They found that elder patients presented higher NIHSS score (at baseline and 7-day) and, at day 7, more neurological worsening and mortality. However, linear regression analysis showed that a higher 7-day NIHSS score was associated with a higher NIHSS score on admission and neurological worsening, but not with age ≥80. In contrast, in one study[22] with 1,064 patients that searched for variables associated with early neurological deterioration (END; defined as an increase of ≥2 in NIHSS score during the first week after stroke), the authors did not find significant differences in age between END and non-END patients. So, it seems that only older ages might be associated with worse acute and subacute outcome, although the association of age with long-term outcome is much clearer.

Another demographic factor that influences stroke outcome is sex. Stroke is a dimorphic disease, and incidence and outcome differences between genders have been reported previously.[23,24] However, due to the difference in lifespan between men and women, age is an important factor to take into consideration when sex influence is analyzed in stroke outcome.[18,19,21,22] Boehme et al.[21] performed a study to analyze the influence of sex and ethnicity on outcomes in which they included a total of 4,925 patients, (27.6% women, 26.9% Afro-Americans). The outcomes of interest were admission NIHSS, END (defined as increase of ≥2 points in NIHSS score within the first 24 hours after hospitalization) and functional outcome. Authors found differences in admission NIHSS and functional outcome depending on sex, although after adjusting by age and glucose on admission, the association was no longer significant. In addition, in a study[22] with 1,064 patients, sex was not associated with END at 1 week. So, it seems that the influence of sex on acute outcome might be dependent on other variables.

Although the role of ethnicity in stroke outcome is not widely considered, there are several studies reporting its influence on long-term outcome. However, in reference to acute outcome, we found two studies[21,25] that fulfilled our inclusion criteria for this review (others did not reach sample size). In both studies, authors found no significant differences in acute outcome between Afro-American patients and Caucasian patients.

Comorbidities

Among the long list of comorbidities that can influence outcome, we included those that have been reported to be associated with stroke outcome: diabetes mellitus (DM), high blood pressure (BP), atrial fibrillation (AF), hyperlipidemia, body mass index (BMI), renal dysfunction, heart failure, prior dementia, and prior disability.

The presence of DM as well as elevated levels of glucose has been associated with worse long-term outcome and acute outcome.[22,26-31] The largest study[27] that reported an association of DM with acute outcome included 1,088 patients. In this study, DM was associated with END (defined as an increase of ≥4 on NIHSS score during the first 72 hours after stroke). These studies indicate that DM or high levels of glucose on admission are associated with worse outcome in the acute and subacute phase.

We found several studies associating BP with acute neurological outcome.[32-43] In one study with 1,116 patients included (210 with END), the authors[37] analyzed the relationship among different measures of BP and END presented within 72 hours after stroke onset. Authors analyzed mean, maximum and minimum systolic and diastolic BP as well as the difference between maximum and minimum, the standard derivation and the coefficient of variation. The statistical analyses showed that all parameters, except diastolic BP mean, were independently associated with END. Moreover, other authors[38] also found a relationship between systolic BP and early neurological outcome (END) in a cohort of 327 ischemic stroke patients. Specifically, authors found that a reduction in systolic BP 24 hours after thrombolysis was independently associated with improvement in NIHSS score at 24 hours after thrombolysis. All those findings suggest that BP (both diastolic and systolic) is influencing outcome, with high systolic BP and diminution of diastolic BP being associated with worse outcome.

AF is one of the major risk factors for stroke[44] and its impact on outcome has been widely studied, although mostly with reference to long-term outcome. Regarding acute outcome, one study[45] analyzed a cohort of 157 patients treated with in-

 https://doi.org/10.5853/jos.2019.01522

travenous thrombolysis (IVT). No association of AF with NIHSS was found at 24 hours or 7 days, and, even AF was significantly associated with 3-month modified Rankin Scale (mRS) and 7-day mortality, this association was no longer significant in the multivariable regression analysis. On the other hand, in a more recent study,[46] authors searched for factors associated with early neurological recovery (ENR; defined as a decrease of ≥8 points in NIHSS score or a score of 0 to 1 at 24 hours) in a cohort of 306 patients treated with IVT. In this case, AF was significantly more present in the non-ENR group; this was also associated with 90-day mortality and 90-day disability.

There is no clear association of hyperlipidemia comorbidity with stroke outcome, neither with acute outcome.[22,47,48] In one study,[47] the authors briefly reported an association between history of hyperlipidemia and 7-day NIHSS score in their cohort of 142 patients. Alternatively, other authors[48] analyzed the relationship between triglycerides (TG) and END (increase in NIHSS score of ≥4 at 24 hours, or ENR, reduction of NIHSS score of ≥4) within a week after stroke onset. Authors include a total of 736 patients in their study. Statistical analysis showed that TG levels had a non-linear J-shape association with END and inverse J-shape association with ENR. Those results suggested that hypoTG and hyperTG were a risk factor for poor early outcome after ischemic stroke. In contrast, other study[22] found that total cholesterol and low density lipoprotein levels were associated with END (defined as an increase of ≥2 on NIHSS score in the first week after stroke) in a cohort of 1,064 patients, but not TG. Consequently, it seems that lipid content may have an influence on acute outcome, but further research is needed to establish a definitive conclusion.

The BMI is associated with cardiovascular diseases, being overweight and obesity well established risk factors.[49] Nevertheless, its influence on stroke outcome has been controversial due to the "obesity paradox."[50] A recent study[51] included 896 patients treated with IVT to determine the association of BMI with 3-month stroke outcome and, as secondary outcome, ENR at 24 hours (defined as an improvement of >40% on NIHSS score). In all cases, BMI was not associated with any studied outcome, even after adjusting for potential confounding factors. Moreover, these findings were observed in another study[22] with 1,064 patients, where no association was found between BMI and 1-week END. In conclusion, it seems that BMI had no impact on acute outcome.

Renal dysfunction is a cardiovascular risk factor commonly found in stroke patients,[52] it is defined as estimated glomerular filtration rate (eGFR) <60 mL/min. We found two studies that analyzed the role of eGFR with acute outcome.[53,54] On one hand, one study[53] analyzed the association of renal dysfunction in a cohort of 229 ischemic stroke patients treated with IVT. The authors found that patients with eGFR <60 mL/min had higher NIHSS scores at baseline and at 24 hours. On the other hand, another study,[54] with 199 patients recruited, did not find any significant difference in NIHSS improvement at 24 hours post-thrombolysis, functional independence at 3 months, nor 30-day mortality between patients with or without renal dysfunction. As a consequence, it is not clear if renal dysfunction could be associated with worse acute outcome. Further research is needed to clarify these controversial results.

For heart failure, prior dementia or prior disability we did not find any study related with acute outcome. Nonetheless, other studies had reported the influence of these variables on long-term outcome.[1,55]

Pharmacological treatments prior to stroke

Statins are prescribed for treatment of hypercholesterolemia. In stroke, statins are reported to reduce the risk of cerebrovascular events, and their role in outcome improvement have been highly studied with controversial results (several studies[47,56-60] found that statins improved outcome, although another study[61] did not find this association). Besides, antithrombotic drugs are used to prevent stroke recurrence and several studies have also analyzed the influence of prior antithrombotic treatments on acute outcome.[56,60] In the most recent study,[60] the authors used a cohort of 1,124 patients to examine the association of statins and antiplatelet pretreatments with neurological deterioration after stroke (defined as an increase of 2 points on NIHSS during 10 days after admission). They found that only concomitant use of antiplatelet and statins was associated with a favorable outcome. Moreover, another study[56] with a cohort of 339 patients observed that statins pretreatment was associated with good outcome (mRS score of 0 to 3) at 10 days, as well as the concomitant use of antihypertensive, antiplatelet and statins drugs, but not with the use of antiplatelet drugs alone. Furthermore, Tsivgoulis et al.[59] found that use of statins prior to stroke was independently related with early clinical neurological recovery in their cohort of 1,660 patients, although it was not related with good 3-month outcome.

In summary, prior use of only antiplatelet drugs is not enough to influence acute stroke outcome. In contrast, treatment with statins prior to stroke could be associated with better acute outcome, but not with long-term outcome.

Early outcome variables

Early outcome variables are those that can be described during the first few hours after stroke symptoms onset, such as: leukoaraiosis, stroke severity, acute treatment performed, neuro-

logical/clinical complications, levels of blood constituents, and, in some cases, stroke etiology.

Leukoaraiosis

Leukoaraiosis is a radiological phenomenon which represents white matter lesions and is commonly observed in elderly people. Patients with leukoaraiosis are described as more likely to suffer ischemic stroke and it has been demonstrated that leukoaraiosis is more common and more severe in ischemic stroke patients than in healthy people.[62] However, little is known about the role of leukoaraiosis on acute stroke outcome, and most of the studies have been performed with small cohorts. In the largest study[63] that we found, the authors analyzed the association of leukoaraiosis with ENR (defined as diminution on NIHSS score during the first 24 hours after stroke) in 158 patients and did not find any association. However, larger studies are warranted to clarify the role of leukoaraiosis in acute outcome.

Stroke severity

Stroke severity is one of the variables most strongly correlated with outcome, and its association with acute outcome has been widely reported.[12,17-19,64,65] In the largest study,[64] a cohort of 3,631 ischemic stroke patients was analyzed to describe the influence of clinical variables on 7-, 30-day, and 1-year mortality. The authors found that stroke severity (measured by Canadian Neurological Scale) was independently associated with mortality at all three time points, as well as neurological deterioration during hospitalization. Stroke severity therefore seems to be more highly related with acute outcome than other factors, as its association is always reported as an independent association after multivariable regression analysis.

Stroke etiology

In one study,[66] authors observed that cardioembolic (CE) stroke patients were more likely to have ENR (NIHSS score improvement of ≥4 points at 24 hours) than large vessel disease (LVD) etiology in their cohort of 557 ischemic stroke patients; no differences were found between LVD and the other Trial of Org 10172 in Acute Stroke Treatment (TOAST) categories. Moreover, another study[67] found similar results in their cohort of 122 ischemic stroke patients when analyzing 7-day neurological improvement depending on stroke subtype. So, it seems that LVD is associated with worse acute outcome in terms of recovery compared with CE stroke etiology. However, both studies were performed on patients undergoing thrombolysis, so further research is needed to clarify the role of stroke etiology in acute stroke outcome.

Acute treatments

By acute treatments we refer to those treatments performed to treat ischemic stroke during the acute phase, commonly thrombolysis and/or thrombectomy. We found several studies about their influence on acute outcome.[68-70] In one study,[69] authors tested the efficacy of mechanical thrombectomy after IVT compared to the use of IVT alone. A total of 196 patients underwent randomization, with 98 patients in each group; there were no significant differences in demographic or clinical characteristics between groups. The primary outcome of the study was functional outcome at 3 months, and secondary outcome was NIHSS changes at 27 hours. They found that combined treatment had a significantly better outcome at 3 months and a higher decrease in NIHSS score, with a better neurological status at 27 hours. Moreover, another study[70] found similar results in a cohort of 206 patients divided into two groups: medical therapy (control group; including IVT when eligible) and medical therapy combined with endovascular therapy by Solitaire stent retriever (thrombectomy group). They found that the thrombectomy group presented a higher rate of ENR (defined as a decrease of 4 points in NIHSS at 24 hours) as well as better 90-day functional and neurological outcome (by Barthel Index score and NIHSS score respectively). In conclusion, mechanical thrombectomy (after IVT or not) is associated with a better acute and long-term outcome.

Neurological complications

By neurological complications we mean those medical complications that may conclude with cognition deficit and could occur during the first days of hospitalization. We focused on hemorrhagic transformation (HT) and edema, due to their prevalence during acute and subacute phase.

HT is defined as an intracranial bleeding commonly detected by imaging (computed tomography or magnetic resonance imaging). In the literature, HT is commonly linked with stroke outcome and frequently more detected in IVT treated patients. Nevertheless, influence of HT on stroke outcome depends on its severity.[71-74] Fiorelli et al.[71] analyzed the influence of HT on ischemic stroke outcome: END (increase of NIHSS score of ≥4 at 24 hours post-stroke onset), 3-month disability (Rankin score ≥1) and 3-month death. Authors used a cohort of 609 patients treated with IVT or placebo, and used the European Cooperative Acute Stroke Study I (ECASS I) protocol for HT classification (hemorrhagic infarct 1 or 2 [HI-1, H1-2]; parenchymal hematoma 1 or 2 [PH-1, PH-2]). They found that PH-2 subtype entailed higher risk of END and 3-month death, independently of age and extent of initial ischemic damage, in placebo and IVT patients. On the other hand, in the most recent

 https://doi.org/10.5853/jos.2019.01522

study[74] authors used a cohort of 339 stroke patients to analyze the influence of HT on stroke outcome at 24 hours after thrombolysis measured by NIHSS. In this case, authors found that PH-2 subtype of HT was associated with worse neurological outcome. The authors concluded that mild to moderate HT should not be considered a complication and might be related with successful treatment and vascular recanalization. In summary, we may conclude that severe HT (i.e., PH-1, PH-2) is associated with worse acute outcome.

Cerebral edema is an accumulation of fluid in brain tissue, commonly observed in the acute phase of stroke. This neurological complication seems to have a more direct effect on stroke long-term outcome than HT.[75] However, we found no references in the literature to its influence on acute or subacute outcome.

Other clinical complications

As clinical complications, we included infections, gastrointestinal bleeding, and dysphagia. These three are the most commonly observed during the first days of hospitalization after stroke.[76] However, in reference to acute or subacute outcome, we only found information about infections. In Boehme et al.,[77] the authors analyzed the influence of infections on acute outcome of ischemic stroke patients for the first time, with END being the primary outcome (NIHSS score increase of ≥2 at 24 hours). They used a cohort of 334 patients, of which 77 had an infection, and classified the infections as present on admission (POA; infection diagnosed within the first 24 hours) and hospital-acquired infections (HAIs; infection diagnosed after 24 hours). Authors found that both POA and HAIs were associated with END, but after adjustment by age, NIHSS at baseline, glucose on admission and IVT treatment, only HAIs remained significant. Thus, as END was defined at 24 hours, and HAIs were posterior to 24 hours, we cannot conclude that infections affect acute outcome. And, as POA were not independently associated with END, it seems that the influence of prior infections on acute outcome is much lower than the influence of other variables, such as stroke severity or acute treatment.

Blood constituents

As there is an important inflammatory response during stroke events, the cells implicated in the immune system are likely to be associated with stroke outcome. In addition, it is reported that neutrophils are related with the blood brain barrier breakdown and their infiltration seems to be associated with higher inflammation and have a role in cerebral ischemia,[78] and higher neutrophil counts before thrombolysis have been associated with worse 3-month outcomes.[79] However, there are few studies analyzing the relation of leukocytes (including neutrophils and lymphocytes) with acute stroke outcome.[80-82] In Nardi et al.,[80] authors aimed to establish whether admission leukocyte count affects early stroke outcome. A total of 811 ischemic stroke patients were included in the study. NIHSS score was measured at baseline and after 72 hours, as well as mRS at discharge, and leukocytes counts were measured within 12 hours post-stroke onset. Authors found that higher leukocytes counts were independently associated with high NIHSS scores at baseline and 72 hours, and with poor functional outcome at discharge. So, it seems that leukocyte counts have an impact on subacute outcome independently of age or NIHSS at baseline.

Blood platelet counts (BPC) were previously associated with ischemic stroke risk,[83] although their influence on outcome is poorly described. The study by Furlan et al.[84] described the association of abnormal BPC with outcome. They analyzed the mortality rate in a cohort of 9,230 patients at 7, 30, and 90 days post-stroke. In a univariate analysis, all variables were associated with BPC, but after adjustment by principal confounders, only 30- and 90-day mortality remains significant. So, it seems that abnormal BPC (such as thrombocytopenia or thrombocytosis) is associated with long term outcome but not with acute outcome. However, further studies are needed to confirm these findings.

Red blood cell counts and hemoglobin levels could influence the reoxygenation during acute ischemic stroke and, in turn, the degree of neurological damage. Turcato et al.[85] analyzed the influence of red blood cell distribution width (RDW) on stroke outcome in a cohort of 316 ischemic stroke patients. Authors analyzed the association of RDW with lack of neurological improvement at 7 days (no NIHSS score of 0, nor NIHSS score ≤4 from baseline). They found that patients with RDW ≥14.5% showed a significantly lower decrease in NIHSS score at 24 hours and 7 days from baseline compared to patients with RDW <14.5%. A more recent study,[86] which included 602 patients, found that RDW was not associated with NIHSS nor NIHSS changes at 24 hours. Nevertheless, RDW was associated with 1-year survival and better 3-month functional outcome in older patients (≥75 years). On the other hand, Furlan et al.[87] analyzed the influence of blood hemoglobin concentration (HGB) on stroke severity and outcome after ischemic stroke in a large cohort of 9,230 ischemic stroke patients. They found that high HGB, but not low HGB, was an independent predictor of increased 7-day mortality compared to normal HGB. In summary, it seems that high oxygen availability in acute phase of stroke is associated with worse acute and subacute outcome, although more research is needed.

Genetic factors

There are several reports searching for the relationship of potential candidate genes with stroke outcome, most of them performed on animal models. However, regarding acute or subacute outcome, we found very few articles that attempted to find genetic factors associated with outcome.[88-90] The genes of interest in those studies were: cytochrome P450 (CYP), cyclooxygenase-2 (COX-2), prostaglandin I2 synthase (PTGIS), thromboxane A synthase 1 (TBXAS1), purinergic receptor P2Y1 (P2RY1), and integrin subunit beta 3 (ITGB3, or GPIIa). In those studies, authors found different single nucleotide polymorphisms (SNPs) independently associated with END at 10 days after stroke (defined as a NIHSS score increase of ≥2). However, no replication was performed in those studies. Only the SNP of COX-2, rs20417,was independently associated to END in two different studies.[89,90] So, it seems that genetic factors could have an influence on END, although further in-depth research is needed in this area. Additionally, in reference to long-term stroke outcome, two GWAS (Genetic contribution to functional Outcome and Disability after Stroke [GODS][7] and Genetics of Ischemic Stroke functional outCOME [GISCOME][8] studies, with 1,791 and 6,165 participants, respectively) have been recently published with remarkable results. One study[7] had found a locus located within a candidate gene, confirmed by an external replication. These studies are beginning to clarify the influence of genetics on patient recovery, which can help us to understand all the mechanisms involved.

Conclusions

Among all the clinical variables that were included in this review, there are few variables strong and clearly associated with acute or subacute stroke outcome (Table 2). These are: glucose levels or DM, BP, presence of AF, prior statin treatment, stroke severity, type of acute treatment performed, sever neurological complications (PH-2), and leukocytes levels. These clinical variables can easily be collected, so might be useful for prognosis of acute outcome. Other clinical variables that might be associated include hyperlipidemia, renal dysfunction, BPCs, and red blood cells (or hemoglobin) levels. For these variables, further research is required to establish a clear association.

It is surprising that age and sex, which are used as covariate in association studies such as GWAS, had a very weak influence on acute and subacute stroke outcome. However, it has been observed that those variables had an important influence on long-term outcome. Likewise, it is interesting that BMI is clearly not associated with acute stroke outcome, in contrast with the controversy observed about the relationship of BMI with long-term outcome. Additionally, it is important to highlight the influence of initial stroke severity on acute outcome, to the extent that it has been reported as an independent predictor in different studies and commonly included as covariate for predictor scales.

Table 2. Reviewed variables classified depending on its association with stroke outcome

Stroke outcome (acute and sub-acute)	Baseline variable	Early outcome variable	Genetic factor
Associated	Glucose levels or diabetes mellitus Blood pressure Atrial fibrillation Prior statin treatment	Stroke severity Type of acute treatment performed Sever neurological complications (PH-2) Leukocyte levels	rs20417 (located in *COX-2* gene)
Might associated	Hyperlipidemia Renal dysfunction	Leukoaraiosis Stroke etiology Prior infections Blood platelet counts Red blood cells or hemoglobin levels	Candidate genes: *CYP* *PTGIS* *TBXAS1* *P2RY1* *ITGB3*
Unknown	Heart failure Prior dementia Prior disability	Cerebral edema Gastrointestinal bleeding Dysphagia	
No associated	Age Sex Ethnicity Body mass index		

PH-2, parenchymal hematoma 2; COX-2, cyclooxygenase-2; *CYP*, cytochrome P450; *PTGIS*, prostaglandin I2 synthase; *TBXAS1*, thromboxane A synthase 1; *P2RY1*, purinergic receptor P2Y1; *ITGB3*, integrin subunit beta 3.

Regarding genetic factors, there are several SNPs reported to be associated with neurological deterioration. However, further studies are needed to validate these data, as there is a lack of replication in most of the studies performed. Only rs20417 (located in *COX-2* gene) was reported to be associated with END in two different studies.[89,90] Alternatively, genetic factors have been found to be associated with long-term outcome,[7,8] providing evidence of the utility of GWAS for exploring the genes associated with stroke outcome. Genetic analyses in this field may be useful to understand the molecular mechanisms behind the acute stroke outcome, and are required as no GWAS are currently reported.

As limitation, we considered that there is a lack of studies with enough statistical power to detect associations and perform consistent replication analysis, and also a no-consensus definition of the acute variable studied (i.e., END) makes impossible to perform meta-analyses, an approach required in order to obtain new qualitative and quantitative findings.

https://doi.org/10.5853/jos.2019.01522

In conclusion, our review provides a "state of the art" of this important field, reporting all the variables consistently associated with stroke early outcome and highlighting the lack of genetic studies. However, further research is required in this field. Analysis of acute and subacute outcome is important to understand the molecular mechanisms behind acute and long-term recovery and, finally, treat or prevent the worsening after stroke.

Disclosure

The authors have no financial conflicts of interest.

Acknowledgments

Alejandro Bustamante is supported by a Juan Rodes research contract from Carlos III Health Institute (JR16/00008). Israel Fernandez-Cadenas is the recipient of a research contract from the Miguel Servet Program from the Carlos III Health Institute (Instituto de Salud Carlos III) (CPII17/00021).

References

1. Saposnik G, Kapral MK, Liu Y, Hall R, O'Donnell M, Raptis S, et al. IScore: a risk score to predict death early after hospitalization for an acute ischemic stroke. *Circulation* 2011;123:739-749.
2. Ibrahim-Verbaas CA, Fornage M, Bis JC, Choi SH, Psaty BM, Meigs JB, et al. Predicting stroke through genetic risk functions: the CHARGE Risk Score Project. *Stroke* 2014;45:403-412.
3. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genome-wide associations. *Stroke* 2012;43:3161-3167.
4. 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* 2012;11:951-962.
5. 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;15:174-184.
6. 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;50:524-537.
7. 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. *Circ Res* 2019;124:114-120.
8. 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:e1271-e1283.
9. Lindgren A, Maguire J. Stroke recovery genetics. *Stroke* 2016;47:2427-2434.
10. Qureshi AI, Kirmani JF, Sayed MA, Safdar A, Ahmed S, Ferguson R, et al. Time to hospital arrival, use of thrombolytics, and in-hospital outcomes in ischemic stroke. *Neurology* 2005;64:2115-2120.
11. Matsuo R, Yamaguchi Y, Matsushita T, Hata J, Kiyuna F, Fukuda K, et al. Association between onset-to-door time and clinical outcomes after ischemic stroke. *Stroke* 2017;48:3049-3056.
12. Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, et al. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: a report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). *Neurology* 1999;53:126-131.
13. Takagi T, Kato T, Sakai H, Nishimura Y. Early neurologic improvement based on the National Institutes of Health Stroke Scale score predicts favorable outcome within 30 minutes after undergoing intravenous recombinant tissue plasminogen activator therapy. *J Stroke Cerebrovasc Dis* 2014;23:69-74.
14. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
15. Jongbloed L. Prediction of function after stroke: a critical review. *Stroke* 1986;17:765-776.
16. Kugler C, Altenhöner T, Lochner P, Ferbert A; Hessian Stroke Data Bank Study Group ASH. Does age influence early recovery from ischemic stroke? A study from the Hessian Stroke Data Bank. *J Neurol* 2003;250:676-681.
17. Boddu DB, Srinivasarao Bandaru VC, Reddy PG, Madhusudan M, Rukmini MK, Suryaprabha T, et al. Predictors of major neurological improvement after intravenous thrombolysis in acute ischemic stroke: a hospital-based study from south India. *Neurol India* 2010;58:403-406.
18. Siegler JE, Boehme AK, Kumar AD, Gillette MA, Albright KC, Beasley TM, et al. Identification of modifiable and nonmodifiable risk factors for neurologic deterioration after acute ischemic stroke. *J Stroke Cerebrovasc Dis* 2013;22:e207-e213.
19. Yeo LL, Paliwal P, Teoh HL, Seet RC, Chan BP, Wakerley B, et al. Early and continuous neurologic improvements after intravenous thrombolysis are strong predictors of favorable long-term outcomes in acute ischemic stroke. *J Stroke Cere-*

brovasc Dis 2013;22:e590-e596.
20. Naess H, Gjerde G, Waje-Andreassen U. Ischemic stroke in patients older and younger than 80 years. *Acta Neurol Scand* 2014;129:399-404.
21. Boehme AK, Siegler JE, Mullen MT, Albright KC, Lyerly MJ, Monlezun DJ, et al. Racial and gender differences in stroke severity, outcomes, and treatment in patients with acute ischemic stroke. *J Stroke Cerebrovasc Dis* 2014;23:e255-e261.
22. Geng HH, Wang Q, Li B, Cui BB, Jin YP, Fu RL, et al. Early neurological deterioration during the acute phase as a predictor of long-term outcome after first-ever ischemic stroke. *Medicine (Baltimore)* 2017;96:e9068.
23. Roth DL, Haley WE, Clay OJ, Perkins M, Grant JS, Rhodes JD, et al. Race and gender differences in 1-year outcomes for community-dwelling stroke survivors with family caregivers. *Stroke* 2011;42:626-631.
24. Giralt D, Domingues-Montanari S, Mendioroz M, Ortega L, Maisterra O, Perea-Gainza M, et al. The gender gap in stroke: a meta-analysis. *Acta Neurol Scand* 2012;125:83-90.
25. Hassaballa H, Gorelick PB, West CP, Hansen MD, Adams HP Jr. Ischemic stroke outcome: racial differences in the trial of danaparoid in acute stroke (TOAST). *Neurology* 2001;57:691-697.
26. Machumpurath B, Davis SM, Yan B. Rapid neurological recovery after intravenous tissue plasminogen activator in stroke: prognostic factors and outcome. *Cerebrovasc Dis* 2011;31:278-283.
27. Roquer J, Rodríguez-Campello A, Cuadrado-Godia E, Giralt-Steinhauer E, Jiménez-Conde J, Dégano IR, et al. Ischemic stroke in prediabetic patients. *J Neurol* 2014;261:1866-1870.
28. Tang H, Zhang S, Yan S, Liebeskind DS, Sun J, Ding X, et al. Unfavorable neurological outcome in diabetic patients with acute ischemic stroke is associated with incomplete recanalization after intravenous thrombolysis. *J Neurointerv Surg* 2016;8:342-346.
29. Yi X, Wang C, Liu P, Fu C, Lin J, Chen Y. Antiplatelet drug resistance is associated with early neurological deterioration in acute minor ischemic stroke in the Chinese population. *J Neurol* 2016;263:1612-1629.
30. Hui J, Zhang J, Mao X, Li Z, Li X, Wang F, et al. The initial glycemic variability is associated with early neurological deterioration in diabetic patients with acute ischemic stroke. *Neurol Sci* 2018;39:1571-1577.
31. Forlivesi S, Micheletti N, Tomelleri G, Bovi P, Cappellari M. Association of hyperglycemia, systolic and diastolic hypertension, and hyperthermia relative to baseline in the acute phase of stroke with poor outcome after intravenous thrombolysis. *Blood Coagul Fibrinolysis* 2018;29:167-171.
32. Vlcek M, Schillinger M, Lang W, Lalouschek W, Bur A, Hirschl MM. Association between course of blood pressure within the first 24 hours and functional recovery after acute ischemic stroke. *Ann Emerg Med* 2003;42:619-626.
33. Castillo J, Leira R, García MM, Serena J, Blanco M, Dávalos A. Blood pressure decrease during the acute phase of ischemic stroke is associated with brain injury and poor stroke outcome. *Stroke* 2004;35:520-526.
34. Pezzini A, Grassi M, Del Zotto E, Volonghi I, Giossi A, Costa P, et al. Influence of acute blood pressure on short- and mid-term outcome of ischemic and hemorrhagic stroke. *J Neurol* 2011;258:634-640.
35. Geeganage C, Tracy M, England T, Sare G, Moulin T, Woimant F, et al. Relationship between baseline blood pressure parameters (including mean pressure, pulse pressure, and variability) and early outcome after stroke: data from the Tinzaparin in Acute Ischaemic Stroke Trial (TAIST). *Stroke* 2011;42:491-493.
36. Kvistad CE, Logallo N, Oygarden H, Thomassen L, Waje-Andreassen U, Naess H. Elevated admission blood pressure and stroke severity in acute ischemic stroke: the Bergen NORSTROKE Study. *Cerebrovasc Dis* 2013;36:351-354.
37. Chung JW, Kim N, Kang J, Park SH, Kim WJ, Ko Y, et al. Blood pressure variability and the development of early neurological deterioration following acute ischemic stroke. *J Hypertens* 2015;33:2099-2106.
38. Gill D, Cox T, Aravind A, Wilding P, Korompoki E, Veltkamp R, et al. A fall in systolic blood pressure 24 hours after thrombolysis for acute ischemic stroke is associated with early neurological recovery. *J Stroke Cerebrovasc Dis* 2016;25:1539-1543.
39. Kellert L, Hametner C, Ahmed N, Rauch G, MacLeod MJ, Perini F, et al. Reciprocal interaction of 24-hour blood pressure variability and systolic blood pressure on outcome in stroke thrombolysis. *Stroke* 2017;48:1827-1834.
40. Kang J, Hong JH, Jang MU, Choi NC, Lee JS, Kim BJ, et al. Change in blood pressure variability in patients with acute ischemic stroke and its effect on early neurologic outcome. *PLoS One* 2017;12:e0189216.
41. Keezer MR, Yu AY, Zhu B, Wolfson C, Côté R. Blood pressure and antihypertensive therapy as predictors of early outcome in acute ischemic stroke. *Cerebrovasc Dis* 2008;25:202-208.
42. Sare GM, Ali M, Shuaib A, Bath PM; VISTA Collaboration. Relationship between hyperacute blood pressure and outcome after ischemic stroke: data from the VISTA collaboration. *Stroke* 2009;40:2098-2103.
43. Zhang YB, Su YY, He YB, Liu YF, Liu G, Fan LL. Early neurological deterioration after recanalization treatment in patients with acute ischemic stroke: a retrospective study. *Chin Med J (Engl)* 2018;131:137-143.

 https://doi.org/10.5853/jos.2019.01522

44. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;22:983-988.
45. Sanák D, Herzig R, Král M, Bártková A, Zapletalová J, Hutyra M, et al. Is atrial fibrillation associated with poor outcome after thrombolysis? *J Neurol* 2010;257:999-1003.
46. Yaghi S, Hinduja A, Bianchi N. Predictors of major improvement after intravenous thrombolysis in acute ischemic stroke. *Int J Neurosci* 2016;126:67-69.
47. Restrepo L, Bang OY, Ovbiagele B, Ali L, Kim D, Liebeskind DS, et al. Impact of hyperlipidemia and statins on ischemic stroke outcomes after intra-arterial fibrinolysis and percutaneous mechanical embolectomy. *Cerebrovasc Dis* 2009;28:384-390.
48. Choi KH, Park MS, Kim JT, Chang J, Nam TS, Choi SM, et al. Serum triglyceride level is an important predictor of early prognosis in patients with acute ischemic stroke. *J Neurol Sci* 2012;319:111-116.
49. Strazzullo P, D'Elia L, Cairella G, Garbagnati F, Cappuccio FP, Scalfi L. Excess body weight and incidence of stroke: meta-analysis of prospective studies with 2 million participants. *Stroke* 2010;41:e418-e426.
50. Andersen KK, Olsen TS. The obesity paradox in stroke: lower mortality and lower risk of readmission for recurrent stroke in obese stroke patients. *Int J Stroke* 2015;10:99-104.
51. Branscheidt M, Schneider J, Michel P, Eskioglou E, Kaegi G, Stark R, et al. No impact of body mass index on outcome in stroke patients treated with IV thrombolysis BMI and IV thrombolysis outcome. *PLoS One* 2016;11:e0164413.
52. Hojs Fabjan T, Hojs R. Stroke and renal dysfunction. *Eur J Intern Med* 2014;25:18-24.
53. Power A, Epstein D, Cohen D, Bathula R, Devine J, Kar A, et al. Renal impairment reduces the efficacy of thrombolytic therapy in acute ischemic stroke. *Cerebrovasc Dis* 2013;35:45-52.
54. Lo WT, Cheung CY, Li CK, Chau KF, Fong WC. Thrombolysis in Chinese ischemic stroke patients with renal dysfunction. *Interv Neurol* 2015;3:101-106.
55. Gao CY, Lian Y, Zhang M, Zhang LL, Fang CQ, Deng J, et al. Association of dementia with death after ischemic stroke: a two-year prospective study. *Exp Ther Med* 2016;12:1765-1769.
56. Yu AY, Keezer MR, Zhu B, Wolfson C, Côté R. Pre-stroke use of antihypertensives, antiplatelets, or statins and early ischemic stroke outcomes. *Cerebrovasc Dis* 2009;27:398-402.
57. Tsai NW, Lin TK, Chang WN, Jan CR, Huang CR, Chen SD, et al. Statin pre-treatment is associated with lower platelet activity and favorable outcome in patients with acute non-cardio-embolic ischemic stroke. *Crit Care* 2011;15:R163.
58. Ní Chróinín D, Callaly EL, Duggan J, Merwick Á, Hannon N, Sheehan Ó, et al. Association between acute statin therapy, survival, and improved functional outcome after ischemic stroke: the North Dublin Population Stroke Study. *Stroke* 2011;42:1021-1029.
59. Tsivgoulis G, Kadlecová P, Kobayashi A, Czlonkowska A, Brozman M, Švigelj V, et al. Safety of statin pretreatment in intravenous thrombolysis for acute ischemic stroke. *Stroke* 2015;46:2681-2684.
60. Yi X, Han Z, Wang C, Zhou Q, Lin J. Statin and aspirin pretreatment are associated with lower neurological deterioration and platelet activity in patients with acute ischemic stroke. *J Stroke Cerebrovasc Dis* 2017;26:352-359.
61. Cappellari M, Deluca C, Tinazzi M, Tomelleri G, Carletti M, Fiaschi A, et al. Does statin in the acute phase of ischemic stroke improve outcome after intravenous thrombolysis? A retrospective study. *J Neurol Sci* 2011;308:128-134.
62. Smith EE. Leukoaraiosis and stroke. *Stroke* 2010;41(10 Suppl):S139-S143.
63. McAlpine H, Churilov L, Mitchell P, Dowling R, Teo S, Yan B. Leukoaraiosis and early neurological recovery after intravenous thrombolysis. *J Stroke Cerebrovasc Dis* 2014;23:2431-2436.
64. Saposnik G, Hill MD, O'Donnell M, Fang J, Hachinski V, Kapral MK, et al. Variables associated with 7-day, 30-day, and 1-year fatality after ischemic stroke. *Stroke* 2008;39:2318-2324.
65. Kim DH, Nah HW, Park HS, Choi JH, Kang MJ, Cha JK. Factors associated with early dramatic recovery following successful recanalization of occluded artery by endovascular treatment in anterior circulation stroke. *J Clin Neurosci* 2017;46:171-175.
66. Schmitz ML, Simonsen CZ, Svendsen ML, Larsson H, Madsen MH, Mikkelsen IK, et al. Ischemic stroke subtype is associated with outcome in thrombolyzed patients. *Acta Neurol Scand* 2017;135:176-182.
67. Forlivesi S, Bovi P, Tomelleri G, Micheletti N, Carletti M, Moretto G, et al. Stroke etiologic subtype may influence the rate of hyperdense middle cerebral artery sign disappearance after intravenous thrombolysis. *J Thromb Thrombolysis* 2017;43:86-90.
68. Ciccone A, Valvassori L, Nichelatti M, Sgoifo A, Ponzio M, Sterzi R, et al. Endovascular treatment for acute ischemic stroke. *N Engl J Med* 2013;368:904-913.
69. Saver JL, Goyal M, Bonafe A, Diener HC, Levy EI, Pereira VM, et al. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. *N Engl J Med* 2015;372:2285-2295.
70. Jovin TG, Chamorro A, Cobo E, de Miquel MA, Molina CA, Rovira A, et al. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *N Engl J Med* 2015;372:2296-2306.

71. Fiorelli M, Bastianello S, von Kummer R, del Zoppo GJ, Larrue V, Lesaffre E, et al. Hemorrhagic transformation within 36 hours of a cerebral infarct: relationships with early clinical deterioration and 3-month outcome in the European Cooperative Acute Stroke Study I (ECASS I) cohort. *Stroke* 1999;30:2280-2284.
72. Kablau M, Kreisel SH, Sauer T, Binder J, Szabo K, Hennerici MG, et al. Predictors and early outcome of hemorrhagic transformation after acute ischemic stroke. *Cerebrovasc Dis* 2011;32:334-341.
73. Dharmasaroja PA, Muengtaweepongsa S, Dharmasaroja P. Early outcome after intravenous thrombolysis in patients with acute ischemic stroke. *Neurol India* 2011;59:351-354.
74. Gill D, Baheerathan A, Aravind A, Veltkamp R, Kar A. Severe hemorrhagic transformation after thrombolysis for acute ischemic stroke prevents early neurological improvement. *J Stroke Cerebrovasc Dis* 2016;25:2232-2236.
75. Clausen BH, Lundberg L, Yli-Karjanmaa M, Martin NA, Svensson M, Alfsen MZ, et al. Fumarate decreases edema volume and improves functional outcome after experimental stroke. *Exp Neurol* 2017;295:144-154.
76. Kumar S, Selim MH, Caplan LR. Medical complications after stroke. *Lancet Neurol* 2010;9:105-118.
77. Boehme AK, Kumar AD, Dorsey AM, Siegler JE, Aswani MS, Lyerly MJ, et al. Infections present on admission compared with hospital-acquired infections in acute ischemic stroke patients. *J Stroke Cerebrovasc Dis* 2013;22:e582-e589.
78. Strecker JK, Schmidt A, Schäbitz WR, Minnerup J. Neutrophil granulocytes in cerebral ischemia: evolution from killers to key players. *Neurochem Int* 2017;107:117-126.
79. Maestrini I, Strbian D, Gautier S, Haapaniemi E, Moulin S, Sairanen T, et al. Higher neutrophil counts before thrombolysis for cerebral ischemia predict worse outcomes. *Neurology* 2015;85:1408-1416.
80. Nardi K, Milia P, Eusebi P, Paciaroni M, Caso V, Agnelli G. Admission leukocytosis in acute cerebral ischemia: influence on early outcome. *J Stroke Cerebrovasc Dis* 2012;21:819-824.
81. Kumar AD, Boehme AK, Siegler JE, Gillette M, Albright KC, Martin-Schild S. Leukocytosis in patients with neurologic deterioration after acute ischemic stroke is associated with poor outcomes. *J Stroke Cerebrovasc Dis* 2013;22:e111-e117.
82. Tian C, Ji Z, Xiang W, Huang X, Wang S, Wu Y, et al. Association of lower leukocyte count before thrombolysis with early neurological improvement in acute ischemic stroke patients. *J Clin Neurosci* 2018;56:44-49.
83. Pósfai É, Marton I, Szőke A, Borbényi Z, Vécsei L, Csomor A, et al. Stroke in essential thrombocythemia. *J Neurol Sci* 2014;336:260-262.
84. Furlan JC, Fang J, Silver FL. Outcomes after acute ischemic stroke in patients with thrombocytopenia or thrombocytosis. *J Neurol Sci* 2016;362:198-203.
85. Turcato G, Cappellari M, Follador L, Dilda A, Bonora A, Zannoni M, et al. Red blood cell distribution width is an independent predictor of outcome in patients undergoing thrombolysis for ischemic stroke. *Semin Thromb Hemost* 2017;43:30-35.
86. Pinho J, Marques SA, Freitas E, Araújo J, Taveira M, Alves JN, et al. Red cell distribution width as a predictor of 1-year survival in ischemic stroke patients treated with intravenous thrombolysis. *Thromb Res* 2018;164:4-8.
87. Furlan JC, Fang J, Silver FL. Acute ischemic stroke and abnormal blood hemoglobin concentration. *Acta Neurol Scand* 2016;134:123-130.
88. Yi X, Lin J, Wang C, Zhou Q. CYP genetic variants, CYP metabolite levels, and neurologic deterioration in acute ischemic stroke in Chinese population. *J Stroke Cerebrovasc Dis* 2017;26:969-978.
89. Yi X, Ming B, Wang C, Chen H, Ma C. Variants in COX-2, PTGIS, and TBXAS1 are associated with carotid artery or intracranial arterial stenosis and neurologic deterioration in ischemic stroke patients. *J Stroke Cerebrovasc Dis* 2017;26:1128-1135.
90. Yi X, Wang C, Zhou Q, Lin J. Interaction among COX-2, P2Y1 and GPIIIa gene variants is associated with aspirin resistance and early neurological deterioration in Chinese stroke patients. *BMC Neurol* 2017;17:4.

 https://doi.org/10.5853/jos.2019.01522

4.3. Artículo 3

Genome-wide association study of white blood cell counts in ischemic stroke patients

RESUMEN

Antecedentes y objetivo: Las células inmunes desempeñan un papel clave en las primeras 24 horas después de un ictus (fase aguda), y se asocian con el pronóstico de esta patología. El objetivo fue encontrar factores de riesgo genéticos asociados con el recuento de leucocitos durante la fase aguda del ictus.

Métodos: Se incluyeron pacientes con ictus isquémico del cual se tenían datos del número de leucocitos durante las primeras 24 h. Se realizaron un Genome-Wide Association Study (GWAS) y estudios de expresión génica *in silico*.

Resultados: Nuestro GWAS, que incluyó a 2.064 (Discovery) y 407 (Replicación) pacientes, reveló un nuevo locus (14q24.3) asociado con los recuentos de leucocitos. Después del análisis conjunto (n = 2.471) cinco polimorfismos más alcanzaron significación genómica (p <5x10-8). El locus 14q24.3 se asoció con el estado neurológico agudo del ictus (rs112809786, p = 0.036) y con la expresión de los genes *ACOT1* y *PTGR2*. Los polimorfismos previos asociados con los recuentos de leucocitos en otras poblaciones no mostraron ninguna asociación significativa en nuestro estudio.

Conclusiones: Hemos encontrado el primer locus asociado con los recuentos de leucocitos en el ictus isquémico durante la fase aguda, también asociado con el pronóstico agudo. El análisis genético de los endofenotipos agudos podría ser útil para encontrar los factores genéticos asociados con el pronóstico del ictus. Nuestros hallazgos sugieren una modulación diferente de las células inmunes en el ictus en comparación con condiciones no patológicas.

GENOME-WIDE ASSOCIATION STUDY OF WHITE BLOOD CELL COUNTS IN ISCHEMIC STROKE PATIENTS

Nuria P Torres-Aguila, MSc[1,2], Caty Carrera, MSc[2], Anne-Katrine Giese, MSc[3], Natalia Cullell, MSc[4], Elena Muiño, MD[1], Jara Cárcel-Marquez, MSc[1], Cristina Gallego-Fabrega, PhD[1,4], Jonathan Gonzalez-Sanchez, MSc[1,4,5], María del Mar Freijo, MD[6], José Álvarez-Sabín, MD, PhD[7], Carlos Molina, MD, PhD[7], Marc Ribó, MD, PhD[7], Jordi Jimenez-Conde, MD, PhD[8], Jaume Roquer, MD, PhD[8], Tomás Sobrino, PhD[9], Francisco Campos, MD[9], José Castillo, MD[9], Lucia Muñoz-Narbona, PhD[10], Elena Lopez-Cancio, MD, PhD[11], Antoni Dàvalos, MD, PhD[10], Rosa Diaz-Navarro, MD[12], Silvia Tur, MD[12], Cristòfol Vives-Bauza, PhD[12], Gemma Serrano-Heras, PhD[13], Tomás Segura, MD[13], Jerzy Krupinski, MD, PhD[5,14], Raquel Delgado-Mederos, MD[15], Joan Martí-Fàbregas, MD, PhD[15], Laura Heitsch, MD[16,17], Laura Ibañez, PhD[18], Carlos Cruchaga, PhD[18], Natalia S Rost, MD, PhD[3], Joan Montaner, MD, PhD[2,19], Jin-Moo Lee, MD, PhD[17], Israel Fernandez-Cadenas, PhD[1,†]

[1.]Stroke Pharmacogenomics and Genetics Laboratory, IR-HSCSP, Barcelona Spain

[2.]Neurovascular Research Laboratory, VHIR, Barcelona, Spain

[3.]Neurology Department, Massachusetts General Hospital, Boston, MA, USA

[4.]Stroke Pharmacogenomics and Genetics Laboratory, FMT, HUMT, Terrassa, Spain

[5.]School of Healthcare Science, MMU, Manchester, UK

[6.]Neurology Department, Hospital de Basurto, Bilbao, Spain

[7.]Stroke Unit, Neurology Department, Hospital Universitari Vall d'Hebron, Barcelona, Spain

[8.]Neurology Department, IMIM-Hospital del Mar; Neurovascular Research Group, IMIM; UAB/ DCEXS-Universitat Pompeu Fabra, Barcelona, Spain

[9.]Clinical Neurosciences Research Laboratory, IDIS, Santiago de Compostela, Spain

[10.]Neurosciences Department, Hospital Germans Trias I Pujol-UAB, Barcelona, Spain

[11.]Stroke Unit, Hospital Universitario Central de Asturias, Oviedo, Spain

[12.]Neuroscience Laboratory, IdISBa, Mallorca, Spain

[13.]Neurology Department, University Hospital of Albacete, Albacete, Spain

[14.]Neurology Service, HUMT, Terrassa, Spain

[15.]Stroke Unit, Neurology Department, HSCSP, Barcelona, Spain

[16.]Emergency Medicine, WUSM, Saint Louis, MO, USA

[†]Corresponding author: Israel Fernández-Cadenas, Address: C/ Sant Antoni M[a]Claret, 167, 08025 Barcelona, Spain, Phone: +34 932746000, israelcadenas@yahoo.es, Twitter: @PharmaEpigenLab.

For the online-only Data Supplement please see https://www.ahajournals.org/journal/str.

[17]Neurology Department, WUSM, Saint Louis, MO, USA

[18]Psychiatry Department, WUSM, Saint Louis, MO, USA

[19]Stroke Research Program, IBiS/Hospital Universitario Virgen del Rocío/CSIC/University of Seville & Department of Neurology, Hospital Universitario Virgen Macarena, Seville, Spain

Abstract

Background and Purpose—Immune cells play a key role in the first 24h post-stroke (acute phase), being associated with stroke outcome. We aimed to find genetic risk factors associated with leukocyte counts during the acute phase of stroke.

Methods—Ischemic stroke patients with leukocyte counts data during the first 24h were included. Genome-Wide Association Study (GWAS) and gene expression studies were performed.

Results—Our GWAS, which included 2,064 (Discovery) and 407 (Replication) patients, revealed a new locus (14q24.3) associated with leukocyte counts. After Joint analysis (n=2,471) five more polymorphisms reached genome-wide significance ($p<5x10^{-8}$). The 14q24.3 locus was associated with acute stroke outcome (rs112809786, $p=0.036$) and with *ACOT1* and *PTGR2* gene expression. Previous polymorphisms associated with leukocyte counts in general-population did not show any significance in our study.

Conclusions—We have found the first locus associated with leukocyte counts in ischemic stroke, also associated with acute outcome. Genetic analysis of acute endophenotypes could be useful to find the genetic factors associated with stroke outcome. Our findings suggested a different modulation of immune cells in stroke compared to healthy conditions.

Keywords

Genome Wide Association Study; leukocyte; ischemic stroke; Genetic; Association Studies; Inflammation; Risk Factors; Functional Genomics

INTRODUCTION

Ischemic stroke (IS) is a complex disease. Even though there are several genes associated with stroke risk, little is known about the genetics behind stroke outcome. Only two genome-wide association studies (GWAS)[1,2] have found two loci associated with 3-month disability post-stroke. The clinical complexity and multifactorial modulation of stroke outcome hinder the discovery of genetic associations. To avoid this problem, one strategy is to individually analyze internal factors of stroke (endophenotypes) associated with stroke outcome.

The acute phase of stroke (first 24h) dramatically influences the short and long-term outcome. During this phase there is an important inflammatory and immune response. Immune cell levels, such as white blood cell counts (WBCc), have been associated with worse outcome[3]. Currently, there are several GWAS of WBCc performed in different populations[4,5] that have found more than 20 loci, but none have been related to stroke.

We aimed to find genetic risk factors associated with WBCc during the acute phase of IS due to its relevant role in post-stroke outcome.

MATERIAL AND METHODS

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Study population

The IS patients included were >18 years old, had a measured neurological deficit on the National Institute of Health Stroke Scale (NIHSS) within 6h post-stroke onset, and underwent blood analysis with WBCc data within 24h post-stroke onset. The outcomes studied were NIHSS at 24h (acute) and modified Rankin Scale (mRS) at 3 months (long-term). Written informed consent and ethical committee approval was obtained for all participants. Additional cohort details are in the online-only Data Supplement (Supplemental Methods).

Genetic analysis

Genotyping, imputation and quality controls are described in the online-only Data Supplement (Supplemental Methods). We performed the SNP association analysis using SNPTEST software and the Joint analysis was performed with METAL software.

To study expression Quantitative Trait Loci (eQTLs) associated with the candidate SNPs, we performed *in silico* analyses with data from the Genotype-Tissue Expression (GTEx) Project. The statistical analyses were performed using SPSS software v17.0 (detailed in Supplemental Methods).

RESULTS

In this study, 2,132 European Caucasian IS patients were recruited in the Discovery cohort and 407 American Caucasian patients were included in the Replication cohort. Detailed characteristics of the Discovery cohort are included in the online-only Data Supplement (Table I).

In the Discovery cohort, WBCc was independently associated with long-term and acute outcome (3-month-mRS, p=3.88x10^{-4}; 24h-NIHSS, p=6.81x10^{-3}). Besides, several variables were associated with WBCc in the univariate analysis (Table I in the online-only Data Supplement). After stepwise linear regression, age, diabetes mellitus, baseline-NIHSS and atherothrombotic etiology remained significant and were included as covariates in the genetic analyses.

In the Discovery GWAS, 68 participants were excluded for lacking some covariate, afterwards 2,064 patients and 7,340,003 SNPs were analyzed. We found one locus in chromosome 14 associated with WBCc (Figure 1). Replication showed the same directionality as in the Discovery cohort for the six suggestive SNPs (p<10^{-6}) of the

14q24.3 locus, three of them with nominal significance ($p<0.05$; Table 1). In the Joint analysis all six SNPs reached genome-wide significance ($p<5x10^{-8}$; Table 1).

We performed an additional genetic association analysis in a subgroup of patients with data collected within the first 6h after ischemic stroke, in order to test the influence of time collection. In this sub-cohort that represents 69.67% of the Discovery cohort of the study (n=1,438), the 14q24.3 locus was associated with WBCc but did not reach genome-wide significance (Table II in the online-only Data Supplement). Additionally, in order to test whether the identified locus was associated with a specific cell type of WBCc, we performed genetic association analysis with the neutrophil and lymphocyte content for a subgroup of patients included in the Discovery cohort (neutrophil counts, n=1,199; lymphocyte counts, n=793). The 14q24.3 locus was not significantly associated specifically with neutrophil or lymphocyte counts (Figure I in the online-only Data Supplement).

We tested the association of each genome-wide significant SNP in the Joint analysis with stroke outcome. For acute outcome, rs112809786 was associated with 24h-NIHSS ($p=0.036$) and four other SNPs showed a trend ($p<0.1$; Table III in the online-only Data Supplement), with the same directionality of WBCc association (allele risk was associated with worse NIHSS score). However, for long-term outcome, no association was found (Table III in the online-only Data Supplement).

We searched previously described SNPs reported in GWAS of WBCc, in different ethnicities and conditions, in our GWAS of WBCc in IS patients. Twenty-eight out of 30 SNPs were present in our GWAS, but none was significantly associated with WBCc after Bonferroni correction ($p>0.002$; Table IV in the online-only Data Supplement).

In silico approaches showed that the 14q24.3 locus was associated with eQTLs of acyl-CoA thioesterase 1 (*ACOT1*) in adrenal gland tissue ($p=2.0x10^{-5}$), and prostaglandin reductase 2 (*PTGR2*) in thyroid tissue ($p=1.3x10^{-5}$). For *ACOT1*, minor alleles were associated with higher gene expression, while for *PTGR2*, minor alleles were associated with lower expression.

DISCUSSION

We have described the first locus (14q24.3) associated with the WBCc during the acute phase of stroke, confirmed by the Joint analysis of two independent populations. Additionally, this locus was not specific to any cell type population (neutrophils or lymphocytes) and was not influenced by the time of collection, at least during the first 6h after stroke. In addition, previous SNPs reported to be associated with WBCc in other healthy populations and diseases were not consistently associated with WBCc in stroke, suggesting a different genetic modulation for leukocyte proliferation/activation. Furthermore, we confirmed the previously reported association of WBCc with stroke outcome: we found an independent association of WBCc with acute and long-term stroke outcome with a positive correlation, whereby higher WBCc was associated with higher NIHSS and mRS, and was consequently associated with a worse outcome. Moreover, the rs112809786 SNP within the 14q24.3 locus was associated with acute stroke outcome, in the same directionality as WBCc (the risk allele group, associated with high WBCc, had higher

24h-NIHSS mean than the no-risk allele group). This finding supports the idea that individually analyzing endophenotypes could be an interesting strategy to find new genetic associations for stroke outcome. Unexpectedly, rs112809786 was not associated with long-term outcome. One reason could be that other factors that occur after the acute phase (infection, rehabilitation, etc.) are influencing 3-month-mRS. Our results suggested that the genetic factors associated with WBCc may have more influence on acute outcome than on long-term outcome.

Besides, *in silico* approaches revealed eQTLs of *ACOT1* and *PTGR2* genes for 14q24.3 locus. *ACOT1*, an acyl-CoA thioesterase family gene member, is described as being involved in the regulation of lipid metabolism by the peroxisome proliferator-activated receptor alpha (PPARα), which participates in the activation and proliferation cascade of some immune cells. Regarding *PTGR2*, it is described for the metabolism of prostaglandins and in the regulation of other PPAR family genes. Interestingly, *PTGR2* has been found downregulated in monocytes of patients with leukemia[6], which is consistent with our *in silico* observation (14q24.3 alleles associated with higher WBCc were associated with lower *PTGR2* expression). Thus, *ACOT1* and *PTGR2* are candidate genes for regulating leukocyte proliferation and contributing to worse acute stroke outcome; however, further studies are necessary to confirm this potential association.

As a limitation, we could not exclude patients with conditions that may influence WBCc. However, this unbiased inclusion suggests that our findings could be applicable to any type of IS patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

NPTA, CC, AKG and NR performed the data analysis, prepared the figures, and drafted a significant proportion of the manuscript. NC, EM, JCM, CGF, JGS, MF, JAS, CM, MR, JJC, JR, TS, FC, JC, ELC, AD, RDN, ST, CVB, GSH, TS, JK, RD, JMF, LH, LI, CC, JM and JML contributed to data acquisition and substantially edited the manuscript, giving their final approval. IFC contributed to the study conception and design and to drafting the manuscript.

SOURCES OF FUNDING

The Carlos III Health Institute, the European Regional Development Fund (ERDF): Redes Temáticas de Investigacion Cooperativa en Salud, Invictus-Plus (RD16/0019/0002, RD16/0019/0010), CPII17/00021, CPII17/00027, CP14/00154. Fundació Marató de TV3 (76/C/2011), NIH (K23 NS099487, and R01NIH NS085419), Spanish Ministry of Science and Innovation (PI051737, PI10/02064, PI12/01238, PI15/00451, PI18/00022), Barnes-Jewish Hospital Foundation.

DISCLOSURES

Dr Lee reports grants from Biogen outside the submitted work. Dr Heitsch reports grants from NIH NINDS K23, grants from AHA Career Development Grant, and grants from Emergency Medicine Foundation Career Development Grant during the conduct of the study; personal fees from Genentech outside the submitted work. Dr Krupinski reports received personal fees for conferences from Boehringer, Bayer, Ferrer. Dr Ribo reports personal fees and other from Anaconda Biomed, personal fees from Medtronic, personal fees from Stryker, personal fees and other from Cerenovus, and personal fees from Apta Targets outside the submitted work.

REFERENCE

1. Mola-Caminal M, Carrera C, Soriano-Tárraga C, Giralt-Steinhauer E, Díaz-Navarro RM, TUR SS, et al. PATJ Low Frequency Variants Are Associated With Worse Ischemic Stroke Functional Outcome. Circ. Res 2019;124:114–120. [PubMed: 30582445]
2. Soderholm 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:e1271–e1283. [PubMed: 30796134]
3. Nardi K, Milia P, Eusebi P, Paciaroni M, Caso V, Agnelli G. Admission leukocytosis in acute cerebral ischemia: Influence on early outcome. J. Stroke Cerebrovasc. Dis 2012;21:819–824. [PubMed: 21703875]
4. Keller MF, Reiner AP, Okada Y, van Rooij FJA, Johnson AD, Chen M-H, et al. Trans-ethnic meta-analysis of white blood cell phenotypes. Hum. Mol. Genet 2014;23:6944–6960. [PubMed: 25096241]
5. Jain D, Hodonsky CJ, Schick UM, Morrison J V, Minnerath S, Brown L, et al. Genome-wide association of white blood cell counts in Hispanic/Latino Americans: the Hispanic Community Health Study/Study of Latinos. Hum. Mol. Genet 2017;26:1193–1204. [PubMed: 28158719]
6. Maffei R, Bulgarelli J, Fiorcari S, Bertoncelli L, Martinelli S, Guarnotta C, et al. The monocytic population in chronic lymphocytic leukemia shows altered composition and deregulation of genes involved in phagocytosis and inflammation. Haematologica. 2013;98:1115–1123. [PubMed: 23349302]

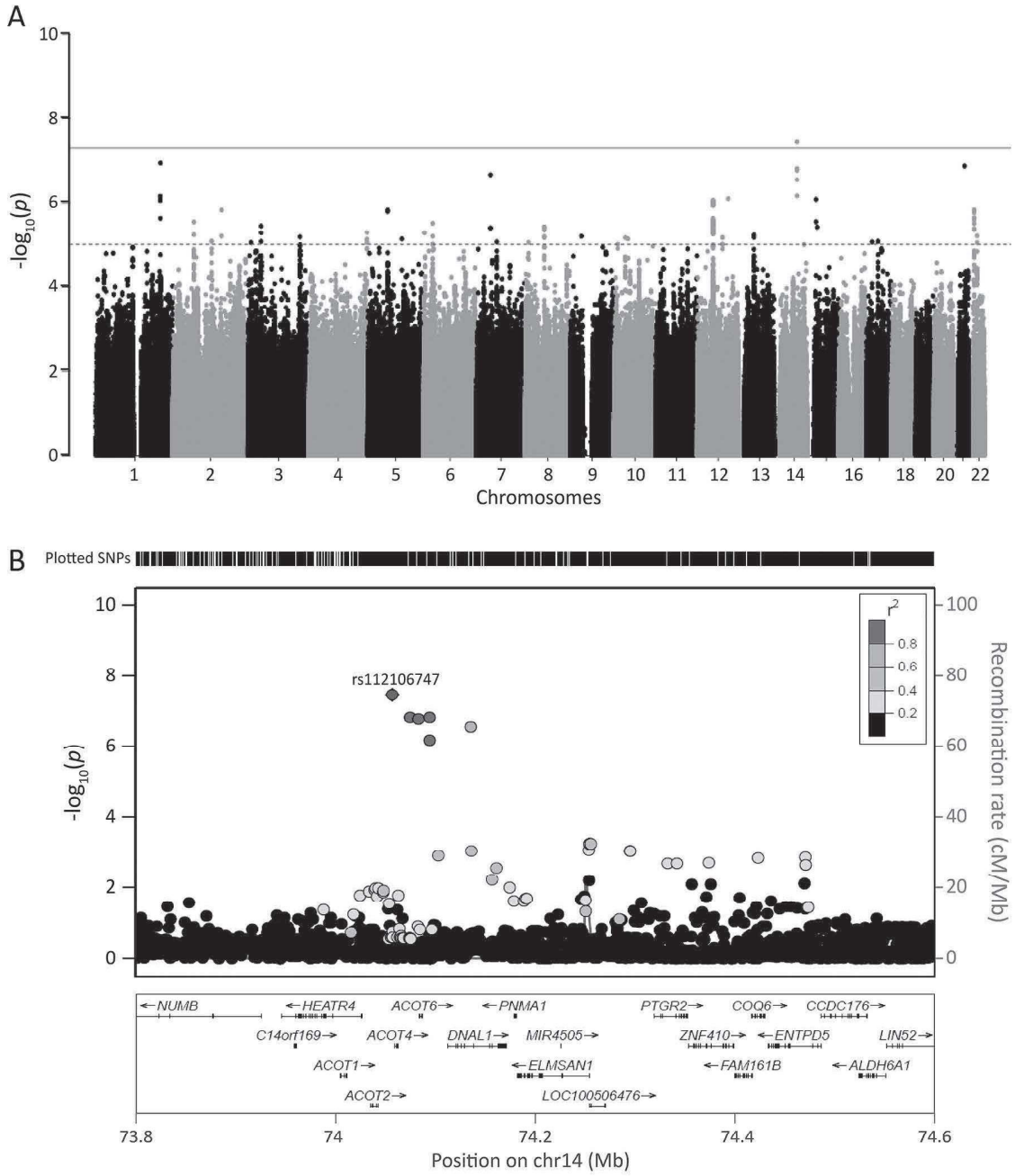

Figure 1. Results of the Discovery analysis.

A) Manhattan plot of the Discovery analysis. Lines represent genome-wide significance ($p<5.0\text{x}10^{-8}$) and suggestive threshold ($p<1.0\text{x}10^{-5}$). B) Zoom plot performed on the LocusZoom portal of the 14q24.3 locus (rs112106747, top SNP). The X and Y axes show chromosome location and negative logarithm to base 10 of *p*-values ($-\log_{10}(p)$), respectively.

Table 1.

List of SNPs in the 14q24.3 locus associated with WBCc.

SNP	A1	A2	Discovery (n=2,064)			Replication (n=407)		Joint (n=2,471)	
			MAF	P value	B±SE	P value	B±SE	P value	B±SE
rs112106747	G	A	0.016	**$3.57x10^{-8}$**	0.67±0.12	0.061	0.43±0.23	**$8.05x10^{-9}$**	0.62±0.1
rs113898499	A	G	0.016	$1.54x10^{-7}$	0.64±0.12	0.055	0.44±0.23	**$2.78x10^{-8}$**	0.60±0.1
rs113492829	A	G	0.016	$1.67x10^{-7}$	0.63±0.12	0.058	0.44±0.23	**$3.16x10^{-8}$**	0.59±0.1
rs112809786	A	G	0.017	$6.89x10^{-7}$	0.59±0.12	0.047	0.49±0.25	**$3.16x10^{-8}$**	0.57±0.1
rs78476982	C	T	0.016	$1.54x10^{-7}$	0.64±0.12	0.018	0.59±0.25	**$8.01x10^{-9}$**	0.63±0.1
rs74995185	A	G	0.016	$2.86x10^{-7}$	0.63±0.12	0.032	0.72±0.33	**$2.51x10^{-8}$**	0.64±0.1

Genome-wide significant P values are in bold. A1 = Allele 1 (minor allele); A2 = Allele 2; MAF = Minor allele frequency; n = total sample size; B = beta coefficient for minor allele; SE = Standard Error.

GENOME-WIDE ASSOCIATION STUDY OF WHITE BLOOD CELL COUNTS IN ISCHEMIC STROKE PATIENTS

SUPPLEMENTAL MATERIAL

SUPPLEMENTAL METHODS

Study population

In our Discovery Cohort we included 2,128 Spanish ischemic stroke patients. For replication analysis, we used two external European cohorts with a total of 407 patients (222 and 185, respectively). Etiologic subgroups of ischemic stroke patients were classified following the TOAST[1] criteria. Ischemic stroke patients were recruited as part of the GENISIS[2] and GODS[3] projects for the Discovery Cohort, and as part of the GASROS[4] project for the Replication Cohort:

GENISIS project: Genetics of Early Neurological Instability after Ischemic Stroke (GENISIS) is an international study currently recruiting patients from four different locations: United States, Finland, Poland, and Spain. The inclusion criteria for the GENISIS study are IS patients (age ≥18 years) collected from 2003 to 2017 with a measurable neurologic deficit on the National Institutes of Health Stroke Scale (NIHSS)within 6 hours of last known normal. Patients who received endovascular thrombectomy, or for whom consent and/or a blood sample could not be obtained were excluded. For our study we only included European Caucasian patients.

GODS project: The Genetic contribution to functional Outcome and Disability after Stroke (GODS) project is a study that aimed to find genetic factors associated with stroke outcome. All participants met the following criteria: (1) European descent, aged >18 years, diagnosis of IS in the anterior vascular territory; (2) assessed by a neurologist during the acute phase of stroke; (3) initial stroke severity >4, according to the NIHSS; (4) information on post-stroke functional status at 3 months (or alternatively between 3 and 6 months); (5) evidence of acute IS in a neuroimaging study; (6) lack of concomitant disease. Individuals with stroke recurrence during the follow-up period were excluded, in addition to posterior vascular territory and lacunar strokes.

GASROS project: Genes Associated with Stroke Risk and Outcomes Study (GASROS) is prospective Institutional Review Board-approved study that enrolled ischemic stroke patients between 2007–2011 with an magnetic resonance imaging (MRI) performed within 48 hours. GASROS is a cross-sectional, hospital based cohort of consecutive adults admitted to neurology service with diagnosis of ischemic stroke confirmed by neuroimaging. Exclusion criteria include inability to obtain informed consent, or a

verified diagnosis of secondary cerebral ischemia (vasculitis, subacute bacterial endocarditis, etc.). All patients were evaluated emergently by a neurologist at the time of admission, and clinical and laboratory data were abstracted from corresponding medical records. The long-term functional outcomes were assessed using the mRS score collected either in person or via a telephone interview at 3–6 months post-stroke. Only patients with data of leukocyte counts during the first 24h post-stroke were included.

Statistical analysis

Statistical Analyses were performed using SPSS statistical package, version 17.0 (IBM, Chicago, US). No-normal distribution of continuous variable was assessed by Kolmogorov-Smirnov and Shapiro-Wilk tests. Correlation of continuous variables were assessed by Spearman Test. The association between continuous variables and categorical variables were assessed by Mann Whitney U test. Univariate General lineal model and Stepwise Linear Regression was performed for test independent association of variables.

Genetic analysis

Genotyping of the Discovery cohort was performed with the Human Core Exome chip (Illumina). For the Replication cohort, 222 participants were genotyped with Axion-Biobank chip (Affimetrix) and 185 with Human Core Exome chip (Illumina).

Quality controls of genotyped data were: test of missingness (≤5%), relatedness (pi-hat > 0.18), heterozygosity, and sex discrepancies, for samples; and for SNPs, genotype call rate (<95%), Hardy-Weinberg equilibrium ($p < 1\times10\text{-}6$), and exclusion of A/T and C/G SNPs. Mitochondrial and sexual chromosomes (X/Y) were also excluded. Cohorts were aligned with 1000G Phase 3 Panel reference and frequency alleles were compared. Multidimensional scaling plots with principal components were performed and were used to identify and remove ethnic outliers. Principal components were also used to adjust for population stratification in the downstream analyses.

Genotype imputation was performed on Michigan Imputation Server Portal[5] using 1000G Phase 3v5 panel. SNPs with r^2 <30%, SNPs with minor allele frequency (MAF) <1%, and insertions and deletions were removed. Finally, a total of 7,340,003 SNPs were included in the Discovery analysis.

We performed a SNP association analysis using SNPTEST[6–8] software, including as covariates age, sex, principal components, and associated cofounders found by stepwise linear regression. Using the METAL software, we performed a joint analysis with the GWAS summary data of the cohorts.

Functional annotation

To study possible eQTLs associated with our candidate SNPs we performed *in silico* analyses using the HaploReg v4.1 web-tool of Broad Institute, and the data available on GTEx portal, from The Genotype-Tissue Expression (GTEx) Project[9].

SUPPLEMENTAL FIGURES

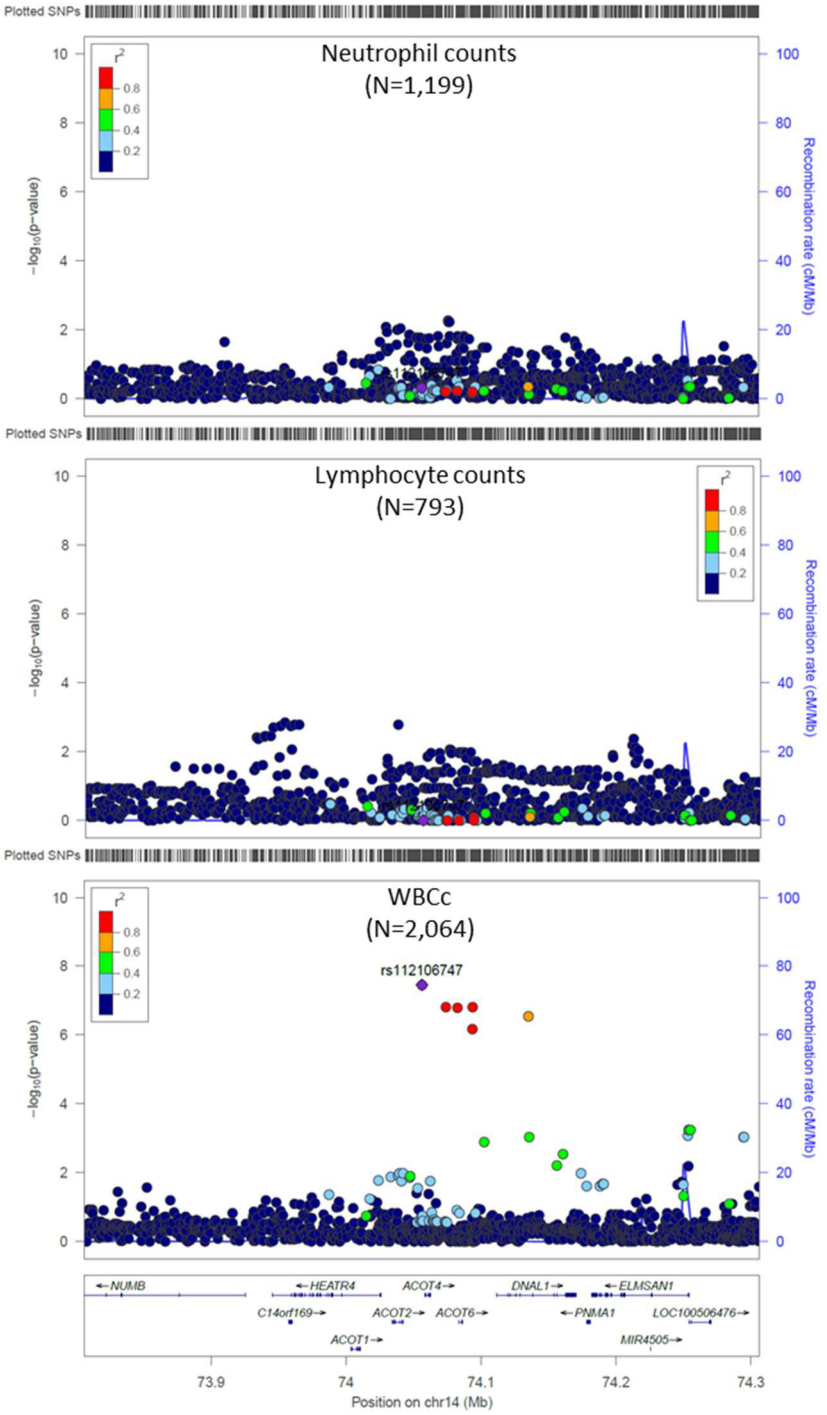

Figure I. Comparative image of the locus 14q24.3 associated with WBCc, neutrophil counts and lymphocyte counts in the Discovery cohort. X and Y axis show chromosome location and negative logarithm to base 10 of p-values (-log10(p)), respectively. N = number of patients included in the analysis. WBCc = white blood cell counts.

SUPPLEMENTAL TABLES

Table I. Detailed characteristics of the Discovery cohort and association of variables with WBCc.

Characteristics	Discovery Cohort (n=2,132)	Association with WBCc (p value)
Age (SD)	73.4 (12.5)	$3.37x10^{-4}$
Woman (%)	959 (45.1)	n.s.
Hypertension (%)	1,457 (68.3)	n.s.
Diabetes mellitus (%)	575 (26.9)	0.012
Atrial fibrillation (%)	676 (31.8)	n.s.
WBCc (SD)	8.8 (3.1)	---
NIHSS at baseline (SD)	10.6 (7.2)	$8.52x10^{-8}$
IV-tPA (%)	1,154 (54.1)	n.s.
TOAST LAA (%)	341 (16.0)	$7.42x10^{-4}$
TOAST CE (%)	906 (42.5)	0.042
TOAST SVS (%)	158 (7.4)	n.s.
Other TOAST categories (%)	729 (34.2)	0.037
24h NIHSS (SD)	7.7 (7.8)	$6.0x10^{-6}$
3-month mRS (SD)	2.5 (2.1)	$1.91x10^{-9}$

WBCc = white blood cell counts; n = number of samples; n.s. = non-significant; SD = standard deviation; % = percentage of presence; IV-tPA = intravenous treatment with tissue plasminogen activator; LAA = Large Artery Atherosclerosis; CE = Cardioembolic stroke; SVS = lacunar or small vessel stroke.

Table II. Detailed p values for the association of the SNPs of 14q24.3 locus with WBCc in the Discovery cohort and a sub-cohort of patients with data collected within first 6h after stroke onset.

SNP	A1	A2	Discovery (n=2,064)			Sub-cohort (n=1,438)	
			MAF	P value	B±SE	P value	B±SE
rs112106747	G	A	0.016	$3.57x10^{-8}$	0.67±0.12	$1.27x10^{-7}$	0.81±0.15
rs113898499	A	G	0.016	$1.54x10^{-7}$	0.64±0.12	$7.73x10^{-7}$	0.76±0.15
rs113492829	A	G	0.016	$1.67x10^{-7}$	0.63±0.12	$8.84x10^{-7}$	0.74±0.15
rs112809786	A	G	0.017	$6.89x10^{-7}$	0.59±0.12	$4.97x10^{-6}$	0.66±0.14
rs78476982	C	T	0.016	$1.54x10^{-7}$	0.64±0.12	$7.73x10^{-7}$	0.76±0.15
rs74995185	A	G	0.016	$2.86x10^{-7}$	0.63±0.12	$1.06x10^{-6}$	0.77±0.16

Table III. Detailed p values for the association of the SNPs of 14q24.3 locus with stroke outcome.

SNPs of 14q24.3 locus	NIHSS at 24h (n = 1,802)			3-month mRS (n = 1,951)		
	AA	Aa	p-value[a]	AA	Aa	p-value[b]
rs112106747	8.13±7.84	9.70±7.83	0.068	2.45±2.05	2.58±2.16	0.610
rs113898499	8.14±7.84	9.58±7.85	0.087	2.45±2.05	2.57±2.15	0.632
rs113492829	8.14±7.85	9.58±7.70	0.067	2.45±2.05	2.56±2.13	0.655
rs112809786	8.13±7.85	9.70±7.49	0.036	2.45±2.05	2.49±2.10	0.836
rs78476982	8.14±7.84	9.58±7.85	0.087	2.45±2.05	2.57±2.15	0.632
rs74995185	8.15±7.84	9.34±7.86	0.171	2.45±2.05	2.53±2.21	0.722

Significant p values are in bold. NIHSS = National Institute of Helath Stroke Scale; mRS = modified Rankin Scale; AA = Homozygore reference allele group; Aa = heterozygote allele group; aa = homozygote allele group. [a]Kruskal-Wallis Test, [b]Chi-Square Test.

Table IV. List of previous described SNPs associated with WBCc in healthy population.

SNP	Locus	Reference	P value Discovery	P value Joint
rs4657616	IFI16	J. Li et al (2013)[10]	0.016	0.023
rs2518564	AIM2	M.F. Keller et al. (2014)[11]	0.221	0.468
rs2814778	DARC	D.R. Crosslin et al. (2011)[12]; D. Jain et al. (2017)[13]	0.961	0.961
rs12075	DARC	D.R. Crosslin et al. (2011)[12]	0.221	0.270
rs114477531	HAAO	D. Jain et al. (2017)[13]	NA	NA
rs6734238	IL1F10	M.F. Keller et al. (2014)[11]	0.609	0.604
rs1449263	ITGA4	M.A. Nalls et al. (2011)[14]	0.094	0.166
rs10932765	ARPC2	M.F. Keller et al. (2014)[11]	0.047	0.029
rs9880192	C3orf27	M.A. Nalls et al. (2011)[14]	0.734	0.835
rs4328821	C3orf27	M.A. Nalls et al. (2011)[14]; Y. Okada et al. (2011)[15]	0.300	0.027
rs549280	CXCL2	M.F. Keller et al. (2014)[11]	0.499	0.142
rs1371799	MTHFD2L, CXCL2	A.P. Reiner et al. (2011)[16]	0.570	0.153
rs2517510	HCG22	M.A. Nalls et al. (2011)[14]	0.211	0.262
rs3094212	CDSN	Y. Kamatani et al. (2010)[17]	0.097	0.076
rs2524079	HLA-C	M.A. Nalls et al. (2011)[14]; D. Jain et al. (2017)[13]	0.007	0.004
rs2853946	HLA-C	M.F. Keller et al. (2014)[11]	0.014	0.016
rs4895441	HBS1L	Y. Kamatani et al. (2010)[17]; M.F. Keller et al. (2014)[11]	0.935	0.244
rs9402686	HBS1L	M.F. Keller et al. (2014)[11]	0.904	0.229
rs445	CDK6	Y. Kamatani et al. (2010)[17]; M.F. Keller et al. (2014)[11]; A.P. Reiner et al. (2011)[16]; Y. Okada et al. (2011)[15]	NA	NA

Table IV. Continuation.

rs2380606	SLCO5A1	D. Jain et al. (2017)[13]	0.223	0.196
rs2163950	CCDC26	M.F. Keller et al. (2014)[11]	0.538	0.428
rs10098310	CCDC26	M.F. Keller et al. (2014)[11]; M.A. Nalls et al. (2011)[14]	0.013	0.085
rs10980800	LPAR1	M.A. Nalls et al. (2011)[14]	0.136	0.233
rs12313946	RAP1B	Y. Kamatani et al. (2010)[17]	0.104	0.152
rs3859192	GSDMA	D.R. Crosslin et al. (2011)[12]	0.168	0.065
rs4065321	PSMD3	D.R. Crosslin et al. (2011)[12]; Y. Kamatani et al. (2010)[17]	0.903	0.404
rs4794822	PSMD3	M.F. Keller et al. (2014)[11]; M.A. Nalls et al. (2011)[14]; Y. Okada et al. (2011)[15]	0.264	0.123
rs8078723	CSF3, PSMD3	M.A. Nalls et al. (2011)[14]	0.278	0.130
rs2227336	MED24	D. Jain et al. (2017)[13]	0.270	0.126
rs10411936	EPS15L1	M.A. Nalls et al. (2011)[14]	0.274	0.242

SUPPLEMENTAL REFERENCES

1. Chung J-W, Park SH, Kim N, Kim W-J, Park JH, Ko Y, et al. Trial of ORG 10172 in Acute Stroke Treatment (TOAST) classification and vascular territory of ischemic stroke lesions diagnosed by diffusion-weighted imaging. J. Am. Heart Assoc. 2014;3.

2. Heitsch L, Ibanez L, Carrera C, Pera J, Jimenez-Conde J, Slowik A, et al. Meta-analysis of Transethnic Association (MANTRA) Reveals Loci Associated With Neurological Instability After Acute Ischemic Stroke. In: International Stroke Conference. 2017.

3. Mola-Caminal M, Carrera C, Soriano-Tárraga C, Giralt-Steinhauer E, Díaz-Navarro RM, TUR SS, et al. PATJ Low Frequency Variants Are Associated With Worse Ischemic Stroke Functional Outcome. Circ. Res. 2019;124:114–120.

4. Wu O, Cloonan L, Mocking SJT, Bouts MJRJ, Copen WA, Cougo-Pinto PT, et al. Role of Acute Lesion Topography in Initial Ischemic Stroke Severity and Long-Term Functional Outcomes. Stroke 2015;46:2438–2444.

5. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat. Genet. 2016;48:1284–1287.

6. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat. Genet. 2007;39:906–913.

7. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:661–678.

8. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat. Rev. Genet. [Internet]. 2010;11:499. Available from: https://doi.org/10.1038/nrg2796

9. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat. Genet. 2013;45:580–585.

10. Li J, Glessner JT, Zhang H, Hou C, Wei Z, Bradfield JP, et al. GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American children. Hum. Mol. Genet. 2013;22:1457–1464.

11. Keller MF, Reiner AP, Okada Y, van Rooij FJA, Johnson AD, Chen M-H, et al. Trans-ethnic meta-analysis of white blood cell phenotypes. Hum. Mol. Genet. 2014;23:6944–6960.

12. Crosslin DR, McDavid A, Weston N, Nelson SC, Zheng X, Hart E, et al. Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network. Hum. Genet. 2012;131:639–652.

13. Jain D, Hodonsky CJ, Schick UM, Morrison J V, Minnerath S, Brown L, et al. Genome-wide association of white blood cell counts in Hispanic/Latino Americans: the Hispanic Community Health Study/Study of Latinos. Hum. Mol. Genet. 2017;26:1193–1204.

14. Nalls MA, Couper DJ, Tanaka T, van Rooij FJA, Chen M-H, Smith A V., et al. Multiple Loci Are Associated with White Blood Cell Phenotypes. PLoS Genet. 2011;7:e1002113.

15. Okada Y, Hirota T, Kamatani Y, Takahashi A, Ohmiya H, Kumasaka N, et al. Identification of Nine Novel Loci Associated with White Blood Cell Subtypes in a Japanese Population. PLoS Genet. 2011;7:e1002067.

16. Reiner AP, Lettre G, Nalls MA, Ganesh SK, Mathias R, Austin MA, et al. Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). PLoS Genet. 2011;7:e1002108.

17. Kamatani Y, Matsuda K, Okada Y, Kubo M, Hosono N, Daigo Y, et al. Genome-wide association study of hematological and biochemical traits in a Japanese population. Nat. Genet. 2010;42:210–215.

DISCUSIÓN

5. Discusión

Los resultados de los trabajos que componen esta tesis indican primero, que es posible encontrar nuevos factores de riesgo genéticos para ictus isquémico utilizando cohortes geográficamente homogéneas y segundo, que el análisis de endofenotipos puede ser una herramienta útil para encontrar nuevos factores de riesgo genéticos asociados al pronóstico del ictus isquémico.

Pese a que tener un gran tamaño muestral se ha demostrado fundamental para encontrar resultados significativos en los estudios GWAS, aumentar la homogeneidad de la muestra, aunque sea reduciendo drásticamente el número de pacientes a analizar, puede ser una herramienta útil para encontrar factores de riesgo genético en determinados casos, por ejemplo, al buscar factores de riesgo genético en poblaciones concretas. Así pues, tanto el uso de cohortes homogéneas a nivel geográfico (hipótesis 1) como analizar una variable menos multifactorial como son los endofenotipos (hipótesis 2), permiten aumentar el poder estadístico de los análisis GWAS al reducir la heterogeneidad y facilita el descubrimiento de nuevos factores genéticos asociados al ictus isquémico.

5.1. Factores genéticos de riesgo en el ictus isquémico

Encontrar nuevos factores de riesgo genéticos asociados al ictus isquémico es importante por varias razones, entre las que destacan el descubrimiento de nuevas dianas terapéuticas, la implementación de nuevos tratamientos resultantes del reposicionamiento de fármacos o la mejora de las predicciones de riesgo añadiendo los factores genéticos a las escalas de predicción clínica como el REGICOR, por ejemplo. Asimismo, para poder alcanzar estos hitos, es necesario entender las implicaciones biológicas que tiene el factor génico de riesgo, ya sea a nivel de regulación genética o a nivel de funcionalidad proteica.

En el artículo 4.1 de la presente Tesis se ha realizado un GWAS con 5.227 controles y 1.752 casos, y se ha utilizado la cohorte internacional del MEGASTROKE[45] como primera replicación y una segunda cohorte española independiente de 1.720 participantes como segunda replicación. Utilizando esta estrategia de GWAS caso-control de tres etapas, se ha descrito por primera vez un factor genético de riesgo específico de población española, cumpliendo así con los objetivos 1 y 2 de la presente Tesis, los cuales estaban enfocados en el descubrimiento de nuevos factores genéticos de riesgo para el ictus isquémico. Concretamente, en este trabajo se describe el SNP rs34324185 como factor de riesgo genético para el ictus de tipo lacunar y, además, se demuestra que dicho factor afecta a los niveles de expresión del gen candidato *MAN2B1* (los individuos con el alelo de riesgo del SNP rs34324185 presentaban mayor expresión de *MAN2B1)*. Además, los estudios funcionales descritos en el manuscrito 8.1, demuestran que la enzima α-manosidasa

lisosomal, codificada en el gen *MAN2B1,* es más activa en los pacientes con ictus lacunar que en pacientes de otras etiologías o individuos sanos (objetivo 3 de la Tesis), abriendo la posibilidad de utilizar esta proteína como biomarcador de ictus lacunar. Estos resultados proporcionan nueva información sobre la fisiopatología del ictus isquémico lacunar.

El ictus lacunar se produce por la obstrucción de un vaso pequeño del sistema circulatorio cerebral. Este subtipo de ictus puede ser producido por mecanismos relacionados o no con los vasos sanguíneos pequeños. Los mecanismos de pequeño vaso conocidos son la disfunción endotelial y la degradación de la barrera hematoencefálica (BBB, del inglés 'blood-brain barrier')[81]. El ictus lacunar causado por estos mecanismos se caracteriza por la presencia de lipohialinosis, el cual causa pequeñas hemorragias, agrandamiento de la pared de los vasos y presencia de depósitos de fibrinoides; además, este tipo de ictus lacunar es probable que esté relacionado con la presencia de híperintensidades de sustancia blanca (WMH, del inglés 'white matter hyperintensities'). Asimismo, los posibles mecanismos no relacionados con los vasos sanguíneos pequeños son la aterosclerosis y la fibrilación auricular[81]. Estos mecanismos suelen causar un ictus lacunar de mayor tamaño que los causados por mecanismos de pequeño vaso, además, la asociación con la presencia de WMH es menos clara.

Por lo que respecta a *MAN2B1,* este gen codifica para la enzima α-manosidasa lisosomal. Esta enzima es miembro de la familia 38 de las glicosil-hidrolasas y es necesaria para el catabolismo de los carbohidratos ligados a nitrógeno liberados durante el recambio de glucoproteínas. La α-manosidasa lisosomal se localiza en el lumen del lisosoma e hidroliza los residuos terminales de α-D-manosa en α-D-manósidos.

Es posible observar una conexión entre lisosoma y enfermedad de pequeño vaso y/o ictus lacunar. Los genes que codifican a Catepsin A (*CTSA*) y adenosina desaminasa 2 (*ADA2*) están relacionados con enfermedades de pequeño vaso. En concreto, mutaciones en estos genes causan las enfermedades monogénicas CARASAL y DADA2, respectivamente, que afectan al sistema de pequeño vaso cerebral. Además, las proteínas codificadas en estos genes también se encuentran localizados en el lumen del lisosoma[82], donde también se localiza la proteína Man2b1. Asimismo, un estudio previo también demostró que la sobreexpresión de *MAN2C1*, otro miembro de la familia 38 de las glicosil-hidrolasas parálogo de *MAN2B1*, provocaba la subglicosilación de las proteínas y la regulación positiva de la vía de degradación asociada al retículo endoplasmático[83] pudiendo afectar a las proteínas estructurales.

Teniendo en cuenta la información previa, es plausible que el papel de *MAN2B1* en el ictus lacunar esté relacionado con los dos mecanismos de pequeño vaso descritos para esta patología. Por una parte, *MAN2B1* puede estar causando un malfuncionamiento del

lisosoma y afectando a las células del endotelio (disfunción endotelial), aumentando así el riesgo de ictus lacunar; y, por otra parte, *MAN2B1* puede estar promoviendo la degradación de la BBB mediante la subglicosilación de las proteínas estructurales de la BBB, actuando de manera similar que *MAN2C1*[83], y desencadenando así el ictus lacunar. No obstante, el papel del lisosoma en las enfermedades de pequeño vaso que no sean enfermedades monogénicas, y/o en el ictus lacunar aún está por determinar.

Por otra parte, durante los últimos años se han descrito diversos biomarcadores para identificar los ictus de tipo lacunar[84]. Sin embargo, actualmente ninguno está siendo aplicado en la práctica clínica.

La actividad enzimática α-manosidasa está claramente diferenciada entre subtipos de ictus, siendo muy elevada en ictus de tipo lacunar, y se puede medir en 30 min en una muestra de plasma del paciente. En base a esto, es posible que el ensayo enzimático de la actividad α-manosidasa sea un buen biomarcador para distinguir a los pacientes con ictus isquémico aterotrombótico y/o cardioembólico de los pacientes con ictus lacunar. Asimismo, esta posible aplicación requiere de más ensayos funcionales que prueben la eficacia, sensibilidad y especificidad de este ensayo en la diferenciación de los diferentes subtipos de ictus. También, puesto que el ictus lacunar comparte características de su patología con el ictus hemorrágico, sería interesante analizar la actividad de esta enzima en el ictus hemorrágico para determinar si también tiene relación con esta patología y si es posible utilizarlo como biomarcador para diferenciar este tipo de ictus.

5.2. Endofenotipos en el pronóstico del ictus isquémico

El pronóstico del ictus isquémico es variado y está influenciado por gran cantidad de factores. En la literatura actual hay gran cantidad de estudios que reportan la asociación de diversas variables clínicas con el pronóstico del ictus isquémico, presentando a veces resultados contradictorios.

Asimismo, una de las variables con mayor influencia en el pronóstico final del paciente es la evolución durante la fase aguda del ictus. Por ello, si se quiere mejorar el pronóstico final de esta enfermedad, analizar los factores que influyen en la evolución de los pacientes durante la fase aguda y subaguda del ictus es relevante desde un punto de vista clínico.

En este trabajo se ha realizado una revisión sistemática de los principales factores, tanto clínicos como genéticos, asociados al pronóstico durante la fase aguda del ictus isquémico (objetivo 4 de la Tesis). Los resultados de dicha revisión se han publicado en el artículo 4.2 de la presente Tesis. Los artículos incluidos en la revisión estudiaban el pronóstico durante la fase aguda del ictus a nivel funcional, medido mediante la escala mRS,

y/o a nivel neurológico, medido mediante la escala NIHSS o la escala neurológica canadiense.

De entre las variables clínicas revisadas, aquellas que tienen una asociación más relevante con el pronóstico durante la fase aguda del ictus son: los niveles de glucosa en sangre (o la diabetes mellitus), la presión sanguínea, la fibrilación auricular, el tratamiento con estatinas previo al ictus, la severidad inicial del ictus, la realización de tratamientos en fase aguda (trombectomía mecánica o trombólisis), padecer una complicación neurológica severa (hematoma parenquimatoso) y los niveles de células inmunes.

Estas variables se pueden obtener fácilmente durante las primeras horas de ingreso del paciente. Así pues, pueden ser de gran utilidad para predecir el estado final del paciente. Concretamente, los niveles de células inmunes se asocian de manera independiente con el estado neurológico del paciente[32] (NIHSS a las 24h post-ictus) y con el deterioro neurológico[85,86] (diferencia de NIHSS entre basal y 24h). Además, al ser una variable que se recoge de manera sistemática en la entrada a urgencias de los pacientes, esto facilita la disponibilidad de los datos de cara a realizar futuros estudios. Sin embargo, aún existen gran cantidad de variables clínicas cuya asociación al pronóstico del paciente no es concluyente, poniendo de manifiesto la necesidad de realizar mayor investigación en este tema.

Respecto a los factores genéticos asociados al pronóstico durante la fase aguda del ictus, estos no han sido estudiados en profundidad en la actualidad. Existen diversos genes candidatos reportados en diferentes estudios basados en estrategias de análisis de gen candidato. Sin embargo, ninguno ha sido adecuadamente validado con excepción del polimorfismo rs20417 situado en el gen *COX-2*, el cual se ha reportado asociado al deterioro neurológico en dos estudios independientes[87,88].

Una mejor estrategia para encontrar nuevos factores genéticos asociados al pronóstico del ictus isquémico es la realización de estudios GWAS, pero actualmente no existe ningún estudio que analice el pronóstico durante la fase aguda del ictus isquémico. No obstante, dos estudios GWAS que analizaron el pronóstico funcional del paciente al tercer mes[48,49], encontraron potenciales genes asociados con el grado de discapacidad post-ictus, demostrando la utilidad de este tipo de estudios genéticos para este campo.

Así pues, la revisión realizada recoge los factores clínicos que intervienen en el pronóstico durante la fase aguda del paciente, determina endofenotipos importantes del mismo, y pone de manifiesto la necesidad de realizar más estudios en este campo, incluyendo estudios genéticos que analicen el genoma completo de los pacientes.

5.3. Factores genéticos para el pronóstico del ictus isquémico

Encontrar nuevos factores genéticos asociados al pronóstico del ictus isquémico tiene una gran utilidad de cara a desvelar las moléculas, proteínas y vías metabólicas implicadas en este proceso. Con esa información sería posible encontrar nuevos fármacos para estimular una mayor recuperación del paciente. Asimismo, la heterogeneidad y el origen multifactorial del pronóstico del ictus dificulta la realización de GWAS sobre esta variable y eleva el número de pacientes necesarios para los análisis genéticos de asociación. Una posible estrategia para solventar este problema es analizar de manera individual factores específicos asociados al pronóstico del ictus (endofenotipos).

En el artículo 4.3 se ha profundizado en la implicación del contenido de células inmunológicas en sangre en el pronóstico durante la fase aguda del ictus isquémico, puesto que es uno de los endofenotipos asociados al mismo. Concretamente, se ha realizado un GWAS del número de leucocitos durante la fase agudo a del ictus en 2.064 pacientes y se han replicado los resultados en una segunda cohorte independiente de 407 pacientes. El objetivo de este estudio era encontrar posibles factores de riesgo genéticos asociados al pronóstico durante la fase aguda del ictus isquémico por estar modificando una variable asociada a dicho pronóstico.

En este artículo se describe la asociación del locus 14q24.3 con número de células inmunes durante la fase aguda del ictus isquémico. Estos resultados representan una novedad en este campo puesto que el locus 14q24.3 no ha sido previamente descrito en los GWAS realizados que analizaban los factores genéticos asociados con el número de leucocitos en población sana o en cohortes que presentaban otro tipo de patologías diferentes al ictus. Además, los análisis *in silico* realizados en este trabajo demuestran la asociación de este locus con loci característicos de expresión cuantitativa o eQTL (del inglés, 'expression Quantitative Trait Loci') de dos genes, acil-CoA tioesterasa 1 (*ACOT1)* y prostaglandina reductasa 2 (*PTGR2)*, convirtiéndolos en dos nuevos genes candidatos para la regulación de la proliferación de leucocitos durante la fase aguda del ictus isquémico.

Concretamente, los individuos portadores del alelo de riesgo asociado a mayor número de células inmunes presentan el gen *ACOT1* sobre-expresado en la glándula suprarrenal, mientras que *PTGR2* se ve infra-expresado en tejido tiroideo. Por una parte, *ACOT1* interviene en la regulación del metabolismo de lípidos, mientras que *PTGR2* está asociado al metabolismo de las prostaglandinas; además, ambos genes interactúan con diferentes miembros de la familia de receptores activados por proliferador de peroxisoma (PPAR, del inglés 'peroxisome proliferator activated receptor'), los cuales han sido descritos en la activación y proliferación de algunas células inmunes[89]. Asimismo, pese a que los eQTL de ambos genes se han encontrado en tejidos diferentes al tejido sanguíneo, en la literatura se

ha descrito la regulación negativa de *PTGR2* en monocitos de pacientes con leucemia[90], consensuando con los resultados obtenidos para este gen.

Adicionalmente, en el artículo 4.3 también se describe la asociación del polimorfismo rs112809786, perteneciente al locus 14q24.3, con el pronóstico neurológico agudo del ictus isquémico. Concretamente, el grupo de pacientes portadores del alelo de riesgo asociado a mayor número de células inmunes presentaba una media de NIHSS a las 24h mayor que el grupo con el alelo de no-riesgo. Pese a que para este polimorfismo no se encontraron eQTL significativamente asociados, los demás polimorfismos del locus 14q24.3 sí se asociaron a eQTL de *ACOT1* y *PTGR2*, por lo que es posible que el SNP rs112809786 también esté alterando la expresión de ambos genes teniendo impacto en el pronóstico neurológico tras el ictus isquémico. La asociación encontrada entre el polimorfismo rs112809786 y el estado neurológico agudo del paciente demuestra la utilidad de los análisis génicos de endofenotipos para encontrar nuevos factores de riesgo genéticos asociados al pronóstico del ictus isquémico (objetivo 5 de la presente Tesis).

Sin embargo, aún es necesario profundizar en la investigación de ambos genes para determinar cómo afectan a la proliferación de los leucocitos en la fase aguda del ictus isquémico y su potencial influencia en el deterioro neurológico del paciente. Asimismo, también es necesario continuar con la investigación respecto al polimorfismo rs112809786 para determinar cómo influye éste en la evolución neurológica del ictus isquémico y si tiene relación con los genes candidatos del locus.

Así pues, los resultados obtenidos en el artículo 4.3 validan la estrategia de análisis génico de endofenotipos asociados al pronóstico del ictus isquémico como una buena estrategia para encontrar nuevos factores genéticos asociados al pronóstico del paciente. Esto abre la posibilidad de realizar otros estudios GWAS centrados en otros endofenotipos relevantes y mencionados en el artículo 4.2, con el objetivo de encontrar nuevos factores genéticos asociados al pronóstico neurológico y funcional del ictus isquémico.

CONCLUSIONES

6. Conclusiones

1. Se ha identificado el polimorfismo rs34324185 como nuevo factor de riesgo específico de población española para ictus de tipo lacunar y se ha validado en población internacional.
   1.1. La utilización de poblaciones homogéneas aumenta el poder estadístico de los análisis genómicos al disminuir la heterogeneidad del trasfondo genético.
2. Los análisis funcionales identifican a *MAN2B1* como el gen candidato del locus asociado con el ictus lacunar.
3. Los endofenotipos claramente asociados al pronóstico durante la fase aguda del ictus isquémico son: los niveles de glucosa en sangre, la presión sanguínea, la fibrilación auricular, el tratamiento con estatinas previo al ictus, la severidad inicial del ictus, la realización de tratamientos en fase aguda, padecer una complicación neurológica severa y el número de leucocitos.
4. Se ha identificado un nuevo factor genético asociado al pronóstico del ictus mediante el análisis genético del número de leucocitos durante la fase aguda del ictus.
   4.1. Hemos descrito un nuevo locus (14q24.3) asociado al número de células inmunes en pacientes con ictus isquémico, no descrito previamente en población sana.
   4.2. El polimorfismo rs20417, perteneciente al locus 14q24.3, se asocia al estado neurológico durante la fase aguda del ictus isquémico.
   4.3. El análisis genético de endofenotipos es una nueva estrategia para descubrir nuevos factores genéticos asociados al pronóstico durante la fase aguda del ictus.

BIBLIOGRAFÍA

7. Bibliografía

1. BioRender. Available at: https://biorender.com.

2. Seshadri, S. & Wolf, P. A. Modifiable Risk Factors and Determinants of Stroke. in *Stroke* 217–233 (Elsevier, 2016). doi:10.1016/B978-0-323-29544-4.00015-3

3. Oliveira Filho, J. & Samuels, O. B. Mechanical thrombectomy for acute ischemic stroke. Available at: https://www.uptodate.com/contents/mechanical-thrombectomy-for-acute-ischemic-stroke. (Accessed: 3rd July 2019)

4. Chung, J.-W. *et al.* Trial of ORG 10172 in Acute Stroke Treatment (TOAST) classification and vascular territory of ischemic stroke lesions diagnosed by diffusion-weighted imaging. *J. Am. Heart Assoc.* **3**, (2014).

5. Mackay, J. & Mensah, G. *The Atlas of Heart Disease and Stroke*. (World Health Organization, 2004).

6. Instituto Nacional de Estadística. (2017). Available at: https://www.ine.es. (Accessed: 3rd July 2019)

7. Institute for Health Metrics and Evaluation. Available at: http://www.healthdata.org/spain. (Accessed: 6th June 2019)

8. Meschia, J. F. *et al.* Guidelines for the Primary Prevention of Stroke. *Stroke* **45**, 3754–3832 (2014).

9. O'Donnell, M. J. *et al.* Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* **376**, 112–123 (2010).

10. D'Agostino, R. B. *et al.* General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* **117**, 743–53 (2008).

11. Amor, A. J. *et al.* Prediction of Cardiovascular Disease by the Framingham-REGICOR Equation in the High-Risk PREDIMED Cohort: Impact of the Mediterranean Diet Across Different Risk Strata. *J. Am. Heart Assoc.* **6**, (2017).

12. Brott, T. *et al.* Measurements of acute cerebral infarction: a clinical examination scale. *Stroke* **20**, 864–870 (1989).

13. Kwah, L. K. & Diong, J. National Institutes of Health Stroke Scale (NIHSS). *J. Physiother.* **60**, 61 (2014).

14. Bonita, R. & Beaglehole, R. Recovery of motor function after stroke. *Stroke* **19**, 1497–

500 (1988).

15. Saposnik, G. *et al.* IScore: A risk score to predict death early after hospitalization for an acute ischemic stroke. *Circulation* **123**, 739–749 (2011).

16. Strimbu, K. & Tavel, J. A. What are biomarkers? *Curr. Opin. HIV AIDS* **5**, 463–466 (2010).

17. Adams, H. P. *et al.* Baseline NIH Stroke Scale score strongly predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). *Neurology* **53**, 126–31 (1999).

18. Saposnik, G. *et al.* Variables associated with 7-day, 30-day, and 1-year fatality after ischemic stroke. *Stroke* **39**, 2318–2324 (2008).

19. Boddu, D. *et al.* Predictors of major neurological improvement after intravenous thrombolysis in acute ischemic stroke: A hospital-based study from south India. *Neurol. India* **58**, 403 (2010).

20. Yeo, L. L. *et al.* Early and continuous neurologic improvements after intravenous thrombolysis are strong predictors of favorable long-term outcomes in acute ischemic stroke. *J. Stroke Cerebrovasc. Dis.* **22**, e590–e596 (2013).

21. Siegler, J. E. *et al.* Identification of modifiable and nonmodifiable risk factors for neurologic deterioration after acute ischemic stroke. *J. Stroke Cerebrovasc. Dis.* **22**, e207-13 (2013).

22. Kim, D.-H. *et al.* Factors associated with early dramatic recovery following successful recanalization of occluded artery by endovascular treatment in anterior circulation stroke. *J. Clin. Neurosci.* **46**, 171–175 (2017).

23. Warren, H. R. *et al.* Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415 (2017).

24. Reiner, A. P. *et al.* Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). *PLoS Genet.* **7**, e1002108 (2011).

25. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).

26. Smith, E. C. & Orkin, S. H. Hemoglobin genetics: recent contributions of GWAS and gene editing. *Hum. Mol. Genet.* **25**, R99–R105 (2016).

27. Sare, G. M., Ali, M., Shuaib, A. & Bath, P. M. W. Relationship between hyperacute blood pressure and outcome after ischemic stroke: Data from the VISTA collaboration. *Stroke* **40**, 2098–2103 (2009).

28. Furlan, J. C., Fang, J. & Silver, F. L. Acute ischemic stroke and abnormal blood hemoglobin concentration. *Acta Neurol. Scand.* **134**, 123–130 (2016).

29. Geng, H.-H. *et al.* Early neurological deterioration during the acute phase as a predictor of long-term outcome after first-ever ischemic stroke. *Medicine (Baltimore).* **96**, e9068 (2017).

30. Anrather, J., Iadecola, C. & Hallenbeck, J. Inflammation and Immune Response. in *Stroke* 129-140.e5 (Elsevier, 2016). doi:10.1016/B978-0-323-29544-4.00010-4

31. Strecker, J. K., Schmidt, A., Schäbitz, W. R. & Minnerup, J. Neutrophil granulocytes in cerebral ischemia – Evolution from killers to key players. *Neurochemistry International* **107**, 117–126 (2017).

32. Nardi, K. *et al.* Admission leukocytosis in acute cerebral ischemia: Influence on early outcome. *J. Stroke Cerebrovasc. Dis.* **21**, 819–824 (2012).

33. Maestrini, I. *et al.* Higher neutrophil counts before thrombolysis for cerebral ischemia predict worse outcomes. *Neurology* **85**, 1408–1416 (2015).

34. National Human Genome Research Institute. The Human Genome Project. Available at: https://www.genome.gov/human-genome-project.

35. Strachan, T. & Read, A. P. Genome projects and model organisms. in *Human molecular genetics* 208–238 (Garland Science, 2004).

36. Ozaki, K. *et al.* Functional SNPs in the lymphotoxin-α gene that are associated with susceptibility to myocardial infarction. *Nat. Genet.* **32**, 650–654 (2002).

37. Yamazaki, K. *et al.* Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum. Mol. Genet.* **14**, 3499–3506 (2005).

38. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

39. Fernandez-Cadenas, I. *et al.* Splicing mosaic of the myophosphorylase gene due to a silent mutation in McArdle disease. *Neurology* **61**, 1432–1434 (2003).

40. Anderson, C. A. *et al.* Data quality control in genetic case-control association studies. *Nat. Protoc.* **5**, 1564–1573 (2010).

41. Coleman, C., Quinn, E. M. & McManus, R. Quality Control Procedures for High-Throughput Genetic Association Studies. *Methods Mol. Biol.* **1326**, 203–215 (2015).

42. Bevan, S. *et al.* Genetic Heritability of Ischemic Stroke and the Contribution of Previously Reported Candidate Gene and Genomewide Associations. *Stroke* **43**, 3161–3167 (2012).

43. 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.* **15**, 174–184 (2016).

44. Traylor, M. *et al.* Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol* **11**, 951–962 (2012).

45. Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537 (2018).

46. Malik, R. *et al.* Genome-wide meta-analysis identifies 3 novel loci associated with stroke. *Ann. Neurol.* **84**, 934–939 (2018).

47. Holliday, E. G. *et al.* Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat. Genet.* **44**, 1147–1151 (2012).

48. Mola-Caminal, M. *et al.* PATJ Low Frequency Variants Are Associated With Worse Ischemic Stroke Functional Outcome. *Circ. Res.* **124**, 114–120 (2019).

49. Söderholm, M. *et al.* Genome-wide association meta-analysis of functional outcome after ischemic stroke. *Neurology* **92**, e1271–e1283 (2019).

50. Lindgren, A. & Maguire, J. Stroke Recovery Genetics. *Stroke* **47**, 2427–34 (2016).

51. Cao, L. *et al.* Impacts of COX-1 gene polymorphisms on vascular outcomes in patients with ischemic stroke and treated with aspirin. *Gene* **546**, 172–176 (2014).

52. Zhang, Z. *et al.* Chromosome 12p13 variants predict recurrence of ischaemic stroke in a Chinese population. *Eur. J. Neurol.* **21**, 1400–1405 (2014).

53. Weinstein, J. R. *et al.* Functional polymorphisms in toll-like receptor 4 are associated with worse outcome in acute ischemic stroke patients. *Neuroreport* 1 (2014). doi:10.1097/WNR.0000000000000140

54. Lv, W. *et al.* Variants of COL3A1 Are Associated with the Risk of Stroke Recurrence and Prognosis in the Chinese Population: a Prospective Study. *J. Mol. Neurosci.* **53**,

196–203 (2014).

55. Yeh, P.-S. *et al.* Prognosis of young ischemic stroke in Taiwan: impact of prothrombotic genetic polymorphisms. *Thromb. Haemost.* **92**, 583–589 (2004).

56. Maguire, J. *et al.* Impact of COX-2 rs5275 and rs20417 and GPIIIa rs5918 Polymorphisms on 90-Day Ischemic Stroke Functional Outcome: A Novel Finding. *J. Stroke Cerebrovasc. Dis.* **20**, 134–144 (2011).

57. Heywood, D. M., Carter, A. M., Catto, A. J., Bamford, J. M. & Grant, P. J. Polymorphisms of the factor VII gene and circulating FVII:C levels in relation to acute cerebrovascular disease and poststroke mortality. *Stroke* (1997). doi:10.1161/01.STR.28.4.816

58. Gromadzka, G., Barańska-Gieruszczak, M., Ciesielska, A., Sarzyńska-Dlugosz, I. & Członkowska, A. APOE Genotype and Serum Cholesterol in Predicting Risk for Early Death from Ischemic Stroke in Men and Women. *Cerebrovasc. Dis.* **20**, 291–298 (2005).

59. Gromadzka, G., Baranska-Gieruszczak, M., Sarzynska-Dlugosz, I., Ciesielska, A. & Czlonkowska, A. The APOE polymorphism and 1-year outcome in ischemic stroke: genotype-gender interaction. *Acta Neurol. Scand.* **116**, 392–8 (2007).

60. Sarzyńska-Długosz, I., Gromadzka, G., Barańska-Gieruszczak, M., Ciesielska, A. & Członkowska, A. APOE does not predict poor outcome 1 year after ischemic stroke. *Neurol. Res.* **29**, 64–69 (2007).

61. Cramer, S. C. & Procaccio, V. Correlation between genetic polymorphisms and stroke recovery: Analysis of the GAIN Americas and GAIN International Studies. *Eur. J. Neurol.* **19**, 718–724 (2012).

62. Carter, A., Catto, A., Bamford, J. & Grant, P. Association of the platelet glycoprotein IIb HPA-3 polymorphism with survival after acute ischemic stroke. *Stroke* (1999).

63. Åberg, N. D. *et al.* Genetic variation at the IGF1 locus shows association with post-stroke outcome and to circulating IGF1. *Eur. J. Endocrinol.* **169**, 759–765 (2013).

64. Hoy, A. *et al.* Myeloperoxidase polymorphisms in brain infarction. Association with infarct size and functional outcome. *Atherosclerosis* (2003). doi:10.1016/S0021-9150(02)00041-2

65. Lövkvist, H. *et al.* Variations in apolipoprotein D and sigma non-opioid intracellular receptor 1 genes with relation to risk, severity and outcome of ischemic stroke. *BMC Neurol.* **14**, 191 (2014).

66. Chang, W. H. *et al.* BDNF Polymorphism and Differential rTMS Effects on Motor Recovery of Stroke Patients. *Brain Stimul.* **7**, 553–558 (2014).

67. Di Lazzaro, V. *et al.* Val66Met BDNF Gene Polymorphism Influences Human Motor Cortex Plasticity in Acute Stroke. *Brain Stimul.* **8**, 92–96 (2015).

68. Mirowska-Guzel, D. *et al.* Impact of BDNF -196 G>A and BDNF -270 C>T Polymorphisms on Stroke Rehabilitation Outcome: Sex and Age Differences. *Top. Stroke Rehabil.* **21**, S33–S41 (2014).

69. Kim, J.-M. *et al.* Associations of BDNF Genotype and Promoter Methylation with Acute and Long-Term Stroke Outcomes in an East Asian Cohort. *PLoS One* **7**, e51280 (2012).

70. Qiu, L.-N. *et al.* Influence of CYP2C19 polymorphisms on platelet reactivity and clinical outcomes in ischemic stroke patients treated with clopidogrel. *Eur. J. Pharmacol.* **747**, 29–35 (2015).

71. Guo, J. *et al.* CRP gene polymorphism predicts post-stroke functional outcome in Han Chinese. *Acta Neurol. Scand.* **129**, 263–268 (2014).

72. Liepert, J., Heller, A., Behnisch, G. & Schoenfeld, A. Catechol- O -Methyltransferase Polymorphism Influences Outcome After Ischemic Stroke. *Neurorehabil. Neural Repair* **27**, 491–496 (2013).

73. Kohen, R. *et al.* Association of Serotonin Transporter Gene Polymorphisms With Poststroke Depression. *Arch. Gen. Psychiatry* **65**, 1296 (2008).

74. Carrera, C. *et al.* Validation of a clinical-genetics score to predict hemorrhagic transformations after rtPA. *Neurology* **93**, e851–e863 (2019).

75. Larsson, S. C., Burgess, S. & Michaëlsson, K. Smoking and stroke: A mendelian randomization study. *Ann. Neurol.* ana.25534 (2019). doi:10.1002/ana.25534

76. Larsson, S. C. *et al.* Type 2 diabetes, glucose, insulin, BMI, and ischemic stroke subtypes. *Neurology* **89**, 454–460 (2017).

77. Hindy, G. *et al.* Role of Blood Lipids in the Development of Ischemic Stroke and its Subtypes. *Stroke* **49**, 820–827 (2018).

78. Gallego-Fabrega, C. *et al.* TRAF3 Epigenetic Regulation Is Associated With Vascular Recurrence in Patients With Ischemic Stroke. *Stroke* **47**, (2016).

79. Soriano-Tárraga, C. *et al.* Ischemic stroke patients are biologically older than their chronological age. *Aging (Albany. NY).* **8**, 2655–2666 (2016).

80. Krupinski, J. *et al.* DNA Methylation in Stroke. Update of Latest Advances. *Comput. Struct. Biotechnol. J.* **16**, 1–5 (2018).

81. Regenhardt, R. W., Das, A. S., Lo, E. H. & Caplan, L. R. Advances in Understanding the Pathophysiology of Lacunar Stroke. *JAMA Neurol.* **75**, 1273 (2018).

82. Lübke, T., Lobel, P. & Sleat, D. E. Proteomics of the lysosome. *Biochim. Biophys. Acta* **1793**, 625–35 (2009).

83. Bernon, C. *et al.* Overexpression of Man2C1 leads to protein underglycosylation and upregulation of endoplasmic reticulum-associated degradation pathway. *Glycobiology* **21**, 363–375 (2011).

84. Riba-Llena, I. Biomarcadores en el ictus lacunar. in *Ictus lacunar* 91–106 (Marge Médica Books, 2012).

85. Kumar, A. D. *et al.* Leukocytosis in patients with neurologic deterioration after acute ischemic stroke is associated with poor outcomes. *J. Stroke Cerebrovasc. Dis.* **22**, e111-7 (2013).

86. Tian, C. *et al.* Association of lower leukocyte count before thrombolysis with early neurological improvement in acute ischemic stroke patients. *J. Clin. Neurosci.* **56**, 44–49 (2018).

87. Yi, X., Wang, C., Zhou, Q. & Lin, J. Interaction among COX-2, P2Y1 and GPIIIa gene variants is associated with aspirin resistance and early neurological deterioration in Chinese stroke patients. *BMC Neurol.* **17**, 1–9 (2017).

88. Yi, X., Ming, B., Wang, C., Chen, H. & Ma, C. Variants in COX-2, PTGIS, and TBXAS1 Are Associated with Carotid Artery or Intracranial Arterial Stenosis and Neurologic Deterioration in Ischemic Stroke Patients. *J. Stroke Cerebrovasc. Dis.* **26**, 1128–1135 (2017).

89. Le Menn, G. & Neels, J. G. Regulation of Immune Cell Function by PPARs and the Connection with Metabolic and Neurodegenerative Diseases. *Int. J. Mol. Sci.* **19**, (2018).

90. Maffei, R. *et al.* The monocytic population in chronic lymphocytic leukemia shows altered composition and deregulation of genes involved in phagocytosis and inflammation. *Haematologica* **98**, 1115–1123 (2013).