

REVISIÓN

Aproximación al conocimiento de las bases genéticas del ictus. Consortio español de genética del ictus

E. Giralt-Steinhauer^{a,*}, J. Jiménez-Conde^a, C. Soriano Tárrega^a, M. Mola^a,
A. Rodríguez-Campello^a, E. Cuadrado-Godía^a, A. Ois^a, I. Fernández-Cádenas^b,
C. Carrera^b, J. Montaner^b, R.M. Díaz Navarro^c, C. Vives-Bauzá^c y J. Roquer^a

^a Grupo de Investigación Neurovascular, Departamento de Neurología, IMIM-Hospital del Mar, Universitat Autònoma de Barcelona, Barcelona, España

^b Laboratorio de Investigación Neurovascular, Institut de Recerca, Hospital Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, España

^c Departamento de Neurología, Hospital Universitari Son Espases, Mallorca, Baleares, España

Recibido el 28 de diciembre de 2012; aceptado el 21 de abril de 2013

Accesible en línea el 5 de julio de 2013

PALABRAS CLAVE

Ictus;
Genética;
Asociación;
Factores de riesgo
genético;
Predicción;
Consortio

KEYWORDS

Stroke;
Genetics;
Association;
Genetic risk factors;
Prediction;
Consortium

Resumen El presente artículo revisa la evolución de los estudios en genética del ictus desde la aproximación por gen candidato hasta los recientes estudios de *genome wide association*. Se destaca la complejidad de esta afección por sus muy variados mecanismos etiopatogénicos, las dificultades que comporta el estudio de su componente genético y las soluciones que se han aportado. Se subraya en especial el valor de las colaboraciones entre distintos centros, ya sea de manera puntual o sobre todo a través de la creación de consorcios estables. Esta estrategia actualmente se hace imprescindible a la hora de realizar estudios de alta calidad científica que permitan seguir avanzando en el conocimiento de las bases genéticas del ictus tanto en etiología, como en tratamiento y prevención.

© 2012 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Exploring the genetic basis of stroke. Spanish stroke genetics consortium

Abstract This article provides an overview of stroke genetics studies ranging from the candidate gene approach to more recent studies by the genome wide association. It highlights the complexity of stroke owing to its different aetiopathogenic mechanisms, the difficulties in studying its genetic component, and the solutions provided to date. The study emphasises the importance of cooperation between the different centres, whether this takes places occasionally or through the creation of lasting consortiums. This strategy is currently essential to the completion of high-quality scientific studies that allow researchers to gain a better knowledge of the genetic component of stroke as it relates to aetiology, treatment, and prevention.

© 2012 Sociedad Española de Neurología. Published by Elsevier España, S.L.U. All rights reserved.

* Autor para correspondencia.

Correo electrónico: 96714@parcdesalutmar.cat (E. Giralt-Steinhauer).

Introducción

El ictus representa hoy en día en España la segunda causa de muerte en la población general y la primera en mujeres (<http://www.ine.es/prensa/np703.pdf>). Además, es la primera causa de discapacidad física en adultos y la segunda de demencia¹.

Sin embargo, y a pesar de las importantes mejoras en prevención, diagnóstico y tratamiento del ictus que se han realizado en las últimas 2 décadas, la Organización Mundial de la Salud prevé un incremento de su incidencia del 27% hasta el 2025². Se han descrito múltiples factores de riesgo asociados a esta afección, aunque a la hora de entender sus causas parece necesario contar con la participación tanto de factores externos como de la predisposición de cada individuo³. En este sentido, se ha demostrado un componente hereditario de la enfermedad y la carga genética desempeñaría un papel muy significativo al respecto⁴. El desarrollo tecnológico ha favorecido en los últimos años llevar a cabo estudios de genotipificación masiva, lo que ha permitido realizar estudios genéticos no sesgados y el descubrimiento de nuevos genes y vías metabólicas asociadas al ictus.

Los objetivos de este trabajo son revisar el estado actual y las perspectivas futuras de la genética en el campo de la enfermedad cerebrovascular.

Desarrollo

Consideraciones genéticas del ictus

A diferencia de las enfermedades monogénicas, las enfermedades complejas se caracterizan por presentar una gran diversidad etiopatológica. Se sabe que la historia familiar es uno de los factores de riesgo más importantes en muchas de las enfermedades complejas, como ocurre en enfermedades cardiovasculares, cáncer o enfermedades autoinmunes⁵. Esto indica que, a pesar de no tratarse de enfermedades mono/oligogénicas, existen variaciones genéticas heredadas. Se cree que la combinación de variantes alélicas con efecto de riesgo pequeño, pero aditivo o multiplicativo y su interacción con el ambiente, pueden desempeñar un papel crucial en el desarrollo de una afección compleja^{5,6}.

En el caso del ictus isquémico, se ha descrito una heredabilidad de entre el 30 y el 50%⁴. Debido a sus múltiples causas, a los numerosos factores de riesgo implicados y a las diferentes formas de manifestación, el ictus podría considerarse como la manifestación final de múltiples enfermedades complejas. Aparte de la clasificación más simple de los ictus entre isquémicos y hemorrágicos, podemos diferenciar varios subtipos etiopatogénicos completamente diferentes. Varios estudios publicados recientemente han demostrado la existencia de heterogeneidad genética, en cuanto a la heredabilidad entre los ictus lacunares, cardioembólicos y ateroscleróticos, siendo estos últimos los más fuertemente asociados a historia familiar^{4,7,8}.

Lo mismo ocurre en las hemorragias intracerebrales (HIC), que representan el 15% de todos los ictus. Las vías patogénicas implicadas en la hemorragia lobar, muy asociada a la angiopatía amiloide cerebral (AAC), pueden diferir

de forma sustancial con las implicadas en las hemorragias hipertensivas^{9,10}.

Así pues, el planteamiento de los estudios genéticos en el ictus debe contemplar estas particularidades en su diseño y contar con las dificultades que plantea la investigación de una enfermedad heterogénea.

Estudios genéticos de asociación en enfermedades complejas. Diferentes técnicas y abordajes

La investigación sobre las causas genéticas del ictus de forma global, prescindiendo de aquellas que siguen un patrón de herencia mendeliana clásica, se fundamenta sobre todo en los estudios de asociación caso-control. Estos se basan en estudiar la diferente frecuencia de variantes genéticas en enfermos respecto a no enfermos. El estudio de variantes comunes o polimorfismos de nucleótido simple (SNP) se basa principalmente en diseños de «*candidate-gene*» y de *genome wide association*, mientras que el reciente desarrollo de técnicas de *secuenciación completa del exoma* permite estudiar variantes infrecuentes. Otros abordajes, como el análisis de la *variación de número de copias* o el estudio del *epigenoma*, son tratados también más adelante.

Estudios mediante «*candidate-gene*»

Durante varias décadas, y hasta hace pocos años, el abordaje más extendido fue el de «*candidate-gene*» o estudios de gen candidato. Esta aproximación consiste en seleccionar aquellos genes que podrían asociarse a una enfermedad basándose en la fisiopatología de esta y en las vías que se supone están implicadas. Se identifican los SNP previamente descritos en la zona de estos genes candidatos a ser analizados y se escogen los más representativos (conocidos como *TagSNP*), generalmente por hallarse en desequilibrio de ligamiento¹¹. Estos *TagSNP* candidatos son genotipificados en una población y, mediante análisis estadísticos, se determina si existe alguna asociación entre las variantes genéticas y el espectro clínico.

Estos estudios presentan varios problemas. Por un lado, el hecho de que la selección de estos genes se sustente en hipótesis generadas a priori puede implicar una inclinación del investigador a buscar diferencias espurias basadas en sus ideas preconcebidas. Por otro lado, deja de analizar posibles genes potencialmente implicados en la enfermedad.

En relación con los estudios de gen candidato, se han asociado múltiples genes con el riesgo de presentar un ictus isquémico. Sin embargo, un estudio reciente demuestra que tras aplicar una corrección estadística adecuada ajustando por el número de SNP analizados, ninguna de estas posibles asociaciones alcanza la significación estadística⁴. Por otro lado, ninguno de los resultados obtenidos mediante este abordaje ha sido replicado posteriormente de forma clara, por lo que esta aproximación ha demostrado ser poco consistente y fiable en cuanto al ictus isquémico^{4,12}.

Aun así, a través de esta aproximación se ha demostrado que mutaciones en los genes COL4A1 y COL4A2 se asocian a HIC. Los heterotrímeros compuestos por colágeno tipo IV alfa 1 (COL4A1) y alfa 2 (COL4A2) son uno de los principales componentes de casi todas las membranas basales del organismo y se han asociado a gran número de enfermedades

no solo cerebrovasculares, sino también oculares, renales y musculares¹³. Inicialmente, mediante estudios de herencia familiar, se encontró que las mutaciones en COL4A1 causaban porencefalia tipo I, demostrando posteriormente causar HIC subcorticales, tanto esporádicas como recurrentes, y también ictus lacunares y leucaraiosis¹⁴⁻¹⁶.

En relación también con la HIC, se ha demostrado recientemente que los pacientes portadores de los haplotipos $\epsilon 2$ y $\epsilon 4$ del gen APOE tienen un riesgo mayor de HIC lobar (*odds ratio* [OR] 1,82 [IC del 95%, 1,50 a 2,23]; $p = 6,6 \times 10^{-10}$) y (OR 2,20 [IC del 95%, 1,85 a 2,63]; $p = 2,4 \times 10^{-11}$, respectivamente), probablemente por el efecto de estas variantes en el riesgo de desarrollar una AAC¹⁷. Posteriormente, el mismo grupo demuestra también que la variante APOE $\epsilon 2$ se asocia a un mayor volumen de la HIC y, como consecuencia del mayor volumen, a un peor pronóstico funcional y mayor mortalidad¹⁸.

Estudios de asociación en genoma completo o *genome wide association studies*

El desarrollo de los llamados estudios de asociación de genoma completo (GWAS) que permiten determinar la frecuencia alélica de cientos de miles de SNP en casos *versus* controles, ha sido posible por 2 grandes motivos. En primer lugar, por la creación del International HapMap Project, que determinó características, equilibrio de ligamiento y frecuencias alélicas de los SNP hallados en el genoma de 270 individuos de 4 orígenes ancestrales diferentes, con el objetivo de crear una base *genome-wide* pública, de acceso abierto a todos los investigadores^{5,19}. Y, en segundo lugar, gracias a los importantes avances técnicos y al desarrollo de plataformas que permiten genotipificar entre 50.000 y 2,7 millones de SNP por individuo a un precio cada vez menor. Los estudios mediante GWAS representan una forma mucho más exhaustiva de estudiar el genoma que los diseños de gen candidato²⁰. Evitan los sesgos de hipótesis preconcebidas y además permiten identificar nuevas asociaciones aun cuando el gen todavía no ha sido implicado en la enfermedad, a diferencia de los estudios anteriores. Pero esta aproximación conlleva otro tipo de limitaciones. Las 2 principales son los grandes tamaños muestrales necesarios en estos estudios y la dificultad en la interpretación de los posibles resultados positivos²¹. Además, implican la necesidad de realizar estrictos controles de calidad, con el fin de detectar errores de genotipificación y valorar la posible estratificación de poblaciones²². Esta última se basa en que hay diferencias en las frecuencias alélicas de un SNP dependiendo de los orígenes étnicos ancestrales de cada individuo (aunque fenotípicamente sean de la misma etnia). Así, se pueden generar resultados falsos si las poblaciones de casos y controles no son comparables en su origen genético y, por tanto, es necesario realizar los ajustes pertinentes. Actualmente, la mejor forma de evitar estas inconsistencias es mediante la replicación de los resultados obtenidos en muestras independientes de gran tamaño²³. Debido a la gran cantidad de análisis a realizar (al menos uno por cada SNP genotipificado), se deben aplicar penalizaciones estadísticas por comparaciones múltiples, que obligan a establecer la significación en valores de p de hasta 10^{-8} (en el caso de analizar un millón de SNP)⁶. Así pues, para alcanzar la potencia estadística necesaria, estos estudios

requieren aumentar significativamente el número de individuos analizados.

Los primeros resultados relacionados con enfermedades complejas aparecieron en 2005, descubriendo la implicación del complemento H en la degeneración macular relacionada con la edad²⁴. Se han conseguido descubrimientos interesantes en relación con la diabetes mellitus de tipo 1 (PPARG²⁵, KCNJ11²⁶ y TCF7L2²⁷), la artritis reumatoide y la diabetes mellitus de tipo 2 (PTPN22^{28,29}) y la enfermedad de Crohn (NOD2³⁰), entre muchos otros, descubriendo así un gran número de loci implicados en la predisposición o riesgo de presentarlas.

En contraposición a los resultados desalentadores obtenidos con otras aproximaciones, varias asociaciones a ictus isquémicos han sido replicadas consistentemente. Es el caso de algunas asociaciones descubiertas mediante GWAS para otras enfermedades asociadas a ictus isquémicos, que han demostrado ser también factores de riesgo para ictus³¹. Así, 2 variantes, rs1906591 (OR 1,48 [IC del 95%, 1,28 a 1,71]; $p = 4,7 \times 10^{-6}$) y rs7193343 (OR 1,22 [IC del 95%, 1,10 a 1,35]; $p = 0,00021$), en los cromosomas 4q25 y 16q22, en regiones cercanas a los genes PITX2 y ZFX3, respectivamente, inicialmente identificadas por contribuir al riesgo de presentar una fibrilación auricular, han demostrado también asociarse al riesgo de presentar un ictus de etiología cardioembólica^{32,33}.

Se identificó una región en el cromosoma 9p21, próxima a los genes CDKN2A/B, asociada a infarto de miocardio y enfermedad coronaria. Examinando posteriormente estos loci mediante un estudio de gen candidato, se vio la asociación del SNP rs1537378 (OR 1,21; [IC del 95%, 1,07-1,37]; $p = 0,002$) con ictus de etiología aterotrombótica³⁴. También, en relación con los ictus isquémicos de etiología aterotrombótica, una nueva asociación mediante GWAS ha sido identificada: el polimorfismo rs11984041 (OR 1,42 [IC del 95%, 1,28 a 1,57]; $p = 1,87 \times 10^{-11}$) en la región del gen HDAC9 en el cromosoma 7p21.1, aunque se desconoce el mecanismo patogénico³⁵ (fig. 1).

Como asociaciones nuevas, un grupo japonés ha identificado una relación significativa entre el polimorfismo rs2230500 (OR 1,66 [IC del 95%, 1,33 a 2,09]; $p = 9,84 \times 10^{-6}$) del gen PRKCH en el cromosoma 14 con ictus de etiología lacunar, con resultados en replicación dispares en otras etnias^{4,36}.

Secuenciación del exoma

Una de las debilidades de los GWAS es que únicamente analizan variantes comunes (frecuencia poblacional > 1%) y no contemplan las variantes infrecuentes (*rare variants*). Es posible que existan múltiples variantes infrecuentes que contribuyan al riesgo genético de las enfermedades complejas. Para intentar identificarlas, se están desarrollando nuevas técnicas de secuenciación masiva del exoma o *exome sequencing*. Estas nuevas técnicas analizan la presencia de todo tipo de polimorfismos o mutaciones que se encuentren en las regiones codificantes del genoma, por lo tanto, son capaces de detectar estas variantes infrecuentes que los estudios de GWAS no identifican. Se trata de una aproximación prometedora, aunque su metodología está en desarrollo y todavía están por consensuar las formas más adecuadas de manejar

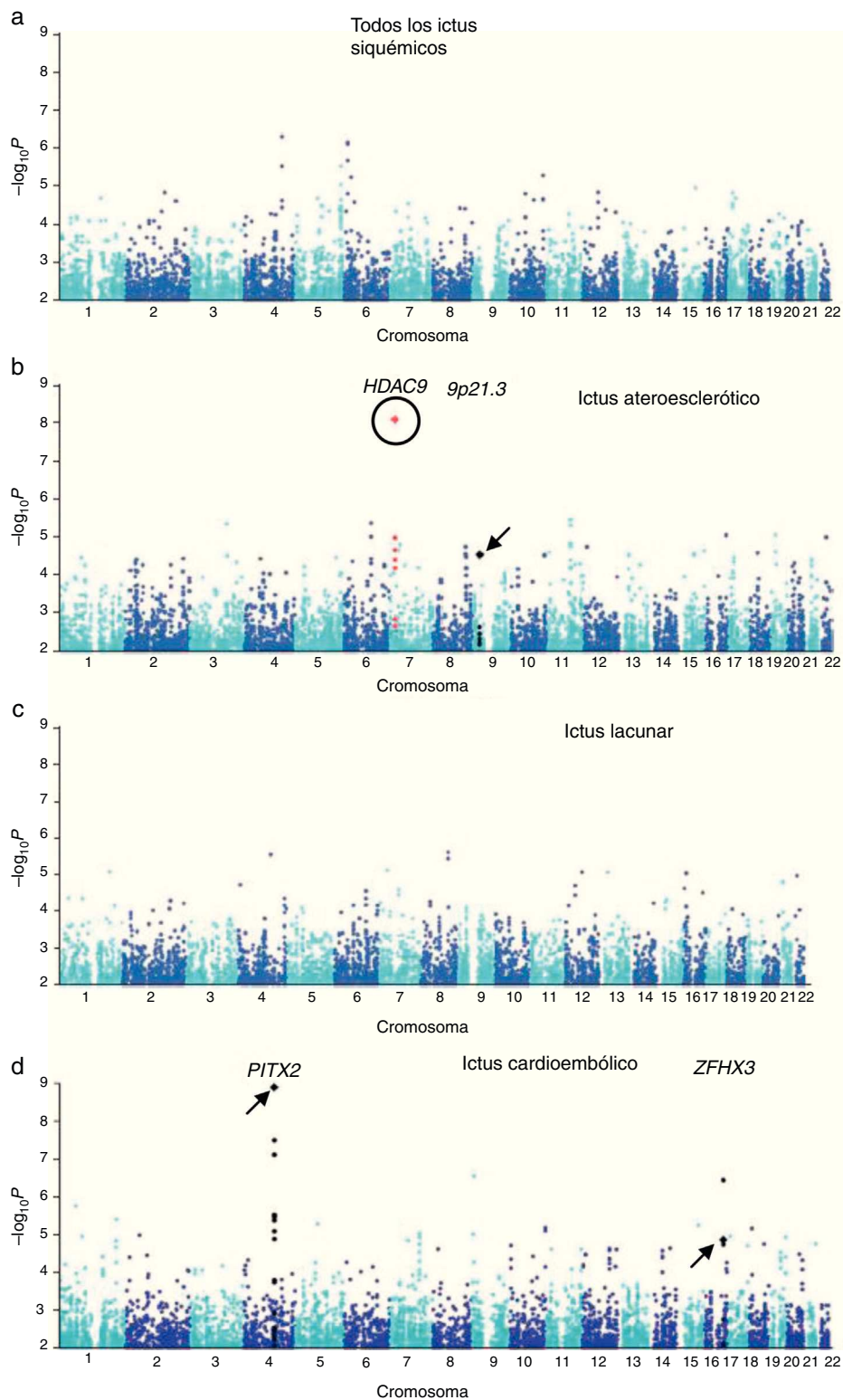


Figura 1 Resultados del GWA de las muestras de Reino Unido y Alemania. *a)* Resultado de todos los ictus isquémicos. *b)* Ictus ateroscleróticos. *c)* Ictus lacunares. *d)* Ictus cardioembólicos. Rodeado por un círculo aparece el nuevo locus *HDAC9*. Señalados con una flecha aparecen los loci previamente descritos en la literatura.

Publicado con el permiso de Macmillan Publishers Ltd: *Nature Genetics*³⁴, copyright 2012.

Adaptada con el permiso de Macmillan Publishers Ltd: *Nature Genetics*³⁴, copyright 2012.

Tabla 1 Principales artículos publicados en revista de gran impacto en el marco del International Stroke Genetics Consortium

Autores, año de publicación y revista	Tipo de ictus y subtipo	N.º casos/ N.º controles	Cromosoma implicado	Conclusión
Gschwendter A. <i>Annals of Neurology</i> , 2009 ³⁴	Ictus isquémico. TOAST: aterosclerótico	1.090/1244 sur de Alemania 758/872 Reino Unido Replicación: 2.528/2.189	Cromosoma 9 p región 21.3: rs1537378	Este locus implicado también en la CI se encuentra relacionado con ictus de etiología aterotrombótica
Biffi A. <i>Annals of Neurology</i> , 2010 ¹⁷	Hemorragias lobares e hipertensivas	Lobares: 931/3744 Profundas: 1.085/3.657	2 variantes en APOE cromosoma 19: rs7412 y rs429358	Fuerte asociación entre variantes APOE e HIC lobar, sobre todo en HIC relacionadas con AAC
Lemmens R. <i>Stroke</i> , 2010 ³²	Ictus isquémico. TOAST: cardioembólico	4.199/3.750	Variante del cromosoma 4q25: rs1906591	Replican la asociación de la variante rs1906591 del cromosoma 4q25 con la FA y el ictus de etiología CE. No se asocia a otros subtipos de ictus
Biffi A. <i>Lancet Neurology</i> , 2011 ¹⁸	HIC lobares	Casos: 865 individuos. Replicación 1: 946 y replicación 2: 214	2 variantes en APOE, cromosoma 19: rs7412 y rs429358	La presencia de alelos APOE ε2 se asocia a un mayor volumen de HIC, probablemente por el papel del ε2 en la severidad de la vasculopatía por AAC
ISGC & WTCCC2. <i>Nature Genetics</i> , 2012 ³⁵	Ictus isquémico. TOAST: aterosclerótico	3.548/5.972 Replicación: 5.859/6.281	Cromosoma 7p21.1 (rs11984041). Gen HDAC9	Replican el locus ya conocido 9p21. Identifican nueva asociación en el cromosoma 7p21.1 El mecanismo no se conoce

AAC: angiopatía amiloide cerebral; CE: cardioembólico; CI: cardiopatía isquémica; FA: fibrilación auricular; HIC: hemorragia intracerebral; TOAST: Trial of Org 10172 in Acute Stroke Treatment.

estadísticamente las cantidades ingentes de datos que generan estos análisis³⁷.

Estudios de variación en el número de copias

La variación en el número de copias o *copy number variation* (CNV) es una variación estructural del genoma, que puede producirse en forma de (micro) deleciones, duplicaciones o inversiones, generando un número anómalo de copias de un fragmento del ADN. Esta variación puede influir tanto en la expresión como en la conformación de las proteínas para las que codifica esa región³⁸⁻⁴⁰. Poco se ha publicado hasta la fecha con respecto al ictus. En un estudio con pacientes con ictus isquémicos no se detectó ninguna CNV asociada, aunque el tamaño muestral era reducido y no se tuvieron en cuenta los distintos subtipos de ictus⁴¹. Sin embargo, hay proyectos de genotipificado masivo en marcha que incluyen en sus plataformas el análisis de múltiples CNV.

Epigenética

Otro campo emergente es el de la epigenética, modalidad que estudia los cambios en la regulación génica heredables que no afectan a la secuencia del ADN. A pesar de que el genoma de cada célula de nuestro organismo es igual, la transcripción génica en cada tejido viene regulada por los distintos mecanismos epigenéticos. Entre los principales, destacan la metilación del ADN, las modificaciones en las histonas y el ARN no codificante⁴². De estos, la metilación del ADN es el más estudiado y

conocido hasta el momento. Los hallazgos más destacables se han obtenido en el estudio de la epigenética de tejidos tumorales. Así, se ha visto que en el caso del glioblastoma multiforme la metilación del gen MGMT inactiva al enzima O (6)-metilguanina-ADN metiltransferasa (MGMT), lo que se ha asociado a una mejor respuesta al tratamiento quimioterapéutico con temozolamida. Esta determinación es aplicada ya en la práctica clínica⁴³. Actualmente, los estudios epigenéticos se están ampliando a otras enfermedades, incluido el ictus. Se ha descrito una asociación entre el ictus y una hipometilación global del ADN, aunque está en fase de replicación⁴⁴. En el marco del Consorcio Español de Genética del Ictus, están en marcha varios estudios en este campo de los que todavía no se dispone de resultados.

Escalas de riesgo y algoritmos genéticos (*genetic networks*)

De cara a integrar la información y la capacidad predictiva de cada uno de los SNP que se han asociado a un fenotipo, están tomando relevancia el diseño de *scores* y *genetic networks*. Los métodos más usados para crear estas escalas o *scores* genéticos son: mediante la suma de alelos de riesgo (1 punto por cada alelo presente), o bien mediante un riesgo ponderado (la puntuación para cada alelo depende de la razón de oportunidades u OR reportada para ese alelo). En un estudio muy reciente, se ha demostrado una asociación entre un *score* conformado por alelos de riesgo para hipertensión arterial y el riesgo de presentar una HIC profunda⁴⁵.

Otra forma de englobar el conocimiento es mediante la creación de redes o algoritmos⁴⁶. Un ejemplo, es la red bayesiana o *Bayesian network*, una herramienta estadística que representa todo un conjunto de variables asociadas teniendo en cuenta las relaciones de dependencia condicionales que se establecen entre ellas. Lo que permite este modelo es estimar la probabilidad de que sucedan las variables no conocidas en función de las que sí conocemos. Las principales ventajas de estas redes bayesianas son, por un lado, que permiten las inferencias bidireccionales (es decir, de los efectos a las causas y de las causas a los efectos) y, por el otro, la rápida actualización de conocimientos⁴⁷.

Genomic convergence

En el estudio genético de las enfermedades complejas está tomando relevancia el concepto de *genomic convergence*. Este se fundamenta en sumar y concentrar la información y los datos que aportan cada uno de los diferentes abordajes para alcanzar así una mayor potencia de descubrimiento y una mayor solidez en los resultados⁴⁸. Como descubrimiento reciente, varios polimorfismos del gen TTC7B han demostrado conferir susceptibilidad en el ictus isquémico⁴⁹.

La creación de consorcios. Consorcio Español de Genética del ictus (GeneStroke)

A partir de la formación de grandes consorcios internacionales, como el Wellcome Trust Case Control Consortium del Reino Unido, se crea en el año 2007, el International Stroke Genetics Consortium (ISGC) (<http://www.strokegenetics.org>). Se han identificado y replicado diversas asociaciones en el marco de este consorcio (ISGC) (tabla 1) y se han generado varios de los proyectos, actualmente en marcha, más ambiciosos en este campo a nivel mundial.

El Consorcio Español de Genética del Ictus (GeneStroke) (<http://www.genestroke.com>), se crea en el año 2009 con el objetivo prioritario de estudiar las bases genéticas del ictus y promover la asociación entre grupos con intereses comunes en este campo. GeneStroke ha puesto en marcha diversos proyectos relacionados con la etiología, la evolución y el pronóstico del ictus, en los que colaboran distintos centros. Los detalles de los proyectos y sus participantes se pueden consultar en su página web (www.genestroke.com).

Desde el Consorcio Español de Genética del Ictus se fomenta el modelo de grandes estudios colaborativos para afrontar los retos que plantea la investigación en el campo de la genética de las enfermedades cerebrovasculares.

Conclusiones

El ictus es, en nuestro entorno, una de las enfermedades más prevalentes y con mayor impacto en morbimortalidad. Presenta una alta heredabilidad, por lo que el conocimiento de sus bases genéticas podrá contribuir a identificar a los individuos de riesgo y a diseñar estrategias de prevención y terapias más eficientes. El estudio de este sustrato genético debería ser una prioridad, así como la colaboración entre diferentes grupos y profesionales con el objetivo de realizar

proyectos de alta calidad científica que puedan traducirse en un futuro beneficio para la sociedad.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Agradecimientos

Agradecer, en primer lugar, a todos los miembros del Consorcio Español de Genética (GeneStroke), que forman parte del Consorcio de forma estable, o bien por su contribución puntual en algún proyecto. Un agradecimiento especial a los miembros del International Stroke Genetics Consortium.

Bibliografía

1. Alvarez Sabin J. In-hospital mortality in stroke patients. *Rev Esp Cardiol*. 2008;61:1007–9.
2. Truelsen T, Piechowski-Jozwiak B, Bonita R, Mathers C, Bogousslavsky J, Boysen G. Stroke incidence and prevalence in Europe: A review of available data. *Eur J Neurol*. 2006;13:581–98.
3. Grau AJ, Weimar C, Buggle F, Heinrich A, Goertler M, Neumaier S, et al. Risk factors, outcome, and treatment in subtypes of ischemic stroke: The German stroke data bank. *Stroke*. 2001;32:2559–66.
4. 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;43:3161–7.
5. International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005;437:1299–320.
6. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996;273:1516–7.
7. Jerrard-Dunne P, Cloud G, Hassan A, Markus HS. Evaluating the genetic component of ischemic stroke subtypes: A family history study. *Stroke*. 2003;34:1364–9.
8. 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;11:951–62.
9. Caplan LR. Intracerebral haemorrhage. *Lancet*. 1992;339:656–8.
10. Flaherty ML, Woo D, Broderick J. The incidence of deep and lobar intracerebral hemorrhage in whites, blacks, and hispanics. *Neurology*. 2006;66:956–7, author reply 956–957.
11. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: Practical considerations. *Nat Rev Genet*. 2002;3:391–7.
12. Dichgans M, Markus HS. Genetic association studies in stroke: Methodological issues and proposed standard criteria. *Stroke*. 2005;36:2027–31.
13. Kuo DS, Labelle-Dumais C, Gould DB. COL4A1 and COL4A2 mutations and disease: Insights into pathogenic mechanisms and potential therapeutic targets. *Hum Mol Genet*. 2012;21(R1):R97–110.
14. Breedveld G, de Coe IF, Lequin MH, Arts WF, Heutink P, Gould DB, et al. Novel mutations in three families confirm a major role of COL4A1 in hereditary porencephaly. *J Med Genet*. 2006;43:490–5.

15. Gould DB, Phalan FC, van Mil SE, Sundberg JP, Vahedi K, Massin P, et al. Role of COL4A1 in small-vessel disease and hemorrhagic stroke. *N Engl J Med*. 2006;354:1489–96.
16. Weng YC, Sonni A, Labelle-Dumais C, de Leau M, Kauffman WB, Jeanne M, et al. COL4A1 mutations in patients with sporadic late-onset intracerebral hemorrhage. *Ann Neurol*. 2012;71:470–7.
17. Biffi A, Sonni A, Anderson CD, Kissela B, Jagiella JM, Schmidt H, et al. Variants at APOE influence risk of deep and lobar intracerebral hemorrhage. *Ann Neurol*. 2010;68:934–43.
18. Biffi A, Anderson CD, Jagiella JM, Schmidt H, Kissela B, Hansen BM, et al. APOE genotype and extent of bleeding and outcome in lobar intracerebral haemorrhage: A genetic association study. *Lancet Neurol*. 2011;10:702–9.
19. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al., International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007;449:851–61.
20. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet*. 2008;9:255–66.
21. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9:356–69.
22. Marchini J, Cardon LR, Phillips MS, Donnelly P. The effects of human population structure on large genetic association studies. *Nat Genet*. 2004;36:512–7.
23. Neale BM, Purcell S. The positives, protocols, and perils of genome-wide association. *Am J Med Genet B Neuropsychiatr Genet*. 2008;147B:1288–94.
24. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005;308:385–9.
25. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, et al. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet*. 2000;26:76–80.
26. Florez JC, Hirschhorn J, Altshuler D. The inherited basis of diabetes mellitus: Implications for the genetic analysis of complex traits. *Annu Rev Genomics Hum Genet*. 2003;4:257–91.
27. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. 2007;316:1336–41.
28. Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am J Hum Genet*. 2004;75:330–7.
29. Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nat Genet*. 2004;36:337–8.
30. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature*. 2001;411:603–6.
31. Markus HS. Stroke genetics. *Hum Mol Genet*. 2011;20(R2):R124–31.
32. Lemmens R, Buyschaert I, Geelen V, Fernandez-Cadenas I, Montaner J, Schmidt H, et al. The association of the 4q25 susceptibility variant for atrial fibrillation with stroke is limited to stroke of cardioembolic etiology. *Stroke*. 2010;41:1850–7.
33. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, et al. A sequence variant in ZFX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet*. 2009;41:876–8.
34. Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, et al. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. *Ann Neurol*. 2009;65:531–9.
35. Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, Burgess AI, Pirinen M, et al., International Stroke Genetics Consortium (ISGC), Wellcome Trust Case Control Consortium 2 (WTCCC2). Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet*. 2012;44:328–33.
36. Kubo M, Hata J, Ninomiya T, Matsuda K, Yonemoto K, Nakano T, et al. A nonsynonymous SNP in PRKCH (protein kinase C eta) increases the risk of cerebral infarction. *Nat Genet*. 2007;39:212–7.
37. Cole JW, Stine OC, Liu X, Pratap A, Cheng Y, Tallon LJ, et al. Rare variants in ischemic stroke: An exome pilot study. *PLoS One*. 2012;7:e35591.
38. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. *Nat Genet*. 2004;36:949–51.
39. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. *Science*. 2004;305:525–8.
40. Pollex RL, Hegele RA. Copy number variation in the human genome and its implications for cardiovascular disease. *Circulation*. 2007;115:3130–8.
41. Matarin M, Simon-Sanchez J, Fung HC, Scholz S, Gibbs JR, Hernandez DG, et al. Structural genomic variation in ischemic stroke. *Neurogenetics*. 2008;9:101–8.
42. Qureshi IA, Mehler MF. Emerging role of epigenetics in stroke: Part 1: DNA methylation and chromatin modifications. *Arch Neurol*. 2010;67:1316–22.
43. Skiriute D, Vaitkiene P, Saferis V, Asmoniene V, Skauminas K, Deltuva VP, et al. MGMT, GATA6, CD81, DR4, and CASP8 gene promoter methylation in glioblastoma. *BMC Cancer*. 2012;12:218.
44. Baccarelli A, Wright R, Bollati V, Litonjua A, Zanobetti A, Tarantini L, et al. Ischemic heart disease and stroke in relation to blood DNA methylation. *Epidemiology*. 2010;21:819–28.
45. Falcone GJ, Biffi A, Devan WJ, Jagiella JM, Schmidt H, Kissela B, et al. Burden of risk alleles for hypertension increases risk of intracerebral hemorrhage. *Stroke*. 2012.
46. Park YK, Bang OS, Cha MH, Kim J, Cole JW, Lee D, et al. SigCS base: An integrated genetic information resource for human cerebral stroke. *BMC Syst Biol*. 2011;14(5 Suppl 2.). S10.0509-5-S2-S10. Epub 2011 Dec 14.
47. Ramoni RB, Himes BE, Sale MM, Furie KL, Ramoni MF. Predictive genomics of cardioembolic stroke. *Stroke*. 2009;40(3 Suppl):S67–70.
48. Bell CG, Finer S, Lindgren CM, Wilson GA, Rakyar VK, Teschendorff AE, et al. Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS One*. 2010;5:e14040.
49. Krug T, Gabriel JP, Taipa R, Fonseca BV, Domingues-Montanari S, Fernandez-Cadenas I, et al. TTC7B emerges as a novel risk factor for ischemic stroke through the convergence of several genome-wide approaches. *J Cereb Blood Flow Metab*. 2012;32:1061–72.