THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

PATJ Low Frequency Variants Are Associated with Worse Ischemic Stroke Functional Outcome: A Genome-Wide Meta-Analysis

Digital Object Identifier (DOI):
10.1161/CIRCRESAHA.118.313533

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Circulation Research

Publisher Rights Statement:
This is the author's peer-reviewed manuscript as accepted for publication.

OPEN ACCESS

PATJ Low Frequency Variants Are Associated with Worse Ischemic Stroke Functional Outcome: A Genome-Wide Meta-Analysis

Genetic contribution to functional Outcome and Disability after Stroke (GODS) Project, NINDS Stroke Genetics Network (SiGN), International Stroke Genetics Consortium (ISGC), and GeneStroke Consortium

Marina Mola-Caminal[1,2]; Caty Carrera[3]; Carolina Soriano-Tárraga[1]; Eva Giralt-Steinhauer[1]; Rosa M. Díaz-Navarro[4]; Sílvia Tur[4]; Carmen Jiménez[4]; Aina Medina-Dols[5]; Natàlia Cullell[6]; Nuria P. Torres-Aguila[6]; Elena Muiño[6]; Ana Rodríguez-Campello[1]; Angel Ois[1]; Elisa Cuadrado-Godia[1]; Rosa M. Vivanco-Hidalgo[1]; Mar Hernandez-Guillamon[3]; Montse Solé[3]; Pilar Delgado[3]; Alejandro Bustamante[3]; Teresa García-Berrocoso[3]; Maite Mendióroz[3]; Mar Castellanos[7]; Joaquín Serena[8]; Joan Martí-Fàbregas[9]; Tomás Segura[10]; Gemma Serrano-Heras[11]; Victor Obach[12]; Marc Ribó[13]; Carlos A. Molina[13]; José Alvarez-Sabín[13]; Ernest Palomeras[14]; Mar Freijo[15]; Maria A. Font[16]; Jonathan Rosand[17,18,19]; Natalia S. Rost[17]; Cristina Gallego-Fabrega[18,19]; Jin-Moo Lee[20]; Laura Heitsch[20,21]; Laura Ibanez[20]; Carlos Cruchaga[20]; Chia-Ling Phuah[20]; Robin Lemmens[22,23,24]; Vincent Thijs[25,26]; Arne Lindgren[27,28]; Jane Maguire[29,30,31]; Kristiina Rannikmae[32,33]; Catherine L. Sudlow[32,33]; Christina Jern[34]; Tara M. Stanne[34]; Erik Lorentzen[35]; Lucía Muñoz-Narbona,[36]; Antonio Dávalos[36]; Elena López-Cancio[37]; Bradford B. Worrall[38]; Daniel Woo[39]; Steven J. Kittner[40]; Braxton D. Mitchell[41,42]; Joan Montaner[3,43,44]; Jaume Roquer[1]; Jurek Krupinski[45]; Xavier Estivill[46,47]; Raquel Rabionet[48], Cristòfol Vives-Bauzá[5]; Israel Fernández-Cadenas[3,6,49]; Jordi Jiménez-Conde*[1]

[1]Neurology, Neurovascular Research Group, Hospital del Mar Medical Research Institute, Universitat Autònoma de Barcelona/DCEXS-Universitat Pompeu Fabra, Barcelona; [2]Genetics, Universitat de Barcelona, Barcelona;[3]Neurovascular Research Laboratory, Vall d'Hebron Institute of Research, Universitat Autònoma de Barcelona, Barcelona;[4]Neurology, Son Espases University Hospital, IdISBa, Palma de Mallorca; [5]Research Unit, Son Espases University Hospital, Institut d'Investigació Sanitària de Les Illes Balears, Palma de Mallorca;m[6]Stroke Pharmacogenomics and Genetics, Fundació Docència i Recerca Mútua Terrassa, Terrassa; [7]A Coruña University Hospital and Biomedical Research Institute, La Coruña; [8]Neurology, Dr. Josep Trueta University Hospital, Girona Institute of Biomedical Investigation, Girona; [9]Stroke Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona; [10]Department of Neurology, Complejo Hospitalario Universitario de Albacete, Albacete; [11]Research Unit, Complejo Hospitalario Universitario de Albacete, Albacete; [12]Neurology, Hospital Clínic i Provincial de Barcelona, Barcelona; [13]Stroke Unit, Department of Neurology, Hospital Universitari Vall d'Hebron, Barcelona; [14]Neurology, Hospital de Mataró, Mataró; [15] Neurology, Hospital de Basurto, Bilbao; [16]Neurology, Hospital de Bellvitge, Hospitalet de Llobregat; [17]J. Philip Kistler Stroke Research Center, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA; [18]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, [19]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA; [20]Neurology, Washington University School of Medicine, St. Louis, MO, USA; [21]The Division of Emergency Medicine, Washington University School of Medicine, St. Louis, MO; [22]KU Leuven - University of Leuven, Department of Neurosciences, Experimental Neurology, Leuven, Belgium; [23]VIB, Center for Brain & Disease Research, Laboratory of Neurobiology, Leuven; [24]University Hospitals Leuven, Department of Neurology, Leuven, [25]Stroke Division, Florey Institute of Neuroscience and Mental Health, University of Melbourne, Heidelberg, Victoria, Australia; [26]Neurology, Austin Health, Heidelberg, Victoria, Australia; [27]Clinical Sciences Lund, Neurology, Lund University, Lund, Sweden; [28]Neurology and Rehabilitation Medicine, Skåne University Hospital, Lund, Sweden;[29]Faculty of Health, University of Technology, Sidney; [30]Hunter Medical Research Institute, University of Newcastle, Newcastle, Australia [31]Priority Research Centre for Stroke and Traumatic Brain Injury, University of Newcastle, Newcastle, Australia; [32]Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh; [33]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh;[34]Institute of

Biomedicine, the Sahlgrenska Academy at University of Gothenburg, Gothenburg; [35]Bioinformatics Core Facility, University of Gothenburg, Gothenburg; [36]Stroke Unit, Department of Neurosciences, Hospital Universitari Germans Trias i Pujol, Universitat Autònoma de Barcelona, Barcelona; [37]Stroke Unit, Hospital Universitario Central de Asturias (HUCA), Oviedo; [38]Neurology & Department of Public Health Sciences, University of Virginia, Charlottesville, VA, [39]Neurology and Rehabilitation Medicine, University of Cincinnati College of Medicine, Cincinnati, OH, [40]Neurology, University of Maryland School of Medicine and Baltimore VAMC, Baltimore, MD;[41]Geriatrics Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD; [42]Medicine, University of Maryland School of Medicine, Baltimore, MD; [43]Institute of Biomedicine of Seville, Hospital Universitario Virgen del Rocío, CSIC, Universidad de Sevilla, Sevilla; [44]Neurology, Hospital Universitario Virgen Macarena, Sevilla; [45]Neurology Unit, Neuroscience Department, Mútua de Terrassa Hospital, Terrassa, [46]Research Department, Sidra Medicine, Doha, Qatar; [47]Genomics Unit, Dexeus Woman's Health, Barcelona; [48]Centre for Genomic Regulation, Barcelona, and; [49]Stroke Pharmacogenomics and Genetics, Institut de Recerca Hospital de la Santa Creu i Sant Pau, Barcelona.

Running title: *PATJ* Variants and Ischemic Stroke Functional Outcome

Address correspondence to:

Dr. Jordi Jiménez-Conde
Hospital del Mar Medical
Research Institute
Room 162
Doctor Aiguader 88
08003 Barcelona
Spain
Tel.+34933160619
jjimenez@imim.es,

Dr. Israel Fernández-Cadenas
Sant Pau Institute of
Research
Sant Antoni Mª Claret 167
08025 Barcelona
Spain
Tel.+34651184139
israelcadenas@yahoo.es

Dr. Marina Mola-Caminal
Hospital del Mar Medical
Research Institute
Room 157
Doctor Aiguader 88
08003 Barcelona
Spain
Tel.+34933160623
marina.mola.c@gmail.com

ABSTRACT

Rationale: Ischemic stroke (IS) is among the leading causes of adult disability. Part of the variability in functional outcome after stroke has been attributed to genetic factors but no locus has been consistently associated with stroke outcome.

Objective: Our aim was to identify genetic loci influencing the recovery process using accurate phenotyping to produce the largest genome-wide association study (GWAS) in IS recovery to date.

Methods and Results: A 12-cohort, two-phase (discovery-replication and joint) meta-analysis of GWAS included anterior-territory and previously independent IS cases. Functional outcome was recorded using 3-month modified Rankin Scale (mRS). Analyses were adjusted for confounders such as discharge NIHSS. A gene-based burden test was performed. The discovery phase (n=1,225) was followed by *open* (n=2,482) and *stringent* joint-analyses (n=1,791). Those cohorts with mRS recorded at timepoints other than 3-month or incomplete data on previous functional status were excluded in the *stringent* analyses. Novel variants in *PATJ* gene were associated with worse functional outcome at 3-month after stroke. The top variant was rs76221407 (G allele, beta=0·40, p=1·70x10^{-9}).

Conclusions: Our results identify a set of common variants in *PATJ* gene associated with 3-month functional outcome at genome-wide significance level. Future studies should examine the role of *PATJ* in stroke recovery and consider *stringent* phenotyping to enrich the information captured to unveil additional stroke outcome loci.

Keywords:
Stroke, functional outcome, SNP, GWAS, ischemic stroke,

Nonstandard Abbreviations and Acronyms

1KGP: 1000 Genomes Project
BASICMAR: BASe de datos de ICtus del hospital del MAR
Chr: Chromosome
DNA: DeoxyriboNucleic Acid
GODS: Genetic contribution to functional Outcome and Disability after Stroke
GWAS: Genome-wide association study
HM: Hospital del Mar
HUSE: Hospital Universitari Son Espases
HUVH: Hospital Universitari Vall d'Hebron
IS: Ischemic Stroke
ISGC: International Stroke Genetics Consortium
LD: Linkage Disequilibrium
MAF: Minor Allele Frequency
mRS: modified Rankin Scale
NIHSS: National Institutes of Health Stroke Scale
PCs: Principal Components
rtPA: recombinant tissue Plasminogen Activator
SE: Standard Error
SNPs: Single Nucleotide Polymorphisms
TOAST: Trial of Org 10172 in Acute Stroke Treatment

INTRODUCTION

Ischemic stroke (IS) is the leading cause of adult disability[1] and the second cause of death worldwide.[2] Approximately 15 million people per year have a stroke, five million result in long-term disability.[3] More than a thousand potential targets for neuronal recovery have been identified,[4] although few have been tested in clinical trials. As no trials had positive results it is of high priority to find new drug targets for clinical practice.

Functional outcome after IS varies between individuals irrespective of clinical factors as initial stroke severity, stroke subtype, and vascular risk factors.[5] Multiple metabolic pathways are important in the response to cerebral ischemic damage and their activity may be modulated by variation in the genes of the involved components. New synaptic connections have been observed in areas surrounding a cerebral infarct within days following a stroke and this response is correlated with functional recovery.[6]

It is reasonable to presume that genetic variation may influence stroke recovery.[7] Candidate-gene studies reported the association between several genes and disability after stroke, but no locus has been found through a hypothesis-free genome-wide approach and most of the reported candidates failed to replicate in other cohorts.[8]

Genome-wide association studies (GWAS) have identified multiple single nucleotide polymorphisms (SNPs) contributing with a small effect to the risk of complex diseases, and different genes have arisen as potential therapeutic targets to reduce the risk of stroke.[9,10] Studying the genetic component of stroke outcome is of great scientific interest but requires very accurate phenotyping and proper attention to potentially confounding factors that may hide the true genetic contribution.

We aimed to find the genetics influencing the stroke recovery process in a dataset of first-ever IS patients, using highly accurate phenotyping and producing the largest GWAS in stroke mid-term functional outcome to date to identify new potential drug targets.

METHODS

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

Detailed Methods are provided in the Supplement.

Study design.
We conducted a two-phase analysis; discovery-replication and joint association of IS cases. The discovery consisted of a meta-analysis of four GWAS. The replication was performed as an *in silico* analysis of the top SNPs ($p<1x10^{-5}$) identified in the discovery. Two final joint meta-analyses according to *open* or *stringent* criteria were performed. The *open* joint meta-analysis (n=2,482) included four discovery and eight replication cohorts and the *stringent* joint meta-analysis (n=1,791) four discovery and three replication cohorts. This joint approach has proved to be more efficient than discovery-replication alone in increasing the statistical power.[11] A gene-based burden test was conducted.

Study sample.
European ancestry patients with a diagnosis of IS according to World Health Organization criteria were selected from the Spanish Stroke Genetics Consortium (GeneStroke) and the International Stroke Genetics Consortium.

Genome-wide association study.

After GWAS quality controls and imputation, association analyses with 3-month mRS were performed using an additive model and a multivariate linear regression. The common variants between the individual GWAS were joined into weighted z-score meta-analyses for 3-month mRS. All the SNPs with $p<5x10^{-8}$ were considered statistically significant for a genome-wide approach.

RESULTS

Genome-wide association study.

GWAS – Discovery phase: We meta-analyzed four individual GWAS with 1,225 individuals and 4,480,015 SNPs in the discovery phase. In the bivariate analysis (Online Table II), 3-month mRS showed association with age at stroke onset, sex, smoking status, stroke subtype by Trial of Org 10172 in Acute Stroke Treatment (TOAST),[12] and discharge NIHSS. The results include 79 variants in 18 independent ($r^2<0·001$) genomic regions associated with impaired 3-month mRS at $p<1\times10^{-5}$ (Figure 1, Online Table III), selected for replication. Global genomic inflation was $\lambda=1·032$. Best evidence was for three SNPs in *PATJ* on chromosome 1 exceeding $p<5x10^{-8}$: rs76221407 ($p=1·087x10^{-8}$, beta=0·42); rs150862264 ($p=1·539x10^{-8}$, beta=0·41); and rs182008837 ($p=1·825x10^{-8}$, beta=0·40). The top variant was imputed with $r^2=89·9\%$ and imputation certainty=99·2%.

GWAS – Replication phase: The SNPs selected for replication were analyzed in eight cohorts (Online Table IV). The *stringent* replication meta-analysis with three cohorts that strictly fulfil our selection criteria showed five SNPs in *PATJ* nominally associated with 3-month mRS ($p<0·05$), including the three top SNPs from the discovery (Online Table V). The effect size in the replication is slightly lower than in the discovery, indicating widespread consistency of results among the cohorts.

GWAS – Joint: A *stringent* joint meta-analysis was performed with the four discovery and the three *stringent* replication cohorts (1,791 individuals). The analysis revealed strong genetic association between 18 low frequency SNPs in *PATJ* and worse 3-month mRS (Table 1). The most striking SNP was rs76221407 ($p=1·72x10^{-9}$, beta=0·40), as shown in the forest plot (Online Figure IV), driven by a variant with low frequency (3%) in our European ancestry cohorts (which is consistent with 1KG data). Figure 3 shows the percentage of patients per group of 3-month mRS depending on the presence of the risk allele. An *open* joint meta-analysis was performed with all 12 cohorts (2,482 individuals) independently of meeting criteria for *stringent* analysis. Results revealed less significant p-values than the *stringent* analysis (rs76221407, $p=1·3x10^{-8}$, beta=0·37) (Figure 2, Online Table VI, Online Figure V). The % of the phenotypic variation in 3-month mRS accounted by the lead SNP is 0·27% in the whole sample (12 cohorts).

The same analyses through an ordinal regression showed loss of statistical power (*stringent* $p=3·60x10^{-5}$; *open* $p=4·33x10^{-5}$). Evidence of the robustness of the genetic association presented in this work is clearly shown in Figure 2, where the top SNP has a consistent effect direction in 11 out of 12 cohorts.

Gene-based study.

The gene-based association test revealed the *PATJ* gene as significantly associated with 3-month mRS. The 4·48 million genetic variants from the discovery GWAS were clustered in 23,972 genes, of which the only significant gene was *PATJ* ($p=1·99x10^{-6}$) (Online Table VII). Significance threshold was set at $p<0·05/23,972=2·086x10^{-6}$.

DISCUSSION

We report the first genetic findings in IS outcome using a GWAS. Our results show a novel association between low frequency genetic variants in *PATJ* gene and worse IS functional outcome measured with 3-month mRS. Genetic studies on IS outcome had previously focused on candidate loci (see Lindgren & Maguire).[7] However, findings had shown contradictory results and failed in consistent replications.[8] We performed a meta-analysis of 12 independent cohorts within the International Stroke Genetics Consortium, applying *open* and *stringent* criteria. Several low frequency variants in *PATJ* were significantly associated with worse functional outcome at 3-months. The lead variant, rs76221407, presents a consistent effect direction in 11 of 12 cohorts, providing convincing evidence of its robust genetic association with stroke outcome. The SNP is located in an intronic region of *PATJ* and shows LD with 17 other variants found through the *stringent* joint meta-analysis ($r^2 > 0.3$, in 1000 Genomes Project for European ancestry individuals). The fact that the variants are intronic may suggest that its effect on protein synthesis is carried out through the regulation of gene expression, similar to other common variants identified by GWAS that are linked to diseases by their modulation of the activity of DNA regulatory elements.[13] No previously established risk locus for stroke[10] has been related to stroke outcome in this work. Considering our restrictive inclusion criteria, which widely differ from the case/control studies, it is reasonable that we do not find overlaps in the top SNPs. None of our significant variants have been associated with another phenotype yet, although other *PATJ* polymorphisms have been related to sleep disturbance[14] and obesity-related traits.[15]

The results show a positive effect (beta=0·4) for the G allele of rs76221407, indicating that an increase of 0·4 points in the mRS score is attributed to each copy of G allele (GG>GA>AA). This means that G allele is related to poor functional outcome at 3-month. The results from the gene-based test also revealed significance for this locus ($p=1.99 \times 10^{-6}$) pointing out *PATJ* as a hot region of accumulated contributing variants. Although our *PATJ* SNPs are low frequency (~3%), the power of the association is strong enough to persist in the replication, evidencing the consistency of the results presented and the suitability of this gene as a therapeutic target.

PATJ, localized at tight junctions and at the apical membrane of epithelial cells, encodes a protein with seven PDZ domains, interaction modules that regulate multiple biological processes like ion channel signaling and transport.

In order to study the genetic component of a complex trait as IS functional outcome, it is key to be precise in characterizing the phenotype. The exclusion of posterior and lacunar strokes was considered necessary because these locations show a poor correlation between infarct size and clinical symptoms, and thus functional outcome.[16] A small lesion can be asymptomatic or show very severe symptoms with great disability depending on a variation of just few millimeters in its location. In these cases, recovery processes and tissue regeneration mechanisms could be masked by this random location effect.

Minor strokes (initial NIHSS≤4) and individuals with a dependent status previous to the stroke event (mRS>2) were also excluded. This was done to permit the analyses of an equivalent recovery process, without taking into consideration the degree of disability before stroke or those patients with only minimal damage to be recovered. In both cases, it would be difficult to evaluate the significance of recovery at 3-month. The achievement of a highly homogeneous sample was one of the main priorities during the study design. This led to the performance of two types of joint meta-analyses. The more permissive inclusion criteria of the *open* analysis led to a larger sample size (1·5 times greater) but less significant results compared to the *stringent* analysis. The reduction of phenotypic heterogeneity by properly defining the study cohort increases statistical power.[17] While the phenotypic homogeneity of the sample is the main strength of our study, it also limited the sample size, as very few cohorts worldwide have the complete data needed for the *stringent* analysis[8] and this may prevent the discovery of other loci.

However, our work demonstrates the value of prioritizing homogeneity and phenotyping accuracy over a larger sample size. The mRS, the most widely used scale in stroke patients to assess functional outcome, has only seven categories but it offers the advantages of being easy to apply and having good inter-observer reproducibility.[18] Analyzing the mRS score as a continuous instead of an ordinal value,[19] according to Rhemtulla et al, is the preferable choice for ordinal data with more than four categories in contrast to robust categorical methodology,[20] and the statistical power is improved.

This project generated extensive genotyping and phenotyping of individual-level data that can help to disentangle the genetic architecture of the stroke recovery process. The use of whole exome or genome sequencing would provide information about rare variants, which may account for a greater proportion of the stroke outcome's genetic component than common variants. Additional functional studies are warranted to establish whether *PATJ* can reveal biological pathways that could be novel therapeutic targets to improve post-IS rehabilitation strategies.

DISCLOSURES
None.

REFERENCES

1   Mozaffarian D, Benjamin EJ, Go AS, et al. Heart Disease and Stroke Statistics--2015 Update: A Report From the American Heart Association. *Circulation* 2015; 131:e535.

2   World Health Organization. The top 10 causes of death. May 2014. http://www.who.int/mediacentre/factsheets/fs310/en/ (accessed June 11, 2017).

3   Feigin VL, Lawes CMM, Bennett DA, Anderson CS. Stroke epidemiology: A review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *Lancet Neurol* 2003; 2:43–53.

4   Mahar M, Cavalli V. Intrinsic mechanisms of neuronal axon regeneration. *Nat Rev Neurosci* 2018; 19:10–3.

5   Roquer J, Ois A, Rodriguez-Campello A, Gomis M, Munteis E, Jimenez-Conde J, Cuadrado-Godia E, Martinez-Rodriguez JE. Atherosclerotic burden and early mortality in acute ischemic stroke. *Arch Neurol* 2007; 64:699–704.

6   Dijkhuizen RM, Singhal AB, Mandeville JB, Wu O, Halpern EF, Finklestein SP, Rosen BR, Lo EH. Correlation between brain reorganization, ischemic damage, and neurologic status after transient focal cerebral ischemia in rats: a functional magnetic resonance imaging study. *J Neurosci* 2003; 23:510–7.

7   Lindgren A, Maguire J. Stroke Recovery Genetics. *Stroke* 2016; 47:2427–34.

8   Maguire JM, Bevan S, Stanne TM, et al. GISCOME – Genetics of Ischaemic Stroke Functional Outcome network: A protocol for an international multicentre genetic association study. *Eur Stroke J* 2017; 2:229–37.

9   NINDS Stroke Genetics Network (SiGN) ISGC (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol* 2015; 15:4–7.

10   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* 2018; 50(4):524–

37.

11 Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 2006; 38:209–13.

12 Adams HP, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* 1993; 24:35–41.

13 Maurano MT, Haugen E, Sandstrom R, Vierstra J, Shafer A, Kaul R, Stamatoyannopoulos JA. Large-scale identification of sequence variants influencing human transcription factor occupancy in vivo. *Nat Genet* 2015; 47:1393–401.

14 Lane JM, Liang J, Vlasac I, et al. Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. *Nat Genet* 2016; 49:274–281.

15 Locke A, Kahali B, Berndt S, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; 518:197–206.

16 Sohn YH, Lee BI, Sunwoo IN, Kim KW, Suh JH. Effect of capsular infarct size on clinical presentation of stroke. *Stroke* 1990; 21:1258–61.

17 Burmeister M, McInnis MG, Zöllner S. Psychiatric genetics: progress amid controversy. *Nat Rev Genet* 2008; 9:527–40.

18 Shinohara Y, Minematsu K, Amano T, Ohashi Y. Modified rankin scale with expanded guidance scheme and interview questionnaire: Interrater agreement and reproducibility of assessment. *Cerebrovasc Dis* 2006; 21:271–8.

19 Nunn A, Bath PM, Gray LJ. Analysis of the modified Rankin scale in randomised controlled trials of acute ischaemic stroke: A systematic review. *Stroke Res Treat* 2016; 2016:9482876.

20 Rhemtulla M, Brosseau-Liard PÉ, Savalei V. When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychol Methods* 2012; 17:354–73.

SNP	Position	Gene	SNP type	MA	MAF	Beta	Beta (80% power)	SE	p
rs76221407	1:62131826	*PATJ*	intronic	G	0·03	0·40	0·44	0·07	$1·72 \times 10^{-9}$
rs150862264	1:62132045	*PATJ*	intronic	C	0·03	0·40	0·44	0·07	$1·76 \times 10^{-9}$
rs182008837	1:62107102	*PATJ*	intronic	C	0·03	0·39	0·43	0·07	$2·93 \times 10^{-9}$
rs117335978	1:62140649	*PATJ*	intronic	T	0·02	0·46	0·54	0·08	$1·96 \times 10^{-8}$
rs137999692	1:62092932	*PATJ*	intronic	A	0·03	0·36	0·43	0·07	$2·98 \times 10^{-8}$
rs75717958	1:62141064	*PATJ*	intronic	A	0·02	0·47	0·55	0·08	$3·71 \times 10^{-8}$
rs7546744	1:62141462	*PATJ*	intronic	A	0·02	0·47	0·55	0·08	$3·71 \times 10^{-8}$
rs7513982	1:62141944	*PATJ*	intronic	C	0·02	0·47	0·55	0·08	$3·72 \times 10^{-8}$
rs17123133	1:62142178	*PATJ*	intronic	A	0·02	0·47	0·55	0·08	$3·72 \times 10^{-8}$
rs7514107	1:62142096	*PATJ*	intronic	C	0·02	0·47	0·55	0·08	$3·72 \times 10^{-8}$
rs141479296	1:62140561	*PATJ*	intronic	G	0·02	0·47	0·55	0·08	$3·77 \times 10^{-8}$
rs11805802	1:62142912	*PATJ*	intronic	C	0·02	0·47	0·55	0·08	$3·78 \times 10^{-8}$
rs10157504	1:62140248	*PATJ*	intronic	G	0·02	0·47	0·55	0·08	$3·80 \times 10^{-8}$
rs74469018	1:62144275	*PATJ*	intronic	T	0·02	0·47	0·56	0·08	$3·86 \times 10^{-8}$
rs77007585	1:62144350	*PATJ*	intronic	A	0·02	0·47	0·56	0·08	$3·88 \times 10^{-8}$
rs11806656	1:62144566	*PATJ*	intronic	C	0·02	0·47	0·56	0·08	$3·88 \times 10^{-8}$
rs7542598	1:62145477	*PATJ*	intronic	T	0·02	0·47	0·56	0·08	$3·95 \times 10^{-8}$
rs118168181	1:62141702	*PATJ*	intronic	A	0·02	0·47	0·56	0·09	$3·98 \times 10^{-8}$

Table 1: Intronic variants in *PATJ* associated with 3-month mRS at $p<5 \times 10^{-8}$ in the *stringent* joint-analysis

SNP, single nucleotide polymorphism; Position, chromosome: genomic coordinates according to human genome reference version 19 (build GRCh38/hg38); MA, minor allele; MAF, minor allele frequency; Beta, beta value for the association; Beta (80% power), beta that could be detected as statistically significant under 80% power; SE, standard error; p, p-value.

FIGURE LEGENDS

Figure 1: Manhattan plot (A) and Q-Q Plot (B) of the discovery results. Association testing was performed using a lineal regression model adjusted for the first two PCs, sex, age, smoking, stroke subtype, and discharge NIHSS. The red line shows $p=1\times10^{-5}$ and the blue line the GWAS significance threshold ($p=5x10^{-8}$) (A). The *X*-axis is the expected $-\log_{10}$ p under the null hypothesis and lambda is the observed median χ^2 test statistic divided by the median expected χ^2 test statistic under the null hypothesis (B).

Figure 2: Forest plot (A) and regional association plot (B) for rs76221407 in the *open* joint-analysis. Plot of the effect size of the association with 3-month mRS across the 12 cohorts (*open* joint-analysis) (A). Association of rs76221407 and other SNPs in the region was plotted with $-\log_{10}$ p-values (left *y*-axis), the estimated local recombination rate in blue (right *y*-axis). The signals are distinguished by color and shape. Linkage disequilibrium (r^2) of nearby SNPs is shown by color gradient (B).

Figure 3: Distribution of 3-month mRS score in the *stringent* joint-analysis. *Joining in the same category the mRS values 5-6 and in the same group the G allele carriers (heterozygous and homozygous) is uniquely suited for facilitating the visualization of this figure, although the analyses are done without joining these categories.

NOVELTY AND SIGNIFICANCE
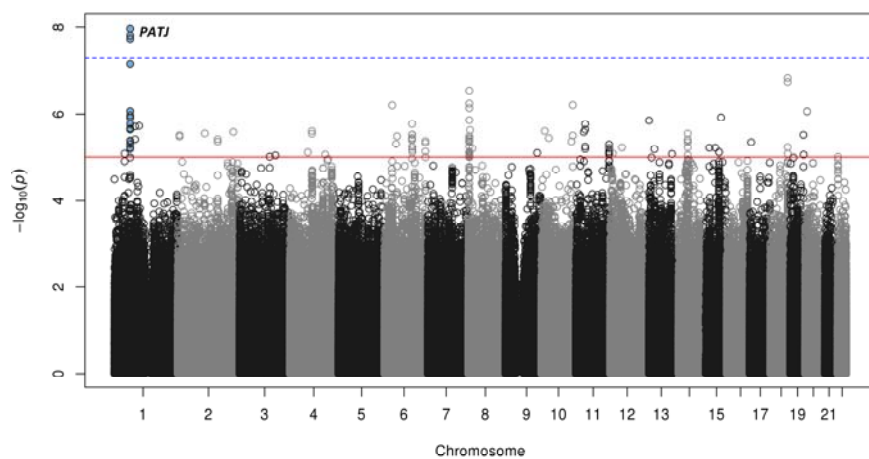
What Is Known?

- Disability due to stroke has a significant impact on public health as it affects the quality of life of both – the patients and the caregivers.

- Irrespective of clinical factors, functional recovery after ischemic stroke (IS) varies widely among individuals.

- Variability in functional outcome has been attributed, in part, to genetic factors, but to date, no locus has been consistently associated with stroke outcome.

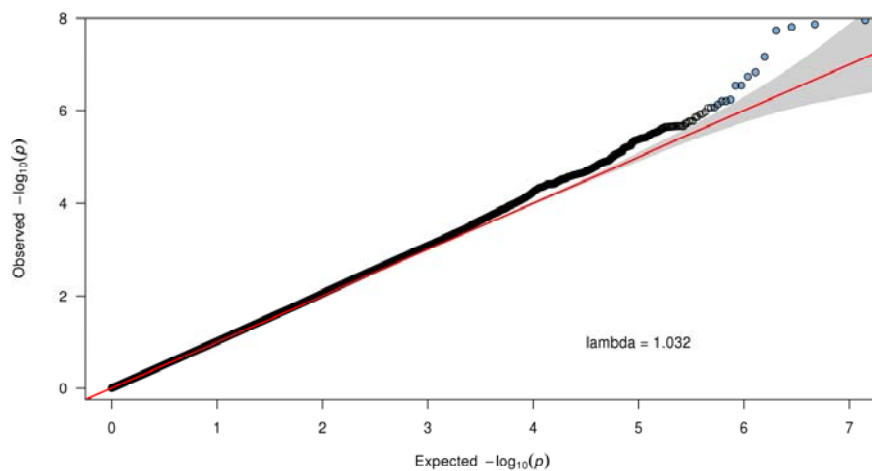### *What New Information Does This Article Contribute?*

- Genetic variants in *PATJ* gene are associated with IS functional outcome at 3 months.

- Association of *PATJ* gene variants with IS functional outcome was discovered through a meta-analysis of genome-wide association studies (GWAS) and was validated in a multiple-cohort replication study.
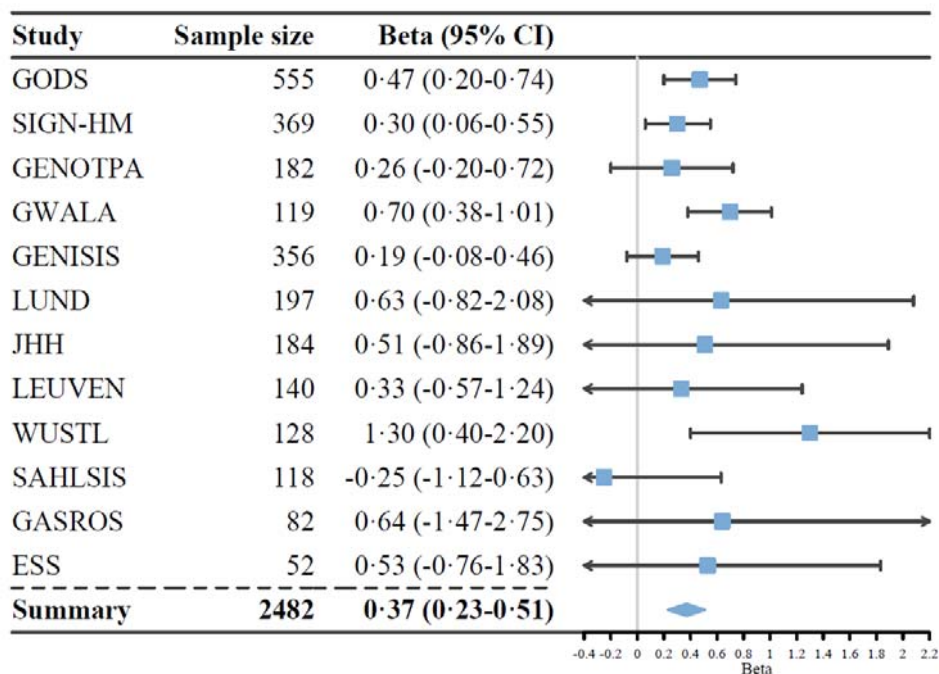
Cerebrovascular disease is the leading cause of adult disability. Mid-term functional recovery after stroke varies significantly among individuals, independent of infarct size, stroke subtype, vascular risk factors, or clinical status after acute treatment. Identifying genetic factors that contribute to this variability requires a hypothesis-free, comprehensive genotyping approach, such as GWAS, as well as accurate phenotyping to identify confounding factors that may obscure the genetic effects. Using a restrictive inclusion criteria to obtain a less heterogeneous patient population, in a multi-cohort GWAS we found that *PATJ* gene variants were associated with functional outcomes in patients with IS. The accumulation of risk alleles in the *PATJ* gene was associated with a worse functional outcome at 3-month after IS. This evidence for a genetic contribution to mid-term stroke prognosis, provides a new platform for understanding the mechanisms that determine functional recovery after stroke, and may also help in better prediction of functional outcomes after IS.

FIGURE 1

A

B

FIGURE 2

Study	Sample size	Beta (95% CI)	
GODS	555	0·47 (0·20-0·74)	
SIGN-HM	369	0·30 (0·06-0·55)	
GENOTPA	182	0·26 (-0·20-0·72)	
GWALA	119	0·70 (0·38-1·01)	
GENISIS	356	0·19 (-0·08-0·46)	
LUND	197	0·63 (-0·82-2·08)	
JHH	184	0·51 (-0·86-1·89)	
LEUVEN	140	0·33 (-0·57-1·24)	
WUSTL	128	1·30 (0·40-2·20)	
SAHLSIS	118	-0·25 (-1·12-0·63)	
GASROS	82	0·64 (-1·47-2·75)	
ESS	52	0·53 (-0·76-1·83)	
Summary	**2482**	**0·37 (0·23-0·51)**	

B

FIGURE 3

3-month mRS

| | 0 | | 1 | | 2 | | 3 | | 4 | | 5-6 |

| AA (N = 1684) | 15,4 | 16,4 | 15,4 | 16,7 | 15,8 | 20,2 |
| AG+GG (N = 107) | 11,7 | 11,7 | 10,7 | 18,4 | 19,4 | 28,2 |

0%   20%   40%   60%   80%   100%

Patients (%)