



UNIVERSITY OF HELSINKI

<https://helda.helsinki.fi>

Multi-ancestry GWAS reveals excitotoxicity associated with outcome after ischaemic stroke

Ibanez, Laura; Heitsch, Laura; Carrera, Caty; Farias, Fabiana H. G.; Del Aguila, Jorge L. ...

2022-07-29

Oxford University Press

<http://hdl.handle.net/10138/355105>

Ibanez, L, Heitsch, L, Carrera, C, Farias, F H G, Del Aguila, J L, Dhar, R, Budde, J, Bergmann, K, Bradley, J, Harari, O, Phuah, C L, Lemmens, R, Viana Oliveira Souza, A A, Moniche, F, Cabezas-Juan, A, Arenillas, J F, Krupinski, J, Cullell, N, Torres-Aguila, N, Muino, E, Carcel-Marquez, J, Marti-Fabregas, J, Delgado-Mederos, R, Marin-Bueno, R, Hornick, A, Vives-Bauza, C, Navarro, R D, Tur, S, Jimenez, C, Obach, V, Segura, T, Serrano-Heras, G, Chung, J W, Roquer, J, Soriano-Tarraga, C, Giralt-Steinhauer, E, Mola-Caminal, M, Pera, J, Lapicka-Bodzioch, K, Derbisz, J, Davalos, A, Lopez-Cancio, E, Munoz, L, Tatlisumak, T, Molina, C, Ribo, M, Bustamante, A, Sobrino, T, Castillo-Sanchez, J, Campos, F, Rodriguez-Castro, E, Arias-Rivas, S, Rodriguez-Yanez, M, Herbosa, C, Ford, A L, Gutierrez-Romero, A, Uribe-Pacheco, R, Arauz, A, Lopes-Cendes, I, Lowenkopf, T, Barboza, M A, Amini, H, Stamova, B, Ander, B P, Sharp, F R, Kim, G M, Bang, O Y, Jimenez-Conde, J, Slowik, A, Stribian, D, Tsai, E A, Burkly, L C, Montaner, J, Fernandez-Cadenas, I, Lee, J M & Cruchaga, C 2022, 'Multi-ancestry GWAS reveals excitotoxicity associated with outcome after ischaemic stroke', *Brain : a journal of neurology*, vol. 145, no. 7, pp. 2394-2406. <https://doi.org/10.1093/brain/awac080>

Downloaded from Helda, University of Helsinki institutional repository. <https://helda.helsinki.fi>

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Multi-ancestry GWAS reveals excitotoxicity associated with outcome after ischaemic stroke

Laura Ibanez,^{1,2} Laura Heitsch,^{3,4} Caty Carrera,⁵ Fabiana H. G. Farias,^{1,2} Jorge L. Del Aguila,^{1,2} Rajat Dhar,³ John Budde,^{1,2} Kristy Bergmann,^{1,2} Joseph Bradley,^{1,2} Oscar Harari,^{1,2,6,7} Chia-Ling Phuah,³ Robin Lemmens,⁸ Alessandro A. Viana Oliveira Souza,^{9,10} Francisco Moniche,¹¹ Antonio Cabezas-Juan,^{11,12} Juan Francisco Arenillas,¹³ Jerzy Krupinski,^{14,15} Natalia Cullell,^{15,16} Nuria Torres-Aguila,^{15,16} Elena Muiño,¹⁶ Jara Cárcel-Márquez,¹⁶ Joan Marti-Fabregas,¹⁶ Raquel Delgado-Mederos,¹⁶ Rebeca Marin-Bueno,¹⁶ Alejandro Hornick,¹⁷ Cristofol Vives-Bauza,¹⁸ Rosa Diaz Navarro,¹⁹ Silvia Tur,¹⁹ Carmen Jimenez,¹⁹ Victor Obach,²⁰ Tomas Segura,²¹ Gemma Serrano-Heras,²¹ Jong-Won Chung,²² Jaume Roquer,²³ Carol Soriano-Tarraga,^{1,2,23} Eva Giralt-Steinhauer,²³ Marina Mola-Caminal,^{23,24} Joanna Pera,²⁵ Katarzyna Lapicka-Bodzioch,²⁵ Justyna Derbisz,²⁵ Antoni Davalos,²⁶ Elena Lopez-Cancio,²⁷ Lucia Muñoz,²⁶ Turgut Tatlisumak,^{28,29} Carlos Molina,⁵ Marc Ribo,⁵ Alejandro Bustamante,²⁶ Tomas Sobrino,³⁰ Jose Castillo-Sanchez,³⁰ Francisco Campos,³⁰ Emilio Rodriguez-Castro,³⁰ Susana Arias-Rivas,³⁰ Manuel Rodríguez-Yáñez,³⁰ Christina Herbosa,³ Andria L. Ford,^{3,6,31} Alonso Gutierrez-Romero,³² Rodrigo Uribe-Pacheco,³² Antonio Arauz,³² Iscia Lopes-Cendes,^{9,10} Theodore Lowenkopf,³³ Miguel A. Barboza,³⁴ Hajar Amini,³⁵ Boryana Stamova,³⁵ Bradley P. Ander,³⁵ Frank R Sharp,³⁵ Gyeong Moon Kim,²² Oh Young Bang,²² Jordi Jimenez-Conde,²³ Agnieszka Slowik,²⁵ Daniel Stribian,³⁶ Ellen A. Tsai,³⁷ Linda C. Burkly,³⁸ Joan Montaner,^{5,11,12} Israel Fernandez-Cadenas,^{5,16} Jin-Moo Lee,^{3,6,31,39,40} and Carlos Cruchaga,^{1,2,3,6,7,41}

1 Department of Psychiatry, Washington University School of Medicine, Saint Louis (63110), Missouri, USA

2 NeuroGenomics and Informatics, Washington University School of Medicine, Saint Louis (63110), Missouri, US

3 Department of Neurology, Washington University School of Medicine, Saint Louis (63110), Missouri, US

4 Department of Emergency Medicine, Washington University School of Medicine, Saint Louis (63110), Missouri, USA

5 Stroke Unit, Vall d'Hebron University Hospital, Universitat de Barcelona, Barcelona (08035), Spain

6 Hope Center for Neurological Disorders, Washington University School of Medicine, Saint Louis (63110), Missouri, USA

7 The Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington University School of Medicine, Saint Louis (63110), Missouri, USA

8 Department of Neuroscience, Katholieke Universiteit Leuven, Campus Gasthuisberg O&N2; Leuven (BE-3000), Belgium

9 Department of Neurology, School of Medical Sciences, University of Campinas (UNICAMP), Cidade Universitaria, Campinas (13083-887), Brazil

- 1 10 Brazilian Institute of Neuroscience and Neurotecnology (BRAINN), R. Tessalia Viera de Camargo,
2 Campinas (13083-887), Brazil
- 3 11 Department of neurology, Hospital Virgen del Rocio, University of Seville, Seville (41013), Spain
- 4 12 Hospital Virgen de la Macarena, University of Seville, Seville (41009), Spain
- 5 13 Department of Neurology, Hospital Clinico Universitario Valladolid, Valladolid University, Valladolid
6 (47003), Spain
- 7 14 Department of Neurology, Mutua Terrassa University Hospital, Universitat de Barcelona, Terrassa
8 (08221), Spain
- 9 15 Fundacio Docencia i Recerca Mutua Terrassa, Universitat de Barcelona, Terrassa (08221), Spain
- 10 16 Department of Neurology, Hospital de la Santa Creu i Sant Pau, Universitat Autonoma de Barcelona,
11 Barcelona (08041), Spain
- 12 17 Department of Neurology, Southern Illinois Healthcare Memorial Hospital of Carbondale, Carbondale
13 (62901), Illinois, USA
- 14 18 Department of Biology, Universitat de les Illes Balears, Palma (07122), Spain
- 15 19 Department of Neurology, Hospital Universitari Son Espases, Universitat de les Illes Balears, Palma
16 (07120), Spain
- 17 20 Department of Neurology, Hospital Clinic de Barcelona, Universitat de Barcelona, Barcelona (08036),
18 Spain
- 19 21 Research Unit, Complejo Hospitalario Universitario de Albacete. Albacete (02008), Spain
- 20 22 Department of Neurology, Samsung Medical Center, Seoul, South Korea
- 21 23 Neurovascular Research Group, Institut Hospital del Mar de Investigacions Mediques, Barcelona
22 (08003), Spain
- 23 24 Department of Surgical Sciences, Orthopedics, Uppsala University, Uppsala, 75185, Sweden
- 24 25 Department of Neurology, Jagiellonian University, Krakow(31-007), Poland
- 25 26 Department of Neurology, Hospital Germans Trias i Pujol, Universitat Autonoma de Barcelona,
26 Badalona (08916), Spain
- 27 27 Department of Neurology, Hospital Universitario Central de Asturias, Oviedo, Spain
- 28 28 Department of Neurology, Sahlgrenska University Hospital, University of Gothenburg, Gothenburg
29 (413 45), Sweden
- 30 29 Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, Sahlgrenska Academy
31 at University of Gothenburg, Gothenburg, Sweden
- 32 30 Clinical Neurosciences Research Laboratory, Health Research Institute of Santiago de Compostela

- 1 (IDIS), Santiago de Compostela (15706), Spain
2 31 Department of Radiology, Washington University School of Medicine, Saint Louis (63110), Missouri,
3 USA
4 32 Instituto Nacional de Neurologia y Neurocirurgia de Mexico, Ciudad de Mexico (14269), Mexico
5 33 Department of Neurology, Providence St. Vincent Medical Center, Portland (97225), Oregon, USA
6 34 Neurosciences Department, Hospital Rafael A. Calderon Guardia, Aranjuez, San José, Costa Rica
7 35 Department of Neurology and MIND Institute, University of California at Davis, Sacramento (95817),
8 California, USA
9 36 Department of Neurology, Helsinki University Hospital, Helsinki (00290), Finland
10 37 Translational Biology, Biogen, Inc, Cambridge (02142), Massachusetts, USA
11 38 Genetics and Neurodevelopmental Disease Research Unit, Biogen, Inc, Cambridge (02142),
12 Massachusetts, USA
13 39 Department of Biomedical Engineering, Washington University School of Medicine, Saint Louis
14 (63110), Missouri, USA
15 40 Stroke and Cerebrovascular Center, Washington University School of Medicine, Saint Louis (63110),
16 Missouri, US
17 41 Department of Genetics, Washington University School of Medicine, Saint Louis (63110), Missouri,
18 USA

19 Correspondence to: Carlos Cruchaga
20 Washington University School of Medicine
21 660 South Euclid Avenue
22 Campus Box 8134
23 Saint Louis, MO 63110, USA
24 E-mail: cruchagac@wustl.edu
25

26 Correspondence may also be addressed to: Jin-Moo Lee
27 Washington University School of Medicine
28 660 South Euclid Avenue
29 Campus Box 8111
30 St. Louis, MO 63110, USA
31 E-mail: leejm@wustl.edu

32 **Running title:** Excitotoxicity and ischaemic stroke
33

1 Abstract

2 During the first hours after stroke onset neurological deficits can be highly unstable: some patients
3 rapidly improve, while others deteriorate. This early neurological instability has a major impact on long-
4 term outcome. Here, we aimed to determine the genetic architecture of early neurological instability
5 measured by the difference between NIH stroke scale (NIHSS) within six hours of stroke onset and NIHSS
6 at 24h (Δ NIHSS). A total of 5,876 individuals from seven countries (Spain, Finland, Poland, United States,
7 Costa Rica, Mexico and Korea) were studied using a multi-ancestry meta-analyses. We found that 8.7%
8 of Δ NIHSS variance was explained by common genetic variations, and also that early neurological
9 instability has a different genetic architecture than that of stroke risk. Eight loci (1p21.1, 1q42.2, 2p25.1,
10 2q31.2, 2q33.3, 5q33.2, 7p21.2, and 13q31.1) were genome-wide significant and explained 1.8% of the
11 variability suggesting that additional variants influence early change in neurological deficits. We used
12 functional genomics and bioinformatic annotation to identify the genes driving the association from
13 each loci. eQTL mapping and SMR indicate that *ADAM23* (log Bayes Factor (LBF)=5.41) was driving the
14 association for 2q33.3. Gene based analyses suggested that *GRIA1* (LBF=5.19), which is predominantly
15 expressed in brain, is the gene driving the association for the 5q33.2 locus. These analyses also
16 nominated *GNPAT* (LBF=7.64) *ABCB5* (LBF=5.97) for the 1p21.1 and 7p21.1 loci. Human brain single
17 nuclei RNA-seq indicates that the gene expression of *ADAM23* and *GRIA1* is enriched in neurons.
18 *ADAM23*, a pre-synaptic protein, and *GRIA1*, a protein subunit of the AMPA receptor, are part of a
19 synaptic protein complex that modulates neuronal excitability. These data provides the first genetic
20 evidence in humans that excitotoxicity may contribute to early neurological instability after acute
21 ischemic stroke.

22 **Keywords:** ischaemic stroke; neuroprotection; genetics; NIHSS

23 **Abbreviations:** AfA=African-American descent; BMI=Body Mass Index; eQTL=Expression Quantitative
24 Trait Loci; EuA=European Descent; GCTA=Genome-wide Complex Traits Analysis; GENISIS=Genetics of
25 Early Neurological InStability after Ischemic Stroke; GTEx=Genotype-Tissue Expression; GWAS=Genome-
26 Wide Association Studies; LBF=Log Bayes Factor; LD=Linkage Disequilibrium; MAF=Minor Allele
27 Frequency; NIH=National Institute of Health; NIHSS=NIH Stroke Scale; SMR=Summary-data-based
28 Mendelian Randomization; SNP=Single Nucleotide Polymorphism

29

1 Introduction

2 Stroke is the second most common cause of death and the most common cause of disability,
3 worldwide.¹ Ischemic stroke, the most common subtype², is caused by the occlusion of an artery in the
4 brain, resulting in the abrupt development of cerebral ischemia and neurological deficits.³ During the
5 first hours after stroke onset, neurological deficits can be highly unstable with some patients
6 demonstrating rapid deterioration, while others rapidly improve.⁴ In fact, early change in neurological
7 deficits have a major influence on long-term outcome. NIH stroke scale (NIHSS) changes from baseline
8 (within 6 hours of stroke onset) to 24 hours after acute ischemic stroke (Δ NIHSS) have a significant and
9 independent association with favorable 90-day outcome, accounting for more than 30% of the explained
10 variance.⁴⁻⁶ A number of mechanisms are thought to account for these early changes including
11 fibrinolysis and reperfusion, hemorrhagic transformation, etiology, and endogenous neuroprotective
12 mechanisms.⁷⁻¹⁴

13 Prior genome wide association studies (GWAS), mostly in populations of European descent, have
14 identified numerous loci associated with stroke risk. In 2018, the MEGASTROKE consortia performed one
15 of the largest GWAS to date, combining most of the available GWAS for stroke risk in a unique multi-
16 ancestry meta-analysis including 67,162 cases and 454,450 controls. This analysis led to the discovery of
17 22 novel loci, bringing the total stroke risk loci to 32. Many loci were previously linked to other vascular
18 traits (blood pressure, cardiac phenotypes, venous thromboembolism); while others had no obvious
19 connection with stroke, warranting further investigation to identify potentially novel mechanisms.¹⁵ A
20 similar approach, used to decipher the genetics of long term disability after ischemic stroke in 6,165
21 non-Hispanic Whites, identified one locus that was been not replicated so far.^{16,17} However, to date
22 there have been no genetic studies examining early neurological change after ischemic stroke.

23 To our knowledge, the Genetics of Early Neurological InStability after Ischemic Stroke (GENISIS) is the
24 largest well-characterized study for early outcomes quantified by Δ NIHSS.¹⁸ To increase the power to
25 detect genetic associations, our study recruited patients from multiple diverse ancestry groups. We
26 leveraged the GENISIS cohort using Δ NIHSS as a quantitative phenotype, to identify novel variants,
27 genes and pathways associated with early neurological instability after ischemic stroke.

28

1 **Materials and methods**

2 **Study design**

3 A detailed description of the acute ischemic stroke patients recruited from 21 sites from seven countries
4 throughout the world, has been published elsewhere.⁶ Briefly, adult acute ischemic stroke patient with
5 measurable deficit on the NIHSS that presented within 6 hours of stroke onset (or last known normal)
6 were enrolled in the study after obtaining informed consent, including patients treated with tPA. All
7 available inpatient data, including history, clinical exam, lab values, diagnostic tests, imaging, and
8 discharge diagnosis were utilized to confirm the diagnosis of ischemic stroke. Patients who underwent a
9 thrombectomy, were enrolled in other treatment trials, or for whom consent and/or a blood sample
10 could not be obtained were excluded. Demographics, co-morbidities, acute treatment variables, imaging
11 data and TOAST classification were collected.

12 To accommodate the difference in the genetic architecture intrinsic to the country of origin, we
13 performed a three-stage analysis (Fig. 1A). First, we used an additive model to perform a GWAS in each
14 country individually, except for the United States, where the population was stratified into European
15 and African ancestry cohorts. We then performed a fixed effects meta-analyses within the same ethnic
16 cohorts. Finally, we used a multi-ancestry Bayesian meta-analysis to collapse all the ethnic backgrounds.
17 Unlike a fixed effect meta-analysis, the Bayesian approach is able to account for population structure
18 differences.¹⁹ Genetic loci that passed multiple test correction, a threshold set at Log Bayes Factor (LBF)
19 $> 5^{19,20}$, were annotated using bioinformatics tools to identify the gene driving the genetic signal (Fig.
20 1B). We used functional annotation, multi-tissue expression quantitative trait loci (eQTL) data, and
21 summary-data-based Mendelian randomization (SMR) to map the genome-wide to specific genes. Single
22 nuclei- RNA-seq data derived from cortex samples was used to determine potential correlation between
23 the transcripts of the identified genes and determine in which brain cell types the genes are expressed.²¹
24 The study was approved by the Institutional Review Boards at every participating site. Written informed
25 consent was obtained from all participants or their family members. All research was performed
26 according to the approved protocols and consents.

27 **Genotyping**

28 All participants were genotyped using Illumina SNP array technology. Samples were genotyped in seven
29 batches during the GENISIS recruitment (see Supplementary Methods). Genotyping quality control and

1 imputation were performed separately for each genotyping round using SHAPEIT²² and IMPUTE2.²³ For
2 each genotyping batch, SNPs with a call rate lower than 98% and autosomal SNPs that were not in
3 Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$) were removed from the dataset. The X chromosome SNPs were
4 used to determine sex based on heterozygosity rates, and samples with discordant inferred sex and
5 reported sex were removed. Only samples with call rate greater than 98% were considered to pass
6 quality control. Finally, the genotype batches were merged in a single file to perform the analyses.
7 Additional QC was performed in the merged dataset. We tested pairwise genome-wide estimates of
8 proportion identity-by-descent, the presence of unexpected duplicates, and cryptically relatedness (PI-
9 HAT > 0.30). Of the pairs of these samples flagged, the sample with higher genotyping rate was kept for
10 downstream analysis. Principal component analysis (PCA) was performed using HapMap as an anchor to
11 remove ethnic outliers and keep the populations as homogeneous as possible for each of the participant
12 countries. Principal components were also used to cluster and identify ancestry populations for US
13 participants with European descent (EuA) and African-American descent (AfA). Samples outside two
14 standard deviations from the center of the Non-Hispanic White or the Asian cluster were considered
15 outliers for Spain, Finland and Poland. We confirmed the ethnicity of the AfA and Hispanic populations,
16 however, due to the genetic heterogeneity present in these populations we did not remove the samples
17 outside two standard deviations from the mean.

18 **Analysis of variance**

19 We used genome-wide complex traits analysis (GCTA) to determine the heritability of Δ NIHSS.²⁴ GCTA
20 estimates the amount of phenotypic variance in a given complex trait explained by all the SNPs and fits
21 the effects of these SNPs as random effects in a linear mixed model. Because it relies on a large,
22 homogeneous populations for accurate results, we only included the individuals with non-Hispanic
23 White ancestry.

24 **Single variant analyses**

25 To mitigate the effects of genetic heterogeneity due to the diverse ancestry of participants enrolled in
26 the GENISIS study, we used a multi-step study design (Fig. 1A). First, we performed single variant
27 analyses each participant country separately. We tested the association of SNPs across the genome with
28 Δ NIHSS using an additive linear model with PLINK 1.9.²⁵ Sex, age, and the two Principal Components
29 calculated for each population were included in the model. Additional covariates include the SNP

1 genotyping batch, TOAST classification (using dummy variables to incorporate all subtypes), tPA, and
2 baseline NIHSS to adjust for stroke severity. The primary focus of the GWAS was on early neurological
3 change; thus, baseline NIHSS was included as a covariate in the model. Although baseline NIHSS was
4 used to calculate Δ NIHSS, it does not fully explain the observed variance in Δ NIHSS; further, there is no
5 multicollinearity between these two variables, permitting their inclusion in the model.²⁶ Second, we
6 meta-analyzed the populations with similar ethnic backgrounds using with fixed effect meta-analyses
7 using METAL.²⁷ We performed two meta-analyses, one for the non-Hispanic Whites (Spain, Finland,
8 Poland, and United States EuA) and one for the Hispanics (Costa Rica and Mexico). Finally, we analyzed
9 the four available ethnicities non-Hispanic Whites (meta-analysis), Hispanics (meta-analysis), Asians
10 (Korea) and African Americans (United States AfA) using MANTRA, a Bayesian-based multi-ancestry
11 meta-analyses.¹⁹ Log Bayes Factor (LBF) greater than 5 was considered to be genome wide significant
12 after multiple test correction.^{19,20}

13 To ensure that the loci were related to Δ NIHSS in all ischemic strokes and were not specific to a stroke
14 subtype (defined by TOAST criteria), we also performed joint analyses for cardioembolic stroke
15 (N=2,149), large-artery atherosclerosis (N=980), small-vessel disease (618), undetermined (N=1,926) and
16 other (N=222). No significant loci were found associated with specific stroke subtypes.

17 As both time to evaluation and time to treatment with tPA have been shown to be predictors of stroke
18 outcome, we conducted sensitivity analyses for subjects that had available information regarding
19 elapsed time to evaluation (n=4,477) and elapsed time to tPA (n=2,312). In both instances, the results of
20 the joint GWAS with and without the time variable of interest demonstrated highly correlated beta and
21 p-values and did not reveal any additional potential loci associated with Δ NIHSS.

22 **Functional annotation**

23 We annotated all the variants with suggestive associations (LBF>4) with ANNOVAR²⁸ and SnpEff²⁹ to
24 identify the nearest gene and to determine if any variant is predicted to change protein sequence (non-
25 synonymous variants) or could affect expression. We also confirmed if any of the SNPs were possible
26 regulatory elements or DNA features using RegulomeDB.³⁰

27 DEPICT³¹ and FUMA³² were used to perform gene ontology and pathways analyses. We also leveraged
28 brain single nuclei RNA expression data (<http://ngi.pub/snuclRNA-seq/>)²¹, to determine if the gene
29 expression of the genes located in each identified loci was expressed in brain. For the ones expressed in
30 brain, we also investigated if they were expressed in any specific brain cells (Fig. 1B). Finally, we
31 accessed blood RNA expression data taken at different times after stroke onset (3h, 5h and 24h) from

1 the CLEAR trial³³ (NCT00250991 at www.Clinical-Trials.gov) to test if the expression of genes located in
2 the identified loci were associated with NIHSS or Δ NIHSS (NIHSS_{5h}-NIHSS_{24h}). We extracted the
3 correlation between Δ NIHSS and gene expression (measured using Affymetrix U133 Plus 2.0 array).³⁴

4 **eQTL mapping Mendelian randomization and colocalization**

5 To identify the most likely functional gene, we accessed available expression quantitative trait (eQTL)
6 datasets: the Genotype-Tissue Expression (GTEx) Project V8 (accessed on 12/09/2021), the Brain eQTL
7 Almanac (Braineac) and an in-house dataset that includes brain expression data for 613 brains³⁵. We
8 used the summary-data-based Mendelian Randomization (SMR)³⁶ and colocalization³⁷ to test for
9 pleiotropic association between the expression level of a gene and a complex trait to evaluate if the
10 effect size of a genetic variant on the phenotype is mediated by gene expression (Fig. 1B). We tested
11 GWAS-significant and -suggestive loci from the Δ NIHSS analysis in two datasets: selected GTEx tissues
12 (brain anterior cortex, cerebellum, brain cerebellar hemisphere, substantia nigra, hippocampus, frontal
13 cortex, and putamen) and the Westra *et al.* dataset³⁸ derived from whole blood. Both SMR and
14 colocalization require effect sizes and the respective standard error to test the causal relationship, but
15 MANTRA does not provide effect sizes. As a consequence, we used the summary statistics from the joint
16 analysis for all populations to perform these analyses that are correlated with the results from MANTRA
17 ($r=-0.57$; $p<1.07\times 10^{-05}$ – data not shown). To complement the Mendelian randomization analyses with
18 the posterior probability of a variant being causal in both GWAS and eQTL studies accounting for the
19 genetic heterogeneity and *linkage disequilibrium* (LD), we used eCAVIAR³⁹ which will consider several
20 variants within the GWAS significant loci to perform the test.

21 **Genetic correlation**

22 We examined similarities in the genetic architectures of stroke early outcomes (Δ NIHSS) and stroke
23 risk¹⁵ using PRSice⁴⁰, LDSC⁴¹ and GNOVA⁴² (Fig. 1B). Briefly, PRSice calculates polygenic risk scores at
24 different p value thresholds by weighting each SNP by their effect size estimates. SNPs present in one
25 dataset, ambiguous SNPs (A/T or C/G) and all SNPs in LD are removed prior to polygenic risk score
26 calculation. LDSC and GNOVA estimate the genetic covariance and the variant-based heritability for two
27 sets of summary statistics, each one corresponding to one trait of interest. These two parameters are
28 used to calculate the genetic correlation and covariance respectively between the two traits. We limited
29 our comparisons to the non-Hispanic White population to keep the population genetically homogeneous

1 and use the 1000 Genomes European population-derived reference dataset. We calculated the genetic
2 correlation between the European ischemic stroke summary statistics of the MEGASTROKE¹⁵ study and
3 the non-Hispanic Whites meta-analysis summary statistics from the GENISIS study. We also determined
4 if traits related to cardiovascular and general health (age at death⁴³, lipid levels⁴⁴ and body mass index
5 (BMI)⁴⁵) are genetically correlated to Δ NIHSS.

6 **Data availability**

7 Summary statistics of the GENISIS dataset used for these analyses along with individual data for the full
8 GENISIS dataset will be uploaded to dbGAP titled: “*Genetics of Early Neurological Instability After*
9 *Ischemic Stroke (GENISIS)*”.

10 **Results**

11 The GENISIS study recruited 5,876 acute ischemic stroke patients from seven countries (Spain, Finland,
12 Poland, United States, Costa Rica, Mexico and Korea). The mean patient age was 73 years; 45% of the
13 patients were females, 54% were treated with tPA, 20% had a prior history of stroke. No significant
14 differences in age or sex were found across sites. The distribution of TOAST classification of stroke
15 etiology was also similar across sites. Significant differences were observed in baseline NIHSS and tPA
16 treatment rates, likely due to differences in practices across the sites (Table 1).⁶ Δ NIHSS approximated a
17 normal distribution, similar to that of each of the ethnic groups (non-Hispanic whites, Hispanics, African
18 descent, and Asians) (Supplementary Fig. 1).

19 **Identification of novel loci associated with stroke early outcomes**

20 We performed single variant analyses for each individual cohort separately; then we combined cohorts
21 with similar ethnic backgrounds; finally, we performed a multi-ancestry meta-analysis with the four
22 ethnic groups available in this study (Non-Hispanic Whites, Hispanics, Asians and African Americans)
23 (Fig. 1A). We identified eight GWAS significant loci (Fig. 2A and Table 2) associated with Δ NIHSS.

24 Three independent loci were identified in chr2. The first locus, tagged by rs58763243 ($MAF_6=0.07$;
25 $LBF=6.58$), was located in a region comprised by several long non-coding RNAs and microRNAs
26 (Supplementary Fig. 2E). For this locus, all of the populations contributed to the association with
27 negative betas, indicating that the minor allele was associated with lower (or more negative) Δ NIHSS. In

1 addition this locus reached genome-wide significance in the US AfA population and was nominally
2 significant in the Finnish population (Supplementary Fig. 2F).

3 The second locus, rs13403787 ($MAF_A=0.16$; $LBF=5.13$), was also located on chr2 in a region with more
4 than 20 genes (Supplementary Fig. 2G). The minor allele was associated with higher (or more positive)
5 $\Delta NIHSS$ in all cohorts (Supplementary Fig. 2H). The last genome wide significant locus in chr2 was
6 rs16838349 ($MAF_A=0.07$; $LBF=5.41$), located in a region that includes *ADAM23*, *CREB1*, *DYTN*, *NRP2*,
7 *MDH1B*, among many others (Fig. 2C). The signal is driven by the Non-Hispanic Whites (meta-analysis
8 $p=8.74 \times 10^{-06}$), but virtually all ethnic groups contributed to this association, as the directionality was
9 consistent across Hispanic, non-Hispanic White and AfA ethnic groups (Fig. 2D and Supplementary Table
10 1). However, the SNPs in this locus were monomorphic in the Asian population. We did not observe any
11 significant correlation between $\Delta NIHSS$ with the genotype in this locus ($R^2=0.063$, $p=0.09$;
12 Supplementary Table 2).

13 Five additional loci were identified outside chr2. Two independent loci in chr1. Both, rs1451040
14 ($MAF_T=0.16$, $LBF=6.56$) and rs9660272 ($MAF_T=0.16$, $LBF=7.64$) were in gene rich regions (Supplementary
15 Fig. 2A and 2C). These two loci are highly significant in the Latino populations (Mexico and Costa Rica),
16 with large negative effect sizes (Supplementary Fig. 2B and 2D). However, they are not significant in any
17 of the other populations, except for rs1451040, that is nominally significant in the Korean population,
18 but has the opposite direction of effect. The locus identified on chr5 is located on a region containing
19 nine genes (*LOC101927134*, *GRIA1*, *FAM114A2*, *SAP30L*, *SAP30L-AS1*, *MFAP3*, *GALNT10*, *HAND1* and
20 *MIR3141*; Fig. 2F). The minor allele for the top hit in this locus, rs17115057 ($MAF_T=0.06$; $LBF=5.19$) was
21 associated with greater (more positive) $\Delta NIHSS$ across most of the cohorts, and was significant in the
22 Spanish ($p=1.35 \times 10^{-07}$) and Finnish cohorts ($p=0.03$; Fig. 2G). Another locus on chr7, tagged by the
23 variant rs10807797 ($MAF_G=0.42$; $LBF=5.97$), and located in a gene rich region with 15 genes, including
24 *TWISTNB*, *MACC1*, *TMEM196*, *ABCB5*, *RPL23P8* (Supplementary Fig. 2I). This locus is tightly
25 encompassed by two recombination sites. The top signal was significant or suggestive in all populations
26 except the Polish and Mexican cohorts. Consistently, the direction of effect was the same in all cohorts
27 except the Mexican cohort (Supplementary Fig. 2J, Supplementary Table 1). Finally, we identified a locus
28 on chr13, tagged by rs9545725 located in a gene desert. No genes were identified in this locus
29 (Supplementary Fig. 2K). The variants in the region were significant in the Latino cohorts but the
30 direction of effect was not consistent across all cohorts (Supplementary Fig. 2L). Moreover, the MAF for
31 these variants ranged between 1% in the Korean population to 15% in the African American and Spanish
32 populations, suggesting that the region is very polymorphic depending on ethnicity. Thus, even though
33 the locus is important for $\Delta NIHSS$, it is possible that it is not the causal variant.

34 **Genetic contribution to early outcomes after ischemic stroke**

35 We used GCTA to quantify the phenotypic variance explained by common SNPs. Because GCTA exploits

1 LD patterns to calculate the explained variance, we restricted our analysis to non-Hispanic Whites. Due
2 to founder effects present in the Finnish population, we also removed this cohort from the variance
3 calculation (Final N=4,573). GCTA revealed that common genetic variants explained 8.7% of the variance
4 of Δ NIHSS ($p=0.001$), confirming that genetic variants and genes are implicated on stroke outcomes.
5 Next we determine what proportion of the genetic component is explained by the GWAS signals.
6 The SNPs comprised within the 8 genome-wide significant loci, defined as 500 bp upstream or
7 downstream of the top signal, explained 1.8% of the total variance ($p=2.18 \times 10^{-04}$) of Δ NIHSS, or just
8 20.7% of the genetic component of Δ NIHSS. This suggests that there are additional loci associated with
9 Δ NIHSS yet to be discovered. Thus, studies with larger sample size and more statistical power are
10 needed to identify these additional loci.

11 **Functional annotation of the genome-wide significant loci**

12 Identifying the likely causal gene from each loci driving the association is a multi-step process (Fig. 1B).
13 We first annotated the suggestive variants ($LBF > 4$), but none of them were predicted to change the
14 protein sequence, comprise a regulatory element, or affect the chromatin architecture. Next, we
15 explored publicly available datasets to investigate if any of the SNPs with suggestive LBFs were eQTLs
16 (Supplementary Table 3). We performed gene-based analyses (Supplementary Table 4) and Mendelian
17 Randomization (MR) analyses to identify possible causal relationships between gene expression and
18 Δ NIHSS (Supplementary Table 5). Summary results can be found in Fig. 3.

19 Gene-based analyses using FUMA suggested that *DYTN* ($p=2.55 \times 10^{-04}$, $Z=3.47$) and *ADAM23* ($p=2.04 \times 10^{-03}$,
20 $Z=2.87$) were the genes driving the association at 2q33.3. Several variants in the *ADAM23* region were
21 strong eQTLs for this gene in multiple tissues, based on the GTEx data (esophageal mucosa: $p=2.00 \times 10^{-06}$;
22 Cultured Fibroblasts: $p=5.90 \times 10^{-05}$). MR analyses indicated that *ADAM23* ($p=0.04$) was the gene
23 driving the association in this locus. Human brain single nuclei RNA-seq data indicate that *ADAM23*
24 expression is enriched in neurons ($p < 2.20 \times 10^{-16}$); compared to all the other brain cell types. More
25 specifically, its expression is enriched in excitatory neurons (Fig. 2E).

26 Gene-based analyses using FUMA revealed that *GRIA1* located in 5q33.2 was the gene most likely driving
27 the association in that region ($p=0.03$, $Z=1.83$). However, Braineac identified several eQTLs for *GALNT10*
28 ($p=3.70 \times 10^{-04}$) in the occipital cortex, but *GALNT10* was less significant in the gene-based analysis
29 ($p=0.04$, $Z=1.79$). GTEx portal and the protein atlas reveals that *GRIA1* is mainly expressed in brain tissue.
30 While *GALNT10* is also expressed in the brain, it has higher expression in other tissues. The human brain
31 single nuclei RNA-seq data confirmed that both *GRIA1* and *GALNT10* are expressed in divergent brain

1 cell types (Fig. 2H and Supplementary 3A). *GRIA1* is highly-expressed in neurons compared to other cell
2 types ($p < 2.20 \times 10^{-16}$), but not expressed in oligodendrocytes ($p < 2.20 \times 10^{-16}$) or astrocytes ($p < 2.20 \times 10^{-16}$).
3 In contrast, *GALNT10* is expressed in microglia, oligodendrocytes and astrocytes, but expression in
4 neurons is low ($p < 2.20 \times 10^{-16}$). *GRIA1* expression in peripheral blood was also nominally associated with
5 Δ NIHSS in the CLEAR trial dataset ($p = 0.002$, $r^2 = 0.22$).

6 Of the remaining six loci, we were able to map five (Supplementary Results). Briefly, eQTL analysis,
7 revealed that 1p21.1 was likely to be driven by *COL11A1* or *AMY2B*. Gene-based and MR analyses
8 suggested that the locus 1q42.2 was driven by *GNPAT*. No eQTLs were identified for 2p25.1, but gene-
9 based analyses suggested that the signal is likely driven by *AGPS* or *TTC30A*. Regarding 2p31.2, it is likely
10 driven by *DFNB59*, while 7p21.1 contains several eQTLs for *TWISTNB* and *ABCB5* (Supplementary
11 Results).

12 Pathway analyses

13 Gene ontology and pathway analyses using DEPICT and summary statistics for Δ NIHSS revealed
14 consistent suggestive associations with functions relating to the brain and central nervous system. The
15 top tissue enrichment from DEPICT identified the cardiovascular system (1.8×10^3) and the central
16 nervous system ($p = 2.0 \times 10^{-03}$), including the brain ($p = 0.01$) and some brain regions: occipital lobe
17 ($p = 2.00 \times 10^{-03}$), cerebral cortex ($p = 4.80 \times 10^{-03}$), and temporal lobe ($p = 6.33 \times 10^{-03}$; Supplementary Table 6).
18 The most significant pathways in the gene-set enrichment were the regulation of the heart contraction
19 ($p = 5.80 \times 10^{-06}$), the sodium ion transmembrane transport ($p = 6.27 \times 10^{-06}$), the circulatory system process
20 ($p = 6.39 \times 10^{-06}$) learning or memory ($p = 7.11 \times 10^{-06}$) and abnormal CNS synaptic transmission ($p = 2.88 \times 10^{-05}$;
21 Supplementary Table 7). Several genome-wide significant candidate genes fell within these networks,
22 of special interest, *GRIA1* (5q33.2, LBF=5.19) in the sodium ion transmembrane transport, which adds
23 evidence to the involvement of *GRIA1* in Δ NIHSS. MAGMA gene-set analyses did not reveal any enriched
24 gene set associated with Δ NIHSS (Supplementary Table 8).

25 Unique genetic architecture of early outcomes after stroke

26 We examined the genetic architecture of Δ NIHSS for shared genetic variation with other cardiovascular
27 and aging-related traits, including stroke risk, age at death, plasma lipid levels and body mass index
28 using PRSice (Supplementary Table 9 and Supplementary Fig. 4), LDSC (Supplementary Table 10) and
29 GNOVA (Supplementary Table 11). Although the p value for PRSice was significant in the comparison of

1 stroke risk and Δ NIHSS, the amount of variance explained was very small ($R^2=0.009$). Additionally, this
2 finding was not supported by LDSC or GNOVA analyses, suggesting that there is no genetic overlap, as
3 reported in a previous work.¹⁸ Similarly, no overlap with age at death, lipid levels, or BMI was identified
4 by LDSC or GNOVA. Even though PRSice found significant correlations with several stroke risk factors,
5 HDL levels ($p=0.01$), TG levels ($p=8.97\times 10^{-04}$), total cholesterol ($p=0.02$), body mass index ($p=1.89\times 10^{-06}$)
6 and age at death ($p=0.01$), the amount of variance explained were all below 0.5%, suggesting that the
7 overlap is minimal. LDSC was unable to calculate the heritability estimate for Δ NIHSS. GNOVA, was
8 successful at estimating the heritability for Δ NIHSS, but could not calculate the genetic correlation
9 estimate. Several of the heritability estimates for Δ NIHSS for overlap were negative, likely due to the low
10 number of variants included in the analyses. Because both GNOVA and LDSC require larger sample sizes,
11 the results of these analyses were inconclusive.

12 Discussion

13 The first 24 hours after stroke onset is a period of great neurological instability, which may
14 reflect brain tissue at risk for infarction but with the potential for salvageability.^{4,46-48} Not only is
15 early neurological change (as reflected by Δ NIHSS) common, it is influenced by known
16 mechanisms involved in early deterioration/improvement, and has a strong influence on long-
17 term functional outcome.⁶ Here, we performed a GWAS using Δ NIHSS as a quantitative
18 phenotype in 5,876 acute ischemic stroke patients. We found that Δ NIHSS is heritable: common
19 SNPs account for 8.7% of its variance. We have found eight genome-wide significant loci that
20 are related to Δ NIHSS. However, they explain only 1.8% of the variance, indicating that 6.9% of
21 the variance is explained by genes below the genome-wide significant threshold. Through
22 functional annotation, we have linked each locus to specific genes, some of which are uniquely
23 expressed in the brain.

24 Of all the loci showing association with Δ NIHSS, functional annotation analyses strongly
25 suggests that *ADAM23* is the functional gene for the locus,2q33.3. *ADAM23* belongs to the
26 ADAM (a disintegrin and metalloproteinase) family of proteins, defined by a single-pass
27 transmembrane structure with a metallopeptidase domain (some inactive). This protein family is
28 involved in cell adhesion, migration, proteolysis and signaling.⁴⁹ *ADAM23* is a transmembrane
29 member without catalytic domain, and is involved in cell-cell and cell-matrix interactions.^{49,50}

1 Previous studies have shown that *ADAM23* is expressed in pre-synaptic membranes, linked by
2 the extracellular protein *LGII* to post-synaptic *ADAM22*.^{51,52} We found that *ADAM23* was
3 expressed primarily in excitatory neurons of the cerebral cortex, based on our human brain
4 single-nuclei transcriptomics dataset²¹, and confirmed by the Human Transcriptomic Cell Types
5 dataset from the Allen Brain Map.⁵³ Several lines of evidence suggest that *ADAM23* is
6 important for pathological synaptic excitability: 1) *adam23* is a common risk gene for canine
7 idiopathic epilepsy;⁵⁴⁻⁵⁶ 2) mutations in its binding partner, *LGII*, cause the neurological
8 syndrome, ADPEAF (autosomal dominant partial epilepsy with auditory features)⁵⁷; 3)
9 autoimmunity against *LGII* (as seen in limbic encephalitis) results in seizures and
10 encephalopathy.⁵⁸

11 Indeed, *ADAM23* is also known to be a binding partner (via *ADAM22* and *PSD95*) of the
12 protein product of another one of our genome-wide significant associated genes, *GRIA1*, which
13 encodes for the α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor subunit 1
14 (AMPA1).⁵² It has long been known that AMPA receptors, along with other glutamate
15 receptors, are mediators of excitotoxic neuronal death, hypothesized to play an important role in
16 ischemic brain injury.^{59,60} The failure of numerous older clinical trials examining the efficacy of
17 anti-excitotoxic drugs has cast doubt on the relevance of excitotoxicity in human acute ischemic
18 stroke, although questions about the quality of these early clinical trials have been raised.⁶¹⁻⁶⁴

19 Thus, the association between *ADAM23* and *GRIA1* with Δ NIHSS provides the first genetic
20 evidence that excitotoxicity may contribute to ischemic brain injury in humans.

21 The plausible roles that *ADAM23* and *GRIA1* play in acute brain ischemia mechanisms lend
22 support to the idea that GWAS using Δ NIHSS as a quantitative phenotype can identify novel
23 mechanisms and potential drug targets to mitigate neurological deterioration or enhance early
24 improvement after stroke. From the CLEAR dataset,^{33,34} expression levels of *GRIA1* in
25 peripheral blood of ischemic stroke patients was associated with Δ NIHSS between 5h and 24h
26 post stroke onset, supporting a link between increased expression of *GRIA1* and improved
27 outcomes. In addition to the two genes discussed above, our GWAS identified six other loci—the
28 functional genes remain to be identified. Acute ischemic stroke patients are extremely well-
29 phenotyped, as part of standard of care, with both clinical assessments and
30 structural/physiological imaging. Thus, there is great potential for additional quantitative
31 phenotypes to expand understanding of the genetic architecture of acute ischemic stroke,

1 promising to identify novel mechanisms and drug targets. Larger and more comprehensive
2 genetic studies of acute ischemic stroke are needed.

3 There are several limitations to this study. GENISIS enrolled a heterogeneous group of stroke
4 patients without regard to underlying etiology, stroke localization and genetic and environmental
5 background. Although we have previously demonstrated that etiology (TOAST criteria) has little
6 influence on Δ NIHSS, it is likely that mechanisms involved in neurological instability may
7 depend on etiology. Stroke localization may also be an important determinant of mechanisms
8 involved in neurological instability. For example, mechanisms in cortical strokes may differ from
9 those in subcortical or brainstem strokes. Furthermore, specific medication information (such as
10 type of anticoagulation medication, if being used for secondary prevention at the time of stroke)
11 were not collected, and therefore cannot be accounted for. False positive findings due to the
12 characteristics of the population is possible, but by using MANTRA we were able to correct by
13 population heterogeneity. Future studies might aim to enroll a more homogeneous cohort of
14 stroke patients to increase power to discover more genetic variants that associate with
15 neurological instability. Finally, most of the patients in GENISIS were enrolled prior to the
16 thrombectomy treatment era, and patients that underwent thrombectomy were excluded from the
17 study to reduce heterogeneity. As a result, genetic interactions with reperfusion are largely
18 unexplored.

19 **Acknowledgements**

20 We would like to thank the patients and their families for making possible all the genetic studies
21 included in this manuscript. We also thank the MEGASTROKE consortium for access to the data (see full
22 list of MEGASTROKE authors in supplementary data), the Genotype-Tissue Expression (GTEx) Project
23 (supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by
24 NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS) and the Brain eQTL Almanac (Braineac) resource to access
25 the UK Brain Expression Consortium (UKBEC) dataset.

26 **Funding**

27 This work was supported by grants from the Emergency Medicine Foundation Career Development

1 Grant; AHA Mentored Clinical & Population Research Award (14CRP18860027); NIH/NINDS-R01-
2 NS085419 (CC, JML); NIH/NINDS-R37-NS107230, NIH/NINDS U24-NS107230 (JML); NIH/NINDS-K23-
3 NS099487 (LH); NIH/NIA-K99-AG062723 (LI),; Barnes-Jewish Hospital Foundation (JML); Biogen (CC,
4 JML); Bright Focus Foundation, US Department of Defense, Helsinki University Central Hospital; Finnish
5 Medical Foundation; Finland government subsidiary funds; Spanish Ministry of Science and Innovation;
6 Instituto de Salud Carlos III (grants “Registro BASICMAR” Funding for Research in Health (PI051737),
7 “GWALA project” from Fondos de Investigación Sanitaria ISC III (PI10/02064, PI12/01238 and
8 PI15/00451), JR18/00004); Fondos FEDER/EDRF Red de Investigación Cardiovascular (RD12/0042/0020);
9 Fundació la Marató TV3; Genestroke Consortium (76/C/2011); Recercaixa’13 (JJ086116). Tomás Sobrino
10 (CPII17/00027), Francisco Campos (CPII19/00020) and Israel Fernandez are supported by Miguel Servet
11 II Program from Instituto de Salud Carlos III and Fondos FEDER. Israel Fernandez is also supported by
12 Maestro project (PI18/01338) and Pre-test project (PMP15/00022) from Instituto de Salud Carlos III and
13 Fondos Feder, Agaur; and Epigenesis project from Marató TV3 Foundation. José Castillo, Joan
14 Montaner, Antonio Dávalos, Joan Martí-Fàbregas, Juan Arenillas and Israel Fernández, are supported by
15 Invictus plus Network (RD16/0019) from Instituto de Salud Carlos III and Fondos Feder. Fundação de
16 Amparo à Pesquisa do Estado de São Paulo (FAPESP-2013/07559-3) (ILC), Sigrid Juselius Foundation. The
17 MEGASTROKE project received funding from sources specified at
18 <http://www.megastroke.org/acknowledgments.html>. Boryana Stamova, Bradley Ander and Frank Sharp
19 are supported by NIH awards: NS097000, NS101718, NS075035, NS079153 and NS106950.

20 **Competing interests**

21 CC receives research support from: Biogen, Eisai, Alector and Parabon, and is a member of the advisory
22 board of ADx Healthcare and Vivid Genomics. JML receives research support from Biogen, and is a
23 consultant for Regenera. EAT and LCB are employed by Biogen. JFA has received speaker or consultant
24 honoraria from Bayer, Boehringer Ingelheim, Pfizer-BMS, Daiichi Sankyo, Amgen and Medtronic. TT
25 receives or has received research support from Bayer, Boehringer Ingelheim and Bristol Myers Squibb;
26 he is a member of advisory boards for Bayer, Boehringer Ingelheim, Bristol Myers Squibb and Portola
27 Pharmaceuticals; and he is granted international patents: new therapeutic uses (method to prevent
28 brain edema and reperfusion injury), and thrombolytic compositions (method to prevent post-
29 thrombolytic hemorrhage formation). The funders of the study had no role in the collection, analysis, or
30 interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

1 Supplementary material

2 Supplementary material is available at *Brain* online

3 References

- 4 1. Organization WH. The Global Burden of Disease: 2004 update. 2008;
- 5 2. Benjamin EJ, Virani SS, Callaway CW, et al. Heart Disease and Stroke Statistics-2018
- 6 Update: A Report From the American Heart Association. *Circulation*. Mar 20 2018;137(12):e67-
- 7 e492. doi:10.1161/CIR.0000000000000558
- 8 3. Farooq MU, Chaudhry AH, Amin K, Majid A. The WHO STEPwise Approach to Stroke
- 9 Surveillance. *Journal of the College of Physicians and Surgeons--Pakistan : JCPSP*. Oct
- 10 2008;18(10):665. doi:10.2008/JCPSP.665665
- 11 4. Saver JL, Altman H. Relationship between neurologic deficit severity and final functional
- 12 outcome shifts and strengthens during first hours after onset. *Stroke*. Jun 2012;43(6):1537-41.
- 13 doi:10.1161/STROKEAHA.111.636928
- 14 5. Tissue plasminogen activator for acute ischemic stroke. The National Institute of
- 15 Neurological Disorders and Stroke rt-PA Stroke Study Group. *N Engl J Med*. Dec 14
- 16 1995;333(24):1581-7. doi:10.1056/NEJM199512143332401
- 17 6. Heitsch L, Ibanez L, Carrera C, et al. Early change in NIH stroke scale after ischemic
- 18 stroke captures brain injury mechanisms and predicts 90-day outcome. *Stroke*. 2020;In Press
- 19 7. Barber PA, Davis SM, Infeld B, et al. Spontaneous reperfusion after ischemic stroke is
- 20 associated with improved outcome. *Stroke*. Dec 1998;29(12):2522-8.
- 21 8. Jorgensen HS, Sperling B, Nakayama H, Raaschou HO, Olsen TS. Spontaneous
- 22 reperfusion of cerebral infarcts in patients with acute stroke. Incidence, time course, and clinical
- 23 outcome in the Copenhagen Stroke Study. *Archives of neurology*. Sep 1994;51(9):865-73.
- 24 9. Bang OY, Saver JL, Buck BH, et al. Impact of collateral flow on tissue fate in acute
- 25 ischaemic stroke. *Journal of neurology, neurosurgery, and psychiatry*. Jun 2008;79(6):625-9.
- 26 doi:10.1136/jnnp.2007.132100
- 27 10. Christoforidis GA, Karakasis C, Mohammad Y, Caragine LP, Yang M, Slivka AP.
- 28 Predictors of hemorrhage following intra-arterial thrombolysis for acute ischemic stroke: the role
- 29 of pial collateral formation. *AJNR American journal of neuroradiology*. Jan 2009;30(1):165-70.
- 30 doi:10.3174/ajnr.A1276
- 31 11. Ali LK, Saver JL. The ischemic stroke patient who worsens: new assessment and
- 32 management approaches. *Reviews in neurological diseases*. Spring 2007;4(2):85-91.
- 33 12. Dreier JP. The role of spreading depression, spreading depolarization and spreading
- 34 ischemia in neurological disease. *Nat Med*. Apr 2011;17(4):439-47. doi:10.1038/nm.2333
- 35 13. Paciaroni M, Agnelli G, Corea F, et al. Early hemorrhagic transformation of brain
- 36 infarction: rate, predictive factors, and influence on clinical outcome: results of a prospective
- 37 multicenter study. *Stroke*. Aug 2008;39(8):2249-56. doi:10.1161/STROKEAHA.107.510321
- 38 14. Torres-Aguila NP, Carrera C, Giese AK, et al. Genome-Wide Association Study of White
- 39 Blood Cell Counts in Patients With Ischemic Stroke. *Stroke*. Dec 2019;50(12):3618-3621.
- 40 doi:10.1161/STROKEAHA.119.026593
- 41 15. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of

- 1 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet.* Apr
2 2018;50(4):524-537. doi:10.1038/s41588-018-0058-3
- 3 16. Soderholm M, Pedersen A, Lorentzen E, et al. Genome-wide association meta-analysis of
4 functional outcome after ischemic stroke. *Neurology.* Mar 19 2019;92(12):e1271-e1283.
5 doi:10.1212/WNL.00000000000007138
- 6 17. Mola-Caminal M, Carrera C, Soriano-Tarraga C, et al. PATJ Low Frequency Variants
7 Are Associated With Worse Ischemic Stroke Functional Outcome. *Circ Res.* Jan 4
8 2019;124(1):114-120. doi:10.1161/CIRCRESAHA.118.313533
- 9 18. Ibanez L, Heitsch L, Dube U, et al. Overlap in the Genetic Architecture of Stroke Risk,
10 Early Neurological Changes, and Cardiovascular Risk Factors. *Stroke.* Jun 2019;50(6):1339-
11 1345. doi:10.1161/STROKEAHA.118.023097
- 12 19. Morris AP. Transethnic meta-analysis of genomewide association studies. *Genetic
13 epidemiology.* Dec 2011;35(8):809-22. doi:10.1002/gepi.20630
- 14 20. Wang X, Chua HX, Chen P, et al. Comparing methods for performing trans-ethnic meta-
15 analysis of genome-wide association studies. *Hum Mol Genet.* Jun 1 2013;22(11):2303-11.
16 doi:10.1093/hmg/ddt064
- 17 21. Del-Aguila JL, Li Z, Dube U, et al. A single-nuclei RNA sequencing study of Mendelian
18 and sporadic AD in the human brain. *Alzheimer's research & therapy.* Aug 9 2019;11(1):71.
19 doi:10.1186/s13195-019-0524-x
- 20 22. Delaneau O, Coulonges C, Zagury JF. Shape-IT: new rapid and accurate algorithm for
21 haplotype inference. *BMC bioinformatics.* Dec 16 2008;9:540. doi:10.1186/1471-2105-9-540
- 22 23. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3.*
23 Nov 2011;1(6):457-70. doi:10.1534/g3.111.001198
- 24 24. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex
25 trait analysis. *Am J Hum Genet.* Jan 7 2011;88(1):76-82. doi:10.1016/j.ajhg.2010.11.011
- 26 25. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation
27 PLINK: rising to the challenge of larger and richer datasets. *GigaScience.* 2015;4:7.
28 doi:10.1186/s13742-015-0047-8
- 29 26. Glymour MM, Weuve J, Berkman LF, Kawachi I, Robins JM. When is baseline
30 adjustment useful in analyses of change? An example with education and cognitive change. *Am J
31 Epidemiol.* Aug 01 2005;162(3):267-78. doi:10.1093/aje/kwi187
- 32 27. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide
33 association scans. *Bioinformatics.* Sep 01 2010;26(17):2190-1.
34 doi:10.1093/bioinformatics/btq340
- 35 28. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and
36 wANNOVAR. *Nature protocols.* Oct 2015;10(10):1556-66. doi:10.1038/nprot.2015.105
- 37 29. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the
38 effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila
39 melanogaster* strain w1118; iso-2; iso-3. *Fly.* Apr-Jun 2012;6(2):80-92. doi:10.4161/fly.19695
- 40 30. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal
41 genomes using RegulomeDB. *Genome research.* Sep 2012;22(9):1790-7.
42 doi:10.1101/gr.137323.112
- 43 31. Pers TH, Karjalainen JM, Chan Y, et al. Biological interpretation of genome-wide
44 association studies using predicted gene functions. *Nature communications.* Jan 19 2015;6:5890.
45 doi:10.1038/ncomms6890
- 46 32. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and

- 1 annotation of genetic associations with FUMA. *Nature communications*. Nov 28 2017;8(1):1826.
2 doi:10.1038/s41467-017-01261-5
- 3 33. Pancioli AM, Broderick J, Brott T, et al. The combined approach to lysis utilizing
4 eptifibatid and rt-PA in acute ischemic stroke: the CLEAR stroke trial. *Stroke*. Dec
5 2008;39(12):3268-76. doi:10.1161/STROKEAHA.108.517656
- 6 34. Stamova B, Xu H, Jickling G, et al. Gene expression profiling of blood for the prediction
7 of ischemic stroke. *Stroke*. Oct 2010;41(10):2171-7. doi:10.1161/STROKEAHA.110.588335
- 8 35. Li Z, Del-Aguila JL, Dube U, et al. Genetic variants associated with Alzheimer's disease
9 confer different cerebral cortex cell-type population structure. *Genome Med*. Jun 8
10 2018;10(1):43. doi:10.1186/s13073-018-0551-4
- 11 36. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies
12 predicts complex trait gene targets. *Nat Genet*. May 2016;48(5):481-7. doi:10.1038/ng.3538
- 13 37. Plagnol V, Smyth DJ, Todd JA, Clayton DG. Statistical independence of the colocalized
14 association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13.
15 *Biostatistics*. Apr 2009;10(2):327-34. doi:10.1093/biostatistics/kxn039
- 16 38. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative
17 drivers of known disease associations. *Nat Genet*. Oct 2013;45(10):1238-1243.
18 doi:10.1038/ng.2756
- 19 39. Hormozdiari F, van de Bunt M, Segre AV, et al. Colocalization of GWAS and eQTL
20 Signals Detects Target Genes. *Am J Hum Genet*. Dec 1 2016;99(6):1245-1260.
21 doi:10.1016/j.ajhg.2016.10.003
- 22 40. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software.
23 *Bioinformatics*. May 01 2015;31(9):1466-8. doi:10.1093/bioinformatics/btu848
- 24 41. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes
25 confounding from polygenicity in genome-wide association studies. *Nat Genet*. Mar
26 2015;47(3):291-5. doi:10.1038/ng.3211
- 27 42. Lu Q, Li B, Ou D, et al. A Powerful Approach to Estimating Annotation-Stratified
28 Genetic Covariance via GWAS Summary Statistics. *Am J Hum Genet*. Dec 7 2017;101(6):939-
29 964. doi:10.1016/j.ajhg.2017.11.001
- 30 43. Pilling LC, Kuo CL, Sicinski K, et al. Human longevity: 25 genetic loci associated in
31 389,166 UK biobank participants. *Aging (Albany NY)*. Dec 6 2017;9(12):2504-2520.
32 doi:10.18632/aging.101334
- 33 44. Liu DJ, Peloso GM, Yu H, et al. Exome-wide association study of plasma lipids in
34 >300,000 individuals. *Nat Genet*. Dec 2017;49(12):1758-1766. doi:10.1038/ng.3977
- 35 45. Turcot V, Lu Y, Highland HM, et al. Protein-altering variants associated with body mass
36 index implicate pathways that control energy intake and expenditure in obesity. *Nat Genet*. Jan
37 2018;50(1):26-41. doi:10.1038/s41588-017-0011-x
- 38 46. Adams HP, Jr., Davis PH, Leira EC, et al. Baseline NIH Stroke Scale score strongly
39 predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment
40 (TOAST). *Neurology*. Jul 13 1999;53(1):126-31.
- 41 47. Yeo LL, Paliwal P, Teoh HL, et al. Early and continuous neurologic improvements after
42 intravenous thrombolysis are strong predictors of favorable long-term outcomes in acute
43 ischemic stroke. *J Stroke Cerebrovasc Dis*. Nov 2013;22(8):e590-6.
44 doi:10.1016/j.jstrokecerebrovasdis.2013.07.024
- 45 48. Sajobi TT, Menon BK, Wang M, et al. Early Trajectory of Stroke Severity Predicts Long-
46 Term Functional Outcomes in Ischemic Stroke Subjects: Results From the ESCAPE Trial

- 1 (Endovascular Treatment for Small Core and Anterior Circulation Proximal Occlusion With
2 Emphasis on Minimizing CT to Recanalization Times). *Stroke; a journal of cerebral circulation*.
3 Jan 2017;48(1):105-110. doi:10.1161/STROKEAHA.116.014456
- 4 49. Goldsmith AP, Gossage SJ, French-Constant C. ADAM23 is a cell-surface glycoprotein
5 expressed by central nervous system neurons. *Journal of neuroscience research*. Dec 1
6 2004;78(5):647-58. doi:10.1002/jnr.20320
- 7 50. Mitchell KJ, Pinson KI, Kelly OG, et al. Functional analysis of secreted and
8 transmembrane proteins critical to mouse development. *Nat Genet*. Jul 2001;28(3):241-9.
9 doi:10.1038/90074
- 10 51. Sagane K, Ishihama Y, Sugimoto H. LGI1 and LGI4 bind to ADAM22, ADAM23 and
11 ADAM11. *International journal of biological sciences*. 2008;4(6):387-96.
12 doi:10.7150/ijbs.4.387
- 13 52. Lovero KL, Fukata Y, Granger AJ, Fukata M, Nicoll RA. The LGI1-ADAM22 protein
14 complex directs synapse maturation through regulation of PSD-95 function. *Proc Natl Acad Sci*
15 *U S A*. Jul 28 2015;112(30):E4129-37. doi:10.1073/pnas.1511910112
- 16 53. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al. An anatomically comprehensive
17 atlas of the adult human brain transcriptome. *Nature*. Sep 20 2012;489(7416):391-399.
18 doi:10.1038/nature11405
- 19 54. Koskinen LL, Seppala EH, Belanger JM, et al. Identification of a common risk haplotype
20 for canine idiopathic epilepsy in the ADAM23 gene. *BMC genomics*. Jun 18 2015;16:465.
21 doi:10.1186/s12864-015-1651-9
- 22 55. Koskinen LL, Seppala EH, Weissl J, et al. ADAM23 is a common risk gene for canine
23 idiopathic epilepsy. *BMC genetics*. Jan 31 2017;18(1):8. doi:10.1186/s12863-017-0478-6
- 24 56. Seppala EH, Koskinen LL, Gullov CH, et al. Identification of a novel idiopathic epilepsy
25 locus in Belgian Shepherd dogs. *PLoS One*. 2012;7(3):e33549.
26 doi:10.1371/journal.pone.0033549
- 27 57. Berkovic SF, Izzillo P, McMahon JM, et al. LGI1 mutations in temporal lobe epilepsies.
28 *Neurology*. Apr 13 2004;62(7):1115-9. doi:10.1212/01.wnl.0000118213.94650.81
- 29 58. van Sonderen A, Schreurs MW, Wirtz PW, Sillevs Smitt PA, Titulaer MJ. From VGKC
30 to LGI1 and Caspr2 encephalitis: The evolution of a disease entity over time. *Autoimmunity*
31 *reviews*. Oct 2016;15(10):970-4. doi:10.1016/j.autrev.2016.07.018
- 32 59. Lee JM, Zipfel GJ, Choi DW. The changing landscape of ischaemic brain injury
33 mechanisms. *Nature*. Jun 24 1999;399(6738 Suppl):A7-14. doi:10.1038/399a007
- 34 60. Palmer CL, Cotton L, Henley JM. The molecular pharmacology and cell biology of
35 alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors. *Pharmacol Rev*. Jun
36 2005;57(2):253-77. doi:10.1124/pr.57.2.7
- 37 61. Chamorro A, Dirnagl U, Urra X, Planas AM. Neuroprotection in acute stroke: targeting
38 excitotoxicity, oxidative and nitrosative stress, and inflammation. *The Lancet Neurology*. Jul
39 2016;15(8):869-881. doi:10.1016/S1474-4422(16)00114-9
- 40 62. Wahlgren NG, Ahmed N. Neuroprotection in cerebral ischaemia: facts and fancies--the
41 need for new approaches. *Cerebrovasc Dis*. 2004;17 Suppl 1:153-66. doi:10.1159/000074808
- 42 63. Neuhaus AA, Couch Y, Hadley G, Buchan AM. Neuroprotection in stroke: the
43 importance of collaboration and reproducibility. *Brain : a journal of neurology*. Aug 1
44 2017;140(8):2079-2092. doi:10.1093/brain/awx126
- 45 64. Castillo J, Davalos A, Noya M. Progression of ischaemic stroke and excitotoxic
46 aminoacids. *Lancet*. Jan 11 1997;349(9045):79-83. doi:10.1016/S0140-6736(96)04453-4

1 **Figure legends**

2 **Figure 1. Study design.** Summarized description of the multi-step approach used to account for the
3 genetic heterogeneity intrinsic to the multi-ancestry nature of the GENISIS study (A). We performed
4 single variant analysis in each of the participating countries separately. Then we meta-analyzes all the
5 non-Hispanic whites (blue) and Hispanic (green) ethnicities. Finally we analyzed the non-Hispanic whites,
6 Hispanics, Korea (orange) and US participants with African Descent (US AfA – yellow) using a Bayesian
7 model. The variants with genome-wide significant or suggestive results were annotated using sequential
8 steps to elucidate the gene driving the association (B). We performed gene-based and pathway
9 analyses, we collected the information available in publicly available datasets and we performed
10 Mendelian randomization. We also performed genetic architecture overlap tests to examine overlap
11 with known genetic risk factors.

12 **Figure 2. Association and annotation results. A.** Manhattan plot shows Log Bayes factor (LBF) values
13 from the multi-ancestry meta-analysis in each genomic location. The red line indicates the GWAS
14 significant threshold (LBF>5) and the blue line the GWAS suggestive threshold (LBF>4). The genome-
15 wide significant loci are highlighted. Local Manhattan plots are shown for rs16838349 (C) and
16 rs17115057 (F) along with the corresponding forest plots (D and G), showing the contribution of each
17 population to the overall signal. As part of the functional gene mapping, we accessed an in-house single
18 nuclei dataset (B) to describe the expression patterns in human brain cortical cell populations of the
19 driving genes identified for rs16838349, *ADAM23* (E) and rs17115057, *GRIA1* (H).

20 **Figure 3. Gene prioritizing summary.** Summary table showing the seven genome-wide significant loci
21 from the multi-ancestry analysis (first column), the total number of genes identified in each of the locus
22 (second column) and gene name for genes for which we have found some kind of evidence (third
23 column). We have included the results from the gene-based analyses, the presence of any eQTL in GTEx
24 portal or Braineac for any of the genome-wide or suggestive variants, if the gene is differentially
25 expressed in any brain region according to the snRNAseq data and the results from Mendelian
26 randomization using Westra dataset (whole blood) or GTEx portal (all tissues). Black dots indicate that
27 the gene was not found, red is that it was found but was not significant, yellow it was moderately
28 significant ($0.05 < p < 1 \times 10^{-3}$) and green shows a significant association ($p < 1 \times 10^{-3}$).

29

30

	Spain (N=3,419)	Finland (N=490)	Poland (N=356)	US-EuA (N=798)	Costa Rica (N=141)	Mexico (N=63)	Korea (N=285)	US-AfA (N=324)	GENISIS (N=5,876)
Age (years) ^a	76.0 (66.0–83.0)	68.0 (58.0–76.0)	71.0 (63.0–80.0)	70.0 (60.0–79.0)	67.0 (56.0–78.0)	67.0 (50.5–75.5)	69.0 (58.0–78.0)	63.0 (54.0–74.3)	73.0 (62.0–81.0)
Sex (Females, %)	1,554 (45.5%)	193 (39.4%)	159 (44.7%)	337 (42.2%)	57 (40.4%)	28 (44.4%)	91.9 (31.9%)	169 (52.2%)	2,588 (44.0%)
Baseline NIHSS ^a	10.0 (5.0–17.0)	5.0 (2.0–9.0)	6.0 (3.0–12.0)	6.0 (3.0–8.2)	13.0 (9.0–18.0)	11.0 (6.0–14.5)	4.0 (2.0–8.0)	7.0 (4.0–12.0)	8.90 (4.0–15.0)
tPA Treatment (%)	48.32%	48.37%	59.55%	73.81%	100%	46.03%	28.07%	75.62%	54.20%
ΔNIHSS ^b	2.77 ± 5.42	2.34 ± 5.68	2.12 ± 3.40	2.11 ± 5.98	6.00 ± 7.14	3.40 ± 4.90	1.17 ± 3.40	2.37 ± 6.29	2.56 ± 5.52
TOAST Classification ^c (%):									
Cardioembolic	38.32%	41.63%	29.21%	37.72%	21.28%	23.81%	30.53%	29.01%	36.50%
Large Artery	17.17%	16.53%	12.36%	13.03%	39.01%	25.40%	24.56%	8.64%	16.76%
Small Vessel Disease	9.15%	6.73%	3.09%	13.16%	12.77%	14.29%	17.89%	16.98%	10.14%
Other	2.46%	8.16%	2.81%	3.13%	2.13%	15.87%	13.68%	3.09%	3.76%
Undetermined	32.90%	26.94%	52.53%	32.96%	24.11%	20.63%	13.33%	42.28%	32.83%

1 **Table I Demographic Characteristics of the GENISIS cohort by country**

2 ^aValues are expressed as median (95% confidence interval).

3 ^bValues are expressed as mean ± Standard Deviation

4 ^cTOAST classification criteria¹

5

1
2
3
4
5
6
7
8
9
10

Table 2 Summary Statistics for the Multi-Ancestry Meta-Analysis top hits by cohort

SNP	rs1451040		rs9660272		rs58763243		rs13403787		rs16838349		rs17115054		rs10807797		rs9545725	
MAF	0.160		0.161		0.070		0.158		0.067		0.059		0.579		0.108	
Effect Allele	T		T		G		A		G		T		A		A	
Chr:Position	1:103158738		1:232253211		2:7762999		2:178459146		2:207427437		5:153074938		7:19995629		13:82056977	
	Beta	P	Beta	P	Beta	P	Beta	P	Beta	P	Beta	P	Beta	P	Beta	P
Non-Hispanic Whites Cohorts																
Spain	-0.095	0.565	-0.060	0.712	-0.413	0.158	0.687	7.50×10^{-5}	0.803	0.001	1.267	1.35×10^{-7}	0.560	1.33×10^{-5}	0.096	0.578
Finland	-0.046	0.905	-0.273	0.481	-1.426	0.003	0.133	0.763	0.674	0.250	1.455	0.032	0.635	0.016	0.015	0.973
Poland	0.360	0.411	-0.427	0.320	-0.285	0.613	0.310	0.588	1.893	0.005	NA ^a	NA ^a	0.127	0.705	0.074	0.892
US EuA	0.254	0.510	-0.586	0.117	-0.844	0.084	NA ^a	NA ^a	0.793	0.130	-0.032	0.963	0.505	0.082	-0.012	0.977
Non-Hispanic White META ^b	-0.004	0.977	-0.187	0.159	0.661	1.41×10^{-3}	0.591	1.43×10^{-4}	0.884	8.74×10^{-6}	1.162	6.41×10^{-8}	0.524	2.82×10^{-7}	0.073	0.614
Hispanic Cohorts																
Costa Rica	-3.236	0.014	-5.203	0.002	0.485	0.691	2.812	0.053	2.674	0.029	0.249	0.893	1.880	0.019	-8.735	1.33×10^{-5}
Mexico	-4.569	7.45×10^{-5}	-6.335	1.17×10^{-6}	0.168	0.825	2.519	0.066	1.451	0.203	0.616	0.678	-0.624	0.343	1.424	6.76×10^{-4}
Hispanic META ^b	-4.131	3.45×10^{-8}	-5.953	1.81×10^{-10}	-0.257	0.690	2.655	6.74×10^{-3}	2.019	0.001	0.473	0.681	0.385	0.444	-6.42	2.07×10^{-8}
Additional Cohorts																
Korea	1.276	0.023	-0.566	0.364	-0.277	0.462	0.827	0.004	NA ^a	NA ^a	NA ^a	NA	0.500	0.084	NA	NA
US AfA	-0.919	0.071	0.078	0.887	-6.491	8.24×10^{-8}	NA ^a	NA ^a	-2.896	0.020	-1.230	0.158	0.416	0.387	-0.514	0.452
	Effect^c	LBF	Effect^c	LBF	Effect^c	LBF	Effect^c	LBF	Effect^c	LBF	Effect^c	LBF	Effect^c	LBF	Effect^c	LBF
MANTRA ^b	---+	6.56	----+	7.64	-+---	6.58	+++?	5.13	+++?	5.41	++?-	5.19	++++	5.97	+++?	5.50

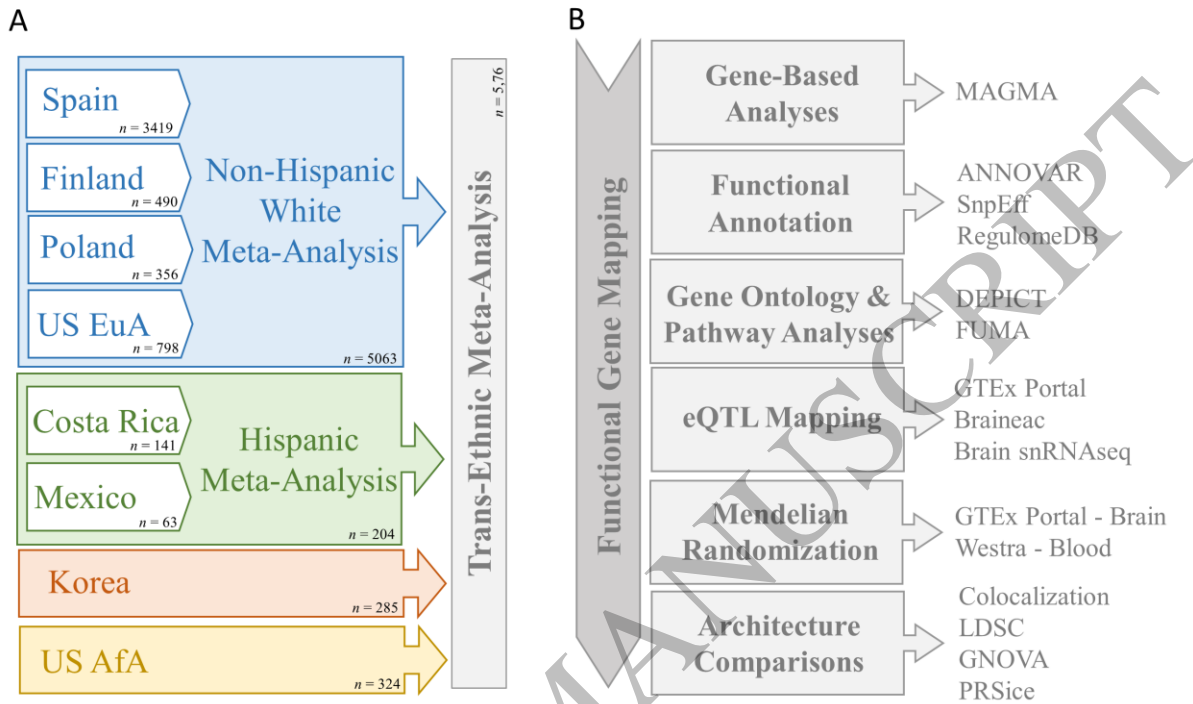
LBF = Log Bayes factor.

^aNA=Not Available due to MAF below the inclusion threshold (0.03) or non-convergence of the statistical model.

^bThe results from the meta-analysis of the single populations.

^cDirection of effect are showed in the following order: Non-Hispanic Whites, Hispanic, Korea and US AfA; +/-/? = positive beta in given population/negative beta in given population/Not present in given population.

1



2

3

4

5

Figure 1
165x103 mm (.15 x DPI)

ACCEPTED MANUSCRIPT

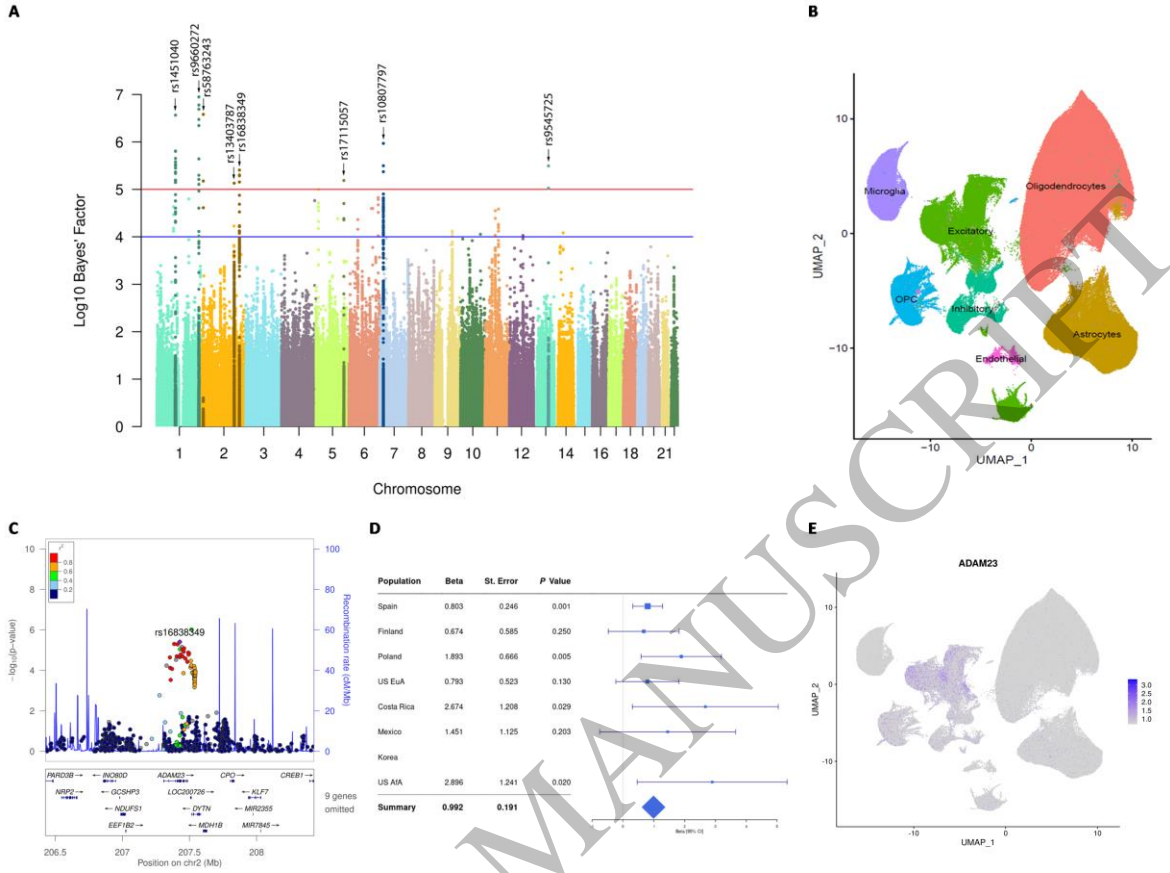


Figure 2
165x115 mm (.15 x DPI)

1
2
3
4

Locus	Total Genes	Gene Name	Gene Based Analysis	eQTL mapping			Mendelian Randomization	
				GTEEx portal	Braineac	snRNAseq	Westra	GTEEx Portal
1p21.1	7	COL11A1	●	●	●	●	●	●
		AMY2B	●	●	●	●	●	
1q42.2	18	GNPAT	●	●	●	●	●	●
		SPRTN	●	●	●	●	●	
		EGLN1	●	●	●	●	●	
		TSNAX	●	●	●	●	●	
		DISC1	●	●	●	●	●	
		NTPCR	●	●	●	●	●	
		EXOC8	●	●	●	●	●	
2p25.1	14	RNF144A	●	●	●	●	●	
		LOC386597	●	●	●	●	●	
2q31.2	24	PLEKHA3	●	●	●	●	●	
		TTC30A	●	●	●	●	●	
		AGPS	●	●	●	●	●	
		DFNB59	●	●	●	●	●	
		AC074286.1	●	●	●	●	●	
		PDE11A	●	●	●	●	●	
2q33.3	24	ADAM23	●	●	●	●	●	
		CREB1	●	●	●	●	●	
		DYTN	●	●	●	●	●	
5q33.2	9	NDUFS1	●	●	●	●	●	
		GRIA1	●	●	●	●	●	
		GALNT10	●	●	●	●	●	
7p21.1	15	TWISTNB	●	●	●	●	●	
		ABCB5	●	●	●	●	●	
13q31.1	0		●	●	●	●	●	

● Not Found ● No Evidence, $P > 0.05$ ● Suggestive Evidence, $0.05 < P < 0.001$ ● Strong Evidence, $P < 0.001$
 ○ Positive Effect ◇ Negative Effect ◻ Unknown Effect

Figure 3
 165x119 mm (.15 x DPI)

1
2
3