

Identification of 20 novel loci associated to ischemic stroke. Epigenome-Wide Association

Study

Carolina Soriano-Tárraga¹ PhD, Uxue Lazcano¹ MSc, Eva Giralt-Steinhauer¹ MD, PhD, Carla Avellaneda-Gómez¹ MD, Ángel Ois¹ MD, PhD, Ana Rodríguez-Campello¹ MD, PhD, Elisa Cuadrado-Godia¹ MD, PhD, Alejandra Gomez-Gonzalez¹ MD, Alba Fernández-Sanlés² MSc, Roberto Elosua^{2,3} MD, PhD, Israel Fernández-Cadenas⁴ PhD, Natalia Cullell^{4,5} Msc, Joan Montaner^{6,7} MD, PhD, Sebastian Moran⁸ PhD, Manel Esteller^{9,10,11,12} MD, PhD, Jordi Jiménez-Conde*¹ MD, PhD, and Jaume Roquer*¹ MD, PhD

1. Neurovascular Research Group, Department of Neurology of Hospital del Mar-IMIM (Institut Hospital del Mar d'Investigacions Mèdiques); Universitat Autònoma de Barcelona/DCEXS-Universitat Pompeu Fabra, Barcelona, Spain.
2. Cardiovascular Epidemiology and Genetics Research Group, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain.
3. CIBER de Enfermedades Cardiovasculares, Barcelona, Spain; Medicine Department, Medical School, University of Vic-Central University of Catalonia (UVic-UCC), Vic, Spain.
4. Stroke Pharmacogenomics and Genetics Group, Institut de Recerca Hospital de la Santa Creu i Sant Pau, Barcelona, Spain.
5. Neurology. Hospital Universitari MútuaTerrassa/Fundacio Docència i Recerca MútuaTerrassa, Spain.
6. Neurovascular Research Laboratory, Vall d'Hebron Institute of Research (VHIR), Universitat Autònoma de Barcelona, Spain.
7. Institute de Biomedicine of Seville, IBiS/Hospital Universitario Virgen del Rocío/CSIC/University of Seville & Department of Neurology, Hospital Universitario Virgen Macarena, Seville, Spain.
8. Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet, Barcelona, Spain.
9. Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Spain.
10. Centro de Investigacion Biomedica en Red Cancer (CIBERONC), Madrid, Spain.
11. Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.
12. Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona (UB), Barcelona, Spain.

*These authors contributed equally

Authors for correspondence: Carolina Soriano-Tárraga

csoriano@imim.es

IMIM, Hospital del Mar Medical Research Institute

DrAiguader 88, 08003 Barcelona, Spain

Word count: 5545

ABSTRACT

Rationale. DNA methylation is dynamic, varies throughout the life course, and its levels are influenced by lifestyle and environmental factors, as well as by genetic variation. The leading genetic variants at stroke risk loci identified to date explain roughly 1–2% of stroke heritability. Most of these single nucleotide polymorphisms are situated within a regulatory sequence marked by DNase I hypersensitivity sites, which would indicate involvement of an epigenetic mechanism.

Objective. To detect epigenetic variants associated to stroke occurrence and stroke subtypes.

Methods and Results. A two-stage case-control epigenome-wide association study was designed. The discovery sample with 401 samples included 218 ischemic stroke (IS) patients, assessed at Hospital del Mar (Barcelona, Spain) and 183 controls from the REGICOR cohort. In two independent samples (N=226 and N=166), we replicated 22 CpG sites differentially methylated in IS in 21 loci, including 2 CpGs in locus *ZFHX3*, which includes known genetic variants associated with stroke. The pathways associated to these loci are inflammation and angiogenesis. The meta-analysis identified 384 differentially methylated CpGs, including loci of known stroke and vascular risk genetic variants, enriched by loci involved in lipid metabolism, adipogenesis, circadian clock, and glycolysis pathways. Stratification analysis by stroke subtypes revealed distinct methylation patterns.

Conclusions. We identified a set of 22 CpGs in 21 loci associated with IS. Our analysis suggests that DNA methylation changes may contribute to orchestrating gene expression that contributes to IS.

Key words: ischemic stroke, DNA methylation, EWAS

Abbreviations

CE	Cardioembolic (stroke etiology)
DNAm	DNA methylation
EWAS	Epigenome-Wide Association Study
GWAS	Genome-Wide Association Study
IS	Ischemic Stroke
LAA	Large-Artery Atherosclerosis (stroke etiology)
MVPs	Methylation-Variable Positions
SNPs	Single Nucleotide Polymorphisms
SVD	Small Vessel Disease (stroke etiology)
TOAST	Trial of Org 10172 in Acute Stroke Treatment
UND	Undetermined (stroke etiology)

INTRODUCTION

Ischemic stroke (IS), the etiology of 80% of all strokes, is a complex and heterogeneous disease with high rates of mortality and long-term disability. Stroke pathogenesis involves a number of different disease processes as well as interactions between environmental, vascular, systemic, genetic, and central nervous system factors ¹.

In recent years, genome-wide association studies (GWAS) have identified many genes robustly associated to IS and its subtypes which show a distinct genetic background, but the underlying genes and pathways are not yet fully understood. Half of the identified genes share a genetic association with other vascular traits; the greatest correlation is with blood pressure. However, the underlying biological pathways do not seem to involve known vascular risk factors and thus the new pathways may constitute new targets for stroke prevention ².

The heritability of stroke, as calculated from genome-wide data, has been estimated to be 30–40% ^{3–5}. The lead variants at stroke risk loci identified to date explain roughly 1–2% of this heritability and their number remains relatively small compared to other common conditions, such as coronary artery disease ⁶. Moreover, most of these leading single nucleotide polymorphisms (SNPs) reside within intergenic or intronic regions, and most are situated within a regulatory sequence marked by DNase I hypersensitivity sites, which would indicate the involvement of an epigenetic mechanism ^{2,7,8}.

DNA methylation (DNAm), an epigenetic mechanism essential for regulation of gene expression, consists of the covalent addition of a methyl group to a cytosine nucleotide, primarily in the context of a CpG dinucleotide. DNAm is dynamic, varies throughout the life course, and its levels are influenced by lifestyle and environmental factors, as well as by genetic variation ⁹.

Stroke epigenetics research has revealed that stroke patients are globally hypomethylated and are *epigenetically* older than healthy controls. This accelerated aging is strongly associated with 3-month IS outcome and mortality, independently of health profile, lifestyle, and genetic characteristics ^{10–13}.

We hypothesized that DNA methylation markers would identify additional stroke risk loci and improve the mapping of genomic variability, providing new insights into stroke pathology. The aim of the present study was to detect epigenetic variants associated to IS occurrence and stroke subtypes.

METHODS

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

Study design

A two-stage case-control epigenome-wide association study (EWAS) was designed, including discovery and replication analyses. The discovery sample consisted of 401 samples (183 controls and 218 IS patients). Two independent samples (N=392,226 from Hospital del Mar [MAR-2] and 166 from Hospital Vall d'Hebron [HVH]) were used to replicate the top 500 methylation-variable positions (MVPs) identified in the discovery sample with an arbitrary p-value $\leq 7.1 \times 10^{-6}$. Joint meta-analysis of all 3 samples was performed (N=793). Stratification analysis by Trial of Org 10172 in Acute Stroke Treatment (TOAST) was performed. Functional and pathway analyses were conducted.

Study Sample

European ancestry controls and patients with a diagnosis of IS according to World Health Organization criteria were selected.

Epigenome-Wide Association Studies

After EWAS quality controls, association analysis was performed in controls and IS samples using a multivariate linear regression. The top MVPs identified in the discovery stage were analyzed in the validation samples. The results were joined in a weighted z-score meta-analysis. Stratification analysis by TOAST criteria was performed using a multivariate linear regression for large-artery atherosclerosis (LAA), cardioembolic (CE), small vessel disease (SVD), and undetermined (UND) etiologies. Functional and pathway analyses were conducted. All MVPs with p-values $< 1.39 \times 10^{-7}$ were considered statistically significant for a genome-wide approach.

RESULTS

Discovery stage. Genome-wide effect of ischemic stroke on methylation status

Of the 865,859 initial CpG sites in Infinium MethylationEPIC Beadchip, a total of 358,709 (41.4%) that passed the quality controls and were common in HumanMethylation450 Beadchip were included in the discovery analysis (**Table S1-2**). Genome-wide DNAm analysis from whole blood was performed in the discovery sample (N=401). Clinical and demographic differences between the IS and control individuals are shown in **Table 1**.

Top 500 CpG sites showing a suggestive association with methylation differences between IS and control samples were selected for replication, with an arbitrary p-value $\leq 7.1 \times 10^{-6}$ (**Supplementary Table S3**). Manhattan and QQ plots are shown in **Figure 1**.

Replication stage

After applying the same quality control steps and normalization as in the discovery analysis, the validation analysis included 392 individuals (226 from MAR-2 and 166 from HVH). The characteristics of the populations included in the replication stage are shown in **Table 1-2**. We replicated 93 CpGs in one replication sample and 114 CpGs in the other; 22 MVPs in 21 loci were replicated in both samples. All of the 22 MVPs were hypomethylated in all the samples, except the one located at *MAPK1*; however, IS samples showed significantly higher methylation levels, compared to controls (**Table 2, Supplementary Figure S1**). Function and traits associated to the 21 loci that harbor the 22 CpGs identified are described in **Supplementary Table S4-7**.

We used the GeneMANIA algorithm on the 21 candidate genes (*CAMSAP3*, *SLC35E1*, *ZFH3*, *PIM3*, *MAPK1*, *LRRC26*, *HIF1A*, *RNF126*, *SENP3*, *ANAPC11*, *PLBD2*, *CCNL2*, *PUM1*, *ITPKB*, *NAPA*, *IL15RA*, *ACSL1*, *JMY*, *PUF60*, *CHSY1*, *BAMBI*) and known IS-related genes². We found that our candidate genes have genetic interactions and are co-expressed with known IS genes (**Supplementary Figure S2**). *FGA* and *ZFH3* are involved in the same pathway, as are *HIF1A* and *FURIN*, *HIF1A* and *MAPK1*. Shared protein domains exist between *HIF1A* and *TWIST1*, *LRRC26* and *LRCH1*, *ANAPC11* and *RNF126*, *ZFH3* and *ZNF318*, and *CHSY1* and *ABO*. *CDK6* shares with *MAPK1* and *PIM3*. Finally, there is co-localization expression of *ITPKB* and *SH3PXD2A*;

NAPA and *RGS7*; *NAPA*, *RGS7*, *ZFH3*, and *MAPK1*; *ZFH3* and *TBX3*; and *IL15RA*, *CASZ1*, and *FOXF2*.

The 22 validated CpGs include 2 CpGs in *ZFH3* loci that harbor known stroke variants. The pathway analysis showed *HIF1A* and *MAPK1* are associated to angiogenesis (FDR p-values=0.014). *IL15RA* and *HIF1A* are involved in inflammatory and interferon gamma response (FDR p-values=0.034). Additionally, *NAPA* and *MAPK1* are involved in protein secretion functions (FDR p-values=0.024) (**Supplementary Table S5**). GTEX data indicate that 20 loci (the exception being *LRRC26*) are expressed in blood vessels and brain (**Supplementary Figure S3**).

These 22 MVPs are associated to CpG islands: 20 are in an enhancer-associated position, 2 are associated to H3K27AC (indicating active gene expression), and 21 are located in a DNase I hypersensitive site, functionally related to transcriptional activity (**Supplementary Table S7**).

Meta-analysis

Meta-analysis results of all 3 samples (N=793) are shown in **Supplementary Table S8**. A total of 384 MVPs were significant after Bonferroni correction (multiple testing). However, 30 (7.8%) of these MVPs were only present in two samples.

The 384 significantly associated CpGs include 4 CpGs in *ZFH3*, *SH2B3* and *WNK1* loci, which harbor known stroke variants. The traits and diseases associated to loci containing these CpGs are shown in **Figure 2**. The functional analysis of the significant loci showed enrichment for genes involved in adipogenesis (FDR p-value= 4.13×10^{-4}), triglyceride and lipid metabolism (FDR p-value=0.044), circadian clock (FDR p-value=0.044), and regulation of glycolysis pathway (FDR p-value=0.037) (**Supplementary Tables S9-11**).

EWAS stratified analysis by IS subtypes

We combined the IS samples from the two Hospital del Mar samples (MAR_1 and MAR_2, N=627) to increase the statistical power to identify MVP sites associated to IS subtypes (TOAST classification), stratified by LAA, CE, SVD, and UND. Clinical and demographic characteristics and EWAS results are shown in **Supplementary Material (Table S12-16)**. Manhattan and QQ plots are in **Supplemental Figure S4**. In the meta-analysis, UND and SVD subtypes were the main

contributors of significant IS-associated CpGs (**Figure 3**).

In LAA subtype, we identified 57 MVPs suggestive of statistical significance (**Supplemental Table S13**). There was an enrichment for genes involved in plasma high-density lipoprotein cholesterol (HDL-C) levels (FDR p-values=0.039) and nominal significance for the adipogenesis pathway (p-values=0.0023) (**Supplementary Tables S17**). In CE subtype, we identified 31 CpGs (**Supplemental Table S14**). The functional analysis showed the insulin signaling pathway as nominally significant (p-values= 1.7×10^{-3}) (**Supplementary Tables S18**). In SVD subtype, we identified 121 MVPs suggestive of association (**Supplemental Table S15**). Enrichment analysis was significant for different traits associated to lipid metabolism (FDR p-values=0.013495) and nominally significant for adipogenous hallmark genes (p-value=0.0024) (**Supplemental Table S19**). In UND subtype, we identified 294 MVPs (**Supplemental Table S16**). Enrichment analysis was significant for traits associated to inflammation (FDR p-values=0.0027) and cellular processes (**Supplementary Tables S20**).

The traits and diseases associated to loci that contain the suggestive CpGs associated to IS subtypes are shown in **Supplementary Tables S20-S24**. Approximately 30% of these MVPs overlap with known loci having genetic variants associated to vascular risk factors (**Figure 4**).

Discussion

Novel findings in the present study include an association between IS and differentially methylated positions of 22 CpGs located in 21 loci (*CAMSAP3*, *SLC35E1*, *ZFHX3*, *PIM3*, *MAPK1*, *LRRC26*, *HIF1A*, *RNF126*, *SENP3*, *ANAPC11*, *PLBD2*, *CCNL2*, *PUM1*, *ITPKB*, *NAPA*, *IL15RA*, *ACSL1*, *JMY*, *PUF60*, *CHSY1*, *BAMBI*). These CpGs are more methylated in IS than in control samples, except the one in the locus *MAPK1*. Two of these CpGs, cg00614832 and cg07786668, are situated in the *ZFHX3* locus, a known stroke gene². Pathway analysis indicated that these 21 loci are involved in inflammation and angiogenesis pathways. Additionally, they share pathways, protein domains, genetic interaction and co-expression with stroke-associated genes identified in previous GWAS studies².

The repercussion of these findings in gene expression has not been studied; however, the 22 validated CpGs are located in CpG island-promoters, 20 of them in enhancer elements, and 21 in a DNase I hypersensitive site, functionally related to transcriptional activity; 12 are associated to H3K27AC, indicating active gene expression. All together, these results suggest a possible effect on gene expression.

The meta-analysis identified 384 MVPs significantly associated to IS, but replication is needed to confirm these results because of possible inflation of p-values. These MVPs were enriched for genes involved in adipogenesis, triglyceride and lipid metabolism, circadian clock, and regulation of glycolysis pathway. We also identified shared MVP associations with 5 genes (*SH2B3*, *ZFHX3*, *WNK1* (beside *NINJ2*), *PCIF1*, *TRIM47*, *YAP1*) previously associated to stroke, to vascular risk factors and to traits such as diabetes mellitus, coagulation, blood pressure, aortic aneurysm, body mass index, white-matter hyperintensity, and coronary heart disease^{2,14-20}.

The overlap between loci containing SNPs associated to vascular risk factors and those related to stroke highlights the importance of epigenetics in identifying possible new biomarkers of stroke risk or/and new causal variants related to stroke. The question of potential independent effects of MVPs and SNPs remains to be studied. Most of the stroke risk SNPs identified are located in intergenic or

intronic regions, and MVPs that were differentially methylated in the meta-analysis are located in the promoter region, making them better candidates to be the causal variant².

The stratification analysis by stroke subtypes revealed distinct methylation patterns that could provide further mechanistic insights; however, replication is needed. The LAA subtype showed enrichment for genes involved in plasma HDL-C levels and adipogenesis pathway which correlates to LAA pathophysiology. Its top MVP, in the *GPS2* locus, has recently been associated with obesity and diabetes mellitus, with the authors suggesting that this gene may contribute to atherogenesis²¹. Additionally, cg17218495 is in the *SMARCA4* locus, previously associated to stroke and lipid metabolism^{22,23}. The CE analysis was enriched by genes involved in the insulin signaling pathway and adipogenous hallmark genes, while SVD analysis was enriched by genes involved in lipid metabolisms. Moreover, two MVPs were located in *CNNM2*, a known variant associated to blood pressure, and *TRIM47*, associated to white-matter hyperintensities which correlates to SVD pathophysiology^{24,25}.

DNAm can respond to changes in the environment and might mediate gene expression and phenotypes induced by prior exposure, such as glycemic exposure, a phenomenon called metabolic memory²⁶⁻²⁸. Based on those studies, DNAm may be one of the mechanisms involved in the metabolic memory. We only reported an association with IS traits, but cannot infer causality of these methylation changes; however, these MVPs could be causal variants, biomarkers of exposure, or stochastic changes. Further investigation is needed.

In this study, we analyzed one of the largest available series of IS patients with genome-wide methylation data. We replicated our results in two independent samples. Moreover, we identified 22 CpG sites related to IS, two of them located in the *ZFHX3* locus, harboring known genetic variants associated with stroke. Stratification by stroke subtype revealed distinct methylation patterns associated to different vascular risk factors; however, these results must be replicated.

EWAS are prone to significant inflation and bias of test statistics. Neither GWAS-based methodology nor confounder adjustment methods completely remove bias and inflation²⁹.

Epigenome data differ from genetic data: they are quantitative measures (whereas genotypes are discrete) that are subject to major confounding effects of batch analysis, whether technical, such as DNA extraction protocols; biological, including cellular heterogeneity; or epigenetic changes caused by the disease process itself^{29–32}. In our study, blood samples were collected in the acute phase of IS; although we adjusted for stroke severity using the NIH scale in an effort to correct the residual confounding, the inflation is evident. Nonetheless, some signals were identified and others require replication.

Some limitations of the study should be considered. First, we measured DNAm in peripheral blood cells. Some authors have suggested that the methylation levels of some CpGs/regions are tissue-specific³³, and we might have lost some signals by not choosing specific tissues that could have a higher impact in DNAm. However, other authors consider methylation patterns of whole blood a good proxy for the methylation levels from a specific site of action. Second, the cross-sectional study design precluded any inference of causality in the reported association between IS and DNAm levels. Third, there is a lack of replication in the meta-analysis. Fourth, we cannot draw conclusions about the effect of these methylation changes in gene expression because we lack gene expression data.

DNAm changes in 22 CpGs play a role in IS pathophysiology. They may orchestrate gene expression that contributes to IS and/or offer a biomarker of IS and vascular risk factors. Further investigation is required to assess the functional effect of DNAm changes and establish causality.

ACKNOWLEDGEMENTS

We are grateful to all neurologists, nurses, and residents from the stroke unit and to technical staff from the IMIM who helped in performing this study. Elaine M. Lilly, PhD, provided English language assistance.

SOURCES OF FUNDING

This project was funded in part by the following sources: Agència de Gestió Ajuts Universitaris de Recerca (2014 SGR 1213; 2017 SGR 222; PERIS SLT002/16/00088); Spain's Ministry of Health (Ministerio de Sanidad y Consumo) through the Carlos III Health Institute (ISCIII-FIS-FEDER-ERDF, PI12/01238, PI15/00451, PI15/0044, PI15/00051, PI18/00017, PI17/02089, PI18/01338, CIBERCV); INVICTUS-PLUS, Instituto de Salud Carlos III RETIC (RD16/0019/0002); and a RecerCaixa 2013 research grant (JJ086116). Fundació la Marató TV3 (76/C/2011). A. Fernández-Sanlés was funded by the Spanish Ministry of Economy and Competitiveness (BES-2014-069718). Epigenesis and Maestro projects.

DISCLOSURES

None

REFERENCES

1. WHO | Global status report on noncommunicable diseases 2014. Available at <http://www.who.int/nmh/publications/ncd-status-report-2014/en/>. Accessed December 2, 2015.
2. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese A-K, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD, Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen W-M, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, Den Hoed M, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu F-C, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jiménez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee J-M, Lemmens R, Leys D, Lewis CM, Lin W-Y, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmäe K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM, Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S, Bis JC, Pastinen T, Ruusalepp A, Schadt EE, Koplev S, Björkegren JLM, Codoni V, Civelek M, Smith NL, Trégouët DA, Christophersen IE, Roselli C, Lubitz SA, Ellinor PT, Tai ES, Kooner JS, Kato N, He J, van der Harst P, Elliott P, Chambers JC, Takeuchi F, Johnson AD, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT, Rofls A, Hata J, Woo D, Rosand J, Pare G, Hopewell

JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Debette S, Dichgans M, Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Ruten-Jacobs L, Giese A-K, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD, Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen W-M, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, Hoed MD, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu F-C, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jiménez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee J-M, Lemmens R, Leys D, Lewis CM, Lin W-Y, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmäe K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM, Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S, Amin N, Aparicio HS, Arnett DK, Attia J, Beiser AS, Berr C, Buring JE, Bustamante M, Caso V, Cheng Y-C, Choi SH, Chowhan A, Culell N, Dartigues J-F, Delavaran H, Delgado P, Dörr M, Engström G, Ford I, Gurpreet WS, Hamsten A, Heitsch L, Hozawa A, Ibanez L, Ilinca A, Ingelsson M, Iwasaki M, Jackson RD, Jood K, Jousilahti P, Kaffashian S, Kalra L, Kamouchi M, Kitazono T, Kjartansson O, Kloss M, Koudstaal PJ, Krupinski J, Labovitz DL, Laurie CC, Levi CR, Li L, Lind L, Lindgren CM, Lioutas V, Liu YM, Lopez OL, Makoto H, Martinez-Majander N, Matsuda K, Minegishi N, Montaner J, Morris AP, Muiño E, Müller-Nurasyid M, Norrving B, Ogishima S, Parati EA, Peddareddygari LR, Pedersen NL, Pera J, Perola M,

- Pezzini A, Pileggi S, Rabionet R, Riba-Llena I, Ribasés M, Romero JR, Roquer J, Rudd AG, Sarin A-P, Sarju R, Sarnowski C, Sasaki M, Satizabal CL, Satoh M, Sattar N, Sawada N, Sibolt G, Sigurdsson Á, Smith A, Sobue K, Soriano-Tárraga C, Stanne T, Stine OC, Stott DJ, Strauch K, Takai T, Tanaka H, Tanno K, Teumer A, Tomppo L, Torres-Aguila NP, Touze E, Tsugane S, Uitterlinden AG, Valdimarsson EM, van der Lee SJ, Völzke H, Wakai K, Weir D, Williams SR, Wolfe CDA, Wong Q, Xu H, Yamaji T, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT, Rolfs A, Hata J, Woo D, Rosand J, Pare G, Hopewell JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Debette S, Dichgans M. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature Genetics*. 2018. doi:10.1038/s41588-018-0058-3.
3. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NLM, Jackson C, Farrall M, Rothwell PM, Sudlow C, Dichgans M, Markus HS. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke; a journal of cerebral circulation*. 2012;43:3161–7.
 4. Dichgans M, Pulit SL, Rosand J. Stroke genetics: discovery, biology, and clinical applications. *The Lancet Neurology*. 2019;18:587–599.
 5. Malik R, Dau T, Gonik M, Sivakumar A, Deredge DJ, Edeleva E V, Götzfried J, van der Laan SW, Pasterkamp G, Beaufort N, Seixas S, Bevan S, Lincz LF, Holliday EG, Burgess AI, Rannikmäe K, Minnerup J, Kriebel J, Waldenberger M, Müller-Nurasyid M, Lichtner P, Saleheen D, International Stroke Genetics Consortium PM, Rothwell PM, Levi C, Attia J, Sudlow CLM, Braun D, Markus HS, Wintropde PL, Berger K, Jenne DE, Dichgans M. Common coding variant in SERPINA1 increases the risk for large artery stroke. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;114:3613–3618.
 6. Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T,

Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang S-J, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikäinen L-P, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith A V, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han B-G, Huang J, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki M-L, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon F-U-R, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim B-J, Kooner JS, Kullo IJ, Lehtimäki T, Loos RJF, Melander O, Metspalu A, März W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature Genetics*. 2015;47:1121–1130.

7. Holliday EG, Maguire JM, Evans T-J, Koblar SA, Jannes J, Sturm JW, Hankey GJ, Baker R, Golledge J, Parsons MW, Malik R, McEvoy M, Biros E, Lewis MD, Lincz LF, Peel R, Oldmeadow C, Smith W, Moscato P, Barlera S, Bevan S, Bis JC, Boerwinkle E, Boncoraglio GB, Brott TG, Brown RD, Cheng Y-C, Cole JW, Cotlarciuc I, Devan WJ, Fornage M, Furie KL, Grétarsdóttir S, Gschwendtner A, Ikram MA, Longstreth WT, Meschia JF, Mitchell BD,

- Mosley TH, Nalls MA, Parati EA, Psaty BM, Sharma P, Stefansson K, Thorleifsson G, Thorsteinsdottir U, Traylor M, Verhaaren BFJ, Wiggins KL, Worrall BB, Sudlow C, Rothwell PM, Farrall M, Dichgans M, Rosand J, Markus HS, Scott RJ, Levi C, Attia J. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nature genetics*. 2012;44:1147–51.
8. Kilarski LL, Achterberg S, Devan WJ, Traylor M, Malik R, Lindgren A, Pare G, Sharma P, Slowik A, Thijs V, Walters M, Worrall BB, Sale MM, Algra A, Kappelle LJ, Wijmenga C, Norrving B, Sandling JK, Rönnblom L, Goris A, Franke A, Sudlow C, Rothwell PM, Levi C, Holliday EG, Fornage M, Psaty B, Gretarsdottir S, Thorsteinsdottir U, Seshadri S, Mitchell BD, Kittner S, Clarke R, Hopewell JC, Bis JC, Boncoraglio GB, Meschia J, Ikram MA, Hansen BM, Montaner J, Thorleifsson G, Stefansson K, Rosand J, de Bakker PIW, Farrall M, Dichgans M, Markus HS, Bevan S. Meta-analysis in more than 17,900 cases of ischemic stroke reveals a novel association at 12q24.12. *Neurology*. 2014;83:678–85.
9. Portela A, Esteller M. Epigenetic modifications and human disease. *Nature biotechnology*. 2010;28:1057–68.
10. Soriano-Tárraga C, Jiménez-Conde J, Giralt-Steinhauer E, Mola M, Ois Á, Rodríguez-Campello A, Cuadrado-Godia E, Fernández-Cadenas I, Carrera C, Montaner J, Elosua R, Roquer J. Global DNA methylation of ischemic stroke subtypes. *PLoS ONE*. 2014;9. doi:10.1371/journal.pone.0096543.
11. Soriano-Tárraga C, Giralt-Steinhauer E, Mola-Caminal M, Vivanco-Hidalgo R, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, Sayols-Baixeras S, Elosua R, Roquer J, Jiménez-Conde J. Ischemic stroke patients are biologically older than their chronological age. *Aging*. 2016;8:2655–2666.
12. Soriano-Tárraga C, Mola-Caminal M, Giralt-Steinhauer E, Ois A, Rodríguez-Campello A,

- Cuadrado-Godia E, Gómez-González A, Vivanco-Hidalgo RM, Fernández-Cadenas I, Cullell N, Roquer J, Jiménez-Conde J. Biological age is better than chronological as predictor of 3-month outcome in ischemic stroke. *Neurology*. 2017;89:830–836.
13. Soriano-Tárraga C, Giralt-Steinhauer E, Mola-Caminal M, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, Fernández-Cadenas I, Cullell N, Roquer J, Jiménez-Conde J. Biological Age is a predictor of mortality in Ischemic Stroke. *Scientific Reports*. 2018;8.
doi:10.1038/s41598-018-22579-0.
14. NINDS Stroke Genetics Network (SiGN) SL, International Stroke Genetics Consortium (ISGC) PF, Wong Q, Malik R, Gwinn K, Achterberg S, Algra A, Amouyel P, Anderson CD, Arnett DK, Arsava EM, Attia J, Ay H, Bartz TM, Battey T, Benavente OR, Bevan S, Biffi A, Bis JC, Blanton SH, Boncoraglio GB, Brown RD, Burgess AI, Carrera C, Chapman Smith SN, Chasman DI, Chauhan G, Chen W-M, Cheng Y-C, Chong M, Cloonan LK, Cole JW, Cotlarciuc I, Cruchaga C, Cuadrado-Godia E, Dave T, Dawson J, Debette S, Delavaran H, Dell CA, Dichgans M, Doheny KF, Dong C, Duggan DJ, Engström G, Evans MK, Pallejà XE, Faul JD, Fernández-Cadenas I, Fornage M, Frossard PM, Furie K, Gamble DM, Gieger C, Giese A-K, Giralt-Steinhauer E, González HM, Goris A, Gretarsdottir S, Grewal RP, Grittner U, Gustafsson S, Han B, Hankey GJ, Heitsch L, Higgins P, Hochberg MC, Holliday E, Hopewell JC, Horenstein RB, Howard G, Ikram MA, Ilinca A, Ingelsson E, Irvin MR, Jackson RD, Jern C, Conde JJ, Johnson JA, Jood K, Kahn MS, Kaplan R, Kappelle LJ, Kardia SLR, Keene KL, Kissela BM, Kleindorfer DO, Koblar S, Labovitz D, Launer LJ, Laurie CC, Laurie CA, Lee CH, Lee J-M, Lehm M, Lemmens R, Levi C, Leys D, Lindgren A, Longstreth WT, Maguire J, Manichaikul A, Markus HS, McClure LA, McDonough CW, Meisinger C, Melander O, Meschia JF, Mola-Caminal M, Montaner J, Mosley TH, Müller-Nurasyid M, Nalls MA, O'Connell JR, O'Donnell M, Ois ángel, Papanicolaou GJ, Paré G, Peddareddygari LR, Pedersén A, Pera J, Peters A, Poole D, Psaty BM, Rabionet R, Raffeld MR, Rannikmäe K, Rasheed A, Redfors P, Reiner AP, Rexrode K, Ribasés M, Rich SS, Robberecht W,

- Rodriguez-Campello A, Rolfs A, Roquer J, Rose LM, Rosenbaum D, Rost NS, Rothwell PM, Rundek T, Ryan KA, Sacco RL, Sale MM, Saleheen D, Salomaa V, Sánchez-Mora C, Schmidt CO, Schmidt H, Schmidt R, Schürks M, Scott R, Segal HC, Seiler S, Seshadri S, Sharma P, Shuldiner AR, Silver B, Slowik A, Smith JA, Söderholm M, Soriano C, Sparks MJ, Stanne T, Stefansson K, Stine OC, Strauch K, Sturm J, Sudlow CL, Tajuddin SM, Talbert RL, Tatlisumak T, Thijs V, Thorleifsson G, Thorsteindottir U, Tiedt S, Traylor M, Trompet S, Valant V, Waldenberger M, Walters M, Wang L, Wassertheil-Smoller S, Weir DR, Wiggins KL, Williams SR, Wloch-Kopec D, Woo D, Woodfield R, Wu O, Xu H, Zonderman AB, Worrall BB, de Bakker PI, Kittner SJ, Mitchell BD, Rosand J, Mitchell BD, Ay H, Gwinn K, Kittner SJ, Lindgren A, Meschia JF, Pulit SL, Sudlow CLM, Thijs V, Woo D, Worrall BBW, Arnett DKA, Benavente O, Cole JW, Dichgans M, Grewal RP, Jern C, Conde JJ, Johnson JA, Kittner SJ, Lee J-M, Levi C, Lindgren A, Markus HS, Melander O, Meschia JF, Rexrode K, Rosand J, Rothwell PM, Rundek T, Sacco RL, Schmidt R, Sharma P, Slowik A, Sudlow CLM, Thijs V, Wassertheil-Smoller S, Woo D, Worrall BB. Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *The Lancet Neurology*. 2016;15:174–184.
15. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, Debette S, Lumley T, Folsom AR, van den Herik EG, Bos MJ, Beiser A, Cushman M, Launer LJ, Shahar E, Struchalin M, Du Y, Glazer NL, Rosamond WD, Rivadeneira F, Kelly-Hayes M, Lopez OL, Coresh J, Hofman A, DeCarli C, Heckbert SR, Koudstaal PJ, Yang Q, Smith NL, Kase CS, Rice K, Haritunians T, Roks G, de Kort PLM, Taylor KD, de Lau LM, Oostra BA, Uitterlinden AG, Rotter JI, Boerwinkle E, Psaty BM, Mosley TH, van Duijn CM, Breteler MMB, Longstreth WT, Wolf PA. Genomewide association studies of stroke. *The New England journal of medicine*. 2009;360:1718–28.
16. Jones GT, Tromp G, Kuivaniemi H, Gretarsdottir S, Baas AF, Giusti B, Strauss E, van't Hof FNG, Webb TR, Erdman R, Ritchie MD, Elmore JR, Verma A, Pendergrass S, Kullo IJ, Ye Z,

Peissig PL, Gottesman O, Verma SS, Malinowski J, Rasmussen-Torvik LJ, Borthwick KM, Smelser DT, Crosslin DR, de Andrade M, Ryer EJ, McCarty CA, Böttinger EP, Pacheco JA, Crawford DC, Carrell DS, Gerhard GS, Franklin DP, Carey DJ, Phillips VL, Williams MJA, Wei W, Blair R, Hill AA, Vasudevan TM, Lewis DR, Thomson IA, Krysa J, Hill GB, Roake J, Merriman TR, Oszkinis G, Galora S, Saracini C, Abbate R, Pulli R, Pratesi C, Saratzis A, Verissimo AR, Bumpstead S, Badger SA, Clough RE, Cockerill G, Hafez H, Scott DJA, Futers TS, Romaine SPR, Bridge K, Griffin KJ, Bailey MA, Smith A, Thompson MM, van Bockxmeer FM, Matthiasson SE, Thorleifsson G, Thorsteinsdottir U, Blankensteijn JD, Teijink JAW, Wijmenga C, de Graaf J, Kiemeny LA, Lindholt JS, Hughes A, Bradley DT, Stirrups K, Golledge J, Norman PE, Powell JT, Humphries SE, Hamby SE, Goodall AH, Nelson CP, Sakalihasan N, Courtois A, Ferrell RE, Eriksson P, Folkersen L, Franco-Cereceda A, Eicher JD, Johnson AD, Betsholtz C, Ruusalepp A, Franzén O, Schadt EE, Björkegren JLM, Lipovich L, Drolet AM, Verhoeven EL, Zebregs CJ, Geelkerken RH, van Sambeek MR, van Sterkenburg SM, de Vries J-P, Stefansson K, Thompson JR, de Bakker PIW, Deloukas P, Sayers RD, Harrison SC, van Rij AM, Samani NJ, Bown MJ. Meta-Analysis of Genome-Wide Association Studies for Abdominal Aortic Aneurysm Identifies Four New Disease-Specific Risk Loci. *Circulation Research*. 2017;120:341–353.

17. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, Ntritsos G, Dimou N, Cabrera CP, Karaman I, Ng FL, Evangelou M, Witkowska K, Tzani E, Hellwege JN, Giri A, Velez Edwards DR, Sun Y V, Cho K, Gaziano JM, Wilson PWF, Tsao PS, Kovesdy CP, Esko T, Mägi R, Milani L, Almgren P, Boutin T, Debette S, Ding J, Giulianini F, Holliday EG, Jackson AU, Li-Gao R, Lin W-Y, Luan J, Mangino M, Oldmeadow C, Prins BP, Qian Y, Sargurupremraj M, Shah N, Surendran P, Thériault S, Verweij N, Willems SM, Zhao J-H, Amouyel P, Connell J, de Mutsert R, Doney ASF, Farrall M, Menni C, Morris AD, Noordam R, Paré G, Poulter NR, Shields DC, Stanton A, Thom S, Abecasis G, Amin N, Arking DE, Ayers KL, Barbieri CM, Batini C, Bis JC, Blake T, Bochud M, Boehnke M,

Boerwinkle E, Boomsma DI, Bottinger EP, Braund PS, Brumat M, Campbell A, Campbell H, Chakravarti A, Chambers JC, Chauhan G, Ciullo M, Cocca M, Collins F, Cordell HJ, Davies G, de Borst MH, de Geus EJ, Deary IJ, Deelen J, Del Greco M F, Demirkale CY, Dörr M, Ehret GB, Elosua R, Enroth S, Erzurumluoglu AM, Ferreira T, Frånberg M, Franco OH, Gandin I, Gasparini P, Giedraitis V, Gieger C, Girotto G, Goel A, Gow AJ, Gudnason V, Guo X, Gyllensten U, Hamsten A, Harris TB, Harris SE, Hartman CA, Havulinna AS, Hicks AA, Hofer E, Hofman A, Hottenga J-J, Huffman JE, Hwang S-J, Ingelsson E, James A, Jansen R, Jarvelin M-R, Joehanes R, Johansson Å, Johnson AD, Joshi PK, Jousilahti P, Jukema JW, Jula A, Kähönen M, Kathiresan S, Keavney BD, Khaw K-T, Knekt P, Knight J, Kolcic I, Kooner JS, Koskinen S, Kristiansson K, Kutalik Z, Laan M, Larson M, Launer LJ, Lehne B, Lehtimäki T, Liewald DCM, Lin L, Lind L, Lindgren CM, Liu Y, Loos RJF, Lopez LM, Lu Y, Lyytikäinen L-P, Mahajan A, Mamasoula C, Marrugat J, Marten J, Milaneschi Y, Morgan A, Morris AP, Morrison AC, Munson PJ, Nalls MA, Nandakumar P, Nelson CP, Niiranen T, Nolte IM, Nutile T, Oldehinkel AJ, Oostra BA, O'Reilly PF, Org E, Padmanabhan S, Palmas W, Palotie A, Pattie A, Penninx BWJH, Perola M, Peters A, Polasek O, Pramstaller PP, Nguyen QT, Raitakari OT, Ren M, Rettig R, Rice K, Ridker PM, Ried JS, Riese H, Ripatti S, Robino A, Rose LM, Rotter JI, Rudan I, Ruggiero D, Saba Y, Sala CF, Salomaa V, Samani NJ, Sarin A-P, Schmidt R, Schmidt H, Shrine N, Siscovick D, Smith A V, Snieder H, Söber S, Sorice R, Starr JM, Stott DJ, Strachan DP, Strawbridge RJ, Sundström J, Swertz MA, Taylor KD, Teumer A, Tobin MD, Tomaszewski M, Toniolo D, Traglia M, Trompet S, Tuomilehto J, Tzourio C, Uitterlinden AG, Vaez A, van der Most PJ, van Duijn CM, Vergnaud A-C, Verwoert GC, Vitart V, Völker U, Vollenweider P, Vuckovic D, Watkins H, Wild SH, Willemsen G, Wilson JF, Wright AF, Yao J, Zemunik T, Zhang W, Attia JR, Butterworth AS, Chasman DI, Conen D, Cucca F, Danesh J, Hayward C, Howson JMM, Laakso M, Lakatta EG, Langenberg C, Melander O, Mook-Kanamori DO, Palmer CNA, Risch L, Scott RA, Scott RJ, Sever P, Spector TD, van der Harst P, Wareham NJ, Zeggini E, Levy D, Munroe PB, Newton-Cheh C, Brown MJ, Metspalu A, Hung AM, O'Donnell CJ, Edwards TL, Psaty BM,

Tzoulaki I, Barnes MR, Wain L V, Elliott P, Caulfield MJ, Million Veteran Program. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nature genetics*. 2018;50:1412–1425.

18. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JRB, Egan JM, Lajunen T, Grarup N, Sparsø T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proença C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccascocca RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Böttcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen Y-DI, Chines P, Clarke R, Coin LJM, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day INM, de Geus EJC, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllenstein U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen A-L, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PR V, Jørgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martínez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orrù M, Pakyz R, Palmer CNA, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AFH, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S,

Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurðsson G, Sijbrands EJG, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvänen A-C, Tanaka T, Thorand B, Tichet J, Tönjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JCM, Yarnell JWG, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Borecki IB, Loos RJF, Meneton P, Magnusson PKE, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis G V, Serrano-Ríos M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WHL, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF, Stumvoll M, Campbell H, Wilson JF, Anders Hamsten on behalf of Procardis Consortium, MAGIC investigators, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BWJH, Boomsma DI, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruukonen A, Jarvelin M-R, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature Genetics*. 2010;42:105–116.

19. Verhaaren BFJ, Debette S, Bis JC, Smith JA, Ikram MK, Adams HH, Beecham AH, Rajan KB, Lopez LM, Barral S, van Buchem MA, van der Grond J, Smith A V., Hegenscheid K, Aggarwal NT, de Andrade M, Atkinson EJ, Beekman M, Beiser AS, Blanton SH, Boerwinkle E, Brickman AM, Bryan RN, Chauhan G, Chen CPLH, Chouraki V, de Craen AJM, Crivello

- F, Deary IJ, Deelen J, De Jager PL, Dufouil C, Elkind MSV, Evans DA, Freudenberger P, Gottesman RF, Guðnason V, Habes M, Heckbert SR, Heiss G, Hilal S, Hofer E, Hofman A, Ibrahim-Verbaas CA, Knopman DS, Lewis CE, Liao J, Liewald DCM, Luciano M, van der Lugt A, Martinez OO, Mayeux R, Mazoyer B, Nalls M, Nauck M, Niessen WJ, Oostra BA, Psaty BM, Rice KM, Rotter JI, von Sarnowski B, Schmidt H, Schreiner PJ, Schuur M, Sidney SS, Sigurdsson S, Slagboom PE, Stott DJM, van Swieten JC, Teumer A, Töglhofer AM, Traylor M, Trompet S, Turner ST, Tzourio C, Uh H-W, Uitterlinden AG, Vernooij MW, Wang JJ, Wong TY, Wardlaw JM, Windham BG, Wittfeld K, Wolf C, Wright CB, Yang Q, Zhao W, Zijdenbos A, Jukema JW, Sacco RL, Kardia SLR, Amouyel P, Mosley TH, Longstreth WT, DeCarli CC, van Duijn CM, Schmidt R, Launer LJ, Grabe HJ, Seshadri SS, Ikram MA, Fornage M. Multiethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI. *Circulation: Cardiovascular Genetics*. 2015;8:398–409.
20. Jian X, Satizabal CL, Smith A V, Wittfeld K, Bis JC, Smith JA, Hsu F-C, Nho K, Hofer E, Hagenaars SP, Nyquist PA, Mishra A, Adams HHH, Li S, Teumer A, Zhao W, Freedman BI, Saba Y, Yanek LR, Chauhan G, van Buchem MA, Cushman M, Royle NA, Bryan RN, Niessen WJ, Windham BG, DeStefano AL, Habes M, Heckbert SR, Palmer ND, Lewis CE, Eiriksdottir G, Maillard P, Mathias RA, Homuth G, Valdés-Hernández MDC, Divers J, Beiser AS, Langner S, Rice KM, Bastin ME, Yang Q, Maldjian JA, Starr JM, Sidney S, Risacher SL, Uitterlinden AG, Gudnason VG, Nauck M, Rotter JI, Schreiner PJ, Boerwinkle E, van Duijn CM, Mazoyer B, von Sarnowski B, Gottesman RF, Levy D, Sigurdsson S, Vernooij MW, Turner ST, Schmidt R, Wardlaw JM, Psaty BM, Mosley TH, DeCarli CS, Saykin AJ, Bowden DW, Becker DM, Deary IJ, Schmidt H, Kardia SLR, Ikram MA, Debette S, Grabe HJ, Longstreth WT, Seshadri S, Launer LJ, Fornage M, neuroCHARGE Working Group M. Exome Chip Analysis Identifies Low-Frequency and Rare Variants in MRPL38 for White Matter Hyperintensities on Brain Magnetic Resonance Imaging. *Stroke*. 2018;49:1812–1819.
21. Drareni K, Ballaire R, Barilla S, Mathew MJ, Toubal A, Fan R, Liang N, Chollet C, Huang Z,

- Kondili M, Foufelle F, Soprani A, Roussel R, Gautier J-F, Alzaid F, Treuter E, Venticlef N. GPS2 Deficiency Triggers Maladaptive White Adipose Tissue Expansion in Obesity via HIF1A Activation. *Cell Reports*. 2018;24:2957–2971.e6.
22. Surakka I, Horikoshi M, Mägi R, Sarin AP, Mahajan A, Lagou V, Marullo L, Ferreira T, Miraglio B, Timonen S, Kettunen J, Pirinen M, Karjalainen J, Thorleifsson G, Hägg S, Hottenga JJ, Isaacs A, Ladenvall C, Beekman M, Esko T, Ried JS, Nelson CP, Willenborg C, Gustafsson S, Westra HJ, Blades M, De Craen AJM, De Geus EJ, Deelen J, Grallert H, Hamsten A, Havulinna AS, Hengstenberg C, Houwing-Duistermaat JJ, Hyppönen E, Karssen LC, Lehtimäki T, Lyssenko V, Magnusson PKE, Mihailov E, Müller-Nurasyid M, Mpindi JP, Pedersen NL, Penninx BWJH, Perola M, Pers TH, Peters A, Rung J, Smit JH, Steinthorsdottir V, Tobin MD, Tsernikova N, Van Leeuwen EM, Viikari JS, Willems SM, Willemsen G, Schunkert H, Erdmann J, Samani NJ, Kaprio J, Lind L, Gieger C, Metspalu A, Eline Slagboom P, Groop L, Van Duijn CM, Eriksson JG, Jula A, Salomaa V, Boomsma DI, Power C, Raitakari OT, Ingelsson E, Järvelin MR, Thorsteinsdottir U, Franke L, Ikonen E, Kallioniemi O, Pietiäinen V, Lindgren CM, Stefansson K, Palotie A, McCarthy MI, Morris AP, Prokopenko I, Ripatti S. The impact of low-frequency and rare variants on lipid levels. *Nature Genetics*. 2015;47:589–597.
23. Dichgans M, Malik R, König IR, Rosand J, Clarke R, Gretarsdottir S, Thorleifsson G, Mitchell BD, Assimes TL, Levi C, O'Donnell CJ, Fornage M, Thorsteinsdottir U, Psaty BM, Hengstenberg C, Seshadri S, Erdmann J, Bis JC, Peters A, Boncoraglio GB, März W, Meschia JF, Kathiresan S, Ikram MA, McPherson R, Stefansson K, Sudlow C, Reilly MP, Thompson JR, Sharma P, Hopewell JC, Chambers JC, Watkins H, Rothwell PM, Roberts R, Markus HS, Samani NJ, Farrall M, Schunkert H. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke; a journal of cerebral circulation*. 2014;45:24–36.

24. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith A V, Tobin MD, Verwoert GC, Hwang S-J, Pihur V, Vollenweider P, O'Reilly PF, Amin N, Bragg-Gresham JL, Teumer A, Glazer NL, Launer L, Zhao JH, Aulchenko Y, Heath S, Söber S, Parsa A, Luan J, Arora P, Dehghan A, Zhang F, Lucas G, Hicks AA, Jackson AU, Peden JF, Tanaka T, Wild SH, Rudan I, Igl W, Milaneschi Y, Parker AN, Fava C, Chambers JC, Fox ER, Kumari M, Go MJ, van der Harst P, Kao WHL, Sjögren M, Vinay DG, Alexander M, Tabara Y, Shaw-Hawkins S, Whincup PH, Liu Y, Shi G, Kuusisto J, Tayo B, Seielstad M, Sim X, Nguyen K-DH, Lehtimäki T, Matullo G, Wu Y, Gaunt TR, Onland-Moret NC, Cooper MN, Platou CGP, Org E, Hardy R, Dahgam S, Palmen J, Vitart V, Braund PS, Kuznetsova T, Uiterwaal CSPM, Adeyemo A, Palmas W, Campbell H, Ludwig B, Tomaszewski M, Tzoulaki I, Palmer ND, Aspelund T, Garcia M, Chang Y-PC, O'Connell JR, Steinle NI, Grobbee DE, Arking DE, Kardina SL, Morrison AC, Hernandez D, Najjar S, McArdle WL, Hadley D, Brown MJ, Connell JM, Hingorani AD, Day INM, Lawlor DA, Beilby JP, Lawrence RW, Clarke R, Hopewell JC, Ongen H, Dreisbach AW, Li Y, Young JH, Bis JC, Kähönen M, Viikari J, Adair LS, Lee NR, Chen M-H, Olden M, Pattaro C, Bolton JAH, Köttgen A, Bergmann S, Mooser V, Chaturvedi N, Frayling TM, Islam M, Jafar TH, Erdmann J, Kulkarni SR, Bornstein SR, Grässler J, Groop L, Voight BF, Kettunen J, Howard P, Taylor A, Guarrera S, Ricceri F, Emilsson V, Plump A, Barroso I, Khaw K-T, Weder AB, Hunt SC, Sun Y V, Bergman RN, Collins FS, Bonnycastle LL, Scott LJ, Stringham HM, Peltonen L, Perola M, Vartiainen E, Brand S-M, Staessen JA, Wang TJ, Burton PR, Soler Artigas M, Dong Y, Snieder H, Wang X, Zhu H, Lohman KK, Rudock ME, Heckbert SR, Smith NL, Wiggins KL, Doumatey A, Shriner D, Veldre G, Viigimaa M, Kinra S, Prabhakaran D, Tripathy V, Langefeld CD, Rosengren A, Thelle DS, Corsi AM, Singleton A, Forrester T, Hilton G, McKenzie CA, Salako T, Iwai N, Kita Y, Ogiwara T, Ohkubo T, Okamura T, Ueshima H, Umemura S, Eyheramendy S, Meitinger T, Wichmann H-E, Cho YS, Kim H-L, Lee J-Y, Scott J, Sehmi JS, Zhang W, Hedblad B, Nilsson P, Smith GD, Wong A, Narisu N, Stančáková A, Raffel LJ, Yao J, Kathiresan S, O'Donnell CJ, Schwartz SM, Ikram MA, Longstreth WT,

Mosley TH, Seshadri S, Shrine NRG, Wain L V, Morken MA, Swift AJ, Laitinen J, Prokopenko I, Zitting P, Cooper JA, Humphries SE, Danesh J, Rasheed A, Goel A, Hamsten A, Watkins H, Bakker SJL, van Gilst WH, Janipalli CS, Mani KR, Yajnik CS, Hofman A, Mattace-Raso FUS, Oostra BA, Demirkan A, Isaacs A, Rivadeneira F, Lakatta EG, Orru M, Scuteri A, Ala-Korpela M, Kangas AJ, Lyytikäinen L-P, Soininen P, Tukiainen T, Würtz P, Ong RT-H, Dörr M, Kroemer HK, Völker U, Völzke H, Galan P, Hercberg S, Lathrop M, Zelenika D, Deloukas P, Mangino M, Spector TD, Zhai G, Meschia JF, Nalls MA, Sharma P, Terzic J, Kumar MVK, Denniff M, Zukowska-Szczechowska E, Wagenknecht LE, Fowkes FGR, Charchar FJ, Schwarz PEH, Hayward C, Guo X, Rotimi C, Bots ML, Brand E, Samani NJ, Polasek O, Talmud PJ, Nyberg F, Kuh D, Laan M, Hveem K, Palmer LJ, van der Schouw YT, Casas JP, Mohlke KL, Vineis P, Raitakari O, Ganesh SK, Wong TY, Tai ES, Cooper RS, Laakso M, Rao DC, Harris TB, Morris RW, Dominiczak AF, Kivimaki M, Marmot MG, Miki T, Saleheen D, Chandak GR, Coresh J, Navis G, Salomaa V, Han B-G, Zhu X, Kooner JS, Melander O, Ridker PM, Bandinelli S, Gyllenstein UB, Wright AF, Wilson JF, Ferrucci L, Farrall M, Tuomilehto J, Pramstaller PP, Elosua R, Soranzo N, Sijbrands EJG, Altshuler D, Loos RJF, Shuldiner AR, Gieger C, Meneton P, Uitterlinden AG, Wareham NJ, Gudnason V, Rotter JI, Rettig R, Uda M, Strachan DP, Witteman JCM, Hartikainen A-L, Beckmann JS, Boerwinkle E, Vasani RS, Boehnke M, Larson MG, Jarvelin M-R, Psaty BM, Abecasis GR, Chakravarti A, Elliott P, van Duijn CM, Newton-Cheh C, Levy D, Caulfield MJ, Johnson T. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011;478:103–9.

25. Verhaaren BFJ, Debette S, Bis JC, Smith JA, Ikram MK, Adams HH, Beecham AH, Rajan KB, Lopez LM, Barral S, Van Buchem MA, Van Der Grond J, Smith A V., Hegenscheid K, Aggarwal NT, De Andrade M, Atkinson EJ, Beekman M, Beiser AS, Blanton SH, Boerwinkle E, Brickman AM, Bryan RN, Chauhan G, Chen CPLH, Chouraki V, De Craen AJM, Crivello F, Deary IJ, Deelen J, De Jager PL, Dufouil C, Elkind MSV, Evans DA, Freudenberger P,

- Gottesman RF, Gunason V, Habes M, Heckbert SR, Heiss G, Hilal S, Hofer E, Hofman A, Ibrahim-Verbaas CA, Knopman DS, Lewis CE, Liao J, Liewald DCM, Luciano M, Van Der Lugt A, Martinez OO, Mayeux R, Mazoyer B, Nalls M, Nauck M, Niessen WJ, Oostra BA, Psaty BM, Rice KM, Rotter JI, Von Sarnowski B, Schmidt H, Schreiner PJ, Schuur M, Sidney SS, Sigurdsson S, Slagboom PE, Stott DJM, Van Swieten JC, Teumer A, Töglhofer AM, Traylor M, Trompet S, Turner ST, Tzourio C, Uh HW, Uitterlinden AG, Vernooij MW, Wang JJ, Wong TY, Wardlaw JM, Windham BG, Wittfeld K, Wolf C, Wright CB, Yang Q, Zhao W, Zijdenbos A, Jukema JW, Sacco RL, Kardia SLR, Amouyel P, Mosley TH, Longstreth WT, DeCarli CC, Van Duijn CM, Schmidt R, Launer LJ, Grabe HJ, Seshadri SS, Ikram MA, Fornage M. Multiethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI. *Circulation: Cardiovascular Genetics*. 2015;8:398–409.
26. Soriano-Tárraga C, Jiménez-Conde J, Giral-Steinhauer E, Mola-Caminal M, Vivanco-Hidalgo RM, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, Sayols-Baixeras S, Elosua R, Roquer J, Krupinski J, Fernández-Cadenas I, Carrera C, Rabionet K, Obach V, Vivas T, Navarro RMD. Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Human Molecular Genetics*. 2016;25. doi:10.1093/hmg/ddv493.
27. Ceriello A. The emerging challenge in diabetes: the “metabolic memory”. *Vascular pharmacology*. 2012;57:133–8.
28. Waki H, Yamauchi T, Kadowaki T. The epigenome and its role in diabetes. *Current diabetes reports*. 2012;12:673–85.
29. van Iterson M, van Zwet EW, Heijmans BT, 't Hoen PAC, van Meurs J, Jansen R, Franke L, Boomsma DI, Pool R, van Dongen J, Hottenga JJ, van Greevenbroek MMJ, Stehouwer CDA, van der Kallen CJH, Schalkwijk CG, Wijmenga C, Zherakova S, Tigchelaar EF, Eline Slagboom P, Beekman M, Deelen J, van Heemst D, Veldink JH, van den Berg LH, van Duijn

- CM, Hofman BA, Isaacs A, Uitterlinden AG, Jhamai PM, Verbiest M, Suchiman HED, Verkerk M, van der Breggen R, van Rooij J, Lakenberg N, Mei H, van Galen M, Bot J, Zhernakova D V., van 't Hof P, Deelen P, Nooren I, Moed M, Vermaat M, Luijk R, Bonder MJ, van Dijk F, Arindrarto W, Kielbasa SM, Swertz MA, 't Hoen PB. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biology*. 2017;18. doi:10.1186/s13059-016-1131-9.
30. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *International journal of epidemiology*. 2012;41:161–76.
31. Soriano-Tárraga C, Jiménez-Conde J, Giralt-Steinhauer E, Ois Á, Rodríguez-Campello A, Cuadrado-Godia E, Fernández-Cadenas I, Montaner J, Lucas G, Elosua R, Roquer J. DNA Isolation Method Is a Source of Global DNA Methylation Variability Measured with LUMA. Experimental Analysis and a Systematic Review. *PLoS ONE*. 2013;8. doi:10.1371/journal.pone.0060750.
32. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K, Irizarry RA. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature reviews Genetics*. 2010;11:733–739.
33. Lowe R, Slodkowitz G, Goldman N, Rakyan VK. The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. *Epigenetics*: official journal of the DNA Methylation Society. 2015. doi:10.1080/15592294.2014.1003744.

FIGURE LEGENDS

Figure 1. Manhattan and QQ-plot in the discovery sample. A) Manhattan plot (red line and blue line for statistically and nominally significance threshold) and B) QQ-plot.

Figure 2. Summary of loci harboring significant CpGs in meta-analysis, overlapping known genome-wide associations to phenotypes and diseases. A) 115 CpG (30%) of the 384 joint CpGs are associated to known variants. Bars represent the number of loci associated to each trait; 37 loci are associated to more than 1 trait. B) List of loci with known associations. In bold, loci validated in discovery.

BMI: Body Mass Index; BP: Blood Pressure; DM: Diabetes Mellitus; CHD: Coronary Heart Disease

Figure 3. Venn Diagrams showing the overlap of methylation-variable positions between meta-analysis and the ischemic stroke subtypes. LAA, Large-Artery Atherosclerosis; CE, Cardiembolic; SVD, Small vessel diseases; UND, Undetermined IS subtypes.

Figure 4. Summary of loci harboring nominally significant CpGs, by TOAST, associated to phenotypes and diseases in genome-wide association studies. One locus can be associated to more than one trait. LAA, Large-Artery Atherosclerosis; CE, Cardiembolic; SVD, Small vessel diseases; UND, Undetermined IS subtypes; BMI, Body Mass Index; BP, Blood Pressure; DM, Diabetes Mellitus; CHD, Coronary Heart Disease.

Table 1. Descriptive characteristics of the samples.

Joint N=793	Discovery MAR_1 sample N=401			Replication					
				MAR_2 sample N=226			HVH sample N=166		
	Controls REGICOR N=183	IS N=218	p-value	Controls N=41	IS N=185	p-value	Controls N=21	IS N=145	p-value
Age, years *	64 (58-69)	71 (62-78)	<0.001	62.8 (14.2)	74.9 (12.6)	<0.001	75 (67-76)	75 (67-81)	0.704
Sex, female, n (%)	95 (51.9)	74 (33.9)	<0.001	20 (48.8)	90 (48.6)	NS	8 (38.1)	45 (31.0)	0.517
Hypertension, n (%)	107 (58.5)	172 (78.9)	<0.001	19 (46.3)	137 (74.1)	0.001	12 (57.1)	94 (64.8)	0.489
Smoking habit, n (%)									
Current smoker	16 (8.7)	72 (33.0)		8 (19.5)	37 (20)		2 (9.5)	26 (17.9)	
Former smoker	2 (1.1)	49 (22.5)	<0.001	12 (29.3)	48 (25.9)	NS	-	-	0.388
Never smoker	165 (90.2)	97 (44.5)		21 (51.2)	99 (53.5)		17 (81.0)	113 (77.9)	
Body mass index, kg/m ² †	28.1 (25.8-31.6)	26.6 (23.7-30.1)	<0.001	28.4 (24.7-31.3)	27.2 (24.2-30.2)	NS	-	-	-
Diabetes mellitus, n (%)	30 (16.4)	81 (37.2)	<0.001	7 (17.1)	69 (37.3)	0.013	6 (28.6)	41 (28.3)	0.001
Hyperlipidemia, n (%)	124 (67.8)	123 (56.4)	0.020	20 (48.8)	87 (47)	NS	-	49/109 (44.9)	-
Atrial fibrillation, n (%)	5 (2.7)	51 (23.4)	<0.001	2 (4.9)	95 (51.4)	<0.001	0	35 (24.1)	0.005
Coronary heart disease, n (%)	5 (2.7)	28 (12.8)	<0.001	1 (2.4)	28 (15.1)	0.027	0	10/82 (12.2)	0.003
NIH stroke scale †	-	3 (1-6)	-	-	5 (3-12.5)	-	-	10 (2-17)	-
Stroke etiology									
Large-artery atherosclerosis		79 (19.7)			18 (8)			54 (37.5)	
Cardioembolic, n (%)		39 (9.7)	-	-	109 (48.2)	-		16 (11)	-
Small-vessel disease, n (%)		36 (9.0)			58 (25.7)			34 (23.4)	
Undetermined, n (%)		64 (16.0)						37 (22.3)	

*Median (IQR)

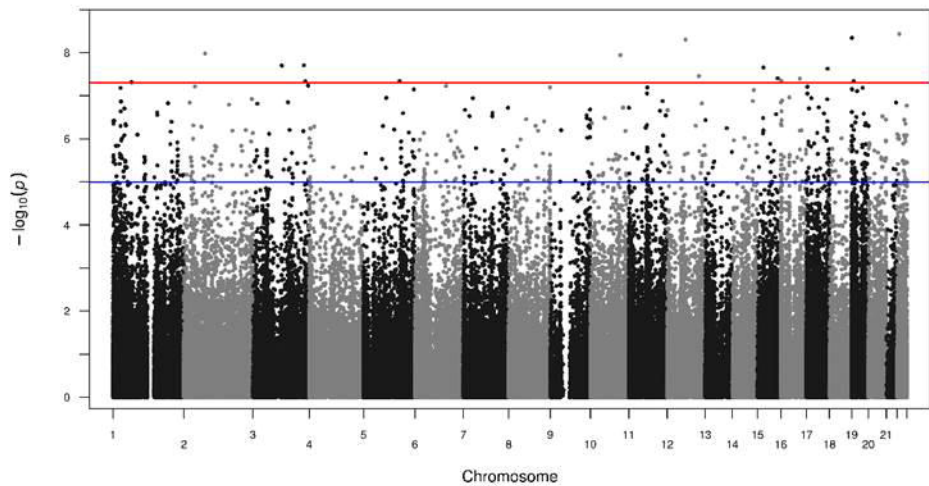
MAR: Hospital del Mar; HVH: Hospital Vall d'Hebron; NIH: National Institutes of Health.

Table 2. CpG sites differentially methylated between ischemic stroke and controls in discovery sample validated in two replication samples and the meta-analysis.

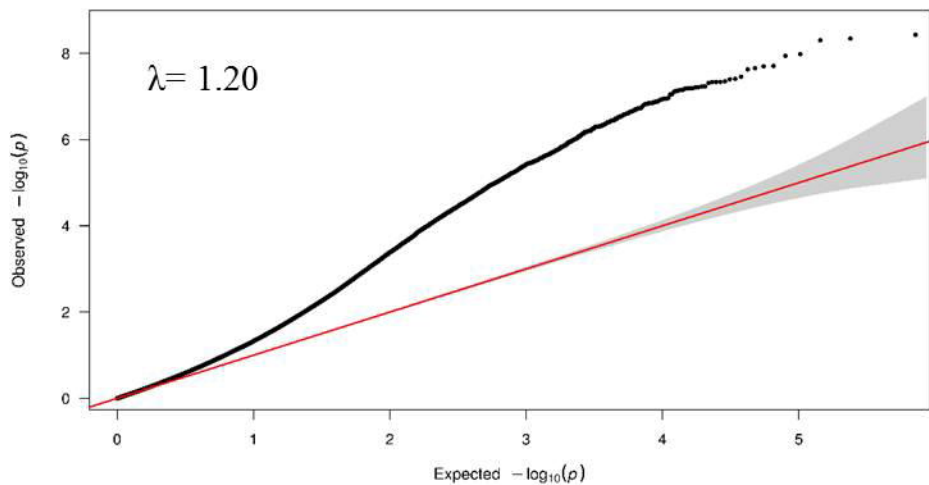
Features					Discovery					Replication 1		Replication 2		Meta-analysis	
CpG	Chr	Pos (hg37)	CpG location	Gene	LogFC	Beta (SD) Control	Beta (SD) IS	DNAm diff	P-value	LogFC	P-value FDR	LogFC	P-value FDR	P-value	Dir
cg09915769	19	7660977	Island	CAMSAP3	0.03	0.07 (0.02)	0.09 (0.02)	3	4.56E-08	4.56E-08	5.08E-05	0.02	5.59E-06	4.26E-16	+++
cg02463426	19	16683387	Island	SLC35E1	0.02	0.06 (0.04)	0.10 (0.06)	2	3.11E-07	3.11E-07	1.34E-05	0.02	2.85E-06	5.19E-16	+++
cg00614832	16	73092394	Island	ZFH3	0.04	0.07 (0.04)	0.11 (0.07)	4	6.41E-07	6.41E-07	0.0001	0.04	2.15E-08	3.23E-16	+++
cg23962478	22	50354086	Island	PIM3	0.03	0.06 (0.03)	0.08 (0.04)	3	8.31E-07	8.31E-07	9.94E-05	0.02	1.14E-06	5.61E-15	+++
cg23681311	22	22221878	Island	MAPK1	0.04	0.78 (0.03)	0.77 (0.03)	4	8.50E-07	8.50E-07	0.0001	0.05	2.18E-05	6.59E-14	+++
cg13696351	9	140063617	Island	LRRC26	0.03	0.05 (0.03)	0.08 (0.04)	3	9.07E-07	9.07E-07	3.60E-06	0.03	1.45E-06	3.03E-16	+++
cg01182555	14	62162064	Island	HIF1A	0.02	0.10 (0.03)	0.13 (0.04)	2	9.57E-07	9.57E-07	0.0001	0.004	2.73E-05	8.18E-14	+++
cg07691609	19	662740	Island	RNF126	0.02	0.06 (0.04)	0.12 (0.08)	2	1.16E-06	1.16E-06	8.56E-05	0.02	6.88E-08	9.84E-16	+++
cg01733795	17	7465439	Island	SENP3	0.02	0.10 (0.04)	0.13 (0.06)	2	1.38E-06	1.38E-06	2.41E-06	0.02	1.19E-06	2.96E-16	+++
cg08184047	17	79849980	Island	ANAPC11	0.02	0.06 (0.04)	0.08 (0.06)	2	1.74E-06	1.74E-06	4.14E-07	0.01	1.00E-04	3.24E-15	+++
cg04355250	12	113796401	Island	PLBD2	0.02	0.07 (0.02)	0.08 (0.04)	2	1.93E-06	1.93E-06	6.47E-05	0.01	1.33E-05	5.78E-14	+++
cg16573386	1	1334508	Island	CCNL2	0.02	0.10 (0.03)	0.12 (0.04)	2	2.67E-06	2.67E-06	1.09E-06	0.02	4.31E-08	2.90E-17	+++
cg23281075	1	31538727	Island	PUM1	0.02	0.08 (0.02)	0.11 (0.05)	2	2.84E-06	2.84E-06	9.06E-05	0.02	6.63E-07	1.38E-14	+++
cg07786668	16	73092391	Island	ZFH3	0.04	0.09 (0.03)	0.11 (0.04)	4	3.06E-06	3.06E-06	3.70E-08	0.06	3.57E-08	1.63E-18	+++
cg04482794	1	226925181	Island	ITPKB	0.02	0.07 (0.04)	0.12 (0.07)	2	3.36E-06	3.36E-06	2.77E-05	0.01	4.32E-06	2.05E-14	+++
cg07806715	19	48018254	Island	NAPA	0.02	0.05 (0.03)	0.09 (0.05)	2	3.58E-06	3.58E-06	1.49E-06	0.02	4.94E-07	3.14E-16	+++
cg08676905	10	6019609	Island	IL15RA	0.02	0.05 (0.02)	0.07 (0.02)	2	3.80E-06	3.80E-06	2.82E-06	0.02	1.69E-06	1.47E-15	+++
cg07619799	4	185747409	Island	ACSL1	0.02	0.04 (0.02)	0.07 (0.04)	2	4.48E-06	4.48E-06	7.15E-05	0.01	1.53E-05	1.79E-13	+++
cg04759220	5	78532560	Island	JMY	0.03	0.07 (0.05)	0.10 (0.07)	2	4.70E-06	4.70E-06	5.26E-06	0.02	2.54E-05	2.66E-14	+++
cg01963056	8	144911482	Island	PUF60	0.01	0.02 (0.01)	0.03 (0.02)	1	6.12E-06	6.12E-06	6.67E-06	0.07	2.45E-12	1.21E-13	+++
cg25869317	15	101792241	Island	CHSY1	0.04	0.07 (0.02)	0.08 (0.03)	4	7.00E-06	7.00E-06	1.67E-06	0.04	1.72E-10	3.06E-19	+++
cg04192862	10	28966472	Island	BAMBI	0.03	0.07 (0.03)	0.09 (0.04)	3	7.10E-06	7.10E-06	2.19E-07	0.01	5.59E-06	6.32E-19	+++

CpG, CpG identification; Chr, chromosome location; pos, genomic position; gene, associated gene and observed mean β -values (standard deviation); LogFC, estimate of the log₂-fold-change corresponding to the effect or contrast; P-value FDR, false discovery rate correcting by multiple testing; Dir, direction of the effect. All p-values were calculated with beta-value.

A

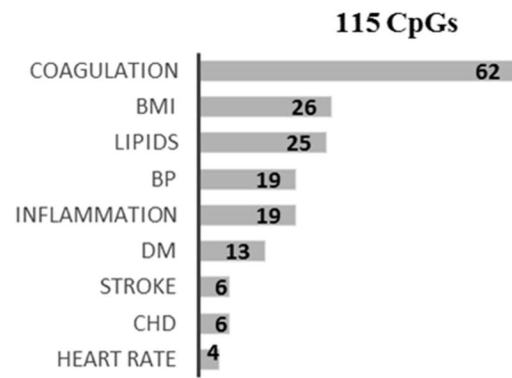
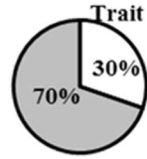


B

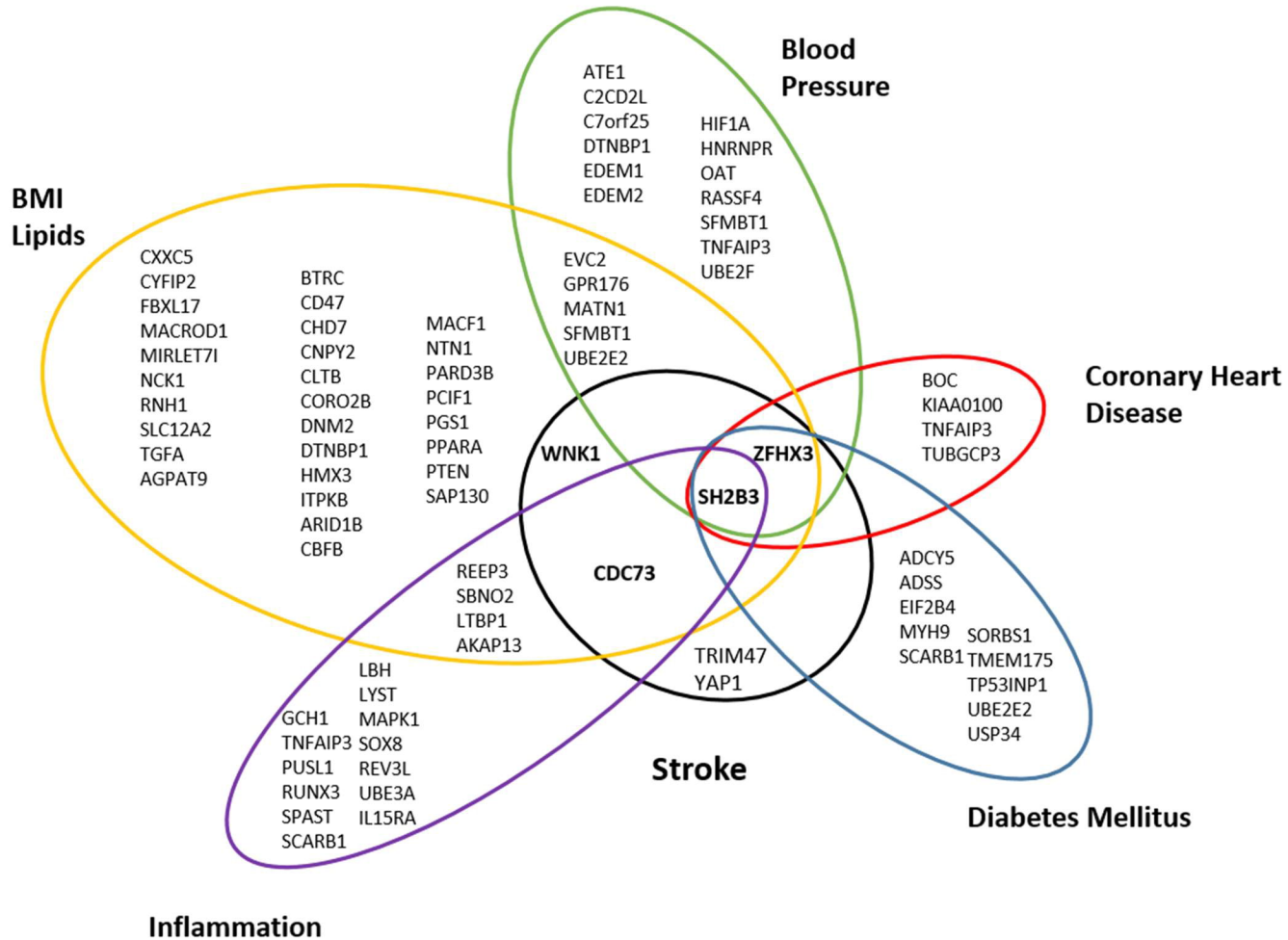


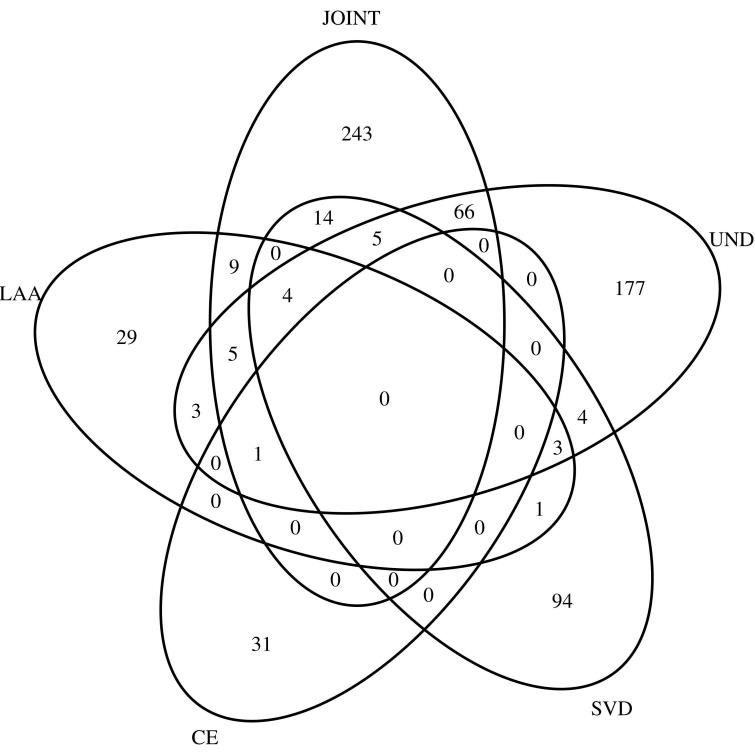
A

Joint
384 CpGs

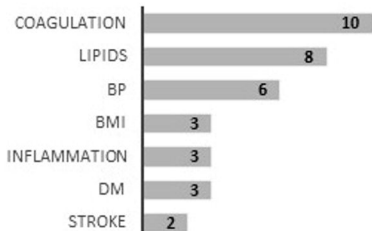


B

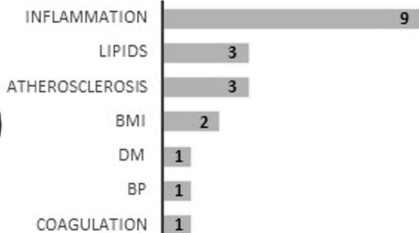
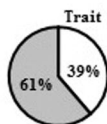




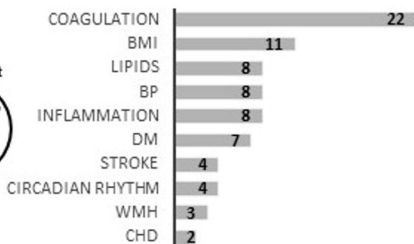
LAA



CE



SVD



UND

