HHS Public Access

Author manuscript

Stroke. Author manuscript; available in PMC 2017 February 01.

Published in final edited form as:

Stroke. 2016 February ; 47(2): 307–316. doi:10.1161/STROKEAHA.115.011328.

Genome-wide association analysis of young onset stroke identifies a locus on chromosome 10q25 near *HABP2*

A full list of authors and affiliations appears at the end of the article.

Abstract

Background and Purpose—Although a genetic contribution to ischemic stroke is well recognized, only a handful of stroke loci have been identified by large-scale genetic association studies to date. Hypothesizing that genetic effects might be stronger for early- versus late-onset stroke, we conducted a two-stage meta-analysis of genome-wide association studies (GWAS), focusing on stroke cases with an age of onset < 60 years old.

Methods—The Discovery stage of our GWAS included 4,505 cases and 21,968 controls of European, South-Asian and African ancestry, drawn from 6 studies. In Stage 2, we selected the lead genetic variants at loci with association $P<5\times10^{-6}$ and performed *in silico* association analyses in an independent sample of up to 1,003 cases and 7,745 controls.

Results—One stroke susceptibility locus at 10q25 reached genome-wide significance in the combined analysis of all samples from the Discovery and Follow-up Stages (rs11196288, OR=1.41, $P=9.5\times10^{-9}$). The associated locus is in an intergenic region between *TCF7L2* and *HABP2*. In a further analysis in an independent sample, we found that two SNPs in high linkage disequilibrium with rs11196288 were significantly associated with total plasma factor VII-activating protease levels, a product of *HABP2*.

Conclusions—*HABP2*, which encodes an extracellular serine protease involved in coagulation, fibrinolysis, and inflammatory pathways, may be a genetic susceptibility locus for early-onset stroke.

Keywords

ischemic stroke; genome-wide analysis; genetics

Corresponding Author: Braxton D Mitchell, PhD, 685 W. Baltimore St., MSTF 302, Baltimore, Maryland 21201, University of Maryland, School of Medicine, bmitchel@medicine.umaryland.edu.

Disclosures

Drs. Kittner, Longstreth, and Worrall are supported by research grants from NIH. Dr. Worrall is Deputy Editor for AAN/Neurology. Dr. Cole is supported by a research grant from the Department of Veterans Affairs. Dr. Boncoraglio is supported by a research grant from the Fondazione IRCCS Istituto Neurologico Carlo Besta. Dr. Metso is supported by grants from the Finnish Medical Foundation, the Orion Farmos Research Foundation, the Maud Kuistila Memorial Foundation, and the Emil Aaltonen Foundation. Dr. Danesh serves on advisory boards for Novartis, Merck Sharp & Dohme UK, Sanofi, the Medical Research Council, and Wellcome Trust, and is a consultant for Takeda. The other authors report no conflicts.

Introduction

Stroke is the 4th leading cause of death in the U.S. and a major cause of long-term disability among adults.[1] Ischemic stroke, the predominant type of stroke, has a multifactorial etiology, with heritability estimated at 37~38%.[2, 3] One approach to gain insights into the molecular basis underlying stroke susceptibility is to identify stroke susceptibility genes and then assess the functions of these genes. A handful of stroke susceptibility loci have recently been identified.[4, 5] Additional stroke susceptibility loci can perhaps be identified by studying special high risk populations and/or focusing on specific subtypes of stroke.

We focus in this study on identifying susceptibility loci associated with early onset stroke, based on the premise that variants associated with younger onset stroke might have higher penetrance and effect sizes than those associated with older onset stroke. Disease-associated variants have been associated with early-onset forms of numerous other complex diseases, including breast cancer, diabetes, and heart disease. While a minority of stroke cases occur at young age, the best available evidence suggests that stroke at younger ages has a stronger genetic basis than stroke occurring at older ages.[4]

To identify loci associated with early-onset stroke, we conducted a 2-stage genome-wide association study (GWAS), in which we meta-analyzed 6 individual studies in the Discovery Stage and then followed up suggestively associated loci in independent samples. In a combined analysis of the Discovery and Follow-up Stages, we detected significant evidence for association to a locus on chromosome 10q25 that met genome-wide significance thresholds. Further analyses of this locus suggest a possible link to *HABP2*, which encodes Factor VII activating protease.

Methods

Genome-wide association meta-analysis

The Discovery stage for the GWAS meta-analysis included 4,505 early-onset ischemic stroke cases and 21,968 controls from 6 case-control studies (Table 1). Our analysis included only early-onset ischemic stroke cases, defined as stroke age of onset <60 years using study-specific diagnostic criteria. Two of the Discovery Stage studies by design included only young onset cases (GEOS and SIFAP). Four of the Discovery studies included individuals of European ancestry only, one included individuals of South-Asian ancestry only, and one included individuals of both European ancestry and African Americans. Three of the studies by design included previously genotyped controls selected from the same geographic region as the cases and genotyped on the same SNP array.

The Follow-up stage included a total of 1,003 independent, early-onset cases and 7,745 controls from three independent case-control studies of European ancestry populations. Study-specific subject inclusion/exclusion criteria are provided in the Supplementary Data, and study characteristics and sample sizes are summarized in Table 1 and Supplemental Tables I–II.

All studies were genotyped on Illumina SNP arrays (San Diego, CA, USA). Following data cleaning, principal component analyses were performed to identify population outliers and to determine population substructure for adjustment in the association analyses. Quality control filters for samples and SNPs were applied within each study prior to imputation (Supplemental Tables III–IV). Imputation was performed in the Discovery Stage samples based on either the 1000 Genomes Phase 1 interim reference panel or the Phase 1 integrated reference panel using IMPUTE2 software (Supplemental Table III).[6] Post-imputation filters further excluded SNPs from association analyses having minor allele frequency (MAF)<1% or imputation quality score (INFO)<0.3. In the Follow-up stage studies, all studies were imputed using 1000 Genomes Phase 1 integrated reference panel prior to association testing.

For the three studies that by design included previously genotyped controls (RACE I, SIFAP, and WTCCC2-UK), additional quality control steps, such as removing SNPs with evidence for differential missingness, were performed to ensure comparability of genotyping performance between cases and controls. Q-Q plots for all studies revealed no evidence for genomic inflation (lambdas ranging from 1.00–1.06) and the meta-analysis was further GC-corrected.

Detailed information about the genotyping platforms, genotype cleaning procedures, imputation process, and association analysis modeling is described in the Supplemental Data and Supplemental Tables III–IV.

GWAS analysis was performed within each study using a logistic regression model with case/control status as the dependent variable and SNP allelic dosage as the independent variable to test for a multiplicative effect of the SNP on risk of ischemic stroke. The study-specific effect of each SNP, expressed as the *ln*-transformed odds ratio (OR) or β, was obtained after adjusting for the effect of population structure and/or additional study-specific variables prior to meta-analysis.

To account for the multi-ethnic composition of studies, we performed a trans-ethnic GWAS meta-analysis that combined association summary statistics from European, South-Asian and African ethnic groups (n=4,505 cases and 21,968 controls from all studies). The Follow-up stage (3 studies, n=1,003 cases and 7,745 controls) consisted of European Caucasians only. Ethnic-specific β estimates (obtained by meta-analyzing studies of the same ethnicity) were meta-analyzed assuming a random effect model to allow for heterogeneity between different ethnic groups. Post-analysis we excluded SNPs with MAF<1% and INFO<0.3. We additionally excluded SNPs with high between-study heterogeneity (I^2>50%) and SNPs present in only 1 ethnic group.

All meta-analyses were performed using the GWAMA software and repeated using the fixed-effect model as well as Han and Eskin's random effects model (implemented with METASOFT software) to evaluate the consistency of associations. Meta-analyses were performed by two independent analysts and the results were consistent.

SNPs with $P<5\times10^{-6}$ in the Discovery stage and having high LD partners that also showed evidence for association were assessed for association in the Follow-up stage via *in silico* look-ups. A joint analysis of the combined Discovery and Follow-up studies was performed

by meta-analyzing the study-specific estimates from all studies to obtain joint ORs and p-values under fixed effect model. We considered a p-value<5×10^{-8} from the joint analysis as our threshold for genome-wide statistical significance.

We also performed a European-only GWAS meta-analysis (Discovery Stage: n=2,567 cases and 17,163 controls from 5 studies), in which study-specific β estimates were meta-analyzed using the inverse-variance weighted approach assuming a fixed-effect model. SNPs absent in more than 2 European studies (European-only meta-analysis) were excluded from analysis.

We estimated that the combined sample provided 80% power to detect an odds ratio of 1.36 in the trans-ethnic meta-analysis at 5% MAF assuming a significance level (α) of 5×10^{-8}. The minimal detectable odds ratio at 80% power in the European-only meta-analysis was 1.48. Power calculations were obtained using the QUANTO software.

Extension to older onset stroke

We tested whether SNPs associated with young onset stroke were also associated with stroke in older populations through *in silico* look-ups in the METASTROKE consortium,[7] a consortium comprising 15 studies of predominantly older stroke cases. METASTROKE studies already included in either the Discovery or Follow-up stage of the early-onset stroke GWAS were excluded from this analysis. Study-specific GWAS results obtained from METASTROKE were provided to Y-C C for meta-analysis. Basic study design features and cohort characteristics of the METASTROKE studies have been published previously[7] and are summarized in the Supplementary Data and Supplemental Table V.

Association of SNPs at 10q25 with FSAP protein levels

A single locus on chromosome 10q25 near hyaluronan binding protein 2 (*HABP2*), the gene encoding Factor VII activating protease (FSAP), met genome-wide statistical significance in the joint meta-analysis. Because we considered *HABP2* to be a strong biological candidate gene for ischemic stroke, we tested for association of the index SNPs at the associated locus with variation in plasma FSAP levels, which were measured in relatively healthy participants of the Sahlgrenska Academy Study on Ischemic Stroke (SAHLSIS)[8] (see Supplemental Data). FSAP levels, measured by an FSAP-specific ELISA as previously described,[9] were logarithmically transformed prior to analysis. Analyses were conducted in normolipidemic control subjects only (n=125) because FSAP antigen and activity levels are increased in both ischemic stroke and hyperlipidemia.[9]

The statistical significance level was set at 0.05 and P-values were two-tailed. Statistical tests were performed using the SPSS Statistics software program.

Results

Genome-wide SNP association with early-onset ischemic stroke

The trans-ethnic meta-analysis included 13,439,215 SNPs having MAF>1% in at least 2 ethnic groups. Approximately 57%, 35% and 8% of the cases were from European, South-Asian and African ancestry, respectively (Table 1; Supplemental Table I). The mean age of

stroke cases ranged between 41 to 52 years. There was evidence for genomic deflation (λ=0.90; Supplemental Figure IB) under the random effect model but no evidence of genomic inflation or deflation under the fixed effect model (λ=1.007). The strongest association was with a cluster of 12 SNPs (P-values<5×10^{-6}) spanning a 22 kb region (chr10: 115042323-115064197bp) (Figure 1B; Supplemental Table VI). The associated cluster fell within an intergenic region between *TCF7L2* and *HABP2* on chromosome 10q25.3 as shown in Figure 2 (lead SNP rs11196288; OR=1.39, P=1.24×10^{-7}, effect allele=G, random effect model). Results of the leading SNP associations for the trans-ethnic analysis remained similar when analyzed using the Han and Eskin's Random Effects model (rs11196288; OR=1.40, P=1.25×10^{-7}) and fixed effect model (rs11196288; OR=1.39, P=1.24×10^{-7}).

The 10q25.3 locus was the only locus with P<5×10^{-6} and was taken forward for follow-up analysis. The index SNP (rs11196288) was not genotyped in the WTCCC Immunochip study. In meta-analysis of the two remaining Follow-up studies (n=502 cases and 2041 controls), this SNP showed significant evidence for significance (OR=1.83, P-value=0.03). In the combined analysis of both Discovery and Follow-up studies, the 10q25.3 locus was genome-wide significant with minimal heterogeneity between studies (rs11196288: between-study heterogeneity I^2=1.6%, P=0.42; OR=1.41, P=9.5×10^{-9} under the fixed effect model and P=1.5×10^{-8} under the random effect model; Supplemental Table VII). The direction of effect for the lead SNP, rs11196288, was consistent in 9 out of the 10 studies (Figure 3). Notably, the study with the inconsistent direction of effect (MILANO; OR=0.93, 95% CI=0.39–2.21, P=0.86), also had the lowest imputation quality score for this SNP (INFO=0.67), while the other 9 studies had excellent imputation quality (INFO ranging between 0.96 and 0.99).

The European-ancestry GWA analysis included 2,567 cases and 17,163 controls from 5 studies in the Discovery stage (CADISP, GEOS European ancestry, MILANO, SIFAP and WTCCC2-UK) and was based on a total of 10,537,953 SNPs with MAF>1% in the study population. Following meta-analysis and after excluding a single isolated SNP showing marginal evidence for association, a total of 21 SNPs from four loci showed suggestive associations with P-values<5×10^{-6} (Figure 1A; Supplemental Figure IA). However, none reached genome-wide significance (P-value<5×10^{-8}). There was no evidence for genomic inflation (λ=1.004). The four suggestively associated loci, each represented by the two most significant SNPs, are listed in Supplemental Table VII. Of note, one of these loci includes *Factor V* (*F5*), a well-known stroke candidate gene.[10, 11] Of the four loci identified in the European-only meta-analysis, none showed strong evidence for association in the Follow-up Stage (all Stage 2 P-values>0.22; Supplemental Table VII).

Secondary analyses of the associations between rs11196288 and ischemic stroke subtypes based on the 8 studies in the Discovery phase revealed nominally significant associations with all three major stroke subtypes under the random effect model: cardioembolic stroke (OR=1.34, P=0.02), large artery atherosclerotic stroke (OR=1.65, P=0.01) and small vessel stroke (OR=1.57, P=0.003). For large artery atherosclerotic stroke, there was substantial heterogeneity in effect size between studies (I^2=40%). In addition, rs11196288 showed a strong association with undetermined stroke, which accounted for 41% of all early-onset

stroke cases in the Discovery phase studies (OR=1.44, $P=3\times10^{-5}$). Evidence for association with stroke of other determined etiology could not be evaluated due to limited number of cases. A further genome-wide meta-analysis of stroke subtypes was performed as an exploratory analysis as it was extremely underpowered; this provided no evidence for subtype-specific associations (Supplemental Table VIII and Supplemental Figures II–III.)

We used the Ensemble Variant Effect Predictor tool to predict function of the index SNPs and SNPs in high LD with them. From these analyses, we identified two SNPs predicted to be regulatory region variants that were in high LD with the index SNP at 10q25. Rs1338423 (r^2=0.88 with rs11196288) and rs4918806 (r^2=1 with rs11196288) both fall in an open chromatin region. Rs4918806 additionally falls in a promoter flanking region and a CTCF binding site.

Associations of previously known stroke-relevant loci and the risk of early-onset ischemic stroke

Using the Discovery stage samples, we examined associations between early-onset ischemic stroke and 8 loci previously associated with ischemic stroke in GWAS of predominantly European older stroke cases.[3, 7, 12–18] We observed nominally significant associations (i.e., P<0.05) with early-onset ischemic stroke at four loci (*ALDH2*, *PITX2*, *CDKN2B-AS1* and *ABO*) in the European-only analyses (Supplemental Table IX). Interestingly, *ABO* was also associated with undetermined stroke, a predominant stroke subtype in early-onset stroke patients. In trans-ethnic meta-analysis, SNPs at *PITX2* and *ABO*, but not *CDKN2B-AS1* or *ALDH2*, showed nominal association with stroke (Supplemental Table IX).

Effect of the 10q25 locus on risk of ischemic stroke in older populations from the METASTROKE consortium

We further examined whether the two lead SNPs identified in our early-onset stroke GWAS were associated with the risk of ischemic strokes at older age of onset by performing an *in silico* look-up of association results from the METASTROKE consortium. Of the 15 discovery studies in METASTROKE, 13 had genotyped rs11196288 and rs4918806, a perfect proxy (r^2=1) for rs61872854. Five of these studies were already included in either the Discovery or Follow-up stage of the early-onset stroke GWAS, leaving 8 studies for the *in silico* look-up. Mean age of stroke in these studies ranged from 57.3 to 81.6 years old. Study level results were utilized because individual level data are not available through METASTROKE. No significant associations were observed for either rs11196288 (OR=1.08, 95% CI=0.96–1.22, P=0.18) or rs4918806 (OR=1.04, 95% CI=0.94–1.14, P=0.47) in the meta-analysis of the 8 studies. However, when ranking these studies by the mean age of stroke cases (Figure 4), the study with youngest mean age (i.e. ARIC, mean age onset=57.3 years) showed the strongest risk associated with the two SNPs (OR=1.51) while the study with oldest mean age (i.e. CHS, mean age onset=81.6 years) showed a non-significant inverse association with the two early-onset stroke associated SNPs (OR=0.67). Results remained similar when the one study with more than 30% of early-onset stroke cases (HPS) was removed from the analyses (rs11196288: OR=1.09, 95% CI=0.96–1.23, P=0.18; rs495366: OR=1.03, 95% CI=0.93–1.14, P=0.55).

Associations with plasma FSAP levels

Of the 12 SNPs at 10q25 showing suggestive associations with early-onset ischemic stroke, two (rs7906302 and rs1338423) had previously been genotyped in SAHLSIS and were tested for association with circulating FSAP levels (n=125 normolipidemic control subjects). The risk alleles associated with stroke susceptibility were significantly associated with higher FSAP levels at both loci (rs7906302: 15.1 vs. 11.7 μg/ml for (AC + CC) vs AA genotypes; P=0.026 and rs1338423: 15.6 vs. 11.7 μg/ml for AG vs AA genotype; P=0.012) (see Figure 5). The risk allele frequencies for rs7906302 (allele C) and rs1338423 (allele G) are 0.04 and 0.03 in SAHLSIS controls, respectively. Both SNPs are in high LD (r^2=0.88, Supplemental Table VI) with the index SNP rs11196288.

Discussion

Although a genetic predisposition to stroke is widely acknowledged, our understanding of genetic basis for ischemic stroke is limited. We have focused in this study on early-onset ischemic stroke, a form of stroke for which there may be a genetic enrichment. Our study included a total of 4,505 early-onset stroke cases and 21,968 controls from three ethnic groups, thus making it the largest genome-wide association study of early-onset stroke carried out to date by far.

Our analysis identified a novel locus at 10q25 strongly associated with early-onset ischemic stroke with an estimated ~1.40-fold increased risk associated with the risk variant. The 10q25 association showed consistent direction of effect in 9 out of the 10 studies, but achieved genome-wide significance only in the combined analyses of discovery and follow-up samples. Further replication of this locus in independent samples is warranted. Despite this limitation, this locus is of great interest due to its proximity to *HABP2*. *HABP2* encodes FSAP, an extracellular serine proteinase that cleaves urinary plasminogen activator, coagulation factor VII, and tissue factor pathway inhibitor, and helps regulate coagulation, tissue remodeling and inflammation.[19–22] Previous studies reported elevated levels of FSAP activity in patients with coronary artery disease[23] and ischemic stroke.[9] Rare genetic variants (e.g. the Marburg I polymorphism) in *HABP2* have also been associated with increased risk of deep venous thrombosis,[24] carotid stenosis,[25] and stroke.[26] Recent animal studies have also suggested that FSAP activity may modulate stroke-associated brain injury.[27]

Although the locus associated with early-onset ischemic stroke in our study is located in an intergenic region ~253 kb away from *HABP2*, several of the associated SNPs are predicted to have regulatory properties and associate with circulating levels of FSAP, implicating a possible mechanism leading to increased risk of early-onset stroke via *HABP2*. Interestingly, the previously known Marburg I polymorphism in *HABP2* (rs7080536), showed no association with early-onset stroke in our analysis (data not shown). Further studies are needed to identify the causal variant(s) tagged by these associated SNPs as well as their functional properties leading to the increased risk of early-onset stroke. In contrast to most of the previously known stroke-relevant loci, our data are consistent with this 10q25 locus being associated with multiple stroke subtypes, a finding consistent with the possible

thrombotic mechanism of *HABP2* that can lead to increased coagulation and therefore susceptibility across the different stroke subtypes.

Our study also suggests that the effects of the 10q25 locus may be specific to (or at least more pronounced in) early-onset stroke because the association was largely absent in METASTROKE studies, which consisted of predominantly older stroke cases. Unfortunately, we did not have access to large case-control cohorts consisting of stroke cases >60 years only, and therefore, we were unable to compare the allele frequency of rs11196288 between younger and older stroke cases directly to assess its effect on age of stroke onset. However, even within METASTROKE, there is a hint of an age-dependency with the locus effect appearing strongest in the study with the youngest mean age of onset (i.e. ARIC, mean age=57.3 years, OR=1.51) and weakest in the study with the oldest mean age of stroke onset (i.e., CHS, mean age=81.6 years, OR=0.67). An age-dependency for genetic effects would be parallel to prior reports that other risk factors may have large effects in younger versus older-onset stroke.[28] Further studies on possible interactions between genetic variants and environmental exposures are needed to elucidate how the genetic risk to stroke changes over the life span and/or in the presence of environmental challenges.

Despite the lack of association between 10q25 and risk of stroke at older age of onset, some of the previously reported stroke loci (identified via studying mainly older stroke cases[3, 7, 13–18]) showed nominal associations with early-onset stroke in our study, including *PITX2* with cardioembolic stroke, *CDKN2B-AS1* with large artery stroke, *ALDH2* with overall ischemic stroke and *ABO* with cardioembolic, large artery stroke and overall ischemic stroke. Our findings are thus consistent with early-onset and late-onset stroke sharing some of the same genetic susceptibility loci.

Our study was powered to detect an odds ratio of 1.36 for a SNP with MAF=0.05, which was slightly smaller than the effect size (OR=1.41) actually observed for the associated SNP. However, even including 4,500 early-onset stroke cases, we were unable to identify any subtype-specific stroke susceptibility loci in this early-onset stroke population using genome-wide association approach, possibly due to the small number of cases within each stroke subtype. Obtaining sufficient early-onset cases within each stroke subtype remains a significant challenge in future studies, and collaborative efforts will be needed to overcome this limitation. In the European GWAS analysis, we failed to replicate *F5*, a key protein in the coagulation pathway, as an early-onset stroke susceptibility locus. Previous evidence from candidate gene studies has long suggested that genetic variants of *F5* predispose to stroke, the most conspicuous example being the Factor V Leiden (FVL) variant.[10, 11, 29] However, a recent large-scale meta-analysis suggested that the reported association between *FVL* and early-onset IS was more pronounced among studies where cases were selected on the basis of having cryptogenic stroke or recruited from a subset of patients referred for a thrombophilic work-up and the *FVL*-stroke association is much smaller among studies with unselected cases.[10] This may explain the lack of consistent results observed in our GWAS because cases included in our analyses were unselected and thus the genetic effect may be less prominent. Given the complexity of the study populations included in this study,

significant phenotypic heterogeneity likely remains among stroke cases even within this early onset stroke group.

In summary, we have identified a novel locus at 10q25 associated with all ischemic strokes in young adult population. This locus is located near *HABP2*, which encodes an extracellular serine protease involved in coagulation, fibrinolysis and even inflammatory pathways, suggesting a plausible biological mechanism leading to increased risk of stroke. This locus did not appear to have a significant effect in older cases, indicating this may be a genetic susceptibility locus for early-onset stroke. Further replication of the 10q25 locus and additional studies investigating the potential age of onset effect are needed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Yu-Ching Cheng, PhD[1,2], Tara M. Stanne, PhD[3], Anne-Katrin Giese, MD[4], Weang Kee Ho, PhD[5], Matthew Traylor, PhD[6], Philippe Amouyel, MD, PhD[7], Elizabeth G. Holliday, PhD[8], Rainer Malik, PhD[9], Huichun Xu, MD, PhD[2], Steven J. Kittner, MD[1,2], John W. Cole, MD[1,2], Jeffrey R. O'Connell, PhD[2], John Danesh, DPhil, FRCP[6,10], Asif Rasheed, MBBS[11], Wei Zhao, MSc[12], Stefan Engelter, MD[13], Caspar Grond-Ginsbach, PhD[14], Yoichiro Kamatani, MD, PhD[15,16], Mark Lathrop, PhD[17,18], Didier Leys, MD, PhD[19], Vincent Thijs, MD, PhD[20,21,22], Tiina M. Metso, MD, PhD[23], Turgut Tatlisumak, MD, PhD[23], Alessandro Pezzini, MD[24], Eugenio A. Parati, MD[25], Bo Norrving, MD, PhD[26], Steve Bevan, PhD[6], Peter M Rothwell, MD, PhD, FRCP[27], Cathie Sudlow, DPhil, FRCP[28], Agnieszka Slowik, MD, PhD[29], Arne Lindgren, MD, PhD[30,31], Matthew R Walters, MD[32], WTCCC-2 Consortium, Jim Jannes, PhD[33], Jess Shen, MSc[34], David Crosslin, PhD[35], Kimberly Doheny, PhD[36], Cathy C. Laurie, PhD[35], Sandip M. Kanse, PhD[37], Joshua C. Bis, PhD[35], Myriam Fornage, PhD[38], Thomas H. Mosley, PhD[39], Jemma C. Hopewell, PhD[40], Konstantin Strauch, PhD[9,41], Martina Müller-Nurasyid, MD, PhD[9,41,42,50], Christian Gieger, PhD[41], Melanie Waldenberger, PhD, MPH[41], Annette Peters, PhD[41,42], Christine Meisinger, MD[41], M. Arfan Ikram, MD, PhD[43], WT Longstreth Jr, MD, MPH[35], James F. Meschia, MD[44], Sudha Seshadri, MD[45], Pankaj Sharma, PhD, MD[46], Bradford Worrall, MD, MSc[47], Christina Jern, MD, PhD[3], Christopher Levi, MBBS[48], Martin Dichgans, MD[9], Giorgio B. Boncoraglio, MD[25], Hugh S. Markus, MD[6], Stephanie Debette, MD, PhD[6,19,49], Arndt Rolfs, MD[4], Danish Saleheen, MBBS, PhD[6,11,12], and Braxton D. Mitchell, PhD[1,2]

Affiliations

[1]Veterans Affairs Maryland Health Care System, Baltimore, Maryland [2]University of Maryland School of Medicine, Baltimore, Maryland [3]The University of Gothenburg, Gothenburg, Sweden [4]University of Rostock, Rostock, Germany [5]University of Nottingham Malaysia Campus, Selangor Darul Ehsa, Malaysia [6]University of Cambridge, Cambridge, UK [7]Institut Pasteur de Lille, F-59000 Lille, France

[8]University of Newcastle, Australia [9]Ludwig-Maximilians-Universität München, Munich, Germany [10]Wellcome Trust Sanger Institute, Cambridge, UK [11]Center for Non-communicable Diseases, Karachi, Pakistan [12]University of Pennsylvania, US [13]Basel University Hospital, Switzerland [14]Heidelberg University Hospital, Germany [15]Centre d'Étude du Polymorphisme Humain, Paris, France [16]RIKEN Center for Integrative Medical Sciences, Yokohama, Japan [17]National Genotyping Center, Evry, France [18]Genome Quebec, McGill University, Montreal, Canada [19]Lille University Hospital, France [20]KU Leuven - University of Leuven, B-3000 Leuven, Belgium [21]Vesalius Research Center, VIB, B-3000 Leuven, Belgium [22]University Hospitals Leuven, B-3000 Leuven, Belgium [23]Helsinki University Central Hospital, Helsinki, Finland [24]Università degli Studi di Brescia, Brescia, Italy [25]Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy [26]University of Lund, Sweden [27]University of Oxford, John Radcliffe Hospital [28]University of Edinburgh, Edinburgh, UK [29]Jagiellonian University Medical College, Krakow, Poland [30]Lund University, Lund, Sweden [31]Skåne University Hospital, Lund, Sweden [32]University of Glasgow, Glasgow, UK [33]University of Adelaide, Australia [34]Mount Sinai Hospital, Toronto, Ontario [35]University of Washington, Seattle, Washington [36]Johns Hopkins University School of Medicine, Baltimore, Maryland [37]University of Oslo, Oslo, Norway [38]The University of Texas Health Science Center at Houston, Houston, Texas [39]University of Mississippi Medical Center, Jackson, MS [40]University of Oxford, Oxford, UK [41]Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany [42]Partner site Munich Heart Alliance, Munich, Germany [43]Erasmus University Medical Center Rotterdam, Netherlands [44]Mayo Clinic, Florida [45]Boston University School of Medicine, Boston, MA [46]Royal Holloway, University of London, UK [47]University of Virginia, Charlottesville, Virginia [48]John Hunter Hospital, Newcastle, Australia [49]University Hospital of Bordeaux, France [50]University Hospital Grosshadern, Munich, Germany

Acknowledgments

Funding Sources

The **Atherosclerotic Risk in Community (ARIC):** NIH contracts: HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, HHSN268201100012C, U01HG004402; HHSN268200625226C, N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, and U01-HL096917. NIH grants: R01HL087641, R01HL59367, R01HL086694, R01HL087641, UL1RR025005, and HL093029.

The **Australian Stroke Genetics Collaborative (ASGC)** grants: the Australian National Health and Medical Research Council (project 569257) and the Australian National Heart Foundation (grant G-04S-1623). Elizabeth Holliday supported by a fellowship (100071) from the Australian Heart Foundation and National Stroke Foundation.

Bio-Repository of DNA in Stroke (BRAINS) support: the Henry Smith Charity, the UK–India Education Research Institutive from the British Council, and a Senior Fellowship from the UK Department of Health awarded to Dr. Pankaj Sharma.

The **Cervical Artery Dissection and Ischemic Stroke Patients (CADISP) study** Institutional: Inserm, Lille 2 University, Institut Pasteur de Lille, and Lille University Hospital. Funding: the ERDF (FEDER funds) and Région

Nord-Pas de Calais in the frame of Contrat de Projets Etat-Region 2007–2013 Région Nord-Pas-de-Calais - Grant N°09120030, Centre National de Genotypage, Emil Aaltonen Foundation, Paavo Ilmari Ahvenainen Foundation, Helsinki University Central Hospital Research Fund, Helsinki University Medical Foundation, Päivikki and Sakari Sohlberg Foundation, Aarne Koskelo Foundation, Maire Taponen Foundation, Aarne and Aili Turunen Foundation, Lilly Foundation, Alfred Kordelin Foundation, Finnish Medical Foundation, Orion Farmos Research Foundation, Maud Kuistila Foundation, the Finnish Brain Foundation, Biomedicum Helsinki Foundation, Projet Hospitalier de Recherche Clinique Régional, Fondation de France, Génopôle de Lille, Adrinord, Basel Stroke-Funds, Käthe-Zingg-Schwichtenberg-Fonds of the Swiss Academy of Medical Sciences, Swiss Heart Foundation. Stéphanie Debette is a recipient of a “Chaire d’Excellence Junior” grant from the Agence Nationale de la Recherche and is supported by a grant from the Fondation Leducq. Vincent Thijs supported by a Fundamental Clinical Research Fellowship from FWO Flanders.

The **Cardiovascular Health Study (CHS)** NIH contracts: HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086. NIH grants: U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, R01AG023629, and R01DK063491. Genotyping: the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124.

The **deCODE CAD/MI Study** NIH grant: R01HL089650. deCODE Genetics supported, in part, through a grant from the European Community’s Seventh Framework Programme (FP7/2007–2013), the ENGAGE project grant agreement HEALTH-F4-2007 to 201413.

The **Genetics of Early-onset Stroke (GEOS) study** NIH grants: U01-HG004436, P30-DK072488, U01-NS069208, R01-NS45012, U01-NS069208, U01-HG004438, U01-HG004446, and the Baltimore Geriatrics Research, Education, and Clinical Center of the Department of Veterans Affairs. Yu-Ching Cheng was supported by a Career Development Award from Department of Veterans Affairs.

Heart Protection Study (HPS; ISRCTN48489393**):** the UK Medical Research Council, British Heart Foundation (BHF), Merck & Co (manufacturers of simvastatin), and Roche Vitamins Ltd (manufacturers of vitamins). Genotyping supported by a grant to Oxford University and Centre National de Génotypage from Merck & Co. Dr. Hopewell supported by the British Heart Foundation (FS/14/55/30806).

The **Heart and Vascular Health Study (HVH)** NIH grants R01-HL085251 and R01-HL073410.

The **Ischemic Stroke Genetics Study (ISGS)/Siblings With Ischemic Stroke Study (SWISS)** NIH grants: R01-NS42733 and R01-NS39987 and NIH intramural project Z01-AG000954. ISGS/SWISS used samples and clinical data from the NIH-NINDS Human Genetics Resource Center DNA and Cell Line Repository (http://ccr.coriell.org/ninds), human subjects protocol numbers 2003-081 and 2004-147. Controls for ISGS/SWISS obtained from the Baltimore Longitudinal Study of Aging with support from the NIA Intramural Research Program (Z01-AG000015-50, human subjects protocol number 2003-078).

The MILANO study supported by Annual Research Funding of the Italian Ministry of Health (Grants: RC-2007/LR6, RC-2008/LR6; RC-2009/LR8; RC-2010/LR8).

The Sahlgrenska Academy Study of Ischemic Stroke (SAHLSIS) the Swedish Research Council, the Swedish state and the Swedish Heart and Lung Foundation. Sandip Kanse acknowledges funding from Behring Roentgen Stiftung, Deutscheforschungsgemeinschaft and Helse Sør-Øst.

Stroke in Young Fabry Patients (SIFAP) NIH grant: U01-HG004436. Control subjects for SIFAP provided by the KORA Study (see below, WTCCC2-Munich).

Risk Assessment of Cerebrovascular Events Study (RACE). Grants: to the University of Cambridge from the Wellcome Trust, British Heart Foundation, UK Medical Research Council, Pfizer, Novartis, and Merck and NIH Grant U01-HG004436.

The **Rotterdam study** Institutional: The Netherlands Organization of Scientific Research (175.010.2005.011), the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research Netherlands Consortium for Healthy Ageing (050-060-810), Nederlandse Hartstichting (2009B102), the Erasmus Medical Center and Erasmus University, Rotterdam, the Netherlands Organization for Health Research and Development, the Research Institute for Diseases in the Elderly, the Ministry of Education, Culture, and Science, the Ministry for Health, Welfare, and Sports, the European Commission, and the Municipality of Rotterdam to the Rotterdam Study.

The Wellcome Trust Case-Control Consortium 2 (WTCCC2)

WTCCC2-UK: the Wellcome Trust (085475/B/08/Z and 085475/Z/08/Z and WT084724MA), The Stroke Association, the Medical Research Council (grants WT095219MA and G1001799), Dunhill Medical Trust,

National Institute of Health Research (NIHR), the NIHR Biomedical Research Centre, the Binks Trust, the Scottish Funding Council and the Chief Scientist Office. PMR has a Wellcome Trust Senior Investigator Award and an NIHR Senior Investigator Award and CS has a Wellcome Trust clinician scientist award.

WTCCC2-Munich: Grants: the German Federal Ministry of Education and Research in the context of the e:Med program (e:AtheroSysMed) and the FP7 European Union project CVgenes-AT-target (261123) to Martin Dichgans and from the Vascular Dementia Research Foundation. The KORA study was initiated and financed by the Helmholtz Zentrum München-German Research Center for Environmental Health. Additional KORA support from the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

WTCCC Immunochip: supported by the WTCCC2. Lund Stroke Register: the Swedish Research Council, The Swedish Heart-Lung Foundation, Region Skåne, the Freemasons Lodge of Instruction EOS in Lund, King Gustaf V and Queen Victoria's Foundation, Lund University, the Swedish Stroke Association, Region Skåne Competence Centre (RSKC Malmö), and Labmedicin Skåne, University and Regional Laboratories Region Skåne, Sweden. Investigator support: Stroke Association Project Grant TSA-2013/01 (Matthew Traylor), the National Research Leading Center, Jagiellonian University, Krakow, Poland (Agnieszka Slowik), NIHR Senior Investigator award (Hugh Markus).

References

1. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Blaha MJ, et al. Heart disease and stroke statistics--2014 update: A report from the American Heart Association. Circulation. 2014; 129:e28–e292. [PubMed: 24352519]
2. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. Stroke. 2012; 43:3161–3167. [PubMed: 23042660]
3. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, et al. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. Nat Genet. 2012; 44:1147–1151. [PubMed: 22941190]
4. Cheng YC, Cole JW, Kittner SJ, Mitchell BD. Genetics of ischemic stroke in young adults. Circ Cardiovasc Genet. 2014; 7:383–392. [PubMed: 24951665]
5. Falcone GJ, Malik R, Dichgans M, Rosand J. Current concepts and clinical applications of stroke genetics. Lancet Neurol. 2014; 13:405–418. [PubMed: 24646874]
6. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009; 5:e1000529. [PubMed: 19543373]
7. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the Metastroke collaboration): A meta-analysis of genome-wide association studies. Lancet Neurol. 2012; 11:951–962. [PubMed: 23041239]
8. Jood K, Ladenvall C, Rosengren A, Blomstrand C, Jern C. Family history in ischemic stroke before 70 years of age: The Sahlgrenska Academy Study on ischemic stroke. Stroke. 2005; 36:1383–1387. [PubMed: 15933254]
9. Hanson E, Kanse SM, Joshi A, Jood K, Nilsson S, Blomstrand C, et al. Plasma factor VII-activating protease antigen levels and activity are increased in ischemic stroke. J Thromb Haemost. 2012; 10:848–856. [PubMed: 22409238]
10. Hamedani AG, Cole JW, Mitchell BD, Kittner SJ. Meta-analysis of factor V leiden and ischemic stroke in young adults: The importance of case ascertainment. Stroke. 2010; 41:1599–1603. [PubMed: 20616326]
11. Xin XY, Song YY, Ma JF, Fan CN, Ding JQ, Yang GY, et al. Gene polymorphisms and risk of adult early-onset ischemic stroke: A meta-analysis. Thromb Res. 2009; 124:619–624. [PubMed: 19660787]
12. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. N Eng J Med. 2009; 360:1718–1728.
13. Kilarski LL, Achterberg S, Devan WJ, Traylor M, Malik R, Lindgren A, et al. Meta-analysis in more than 17,900 cases of ischemic stroke reveals a novel association at 12q24.12. Neurology. 2014; 83:678–685. [PubMed: 25031287]

14. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. Ann Neurol. 2008; 64:402–409. [PubMed: 18991354]
15. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, et al. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. Nat Genet. 2009; 41:876–878. [PubMed: 19597491]
16. Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, et al. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. Ann Neurol. 2009; 65:531–539. [PubMed: 19475673]
17. Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, et al. International Stroke Genetics Consortium, Wellcome Trust Case Control Consortium. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. Nat Genet. 2012; 44:328–333. [PubMed: 22306652]
18. Williams FM, Carter AM, Hysi PG, Surdulescu G, Hodgkiss D, Soranzo N, et al. Ischemic stroke is associated with the ABO locus: The Euroclot study. Ann Neurol. 2013; 73:16–31. [PubMed: 23381943]
19. Kanse SM, Etscheid M. Factor VII activating protease. Single nucleotide polymorphisms light the way. Hamostaseologie. 2011; 31:174–176. 177–178. [PubMed: 21655671]
20. Parahuleva MS, Maj R, Holschermann H, Parviz B, Abdallah Y, Erdogan A, et al. Regulation of monocyte/macrophage function by factor vii activating protease (fsap). Atherosclerosis. 2013; 230:365–372. [PubMed: 24075769]
21. Romisch J, Feussner A, Vermohlen S, Stohr HA. A protease isolated from human plasma activating factor VII independent of tissue factor. Blood Coagul Fibrinolysis. 1999; 10:471–479. [PubMed: 10636458]
22. Kannemeier C, Feussner A, Stohr HA, Weisse J, Preissner KT, Romisch J. Factor VII and single-chain plasminogen activator-activating protease: Activation and autoactivation of the proenzyme. Eur J Biochem. 2001; 268:3789–3796. [PubMed: 11432747]
23. Parahuleva MS, Holschermann H, Zandt D, Pons-Kuhnemann J, Parviz B, Weiskirchen R, et al. Circulating factor VII activating protease (fsap) is associated with clinical outcome in acute coronary syndrome. Circ J. 2012; 76:2653–2661. [PubMed: 22850287]
24. Ahmad-Nejad P, Dempfle CE, Weiss C, Bugert P, Borggrefe M, Neumaier M. The G534E-polymorphism of the gene encoding the factor VII-activating protease is a risk factor for venous thrombosis and recurrent events. Thromb Res. 2012; 130:441–444. [PubMed: 22421107]
25. Willeit J, Kiechl S, Weimer T, Mair A, Santer P, Wiedermann CJ, et al. Marburg i polymorphism of factor VII--activating protease: A prominent risk predictor of carotid stenosis. Circulation. 2003; 107:667–670. [PubMed: 12578864]
26. Trompet S, Pons D, Kanse SM, de Craen AJ, Ikram MA, Verschuren JJ, et al. Factor VIIi activating protease polymorphism (G534E) is associated with increased risk for stroke and mortality. Stroke Res Treat. 2011; 2011:424759. [PubMed: 21789270]
27. Joshi AU, Orset C, Engelhardt B, Baumgart-Vogt E, Gerriets T, Vivien D, et al. Deficiency of factor VII activating protease alters the outcome of ischemic stroke in mice. Eur J Neurosci. 2015; 41:965–975. [PubMed: 25615590]
28. Asplund K, Karvanen J, Giampaoli S, Jousilahti P, Niemela M, Broda G, et al. Relative risks for stroke by age, sex, and population based on follow-up of 18 European populations in the Morgam project. Stroke. 2009; 40:2319–2326. [PubMed: 19520994]
29. Hamedani AG, Cole JW, Cheng Y, Sparks MJ, O'Connell JR, Stine OC, et al. Factor V leiden and ischemic stroke risk: The genetics of early onset stroke (GEOS) study. J Stroke Cerebrovasc Dis. 2013; 22:419–423. [PubMed: 22100829]

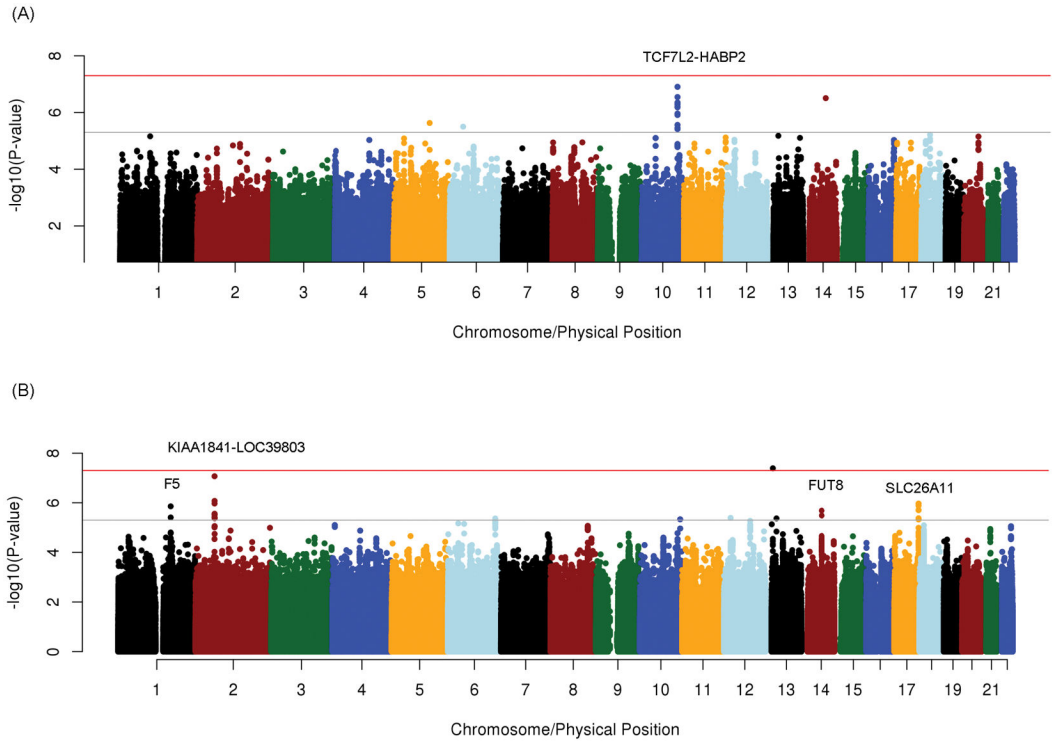

Figure 1.

Genome-wide association results of early-onset ischemic stroke based on (A) Trans-ethnic meta-analysis and (B) European-only meta-analysis. Red line: $P=5\times10^{-8}$; grey line: $P=5\times10^{-6}$.

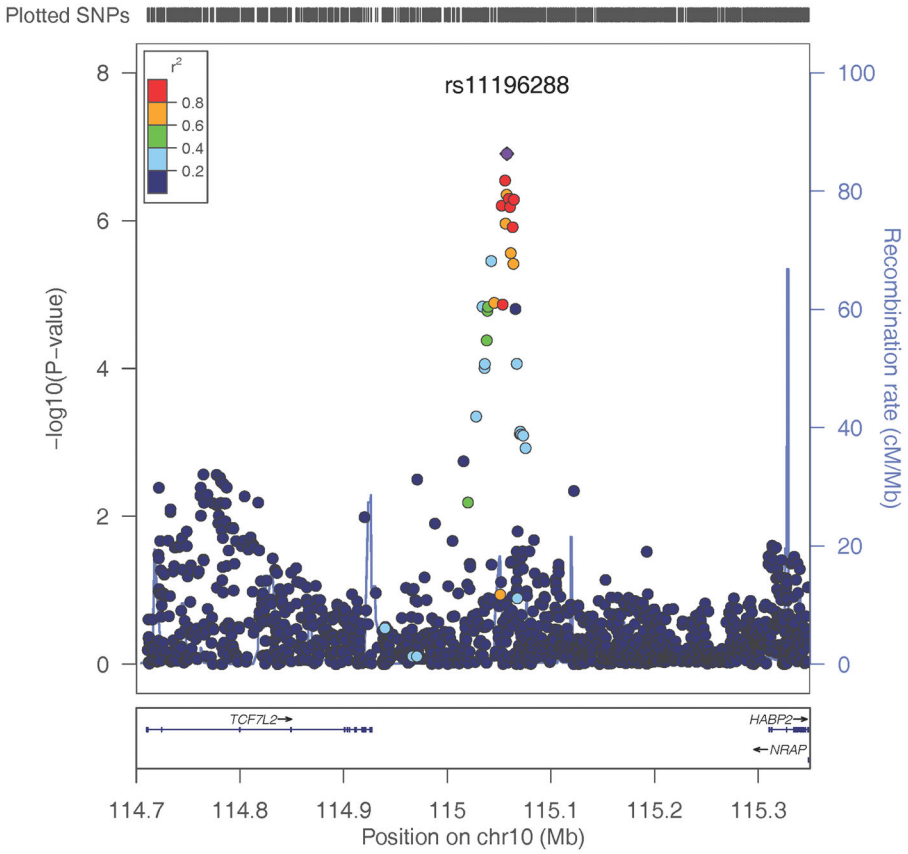

Figure 2.

Regional plot of the chr10q25.3 locus from trans-ethnic association analysis in the Discovery stage. Recombination rate was based on 1000 Genomes EUR data. Plot was generated using LocusZoom software.

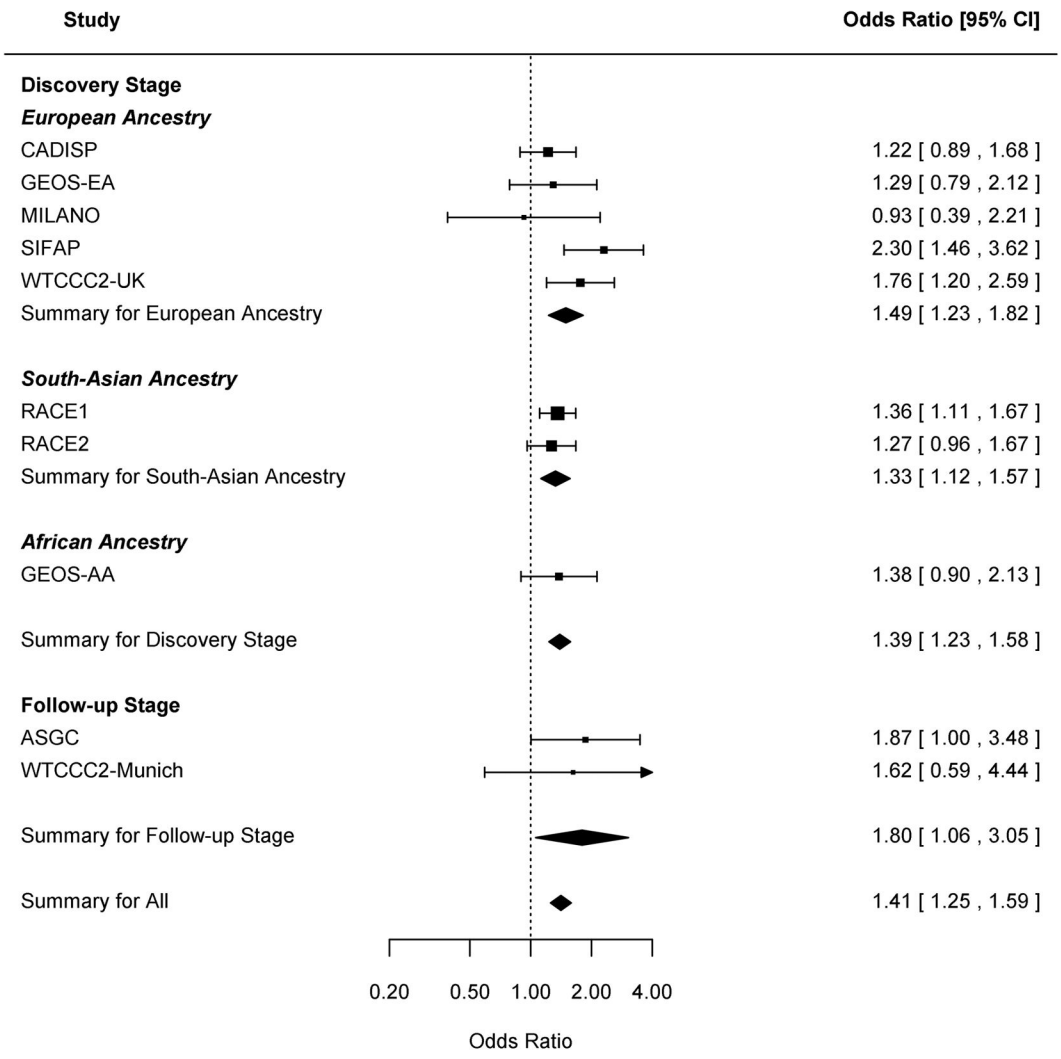

Figure 3.

Forest plot of rs11196288 at 10q25.3, including Discovery and Follow-up studies, in trans-ethnic analysis.

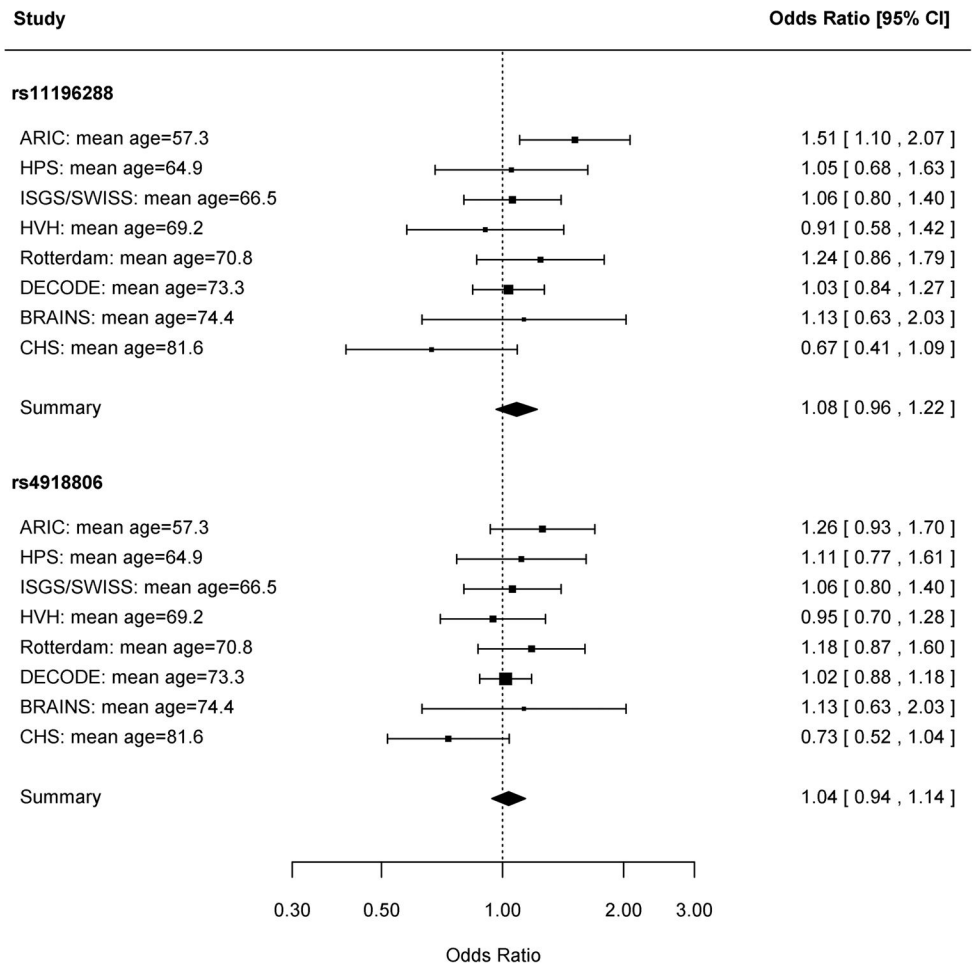

Figure 4.

Associations of rs11196288 and rs4918896 at 10q25.3 with risk of stroke in METASTROKE studies, by mean age of stroke cases.

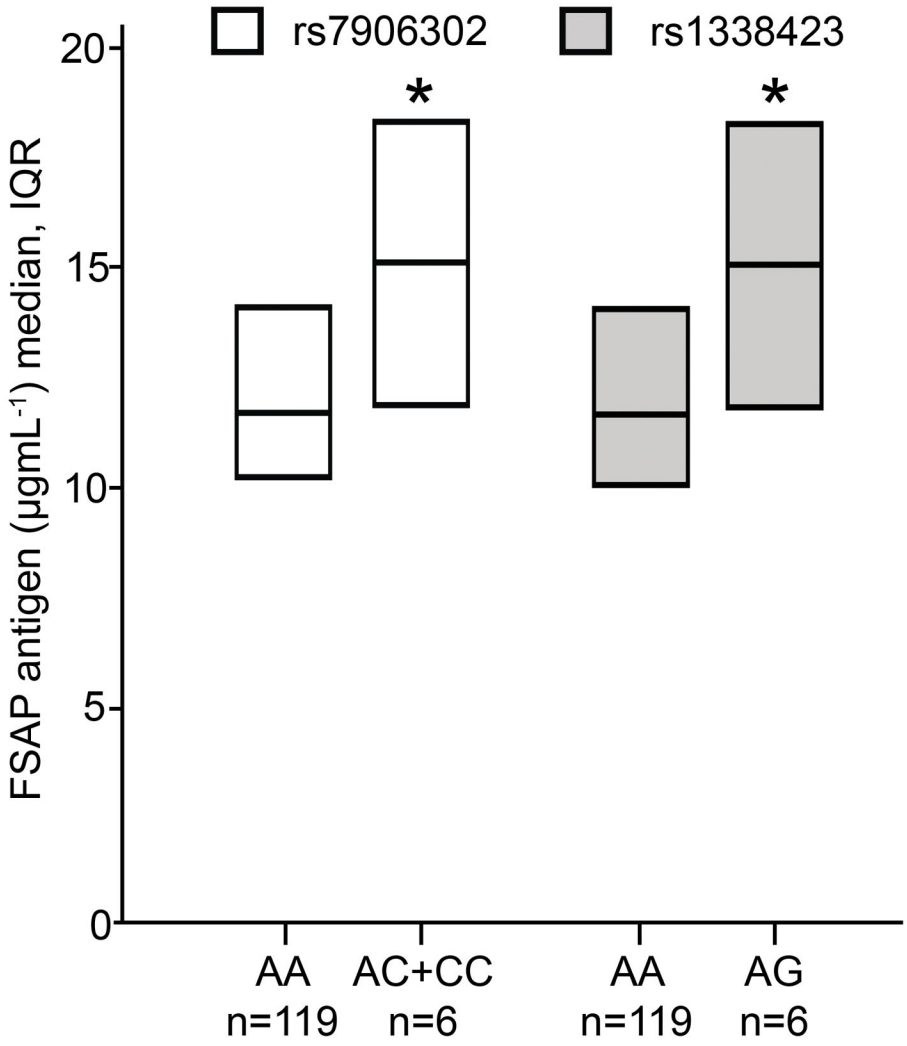

Figure 5.

Plasma FSAP levels are associated with rs7906302 and rs1338423 in normolipidemic controls in SAHLSIS. Median and interquartile range (IQR) of FSAP according to genotype in controls. Differences in logFSAPag levels were calculated with Student's t-test. * P<0.05

Table 1

Characteristics of Participating Studies

	Cases			Controls					
Study	**Subjects, n**	**Age, mean (SD)**	**Male, n (%)**	**Subjects, n**	**Age, mean (SD)**	**Male, n (%)**	**External control**	**Ancestry**	**Country**
Stage 1: Discovery Stage									
CADISP	555	43.73 (9.9)	339 (61.1)	9259	N/A	N/A	No	EA	Belgium, France, Germany, Italy, Switzerland and Finland
GEOS EA	448	41.0 (7.0)	275 (61.4)	498	39.5 (6.7)	282 (56.6)	No	EA	USA
GEOS AA	381	41.9 (6.8)	207 (54.3)	352	40.0 (6.8)	196 (55.7)	No	AA	USA
MILANO	201	45.0 (10.4)	120 (60.9)	407	50.8 (8.1)	357 (87.8)	No	EA	Italy
RACE 1	1218	50.1 (9.9)	638 (52.4)	1158	51.9 (7.9)	613 (53)	PROMIS	South Asian	Pakistan
RACE 2	339	50.2 (9.2)	272 (80.4)	3295	60.9 (13.2)	1838 (55.8)	PROMIS	South Asian	Pakistan
SIFAP	981	41.7 (7.4)	599 (61.1)	1824	55.2 (11.6)	899 (49.3)	KORA	EA	Germany
WTCCC2-UK	382	51.9 (7.3)	228 (59.7)	5175	52	2611 (50.5)	British Birth Cohort & UK Blood Service Control	EA	UK
Total	4505			21968					
Stage 2: Follow-up Stage									
ASGC	227	52.3 (8.7)	151 (66.5)	1244	66.3 (7.5)	625 (50.2)	No	EA	Australia
WTCCC2-Munich	275	48.8 (8.5)	188 (68.4)	797	62.7(10.9)	410 (51.4)	KORA	EA	Germany
Immunochip	501	48.7 ~ 53.4.	58% ~ 73%	5704	N/A	27% ~ 50%	No	EA	Belgium, Germany, Poland, Sweden and UK
Total	1003			7745					