DNA methylation analyses identify an intronic ZDHHC6 locus associated with time to recurrent stroke in the Vitamin Intervention for Stroke Prevention (VISP) clinical trial

Nicole M. Davis Armstrong1, Wei-Min Chen2,3, Fang-Chi Hsu4, Michael S. Brewer1, Natalia Cullell5,6, Israel Fernández-Cadenas5,6, Stephen R. Williams7, Michèle M. Sale2,3, Bradford B. Worrall5,7, Keith L. Keene1,8*

1 Department of Biology, East Carolina University, Greenville, NC, United States of America, 2 Center for Public Health Genomics, University of Virginia, Charlottesville, VA, United States of America, 3 Department of Public Health Sciences, University of Virginia, Charlottesville, VA, United States of America, 4 Department of Biostatistics and Data Science, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, United States of America, 5 Stroke Pharmacogenomics and Genetics, Fundació Doccència i Recerca Mutua Terrassa, Hospital Universitari Mutua de Terrassa, Terrassa, Barcelona, Spain, 6 Stroke Pharmacogenomics and Genetics, Sant Pau Institute of Research, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, 7 Department of Neurology, University of Virginia, Charlottesville, VA, United States of America, 8 Center for Health Disparities, Brody School of Medicine, East Carolina University, Greenville, NC, United States of America

* keenek@ecu.edu

Abstract

Aberrant DNA methylation profiles have been implicated in numerous cardiovascular diseases; however, few studies have investigated how these epigenetic modifications contribute to stroke recurrence. The aim of this study was to identify methylation loci associated with the time to recurrent cerebro- and cardiovascular events in individuals of European and African descent. DNA methylation profiles were generated for 180 individuals from the Vitamin Intervention for Stroke Prevention (VISP) clinical trial using Illumina HumanMethylation 450K BeadChip microarrays, resulting in beta values for 470,871 autosomal CpG sites. Ethnicity-stratified survival analyses were performed using Cox Proportional Hazards regression models for associations between each methylation locus and the time to recurrent stroke or composite vascular endpoint. Results were validated in the Vall d’Hebron University Hospital cohort from Barcelona, Spain. Network analyses of the methylation loci were generated using weighted gene coexpression network analysis. Primary analysis identified four significant loci, cg04059318, ch.2.81927627R, cg03584380, and cg24875416, associated with time to recurrent stroke. Secondary analysis identified three loci, cg00076998, cg16758041, and cg02365967, associated with time to composite vascular event. Results were validated in the Vall d’Hebron University Hospital cohort from Barcelona, Spain. Network analyses of the methylation loci were generated using weighted gene coexpression network analysis. Primary analysis identified four significant loci, cg04059318, ch.2.81927627R, cg03584380, and cg24875416, associated with time to recurrent stroke. Secondary analysis identified three loci, cg00076998, cg16758041, and cg02365967, associated with time to composite vascular endpoint. Locus cg03584380, which is located in an intron of ZDHHC6, was replicated in the Vall d’Hebron University Hospital cohort. The results from this study implicate the degree of methylation at cg03584380 is associated with the time of recurrence for stroke or composite vascular events across two ethnically diverse groups. Furthermore, modules of loci were associated with clinical traits and blood biomarkers including previous number of strokes, prothrombin fragments 1 + 2, thrombomodulin, thrombin-antithrombin complex, triglyceride levels, and tissue
plasminogen activator. Ultimately, these loci could serve as potential epigenetic biomarkers that could identify at-risk individuals in recurrence-prone populations.

Introduction

Ischemic strokes account for 87% of all strokes and are heterogeneous, multifactorial diseases comprising genetic and environmental contributions. Of the approximately 795,000 incident stroke cases annually in the United States, approximately 25% are recurrent events [1]. Individuals experiencing an ischemic stroke are at high risk of having a recurrent stroke or developing cardiovascular diseases including myocardial infarction (MI), coronary heart disease, or vascular death. Only a proportion of the estimated 37.9% heritability for ischemic stroke [2] is accounted for from genetic variants identified in genome-wide association studies, suggesting other mechanisms, such as epigenetic modifications, could comprise some of the remaining heritability for ischemic stroke and stroke recurrence risk.

Epigenetics, such as DNA methylation, refer to chemical modifications of DNA structure that can be maintained over cellular generations [3] and serve to propagate cellular memory [4]. Abnormal DNA methylation patterns have been implicated in a number of cardiovascular diseases [5, 6]; however, few studies have investigated these epigenetic contributions to stroke recurrence [7–9]. The aim of this study was to elucidate DNA methylation loci associated with the time to vascular events, including recurrent stroke, MI, and death. Single locus and comprehensive loci networks were analyzed in 180 individuals from the Vitamin Intervention for Stroke Prevention (VISP) clinical trial and further validated in the independent Vall d’Hebron University Hospital cohort. Findings from this study support the utility of epigenetic marks as potential biomarkers and may lead to improved prognosis or prevention of recurrent stroke and cardiovascular events.

Methods

Ethics statement

The institutional review boards (IRBs) of Wake Forest University School of Medicine, the University of North Carolina at Chapel Hill School of Medicine, and individual recruitment sites approved the VISP clinical trial study protocol. All VISP participants provided written, informed consent. A subset of 2,100 participants agreed to be included in subsequent genetic studies. IRB approval from the University of Virginia and East Carolina University was obtained for the genetic and epigenetic components.

Discovery cohort: Vitamin Intervention for Stroke Prevention (VISP) clinical trial

VISP was a multi-centered, double-blinded, randomized and controlled clinical trial that enrolled participants aged 35 years or older with baseline homocysteine levels at or above the 25th percentile and was designed to determine whether pyridoxine (B6), cyanocobalamin (B12), and folic acid (B9) supplementation reduced recurrent cerebral infarction, MI, or fatal coronary heart disease (CHD) [10]. Participants were enrolled within 120 days of suffering a non-disabling cerebral infarction, assigned a daily B vitamin high-dose or low-dose formulation, and followed for two years. While ischemic subtype was not adjudicated, based on inclusion/exclusion criteria [10], the VISP enrollment stroke was most likely a small-vessel (lacunar) infarct.
A subset of 2,100 VISP participants consented for inclusion in genetic studies of which, methylation data were generated for 204 participants. Upon quality control (QC), the methylation profiles of 180 individuals were used in subsequent analyses, including 76 individuals of African descent (AFR) and 104 of European descent (EUR). Individual level genetics and epigenetics data is considered sensitive controlled data and cannot be shared publicly, as specified by the IRBs of Wake Forest University School of Medicine, the University of North Carolina at Chapel Hill School of Medicine, and the University of Virginia School of Medicine, along with the NIH Data Access Committee. Controlled-access data can only be obtained if a user has been authorized by the appropriate Data Access Committee. The individual level Genomics and Randomized Trials Network (GARNET) VISP data are available in the database of Genotypes and Phenotypes (dbGaP) (Accession: phs000343.v3.p1) and can be requested through the dbGaP Authorized Access System (https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?page=login). The authors will also share the data on request.

**Replication cohort: The Vall d’Hebron University Hospital cohort**

From a cohort of 1,900 patients with stroke from Vall d’Hebron University Hospital (Barcelona, Spain), 28 subjects with composite vascular recurrence events were selected. Of these participants, 18 had a recurrent ischemic stroke. Composite vascular recurrence was described as new ischemic stroke, MI, peripheral vascular disease, or cardiovascular death and was detected through telephone interviews every three months or direct clinical visit. The epigenome-wide methylation profiles were generated using Illumina Infinium 450k BeadChip microarrays. The profiles were processed in a single working batch and preprocessing, correction, normalization steps, and QC was performed using R, as previously described [8]. The epigenetics and genetics summary data used for replication will be shared for research studies on approval of the principal investigators of the GRECOS (Genotyping Recurrence Risk of Stroke) cohort.

**Stroke recurrence and composite vascular event definition**

The primary vascular endpoint analyzed in the current study was an incident recurrent stroke during the VISP trial (VISP recurrent stroke). VISP recurrent stroke was defined as an acute neurological ischemic event, requiring a sudden onset of symptoms lasting at least 24 hours, as reported on the Follow-up Stroke Symptoms (FSS) form or as determined by the VISP Endpoint Review Committee [10].

A composite vascular event was examined as a secondary endpoint and was defined as fatal coronary heart disease, a non-fatal hospitalized MI and resuscitation for cardiac collapse, coronary bypass surgery, coronary angioplasty, or a VISP recurrent cerebrovascular event.

**Methylation data generation**

Genomic DNA was extracted from whole blood samples of the VISP participants, as previously described [9]. Briefly, the DNA was denatured and bisulfite-converted using Zymo EZ DNA Methylation Kits (Zymo Research Corp., Irvine, CA). Illumina Infinium Human Methylation450k BeadChip microarrays were used to interrogate CpG sites across the genome and the resulting intensity files were analyzed using GenomeStudio. Beta scores were generated as the ratio of methylated intensities divided by the sum of the methylated and unmethylated intensities (i.e. the proportion of total signal from the methylation-specific probe). Probes were removed if they failed to hybridize (detection p>5%) or if they were located on sex chromosomes. The filtered beta scores underwent stratified quantile normalization with the minfi package in R [11–13]. A total of 470,871 autosomal CpGs were used in subsequent analyses.
Covariate generation

To account for population substructure, a principal component analysis was performed using KING software [14] on VISP genotype data. The top four and ten principal components (PCs) were generated and used as covariates in the survival analysis for the EUR population and AFR population, respectively. Batch effect was adjusted for using factor variables indicative of methylation data generation round, while cellular heterogeneity due to the variation in cell population proportions within whole blood samples was controlled for using cell proportion estimates generated by the estimateCellCounts function in minfi [12, 15–17]. This function generates cell counts for B-lymphocytes, CD4+ and CD8+ T-lymphocytes, natural killer cells, granulocytes, and monocytes.

Statistical analysis

Baseline characteristics of the study participants with and without recurrent stroke were compared using t-tests and \( \chi^2 \) tests for continuous and categorical traits, respectively, by each ethnic stratum.

Survival analyses utilizing Cox Proportional Hazards (PH) regression models were performed for the time in days to VISP recurrent stroke or composite vascular event for both AFR and EUR participants independently. The degree of methylation was the exposure and the models adjusted for age, sex, PCs, batch effect, treatment arm, and cellular proportions. The replication cohort adjusted for age, sex, the top two PCs, cell heterogeneity proportions, and batch effect. Statistical significance was calculated at \( p \leq 1.06e^{-07} \) (= 0.05/470871 total number of loci) for the discovery cohort. Statistical power ranged from 0.5847 to 0.7466 based on analysis phenotype [18]. A significance threshold of \( p \leq 7.14e^{-03} \) (= 0.05/7) was determined for the look-up analysis in the replication cohort.

Weighted Gene Co-expression Network Analysis (WGCNA) [19] was used to identify networks or modules of highly correlated DNA methylation loci and blood biomarkers or clinical traits (n = 27 traits; S1 Table). Modules were calculated using the blockwise module function in the WGCNA R package. Outlier samples were identified and removed to ensure reliable network construction. An appropriate soft threshold power was calculated for each stratum. A soft thresholding power of 20 in the AFR cohort and 26 in the EUR cohort was determined at a threshold \( \geq 0.8 \). These parameters were used in a signed-hybrid network model with the minimum number of loci set to 30 and a maximum block size set to 10,000. The loci comprising each module were represented by a weighted average, which is indicative of the first principle component in the analysis.

Statistical significance was calculated as \( p \leq 3.93e^{-09} \) (= 0.05/(470871 loci + 27 traits)), while the suggestive threshold was set at \( p < 1.00e^{-03} \).

Gene ontology (GO) term enrichment was performed using GOriilla (Gene Ontology enRichment anaLysis and visuaLizAtion) [20, 21] for the seven significant loci identified in the survival analysis with \( p \leq 1.06e^{-07} \). Statistical significance for GO term enrichment included terms with a false discovery rate (FDR) \( q \)-value \( \leq 0.05 \).

To further evaluate the biological mechanisms of statistically and marginally significant \( (p < 1.00e^{-06}) \) CpG sites from the AFR VISP recurrent stroke and composite vascular event analyses, the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA GWAS) version 1.3.6a [22] online platform and the GENE2FUNC process was utilized. The EUR composite vascular event analysis resulted in two CpGs that met or exceeded the suggestive threshold and therefore, functional mapping was not performed. A total of 36 (AFR VISP recurrent stroke) and 30 (AFR VISP composite vascular event) gene names were combined, resulting in 57 unique gene names uploaded into FUMA for evaluation of gene
expression and enrichment of differentially expressed gene sets in tissues from the Genotype-Tissue Expression (GTEx) 8 RNA sequencing data [23]. Upon filtering by ENSEMBL identifiers, 53 genes were used in the GTEx analyses. A multiple testing correction was performed using a Benjamini-Hochberg adjustment. Statistical significance was calculated using a p-threshold of p<0.05.

Results

The baseline characteristics of 76 AFR and 104 EUR VISP study participants with and without recurrent stroke were compared using χ² and t-tests for categorical and continuous variables, respectively, and are presented in Table 1. Of these individuals, 28 AFR and 32 EUR participants had an incident VISP recurrent stroke. In the AFR individuals, the average baseline age of those experiencing a recurrent stroke was five years older compared to those not having a recurrent stroke (65 years versus 60 years; p = 0.047). Of particular interest, those individuals having a VISP recurrent stroke had a more severe enrollment stroke compared to the non-recurrent control group, as measured on the modified Rankin stroke scale (RSS). While the enrollment criteria for VISP included mild-to-moderate RSS scores of 0–3, 35.7% of those suffering a recurrent stroke during VISP follow-up experienced an enrollment stroke with a RSS of 3, which is indicative of moderate disability, while 10.4% of the non-recurrent individuals had a similar enrollment stroke severity. Although not as extreme, VISP EUR participants had similar baseline age differences (70 years versus 68 years for VISP recurrence and non-recurrence, respectively), as well as enrollment stroke severity compared with AFR participants. Approximately 44% of EUR VISP participants who experienced a recurrent stroke had RSS of 0 or 1, compared to nearly 71% of participants in the non-recurrent group having a similar score indicative of no significant disability (Table 1).

Ethnicity-stratified Cox PH analyses identified a total of seven methylation loci associated with time to event for recurrent stroke or composite vascular event (Tables 2 and 3). Four statistically significant loci were identified for days to VISP recurrent stroke. The most significant association was observed for cg04059318 (HR [95% CI] = 7.19 [3.55–14.57]; p = 4.52e-08), located on chromosome 10. Three additional loci, ch.2.81927627R (HR = 2.72 [1.89–3.93]; p = 9.11e-08), cg03584380 (HR = 5.41 [2.91–10.06]; p = 9.66e-08), and cg24875416 (HR = 2.43 [1.75–3.37]; p = 9.82e-08) were also implicated in the AFR recurrent stroke analysis (Table 2; Fig 1). There were no statistically significant loci in the days to VISP stroke recurrence analysis in EUR. cg03584380, an intronic locus of ZDHHC6, was the only locus to remain significant in the replication cohort from Vall d’Hebron University in Barcelona (HR = 1.83(1.21–2.77); p = 4.00e-03; Table 2). The ENCODE annotation and gene position for cg03584380 is presented in Fig 2. For the time to composite vascular event analysis, cg00076998 (HR = 5.58 [2.98–10.44]; p = 7.87e-08) and cg16758041 (HR = 3.44 [2.18–5.43]; p = 1.04e-07) were identified in AFR (Fig 3; Table 3), while cg02365967 (HR: 0.42 [0.31–0.58]; p = 8.08e-08) was identified in the EUR stratum (Fig 4; Table 3). Using Schoenfeld residual tests, we did not observe any evidence of violation to the proportional hazards assumptions for any of the base models indicated by a global p<0.05 (p = 0.084 for AFR VISP recurrent stroke, p = 0.059 for AFR composite vascular endpoint, p = 0.320 for EUR VISP recurrent stroke, and p = 0.121 for EUR composite vascular endpoint analyses).

GO analysis was performed on the eight genes located closest to the significant loci identified in the survival analyses (PTEN, KLLN, PIK3CB, HERC2, CTNNA2, ZDHHC6, STRIP1, and NDUFB6). Nine terms describing biological processes and molecular function met a modest significance threshold (p≤1.00e-03), including prepulse inhibition (GO:0060134; p = 8.22e-06), brain morphogenesis (GO: 0048854, p = 2.87e-05), and axonogenesis (GO:
Table 1. Baseline demographics for VISP participants.

|                      | AFR | EUR | p* | AFR | EUR | p* |
|----------------------|-----|-----|----|-----|-----|----|
| N                    | 28  | 48  |    | 32  | 72  |    |
| Age, yrs*            | 65.18 (11.54) | 60.40 (8.94) | 0.047 | 70.03 (11.48) | 68.01 (10.81) | 0.391 |
| Sex                  |     |     |    |     |     |    |
| Male (%)             | 17 (60.7) | 29 (60.4) | 1.000 | 18 (56.2) | 39 (54.2) | 1.000 |
| Female (%)           | 11 (39.3) | 19 (39.6) | 14 (43.8) | 33 (45.8) |     |    |
| Treatment Arm        |     |     |    |     |     |    |
| High-dose (%)        | 14 (50.0) | 22 (45.8) | 0.910 | 14 (43.8) | 40 (55.6) | 0.368 |
| Low-dose (%)         | 14 (50.0) | 26 (54.2) | 18 (56.2) | 32 (44.4) |     |    |
| Current Smoker (%)   | 8 (28.6) | 13 (27.1) | 1.000 | 3 (9.4) | 9 (12.5) | 0.898 |
| Body Mass Index, kg/m² | 29.01 (5.45) | 30.35 (6.65) | 0.370 | 29.64 (7.21) | 28.52 (5.68) | 0.403 |
| Diabetes mellitus (%)| 9 (32.1) | 20 (41.7) | 0.562 | 13 (40.6) | 21 (29.2) | 0.356 |
| Myocardial infarction (%) | 3 (10.7) | 1 (2.1) | 0.274 | 0 (0.0) | 4 (5.6) | 0.419 |
| Recurrent stroke ever (%) | 28 (100.0) | 16 (33.3) | <0.001 | 32 (100.0) | 15 (20.8) | <0.001 |
| PNS                  |     |     |    |     |     |    |
| 0                    | 12 (44.4) | 32 (66.7) | 0.139 | 19 (61.3) | 57 (79.2) | 0.212 |
| 1                    | 8 (29.6) | 13 (27.1) | 6 (19.4) | 10 (13.9) |     |    |
| 2                    | 4 (14.8) | 2 (4.2) | 5 (16.1) | 4 (5.6) |     |    |
| 3                    | 2 (7.4) | 1 (2.1) | 1 (3.2) | 1 (1.4) |     |    |
| 4                    | 1 (3.7) | 0 (0.0) | 0 (0.0) |     |     |    |
| RSS                  |     |     |    |     |     |    |
| 0                    | 2 (7.1) | 7 (14.6) | 0.001 | 4 (12.5) | 17 (23.6) | 0.071 |
| 1                    | 16 (57.1) | 20 (41.7) | 10 (31.2) | 34 (47.2) |     |    |
| 2                    | 0 (0.0) | 16 (33.3) | 14 (43.8) | 16 (22.2) |     |    |
| 3                    | 10 (35.7) | 5 (10.4) | 4 (12.5) | 5 (6.9) |     |    |
| Hypertension (%)     | 24 (85.7) | 42 (87.5) | 1.000 | 28 (87.5) | 56 (77.8) | 0.373 |
| Systolic Blood Pressure, mmHg | 144.34 (20.43) | 143.65 (20.16) | 0.886 | 146.77 (19.29) | 140.11 (19.67) | 0.112 |
| Diastolic Blood Pressure, mmHg | 80.09 (8.99) | 82.05 (10.12) | 0.414 | 81.34 (8.67) | 78.14 (9.81) | 0.115 |
| Creatinine, mg/dL    | 1.16 (0.39) | 1.41 (0.68) | 0.076 | 1.26 (0.55) | 1.11 (1.00) | 0.44 |
| Total cholesterol, mg/dL | 190.09 (53.10) | 212.85 (55.33) | 0.107 | 206.48 (37.67) | 201.93 (48.04) | 0.641 |
| High-density lipoprotein, mg/dL | 42.17 (10.76) | 51.76 (16.99) | 0.016 | 45.17 (13.78) | 46.59 (16.10) | 0.676 |
| Triglycerides, mg/dL | 146.95 (105.06) | 162.46 (138.70) | 0.65 | 193.83 (110.43) | 201.89 (193.05) | 0.834 |
| Total plasma homocysteine, μmol/L | 15.20 (5.16) | 14.95 (6.62) | 0.862 | 15.58 (6.85) | 16.82 (12.82) | 0.607 |
| B6, nmol/L           | 36.08 (33.36) | 22.33 (20.91) | 0.042 | 29.99 (19.26) | 43.03 (30.34) | 0.18 |
| B12, pmol/L          | 421.63 (226.23) | 418.30 (190.21) | 0.947 | 354.97 (132.03) | 313.50 (124.92) | 0.144 |
| Folate, ng/mL        | 31.92 (32.59) | 21.85 (9.26) | 0.057 | 20.84 (10.59) | 24.65 (11.25) | 0.122 |
| C-reactive protein, mg/L | 12.49 (8.92) | 12.64 (10.26) | 0.953 | 13.50 (8.41) | 11.37 (8.92) | 0.277 |
| Prothrombin fragments 1+2, nmol/L | 0.99 (0.56) | 1.25 (1.42) | 0.386 | 1.14 (0.63) | 1.14 (1.28) | 0.996 |
| Thrombin-antithrombin complex, μg/L | 9.11 (12.52) | 5.26 (3.49) | 0.099 | 6.47 (8.94) | 6.42 (6.73) | 0.98 |
| Thrombomodulin, ng/mL | 10.33 (18.47) | 11.74 (25.00) | 0.84 | 10.66 (22.63) | 8.34 (12.82) | 0.588 |
| Tissue plasminogen activator, ng/mL | 15.42 (10.61) | 15.78 (19.58) | 0.933 | 14.32 (15.14) | 18.07 (21.12) | 0.419 |
To further elucidate any biological implications of our findings, we performed functional annotation and mapping of the 57 unique genes with suggestive (p<1.00E-06) CpG loci from the Cox PH results on AFR VISP recurrent stroke and composite vascular endpoint analyses. Tissue analysis on 53 specific types from the GTEx project, revealed statistically significant differential down-regulated expression in the pancreas (p\text{adj}=3.05e-08), putamen basal ganglia (p\text{adj}=2.42e-06), left ventricle of the heart (p\text{adj}=1.54e-05), liver (p\text{adj}=8.75e-05), caudate basal ganglia (p\text{adj}=2.30e-04), hippocampus (p\text{adj}=3.52e-04), nucleus accumbens basal ganglia (p\text{adj}=9.30e-04), anterior cingulate cortex BA24 (p\text{adj}=5.93e-04), substantia nigra (p\text{adj}=6.02e-03), skeletal muscle (p\text{adj}=2.72e-02), whole blood (p\text{adj}=2.91e-02), hypothalamus (p\text{adj}=2.93e-02), and the cortex (p\text{adj}=3.89e-02) (Fig 5, Table 4), when compared to the background gene set from GTEx v8.

To evaluate comprehensive networks of methylation loci associated with stroke related clinical traits, WGCNA was performed in the two ethnic strata upon outlier removal, which resulted in 100 EUR and 73 AFR participants included. Twenty-seven traits, including baseline biomarker levels, stroke risk factors, and outcome statuses were included in the analyses (Table 1; S1 Table). In the AFR stratum, 106 modules were produced, of which six were observed in significant module-clinical trait associations (p<3.93e-09; Table 5). The significant associations observed were between modules and the previous number of strokes experienced before VISP enrollment (r = -0.90; 5.00e-27), prothrombin fragments 1 + 2 (r = -0.79; p = 8.00e-17), thrombomodulin (r = -0.77; 1.00e-15), thrombin-antithrombin complex (r = -0.73; p = 3.00e-13), triglyceride levels (r = -0.71; p = 2.00e-12), and tissue plasminogen activator (r = -0.68; p = 3.00e-11). Additional module associations in AFR were detected with traits and outcomes including MI, hypertension, thrombomodulin, and thrombin-antithrombin complex. In the EUR analysis, 27 modules were produced; however, no module-trait associations observed met the significance threshold. Notable nominal associations in this stratum included modules correlated with composite vascular endpoint, VISP recurrent stroke, and

Table 1. (Continued)

| Locus | CHR:BP | Location | Gene | Mean Beta Value | HR (95% CI) | p | p |
|-------|--------|----------|------|-----------------|-------------|---|---|
| cg04059318 | 10:89622526 | TSS1500; Exon 1 | PTEN; KLLN | 0.07 | 7.19 (3.55–14.57) | 4.52E-08 | 0.665 |
| ch.2.81927627R | 2:82074116 | Downstream | CTNNA2 | 0.09 | 2.72 (1.89–3.93) | 9.11E-08 | 0.178 |
| cg03584380 | 10:114206433 | Intron 1 | ZDHHC6 | 0.06 | 5.41 (2.91–10.06) | 9.66E-08 | 0.004 |
| cg24875416 | 1:110577936 | Intron 1 | STRIP1 | 0.08 | 2.43 (1.75–3.37) | 9.82E-08 | 0.784 |

aDiscovery model adjusts for age, sex, top 10 genetic principal components, treatment arm, batch effect, and estimated cellular proportions.

Abbreviations: CHR- chromosome; BP-base position (hg19); TSS1500- covers from 200 to 1500 nucleotides upstream of transcriptional start site (TSS); HR-hazard ratio; CI-confidence interval.

https://doi.org/10.1371/journal.pone.0254562.t002
total plasma homocysteine. Only two loci identified in the Cox PH models were included in any of the modules. While identified in the AFR survival analyses, cg04059318 and cg24875416 were included in the turquoise module for EUR. Interestingly, this was one of two modules nominally associated with VISP recurrent stroke or composite vascular endpoint ($r = 0.43, p = 8.00\times10^{-6}$ and $r = 0.44, p = 5.00\times10^{-6}$, respectively) (Table 5).

**Discussion**

To the best of our knowledge, this study represents the first epigenome-wide association study evaluating association between time to recurrent stroke (or composite vascular event) and the degree of methylation. This study was performed in a subset of the VISP clinical trial

| Strata | Locus  | CHR:BP       | Locus Location | Gene   | Mean Beta Value | HR (95% CI)   | P     | P     |
|--------|--------|--------------|----------------|--------|----------------|---------------|-------|-------|
| AFR    | cg00076998 | 3:138553170  | Upstream       | PIK3CB | 0.06           | 5.58 (2.98–10.44) | 7.87E-08 | 0.608 |
| AFR    | cg16758041 | 9:32573371   | TSS200         | NDUF6b | 0.07           | 3.44 (2.18–5.43)  | 1.04E-07 | 0.738 |
| EUR    | cg02365967 | 15:28473380  | Exon 35       | HERC2  | 0.94           | 0.42 (0.31–0.58)  | 8.08E-08 | 0.163 |

*Discovery model adjusts for age, sex, genetic principal components (first 4 in EUR; first 10 in AFR), treatment arm, batch effect, and estimated cellular proportions.

**Abbreviations:** CHR- chromosome; BP-base position (hg19); TSS200- region from the transcriptional start site (TSS) to 200 nucleotides upstream of TSS; HR-hazard ratio; CI-confidence interval.

https://doi.org/10.1371/journal.pone.0254562.t003

Fig 1. Epigenome-wide Manhattan plot for time (days) to VISP recurrent stroke survival analysis in AFR. Each point corresponds to the -log$_{10}$(P-value) for a CpG site at its specific chromosome location (y-axis). Horizontal line is indicative of epigenome-wide significance threshold ($p \leq 1.06\times10^{-7}$).

https://doi.org/10.1371/journal.pone.0254562.g001
participants, providing a diverse cohort of individuals of both AFR and EUR descent. The inclusion of individuals of AFR descent is a strength of this study, since this population is 60% more likely to experience a recurrent stroke within two years compared to individuals of EUR descent, albeit likely mediated by stroke risk factors and comorbidities [24]. Seven loci reached
or exceeded a Bonferroni corrected significance threshold, with cg03584380 statistically significant in AFR stroke recurrence and validated upon replication in the Vall d’Hebron University Hospital cohort. While similar trends were observed across phenotypes due to the strong correlation between the primary (VISP recurrent stroke) and secondary (composite vascular) outcomes of interest at these loci, results were not significant across separate cohorts, suggesting potential ethnic disparities, as demonstrated by the significant association between

Fig 4. Epigenome-wide Manhattan plot for EUR time to composite vascular event. Horizontal line is indicative of epigenome-wide significance threshold ($p < 1.06e-07$).

https://doi.org/10.1371/journal.pone.0254562.g004

Fig 5. Differentially expressed down-regulated genes from FUMA. Red bars are indicative of statistically significant differentially expressed tissues based on Benjamini-Hochberg adjustment. Significance threshold $p_{adj} < 0.05$.

https://doi.org/10.1371/journal.pone.0254562.g005
DNA methylation analyses identify an intronic locus associated with time to recurrent stroke in VISP
cg03584380 and stroke recurrence in the AFR and Spanish cohorts, but not in EUR (p = 4.07e-01). Within the AFR stratum, an increase in the beta values at the six significant loci were associated with shorter duration to recurrent stroke or composite vascular event, thus suggesting

Table 4. Differentially expressed genes from FUMA using GTEx version 8 tissue expression.

| General Tissue Type | Specific Tissue Type from GTEx | N genes* | N overlapb | p<sup>c</sup> | p<sub>adj</sub>d |
|---------------------|-------------------------------|----------|------------|--------------|--------------|
| Pancreas            | Pancreas                      | 9668     | 37         | 5.64E-10     | 3.05E-08     |
| Brain               | Putamen basal ganglia         | 7853     | 31         | 4.49E-08     | 2.42E-06     |
| Heart               | Left ventricle               | 9531     | 33         | 2.85E-07     | 1.54E-05     |
| Liver               | Liver                        | 8059     | 29         | 1.62E-06     | 8.75E-05     |
| Brain               | Amygdala                     | 7751     | 28         | 2.77E-06     | 1.50E-04     |
| Brain               | Caudate basal ganglia        | 6927     | 26         | 4.26E-06     | 2.30E-04     |
| Brain               | Hippocampus                  | 7571     | 27         | 6.53E-06     | 3.52E-04     |
| Brain               | Nucleus accumbens basal ganglia | 6479   | 24         | 1.72E-05     | 9.30E-04     |
| Brain               | Anterior cingulate cortex BA24 | 6725   | 23         | 1.10E-04     | 5.93E-03     |
| Brain               | Substantia nigra             | 7223     | 24         | 1.11E-04     | 6.02E-03     |
| Muscle              | Skeletal                     | 6908     | 22         | 5.03E-04     | 2.72E-02     |
| Blood               | Whole blood                  | 6940     | 22         | 5.38E-04     | 2.91E-02     |
| Brain               | Hypothalamus                 | 5967     | 20         | 5.43E-04     | 2.93E-02     |

*The number of background genes expressed in specific tissue from GTEx version 8.

bThe number of overlap occurring between input genes and background.

Unadjusted p-value from the hypergeometric test performed in FUMA.

Benjamini-Hochberg adjusted p-value.

https://doi.org/10.1371/journal.pone.0254562.t004

Table 5. Significant WGCNA module-trait associations.

| Strata | Module | Trait                              | r    | p<sup>a</sup> |
|--------|--------|------------------------------------|------|--------------|
| AFR    | bisque4| Previous number of strokes         | -0.90| 5.00E-27     |
| AFR    | coral2 | Prothrombin fragments 1 + 2        | -0.79| 8.00E-17     |
| AFR    | pink4  | Thrombomodulin                     | -0.77| 1.00E-15     |
| AFR    | coral  | Thrombin-antithrombin complex      | -0.73| 3.00E-13     |
| AFR    | yellowgreen | Triglycerides                  | -0.71| 2.00E-12     |
| AFR    | lightsteelblue | Tissue plasminogen activator   | -0.68| 3.00E-11     |
| AFR    | darkturquoise | Myocardial infarction        | 0.51 | 4.00E-06     |
| EUR    | turquoise | Composite endpoint               | 0.44 | 5.00E-06     |
| EUR    | turquoise | VISP recurrent stroke            | 0.43 | 8.00E-06     |
| AFR    | honeydew1 | Myocardial infarction           | 0.49 | 9.00E-06     |
| AFR    | blue2  | Myocardial infarction            | -0.49| 1.00E-05     |
| AFR    | coral1 | Myocardial infarction            | -0.48| 2.00E-05     |
| AFR    | lightcyan1 | Myocardial infarction         | -0.48| 2.00E-05     |
| AFR    | lightslateblue | Myocardial infarction        | 0.48 | 2.00E-05     |
| AFR    | coral  | Thrombomodulin                    | -0.43| 1.00E-04     |
| AFR    | mediumpurple2 | Thrombin-antithrombin complex | -0.44| 1.00E-04     |
| AFR    | royalblue | Myocardial infarction          | -0.42| 3.00E-04     |
| EUR    | cyan   | Composite endpoint               | 0.34 | 5.00E-04     |
| AFR    | thistle3 | Hypertension                     | 0.39 | 7.00E-04     |
| EUR    | cyan   | Total plasma homocysteine        | 0.33 | 8.00E-04     |

*Statistical significance (bold): p≤ 3.93e-09.

https://doi.org/10.1371/journal.pone.0254562.t005
that increased methylation at these loci could be indicative of earlier event recurrence. Three of these loci have implications in the phosphatidylinositol pathway, while the remaining three loci were located in genes or gene families linked to cardiovascular traits including the pathogenesis of cardiac disease and diabetes [25], regulation of hyperlipidemia and arteriosclerosis [26], and insulin sensitivity [27].

cg03584380 was the only locus replicated in our analyses. This methylation site is located in the first intron of ZDHHC6, the gene which encodes palmitoyltransferase ZDHHC6 [28] and mediates the palmitoylation of critical endoplasmic reticulum (ER) proteins, including calnexin and the inositol 1,4,5-trisphosphate receptor (ITPR1) [29]. Calnexin is an ER chaperone that has been implicated in cardiomyocyte viability and in ER stress, a prominent clinical feature of cardiovascular disease [29], while ITPR1 mediates the influx and release of intracellular Ca^{2+} and is regulated by the palmitoylation cascade of ZDHHC16/ZDHHC6 [30, 31]. Additionally, ITPR1 can be phosphorylated by Akt kinase and further regulated by phosphatidylinositol 3-kinase (PI3K) [32]. Results from the ENCODE annotation suggest this region around cg03584380 is an active regulatory site. There are several active promoters for both the ZDHHC6 and VTI1A genes as indicated by the red bars in Fig 2 in the chromatin state segmentation by the Hidden Markov Model (HMM) track, as well as a CpG island just downstream of this locus. Furthermore, there is evidence of increased H3K27 (lysine 27 of the H3 histone) acetylation, which also indicates enhanced transcription. Therefore, it is plausible that cg03584380 could regulate chromatin and histone states related to transcription and transcription factor binding, although future gene expression and functional work is needed.

Although the association between cg04059318 and AFR stroke recurrence was not validated in the EUR only analysis (p = 7.34e-02) or look-up efforts (p = 6.65e-01), it was the most statistically significant association detected in the VISP survival analysis. cg04059318 is located within the 5' untranslated region of PTEN and within an intron of KLLN. KLLN encodes KILLIN, a DNA-binding protein that inhibits DNA synthesis and mediates p53/TP53-induced apoptosis [28]. PTEN encodes PTEN, a tumor suppressor that regulates angiogenesis [33] and has been associated with a number of stroke-related clinical traits including triglycerides [34] and type 2 diabetes [35]. PTEN negatively regulates the Akt signaling pathway through intracellular phosphatidylinositol 3-phosphate (PI3P), which is of interest due to the neuroprotective properties of Akt against ischemia-induced damage [36, 37]. A third locus involved in the phosphatidylinositol pathway was identified in the composite vascular event analysis. cg00076998 is located upstream of PIK3CB, which encodes the catalytic subunit of phosphatidylinositol 3-kinase, beta (PI3Kβ). PI3K activates cellular signaling cascades through the generation of phosphatidylinositol (3,4,5)-trisphosphate and recruitment of Akt and phosphoinositide-dependent kinase-1 [38]. PI3K is involved in platelet activation signaling triggered by G-protein coupled receptors [28] and is required for platelet-induced aggregation induced by thrombin and thromboxane A2 (TxA2) [39]. PI3K/Akt signaling has been observed in regulation of vascular tone, or the degree of vasoconstriction experienced by a blood vessel, in both vascular endothelium and smooth muscle cells [40]. Exposure to homocysteine, a well-documented risk factor for atherosclerosis and stroke [41, 42], in endothelial cells can form the intermediate S-nitroso-homocysteine, a vasodilator implicated in reducing vascular tone and altering arteriole calcium levels [40]. Collectively, these associations have identified a novel association with ZDHHC6 and implicate PI3K/Akt signaling in time to stroke and/or vascular event following stroke [43, 44]. These findings could in part, help explain the ethnic disparities in stroke severity and recovery seen in African American patients; however, further functional analyses are needed to confirm the biological significance of our results.
cg02365967 was the only locus identified in our EUR cohort (composite vascular event; AFR composite p = 2.54e-01), and was the sole locus that exhibited hypermethylated beta values (mean $\beta = 0.9411$). Increased beta values at this locus were associated with longer duration to event, which differed from the other identified loci. cg02365967 is located in exon 35 of \textit{HERC2}, which encodes the ubiquitin-protein ligase \textit{HERC2}. This protein regulates ubiquitin-dependent retention of repair proteins and is a binding protein with SIRT1, a NAD-dependent protein deacetylase that plays a vasoprotective role in endothelial cells [45]. Increased levels of SIRT1 have been described to promote endothelial angiogenesis, enhance vasodilation, and suppress vascular inflammation [45]. Overexpression of SIRT1 in murine endothelial cells prevented hypertension and adverse arterial remodeling; however, a knockdown of \textit{HERC2} abolished any beneficial effects of SIRT1 [45], suggesting a neuroprotective regulatory role of \textit{HERC2}.

Network analyses using WGCNA, identified six statistically significant associations between modules and clinical traits in the AFR stratum, including a correlation between a module comprised of 58 methylation loci within 38 genes and the previous number of strokes prior to VISP enrollment variable. Of these genes, \textit{A2ML1}, \textit{AGGF1}, \textit{CBS}, \textit{ECE1}, \textit{GABRB3}, \textit{GALNT2}, \textit{MRPS6/SLC5A3}, and \textit{SPG7} had documented associations with hypertension, ischemic stroke and methionine metabolism, atherosclerosis, and early-onset MI [46–48]. GO term enrichment of the genes identified in survival analyses, resulted in terms associated with prepulse inhibition (GO: 0060134), brain morphogenesis (GO: 0048854), axonogenesis (GO: 0007409), and postsynaptic assembly, organization, and regulation (GO: 0098698, GO: 2000463, GO: 0098815, and GO: 0099084). Prepulse inhibition (PPI) describes the regulated transmission of sensory information, while disrupted PPI has been linked to neurological disorders including Tourette’s syndrome and Schizophrenia [49]. Brain morphogenesis generates and organizes anatomical structures of the brain, which can be a crucial process after ischemia. Axonogenesis, or the \textit{de novo} generation of a neuron’s long process [50], requires the melatonin MT2 receptor [51]. Deficits in MT2 signaling have been observed in a number of neurological disorders, including Alzheimer’s disease, suggesting that axonogenesis may be beneficial in regard to the outcome of these disorders [51]. The GO results, in addition to the differential gene expression analysis in GTEx tissues, provide evidence of altered cerebro- and cardiovascular regulation.

Recurrent stroke is a vastly understudied phenotype and its etiology is not well understood. The utilization of this phenotype is a strength of this study; however, one limitation is the lack of formally adjudicated ischemic stroke subtype. We can conclude that due to the inclusion and exclusion criteria of VISP, the enrollment strokes were most likely lacunar or small vessel infarctions. DNA methylation profiles were generated from whole blood samples. Although not optimal due to cellular heterogeneity, whole blood provides a valuable resource that is potentially available for replication studies and represents a minimally invasive source for potential biomarker testing. To overcome this limitation, cellular proportions were calculated \textit{in silico}, and used as covariates in the Cox PH models. Although our analyses detected seven statistically significant loci and replicated cg03584380 in the Vall d’Hebron University Hospital cohort, one constraint of this study that could not be fully addressed was the limited statistical power due to the modest sample sizes of both the discovery cohort and replication cohort. Furthermore, previous studies have shown that global methylation patterns differ across ethnicities, as well as within ethnic subgroups. In a study of 573 individuals from diverse Latino ethnic sub-groups, genetic ancestry explained approximately 75% of the variation in methylation between the sub-groups [52]. Therefore, the use of cohorts from different races or ethnicities should allow for the identification of methylation sites with global implications (affecting most ethnicities similarly) but will limit our ability to identify race/ethnicity-specific...
associations. In conclusion, findings from this study provide insight into the relationship between the degree of DNA methylation and the duration to recurrent stroke and vascular event following a stroke and lay the foundation for further studies investigating these outcomes in diverse populations.

Supporting information

S1 Table. List of blood biomarkers and clinical traits used in WGCNA.
(DOCX)

S2 Table. Gene ontology results from significant methylation loci.
(DOCX)

Acknowledgments

The authors would like to thank the individuals who participated in the VISP study. Co-author Michele M. Sale passed away before the submission of the final version of this manuscript. The corresponding author (KLK) accepts responsibility for the integrity and validity of the data collected and analyzed.

Author Contributions

Conceptualization: Nicole M. Davis Armstrong, Stephen R. Williams, Michèle M. Sale, Bradford B. Worrall, Keith L. Keene.

Data curation: Wei-Min Chen, Fang-Chi Hsu, Michèle M. Sale.

Formal analysis: Nicole M. Davis Armstrong, Fang-Chi Hsu, Michael S. Brewer, Natalia Cullell, Keith L. Keene.

Funding acquisition: Michèle M. Sale, Bradford B. Worrall, Keith L. Keene.

Methodology: Nicole M. Davis Armstrong, Wei-Min Chen, Fang-Chi Hsu, Michael S. Brewer, Stephen R. Williams, Bradford B. Worrall, Keith L. Keene.

Resources: Michèle M. Sale, Bradford B. Worrall, Keith L. Keene.

Supervision: Keith L. Keene.

Validation: Natalia Cullell, Israel Fernández-Cadenas.

Writing – original draft: Nicole M. Davis Armstrong, Keith L. Keene.

Writing – review & editing: Nicole M. Davis Armstrong, Wei-Min Chen, Fang-Chi Hsu, Michael S. Brewer, Natalia Cullell, Israel Fernández-Cadenas, Stephen R. Williams, Bradford B. Worrall, Keith L. Keene.

References

1. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics–2020 Update: A Report From the American Heart Association, Circulation. 2020; 141 (9):e139–e596. https://doi.org/10.1161/CIR.0000000000000757 PMID: 31992061

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(12):3161–7. https://doi.org/10.1161/STROKEAHA.112.665760 PMID: 23042660

3. Reid MA, Dai Z, Locasale JW. The impact of cellular metabolism on chromatin dynamics and epigenetics. Nat Cell Biol. 2017; 19(11):1298–306. https://doi.org/10.1038/ncb3629 PMID: 29058720
PLOS ONE

DNA methylation analyses identify an intronic locus associated with time to recurrent stroke in VISP

4. Lappalainen T, Greally JM. Associating cellular epigenetic models with human phenotypes. Nat Rev Genet. 2017; 18(7):441–51. https://doi.org/10.1038/nrg.2017.32 PMID: 2855657

5. Endres M, Meisel A, Biniszkiewicz D, Namura S, Prass K, Ruscher K, et al. DNA methyltransferase contributes to delayed ischemic brain injury. J Neurosci. 2000; 20(9):3175–81. https://doi.org/10.1523/JNEUROSCI.20-09-03175.2000 PMID: 10777781

6. Zaina S, Heyn H, Carmona FJ, Varol N, Sayols S, Condom E, et al. DNA methylation map of human atherosclerosis. Circ Cardiovasc Genet. 2014; 7(5):692–700. https://doi.org/10.1161/CIRCGENETICS.113.00441 PMID: 25091541

7. Gallego-Fabrega C, Carrera C, Reny JL, Fontana P, Slowik A, Pera J, et al. PPM1A Methylation Is Associated With Vascular Recurrence in Aspirin-Treated Patients. Stroke. 2016; 47(7):1926–9. https://doi.org/10.1161/STROKEAHA.115.012237 PMID: 2706631

8. Gallego-Fabrega C, Carrera C, Reny JL, Fontana P, Slowik A, Pera J, et al. TRAF3 Epigenetic Regulation Is Associated With Vascular Recurrence in Patients With Ischemic Stroke. Stroke. 2016; 47(5):1180–6. https://doi.org/10.1161/STROKEAHA.115.012237 PMID: 2706631

9. Davis Armstrong NM, Chen WM, Brewer MS, Williams SR, Sale MM, Worrall BB, et al. Epigenome-Wide Analyses Identify Two Novel Associations With Recurrent Stroke in the Vitamin Intervention for Stroke Prevention Clinical Trial. Front Genet. 2018; 9:358. https://doi.org/10.3389/fgene.2018.00358 PMID: 30237808

10. Toole JF. Vitamin intervention for stroke prevention. J Neurol Sci. 2002;203–204:121–4. https://doi.org/10.1016/s0022-510x(02 )00265-4 PMID: 12417369

11. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013; 14:293. https://doi.org/10.1186/1471-2164-14-293 PMID: 23631413

12. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Illumina DNA methylation microarrays. Bioinformatics. 2014; 30(10):1363–9. https://doi.org/10.1093/bioinformatics/btu049 PMID: 24478339

13. Touleimat N, Tost J. Complete pipeline for Infinium(R) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics. 2012; 4(3):325–41. https://doi.org/10.2217/epi.11.21 PMID: 22690668

14. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010; 26(22):2867–73. https://doi.org/10.1093/bioinformatics/btq255 PMID: 20926424

15. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 2014; 15(2):R31. https://doi.org/10.1186/gb-2014-15-2-r31 PMID: 24495553

16. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012; 13:86. https://doi.org/10.1186/1471-2105-13-86 PMID: 22568884

17. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlen SE, Greco D, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS One. 2012; 7(7):e41361. https://doi.org/10.1371/journal.pone.0041361 PMID: 22848742

18. Qiu W CJ, Lazarus R, Rosner B, Ma J. powerSurvEpi: Power and Sample Size Calculation for Survival Analysis of Epidemiological Studies. R package version 0.1.0. Available at: https://CRAN.R-project.org/package=powerSurvEpi. 2018.

19. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. https://doi.org/10.1186/1471-2105-9-559 PMID: 1914008

20. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol. 2007; 3(3):e39. https://doi.org/10.1371/journal.pcbi.0030039 PMID: 17381235

21. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GORilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009; 10:48. https://doi.org/10.1186/ 1471-2105-10-48 PMID: 19192299

22. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017; 8(1):1826. https://doi.org/10.1038/s41467-017-01261-5 PMID: 29184056

23. Consortium GT, Laboratory DA, Coordinating Center - Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC, et al. Genetic effects on gene expression across human tissues. Nature. 2017; 550(7675):204–13. https://doi.org/10.1038/nature24277 PMID: 29022597

24. Park JH, Ovbiagiele B. Association of black race with recurrent stroke risk. J Neurol Sci. 2016; 365:203–6. https://doi.org/10.1016/j.jns.2016.04.012 PMID: 27206907
25. Hwang J, Pallas DC. STRIPAK complexes: structure, biological function, and involvement in human diseases. Int J Biochem Cell Biol. 2014; 47:118–48. https://doi.org/10.1016/j.biocel.2013.11.021 PMID: 24333164

26. Fox CS, Heard-Costa N, Cuppies LA, Dupuis J, Vasan RS, Atwood LD. Genome-wide association to body mass index and waist circumference: the Framingham Heart Study 100K project. BMC Med Genet. 2007; 8 Suppl 1:S18. https://doi.org/10.1186/1741-2350-8-S1-S18 PMID: 17903300

27. Ling C, Poulsen S, Simonsson S, Rønn T, Holmkrøtt J, Almgren P, et al. Genetic and epigenetic factors are associated with expression of respiratory chain component NDUFB6 in human skeletal muscle. J Clin Invest. 2007; 117(11):3427–35. https://doi.org/10.1172/JCI30938 PMID: 17948130

28. UniProt Consortium T. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 2018; 46 (5):2699. https://doi.org/10.1093/nar/gky092 PMID: 29425356

29. Boussette N, Abbassi C, Chis R, Gramolini AO. Calnexin silencing in mouse neonatal cardiomyocytes induces Ca2+ cycling defects, ER stress, and apoptosis. J Cell Physiol. 2014; 229(3):374–83. https://doi.org/10.1002/jcp.24459 PMID: 24037923

30. Abrami L, Dallavilla T, Sandoz PA, Demir M, Kunz B, Savoglidis G, et al. Identification and dynamics of the human ZDHHC16-ZDHHC6 palmitoylation cascade. Elife. 2017; 6. https://doi.org/10.7554/eLife.27826 PMID: 28826475

31. Fredericks GJ, Hoffmann FW, Rose AH, Osterheld HJ, Hess FM, Mercier F, et al. Stable expression of human recombinant MTH1 in Saccharomyces cerevisiae. FEBS Lett. 2003; 554(3):450–4. https://doi.org/10.1016/s0014-5793(03)01225-0 PMID: 14623110

32. Khan MT, Wagner L 2nd, Yule DI, Bhanumathy C, Joseph SK. Akt kinase phosphorylation of inositol 1,4,5-triphosphate receptor requires palmitoylation by a DHHC6/selenoprotein K complex. Proc Natl Acad Sci U S A. 2014; 111(46):16478–83. https://doi.org/10.1073/pnas.1417176111 PMID: 25368151

33. Zhu Y, Wloch A, Wu Q, Peters C, Pagentecher A, Bertalanffy H, et al. Involvement of PTEN promoter methylation in cerebral cavernous malformations. Stroke. 2009; 40(3):820–6. https://doi.org/10.1161/STROKEAHA.108.526376 PMID: 19118244

34. Kathiresan S, Manning AK, Demissie S, D’Agostino RB, Surti A, Guiducci C, et al. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. BMC Med Genet. 2007; 8 Suppl 1:S17. https://doi.org/10.1186/1471-2350-8-S1-S17 PMID: 17903299

35. Ishihara H, Sasaoka T, Kagawa S, Murakami S, Fukui K, Kawagishi Y, et al. Association of the polymorphisms in the 5’-untranslated region of PTEN gene with type 2 diabetes in a Japanese population. FEBS Lett. 2003; 554(3):450–4. https://doi.org/10.1016/s0014-5793(03)01225-0 PMID: 14623110

36. Liu B, Li L, Zhang Q, Chang N, Wang D, Shan Y, et al. Preservation of GABAA receptor function by PTEN inhibition protects against neuronal death in ischemic stroke. Stroke. 2010; 41(5):1018–26. https://doi.org/10.1161/STROKEAHA.110.579011 PMID: 20360540

37. Zhang L, Kendler KS, Chen X. Association of the phosphatase and tensin homolog gene (PTEN) with smoking initiation and nicotine dependence. Am J Med Genet B Neuropsychiatr Genet. 2008; 141B (1):10–4. https://doi.org/10.1002/ajmg.b.30240 PMID: 16331670

38. Gene. National Library of Medicine, National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/gene/, 2004.

39. Estevez B, Du X. New Concepts and Mechanisms of Platelet Activation Signaling. Physiology (Bethesda). 2017; 32(2):162–77. https://doi.org/10.1152/physiol.00020.2016 PMID: 28228483

40. An YM, Feng H, Zhang XZ, Cong X, Zhao Q, Wu LL, et al. Homocysteine ameliorates the endothelium-independent hypoxic vasoconstriction via the suppression of phosphatidylinositol 3-kinase/Akt pathway in porcine coronary arteries. Biochem Biophys Res Commun. 2017; 486(1):371–7. https://doi.org/10.1016/j.bbrc.2017.03.022 PMID: 28285136

41. Shi Z, Guan Y, Huo YR, Liu S, Zhang M, Lu H, et al. Elevated Total Homocysteine Levels in Acute Ischemic Stroke Are Associated With Long-Term Mortality. Stroke. 2015; 46(9):2419–25. https://doi.org/10.1161/STROKEAHA.115.009136 PMID: 26199315

42. Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. BMJ. 2002; 325(7374):1202. https://doi.org/10.1136/bmj.325.7374.1202 PMID: 12446535

43. Jin R, Song Z, Yu S, Piazza A, Nanda A, Penninger JM, et al. Phosphatidylinositol-3-kinase gamma plays a central role in blood-brain barrier dysfunction in acute experimental stroke. Stroke. 2011; 42(7):2033–44. https://doi.org/10.1161/STROKEAHA.110.601369 PMID: 21546487

44. Kim YS, Yoo A, Son JW, Kim HY, Lee YJ, Hwang S, et al. Early Activation of Phosphatidylinositol 3-Kinase after Ischemic Stroke Reduces Infarct Volume and Improves Long-Term Behavior. Mol Neurobiol. 2017; 54(7):5375–84. https://doi.org/10.1007/s12035-016-0663-4 PMID: 27590139
45. Bai B, Man AW, Yang K, Guo Y, Xu C, Tse HF, et al. Endothelial SIRT1 prevents adverse arterial remodeling by facilitating HERC2-mediated degradation of acetylated LKB1. Oncotarget. 2016; 7(26):39065–81. https://doi.org/10.18632/oncotarget.9687 PMID: 27259994

46. Surendran P, Drenos F, Young R, Warren H, Cook JP, Manning AK, et al. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. Nat Genet. 2016; 48(10):1151–61. https://doi.org/10.1038/ng.3654 PMID: 27618447

47. Sui R, He Z. Polymorphisms of ECE1 may contribute to susceptibility to ischemic stroke in Han Chinese of Northern China. Cell Biochem Biophys. 2014; 69(2):237–46. https://doi.org/10.1007/s12013-013-9789-z PMID: 24595843

48. Williams SR, Yang Q, Chen F, Liu X, Keene KL, Jacques P, et al. Genome-wide meta-analysis of homocysteine and methionine metabolism identifies five one carbon metabolism loci and a novel association of ALDH1L1 with ischemic stroke. PLoS Genet. 2014; 10(3):e1004214. https://doi.org/10.1371/journal.pgen.1004214

49. Mena A, Ruiz-Salas JC, Puentes A, Dorado I, Ruiz-Veguilla M, De la Casa LG. Reduced Prepulse Inhibition as a Biomarker of Schizophrenia. Front Behav Neurosci. 2016; 10:202. https://doi.org/10.3389/fnbeh.2016.00202 PMID: 27803654

50. QuickGO. Embl-EBI. http://www.ebi.ac.uk/QuickGO/. 2017.

51. Liu D, Wei N, Man HY, Lu Y, Zhu LQ, Wang JZ. The MT2 receptor stimulates axonogenesis and enhances synaptic transmission by activating Akt signaling. Cell Death Differ. 2015; 22(4):583–96. https://doi.org/10.1038/cdd.2014.195 PMID: 25501601

52. Galanter JM, Gignoux CR, Oh SS, Torgerson D, Pino-Yanes M, Thakur N, et al. Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. Elife. 2017; 6. https://doi.org/10.7554/eLife.20532 PMID: 28044981