The Soluble Epoxide Hydrolase Gene Harbors Sequence Variation Associated with Susceptibility to and Protection from Incident Ischemic Stroke

Myriam Fornage, Ph.D.[1],[*], **Craig R. Lee, Pharm. D.**[2],[3], **Peter A. Doris, Ph.D.**[1], **Molly S. Bray, Ph.D.**[4], **Gerardo Heiss, M.D., Ph.D.**[5], **Darryl C. Zeldin, M.D.**[3], and **Eric Boerwinkle, Ph.D.**[1]

1*Institute of Molecular Medicine for the Prevention of Human Diseases, University of Texas Health Science Center, Houston, Texas*

2*Division of Pharmacotherapy and Experimental Therapeutics, School of Pharmacy, the University of North Carolina at Chapel Hill*

3*Division of Intramural Research, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, North Carolina*

4*Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, Texas*

5*Department of Epidemiology, School of Public Health, the University of North Carolina at Chapel Hill.*

Abstract

Stroke is the leading cause of severe disability and the third leading cause of death, accounting for 1 of every 15 deaths in the United States. We investigated the association of polymorphisms in the soluble epoxide hydrolase gene (*EPHX2*) with incident ischemic stroke in African-Americans and Whites. Twelve single nucleotide polymorphisms (SNPs) spanning *EPHX2* were genotyped in a case-cohort sample of 1,336 participants from the Atherosclerosis Risk in Communities (ARIC) study. In each racial group, Cox proportional hazard models were constructed to assess the relationship between incident ischemic stroke and *EPHX2* polymorphisms. A score test method was used to investigate the association of common haplotypes of the gene with risk of ischemic stroke. In African-Americans, two common *EPHX2* haplotypes with significant and opposing relationships to ischemic stroke risk were identified. In Whites, two common haplotypes showed suggestive indication of an association with ischemic stroke risk but, as in African-Americans, these relationships were in opposite direction. These findings suggest that multiple variants exist within or near the EPHX2 gene, with greatly contrasting relationships to ischemic stroke incidence; some associated with a higher incidence, others with a lower incidence.

INTRODUCTION

Stroke is the leading cause of severe disability and the third leading cause of death, accounting for 1 of every 15 deaths in the United States. Stroke mortality and morbidity are greater among certain sub-groups, such as the elderly and African-Americans. Ischemic stroke accounts for about 80% of all strokes (1). While the genetic basis of some Mendelian forms of stroke has been elucidated (2–5), the mutations implicated in these rare disorders likely contribute little to the overall prevalence of stroke in the population-at-large. Thus, identification and

---

[*]**Correspondence to**: Myriam Fornage, PhD, Institute of Molecular Medicine, The University of Texas Health Science Center – Houston, 2121 W. Holcombe Blvd., Houston, TX 77030, (713) 500–2463 (tel.), (713) 500–2424 (fax.). myriam.fornage@uth.tmc.edu

characterization of genes implicated in the more common forms of stroke is critical to improve understanding of the pathophysiologic processes leading to its development.

Epoxyeicosatrienoic acids (EETs) are lipid metabolites of arachidonic acid that are synthesized in vascular endothelial cells by the cytochrome P450 system (6,7). They function as potent vasodilators and participate in the regulation of vascular tone (8). They have anti-inflammatory properties (9), modulate platelet function during hemostasis (10), and promote cell proliferation (11). In the brain, both vascular endothelial cells and astrocytes provide a carefully regulated supply of EETs to the cerebral microvasculature (12,13). EETs have been shown to regulate cerebral blood flow (12) and, through their mitogenic properties, may contribute to angiogenesis in the brain (14). Hence, they may play a role in predisposition to and/or recovery from cerebrovascular injury.

Hydrolysis of the EETs to their corresponding diols by epoxide hydrolases regulates EETs levels and represents a major mechanism by which the biological effects of EETs are attenuated (7). Inactivation of EETs by EPHX2 has been proposed to contribute to ischemic brain injury in rodent models, as both EPHX2 gene deletion and pharmacological inhibition of the enzyme reduced infarct size after focal cerebral ischemia (15). We have reported that sequence variation and altered levels of expression of the EPHX2 gene may contribute to kidney and brain injury in a rat model of genetic stroke (16). Multiple polymorphisms in the human EPHX2 gene have been identified, which result in amino-acid substitutions in the protein (17,18). Significant variations in enzyme activity and stability have been shown to be associated with these variants in *in vitro* assays (17,19). We have recently described an association between a functional polymorphism of the human EPHX2 gene with coronary artery calcification, a marker of atherosclerosis and independent predictor of stroke risk (20). In light of the mounting evidence of the role of endogenous EETs in processes intimately connected to cerebrovascular function, we have examined the association of polymorphisms in the EPHX2 gene with incidence of ischemic stroke in the large prospective Atherosclerosis Risk in Communities (ARIC) study.

RESULTS

Means, standard errors and proportions of established stroke risk factors are presented in Table 1 for incident ischemic stroke cases and the cohort random sample (CRS). Mean values for all variables except triglycerides, HDL cholesterol, and LDL cholesterol were significantly different between the 2 groups. Moreover, stroke cases had a significantly greater frequency of African-Americans, men, participants with hypertension and diabetes.

Figure 1 shows the gene location and race-specific allele frequencies (calculated in the CRS) for each of the polymorphisms genotyped in this study. Genotype frequency distributions were in accordance to Hardy-Weinberg equilibrium expectations in each race for all 12 polymorphisms (not shown). Allele and genotype frequency distributions for most of the polymorphisms significantly differed between the 2 racial groups, with the exception of SNP6 (R287Q) (Figure 1). SNP2 (R103C), SNP3, and SNP4 were moderately frequent in African-Americans but rare in Whites. SNP10 (E470G) was not polymorphic in Whites. Allele frequencies of each of the 12 polymorphisms were similar to those observed in other African-American and white populations ((20) and unpublished data).

A summary of the estimates of pairwise linkage disequilibrium (LD), as measured by D' and r^2, between each of the polymorphisms is presented by race in Figure 2. SNP 2, 3, 4, and 10 were not tested in Whites because of their low frequencies. LD levels were generally lower in African-Americans as compared to Whites, consistent with previous reports (21,22).

Table 2 presents the results of analyses of association between individual non-synonymous SNPs of the EPHX2 gene and risk for incident ischemic stroke. There was no statistically

significant association between each of the 5 non-synonymous polymorphisms in the EPHX2 gene and incident ischemic stroke in this sample of African-American and White middle aged-individuals. A rare non-synonymous variant, E470G, showed a trend toward a positive association of the A allele (470G) with a higher risk for ischemic stroke (HRR=2.9; P=0.13) in African-Americans. This trend was observed in all 3 analytical models. The E470G variant was not observed in Whites. Given the rare frequency of this variant, very large sample sizes are required to obtain statistically meaningful numbers of individuals carrying the 470G allele. The numbers of individuals carrying the 470G allele in this sample was 11, including 5 stroke cases.

We next investigated whether the combined effects of common SNPs in the EPHX2 gene significantly influence incident clinical ischemic stroke risk in African-Americans and Whites. To this end, haplotypes were constructed from a selected subset of SNPs. Consistent with approaches to reduce risk of false positive findings related to low frequency haplotypes (23), only haplotypes with a frequency greater or equal to 3% were considered in the association analyses. Nine common haplotypes, accounting for 93% of all African-American chromosomes, and 5 common haplotypes, accounting for nearly 98% of all white chromosomes, met this criterion. Table 3 presents the haplotype frequencies and results of the analyses of association between common haplotypes of the EPHX2 gene and ischemic stroke case status by race. In African-Americans, two common haplotypes showed significant associations with risk for ischemic stroke: Haplotype ATACGGT (frequency=3.2%) was associated with a statistically significantly lower risk of ischemic stroke (P=0.03), while haplotype ACACAGT (frequency=11.5%) was associated with a statistically significantly higher risk of ischemic stroke (P=0.02). Adjustment for covariates did not significantly affect these results. Cox proportional hazard analyses modeling the ratios of hazard rates of incident clinical ischemic stroke between individuals carrying 0, 1, or 2 copies of these haplotypes, and appropriately weighting for the sampling design, corroborated these findings. The ratio of hazards of incident clinical ischemic stroke per copy of haplotype ATACGGT was 0.34 (95% CI: 0.11; 1.00), while that of haplotype ACACAGT was 1.74 (95%CI: 1.11; 2.72) (Table 4). Haplotype ATACGGT was not uniquely tagged. Haplotype ACACAGT was uniquely tagged by SNP9, which encodes a synonymous substitution at Alanine 425. We then directly evaluated the association of SNP9 with stroke incidence in African-Americans. Consistent with the haplotype analyses results, individuals carrying at least one copy of the A allele had a 1.70 fold increased hazard rate of ischemic stroke (95% CI=1.06, 2.75; P=0.03). This association was statistically significant in unadjusted and minimally adjusted models, but was no longer statistically significant after adjusting for known stroke risk factors (HRR=1.60; 95% CI=0.90, 2.90; P=0.10). In particular, adjustment for hypertension status significantly diminished the association of SNP9 with ischemic stroke (HRR=1.40; 95% CI=0.84, 2.85; P=0.19), although no association of this SNP with hypertension status, systolic blood pressure, or diastolic blood pressure was detected in either the CRS, or the stroke cases (not shown).

In Whites, there was no statistically significant association between common patterns of variation in the EPHX2 gene and incident ischemic stroke (Table 3.B). However, haplotype AAGTA, the most frequent haplotype in this race group, showed a trend toward an association with a higher risk for ischemic stroke (P=0.05–0.08). This haplotype is uniquely tagged by SNP5. In single variant analyses, this variant also showed a trend toward a positive association with incident ischemic stroke in Whites (HRR AA vs GG= 1.6 (95%CI: 0.98; 2.6; *P*=0.06)). Similar to the dual associations of EPHX2 with ischemic stroke risk observed in African-Americans, a common haplotype (AGGTA) showed trend toward an association with lower risk for ischemic stroke in Whites (P=0.06). This haplotype was not uniquely tagged.

DISCUSSION

There is increasing evidence that the genetic architecture underlying susceptibility to common complex diseases reflects the influence of both common and rare DNA variants (24,25), but the relative contribution of common versus rare variants remains to be determined. Multiple, generally rare, polymorphisms in the EPHX2 gene have been identified, which result in amino acid substitutions in the EPHX2 protein (17,18). Significant variation in enzyme activity and stability has been shown to be associated with these variants in experimental assays *in vitro* (17,19). In this study, we have genotyped 5 non-synonymous polymorphisms of the EPHX2 gene, with established or potential relevance to enzyme structure and function, as well as 7 additional common variants spanning the EPHX2 gene sequence. We have used a combination of single polymorphism- and haplotype-based methodologies to investigate the association of variation in the EPHX2 gene with incident ischemic stroke in African-Americans and Whites from the ARIC cohort.

An association between the E470G variant with incident ischemic stroke in African-Americans is suggested. Adjusting for effects of known stroke risk factors, individuals carrying the 470G allele had a 2 to 3 fold hazard rates of incident ischemic stroke compared to those who did not, although statistical significance for this effect was not achieved. Because this variant is rare in African-Americans and not present in our sample of Whites, our power to detect a statistically significant effect was limited, and very large sample sizes are required to obtain statistically meaningful numbers of individuals carrying the 470G allele. Future studies in the entire ARIC cohort will clarify the role of this variant in ischemic stroke. Interestingly, *in vitro* enzymatic assays showed that the 470G EPHX2 isoform had a significantly greater hydrolase activity but similar phosphatase activity as compared to the wild type isoform (17). Such an altered function, if present *in vivo*, may have important consequences on the regulation of EETs levels in the brain and vasculature, and may contribute to cerebrovascular injury (12). None of the other non-synonymous polymorphisms, with known relevance to enzyme structure or function, was independently associated with incident ischemic stroke.

The relationship between EPHX2 sequence variation and ischemic stroke may be exerted through higher-order combinations of polymorphisms, rather than through their individual effects. Indeed, it has been suggested that haplotypes, rather than individual variants, define functional units of genes (26, 27). Thus, we next investigated the association of common EPHX2 haplotypes with stroke risk. In African-Americans, two common haplotypes with opposing effects on ischemic stroke risk were identified. Haplotype ATACGGT was associated with a significant decreased hazard rate for incident ischemic stroke, either alone or after adjusting for other known stroke risk factors. The variant(s) responsible for this effect is (are) unknown. Additional markers that uniquely tag this haplotype may be helpful in the identifying the underlying causative variant(s). Haplotype ACACAGT was associated with a significant increased hazard rate of incident ischemic stroke, and this effect was largely independent from the effect of other stroke risk factors. This haplotype is uniquely tagged by SNP9 (Ala425Ala), which was significantly associated with incidence of ischemic stroke in single polymorphism analyses, suggesting that this variant itself, or a variant(s) in LD with it, contributes to ischemic stroke risk in African-Americans. On average, synonymous SNPs are more frequent than non-synonymous SNPs and it is generally thought that most synonymous mutations may be functionally neutral. However, studies in *E.coli*, yeast, and *Drosophila* support evidence for translation selection for major codons (28,29), and it has been shown that there is a strong correlation between bias in synonymous codon usage and gene expression levels (30). Moreover, given its close proximity to the junction of exon and intron 14, it is also possible that SNP9 may affect RNA splicing. Predictions from computational analysis (31, 32) revealed that the G to A substitution at SNP9 may lead to the activation of an exonic splicing enhancer (ESE), with the A allele having a 6.8 fold greater predicted SF2/ASF binding affinity than the

G allele (see Supplementary Figure 1). To our knowledge, alternate splice variants of this gene have not been characterized in humans. Further studies are needed to determine whether the Ala425Ala polymorphism (SNP9) may produce such variants.

The strong LD between E470G and SNP9 raises the possibility that association of SNP9 with ischemic stroke may represent, at least in part, the association of E470G with the disease. Several lines of evidence suggest that this is not so. First, the haplotype associated with greater stroke incidence and uniquely tagged by SNP9 carries the E470 allele not the G470 allele. Second, the association of one SNP with incident ischemic stroke is not affected when controlling for the effects of the other (not shown), suggesting independent associations of the two SNPs with the disease.

Although we could not find conclusive evidence for an association between variation in the EPHX2 gene and ischemic stroke in Whites, there was a suggestive indication for an effect of a common EPHX2 haplotype, AAGTA, on higher ischemic stroke risk. This haplotype was associated with a modestly increased ratio of hazard rates of incident ischemic stroke, and was uniquely tagged by a SNP in intron 5, which itself showed some suggestion of an association with increased ischemic stroke incidence. It is not clear whether this SNP itself, or one or more SNP in LD with it, is responsible for the observed effects. As in African-Americans, in addition to contributing to increased ischemic stroke risk, the EPHX2 gene may also encompass sequence variation that contributes to lower ischemic stroke risk in Whites, as suggested by the trend toward an association of the AGGTA haplotype with lower ischemic stroke risk. We have recently shown that this haplotype is also significantly associated with an increased risk for calcified plaque in the CARDIA cohort (Fornage et al, in preparation). Plaques vulnerable to rupture and, thus, prone to cause stroke, have been rarely found to contain calcium, and there is accumulating evidence that plaque calcification may confer plaque stability (33). Moreover, patients with calcified plaque are less likely to have symptomatic disease than those with non-calcified plaque (33–36). Our findings of a common EPHX2 haplotype associated with both increased susceptibility to calcified plaque and decreased risk for stroke are consistent with such data and justify further investigation about the role of this gene in the clinical manifestations of atherosclerosis.

Taken together, these data provide evidence that both common and rare variants are likely to play a role in the relationship between the EPHX2 gene and ischemic stroke. Moreover, they provide evidence that multiple variants exist within or near the EPHX2 gene, with greatly contrasting effects on ischemic stroke incidence; some associated with a higher incidence, others with a lower incidence. Several hypotheses may be put forth to explain these results. First, multiple functional variants of the enzyme have been identified with contrasting enzymatic activity and/or stability (17, 19). Hence, different haplotypes carrying different such variants may influence disease risk in a contrasting manner. Second, it has been recently shown that the EPHX2 is a bifunctional enzyme composed of an epoxide hydrolase domain at its C-terminus and a phosphatase domain at its N-terminus (37, 38). Thus, the presence of two functionally distinct domains in the EPHX2 gene provides the opportunity for functional variants of these domains to influence the phenotype in a distinct fashion. Third, it is also conceivable that the observed associations may not stem solely from effects of polymorphisms in the EPHX2 gene, but also from those of polymorphisms in a nearby gene, whose relationship to ischemic stroke may differ from and oppose that of EPHX2. However, recent data on whole-genome patterns of common DNA variation (39, 40) do not support the possibility that LD extends to the characterized genes near the EPHX2.

This is the first study to describe the common patterns of EPHX2 gene variation associated with incident ischemic stroke in a well-characterized epidemiological setting, such as the ARIC cohort. However, the large number of statistical tests performed warrants caution interpretation

of the findings, and underscores the need for additional studies to independently confirm the data reported here. Dual associations of EPHX2 gene variation with risk of coronary heart disease have been observed in a separate study (DC Zeldin, personal communication), lending further credibility to the hypotheses presented here. Additional studies are also needed to further investigate the functional relationships of the EPHX2 sequence variation with ischemic cerebrovascular disease. Identification of the variants influencing susceptibility to ischemic stroke is a necessary first step toward understanding the biological basis of the associations detected in this study.

MATERIAL AND METHODS

Study Population

Participants were selected from the Atherosclerosis Risk in Communities (ARIC) study, an ongoing prospective study of atherosclerosis and its sequelae. The cohort consists of a probability sample of 15,792 men and women aged 45 to 64 years at the time of initial examination. Subjects were recruited from four US communities (Suburban Minneapolis, Minnesota; Washington County, Maryland; Forsyth County, North Carolina, and Jackson, Mississippi). In the first three communities, the sample reflects the demographic composition of the community. In Jackson, only black residents were enrolled. The baseline examination, conducted from 1987 to 1989, included a home interview to ascertain demographic, socioeconomic, and cardiovascular risk factors, and medical history; a clinical examination; and a blood draw for laboratory determination. The cohort was followed by annual phone interviews, a clinic examination every 3 years, and hospital and death certificate surveillance. A detailed description of the ARIC study design and methods has been previously published (41). Subjects with prevalent coronary heart disease, stroke, or transient ischemic attacks at the baseline visit were excluded from these analyses (N=1,434).

Incidence of clinical ischemic stroke was determined by review of hospital records and survey of discharge lists from local hospitals and death certificates from state vital statistics offices for potential cerebrovascular events (41,42). Validation of stroke hospitalizations in ARIC is described elsewhere (42). Briefly, a hospitalization was eligible for validation if there was: 1) a discharge diagnosis indicative of stroke; 2) a discharge summary indicative of a possible stroke such as: transient ischemic attack (TIA), cerebral vascular disease, cerebral hemorrhage, cerebral infarction (list not exhaustive); and/or 3) a CT or MRI with cerebral vascular findings or admission to the neurological intensive care unit. Records for eligible hospitalizations were abstracted at a central center by a trained nurse using a standardized computerized algorithm. Medical records were also reviewed by a trained physician. Disagreement between computer diagnosis and physician review were adjudicated by a second physician. 315 validated incident clinical ischemic strokes were identified through 2001. A stratified random sample of the ARIC cohort (Cohort Random Sample (CRS), n=1,021) was used as the comparison group for the clinical cases. Selection of this cohort random sample (CRS) was stratified on the basis of ultrasound examination of carotid arteries, age, and sex.

Data Collection

Fasting levels of total triglycerides, total and HDL cholesterol were measured by enzymatic methods. LDL cholesterol was calculated. Seated blood pressure was measured three times with a random-zero sphygmomanometer and the last 2 measurements were averaged. Hypertension status was defined as a systolic blood pressure level $\geq$ 140 mm Hg, a diastolic blood pressure level $\geq$ 90 mm Hg, and/or treatment with antihypertensive medication. Diabetes was defined as a fasting blood glucose levels $\geq$ 126 mg/dl, a non-fasting blood glucose level $\geq$ 200 mg/dl, and/or a history of or treatment for diabetes. Ratio of waist (umbilical level) and hip (maximum buttocks) circumferences were calculated as a measure of adiposity. Cigarette

smoking status was analyzed by comparing current smokers to individuals who had formerly or never smoked.

EPHX2 gene polymorphisms and genotyping

Twelve polymorphisms spanning the EPHX2 gene sequence were selected from previously published work, public databases, as well as from our own resequencing efforts. Polymorphisms were selected for genotyping based on their known or likely functional significance (non-synonymous SNPs), frequency (>5%), and/or their ability to tag haplotype blocks based on r^2 values (43). Genotyping was performed using the TaqMan assay (Applied Biosystems) or the MassARRAY assay (Sequenom) as previously described (44). Genotype determination was done blinded to case-cohort status. For each polymorphism, race-specific genotype frequencies were estimated by gene counting in the CRS. Agreement of the genotype frequencies with Hardy-Weinberg equilibrium expectations was tested using a χ^2 goodness-of-fit test.

Statistical Methods

Weighted analyses were used to draw inferences to the entire ARIC cohort from whom the CRS was drawn. The weight for a given stratum of the CRS equaled the total number of eligible participants in the study population stratum divided by the total number of participants sampled from that stratum. The proportions, means, and standard errors of established stroke risk factors were reported as weighted results for incident clinical ischemic stroke cases and the CRS.

Individual SNPs analyses

For each polymorphism, Cox proportional hazards models were used to estimate the ratios of hazard rates of incident clinical ischemic stroke between individuals in the different genotype categories. The method of Barlow was used to adjust for the sampling strategy in the Cox proportional hazards models (45). Three general models were estimated: Model 1 was an unadjusted model containing only the genotype category; Model 2 included age, sex, and field center as covariates; Model 3 included covariates in model 2 and hypertension status, diabetes status, smoking status, and waist-hip ratio. For low frequency polymorphisms, heterozygous and homozygous for the minor allele were combined into a single category. All analyses were performed separately by race.

Haplotype analyses

Within each race, only polymorphisms with a minor allele frequency ≥ 10% were further considered for the haplotype analyses. SNPs with pairwise $r^2 > 0.8$ were considered redundant and only one of them was included in the haplotype analyses. Hence, SNPs included in the haplotype analyses constituted a minimal set of highly informative markers (tagSNPs) while minimizing redundant data. This set was composed of SNP 1, 2, 4, 7, 9, 11, and 12 in African-Americans, and of SNP 1, 5, 6, 7, and 11 in Whites.

Significance of haplotype effects on each of the traits was evaluated using a regression method for unphased haplotypes based on score statistics (46,47). This method uses the Generalized Linear Model (GLM) framework but accounts for haplotype ambiguity by modeling the probabilities of the possible haplotype pairs per subject (46). Score statistics were constructed to test the null hypothesis of no haplotype effects on the probability of being a stroke case. Statistical significance was evaluated by a permutation test that takes into account the number of haplotypes and, therefore, adjusts for multiple comparisons. Adjustment for covariates followed the same scheme as described above. Only individuals whose genotype was non-missing in 80% of the selected tagSNPs were included in haplotype-based analyses (N=439

and 813, for African-Americans and Whites, respectively). All analyses were performed separately by race.

While the score method has the advantage that permutation P values for significance tests can be easily computed, it does not estimate the regression parameters, hence, the magnitude of effects of haplotypes on traits. Each individual was assigned the most probable pair of haplotypes, as estimated by EM algorithm implemented in the score method above. For those haplotypes showing a significant association with ischemic stroke by the score test, haplotype effects size was estimated by the Barlow method described above, modeling the ratios of hazard rate of incident clinical ischemic stroke between individuals carrying 0, 1, or 2 copies of the given haplotype.
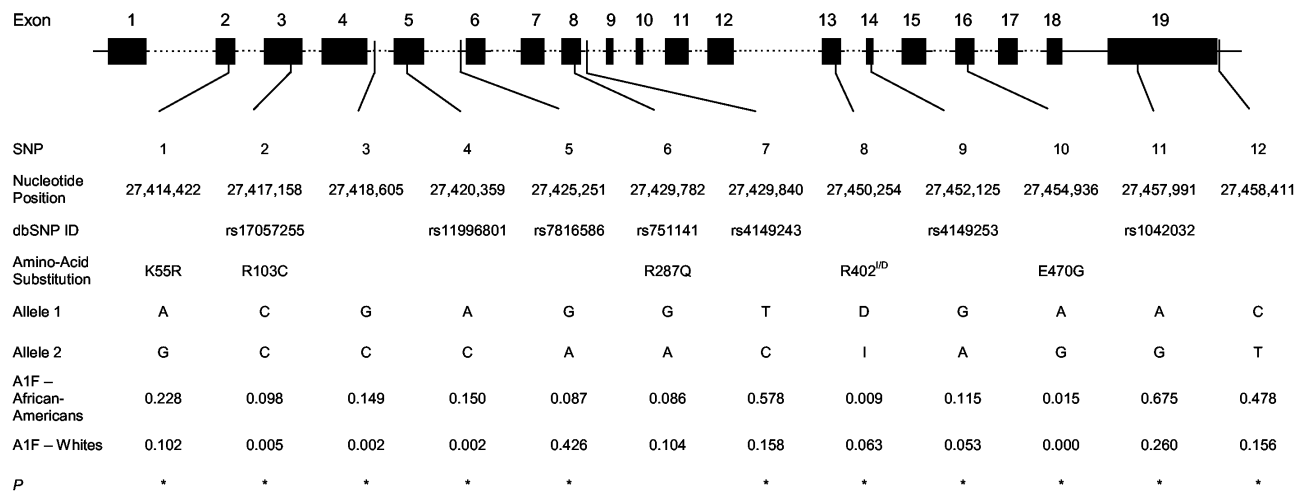
REFERENCES

1. AHA. Heart Disease and Stroke Statistics - Update. American Heart Association; Dallas: 2005.

2. Levy E, Carman MD, Fernandez-Madrid IJ, Power MD, Lieberburg I, van Duinen SG, Bots GT, Luyendijk W, Frangione B. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. Science 1990;248:1124–1126. [PubMed: 2111584]

3. Craig HD, Gunel M, Cepeda O, Johnson EW, Ptacek L, Steinberg GK, Ogilvy CS, Berg MJ, Crawford SC, Scott RM, et al. Multilocus linkage identifies two new loci for a mendelian form of stroke, cerebral cavernous malformation, at 7p15-13 and 3q25.2-27. Hum. Mol. Genet 1998;7:1851–1858. [PubMed: 9811928]

4. Joutel A, Corpechot C, Ducros A, Vahedi K, Chabriat H, Mouton P, Alamowitch S, Domenga V, Cecillion M, Marechal E, et al. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. Nature 1996;383:707–710. [PubMed: 8878478]

5. Kobayashi Y, Momoi MY, Tominaga K, Shimoizumi H, Nihei K, Yanagisawa M, Kagawa Y, Ohta S. Respiration-deficient cells are caused by a single point mutation in the mitochondrial tRNA-Leu (UUR) gene in mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes (MELAS). Am. J. Hum. Genet 1991;49:590–599. [PubMed: 1715668]

6. Capdevila JH, Falck JR, Estabrook RW. Cytochrome P450 and the arachidonate cascade. Faseb J 1992;6:731–736. [PubMed: 1537463]

7. Zeldin DC. Epoxygenase pathways of arachidonic acid metabolism. J. Biol. Chem 2001;276:36059–36062. [PubMed: 11451964]

8. Imig JD, Navar LG, Roman RJ, Reddy KK, Falck JR. Actions of epoxygenase metabolites on the preglomerular vasculature. J. Am. Soc. Nephrol 1996;7:2364–2370. [PubMed: 8959626]

9. Node K, Huo Y, Ruan X, Yang B, Spiecker M, Ley K, Zeldin DC, Liao JK. Anti-inflammatory properties of cytochrome P450 epoxygenase-derived eicosanoids. Science 1999;285:1276–1279. [PubMed: 10455056]

10. Heizer ML, McKinney JS, Ellis EF. 14,15-Epoxyeicosatrienoic acid inhibits platelet aggregation in mouse cerebral arterioles. Stroke 1991;22:1389–1393. [PubMed: 1750047]

11. Davis BB, Thompson DA, Howard LL, Morisseau C, Hammock BD, Weiss RH. Inhibitors of soluble epoxide hydrolase attenuate vascular smooth muscle cell proliferation. Proc. Natl. Acad. Sci. USA 2002;99:2222–2227. [PubMed: 11842228]

12. Alkayed NJ, Birks EK, Hudetz AG, Roman RJ, Henderson L, Harder DR. Inhibition of P-450 arachidonic acid epoxygenase decreases baseline cerebral blood flow. Am. J. Physiol 1996;271:H1541–1546. [PubMed: 8897950]

13. Medhora M, Narayanan J, Harder D. Dual regulation of the cerebral microvasculature by epoxyeicosatrienoic acids. Trends Cardiovasc. Med 2001;11:38–42. [PubMed: 11413051]

14. Zhang C, Harder DR. Cerebral capillary endothelial cell mitogenesis and morphogenesis induced by astrocytic epoxyeicosatrienoic Acid. Stroke 2002;33:2957–2964. [PubMed: 12468797]

15. Otsuka T, Sugo N, Koehler RC, Hurn PD, Traystman RJ, Alkayed NJ. Soluble Epoxide Hydrolase gene deletion is protective against experimental cerebral ischemia. Stroke 2003;34:301.

16. Fornage M, Hinojos CA, Nurowska BW, Boerwinkle E, Hammock BD, Morisseau CH, Doris PA. Polymorphism in soluble epoxide hydrolase and blood pressure in spontaneously hypertensive rats. Hypertension 2002;40:485–490. [PubMed: 12364351]

17. Przybyla-Zawislak BD, Srivastava PK, Vazquez-Matias J, Mohrenweiser HW, Maxwell JE, Hammock BD, Bradbury JA, Enayetallah AE, Zeldin DC, Grant DF. Polymorphisms in human soluble epoxide hydrolase. Mol. Pharmacol 2003;64:482–490. [PubMed: 12869654]

18. Sandberg M, Hassett C, Adman ET, Meijer J, Omiecinski CJ. Identification and functional characterization of human soluble epoxide hydrolase genetic polymorphisms. J. Biol. Chem 2000;275:28873–28881. [PubMed: 10862610]

19. Srivastava PK, Sharma VK, Kalonia DS, Grant DF. Polymorphisms in human soluble epoxide hydrolase: effects on enzyme activity, enzyme stability, and quaternary structure. Arch. Biochem. Biophys 2004;427:164–169. [PubMed: 15196990]

20. Fornage M, Boerwinkle E, Doris PA, Jacobs D, Liu K, Wong ND. Polymorphism of the soluble epoxide hydrolase is associated with coronary artery calcification in African-American subjects: The Coronary Artery Risk Development in Young Adults (CARDIA) study. Circulation 2004;109:335–339. [PubMed: 14732757]

21. Shifman S, Kuypers J, Kokoris M, Yakir B, Darvasi A. Linkage disequilibrium patterns of the human genome across populations. Hum. Mol. Genet 2003;12:771–776. [PubMed: 12651872]

22. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, et al. Linkage disequilibrium in the human genome. Nature 2001;411:199–204. [PubMed: 11346797]

23. Comeron JM, Kreitman M, De La Vega FM. On the power to detect SNP/phenotype association in candidate quantitative trait loci genomic regions: a simulation study. Pac. Symp. Biocomput 2003:478–489. [PubMed: 12603051]

24. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat. Genet 2003;33:177–182. [PubMed: 12524541]

25. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 2004;305:869–872. [PubMed: 15297675]

26. Clark AG. The role of haplotypes in candidate gene studies. Genet. Epidemiol 2004;27:321–333. [PubMed: 15368617]

27. Schaid DJ. Evaluating associations of haplotypes with traits. Genet. Epidemiol 2004;27:348–364. [PubMed: 15543638]

28. Comeron JM, Aguade M. An evaluation of measures of synonymous codon usage bias. J. Mol. Evol 1998;47:268–274. [PubMed: 9732453]

29. Comeron JM, Kreitman M. The correlation between synonymous and nonsynonymous substitutions in Drosophila: mutation, selection or relaxed constraints. Genetics 1998;150:767–775. [PubMed: 9755207]

30. Akashi H. Gene expression and molecular evolution. Curr. Opin. Genet. Dev 2001;11:660–666. [PubMed: 11682310]

31. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 2003;31:3568–3571. [PubMed: 12824367]

32. Nalla VK, Rogan PK. Automated splicing mutation analysis by information theory. Hum. Mutat 2005;25:334–342. [PubMed: 15776446]

33. Beckman JA, Ganz J, Creager MA, Ganz P, Kinlay S. Relationship of clinical presentation and calcification of culprit coronary artery stenoses. Arterioscler. Thromb. Vasc. Biol 2001;21:1618–1622. [PubMed: 11597935]

34. Nandalur KR, Baskurt E, Hagspiel KD, Phillips CD, Kramer CM. Calcified carotid atherosclerotic plaque is associated less with ischemic symptoms than is noncalcified plaque on MDCT. Am. J. Roentgenol 2005;184:295–298. [PubMed: 15615991]

35. Hunt JL, Fairman R, Mitchell ME, Carpenter JP, Golden M, Khalapyan T, Wolfe M, Neschis D, Milner R, Scoll B, et al. Bone formation in carotid plaques: a clinicopathological study. Stroke 2002;33:1214–1219. [PubMed: 11988593]

36. Shaalan WE, Cheng H, Gewertz B, McKinsey JF, Schwartz LB, Katz D, Cao D, Desai T, Glagov S, Bassiouny HS. Degree of carotid plaque calcification in relation to symptomatic outcome and plaque inflammation. J. Vasc. Surg 2004;40:262–269. [PubMed: 15297819]

37. Cronin A, Mowbray S, Durk H, Homburg S, Fleming I, Fisslthaler B, Oesch F, Arand M. The N-terminal domain of mammalian soluble epoxide hydrolase is a phosphatase. Proc. Natl. Acad. Sci. USA 2003;100:1552–1557. [PubMed: 12574508]

38. Newman JW, Morisseau C, Harris TR, Hammock BD. The soluble epoxide hydrolase encoded by EPXH2 is a bifunctional enzyme with novel lipid phosphate phosphatase activity. Proc. Natl. Acad. Sci. USA 2003;100:1558–1563. [PubMed: 12574510]

39. Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA, Cox DR. Whole-genome patterns of common DNA variation in three human populations. Science 2005;307:1072–1079. [PubMed: 15718463]

40. . The International HapMap Project. Nature 2003;426:789–796. [PubMed: 14685227]

41. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am. J. Epidemiol 1989;129:687–702. [PubMed: 2646917]

42. Rosamond WD, Folsom AR, Chambless LE, Wang CH, McGovern PG, Howard G, Copper LS, Shahar E. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. Stroke 1999;30:736–743. [PubMed: 10187871]

43. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am. J. Hum. Genet 2004;74:106–120. [PubMed: 14681826]

44. Fornage, M.; Doris, PA. Single-Nucleotide Polymorphism Genotyping. In: Fennell, JP.; Baker, AH., editors. Hypertension - Methods and Protocols. Humana Press; Totowa, NJ: 2005. p. 159-172.

45. Barlow WE. Robust variance estimation for the case-cohort design. Biometrics 1994;50:1064–1072. [PubMed: 7786988]

46. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am. J. Hum. Genet 2002;70:425–434. [PubMed: 11791212]

47. Lake SL, Lyon H, Tantisira K, Silverman EK, Weiss ST, Laird NM, Schaid DJ. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. Hum. Hered 2003;55:56–65. [PubMed: 12890927]

Figure 1.

Exon	1	2	3	4	5	6	7 8	9 10 11 12	13 14 15	16 17 18	19

SNP	1	2	3	4	5	6	7	8	9	10	11	12
Nucleotide Position	27,414,422	27,417,158	27,418,605	27,420,359	27,425,251	27,429,782	27,429,840	27,450,254	27,452,125	27,454,936	27,457,991	27,458,411
dbSNP ID		rs17057255		rs11996801	rs7816586	rs751141	rs4149243		rs4149253		rs1042032	
Amino-Acid Substitution	K55R	R103C				R287Q		R402$^{I/D}$		E470G		
Allele 1	A	C	G	A	G	G	T	D	G	A	A	C
Allele 2	G	C	C	C	A	A	C	I	A	G	G	T
A1F – African-Americans	0.228	0.098	0.149	0.150	0.087	0.086	0.578	0.009	0.115	0.015	0.675	0.478
A1F – Whites	0.102	0.005	0.002	0.002	0.426	0.104	0.158	0.063	0.053	0.000	0.260	0.156
P	*	*	*	*	*		*	*	*	*	*	*

Figure 1.
Gene location and minor allele frequency (in CRS only) of 12 genotyped polymorphisms in the EPHX2 gene by race. Exons are indicated by boxes and are numbered. Polymorphism position is given relative to the NCBI35 chromosome 8 sequence. Non-synonymous substitutions are shown. I/D corresponds to the presence or absence of 3 nucleotides (TCG) at position 27,450,254 resulting in the insertion of an Arginine residue at position 402 of the enzyme. A1F: frequency of allele 1. *: P<0.05 for difference in genotype distributions between the two races.
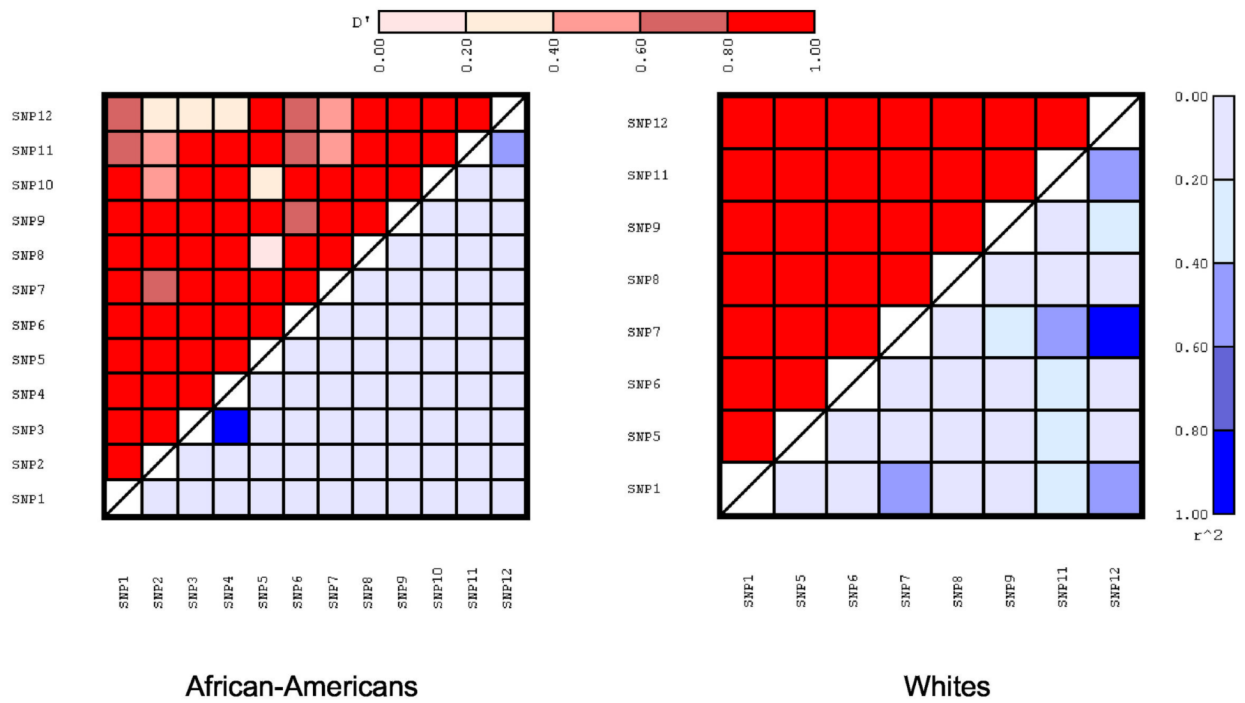
Figure 2.

Figure 2.
Pairwise estimates of linkage disequilibrium by race (African-Americans, left panel; Whites, right panel). D' is shown in the upper diagonal and r^2 is shown in the lower diagonal. SNP10 (E470G) was not included in Whites.

Table 1.

Baseline characteristics of the study sample.

	Incident Ischemic Stroke Cases N=315	Cohort Random Sample N=1,021	P
Age at Visit 1 (years)	57.3(0.6)	53.8(0.10)	<0.0001
Waist-Hip ratio	0.95(0.007)	0.92(0.002)	0.0001
HDL cholesterol (mmol/l)	1.33(0.05)	1.36(0.01)	0.53
LDL cholesterol (mmol/l)	3.58(0.14)	3.48(0.03)	0.46
Triglycerides (mmol/l)	1.55(0.10)	1.46(0.03)	0.38[*]
%African-Americans	47.0	26.1	0.0001
%men	53.3	43.4	0.001
%current smoker	34.1	24.8	0.001
%hypertensive	63.8	29.8	<0.0001
%diabetic	32.3	12.0	0.0004

Data are presented as means (standard error) or proportions

HDL: high density lipoprotein

LDL: low density lipoprotein

[*] P value for difference in geometric means

Table 2.

Analysis of association between non-synonymous polymorphisms of the EPHX2 gene and incident ischemic stroke.

	African-Americans			Whites		
	Hazard Ratios		*P*	Hazard Ratios		*P*
K55R	**GG+AG**	**AA**		**GG+AG**	**AA**	
Model 1[*]	0.9(0.6;1.5)	Ref.	0.81	0.9(0.6;1.4)	Ref.	0.61
Model 2[†]	0.9(0.6;1.4)	Ref.	0.87	0.8(0.5;1.3)	Ref.	0.42
Model 3[‡]	0.9(0.5;1.6)	Ref.	0.61	0.8(0.5;1.4)	Ref.	0.45
R103C	**TT+TC**	**CC**		**TT+TC**	**CC**	
Model 1	0.8(0.5;1.3)	Ref.	0.38	1.5(0.4;5.7)	Ref.	0.51
Model 2	0.8(0.5;1.4)	Ref.	0.45	1.7(0.5;6.3)	Ref.	0.43
Model 3	0.8(0.5;1.5)	Ref.	0.49	2.0(0.5;8.5)	Ref.	0.33
R287Q	**AA+AG**	**GG**		**AA+AG**	**GG**	
Model 1	1.1(0.6;1.8)	Ref.	0.85	0.8(0.5;1.3)	Ref.	0.49
Model 2	1.1(0.6;1.8)	Ref.	0.84	1.0(0.6;1.6)	Ref.	0.95
Model 3	1.1(0.5;2.0)	Ref.	0.75	1.1(0.6;1.8)	Ref.	0.75
R402$^{I/D}$	**R402$^{I/D}$**	**R402$^{D/D}$**		**R402$^{I/I}$+R402$^{I/D}$**	**R402$^{D/D}$**	
Model 1	1.2(0.2;6.9)	Ref.	0.82	0.8(0.4;1.3)	Ref.	0.29
Model 2	1.0(0.2;6.0)	Ref.	0.98	0.8(0.5;1.5)	Ref.	0.54
Model 3	0.7(0.1;5.3)	Ref.	0.68	0.9(0.5;1.7)	Ref.	0.75
E470G	**AG(470E/G)**	**GG(470E)**				
Model 1	2.2(0.6;8.2)	Ref.	0.12			
Model 2	2.7(0.8;9.8)	Ref.	0.12			
Model 3	2.9(0.7;11.9)	Ref.	0.13			

[*] unadjusted

[†] adjusted for age, sex, and center

[‡] adjusted for age, sex, center, hypertension status, diabetes status, smoking, and waist-hip ratio.

Table 3.

Frequency distribution and results of analyses of the association between common haplotypes of the EPHX2 gene and ischemic stroke case status by race. SNPs that uniquely tag common haplotypes are shown in bold.

3.A. African-Americans

Haplotype[*]	E(Freq)	Model 1		Model 2		Model 3	
		Score[†]	P	Score	P	Score	P
A T A C G G T	0.032	-2,19	**0.03**	-2.08	**0.04**	-2.49	**0.02**
A C C C G G C	0.091	-0.53	0.6	-0.52	0.6	0.98	0.34
A C A T G G T	0.092	-0.52	0.6	-0.52	0.6	-0.06	0.94
A C C C G G T	0.043	-0.49	0.61	-0.45	0.65	0.03	0.98
G C A C G G T	0.193	-0.33	0.71	-0.4	0.7	-1.11	0.26
A C A T G G C	0.094	0.27	0.77	0.2	0.86	0.18	0.86
A C A T G A C	0.207	0.19	0.85	0.24	0.82	-0.55	0.6
A T A C G A C	0.06	0.41	0.68	0.35	0.71	0.77	0.47
A C A **C A** G T	0.115	2.33	**0.02**	2.32	**0.02**	2.21	**0.04**

3.B. Whites

Haplotype[*]	E(Freq)	Model 1		Model 2		Model 3	
		Score[†]	P	Score	P	Score	P
A G G T A	0.314	-1.45	0.14	-1.85	0.08	-1.85	0.06
G G G C G	0.092	-0.68	0.50	-0.82	0.40	-0.95	0.34
A G **A** T G	0.092	-0.68	0.48	-0.22	0.82	-0.01	0.99
A G G C G	0.056	0.96	0.33	1.28	0.23	1.41	0.16
A **A** G T A	0.422	1.91	0.06	1.99	**0.05**	1.73	0.08

[*] Haplotypes are designated by the corresponding nucleotide at SNP1, 2, 4, 7, 9, 11, and 12 in African-Americans, and SNP1, 5, 6, 7 and 11 in Whites, in that order.

[†] A negative score represents a lower than average risk for stroke; a positive score represents a higher than average risk. Model 1, 2, and 3 have been defined in Table 2.

Table 4.

Association between selected common haplotypes and incident ischemic stroke in African-Americans.

Haplotype[*]	Hazard ratio per copy of haplotype[†] (95% confidence interval)	P
A T A C G G T		
Model 1	0.34 (0.11; 1.00)	**0.05**
Model 2	0.35 (0.12; 1.02)	0.06
Model 3	0.29 (0.09; 0.91)	**0.03**
A C A C A G T		
Model 1	1.74 (1.11; 2.72)	**0.01**
Model 2	1.63 (1.05; 2.54)	**0.03**
Model 3	1.74 (0.99; 3.00)	**0.05**

[*] Haplotypes are designated by the corresponding nucleotide atSNP1, 2, 4, 7, 9, 11, and 12, in that order

[†] relative to individuals not carrying the given haplotype. SNPs that uniquely tag common haplotypes are shown in bold.