A Whole-Genome Scan for Stroke or Myocardial Infarction in Family Blood Pressure Program Families

Richard Sherva, PhD; Michael B. Miller, PhD; James S. Pankow, PhD; Steven C. Hunt, PhD; Eric Boerwinkle, PhD; Thomas H. Mosley, PhD; Alan B. Weder, MD; J. David Curb, MD; Amy Luke, MD; Alanna C. Morrison, PhD; Myriam Fornage, PhD; Donna K. Arnett, PhD

Background and Purpose—Atherothrombotic diseases, including stroke and myocardial infarction, share a common pathogenesis. Chromosomal regions have been linked to atherothrombotic diseases in family studies, and association studies have identified candidate gene polymorphisms that affect the risk of stroke and/or myocardial infarction. Using data from the Family Blood Pressure Program, we tested for chromosomal regions linked to the composite phenotype of stroke or myocardial infarction in a large set of hypertensive families.

Methods—Nonparametric linkage analysis was implemented in MERLIN, which tests for excess allele-sharing among affected siblings. Empirical distributions based on gene dropping simulations were constructed for each test statistic, and the $-\log_{10}$ of the associated probability values were compared.

Results—Analyses were based on 9607 individuals in 226 black, 395 Hispanic, and 207 white families; 106 families had multiple affected individuals. Several regions showed linkage to stroke or myocardial infarction, most significantly in Hispanics on chromosomes 2p21 ($-\log_{10} P=3.0$) and 7q21.1 ($-\log_{10} P=2.8$), 9q32 in blacks and Hispanics ($-\log_{10} P=3.0$), 11p13 in blacks ($-\log_{10} P=2.1$), and 12q24.33 in whites ($-\log_{10} P=3.0$).

Conclusions—There is statistically significant evidence for loci affecting stroke or myocardial infarction on chromosomes 2, 9, and 12. (Stroke. 2008;39:1115-1120.)

Key Words: cerebrovascular accident ■ epidemiology ■ linkage (genetics) ■ LOD score ■ myocardial infarction

Stroke and myocardial infarction (MI) are common manifestations of atherothrombotic disease and together are the leading cause of death and disability in the United States. There are approximately 865,000 new and recurrent MIs in the United States each year and 750,000 new and recurrent cases of stroke. Ischemic blockage accounts for approximately 80% of stroke events; approximately 10% are due to intracerebral hemorrhage; 10% result from subarachnoid hemorrhage.

Ischemic stroke and MI are similar phenotypes despite occurring in different locations. They share several pathophysiological mechanisms (e.g., inflammation, atherosclerosis, plaque rupture, thrombosis, ischemic blockage) and risk factors (hypertension, smoking, dyslipidemia, no or high alcohol consumption, inflammation, homocystinemia, hypercoagulable states, diabetes). Both phenotypes have a substantial familial component. Monozygotic twins are 1.6 times more likely to be concordant for stroke than dizygotic twins; several studies show a small, significant increase in stroke risk in those with a family history of stroke. MIs are 2.48-fold more common in individuals with a sibling who has had an MI compared with those without. The heritability of arterial and venous thrombosis is estimated at 60% using liability threshold models. Past genome scans for stroke and MI have identified potential regions and genes of interest in various populations.

Combining stroke and MI allows detection of genes common to the pathogenesis of both and may increase power for detection of susceptibility loci. Therefore, we performed a genome scan for stroke or MI in black, Hispanic, and white families enrolled in the Family Blood Pressure Program (FBPP), the largest sample of families with data available for linkage analysis on stroke and MI. A combination of liability threshold and affected sibpair (ASP) models was used to detect linked regions.
Materials and Methods

Recruitment and Participants
FBPP data were obtained by pooling data from 4 National Heart Lung and Blood Institute hypertension genetics studies, The Genetic Epidemiology Network (GenNet), the Genetic Epidemiology Network of Atherosclerosis (GENOA), the Hypertension Genetic Epidemiology Network (HyperGEN), and Stanford Asian Pacific Program in Hypertension and Insulin Resistance (SAPPHIRE). Recruitment and inclusion criteria varied across networks and are published elsewhere. Informed consent was obtained from all participants.

Data Collection
All networks measured a standard set of 95 phenotypes. Stroke and MI data were obtained through self-report based on the following questions: “Have you ever been told by a doctor that you had a heart attack?” and “Have you ever been told by a doctor that you had a stroke or transient ischemic attack?” Diabetes status was also assessed in this manner. Lifestyle factors such as smoking status, alcohol use, and family structure were ascertained by interview. Anthropomorphic data were recorded according to standardized protocols.

Other Measurements
Systolic and diastolic blood pressures were measured using an automated device with a consistent protocol across networks. Blood pressure was measured 3 times on each participant and averaged for analysis. Lipids were measured spectrophotometrically using enzymatic methods (Cobas Mira analyzer; Roche Diagnostics) in HyperGEN, and an enzymatic method (COBAS FARA centrifugal analyzer; Roche Diagnostics) in SAPPHIRE. Other measurements included body mass index, cholesterol, triglycerides, and smoking status.

Genotyping
Genotyping was done by the National Heart, Lung and Blood Institute Mammalian Genotyping Service (Marshfield, Wis). Genotypes for 391 short tandem repeat polymorphisms were determined by automated polymerase chain reaction and scanning fluorescence detection.

Statistical Analysis
Race-specific and combined Kong-Cox logarithm of the odds (LOD) scores were calculated using the nonparametric LOD feature in MERLIN, which tests for excess allele sharing among an ASP. LOD scores were calculated within racial groups and summed to obtain whole-sample estimates. In MERLIN, empirical probability values were also calculated for Kong-Cox LOD scores based on gene-dropping simulations in which a single, fully informative marker is repeatedly, randomly “dropped” from parents to offspring in pedigrees with the same structure and affection status as the actual families. A LOD score was computed for each simulation under the null hypothesis of no linkage between marker and trait loci.

Results
The overall FBPP characterized 9607 individuals. There were 226 black, 395 Hispanic, and 207 white families; 880 individuals reported a history of stroke or MI and 83 individuals reported a history of both events. Table 1 shows baseline characteristics.

Table 1. Descriptive Characteristics of the FBPP Sample by Race (N=9607)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Black (SD)</th>
<th>Hispanic (SD)</th>
<th>White (SD)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>50.5 (13.4)</td>
<td>55.2 (11.9)</td>
<td>53.3 (13.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Body mass index, kg/m^2</td>
<td>31.2 (7.4)</td>
<td>30.9 (6.2)</td>
<td>29.7 (6.2)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>130 (22.6)</td>
<td>131 (22.5)</td>
<td>127 (20.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Diastolic blood pressure, mm Hg</td>
<td>73 (12.5)</td>
<td>71 (10.6)</td>
<td>72 (11.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cholesterol, mg/dL</td>
<td>199.8 (44.4)</td>
<td>205.2 (47.1)</td>
<td>204.4 (38.2)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Drinks per week</td>
<td>2.9 (8.9)</td>
<td>8.9 (14.7)</td>
<td>3.4 (6.6)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Stroke, N</td>
<td>189 (4.6)</td>
<td>76 (4.6)</td>
<td>117 (3.0)</td>
<td>0.0002</td>
</tr>
<tr>
<td>MI, N</td>
<td>188 (4.6)</td>
<td>111 (6.8)</td>
<td>252 (6.5)</td>
<td>0.0005</td>
</tr>
<tr>
<td>Diabetes, %</td>
<td>17.9</td>
<td>60.9</td>
<td>10.7</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Current smoker, %</td>
<td>27.7</td>
<td>17.2</td>
<td>13.9</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*p values are for analysis of variance tests for overall differences in outcome variables by race.

Discussion
In the FBPP, we found 3 significant linkage regions for the composite phenotype that are consistent with prior reports (summarized in Table 3). Although our LOD scores in these previously identified linkage regions do not achieve genome-wide significance, nor were they our most significant results, we show −log_{10} probability values of 1.9, 1.9, and 1.0 at 2q36-q37.3, 2q21.1-22, and 14q32.2, respectively. The previous 2q36-q37.3 signal was for MI or coronary heart disease in a small sample of Australian ASPs, whereas our signal was mainly in blacks. Genes of interest in the vicinity include...
insulin receptor substrate-1, the high-density lipoprotein cholesterol-binding protein, and calpain 10. Similarly, our modest signal at 14q32.2 was driven mainly by black and Hispanic families, whereas the original finding was in German families. Our replication of linkage on 2q21.1-22 was also observed in both black and Hispanic families, but not in whites. The original finding was in Finnish families ascertainment for premature coronary heart disease. This region

Figure. Kong-Cox LOD plots for stroke or MI in affected sibpairs.
Table 2. Chromosomal Regions With $-\log_{10} P$ Values of Combined Race Kong-Cox LOD Scores Greater Than 2.0

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Location (cM)</th>
<th>$-\log_{10} \text{Kong-Cox}$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>243</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>94</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>99</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>45</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>166</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>97</td>
<td>2.1</td>
<td></td>
</tr>
</tbody>
</table>

contains the gene thrombospondin type-1 domain-containing protein 7B precursor. The protein’s function has not been characterized, although thrombospondin type-1, in addition to its role in platelet aggregation and coagulation, interacts with platelet-derived growth factor to stimulate smooth muscle cell proliferation\cite{19} and with platelet glycoprotein IV (CD36) to inhibit angiogenesis.\cite{20} The latter findings are of particular interest because CD36, discussed in more detail subsequently, lies directly under our chromosome 7 peak. If the thrombospondin type-1 domain on this protein has similar functions, it could conceivably affect atherogenesis, stroke, and MI through multiple pathways.

In addition to the genes located in previously identified regions, our novel linkage regions also contain interesting candidates. One of these lies directly under the peak at 20 cM on chromosome 4. HS3ST1 encodes the rate-limiting enzyme heparan sulfate 3-O-sulfotransferase-1, which in turn controls production of the anticoagulant heparan sulfate (HS\textsuperscript{act}) in endothelial cells. HajMohammadi et al created an HS3ST1 double knockout mouse strain; although their tissue HS\textsuperscript{act} expression was dramatically reduced, the knockout mice showed a similar hemostatic profile as wild-type mice. The authors suggest that deficiency in the activity of the human HS3ST1 gene product might exert an affect only when combined with deficiencies for other HS3ST1 isozymes or other anticoagulants/fibrinolytics.\cite{21}

Probably the most promising candidate gene in a novel linkage region is located directly under the 93 cM peak on chromosome 7. CD36 is a receptor for thrombospondin type-1, which affects the adhesion of platelets to collagen.\cite{22} Substantial research has been conducted on this gene, mostly in the context of metabolic syndrome and cellular fatty acid uptake. Griffin et al found an increase in macrophage CD36 transcripts under high glucose conditions in human vascular lesion cells, suggesting this as a mechanism for accelerated atherosclerosis in patients with diabetes.\cite{23} Also, CD36 is a scavenger receptor specific to oxidized low-density lipoprotein.\cite{24} Animal models collectively show that CD36 deficiency underlies insulin resistance, defective fatty acid metabolism, and hypertriglyceridemia.\cite{25-27} Finally, the evidence suggesting this gene may interact with a domain of a protein located in another of our stroke/MI-linked regions may warrant molecular study of both genes.

We found racial/ethnic differences in our findings, consistent with the observed variation in lifestyle and, perhaps, genetic factors that underlie these phenotypes across these groups. Stroke and MI rates vary by race/ethnicity. Blacks and Hispanic Americans have higher stroke rates than whites, although stroke rate differences for Hispanics are mostly attributable to Mexican Americans having an increased incidence of intracerebral hemorrhage and subarachnoid hemorrhage.\cite{28} In contrast, MI rates are similar across the ethnic groups.\cite{1} Although we adjusted for lifestyle and other risk factors in our analysis, it is likely that these adjustments did not fully capture the nongenetic sources of variation among ethnic groups. Additionally, gene–environment interactions could explain the race-specific findings.

The different recruitment and inclusion protocols that gave rise to the samples in each ethnic group could also have influenced the results. Although similar inclusion criteria were used in black and white families, Hispanic families were eligible only if they had 2 or more members with adult-onset diabetes. This fact could also imply different genetic mech-

Table 3. Linkage Peaks From Previous Genome Scans for Atherothrombotic Disease End Points and $-\log_{10} P$ Values at Corresponding Locations in the Current Study

<table>
<thead>
<tr>
<th>Chromosomal Location</th>
<th>LOD Score</th>
<th>Phenotype</th>
<th>Population</th>
<th>$-\log_{10} P$ Value From Current Study (Δ location)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q36-q37.3</td>
<td>2.63</td>
<td>MI/CHD</td>
<td>53 Australian sibpairs\cite{2}</td>
<td>1.9 (15 cM)</td>
</tr>
<tr>
<td>2q21.1-22</td>
<td>3.0</td>
<td>Premature CHD</td>
<td>156 Finnish families\cite{3}</td>
<td>1.9 (5 cM)</td>
</tr>
<tr>
<td>5q12</td>
<td>4.4</td>
<td>Ischemic stroke</td>
<td>476 Icelandic patients and 438 relatives\cite{13}</td>
<td>0.7</td>
</tr>
<tr>
<td>5q12</td>
<td>2.06</td>
<td>Ischemic stroke</td>
<td>56 Swedish families\cite{14}</td>
<td>0.7</td>
</tr>
<tr>
<td>10q23</td>
<td>2.06</td>
<td>CHD</td>
<td>99 Indo-Mauritian families\cite{3}</td>
<td>0.0</td>
</tr>
<tr>
<td>13q12-13</td>
<td>2.86</td>
<td>MI</td>
<td>Icelandic females\cite{11}</td>
<td>0.1</td>
</tr>
<tr>
<td>14q32.2</td>
<td>3.9</td>
<td>MI</td>
<td>513 western European families\cite{10}</td>
<td>1.0</td>
</tr>
<tr>
<td>16p13</td>
<td>3.06</td>
<td>CHD</td>
<td>99 Indo-Mauritian families\cite{3}</td>
<td>0.8</td>
</tr>
<tr>
<td>Xq23-26</td>
<td>2.46</td>
<td>Premature CHD</td>
<td>156 Finnish families\cite{3}</td>
<td>NA</td>
</tr>
</tbody>
</table>

Δ location indicates difference between location of previously reported LOD score and location of peak in the current study; difference equals zero if not listed.

CHD indicates coronary heart disease; NA, not applicable.
organisms caused stroke and MI in different ethnic groups and may explain the differences in linkage results. As an example, the signal in the region of the CD36 gene was observed only in Hispanics, 60% of whom are diabetic.

We confirmed our results using a liability threshold model implemented in SOLAR, calculated an empirical LOD score using 10,000 simulations, and compared the agreement between the magnitudes of the adjusted liability threshold LOD and Kong-Cox LOD scores. There was evidence for linkage using the liability threshold model at all the locations identified using ASP models, and the genomewide correlation between the $-\log_{10}$ of the probability values for LOD scores calculated in threshold and at the corresponding location in ASP analysis was 0.44, suggesting that the findings reported here have good agreement with another independent method of linkage analysis.

**Study Limitations**

Phenotypic measurement methods were not optimal in this study. First, stroke and MI history data were obtained through self-report and were not verified. Also, people with a history of stroke or MI that resulted in failure to meet inclusion criteria (including fatal stroke or MI) were not enrolled. In addition, stroke subtypes were not differentiated. Knowing that approximately 80% of all strokes are ischemic, any genes unique to this subtype would likely be overrepresented.

Combining stroke and MI events may or may not be a weakness of these analyses. Although much evidence exists that stroke and MI are influenced by similar pathophysiological mechanisms, there may still be significant heterogeneity in the genetic mechanisms that influence risk, and certain risk factors have stronger effects on either stroke or MI. For example, serum cholesterol appears to be a more important MI risk factor, and its effect seems to be opposite on ischemic and hemorrhagic stroke risk. The importance of other risk factors also differs between stroke and MI. Pulse pressure is a significant predictor of MI but not stroke, whereas overall blood pressure predicts stroke rate but not pulse pressure-adjusted MI. LOD scores from multivariate-adjusted liability threshold models were very similar to those calculated in threshold and at the corresponding location in ASP analysis.

In conclusion, we presented results of a whole-genome scan for stroke or MI conducted in a large, ethnically diverse pool of families. We used 2 fundamentally different linkage models to show statistically significant evidence for loci affecting stroke or MI on chromosomes 2, 9, and 12.

**Source of Funding**

Support for this research provided by The National Heart Lung and Blood Institute cardiovascular disease genetics training grant 5 T32 HL007972.

**Disclosures**

None.

**References**


A Whole-Genome Scan for Stroke or Myocardial Infarction in Family Blood Pressure Program Families  

*Stroke*. 2008;39:1115-1120; originally published online March 6, 2008;  
doi: 10.1161/STROKEAHA.107.490433  
*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231  
Copyright © 2008 American Heart Association, Inc. All rights reserved.  
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:  
http://stroke.ahajournals.org/content/39/4/1115

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Stroke* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:  
http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Stroke* is online at:  
http://stroke.ahajournals.org/subscriptions/