Cryptogenic Stroke in Relation to Genetic Variation in Clotting Factors and Other Genetic Polymorphisms Among Young Men and Women

Harland Austin, DSc; Marc I. Chimowitz, MBChB; Holly A. Hill, MD, PhD; Seemant Chaturvedi, MD; Lawrence R. Wechsler, MD; Robert J. Wityk, MD; Elizabeth Walz, MD; Janet L. Wilterdink, MD; Bruce Coull, MD; Cathy A. Sila, MD; Panos Mitsias, MD; Bruce Evatt, MD; W. Craig Hooper, PhD; for the Genetics and Stroke in the Young Study Group

Background and Purpose—The purpose of the present study was to compare the prevalences of genetic polymorphisms in persons with cryptogenic stroke with those among stroke patients with evidence of large-artery occlusive disease or an unequivocal cardioembolic source (noncryptogenic stroke).

Methods—We compared the prevalences of genetic polymorphisms thought to be related to thrombi formation in young stroke patients with evidence of large-artery occlusive disease or an unequivocal cardioembolic source (noncryptogenic stroke; controls; n=79) with those in young stroke patients without such sources (cryptogenic stroke; cases; n=67). Common variations in the genes encoding factor V, prothrombin, angiotensin I–converting enzyme, 5,10-methylene-tetrahydrofolate reductase, endothelial cell nitric oxide synthase, tissue plasminogen activator, plasminogen activator inhibitor–1, and fibrinogen were evaluated. We also compared the allele prevalence of these genes among all stroke patients with those among a large pool of historical controls assayed for these genes.

Results—None of these genetic polymorphisms was statistically significantly related to cryptogenic stroke. With respect to a comparison of all ischemic stroke with historical controls, only the prevalence of tissue plasminogen activator D allele among stroke subjects was statistically significantly higher than that of the historical controls (P=0.0014).

Conclusions—These findings generally do not support the hypothesis that genes associated with a prothrombotic state are risk factors among a subgroup of young people with stroke of undetermined cause. Except for the D tissue plasminogen activator allele, the findings also indicated that these genetic factors are unrelated, or only weakly related, to all ischemic stroke. (Stroke. 2002;33:2762-2769.)

Key Words: coagulation ■ factor V ■ genetics ■ stroke

Approximately 30% of ischemic stroke is considered cryptogenic,1 ie, no specific cause can be identified, and cryptogenic stroke is more common among young patients.2–4 These young patients frequently have a large infarct in the territory of a major intracranial artery, yet cerebral angiography and transesophageal echocardiography do not reveal large-artery occlusive disease or an unequivocal cardioembolic source. Because the vascular imaging studies are unrevealing, it is thought that many of these patients are in a prothrombotic state. However, hemostatic markers of a prothrombotic tendency—eg, deficiencies in protein C, protein S, and antithrombin III—are rarely found.5

A mutation in the factor V gene (factor V Leiden) has been implicated as a cause of venous thrombosis.6,7 A recently described polymorphism in the prothrombin gene has been linked to increased levels of prothrombin and increased risk of venous thrombosis.8–10 The purpose of the present study was to compare the prevalences of these genetic polymorphisms in persons with cryptogenic stroke with those among stroke patients with evidence of large-artery occlusive disease or an unequivocal cardioembolic source (noncryptogenic stroke). We hypothesized that strokes of undetermined cause...
cause are more likely to be related to a genetic prothrombotic state than are strokes arising in the context of cardiac or arterial disease. We studied young stroke patients because \( \approx 50\% \) of strokes are cryptogenic in the young. In addition to the factor V and prothrombin genes, we evaluated the angiotensin I–converting enzyme (ACE) gene, 5,10-methylene tetrahydrofolate reductase gene (MTHFR), endothelial cell nitric oxide synthase (eNOS) gene, tissue plasminogen activator (tPA) gene, plasminogen activator inhibitor-1 (PAI-1) gene, and HaeIII polymorphisms of the \( \beta \)-fibrinogen gene. Each of these genes is important in maintaining vascular tone and/or hemostasis and hence may be important in the etiology of thrombotic disease and stroke.

**Materials and Methods**

**Study Subjects**

Persons 18 to 50 years of age with a first ischemic stroke diagnosed between August 1995 and November 1998 and treated at 1 of 12 medical centers in the United States were the study subjects. All stroke patients were enrolled in the study within 4 months of their diagnosis.

Age-eligible stroke patients were invited to participate at each site. Participation involved providing a blood sample, completing a questionnaire, and granting permission for review of medical records. Proxy responses to the questionnaire (usually from a spouse) were allowed for subjects unable to complete the questionnaire. Informed consent was obtained from all study subjects, and the study protocol was reviewed and approved by the appropriate human investigation committees.

**Diagnostic Evaluation for Ischemic Stroke**

All study subjects had an extensive diagnostic evaluation to determine the cause of stroke. The evaluation included a physical examination; CT or MR imaging of the brain; cerebrovascular imaging (typically carotid and transcranial Doppler ultrasound, MR angiography, or cerebral angiography); transesophageal or transthoracic echocardiography with bubble contrast; determination of cholesterol, triglycerides, and high- and low-density lipoprotein; and a complete blood count.

Study subjects were classified into 1 of 2 stroke categories: cryptogenic or noncryptogenic. Stroke of determined cause (noncryptogenic) included patients with carotid or vertebralbasilar occlusive disease (\( \geq 50\% \) stenosis or occlusion ipsilateral to stroke), craniovascular artery dissection, moyamoya disease, a clear-cut cardioembolic source, or penetrating artery disease. Clear-cut cardioembolic sources included atrial fibrillation, mitral stenosis, prosthetic valve, myocardial infarction within the preceding 6 weeks, intracardiac clot, left atrial spontaneous echo contrast, \( \text{ventricular aneurysm, and bacterial endocarditis. Penetrating artery disease was diagnosed in patients with a history of hypertension or diabetes who were presented with pure motor hemiparesis or pure sensory stroke, a subcortical infarct } < 1.5 \text{ cm that correlated with the clinical defect, and no evidence of proximal artery occlusive disease or a cardioembolic source.} \)

Cryptogenic stroke included patients with a normal diagnostic evaluation and those who may have had some risk factors for stroke but in whom the pathophysiology of the stroke was uncertain. The latter group included patients with a right-to-left interarterial shunt without overt deep vein thrombosis, an atrial septal aneurysm, ventricular dyskinesia or akinesia without mural thrombus, elevated anticardiolipin antibodies, oral contraceptive use, migraine, and low level of protein S, protein C, or antithrombin III. The diagnostic workup results were reviewed by a single neurologist (M.I.C.) for consistency in classifying the type of stroke. This review was performed with the neurologist blinded to the genetic laboratory results.

**Laboratory Methods**

DNA was extracted from 10 mL whole blood according to the manufacturer’s protocol using the Puregene kit from Gentra Systems, Inc, and stored at \(-20^\circ\text{C}\). Polymerase chain reaction techniques were used at the Centers for Disease Control (CDC) genetic laboratory to amplify DNA fragments in the factor V, prothrombin, ACE, MTHFR, eNOS, tPA, PAI-1, and \( \beta \)-fibrinogen genes.

The presence of the factor V Leiden (denoted as A, wild-type allele as G) was determined by restriction enzyme analysis using Mnl. Restriction enzyme analysis using HindIII was used to determine the presence of prothrombin 20210 mutation. Restriction enzyme analysis using HinfI was used to determine the polymorphisms of the C677T MTHFR gene (with T denoting the rare allele and C the wild-type allele). The PAI-1 polymorphisms were identified by a BseRI restriction site using a G-to-A substitution (by primer) at nucleotide 2489. HaeIII restriction enzyme analysis was used to identify the A \( ^{16} \) allele (denoted as H2, wild-type allele as H1) of the \( \beta \)-fibrinogen gene.

Analyses for the tPA insertion/deletion (I/D) and ACE I/D polymorphisms did not require restriction enzyme procedures because the mutations are insertions rather than point mutations. The I/D polymorphisms of the ACE gene were determined by sizing the amplified product in an ethidium bromide–stained 1.5% agarose gel. For tPA, if the insertion was present, a fragment size of 412 bp was seen, whereas the fragment size for a deletion was 112 bp.

For eNOS, the area of tandem repeats in intron 4 was amplified by polymerase chain reaction as described by Wang et al. Heterozygosity or homozygosity was determined by the size of the amplified segment. For this analysis, the wild-type allele is 420 bp, and the rare allele is 393 bp.

**Statistical Methods**

The study is a case-control design. The cases are patients whose stroke was of unknown cause (cryptogenic); the controls are patients with stroke of determined cause (noncryptogenic).

Internal comparisons of subjects with cryptogenic and noncryptogenic stroke were based on the distribution of genotypes (NN, MN, and MM, where N is a normal allele and M is a mutant allele). The odds of cryptogenic stroke for those heterozygous for the mutation (MN) or those homozygous for the mutation (MM) genotypes divided by the odds for those with the wild-type genotype (NN) were obtained. These odds ratios (ORs) were compared for whites and blacks, and if they were not statistically different, we report race-adjusted ORs. We computed a race-adjusted OR for a dominant genetic model (NN versus MN and MM) and, where possible, a recessive genetic model (NN versus MN and MM). All of these analyses were done by conditional (on race) logistic regression as implemented by Stata.

We also obtained the overall prevalence of each “mutant” allele among all stroke subjects. We compared this prevalence in all stroke subjects with that of historical controls assayed for these alleles at the CDC genetic laboratory. That is, these external analyses are based on alleles, not people. Some of these CDC subjects were the controls in 1 of 2 case-control studies of venous thrombosis (1 published and 1 currently active). Most of these CDC historical controls were members of a large health plan and were enrolled by the CDC at their annual physical examination. We restricted the historical controls to white and black men and women between 20 and 55 years of age.

We first evaluated whether the prevalence of the allele was similar among the 3 control study populations. We then evaluated Hardy-Weinberg equilibrium among control subjects and stroke subjects. Finally, we compared the prevalence of each allele of interest among all stroke subjects and the historical controls conditional on categories of age (20 to 29, 30 to 39, 40 to 49, and 50 to 54 years), sex, and race (white versus black). These statistical analyses were done with conditional logistic regression on the allele frequencies. The prevalences of each mutant allele (with exact 95% CIs) are reported for stroke subjects and control subjects.
Diagnostic Evaluation and Stroke Subtypes

One hundred sixty-two patients were enrolled in the study, but 16 were excluded because of inadequate blood samples for genetic testing. Of the remaining 146 patients, brain imaging consisted of CT and MRI in 97 patients (66%), CT alone in 23 patients (16%), and MRI alone in 22 patients (15%); the procedures were not specified in 4 patients (3%). The strokes involved the anterior circulation in 89 patients (61%), posterior circulation in 46 patients (32%), and both circulations in 8 patients (5%); it was not specified in 3 patients (2%). Of the 97 patients with anterior circulation strokes, extracranial carotid imaging was performed in 95 patients (98%), and imaging of the ipsilateral intracranial arteries was performed in 88 patients (91%). In 54 patients with posterior circulation strokes, imaging of the vertebrobasilar system was performed in all 54 patients (100%). Echocardiography was performed in 134 of 146 patients (92%); transesophageal echo was performed in 89 patients (61%); and transthoracic echo alone was performed in 45 patients (31%). On the basis of the results of these diagnostic studies, cause of stroke was classified as cryptogenic in 67 patients and noncryptogenic in 79 patients.

The distribution of study subjects according to demographic factors and stroke diagnosis is shown in Table 1. No study subject was homozygous for factor V Leiden. The odds of having 1 defective allele was slightly higher among subjects with cryptogenic stroke compared with subjects with noncryptogenic stroke. However, this difference is small and not statistically significant ($P>0.20$). The prevalence of factor V Leiden was slightly higher among all ischemic stroke subjects compared with the historical controls (Table 3), but the difference was not statistically significant. Of 20 subjects with right-to-left shunt (19 of which were classified as cryptogenic), 2 (10%) had factor V Leiden. Among the remaining 126 subjects, 7 (6%) had factor V Leiden. This difference is not statistically significant. In general, the findings reported here changed little if we reclassified those 19 cryptogenic stroke subjects with right-to-left shunt as noncryptogenic.

Of the 146 study subjects, 5 (3.4%) were heterozygous for the prothrombin mutation (3 cryptogenic, 2 noncryptogenic; data not shown). Although the prevalence of heterozygosity for the mutation was slightly higher among subjects with cryptogenic stroke (OR, 1.8; 95% CI, 0.34 to 8.9), this finding is not statistically significant ($P>0.20$). Furthermore, the prevalence of the mutant allele among all stroke subjects was nearly identical to that among the historical controls (Table 3).

There is little or no association between type of stroke and the ACE I/D allele. The OR is not much different from its null value of unity in both the dominant and the recessive models. The prevalence of the D allele among all stroke subjects is not statistically significantly different from that among the historical controls (Table 3). There is little difference between type of stroke and the MTHFR genotypes. However, among all stroke subjects, the prevalence of the T allele is higher than that among the historical controls (Table 3). The finding is of borderline statistical significance ($P=0.08$).

Because the eNOS 393 allele was rare in our study subjects, only a dominant allele model could be evaluated. The odds of cryptogenic stroke was not meaningfully higher among subjects with a 393 allele compared with those without this mutation. In addition, the prevalence of the 393 allele was somewhat lower among all stroke subjects compared with the controls, but the difference is consistent with chance (Table 3).

The tPA D allele was slightly less common among subjects with cryptogenic versus noncryptogenic stroke, but the distribution of the tPA I/D genotypes did not differ significantly between the 2 groups. On the other hand, the odds of the D allele among all stroke subjects was 50% higher than that of the historical controls, and this difference was highly statistically significant ($P < 0.0014$; Table 3). A dominant genetic model for the tPA D allele suggests that the risk of stroke is nearly doubled for persons with at least 1 D allele (relative risk, 1.8; 95% CI, 1.1, 2.8; $P=0.01$) compared with persons with no D allele.

The odds of heterozygosity or homozygosity for the PAI-1 4G allele was about two thirds higher among cryptogenic
TABLE 2. Distribution of Subjects by Race According to Stroke Classification and Selected Genotypes, ORs, and 95% CIs

<table>
<thead>
<tr>
<th></th>
<th>Whites</th>
<th>Blackmost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cryptogenic Stroke</td>
<td>Cryptogenic Stroke</td>
</tr>
<tr>
<td></td>
<td>Yes (n=55)</td>
<td>No (n=58)</td>
</tr>
<tr>
<td></td>
<td>Adjusted OR</td>
<td></td>
</tr>
<tr>
<td>Factor V Leiden</td>
<td>G/G</td>
<td>92.7* 93.1</td>
</tr>
<tr>
<td></td>
<td>G/A</td>
<td>7.3 6.9</td>
</tr>
<tr>
<td></td>
<td>A/A</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Dominant model: adjusted OR 1.42</td>
<td>(95% CI, 0.37–5.5)</td>
</tr>
<tr>
<td>ACE (deletion)</td>
<td>I/I</td>
<td>14.6 20.7</td>
</tr>
<tr>
<td></td>
<td>I/D</td>
<td>49.1 50.0</td>
</tr>
<tr>
<td></td>
<td>D/D</td>
<td>36.4 29.3</td>
</tr>
<tr>
<td></td>
<td>Likelihood ratio test† (2 df): P&gt;0.20, Trend‡ P&gt;0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dominant model: adjusted OR 1.19</td>
<td>(95% CI, 0.51–2.8)</td>
</tr>
<tr>
<td></td>
<td>Recessive model: adjusted OR 1.08</td>
<td>(95% CI, 0.54–2.2)</td>
</tr>
<tr>
<td>MTHFR genotypes</td>
<td>C/C</td>
<td>32.7 43.1</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>54.5 37.9</td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>12.7 19.0</td>
</tr>
<tr>
<td></td>
<td>Likelihood ratio test† (2 df): P&gt;0.20, Trend‡ P&gt;0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dominant model: adjusted OR 1.28</td>
<td>(95% CI, 0.64–2.5)</td>
</tr>
<tr>
<td></td>
<td>Recessive model: adjusted OR 0.76</td>
<td>(95% CI, 0.28–2.0)</td>
</tr>
<tr>
<td>ecNOS</td>
<td>420/420</td>
<td>76.4 75.9</td>
</tr>
<tr>
<td></td>
<td>420/393</td>
<td>23.6 24.1</td>
</tr>
<tr>
<td></td>
<td>393/393</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Dominant model: adjusted OR 1.16</td>
<td>(95% CI, 0.55–2.4)</td>
</tr>
<tr>
<td>PA</td>
<td>I/I</td>
<td>21.8 19.0</td>
</tr>
<tr>
<td></td>
<td>I/D</td>
<td>49.1 50.0</td>
</tr>
<tr>
<td></td>
<td>D/D</td>
<td>29.1 31.0</td>
</tr>
<tr>
<td></td>
<td>Likelihood ratio test† (2 df): P&gt;0.20, trend‡ P&gt;0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dominant model: adjusted OR 0.72</td>
<td>(95% CI, 0.31–1.7)</td>
</tr>
<tr>
<td></td>
<td>Recessive model: adjusted OR 0.68</td>
<td>(95% CI, 0.34–1.4)</td>
</tr>
<tr>
<td>PAI</td>
<td>5G/5G</td>
<td>21.8 26.3</td>
</tr>
<tr>
<td></td>
<td>4G/5G</td>
<td>54.5 52.6</td>
</tr>
<tr>
<td></td>
<td>4G/4G</td>
<td>23.6 21.1§</td>
</tr>
<tr>
<td></td>
<td>Likelihood ratio test† (2 df): P&gt;0.20, trend‡ P&gt;0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dominant model: adjusted OR 1.7</td>
<td>(95% CI, 0.78–3.5)</td>
</tr>
<tr>
<td></td>
<td>Recessive model: adjusted OR 1.3</td>
<td>(95% CI, 0.55–3.1)</td>
</tr>
<tr>
<td>Haefll</td>
<td>H1/H1</td>
<td>61.8 67.2</td>
</tr>
<tr>
<td></td>
<td>H1/H2</td>
<td>30.9 32.8</td>
</tr>
<tr>
<td></td>
<td>H2/H2</td>
<td>7.3 0</td>
</tr>
<tr>
<td></td>
<td>Dominant model: adjusted OR 1.12</td>
<td>(95% CI, 0.53–2.4)</td>
</tr>
</tbody>
</table>

*Percentage of total.
†Overall race-adjusted test of association between genotypes and stroke classification.
‡Race-adjusted trend test for the ORs.
§n=57.
stroke subjects compared with noncryptogenic stroke subjects, but this finding is not statistically significant ($P=0.18$). In contrast, the prevalence of the 4G allele among all stroke subjects was somewhat lower than that of the historical controls, but the difference is small and not statistically significant (Table 3).

Only 1 subject was homozygous for the HaeIII H2 allele, so only the dominant genetic model was evaluated for this fibrinogen-regulating gene. The odds of cryptogenic stroke was nearly identical among subjects with at least 1 H2 allele compared with subjects without this allele. The prevalence of the H2 allele was lower among all stroke subjects compared with controls, but for this polymorphism, the data among the controls are sparse, and the difference is not statistically significant.

For the internal analysis of cryptogenic versus noncryptogenic stroke, we classified each subject according to the total number of potentially harmful genetic variants (assuming a dominant genetic model for factor V Leiden, the prothrombin mutation, and ecNOS 393 variant, and a recessive genetic model for the other 5 polymorphisms). The average number of defects was 1.69 and 1.63 for cryptogenic and noncryptogenic stroke, respectively. After adjustment for race, the difference was 0.04 and was not statistically significant ($P>0.20$). There was no increase in the odds of classification as cryptogenic stroke among stroke patients with $\geq 3$ potential genetic defects.

For the external allelic analysis, we excluded the fibrinogen gene because of many missing data. The average number of potentially harmful alleles was 2.11 and 2.00 for stroke subjects and historical controls, respectively. After adjustment for age, sex, and race, the difference was 0.12 ($P=0.12$). However, this difference is due almost entirely to the tPA D allele. When this polymorphism was excluded, the difference in potentially harmful alleles for stroke subjects and historical controls was essentially zero.

**Discussion**

This study has a number of strengths that together suggest the absence of bias in its findings. Each diagnosis was based on the best available medical information, and each was reviewed by a single neurologist (M.I.C.) to ensure consistency in the classification of stroke. The genotypic assays were done by researchers blinded to stroke classification. Thus, the findings with respect to cryptogenic stroke versus noncryptogenic stroke are very likely to be valid.

The comparison of all ischemic stroke cases with the external historical CDC controls is subject to more bias, largely because the study was not designed for this purpose. However, the genotypic assays for these stroke cases and all the historical controls were done in the same laboratory with the same techniques. Furthermore, this laboratory has done extensive work on these genes in many diverse study populations and has found a remarkable degree of homogeneity both among whites and blacks of the prevalence of the alleles reported in this article. Also, we restricted the historical controls to subjects comparable in age to the stroke subjects and adjusted all analyses for categories of age, sex, and race. For all these reasons, we believe that a comparison of these stroke patients with the historical CDC controls is valid.

An obvious limitation of this study is its small size for a contrast of cryptogenic to noncryptogenic stroke and, to a lesser extent, for a contrast of all ischemic stroke to the historical controls. For example, for the rare genetic variants (factor V Leiden, the prothrombin mutation, and ecNOS 393) the comparison of cryptogenic and noncryptogenic stroke with our sample sizes required ORs between 6 and 8 to have $\approx 80\%$ power. The more common genetic variants (ACE, MTHFR, tPA, and PAI-1) required ORs between 2 and 3 for $\approx 80\%$ power. On the other hand, the precision of our study also can be evaluated by the upper confidence limit of the ORs. For the more common genetic variants, the upper confidence limit is between 2 and 3 (see Table 2). Thus, this study does effectively exclude large associations between cryptogenic stroke and the more common genotypes. Another limitation of the small study size is that some subtypes of stroke may be related to these genetic factors, and such associations probably would have been missed.

The external allelic comparison of all stroke subjects with the historical controls had better statistical power. We esti-

---

**TABLE 3. Allele Prevalence for Selected Genetic Variants Among All Stroke Subjects and Historical Controls According to Race**

<table>
<thead>
<tr>
<th>Allele / Gene</th>
<th>All Stroke Subjects</th>
<th>Controls†</th>
<th>All Stroke Subjects</th>
<th>Controls</th>
<th>OR‡ (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor V Leiden</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n Frequency</td>
<td>113 0.035 0.015–0.069</td>
<td>2070 0.029 0.024–0.035</td>
<td>33 0.015 0.0004–0.08</td>
<td>268 0.007 0.002–0.019</td>
<td>1.22 (0.60–2.5)</td>
<td>&gt;0.20</td>
</tr>
<tr>
<td>Prothrombin/20210A allele</td>
<td>113 0.018 0.005–0.045</td>
<td>2052 0.017 0.013–0.022</td>
<td>33 0.015 0.0004–0.08</td>
<td>267 0.004 0.0005–0.013</td>
<td>0.99 (0.39–2.5)</td>
<td>&gt;0.20</td>
</tr>
<tr>
<td>ACE/D allele</td>
<td>113 0.58 0.51–0.64</td>
<td>1597 0.55 0.53–0.56</td>
<td>33 0.59 0.46–0.71</td>
<td>234 0.57 0.53–0.62</td>
<td>1.14 (0.89–1.5)</td>
<td>&gt;0.20</td>
</tr>
<tr>
<td>MTHFR/T allele</td>
<td>113 0.39 0.33–0.46</td>
<td>3188 0.34 0.33–0.35</td>
<td>33 0.14 0.06–0.24</td>
<td>336 0.11 0.09–0.14</td>
<td>1.26 (0.97–1.6)</td>
<td>0.08</td>
</tr>
<tr>
<td>ecNOS/393 allele</td>
<td>113 0.12 0.08–0.17</td>
<td>1513 0.14 0.13–0.16</td>
<td>33 0.26 0.16–0.38</td>
<td>227 0.32 0.28–0.37</td>
<td>0.81 (0.57–1.1)</td>
<td>&gt;0.20</td>
</tr>
<tr>
<td>tPA/D allele</td>
<td>113 0.55 0.48–0.61</td>
<td>1600 0.43 0.41–0.45</td>
<td>33 0.68 0.56–0.79</td>
<td>234 0.60 0.55–0.64</td>
<td>1.50 (1.2–1.9)</td>
<td>0.0014</td>
</tr>
<tr>
<td>PA-1/4G allele</td>
<td>112 0.49 0.42–0.56</td>
<td>1548 0.54 0.52–0.56</td>
<td>33 0.21 0.12–0.33</td>
<td>229 0.23 0.19–0.27</td>
<td>0.83 (0.65–1.1)</td>
<td>0.16</td>
</tr>
<tr>
<td>Fibrinogen/H2 allele</td>
<td>113 0.19 0.15–0.25</td>
<td>95 0.22 0.16–0.28</td>
<td>33 0.03 0.004–0.11</td>
<td>156 0.06 0.04–0.10</td>
<td>0.78 (0.48–1.3)</td>
<td>&gt;0.20</td>
</tr>
</tbody>
</table>

*Cryptogenic and noncryptogenic stroke subjects divided by the odds of “mutant” allele among the controls adjusted for age, race, and sex (see text) with 95% confidence interval.

†CDC historical controls (see text).

‡Odds of “mutant” allele among stroke subjects divided by the odds of “mutant” allele among the controls adjusted for age, race, and sex (see text) with 95% confidence interval.
imate that this aspect of the study had ≈80% statistical power to detect an allelic OR of 2.5 for factor V Leiden; 3.0 for prothrombin mutation; ≈1.5 for the ACE D, MTHFR T, tPA D, and PAI-1 4G allele; 1.7 for the ecNOS 393 allele; and 1.9 for the fibrinogen H2 allele. Thus, with the exception of the rare mutations, the external analysis had reasonable statistical power.

We restricted the cases to young stroke victims because we believed that the proportion of cryptogenic stroke among the young would be relatively high compared with that among older stroke victims and that it is more likely that genetic disorders would have greater influence on stroke risk in younger rather than older subjects. Thus, our findings are not readily generalizable to older stroke victims, although it seems likely that these genetic factors probably are not related to cryptogenic stroke or to stroke overall in older subjects as well. We do not know if the moderate positive association between all stroke and the tPA polymorphism among our young stroke subjects would pertain to older stroke victims.

We found no association between factor V Leiden and ischemic stroke and between factor V Leiden and cryptogenic versus noncryptogenic stroke. Most epidemiological studies of adults indicate that factor V Leiden is not a cause of stroke. In contrast, among neonates and children, factor V Leiden appears to increase the risk of ischemic stroke by ≈5-fold.21-24

The odds of the prothrombin mutation was ≈2-fold higher among those with cryptogenic stroke compared with those with noncryptogenic stroke, but the finding is not statistically significant because of sparse data. Using our historical CDC controls, we estimate that the relative risk of stroke for those with the prothrombin mutation compared with those without the mutation is 1.25—a small, imprecise relative risk with an upper 95% confidence limit of ≈3.1. Two other epidemiological studies25-26 of ischemic stroke report a modest but not statistically significant relation with the prothrombin mutation, and 1 study of children24 reports a statistically significant 5-fold increase risk. In the aggregate, these studies and our study suggest that the prothrombin mutation is associated with no or only a small increased risk of ischemic stroke in adults.

In our study, the prevalence of the ACE D allele did not differ meaningfully according to our stroke classification (59% and 57% for cryptogenic and noncryptogenic, respectively). The prevalence of the D allele among all our stroke subjects was slightly higher that it was among our historical controls, but the difference is very small and not statistically significant. Three case-control studies from the United Kingdom17-19 report a higher prevalence of the ACE D allele among stroke patients compared with controls, but in only 1 was the difference statistically significant.27 The evidence to date regarding ACE and stroke is equivocal. However, it appears that the D allele may be slightly more common among stroke victims, although it is likely that the association is weak.

Cryptogenic stroke was not associated with the MTHFR genotypes in our study, and the T allele was associated only weakly and not statistically significantly so (P=0.08) with stroke overall. Elevated homocysteine levels have been related to an increased risk of arterial disease, including stroke,20 and the TT MTHFR genotype has been related to elevated plasma homocysteine in most studies. Yet the MTHFR genotype has not been consistently related to stroke.21 Thus, our findings for stroke overall are consistent with the absence of an association with the MTHFR genotype.

We found no association between the ecNOS and the HaeIII genotypes for either cryptogenic stroke or all ischemic stroke. We are not aware of any other studies of these polymorphisms for cryptogenic stroke specifically or ischemic stroke overall. We also found little or no association between the PAI-1 genotypes and cryptogenic or all ischemic stroke. Endler et al22 reported that the 4G/4G genotype was not a risk factor for minor stroke and transient ischemic attack. On the other hand, we found a statistically significant higher prevalence of the tPA D allele among all ischemic stroke subjects compared with the CDC controls. Because of the very small probability value, this finding is not likely due to chance. This probability value is small even in light of the large number of statistical tests that we report. Elevated tPA activity is related to increased fibrinolytic activity, whereas increased TPA mass concentration has been related to increased risk of stroke in 2 prospective studies.33,34 In a European nested case-control study, tPA levels and tPA/PAI-1 complex level were higher among stroke cases (especially hemorrhagic stroke) than among matched controls.35 In a German case-control study, plasma tPA antigen, PAI-1 antigen, and PAI activity were higher among subjects with ischemic stroke compared with control subjects.36 However, tPA plasma antigen levels have not yet been related to tPA genotype. For example, we recently reported that tPA genotype was unrelated to tPA plasma antigen among our controls in a case-control study.13 Because stroke patients often have increased tPA plasma antigen, our finding of a higher prevalence of the D tPA allele among ischemic stroke cases compared with the historical controls is biologically plausible. However, the lack of a demonstrated association between genotype and antigen level and the general paucity of data on tPA genotype and stroke warrant that our positive finding be interpreted with caution.

In summary, our findings do not support the hypothesis that genes associated with a prothrombotic state are risk factors among a subgroup of young people with stroke of undetermined cause. With respect to all ischemic stroke, our study implicates the tPA D allele as a risk factor.

Appendix

Participants in the Genetics and Stroke in the Young Study in Order of Recruitment
Emory University (26 patients): M. Chimowitz, B. Stern, H. Howlett-Smith; Wayne State University (24 patients): S. Chaturvedi, B. Bertasio; University of Pittsburgh (20 patients): L. Wechsler, L. Massaro; Johns Hopkins University (17 patients): R. Wityk, J. Wemmer, K. Lane; Ohio State University (14 patients): E. Walz, M. Notestine; Rhode Island Hospital (13 patients): J. Wilterdink, J. Sarafin; University of Arizona (11 patients): B. Coull, D. Bruck; Cleveland Clinic (11 patients): C. Silia, M. Horvat; Henry Ford Hospital (10 patients): P. Mitsias, P. Marchese, K. Jones; Yale
References


Genetic Association Studies of Stroke: Hope, Signal, and Noise

With human molecular genetics, hope springs eternal. The daily revisions of the working draft of the human genome, the burgeoning catalogues of human genetic variants and the crafting of slick new bioinformatic tools together give hope that we are on the verge of an explosion of new understanding of health and disease. Probing the genome provides an express route—or so it is hoped—that will accelerate the discovery of novel biological and biochemical pathways specifying new molecular targets and directing development of novel strategies for diagnosis and prevention. For monogenic diseases, or diseases that have a strong single-gene component, this hope is being realized. However, for the complex and common chronic diseases of adulthood, such as stroke, hope has to substitute in large measure for actual results to date.

One strategy for finding causative genes for stroke is illustrated by use of the positional cloning approach, in which relatives concordant for the stroke phenotype are studied.1 In this study design, a statistical excess of sharing of specific chromosomal regions among affected family members specifies those regions as harboring likely candidate genes. Choosing samples from communities or populations with low background genetic and/or environmental heterogeneity can improve the signal-to-noise ratio for such an analysis. The linked genes are candidates for disease causation strictly because of their position within the genome and usually have no evidence of a functional role in pathophysiology. In vivo and in vitro functional testing comes later, after a specific gene has been implicated as a positional candidate gene by statistics.

A different, more intuitive strategy for clinicians trained to find the most parsimonious link between mechanism and clinical presentation is the functional candidate gene approach. Here, proteins whose functional relevance has already been demonstrated—often in a bygone era using biochemical or cellular experiments—specify the genes for study. Ideally, these functional candidate genes can have common variants that encode changes in protein structure and function in vitro and in vivo, thus providing mechanistic ballast for hypothesis testing related to disease association. The study by Austin et al represents application of this strategy.2

The authors report the results of a genetic association study of 67 young women and men with “cryptogenic” stroke, 79 controls with an unequivocal cardioembolic source, and a pool of historical controls. Genotyping was performed for common variants of 8 genes, namely the genes encoding clotting factor V, prothrombin, tissue plasminogen activator (tPA), plasminogen activator inhibitor-1, the beta chain of fibrinogen, methylenetetrahydrofolate reductase, the angiotensin converting enzyme, and endothelial nitric oxide synthase. Comparisons were subdivided according to ethnic background (white or black). There were no differences in allele frequencies for any marker tested between cases and controls. The authors then performed a second analysis, pooling the cases and controls, and comparing allele frequencies with those from ethnic reference control samples of white or black background. Of the comparisons done this way, only those involving the tPA deletion allele showed it to be more common in combined stroke subgroups than in controls. The authors concluded that the genetic factors are only weakly related to ischemic stroke.

The study is of interest for several reasons. First, genetic association studies performed for any subphenotype of stroke are relatively uncommon, and thus any new information is welcome. Furthermore, by focusing on the highly specific subphenotype of cryptogenic stroke, the authors have “stacked the deck” and increased the chance of a favorable signal-to-noise ratio for a potential genetic association. Second, implicit in the focus on young study subjects was the fact that genetics would have played a relatively more important role in their disease presentation, similarly increasing the signal-to-noise ratio. Third, the markers selected were among the most widely characterized and examined genetic markers, and some actually have been proven to have a functional consequence. The mechanistic impact of a genetic variant with known functional consequence would also increase the chance of detecting a genetic association signal. Thus, the study had at least some of the desirable attributes for a genetic association study.3

While the study provides useful negative data, there are caveats—factors that can obscure the genetic signal with noise. Despite the Herculean effort to recruit subjects, the resulting sample size of the study was relatively small, as would be expected for an extreme and specific phenotype, such as cryptogenic stroke in young adults. Thus, the largely negative results might simply indicate that any true association of the tested markers with cryptogenic stroke could not rise above the background noise inherent to both the case-control study design and the limited sample size.4 This trait could simply be too etiologically or genetically heterogeneous to hope to detect an association signal without using a much larger sample. Alternatively, there might still be a major genetic determinant of cryptogenic stroke in young people, but it was not among those tested in this report. It is this hope that propels continued research to identify new genetic loci and variants and then to test these in future studies.

Robert A. Hegele, MD, FRCP, FACP, Guest Editor
Robarts Research Institute
London, Ontario, Canada

References
Cryptogenic Stroke in Relation to Genetic Variation in Clotting Factors and Other Genetic Polymorphisms Among Young Men and Women
Harland Austin, Marc I. Chimowitz, Holly A. Hill, Seemant Chaturvedi, Lawrence R. Wechsler, Robert J. Wityk, Elizabeth Walz, Janet L. Wilterdink, Bruce Coull, Cathy A. Sila, Panos Mitsias, Bruce Evatt and W. Craig Hooper
for the Genetics and Stroke in the Young Study Group

Stroke. 2002;33:2762-2768; originally published online November 21, 2002;
doi: 10.1161/01.STR.0000038094.79901.3B
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2002 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/33/12/2762

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/