The Relationship Between Smoking and Replicated Sequence Variants on Chromosomes 8 and 9 With Familial Intracranial Aneurysm

Ranjan Deka, PhD; Daniel L. Koller, PhD; Dongbing Lai, MS; Subba Rao Indugula, PhD; Guangyun Sun, PhD; Daniel Woo, MD; Laura Sauerbeck, MS; Charles J. Moomaw, PhD; Richard Hornung, DPH; E. Sander Connolly, MD; Craig Anderson, MD; Guy Rouleau, PhD; Irene Meissner, MD; Joan E. Bailey-Wilson, PhD; John Huston III, MD; Robert D. Brown, MD; Dawn O. Kleindorfer, MD; Matthew L. Flaherty, PhD; Carl D. Langefeld, PhD; Tatiana Foroud, PhD; Joseph P. Broderick, MD; the FIA Study Investigators

Background and Purpose—The purpose of this study was to replicate the previous association of single nucleotide polymorphisms (SNPs) with risk of intracranial aneurysm (IA) and to examine the relationship of smoking with these variants and the risk of IA.

Methods—White probands with an IA from families with multiple affected members were identified by 26 clinical centers located throughout North America, New Zealand, and Australia. White control subjects free of stroke and IA were selected by random digit dialing from the Greater Cincinnati population. SNPs previously associated with IA on chromosomes 2, 8, and 9 were genotyped using a TaqMan assay or were included in the Affymetrix 6.0 array that was part of a genomewide association study of 406 IA cases and 392 control subjects. Logistic regression modeling tested whether the association of replicated SNPs with IA was modulated by smoking.

Results—The strongest evidence of association with IA was found with the 8q SNP rs10958409 (genotypic \( P = 9.2 \times 10^{-5} \); allelic \( P = 1.3 \times 10^{-5} \); OR = 1.86, 95% CI: 1.40 to 2.47). We also replicated the association with both SNPs on chromosome 9p, rs1333040 and rs10757278, but were not able to replicate the previously reported association of the 2 SNPs on chromosome 2q. Statistical testing showed a multiplicative relationship between the risk alleles and smoking with regard to the risk of IA.

Conclusion—Our data provide complementary evidence that the variants on chromosomes 8q and 9p are associated with IA and that the risk of IA in patients with these variants is greatly increased with cigarette smoking. (Stroke. 2010;41:1132-1137.)

Key Words: familial ▪ genomewide association studies ▪ intracranial aneurysm ▪ smoking

Recent genomewide association studies (GWAS) have identified several common sequence variants on chromosome 9p21 that are associated with myocardial infarction, coronary artery disease, abdominal aortic aneurysms, and intracranial aneurysms (IAs).1-6 In addition, variants on chromosomes 2q33 and 8q11 have been associated with IA in distinct populations that attain genomewide significance.1 Smoking is the most powerful environmental risk factor for ruptured and unruptured IA, with 70% to 80% of patients reporting a history of smoking, similar to the strength of the relationship between smoking and lung cancer.7-12 The relationship between sequence variants associated with the risk of IA and smoking has yet to be explored.

We sought to replicate the association of 6 variants identified in the aforementioned studies with IA in an independent case–control sample from the Familial Intracranial Aneurysm Study (FIA; www.FIAStudy.com) as well to examine the relationship of smoking with these variants and the risk of IA.

Methods

Probands with an IA were identified by 26 clinical centers (41 recruitment sites) located throughout North America, New Zealand, and Australia. To be eligible for inclusion in the FIA study,13-15 the proband was required to have additional family members who also had an IA. Exclusion criteria were used to eliminate subjects who...
had an IA due to a known genetic cause such as Ehlers-Danlos or polycystic kidney disease or as a secondary phenotype such as an association with an arteriovenous malformation. One case was selected from each of the white multiplex IA families.

White control subjects free of stroke and known IA were selected from the Greater Cincinnati/Northern Kentucky population. The methodology for control identification and enrollment has been previously published.8,16 In short, the University of Cincinnati Institute for Policy Research used random digit dialing telephone survey techniques to identify control subjects of the same sex, race, and age for comparison with cases of subarachnoid and intracerebral hemorrhage in an ongoing National Institute of Neurological Diseases and Stroke-funded study. After informed consent was obtained, each control subject or a proxy was interviewed face to face in a highly structured and identical manner. In addition to the interview, which included detailed questions about past and present cigarette smoking, blood pressure measurements and blood samples for DNA extraction were obtained. FIA cases were interviewed in an identical manner.

The cases (N=410) and control subjects (N=393) were previously genotyped using the Affymetrix 6.0 array (data not shown). These data were used to test for cryptic relatedness among the reportedly unrelated cases and control subjects and to ensure that the association analysis was not confounded by the effect of population substructure. A principal component-based analysis was performed in PLINK17 to cluster these samples along with HapMap reference samples (CEU, YRI, CHB, and JPT) to verify that the samples used in this study were derived from European ancestry (Supplemental Figure I; available at http://stroke.ahajournals.org). Five subjects (4 cases and 1 control) who did not cluster with the white samples and a CEU reference sample were excluded from further analysis. The final analysis sample consisted of 406 IA cases and 392 control subjects.

Six single nucleotide polymorphisms (SNPs; rs1429412, rs700651 on 2q33; rs10958409, rs2988506 on 8q11; rs1333040, rs10757278 on 9p21) previously associated with arterial diseases1–3 were genotyped in our sample of cases and control subjects. The rs10958409 SNP, included in the Affymetrix 6.0 array, was genotyped as part of our previously completed GWAS; the other 5 SNPs were genotyped using the TaqMan (fluorogenic 5′ nuclease) assay, and the end point results were scored on the ABI 7900HT Sequence Detection System. Blind duplicates and known control samples were included in test plates for quality assurance. Quality control metrics and SNP descriptive statistics were computed for each of the 6 SNPs. Completeness of genotyping and allele frequencies were calculated using all genotyped subjects; Hardy-Weinberg equilibrium was assessed using only the data from the control subjects. Genotype proportions at all SNPs conformed to the expectation of Hardy-Weinberg equilibrium (data not shown). Two tests of association for each SNP with IA were performed in PLINK: a 2 degree of freedom genotypic test for differences between cases and control subjects with any of the 3 observed genotypes and an allelic test (1 degree of freedom) comparing the minor allele frequency between cases and control subjects. Pairwise linkage disequilibrium between all pairs of SNPs in the 3 chromosomal regions was computed based on the \( r^2 \) statistic. We examined 6 SNPs in 3 regions, suggesting that a conservative correction for multiple testing would require a 0.05/3 (0.017) threshold for replication. We had 80% power to detect ORs of 1.42 to 1.45 across the range of minor allele frequencies typed at the \( P=0.017 \) significance threshold.

For SNPs for which we were able to replicate evidence of association, we augmented the available data with the SNP genotypes generated as part of the previously completed GWAS in the same samples. We added to the data set all SNPs within 250 kilobases (kb) upstream and downstream of a replicated SNP. These SNPs underwent similar quality review to ensure genotypic completeness and lack of deviation from Hardy-Weinberg equilibrium.

Because smoking is such an important risk factor in IA, we next performed logistic regression analyses to test whether the association of the replicated SNPs on chromosomes 8 and 9 was modulated by smoking. We used the same case–control design and a logistic regression model to test each SNP. Each model included the presence of SNP risk alleles, scored as 0=no risk allele, 1=1 risk allele (heterozygous), and 2=2 risk alleles (homozygous). The risk allele was defined as the allele more common in cases than control subjects. A log of pack-years smoked was used to evaluate the effect of smoking. For purposes of the logistic regression, persons without any history of smoking were defined as having 0.05 pack-years. Each model was adjusted for age and the data presented as OR and 95% CIs. An explicit interaction between log of pack-years and the risk allele score was tested to determine whether there was a deviation from the multiplicative effect on risk that is modeled by the logistic regression (ie, closer to additive effects on risk or greater than multiplicative interaction). We also compared the geometric mean of the log of pack-years smoked for those subjects with 1 IA as compared with those with multiple IAs.

Results

Of the 406 cases, 46.7% were male compared with 54.3% of the 392 control subjects (\( P=0.0004 \)). The mean±SD age of the cases at time of diagnosis was 50.5±11.6 versus 63.4±15.1 at the time of interview for the control subjects (\( P<0.0001 \)). At diagnosis, 47.3% of cases were current smokers and 35.2% were prior smokers versus 16.6% and 35.7%, respectively, at interview for the control subjects.
Table 1. Association of 6 SNPs on Chromosomes 2, 8, and 9 With Familial IA

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>SNP</th>
<th>Position, Mb</th>
<th>Genotype Frequencies, %</th>
<th>P Value Genotypic</th>
<th>RAF</th>
<th>P Value Allelic</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q33</td>
<td>rs1429412</td>
<td>197.9</td>
<td>GG 41.9, GA 41.9, AA 42.2</td>
<td>0.178</td>
<td>0.178</td>
<td>0.268</td>
<td>1.18 (0.88–1.56)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 13.2, Control subject 10.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2q33</td>
<td>rs700651</td>
<td>198.3</td>
<td>GG 41.9, GA 41.9, AA 44.5</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>1.33 (1.09–1.62)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 12.5, Control subject 9.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q11</td>
<td>rs10958409</td>
<td>55.5</td>
<td>AA 31.3, AG 44.5, GG 44.5</td>
<td>9.2×10⁻⁵</td>
<td>0.049</td>
<td>0.005</td>
<td>5.04 (2.00–13.16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 4.3, Control subject 1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q11</td>
<td>rs9298506</td>
<td>55.6</td>
<td>AA 67.5, AG 20.9, GG 22.1</td>
<td>0.522</td>
<td>0.522</td>
<td>0.039</td>
<td>1.24 (1.01–1.52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 2.4, Control subject 1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9p21</td>
<td>rs1333040</td>
<td>22.1</td>
<td>CC 67.5, CT 20.9, TT 12.1</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>1.33 (1.09–1.62)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 26.9, Control subject 19.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RAF indicates frequency of the more common allele in cases compared with control subjects.

(P<0.0001). Figure 1 shows the plot of the all cases who were smokers by pack-years of smoking and age at diagnosis of IA. Of the 406 cases of IA, 159 were ruptured IAs.

The association analyses, which include the genotype frequencies as well as the frequency of the risk allele, are presented in Table 1. The strongest evidence of an association with IA was found with the 8q SNP rs10958409 (genotypic OR=1.48, allelic OR=1.86, 95% CI: 1.40 to 2.47). We also found evidence of an association with both SNPs on chromosome 9p, rs1333040 and rs10757278, with rs1333040 meeting our corrected level of significance. We were not able to replicate the association of the 2 SNPs on chromosome 2q reported by Bilvugar and colleagues.1

As shown in Figure 2A, substantial support for the association to chromosome 8 was provided by the SNPs genotyped in the GWAS. Six of these SNPs, located on both sides of the index replication SNP rs10958409, achieved association probability values <0.001. These results suggest the presence of a substantial linkage disequilibrium block near the 55.5 megabase position on chromosome 8 that contains a variant associated with IA. The significance of the probability value obtained for rs10958409, as compared with that of the surrounding GWAS SNPs, suggests the frequency of the IA-predisposing allele is near that of rs10958409 (minor allele frequency = 0.199 in cases and 0.118 in control subjects). In contrast, as shown in Figure 2B, we were able to improve the evidence for an association through examination of the results from GWAS SNPs surrounding rs10757278 on chromosome 9p. Probability values as small as 4×10⁻⁴ were observed in this region in the GWAS with P=0.005 for rs10757278, the SNP reported in the previous studies. The GWAS SNP providing strongest support for an association to chromosome 9p in our study, rs2891168, surpassed the α=0.05 significance threshold corrected by the simpleM method18 for the SNPs in the 500-kb region considered. These results clearly illustrate the benefit of denser SNP coverage over a range of allele frequency values.

Statistical testing was consistent with a multiplicative relationship between the at-risk alleles and smoking (Table 2). For example, a nonsmoker with 1 risk allele for SNP rs10958409 on chromosome 8 would have an OR of 1.48 for having an IA, a 20-year smoker with no copies of this risk allele would have an OR of 5.04, and a 20-pack-year smoker with 1 at-risk allele would have an OR of 7.46 (see footnote in Table 2 for explanation). We tested for interactions of each SNP with the log of pack-years of smoking to determine if they differed from the multiplicative relationship inherent in the logistic regression models. No test reached statistical significance indicating a multiplicative relationship provided a good fit to the data (or no evidence for supramultiplicative or less than multiplicative relationship).

There was a significantly greater geometric mean of log(pkys) for smoking among the 147 subjects with >1 aneurysm (11.19) as compared with the 254 subjects with 1 aneurysm (5.93; t=2.59, P=0.010). We did not find a significant difference in prevalence of risk alleles for chromosomes 8 (rs10958409) and 9 (rs1333040, rs10757278) in those subjects with 1 aneurysm and those with >1 aneurysm.

Discussion

This study replicates the associations of SNPs on chromosomes 8 and 9 and demonstrates the powerful effect of smoking on the risk of IA in persons with these risk variants. For example, in the logistic regression model, a nonsmoker with 2 rs10757278 alleles on chromosome 9 has an OR of 1.96 for the presence of IA, whereas someone with 40
pack-years would have an OR of 13.78. Our replication of SNPs on chromosomes 8 and 9 shows that we are getting close to identifying the causal variants associated with IA. Although the exact gene variants that are associated with IA have yet to be found, it is clear that smoking greatly enhances their effect and that cessation of smoking would have a tremendous impact in prevention of IA, particularly in those at increased genetic risk. The fact that 82.5% of our IA cases were smokers at some point and 47% were current smokers speaks strongly to the opportunity for prevention.

The associations of the 2 8q SNPs with IA were initially found in 2 European cohorts from Finland and The Netherlands. However, in a Japanese sample, only rs10958409 was replicated; rs9298506 was not. In our study cohort, comprised only of white samples, the strongest evidence of association was with rs10958409 (genotypic \( P = 9.2 \times 10^{-5} \); allelic \( P = 1.3 \times 10^{-5} \); OR = 1.86, 95% CI: 1.40 to 2.47). Like in the Finnish, Dutch, and Japanese cohorts, we found the association of the same risk allele in this SNP in our sample of IA cases (Table 1). We did not find evidence of an association with rs9298506. However, with our relatively small sample size, it is premature to exclude the involvement of this SNP influencing the risk of IA. As described by Bilguvar et al., SOX17 is the closest gene within the interval of the 8q variants, which is involved in endothelium formation and maintenance.

A series of GWAS has reported association of sequence variants on 9p21 with myocardial infarction, coronary artery
Table 2. Logistic Regression Models for 3 SNPs and Smoking (Pack-Years) on Chromosomes 8 and 9 (Models Adjusted for Age)

<table>
<thead>
<tr>
<th>SNP</th>
<th>OR</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10958409 (chromosome 8q)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score per risk allele</td>
<td>1.48</td>
<td>1.06–2.07</td>
<td>0.023</td>
</tr>
<tr>
<td>20 pack-years of smoking</td>
<td>5.04</td>
<td>3.50–7.61</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age per year</td>
<td>0.93</td>
<td>0.92–0.94</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>rs10757278 (chromosome 9p)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score per risk allele</td>
<td>1.40</td>
<td>1.10–1.78</td>
<td>0.007</td>
</tr>
<tr>
<td>20 pack-years of smoking</td>
<td>5.75</td>
<td>3.99–8.29</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age per year</td>
<td>0.93</td>
<td>0.92–0.94</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>rs1333040 (chromosome 9p)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score per risk allele</td>
<td>1.37</td>
<td>1.08–1.74</td>
<td>0.012</td>
</tr>
<tr>
<td>20 pack-years of smoking</td>
<td>5.16</td>
<td>1.88–2.71</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age per year</td>
<td>0.93</td>
<td>0.92–0.94</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: The OR for any given No. of K pack-years can be calculated using the following equation: OR = \( \exp([\beta_{\text{Ln(K)}} + 2.99]) \) where \( \beta \) = the regression coefficient for log(pkyrs), \( \beta = 0.270 \) for rs10958409, \( \beta = 0.292 \) for rs10757278, and \( \beta = 0.278 \) for rs1333040. For example, the OR for 40 pack-years of smoking for subjects in the model of the rs10757278 risk allele = \( \exp(0.292[\text{Ln(40)}] + 2.99) = 7.03 \). To determine the OR for presence of 2 risk alleles of rs10757278 (homozygous state) and 40 pack-years of smoking, one would multiply \((1.40)^2 = 1.96 \) (2 risk alleles) \( \times 7.03 \), which equals an OR of 13.78.

Subjects limits the power to detect genes of smaller effect size. Nevertheless, our cases represent a unique group in that all are from families with strong familial aggregation of IA as compared with prior GWAS studies of IA, in which the majority of subjects did not have a positive family history. We believe that such a cohort may provide greater ability to detect genetic risk factors for IA and eventually will facilitate identification of potentially causal gene variants within our larger FIA families.

Another limitation of our analysis is that our control group was not perfectly matched to our IA cases. Consequently, the control subjects were older than the cases and there was a slightly greater proportion of men among our control subjects. Although the control group was identified by random digit dialing from the entire population as part of an ongoing National Institute of Neurological Diseases and Stroke-funded study, cases of intracerebral hemorrhage comprised approximately two thirds of the cases of intracranial hemorrhage that were used for identification of matched control subjects. People with intracerebral hemorrhage are more likely to be older and more likely to be male than those with subarachnoid hemorrhage, which accounts for the older age of available control subjects. That older control subjects make it more likely that these individuals do not have or would not develop an IA can be considered a strength of genetic studies. However, an age difference between cases and control subjects is a potential disadvantage when environmental covariates are correlated with age. The frequency of current smoking is inversely associated with advancing age because some people stop smoking for health reasons as they enter middle and older age. Pack-years and history of smoking, which look at lifetime exposure rather than current smoking state, help ameliorate this concern as does the adjustment for age in the logistic regression models. Finally, the lack of information about second-hand smoking for cases and control subjects limits the precision of true exposure to smoking among cases and control subjects.

In summary, we analyzed 6 sequence variants reported to be significantly associated with several vascular diseases in a sample of unrelated cases from multiplex families affected with IA and population control subjects. Our data provide complementary evidence that the variants on chromosome 8q are strongly associated with IA and variants on chromosome 9p are moderately associated with IA and that the associated risks of IA in patients with these variants are greatly increased with cigarette smoking. We did not find a significant association of the 2q variants.

Sources of Funding
This study was funded by grants from the National Institute of Neurological Diseases and Stroke (NINDS R01 NS39512; R-01-NS 36695), National Institutes of Health, Bethesda, Md; the State of Ohio TECH 04-042, Ohio Department of Development, Wright Centers of Innovation Program Computational Medicine Center for the “Cincinnati Control Cohort Study”; and by the Intramural Research Program of the National Institutes of Health, National Cancer Institutem and National Human Genome Research Institute.

Disclosures
None.
References


The Relationship Between Smoking and Replicated Sequence Variants on Chromosomes 8 and 9 With Familial Intracranial Aneurysm


Stroke. 2010;41:1132-1137; originally published online February 26, 2010;
doi: 10.1161/STROKEAHA.109.574640

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2010 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/41/6/1132

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/