Conclusions—region SNPs capture some but not all of the effect of KIV2 repeat length on Lp(a) level. There are LPA for CAAD and ischemic stroke.9,10 The apolipoprotein(a) covalently bound to apolipoprotein(a), confer increased risk preventative therapies. data may help to refine risk estimates and thus better target /H11001 by exons 4 through 5 /H11002 1) of morphism of the kringle IV type 2 domain (KIV2) encoded protein varies in size because of a copy number poly-\[98195-7720. E-mail pair@u.washington.edu\]/H11022 number of patients needed to treat and the cost to prevent one /H11017 copies. The KIV2 polymorphism has been reported to ac-\[294x695\]count for 69% of the variability in Lp(a) level,12 with increased repeat number leading to impaired liver secretion,13 lower plasma Lp(a) level,11,12,14,15 and decreased risk of myocardial infarction.16,17 Although many risk factors overlap between CAAD and coronary artery disease, the relative importance of nongenetic risk factors varies between these two disease processes,18 suggesting the possibility of differential underlying genetic diatheses. In addition, although many well-established ge-\[324x344\]netic risk factors for coronary artery disease also increase risk for CAAD,19,20 some loci, such as that on chromosome 10q11.21, appear to increase risk for coronary artery disease but not stroke;20 however, other genetic factors, such as dyslipidemia genetic risk scores, display much weaker effects on CAAD.21 Therefore, given the known association between Lp(a) levels and CAAD, and the reported association between LPA polymorphisms and coronary artery disease,22 we evaluated the role of LPA region polymorphisms in predicting Lp(a) level and CAAD risk. For polymorphisms associated

Received May 20, 2010; accepted July 29, 2010.
From the Departments of Medicine, Divisions of Medical Genetics (J.R., R.R., F.C., A.S.N., G.P.J.) and Metabolism, Endocrinology, and Nutrition (S.M.), Genome Sciences (A.S.N.), Vascular Surgery (T.H., G.P.J.), and Biostatistics (P.H.), University of Washington, Seattle, Wash; Department of Surgery (T.K.), Veterans Affairs Puget Sound Health Care System, Seattle, Wash.
J.R., R.R., and F.C. contributed equally to this work.
J.R., R.R., F.C., and G.P.J. analyzed the data and wrote the paper.
The online-only Data Supplement is available at http://stroke.ahajournals.org/cgi/content/full/STROKEAHA.110.591230/DC1.
Correspondence to Gail P. Jarvik, MD, PhD, Head, Division of Medical Genetics, University of Washington Medical Center Box 357720, Seattle, WA 98195-7720, E-mail pair@u.washington.edu
© 2010 American Heart Association, Inc.

Stroke is available at http://stroke.ahajournals.org

DOI: 10.1161/STROKEAHA.110.591230
with Lp(a), we determined whether these effects were attributable to linkage disequilibrium (LD) with the KIV2 repeat. Because of the labor-intensive, low-throughput assays required for accurate KIV2 genotyping,23 we also sought to determine whether LPA polymorphisms in LD with the KIV2 copy number polymorphism could act as an adequate surrogate for KIV2 genotype in predicting Lp(a) level. Finally, we sought to determine whether LPA polymorphisms contribute to CAAD risk because of, or in addition to, their effects on Lp(a) level.

Materials and Methods

Study Population

All study participants were ascertained at four Seattle medical centers and gave written informed consent. The University of Washington, Virginia Mason Medical Center, and the Puget Sound Veterans Affairs Health Care System human subject review boards approved this study. Characteristics of study participants, clinical covariates, and phenotyping are shown in Table 1 and are described.

Table 1. Study Participant Demographics

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(&gt;80%) Internal</td>
<td>(&lt;15%) Internal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Carotid Stenosis</td>
<td>Carotid Stenosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n=530)</td>
<td>(n=196)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>10.9%</td>
<td>17.3%</td>
<td>0.0025</td>
</tr>
<tr>
<td>Censored age</td>
<td>66.6</td>
<td>68.1</td>
<td>64.7</td>
</tr>
<tr>
<td>Mean Lp(a), nmol/L</td>
<td>91 (38)</td>
<td>93 (42)</td>
<td>47 (23)</td>
</tr>
<tr>
<td>Log. Lp(a), nmol/L</td>
<td>3.6</td>
<td>3.6</td>
<td>3.0</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>27.8</td>
<td>27.5</td>
<td>28.5</td>
</tr>
<tr>
<td>Diabetes</td>
<td>30.8%</td>
<td>27.6%</td>
<td>12.8%</td>
</tr>
<tr>
<td>Antihypertensive therapy</td>
<td>86.9%</td>
<td>79.7%</td>
<td>47.7%</td>
</tr>
<tr>
<td>Current smoker</td>
<td>28.0%</td>
<td>19.5%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Pack-years smoked</td>
<td>45.5</td>
<td>41.4</td>
<td>24.6</td>
</tr>
<tr>
<td>Lipid-lowering therapy</td>
<td>71.7%</td>
<td>63.6%</td>
<td>22.9%</td>
</tr>
<tr>
<td>HDL, mg/dL</td>
<td>46.4</td>
<td>47.0</td>
<td>51.9</td>
</tr>
<tr>
<td>LDL, mg/dL</td>
<td>99.6</td>
<td>106.1</td>
<td>117.4</td>
</tr>
</tbody>
</table>

HDL indicates high-density lipoprotein; LDL, low-density lipoprotein; Lp(a), lipoprotein(a).

Methods and Procedures

Genotyping and SNP Selection

Genotyping was performed using the Illumina HumanCVD BeadChip.28 We excluded any individual who showed >10% non-white ancestry.

Lp(a) Measurement and KIV2 Copy Number

Lp(a) concentrations in nmol/L were measured on fasting plasma samples by a double monoclonal antibody-based enzyme-linked immunosass as described.29 The apo(a) size isoforms, defined by the relative number of KIV2 repeats, were determined in a subset of 90 subjects by high-resolution sodium dodecyl sulfate-agarose gel electrophoresis followed by immunoblotting, as previously described.30,31 Subjects were oversampled for infrequent SNP genotypes to increase genetic diversity in the subsample. The repeat number of the predominantly expressed isoform was used as the analysis trait;26 in 8 individuals with equally expressed isoforms the two repeat numbers were averaged.

Statistical Analyses

All analyses were performed in R32 and PLINK. Lp(a) level was natural log-transformed because of right skew; in all analyses, Lp(a) level refers to the log-transformed values. Unless otherwise specified, Lp(a) levels were adjusted for gender and age at the time of blood draw. CAAD was adjusted for gender, censored age, body mass index, type 2 diabetes, current smoking status, and total pack-years smoked. Cholesterol levels, use of lipid-lowering medications, and use of antihypertensive medications were not included as covariates for CAAD regression models because, consistent with guidelines, individuals with CAAD in our study are generally treated to lower cholesterol38 and blood pressure39 targets than individuals without disease. Thus, including lipid and blood pressure measurements or treatment status would lead to a problematic model in which the dependent variable, case versus control status, is causal for these independent variables. Linear and logistic regressions were used to test for associations between the minor allele dose of each SNP and Lp(a) levels and CAAD status, respectively. Haplotype-based tests of association were performed in PLINK. The proportion of Lp(a) level variance explained was estimated by the difference in \( r^2 \) values between models, which accounts for the number of explanatory terms.

To account for multiple testing, we performed Bonferroni corrections or estimated false discovery rates (FDR) using the Benjamini Hochberg procedure.40 Bonferroni corrections assume independence of tests and are conservative in the presence of correlated tests, whereas FDRs estimate a family-wise error rate in the setting of LD between SNP pairs. We performed permutation tests for associations between the minor allele dose of each SNP and Lp(a) levels and CAAD status, respectively. We used the randomization test associated with PLINK to estimate the number of permutations needed to achieve a specified FDR.

We assessed the power of our sample size to detect association between SNPs and CAAD case versus control status, conservatively assuming an underlying variant that confers a relative risk of 2.0 for homozygotes and 1.5 in the heterozygotes.22,24,50 As shown in Supplementary Table 1 (available online at http://stroke.ahajournals.org), our panel of ancestry informative markers on the Illumina HumanCVD BeadChip.28 We excluded any individual who showed >10% non-white ancestry.
than in controls (P<0.05, corresponding to a FDR of 0.067 (Table 2). Using backward stepwise regression identify a multi-SNP model in which redundant associations attributable to LD were discarded, 9 SNPs were jointly associated with Lp(a) at an inclusion P=0.05, corresponding to a FDR of 0.0064. These 9 SNPs, shown in order of inclusion in Figure 1, explained 30% of the variance in loge-transformed Lp(a) levels in the complete cohort and 27% in controls.

Multiple SNPs individually and jointly predict Lp(a) level. We found that 20 out of 28 tagSNPs were associated with Lp(a) level at P=0.05, corresponding to a FDR of 0.067 (Table 2). Using backward stepwise regression identify a multi-SNP model in which redundant associations attributable to LD were discarded, 9 SNPs were jointly associated with Lp(a) at an inclusion P=0.05, corresponding to a FDR of 0.0064. These 9 SNPs, shown in order of inclusion in Figure 1, explained 30% of the variance in loge-transformed Lp(a) levels in the complete cohort and 27% in controls.

Haplotypes formed by these 9 SNPs showed significant association with Lp(a) levels but did not improve the percent variance explained compared to the multi-SNP model. The minor alleles of rs10455872 and rs3798820 mark haplotypes carrying short KIV2 alleles and are associated with elevated Lp(a) levels. In a global search of all 2- and 3-SNP haplo-
types, the minor allele haplotypes (GA and AG) of these 2 SNP were the strongest predictors of elevated Lp(a) (Table 3). To determine whether a haplotype associated with elevated Lp(a) on the major allele backgrounds of rs10455872 and rs3798820 could be identified, we analyzed all possible combinations of 2 and 3 SNP haplotypes formed from the 7 remaining SNPs. One haplotype marked by several SNPs, including GA of rs6919346 and rs2295368 and GCA of rs6919346, rs3798221, and rs2295368, was associated with higher Lp(a) level (Bonferroni-corrected \(P = 0.05\)). However, models explicitly including this haplotype did not account for much more variance in Lp(a) level (25%) than the model that included the underlying SNPs as independent additive covariates (24%).

**KIV2 Copy Number Effects**

We assayed KIV2 repeat length in a subset of 90 individuals who were approximately representative of 1496 individuals for whom Lp(a) levels were measured. The distribution of Lp(a) level was indistinguishable between these 90 individuals selected for KIV2 typing compared with the complete sample (Kolmogorov-Smirnov \(P = 0.64\)), as were the frequencies of CAAD cases and controls (Fisher test \(P = 0.66\)).

![Figure 1. Percentage of variance in lipoprotein (a) (Lp(a)) level explained by single nucleotide polymorphisms (SNPs) and kringle IV type 2 (KIV2) copy number. Curves show the percent variation in Lp(a) level as SNPs are added sequentially in backward stepwise regression (bold indicates \(P < 0.05\)). The upper gray curve and SNPs listed along the top correspond to a model including KIV2 number. The leftmost point on the upper curve corresponds to KIV2 alone.](image)

### Table 3. Haplotype Analyses for Lp(a) and CAAD

<table>
<thead>
<tr>
<th>Trait</th>
<th>SNP(s)</th>
<th>High-Risk Haplotypes</th>
<th>Frequency</th>
<th>Lp(a) OR or CAAD OR</th>
<th>(P) ([Lp(a) % Variance]†)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lp(a)</td>
<td>rs3798221/rs10455872</td>
<td>GA</td>
<td>0.017</td>
<td>1.8</td>
<td>1.3 \times 10^{-76} (21%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.076</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>Lp(a)</td>
<td>rs6919346/rs3798221/rs10455872</td>
<td>GGA</td>
<td>0.017</td>
<td>2.1</td>
<td>5.0 \times 10^{-82} (23%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAG</td>
<td>0.075</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAA</td>
<td>0.75</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Lp(a)</td>
<td>rs6919346/rs3798221/rs10455872/rs2295368</td>
<td>GAAA</td>
<td>0.34</td>
<td>0.62</td>
<td>4.1 \times 10^{-12} (25%)</td>
</tr>
<tr>
<td>Lp(a)</td>
<td>rs6919346/rs3798221/rs10455872/rs2295368</td>
<td>GACAA</td>
<td>0.27</td>
<td>0.94</td>
<td>8.3 \times 10^{-18} (25%)</td>
</tr>
<tr>
<td>CAAD</td>
<td>rs6919346/rs10455872</td>
<td>GG</td>
<td>0.075</td>
<td>2.0</td>
<td>0.020</td>
</tr>
<tr>
<td>CAAD</td>
<td>rs3123629/rs6919346/rs10455872</td>
<td>AGG</td>
<td>0.066</td>
<td>2.1</td>
<td>0.054</td>
</tr>
</tbody>
</table>

CAAD indicates carotid artery disease; Lp(a), lipoprotein(a); SNPs, single nucleotide polymorphisms; OR, odds ratio.

Haplotype analyses were performed using covariates as described in Materials and Methods.

*Haplotypes among the set tested with the largest \(\beta\) for Lp(a) or OR for CAAD.

†\(P\) are for the global test of no difference between haplotypes vs any differences and are Bonferroni-corrected for the number of haplotypes tested.

‡Conditional on the major allele (AA) background of rs3798220 and rs10455872.
distributions of all nongenetic covariates were similar between the subsample and complete sample (data not shown). Although 6 SNPs showed significantly greater minor allele frequencies in the subsample at \( P < 0.05 \) (Supplementary Table II available online at http://stroke.ahajournals.org), these SNPs did not tend to be better predictors of Lp(a) level conditional on KIV2 effects (Figure 1).

Although the KIV2 copy number polymorphism alone accounted for a greater percentage of Lp(a) level variation than the 9 SNPs described, SNPs added independent explanatory power not captured by the kringle genotype (upper curve in Figure 1). The dominant KIV2 repeat length (see Materials and Methods) showed significant association with Lp(a) level (\( P < 0.005 \)) and, conditional on this dominant repeat, there was no further information for Lp(a) prediction provided by either the short (\( P = 0.58 \)) or the long (\( P = 0.89 \)) isoform. The dominant KIV2 repeat length combined with 6 SNPs included in backward stepwise regression at \( P = 0.05 \) (FDR = 0.16) accounted for 60% of the variation in loge Lp(a) in the subset of 90 individuals for whom KIV2 genotypes were assayed and for 57% in controls alone. Although the FDR is rather liberal, the 2 weakest predictors, rs3798220 and rs10455872 (\( P = 0.0083 \) and \( P = 0.034 \), respectively), have previously been shown to predict Lp(a) level conditional on KIV2 repeat number.\cite{41,42}

Prediction of CAAD by SNPs

Five SNPs, spanning 3 separate LD blocks, showed significant association with CAAD at \( P = 0.05 \), corresponding to a FDR of 0.19 (Figure 2 and Table 2). All 5 SNPs showed significant association with Lp(a) level, and for each SNP the same allele associated with increased Lp(a) was also associated with increased risk of CAAD. At a more stringent \( P \) threshold of 0.01, 3 SNPs were significantly associated with CAAD (FDR = 0.065). The most significant SNP, rs10455872, whose minor allele confers an OR of 2.1 per allele (95% CI, 1.3–3.2; \( P = 0.0013 \)), was significant at a family-wise error rate of 0.05 (corresponding to \( P = 0.0024 \)) estimated by permutation testing.

An additive genetic risk score in which alleles of each of the 9 SNPs associated with Lp(a) level were weighted according to their effect on Lp(a) level improved prediction of CAAD risk (OR, 1.5 per unit increase in risk score; \( P = 7.9 \times 10^{-5} \)). We also performed a comprehensive search of all 2- and 3-SNP haplotypes that could be formed from the 28 tagSNPs in the region. A 2-SNP haplotype formed by rs10455872 and rs6919346 was significantly associated with CAAD (nominal \( P = 0.00021 \); \( P = 0.0024 \) after Bonferroni correction for 378 2-SNP haplotypes tested). Among 3-SNP haplotypes, the rs10455872, rs6919346, and rs3123629 system was nearly significant (nominal \( P = 1.7 \times 10^{-5} \); Bonferroni-corrected \( P = 0.054 \) for 3276 haplotypes tested).

Prediction of CAAD by SNPs Conditional on Lp(a)

After accounting for Lp(a) level, CAAD status was no longer associated with genetic variation in the 6q26-27 region. When Lp(a) level was included as a covariate in CAAD analyses, the minimum \( P \) value across the 28 tagSNPs analyzed was 0.088 (FDR = 0.1). In a model including both Lp(a) level and the genetic risk score described, Lp(a) level (OR 1.40 per unit increase; \( P = 2.6 \times 10^{-5} \)), but not the risk score (OR 1.08; \( P = 0.52 \)), was associated with CAAD. Neither the 2-SNP haplotype system formed by rs10455872 and rs6919346 (nominal \( P = 0.37 \)) nor the 3-SNP haplotype system formed by rs10455872, rs6919346, and rs3123629 system was significantly associated with CAAD (nominal \( P = 0.24 \); Bonferroni-corrected \( P = 0.15 \) for 3276 haplotypes tested).
by the rs10455872, rs6919346, and rs3123629 (nominal
$P=0.313$) was a significant predictor of CAAD after account-
ing for Lp(a) level. Among all possible 2- and 3-SNP haplo-
types, the minimum $P$ values observed after condition-
ing on Lp(a) level ($P=0.044$ and 0.0024, respectively) were
well within the expectation for the null distribution.

Discussion

We identified 9 SNPs that are significantly associated with
Lp(a) level when analyzed jointly and a partially overlapping
set of 6 SNPs that were significantly associated with Lp(a)
level when analyzed conditionally on KIV2 repeat length.
Among the 9 SNPs predictive of Lp(a) independent of KIV2
length, 5 overlapped with the set of 7 SNPs described by
Clarke et al. Among the 6 SNPs predictive of Lp(a)
conditional on KIV2 repeat, 4 overlapped with previous
reports. In particular, rs10455872, which marks a haplo-
type carrying a low KIV2 copy number, and rs3798220,
which is in LD with both the KIV2 copy number poly-
morphism and the $SLC22A3$-$LPAL2$-$LPA$ haplotype
associated with coronary artery disease, are strongly
associated with Lp(a) level. Interestingly, both of these SNPs
continued to be associated with Lp(a) level after accounting for
KIV2 repeat length, consistent with previous observations.
It has been proposed that the nonsynonymous SNP rs3798220
may affect protein stability, whereas rs10455872 may be in
LD with regulatory variants. Alternatively, their continued
association with Lp(a) level after regressing on KIV2 effect
may reflect inaccuracies in measurement of or statistical
modeling of KIV2 repeat length.

Although we identified multiple SNPs that jointly predict
Lp(a) level, these SNPs accounted for just over half of the
variance in Lp(a) level explained by KIV2 alone. Our
ability to only partially capture KIV2 copy number effects
using SNP markers is consistent with the theoretical expe-
tation of high expansion and contraction mutation rates for
copy number variants and repeat polymorphisms, as well
as previous empirical data suggesting weak LD between the
KIV2 copy polymorphism and surrounding variation.
Moreover, a pentanucleotide repeat polymorphism upstream
of $LPA$ reportedly accounts for 10% to 14% of the variation
in Lp(a) level. In addition to the possibility that recurrent
mutation has placed the same KIV2 and pentanucleotide
repeat alleles on multiple haplotype backgrounds, another
plausible explanation for the observation that multiple tag-
SNPs are jointly associated with Lp(a) level is the existence
of numerous rare variants in the region, which can lead to
high rates of synthetic association.

Our finding that tagSNPs improve prediction of Lp(a) level
when combined with KIV2 genotype is consistent with a
recent report by Lanktree et al. However, in white individ-
uals, Lanktree et al report that only 36% of the variance in
Lp(a) level can be explained by SNPs and KIV2 copy
number. In contrast, in a large coronary artery disease cohort,
Clarke et al report that 36% of Lp(a) level can be explained
by rs10455872 and rs3798220 alone. In our CAAD cohort,
we found that these two SNPs explain only 22% (20% in
controls) of Lp(a) variance, whereas SNPs and the KIV2
repeat explain 60% (57% in controls). One explanation for
such heterogeneous estimates is that ascertainment on vascu-
lar disease and for KIV2 typing introduces an upward bias
into estimates of percent variance explained in Lp(a); these
limitations require that our estimates of the percent variation
in Lp(a) should be interpreted with some caution. However,
with regard to ascertainment on CAAD, the differences
between the percent variations explained in our total sample
compared to controls were only modest. In randomly ascer-
tained families in which variance explained would be ex-
pected to be less, Boerwinkle et al estimate that 69% of the
variance in Lp(a) is attributable to KIV2 copy number and
90% of the variance in Lp(a) is attributable to variation at the
LPA locus in general. These mutually inconsistent results
highlight the difficulties in accurately assaying and modeling
Lp(a) and KIV2 copy number, and underscore the importance
of identifying alternative genetic markers in the region.

Our observation that $LPA$ region SNPs and haplotypes
predict CAAD extends that of Clarke et al, who found that
6q26-27 SNPs are associated with coronary artery disease. In
fact, rs10455872 confers a higher risk of carotid disease (OR
2.1) than any SNP among a panel of 34 markers associated
with elevated low-density lipoprotein or decreased high-
density lipoprotein (highest OR 1.8). In general, the genetic
effects on CAAD that we observe are consistent with those
reported by Clarke et al and others for the $LPAL2$-$LPA$-$PLG$
locus. Although because of its low minor allele frequency
rs3798220 failed to achieve statistical significance in our
study, the OR for the minor allele (1.8) is within the
confidence interval for the effect of this variant on coronary
artery disease risk reported in other studies. We also
found, in agreement with Clarke et al, that conditioning on
the Lp(a) level eliminated the association signal in the
6p26-27 region. This has several implications. First, because
SNP genotypes undergo Mendelian randomization, their use
as an instrumental variable indicates that Lp(a) levels are
causal for CAAD as well as for coronary artery disease, as
previously suggested, rather than the reverse. Second,
because conditioning on Lp(a) level abolished any evidence
of association between tagSNPs or haplotypes and CAAD,
this argues against the presence of common structural varia-
tion, such as nonsynonymous changes, within $LPA$ that effect
protein function without affecting protein level. Third, de-
spite a lack of standardization, direct measurement of Lp(a)
surpasses use of genetic polymorphisms as a surrogate
marker.

Conclusion

In summary, we identified 9 SNPs that jointly accounted for
30% of the variation in Lp(a) levels. This was less explained
variance than by use of KIV2 repeat number alone; however,
by combining 6 SNPs with KIV2 repeat number, we ac-
counted for 60% of the variance in Lp(a) levels. Five SNPs,
al of which were strongly associated with Lp(a) levels,
predicted CAAD risk, but these associations, as well as
haplotype-based CAAD association signals within the
$LPA$-$LPA$-$PLG$ region, appeared to be driven exclusively
by interindividual variation in Lp(a) levels.
Acknowledgments

The authors thank the study participants. The authors also thank the following people for their technical assistance: Tamara Bacus, Julieann Marshall, Laura McKinstry, Loida Erhard, Karen Nakayama, Jane Ranchalis, and Jeff Rodenbaugh.

Sources of Funding

This work was funded by the National Institutes of Health (NIH) PO1 HL 30086 and RO1 HL67406. This work was supported in part by resources from the VA Puget Sound Health Care System, Seattle, Washington, including the Veteran Affairs Epidemiology Research and Information Center Program (award CSP 701S).

Disclosures

The authors have no conflicts of interest to declare.

References


Genetic Variation in LPAL2, LPA, and PLG Predicts Plasma Lipoprotein(a) Level and Carotid Artery Disease Risk

James Ronald, Ramakrishnan Rajagopalan, Felecia Cerrato, Alex S. Nord, Thomas Hatsukami, Ted Kohler, Santica Marcovina, Patrick Heagerty and Gail P. Jarvik

Stroke. 2011;42:2-9; originally published online December 2, 2010; doi: 10.1161/STROKEAHA.110.591230

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/42/1/2

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2010/12/02/STROKEAHA.110.591230.DC1

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/
Supplementary Table 1. Power to detect association between genetic variants and CAAD.

<table>
<thead>
<tr>
<th>Relative Risk</th>
<th>Risk Allele Frequency</th>
<th>Linkage Disequilibrium (r^2)</th>
<th>Power (p&lt;0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0 (1.5)</td>
<td>0.05</td>
<td>1</td>
<td>0.72</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.1</td>
<td>1</td>
<td>0.91</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.2</td>
<td>1</td>
<td>0.99</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.3</td>
<td>1</td>
<td>0.99</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.05</td>
<td>0.8</td>
<td>0.63</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.1</td>
<td>0.8</td>
<td>0.85</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.2</td>
<td>0.8</td>
<td>0.96</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.3</td>
<td>0.8</td>
<td>0.98</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.05</td>
<td>0.64</td>
<td>0.55</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.1</td>
<td>0.64</td>
<td>0.74</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.2</td>
<td>0.64</td>
<td>0.92</td>
</tr>
<tr>
<td>2.0 (1.5)</td>
<td>0.3</td>
<td>0.64</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Power estimates are based on 1000 simulations per parameter set for a sample size of 530 CAAD cases and 770 controls, assuming a population prevalence of 0.04. Association tests were performed using logistic regression as described in Methods.
Supplementary Table 2. Minor allele frequencies in the sub-sample typed for KIV2 repeat and the complete sample.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Complete sample minor allele freq</th>
<th>Sub-sample minor allele freq</th>
<th>Fisher’s exact p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3123629</td>
<td>0.30</td>
<td>0.30</td>
<td>1.00</td>
</tr>
<tr>
<td>rs7754014</td>
<td>0.23</td>
<td>0.27</td>
<td>0.20</td>
</tr>
<tr>
<td>rs7449650</td>
<td>0.33</td>
<td>0.34</td>
<td>0.68</td>
</tr>
<tr>
<td>rs6919346</td>
<td>0.15</td>
<td>0.18</td>
<td>0.29</td>
</tr>
<tr>
<td>rs3798220</td>
<td>0.02</td>
<td>0.01</td>
<td>0.37</td>
</tr>
<tr>
<td>rs11751605</td>
<td>0.16</td>
<td>0.13</td>
<td>0.40</td>
</tr>
<tr>
<td>rs4708871</td>
<td>0.03</td>
<td>0.02</td>
<td>0.48</td>
</tr>
<tr>
<td>rs7761293</td>
<td>0.47</td>
<td>0.42</td>
<td>0.25</td>
</tr>
<tr>
<td>rs6415084</td>
<td>0.47</td>
<td>0.48</td>
<td>0.94</td>
</tr>
<tr>
<td>rs3798221</td>
<td>0.20</td>
<td>0.18</td>
<td>0.63</td>
</tr>
<tr>
<td>rs10455872</td>
<td>0.08</td>
<td>0.07</td>
<td>1.00</td>
</tr>
<tr>
<td>rs1367209</td>
<td>0.28</td>
<td>0.29</td>
<td>0.73</td>
</tr>
<tr>
<td>rs1321195</td>
<td>0.15</td>
<td>0.11</td>
<td>0.23</td>
</tr>
<tr>
<td>rs783149</td>
<td>0.17</td>
<td>0.13</td>
<td>0.26</td>
</tr>
<tr>
<td>rs1358753</td>
<td>0.13</td>
<td>0.18</td>
<td>0.12</td>
</tr>
<tr>
<td>rs9458005</td>
<td>0.22</td>
<td>0.31</td>
<td>0.01</td>
</tr>
<tr>
<td>rs1950562</td>
<td>0.42</td>
<td>0.51</td>
<td>0.03</td>
</tr>
<tr>
<td>rs9458011</td>
<td>0.06</td>
<td>0.07</td>
<td>0.62</td>
</tr>
<tr>
<td>rs1853017</td>
<td>0.30</td>
<td>0.27</td>
<td>0.50</td>
</tr>
<tr>
<td>rs2295368</td>
<td>0.40</td>
<td>0.41</td>
<td>0.94</td>
</tr>
<tr>
<td>rs9456577</td>
<td>0.03</td>
<td>0.01</td>
<td>0.32</td>
</tr>
<tr>
<td>rs4252129</td>
<td>0.01</td>
<td>0.02</td>
<td>0.35</td>
</tr>
<tr>
<td>rs813641</td>
<td>0.16</td>
<td>0.24</td>
<td>0.0046</td>
</tr>
<tr>
<td>rs1317026</td>
<td>0.02</td>
<td>0.02</td>
<td>1.00</td>
</tr>
<tr>
<td>rs2859879</td>
<td>0.34</td>
<td>0.48</td>
<td>0.00018</td>
</tr>
<tr>
<td>rs4252170</td>
<td>0.07</td>
<td>0.13</td>
<td>0.0090</td>
</tr>
<tr>
<td>rs783182</td>
<td>0.49</td>
<td>0.53</td>
<td>0.32</td>
</tr>
<tr>
<td>rs783176</td>
<td>0.17</td>
<td>0.25</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Bold indicates 6 SNPs which show significantly different minor allele frequencies in the sub-sample of 90 individuals typed for the KIV2 repeat compared with the complete sample at p < 0.05 (FDR = 0.14).