Background and Purpose—Previous studies examining genetic associations with MRI-defined brain infarct have yielded inconsistent findings. We investigated genetic variation underlying covert MRI infarct in persons without histories of transient ischemic attack or stroke. We performed meta-analysis of genome-wide association studies of white participants in 6 studies comprising the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium.

Methods—Using 2.2 million genotyped and imputed single nucleotide polymorphisms, each study performed cross-sectional genome-wide association analysis of MRI infarct using age- and sex-adjusted logistic regression models.
Study-specific findings were combined in an inverse-variance-weighted meta-analysis, including 9401 participants with mean age 69.7 (19.4% of whom had ≥1 MRI infarct).

**Results**—The most significant association was found with rs2208454 (minor allele frequency, 20%), located in intron 3 of **MACRO domain containing 2** gene and in the downstream region of **fibronectin leucine-rich transmembrane protein 3** gene. Each copy of the minor allele was associated with lower risk of MRI infarcts (odds ratio, 0.76; 95% confidence interval, 0.68–0.84; \( P=4.64 \times 10^{-7} \)). Highly suggestive associations (\( P<1.0 \times 10^{-5} \)) were also found for 22 other single nucleotide polymorphisms in linkage disequilibrium (\( r^2 > 0.64 \)) with rs2208454. The association with rs2208454 did not replicate in independent samples of 1822 white and 644 black participants, although 4 single nucleotide polymorphisms within 200 kb from rs2208454 were associated with MRI infarcts in the black population sample.

**Conclusions**—This first community-based, genome-wide association study on covert MRI infarcts uncovered novel associations. Although replication of the association with top single nucleotide polymorphisms failed, possibly because of insufficient power, results in the black population sample are encouraging, and further efforts at replication are needed. *(Stroke. 2010;41:210-217.)*

**Key Words:** brain infarction ■ cohort study ■ genome-wide association study ■ meta-analysis

**V**ascular disease of the brain is a leading cause of long-term disability and death. The burden of brain vascular disease is far greater than suggested by occurrence of acute neurological events such as stroke.1 Brain imaging techniques, especially MRI, have revealed that brain infarcts are common in the elderly, especially small subcortical infarcts.2 Although the majority of these MRI infarcts do not produce acute clinical symptoms leading to a diagnosis of stroke, they cannot be considered benign, silent, or asymptomatic, because they are associated with an increased risk for cognitive deficits, motor impairments, and future stroke.3 The pathogenesis of these covert brain infarcts remains poorly understood.

Whereas several monogenic disorders are known to cause brain infarcts, the genes underlying brain infarcts in the general population remain undetermined.3 A genetic component is suggested by increased risk of covert MRI infarcts among individuals whose parents or siblings have experienced clinically overt infarcts.4,5 Previous candidate gene studies of covert MRI infarcts have yielded inconsistent findings. Genome-wide association studies (GWAS) of MRI infarcts are lacking and would permit an unbiased search for genetic variants associated with this phenotype, without relying on a priori hypotheses about underlying pathophysiology.6 To study genetics of these infarcts, we adapted an analytic approach used in a previous study7 of stroke from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium and combined GWAS from 6 prospective population-based cohort studies: the Aging Gene-Environment Susceptibility-Reykjavik Study (AGES-Reykjavik); the Atherosclerosis Risk in Communities (ARIC) Study; the Austrian Stroke Prevention Study (ASPS); the Cardiovascular Health Study (CHS); the Framingham Heart Study (FHS); and the Rotterdam Study. We present results from this meta-analysis that included 9401 stroke-free white participants.

**Materials and Methods**

**Consortium**
The CHARGE consortium includes large prospective community-based cohort studies having genome-wide variation data coupled with extensive data on multiple phenotypes.8 All participating studies agreed on phenotype harmonization, covariate selection, prespecified analytic plans for within-study analyses, and meta-analysis of results. Each study secured approval from Institutional Review Boards, and all participants provided written informed consent for study participation, MRI scanning, and use of DNA for genetic research.

**Setting**
Details of cohort selection, risk factor assessment, and outcome determination in the 6 studies have been reported previously (supplemental Appendix, section 1, available online at http://stroke.ahajournals.org). Briefly, the AGES-Reykjavik Study is a single-center, prospective continuation of the Reykjavik Study, which included persons born 1907 to 1935 and living in Reykjavik, Iceland, in 1967, when the study was started. In 2002 to 2006, 5764 participants from the cohort were reexamined as a part of the AGES-Reykjavik Study.9 The ARIC study enrolled adults aged 45 and 64 years, from 4 US communities (N=15 792, including 11 478 whites). The baseline examination was in 1987 to 1989.10 The ASPS enrolled 2007 inhabitants of Graz, Austria, who lacked neuropsychiatric disease. Between 1991 and 1994 and 1999 and 2003, an extended diagnostic work-up including neuroimaging was performed in a subset of participants aged 45 to 85 years.11,12 The CHS enrolled adults who were 65 years or older and from 4 US communities (N=5888, including 4925 whites). The baseline examination was either in 1989 to 1990 or 1992 to 1993.13 The FHS is a US-based, single-site study that comprises 3 generations of participants. Members of the original cohort followed-up since 1948 (N=5209),14,15 and the members of the offspring cohort followed-up since 1971 (N=5124),16 were invited to undergo an initial brain MRI in 1999 to 2005. The Rotterdam Study enrolled inhabitants from a district of Rotterdam (Ommoord), The Netherlands, aged 55 years or older (N=7983), at the baseline examination in 1990 to 1993 (Rotterdam Study I).17 In 2001, the Rotterdam Study cohort was expanded by 3011 newly eligible persons (Rotterdam Study II).

**MRI Scans**
In each study, eligible participants were invited to undergo MRI scans, which were performed and interpreted in a standardized fashion without knowledge of demographic or clinical information (supplemental Appendix, section 2). Infarct on MRI scan was defined as an area of abnormal signal intensity in a vascular distribution that lacked mass effect. Infarcts had to be ≥3 to 4 mm. Efforts were made in all studies to distinguish infarcts from dilated perivascular spaces. All participants were categorized as having or not having at least 1 MRI infarct.

**Genotyping**
The consortium was formed after individual studies had finalized their GWAS platforms, which differed across studies. All studies
Table 1. Characteristics of Study Participants in Analysis of Covert MRI-Defined Brain Infarcts

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>AGES-Reykjavik</th>
<th>ARIC</th>
<th>ASPS</th>
<th>CHS</th>
<th>FHS</th>
<th>Rotterdam Study I</th>
<th>Rotterdam Study II</th>
</tr>
</thead>
<tbody>
<tr>
<td>N with MRI and genotyping</td>
<td>2866</td>
<td>751</td>
<td>787</td>
<td>2122</td>
<td>2291</td>
<td>481</td>
<td>591</td>
</tr>
<tr>
<td>Excluded for TIA or stroke</td>
<td>310</td>
<td>21</td>
<td>0†</td>
<td>0†</td>
<td>87</td>
<td>46</td>
<td>24</td>
</tr>
<tr>
<td>N in these analyses</td>
<td>2556</td>
<td>730</td>
<td>787</td>
<td>2122</td>
<td>2204</td>
<td>435</td>
<td>567</td>
</tr>
<tr>
<td>N (%) with MRI infarcts</td>
<td>714 (27.9%)</td>
<td>65 (8.9%)</td>
<td>88 (11.2%)</td>
<td>555(26.2%)</td>
<td>252 (11.4%)</td>
<td>89 (20.4%)</td>
<td>59 (10.4%)</td>
</tr>
<tr>
<td>N (%) with lacunar infarcts</td>
<td>NA</td>
<td>58 (8.0%)</td>
<td>73 (9.3%)</td>
<td>482 (22.7%)</td>
<td>211 (9.6%)</td>
<td>82 (18.9%)</td>
<td>48 (8.5%)</td>
</tr>
<tr>
<td>Mean age (± SD) at MRI</td>
<td>76.2 (5.4)</td>
<td>63.2 (± 4.4)</td>
<td>65.3 (± 8.0)</td>
<td>71.7 (± 4.8)</td>
<td>63.9 (± 11.3)</td>
<td>72.9 (± 7.9)</td>
<td>67.2 (± 5.3)</td>
</tr>
<tr>
<td>Women, %</td>
<td>1510 (59.1%)</td>
<td>433 (59.3%)</td>
<td>449 (57.1%)</td>
<td>1304 (61.5%)</td>
<td>1195 (54.2%)</td>
<td>224 (51.5%)</td>
<td>284 (60.1%)</td>
</tr>
<tr>
<td>Dementia at MRI, %</td>
<td>107 (4.2%)</td>
<td>0</td>
<td>0†</td>
<td>77 (3.6%)</td>
<td>6 (0.3%)</td>
<td>0†</td>
<td>0†</td>
</tr>
</tbody>
</table>

Cardiovascular risk factor at MRI*

<table>
<thead>
<tr>
<th></th>
<th>Systolic BP (mean±SD)</th>
<th>Hypertension (mean±SD)</th>
<th>Diabetes mellitus, %</th>
<th>Current smoker, %</th>
<th>Prevalent CVD at MRI, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>142 (20)</td>
<td>2025 (79.2%)</td>
<td>271 (10.6%)</td>
<td>320 (12.5%)</td>
<td>397 (15.5%)</td>
</tr>
<tr>
<td></td>
<td>119 (±17)</td>
<td>193 (26.7%)</td>
<td>74 (10.2%)</td>
<td>136 (18.7%)</td>
<td>41 (5.8%)</td>
</tr>
<tr>
<td></td>
<td>143 (±22.6)</td>
<td>524 (66.6%)</td>
<td>75 (9.5)</td>
<td>92 (11.7)</td>
<td>269 (34.2%)</td>
</tr>
<tr>
<td></td>
<td>133.7 (±20.5)</td>
<td>1043 (49.2%)</td>
<td>203 (9.7%)</td>
<td>218 (10.3%)</td>
<td>0†</td>
</tr>
<tr>
<td></td>
<td>127 (±19)</td>
<td>943 (43.6%)</td>
<td>261 (12.2%)</td>
<td>252 (11.7%)</td>
<td>242 (11.1%)</td>
</tr>
<tr>
<td></td>
<td>146 (±20)</td>
<td>310 (71.3%)</td>
<td>18 (4.1%)</td>
<td>81 (18.6%)</td>
<td>31 (7.2%)</td>
</tr>
<tr>
<td></td>
<td>145 (±18)</td>
<td>391 (69%)</td>
<td>51 (9%)</td>
<td>167 (30%)</td>
<td>38 (6.7%)</td>
</tr>
</tbody>
</table>

AGES-Reykjavik indicates Aging Gene-Environment Susceptibility-Reykjavik Study; ARIC, Atherosclerosis Risk in Communities Study; ASPS, Austrian Stroke Prevention Study; BP, blood pressure; CHS, Cardiovascular Health Study; CVD, cardiovascular disease other than transient ischemic attack or stroke, including clinically evident coronary artery disease, congestive heart failure and peripheral vascular disease; FHS, Framingham Heart Study; MRI, magnetic resonance imaging; NA, not available; SD, standard deviation; TIA, transient ischemic attack.

*Definition of baseline characteristics was uniform across all studies. Hypertension was defined using standard criteria as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or being on antihypertensive treatment; diabetes mellitus was defined as a casual or 2 hours postprandial blood glucose ≥200 mg/dL (11 mmol/L), a fasting blood glucose ≥126 mg/dL (7 mmol/L), or use of insulin or oral hypoglycemic agents; CVD (cardiovascular disease) was defined as presence of congestive heart failure, coronary heart disease, or intermittent claudication.

†In ASPS and CHS, participants with prevalent TIA or stroke were not included. In CHS, participants with other prevalent CVD were also not included.

‡In ASPS and Rotterdam studies, participants with dementia did not undergo cranial MRI.

Table 1. Characteristics of Study Participants in Analysis of Covert MRI-Defined Brain Infarcts

Study Population

Participants were eligible for these analyses if they had genotyping, an MRI, and lacked a history of transient ischemic attack or stroke before their MRI (covert MRI infarcts). Participants were either entirely or almost entirely European whites in the AGES-Reykjavik Study, ASPS, FHS, and Rotterdam Study, so black participants from ARIC and CHS were not included in these analyses. In addition, CHS did not genotype participants with any form of clinical cardiovascular disease at baseline. By design, ASPS did not perform MRI scans in patients with transient ischemic attack or stroke. Also, ASPS and Rotterdam Study did not perform MRI scans in participants with dementia. The number and characteristics of participants from each cohort are shown in Table 1.

Statistical Analyses Within Studies

Each study fit an additive genetic model with a 1-degree-of-freedom trend test relating genotype dosage, 0 to 2 copies of the minor allele, to having or not having at least 1 MRI infarct. We used logistic regression models to calculate odds ratios (OR) with corresponding 95% confidence intervals (CI). Initial analyses were adjusted only for age and sex to avoid adjusting for covariates that might lie along a causal pathway. In addition, ARIC and CHS also adjusted for study site, and FHS adjusted for familial structure. To explore potential mechanisms, we additionally adjusted our most significant association in one model for systolic blood pressure and in another model for the presence or absence of hypertension, defined as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or use of antihypertensive medications. All studies screened for latent population substructure, which was negligible (supplemental Appendix, section 4).

Meta-Analysis

We conducted a fixed-effects meta-analysis of results from the 6 studies, with 7 cohorts counting the Rotterdam Study II, using inverse-variance weighting. After quality control, filtering, and imputation within each study, we restricted our meta-analysis to 2,217,889 autosomal SNP that were common to all studies and had an average minor allele frequency >2%. Details on the meta-analysis strategy and functional annotation of SNP are available (supplemental Appendix, section 5). As suggested by others, we decided a priori on a genome-wide significant threshold of 5×10⁻⁸. SNP with 5×10⁻⁸<P<1×10⁻⁶ were considered highly suggestive associations. As available in meta-analysis, we also examined associa-
tions with candidate SNP, or their proxies, previously reported to be significantly associated with covert MRI infarcts.

Replication

We attempted to replicate findings for our top SNP by genotyping it in 1822 elderly white participants from the 3C-Dijon study and 644 black participants from the ARIC study. We also explored 59 SNP within 300 kb of our top SNP using in-silico replication in the black sample (supplemental Appendix, section 7). We set the threshold for replication at 1-sided \( P \leq 0.05 \).

Results

Among 9401 participants whose mean age was 69.7 years and 53.4% women, 1822 (19.4%) had at least 1 MRI infarct (Table 1). After meta-analysis the genomic inflation factor \( \lambda \) was 0.996, indicating no significant inflation of probability values. Figure 1 shows the genome-wide plot of probability values for individual SNP against their genomic position. None of the peaks cleared the threshold for genome-wide significance, but 51 SNP had highly suggestive associations, with \( P < 1 \times 10^{-5} \) (Table 2, supplemental Appendix, section 6, Table 1). There was no significant heterogeneity across studies for the association with these SNP (supplemental Appendix, section 6).

The most significant association was found on chromosome 20p12 with SNP rs2208454, located in intron 3 of MACRO domain containing 2 (MACROD2) gene and in the downstream region of fibronectin leucine-rich transmembrane protein 3 (FLRT3) gene. The OR for MRI infarcts was 0.76 (95% CI, 0.68–0.84; \( P = 4.64 \times 10^{-7} \)). Additional adjustment for systolic blood pressure (OR, 0.76; 95% CI, 0.68–0.85) or hypertension (OR, 0.76; 95% CI, 0.68–0.84) did not change results. Figure 2 shows a forest plot of risk estimates for rs2208454 across the 7 cohorts. Twenty-two other SNP in intron 3 of MACROD2 were also associated with MRI infarcts, with \( P < 1 \times 10^{-5} \) (Table 2, supplemental Appendix, Table 1). All were in linkage disequilibrium with rs2208454: \( r^2 > 0.64 \) for all and \( r^2 > 0.8 \) for 17 of the SNP. Of these 22 SNP, 1 was intronic within FLRT3 (rs6110247) and 3 were potential transcription factor binding sites (rs6110247, rs743216, and rs3789335). Figure 3 shows all SNP within a 200-kb region on either side of the top hit, together with probability values, recombination rates, and known genes in that region.

Although estimated quality of imputation for rs2208454 was excellent in most studies (observed/expected [O/E] ratio, >0.95), it was poor in CHS (O/E ratio, 0.48). Therefore, rs2208454 was genotyped in CHS. When incorporating this result in the meta-analysis instead of the imputed data, the OR was the same at 0.76, but the probability value was smaller at 1.44\( \times 10^{-7} \).

We failed to replicate findings for the top SNP either in 1822 white participants from 3C-Dijon for whom mean age was 72.5 years and 9.4% had at least 1 MRI infarct (minor allele frequency, 22%; OR, 0.99; 95% CI, 0.76–1.29; \( P = 0.92 \)) or in 644 black participants from the ARIC study, for whom mean age was 61.5 years and 15.5% had at least 1 MRI infarct (minor allele frequency, 6.5%; OR, 1.26; 95% CI, 0.74–2.15; \( P = 0.40 \)). However, 4 other SNP in intron 3 of MACROD2 within 50 to 151kb from rs2208454 were significantly associated with MRI infarcts in the black sample (rs7268327, \( P = 0.045 \); rs1998237, \( P = 0.006 \); rs4464346,
Table 2. Strongest SNP Phenotype Associations in Meta-Analysis for Covert MRI-Defined Brain Infarcts

<table>
<thead>
<tr>
<th>SNP rs</th>
<th>SNP Function</th>
<th>Chr Position</th>
<th>Minor Allele</th>
<th>MAF</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>PAR</th>
<th>Closest Gene</th>
<th>Distance</th>
<th>Second Closest Gene</th>
<th>Distance</th>
<th>Additional SNP at P&lt;10^-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2208454</td>
<td>Intrinsic</td>
<td>20:14213415</td>
<td>T</td>
<td>0.20</td>
<td>0.76 (0.68–0.84)</td>
<td>4.64E–07</td>
<td>0.12</td>
<td>MACROD2</td>
<td>In gene</td>
<td>FLRT3</td>
<td>39.2</td>
<td>22 intronic</td>
</tr>
<tr>
<td>rs1834018</td>
<td>Intrinsic</td>
<td>16:56864743</td>
<td>G</td>
<td>0.12</td>
<td>1.32 (1.18–1.48)</td>
<td>1.59E–06</td>
<td>0.07</td>
<td>CDCC113</td>
<td>In gene</td>
<td>KLKBL4</td>
<td>6.7</td>
<td>1 intronic, 2 upstream</td>
</tr>
<tr>
<td>rs2869036</td>
<td>Intergenic</td>
<td>15:78454627</td>
<td>G</td>
<td>0.22</td>
<td>0.76 (0.68–0.85)</td>
<td>3.36E–06</td>
<td>0.13</td>
<td>TGFB1</td>
<td>27.0</td>
<td>IREB2</td>
<td>62.9</td>
<td>0</td>
</tr>
<tr>
<td>rs1471895</td>
<td>Intrinsic</td>
<td>11:11297762</td>
<td>A</td>
<td>0.06</td>
<td>1.44 (1.23–1.68)</td>
<td>3.90E–06</td>
<td>0.05</td>
<td>GALNT4</td>
<td>In gene</td>
<td>EIF4G2</td>
<td>510.6</td>
<td>1 intronic</td>
</tr>
<tr>
<td>rs4335430</td>
<td>Intrinsic</td>
<td>1:215166854</td>
<td>T</td>
<td>0.16</td>
<td>1.27 (1.15–1.41)</td>
<td>4.31E–06</td>
<td>0.08</td>
<td>ESRRG</td>
<td>In gene</td>
<td>USH2A</td>
<td>503.5</td>
<td>0</td>
</tr>
<tr>
<td>rs17695069</td>
<td>Intrinsic</td>
<td>18:54801916</td>
<td>G</td>
<td>0.11</td>
<td>1.34 (1.16–1.52)</td>
<td>4.54E–06</td>
<td>0.07</td>
<td>ZNF532</td>
<td>In gene</td>
<td>SEC11C</td>
<td>156.2</td>
<td>0</td>
</tr>
<tr>
<td>rs2284038</td>
<td>Intrinsic</td>
<td>22:35965001</td>
<td>G</td>
<td>0.36</td>
<td>1.20 (1.11–1.30)</td>
<td>5.98E–06</td>
<td>0.13</td>
<td>RAC2</td>
<td>In gene</td>
<td>SST3</td>
<td>26.7</td>
<td>3 intronic</td>
</tr>
<tr>
<td>rs1285474</td>
<td>Intrinsic</td>
<td>14:24451217</td>
<td>A</td>
<td>0.12</td>
<td>0.75 (0.66–0.85)</td>
<td>6.90E–06</td>
<td>0.08</td>
<td>STXB6</td>
<td>In gene</td>
<td>GZMB</td>
<td>277.9</td>
<td>13 intronic</td>
</tr>
<tr>
<td>rs1174692</td>
<td>Intrinsic</td>
<td>5:149114067</td>
<td>A</td>
<td>0.27</td>
<td>0.81 (0.73–0.89)</td>
<td>8.35E–06</td>
<td>0.12</td>
<td>PPARC16B</td>
<td>In gene</td>
<td>PDE6A</td>
<td>103.6</td>
<td>0</td>
</tr>
</tbody>
</table>

The reference SNP number (rs), function, and chromosome (chr) position are listed. OR, 95% CI, and P as powers of 10 (E) are based on the meta-analysis. Each row lists only the SNP phenotype association with the lowest P for that locus. The last column shows the number of additional SNP at the same locus, within 250 kb of the specified SNP, that were also associated with the phenotype with a P<10^-5. Complete details for these additional SNP are provided online in Table 1. Alleles were identified based on the plus strand of the NCBI build 36. The minor allele was also the coded allele, and minor allele frequency (MAF) is based on allele frequency distribution pattern in this region differs substantially between white and black populations (Figure 4). Details on the replication effort are contained in the supplemental Appendix, section 7.

We repeated these analyses including, rather than excluding, participants with a history of a transient ischemic attack or stroke (Table 1). The results were similar, although the associations were slightly weaker in general (data not shown). We also examined associations with previously reported candidate SNP, or their proxies, none of which was significant after correction for multiple testing (supplemental Appendix, section 9). In addition, we examined associations with top SNP in the Ninjurin-2 gene associated with ischemic stroke in a recent CHARGE GWAS meta-analysis7 and found none significantly associated with MRI infarcts. Finally, we explored relations between ischemic stroke and the 23 SNP in MACROD2 (Table 2) in the recent CHARGE GWAS meta-analysis of ischemic stroke.2 No significant association of these SNP with ischemic stroke was identified.

**Discussion**

This meta-analysis of GWAS data on covert MRI-defined brain infarcts included 9401 participants without a history of transient ischemic attack or stroke from 6 community-based studies. The most significant association (P=4.64x10^{-7}) was found for SNP rs2208454 on chromosome 20p12, located in intron 3 of MACROD2 and in the downstream region of FLRT3. The less common allele was associated with a lower risk. Twenty-two SNP in linkage disequilibrium with rs2208454 were also associated with MRI infarcts, with P<1.0x10^{-5}, as were 28 other SNP in 8 different loci. No association reached our preset threshold for genome-wide significance of 5.0x10^{-8}. In 2 replication samples of 1822 white participants from 3C-Dijon and 644 black participants from ARIC, we did not observe an association with rs2208454, although 4 SNP within 200 kb from rs2208454 were associated with MRI infarcts in the black sample. Finally, we did not observe an association with SNP previously reported to be associated significantly with covert MRI infarcts in candidate genes studies.
The function of the protein encoded by MACROD2 is poorly understood. It contains a macro domain that is evolutionarily conserved and expressed in fetal and adult human brain.\textsuperscript{20,21} Macro domains bind ADP-ribose, suggesting a role in ADP-ribosylation, a post-translational modification involved in many processes including DNA repair, transcriptional activation, and repression, and telomere and chromatin biology.\textsuperscript{22} Nested in intron 3 of MACROD2, FLRT3 encodes

\begin{figure}[!h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Regional plot for associations in region centered on top hit. All SNP (triangles) are plotted with their meta-analysis probability values against their genomic position. The color of the triangles represents the linkage disequilibrium between SNP: purple, $r^2 \leq 0.05$; light blue, $0.05 < r^2 \leq 0.10$; green, $0.10 < r^2 \leq 0.30$; yellow, $0.30 < r^2 \leq 0.60$; orange, $0.60 < r^2 \leq 0.80$; red, $r^2 > 0.80$. Light blue line represents estimated recombination rates. Genes are shown as dark green arrows.}
\end{figure}

\begin{figure}[!h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Linkage disequilibrium (LD) plot of region in MACROD2 including top hit (rs2208454), 22 other SNP with $P < 10^{-5}$ in meta-analysis, and 4 SNP with $P < 0.05$ in black sample, using HapMap release 22. Plot on top depicts LD in European population (CEU), whereas plot on bottom depicts LD in black population (YRI). The color scheme is white for $D^* = 1$ and logarithm of odds (LOD) $< 2$, blue for $D^* = 1$ and LOD $\leq 2$, shades of pink and red for $D^* < 1$ and LOD $\geq 2$, and bright red for $D^* = 1$ and LOD $> 2$. The SNP marked in red is rs2208454, the 4 SNP marked in blue are SNP with $P < 0.05$ in the black sample; from left to right: rs7268327, rs1998237, rs4643436, and rs8116105. LD could not be measured for 3 SNP in the YRI population because of a minor allele frequency $< 0.001$ (rs6135125, rs6110247, rs12624446). Green lines represent genes in region.}
\end{figure}
fibronectin leucine-rich transmembrane protein 3.23 The gene is expressed in various tissues, including brain, and is well-conserved across species.24 The protein it encodes modulates homotypic cell adhesion and promotes fibroblast growth factor signaling,24 which is potentially involved in angiogenesis and neurogenesis.25 In animal experiments, FLRT3 was shown to promote neurite outgrowth after axonal injury.26,27

Intriguingly, even though both FLRT3 and Ninjurin-2 appear to modulate response to neuronal injury, the Ninjurin-2 SNP identified in the recently published ischemic stroke GWAS within the CHARGE consortium were not significantly associated with covert MRI infarcts, and the MACROD2 SNP identified through the present analysis were not associated with overt ischemic stroke. A possible explanation for this discrepancy could be that MRI infarcts comprise mainly small subcortical infarcts, whereas ischemic stroke represents a much more heterogeneous entity. Further investigations are needed, beginning with replication of the associations in external cohorts of MRI-defined infarcts, ischemic stroke, and subtypes of ischemic stroke.

None of the 23 SNP in MACROD2 with \( P<10^{-5} \) was significantly associated with gene expression in publicly available genome-wide expression quantitative trait loci data sets (supplemental Appendix, section 6), but this finding should be interpreted cautiously because less than half of these SNP were present on any of the genotyping arrays used in these studies and expression quantitative trait loci may be tissue and insult specific.

This meta-analysis has strengths. It included 6 large cohort studies with similar MRI protocols. The analyses were restricted to white participants to minimize the risk of population stratification. Genotyping was subjected to rigorous quality control. Our study also has limitations. Despite having close to 10 000 participants, of whom almost 2000 had covert MRI infarcts, we had limited power to detect associations with small effect sizes and associations with rare variants. An important caution is that the association with the top SNP was not directly replicated in 2 independent samples. Hence, we cannot exclude the possibility that this association was a chance finding. However, the power to detect an association with rs2208454 in the 2 replication samples was relatively low (61% for the white participants from 3C-Dijon and 18% for the black participants from ARIC), assuming the relatively low (61% for the white participants from 3C-Dijon association with rs2208454 in the 2 replication samples was hence, we cannot exclude the possibility that this association is population-based, the samples are not perfectly representative of the total cohort, because they include subsets of individuals who were able to undergo brain MRI and agreed to do so.

Conclusion

This meta-analysis of GWAS shows highly suggestive association of covert MRI infarcts with rs2208454 on chromosome 20p12 (\( P=4.64 \times 10^{-7} \)). Attempted replication of the top SNP in an independent white sample failed, and additional attempts in larger samples to replicate this finding and associations with other suggestive loci are needed. Extending replication efforts not only in white but also in black populations may prove useful for fine mapping purposes. The molecular, clinical, and epidemiological correlates of confirmed associations may permit new insights into the pathophysiology and prevention of covert brain infarcts.

Sources of Funding

Aging Gene-Environment Susceptibility-Reykjavik Study: The research has been funded by NIA contract N01-AG-12100 with contributions from NEI, NIDCD, and NHLBI, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

The Atherosclerosis Risk in Communities Study: The research is performed as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, and grants RO1HL087641 and ROI1HL093029; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. Infrastructure was partly supported by grant number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

The Austrian Stroke Prevention Study: The research reported in this article was funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180. The Medical University of Graz supports the databank of the ASPS.

Cardiovascular Health Study: The research reported in this article was supported by contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01-HC-15103, N01-HC-55222, N01-HC-75150, and N01-HC-45133, and grant numbers U01 HL080295 and R01 HL087652 from the National Heart, Lung, and Blood Institute, with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01RR00425 to the Cedars-Sinai General Clinical Research Center Genotyping core and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center.
Framingham Heart Study: From the Boston University School of Medicine and the Framingham Heart Study. This work was supported by the National Heart, Lung, and Blood Institute’s Framingham Heart Study (contract number N01-HC-25195) and its contract with Affymetrix, Inc, for genotyping services (contract number N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinCGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. Analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARE) project. This study was also supported by grants from the National Institute of Neurological Disorders and Stroke (NS17950) and the National Institute of Aging (AG08122, AG16495, and AG0033193).

Rotterdam Study: The generation and management of genome-wide association study genotype data for the Rotterdam Study is supported by the Netherlands Organization of Scientific Research Investments (no. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGL)/Netherlands Organization for Scientific Research (project no. 050-060-810), and NWO grants (918-46-615, 948-00-010). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands, Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture, and Science, the Ministry for Health, Welfare, and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

3C-Dijon Study: The Three-City (3C) Study is conducted under a partnership agreement between the Institut National de la Santé et de la Recherche Médicale (INSERM), the Victor Segalen–Bordeaux II University, and Sanofi-Aventis. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study is also supported by the Caisse Nationale Maladie des Travailleurs Salariés, Direction Générale de la Santé, Mutuelle Générale de l’Education Nationale (MGEN), Institut de la Longévité, Conseils Régionaux of Aquitaine and Bourgogne, Fondation de France, and Ministry of Research–INSERM Programme “Cohortes et collections de données biologiques.”

Disclosures

None.

References

Genome-Wide Association Studies of MRI-Defined Brain Infarcts: Meta-Analysis From the CHARGE Consortium


Stroke. 2010;41:210-217; originally published online December 31, 2009;
doi: 10.1161/STROKEAHA.109.569194

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 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/41/2/210

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2016/03/31/STROKEAHA.109.569194.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/
磁共振诊断的脑梗死的全基因组关联研究
来自 CHARGE 协作组的荟萃分析

Genome-Wide Association Studies of MRI-Defined Brain Infarcts
Meta-Analysis From the CHARGE Consortium

Stéphanie Debette, MD, PhD; Joshua C. Bis, PhD; Myriam Fornage, PhD; Helena Schmidt, MD, PhD; M. Arfan Ikram, MD, PhD; Sigurdur Sigurdsson, MSc; Gerardo Heiss, MD; Maksim Struchalin, MSc; Albert V. Smith, PhD; Aad van der Lugt, MD, PhD; Charles DeCarli, MD; Thomas Lumley, PhD; David S. Knopman, MD; Christian Enzinger, MD; Gudny Eiriksdottir, MSc; Peter J. Koudstaal, MD, PhD; Anita L. DeStefano, PhD; Bruce M. Psaty, MD, PhD; Carole Dubois, PhD; Diane J. Catellier, DrPH; Franz Fazekas, MD; Thor Aspelund, PhD; Yuri S. Aulchenko, PhD; Alexa Beiser, PhD; Jerome I. Rotter, MD; Christophe Tzourio, MD, PhD; Dean K. Shibata, MD; Maria Tscherne, MD; Tamara B. Harris, MD, MS; Fernando Rivadeneira, MD, PhD; Larry D. Atwood, PhD; Kenneth Rice, PhD; Rebecca F. Gottesman, MD, PhD; Mark A. van Buchem, MD, PhD; Andre G. Uitterlinden, PhD; Margaret Kelly-Hayes, RN, DEd; Mary Cushman, MD, MSc; Yicheng Zhu, MD; Eric Boerwinkle, PhD; Vilmundur Gudnason, MD, PhD; Albert Hofman, MD, PhD; Jose R. Romero, MD; Oscar Lopez, MD; Cornelia M. van Duijn, PhD; Rhoda Au, PhD; Susan R. Heckbert, MD, PhD; Philip A. Wolf, MD; Thomas H. Mosley, PhD; Sudha Seshadri, MD; Monique M. B. Breteler, MD, PhD; Reinhold Schmidt, MD; Lenore J. Lauener, PhD; W.T. Longstreth, Jr, MD, MPH

背景和目的: 先前关于磁共振诊断的脑梗死的基因相关性研究得出的结果不一致。本项研究旨在探讨经MRI诊断的既往无短暂性脑缺血发作或卒中病史的隐匿性脑梗死患者基因的差异。我们对心脏和衰老基因流行病学队列研究(Cohorts for Heart and Aging Research in Genomic Epidemiology, CHARGE)协作组的6项涉及白人群的全基因组关联研究进行了荟萃分析。

方法: 在各项研究中,应用220万个基因型和插补的单核苷酸多态性及年龄、性别校正的logistic回归模型,针对MRI诊断的脑梗死进行横断面的全基因组关联分析。结合各研究的特殊发现,进行反变量加权荟萃分析,共纳入9401个参与者,平均年龄69.7岁(其中19.4%存在≥1个MRI诊断的梗死灶)。

结果: 相关性最强的是rs2208454(较小等位基因频率为20%),位于MACROD2 (MACRO domain containing 2)基因的内含子3中,在FLRT3 (富含亮氨酸的跨膜蛋白3)基因的下游。每个较小等位基因的复制与MRI诊断的脑梗死的风险降低有关(OR, 0.76; 95% CI 0.68–0.84; P=4.64×10^{-7})。另外,高度相关的还有22个与rs2208454连锁不平衡(r^2>0.64)的单核苷酸多态性(SNP) (P<1.0×10^{-5})。虽然在黑人样本中,来自rs2208454的4个200 kb以内的SNP与MRI诊断的脑梗死存在相关性,但与rs2208454的关联性并未在两个独立样本(1822例白人样本和644例黑人样本)中得到验证。

结论: 这个首次基于社区的关于MRI诊断的隐匿性脑梗死的全基因组关联研究得出了新的相关性。受到研究效能的限制,虽然未能成功复制常见SNP的关联性,但在黑人样本中得出了鼓舞人心的结果,此结果仍需进一步验证。

关键词: 脑梗死,队列研究,全基因组关联研究,荟萃分析

(Stroke. 2010;41:210-217. 濮月华 译 刘丽萍 校)
也不能被认为是良性的、静止的或非症状性的，因为它们与认知缺陷、运动障碍及将来的卒中风险增高相关[2]。这些隐匿性脑梗死的发病机制仍不清楚。

尽管已知一些单基因疾病可导致脑梗死，但在广大人群中潜藏的引起脑梗死的基因尚未明确[3]。对于父母或兄弟姐妹已患临床上明确诊断的脑梗死的个体，某个基因的组成部分可增加其患MRI诊断的隐匿性脑梗死的风险[4, 5]。以前关于MRI诊断的隐匿性脑梗死的候选基因研究得出了不一致的结果。目前仍缺少针对MRI诊断的隐匿性脑梗死的全基因组关联研究(Genome-wide association studies, GWAS)，GWAS能够帮助人们在不依赖潜在的病理生理学机制的先验假设的情况下，无偏倚的寻找与这种显型相关的基因变异[6]。

为了研究这些梗死的遗传学，我们调整了一个先前卒中研究[7]的分析方法，这一方法源于心脏和衰老基因流行病学队列研究(Cohorts for Heart and Aging Research in Genomic Epidemiology, CHARGE)协作组，并结合来自6项基于人群的前瞻性队列研究的GWAS。这6项研究分别是：衰老基因环境易感性雷克亚维克研究(AGES Reykjavik)、社区动脉粥样硬化风险研究(ARIC)、奥地利卒中预防研究(ASPS)、心血管健康研究(CHS)、弗雷明汉研究(FHS)和鹿特丹(Rotterdam)研究。本项研究对9401例未患卒中的白人受试者进行了荟萃分析。

## 材料和方法

### 协作组

CHARGE协作组包括几个大型前瞻性基于社区的队列研究，这些研究包含全基因组的变异数据和关于多重表型的大量资料[8]。所有入选的研究需在下列方面达成一致，即表型协调、协变量选择、为特定研究预定的分析计划及荟萃分析结果。每项研究都获得伦理委员会的批准，所有受试者都同意MRI扫描及DNA用于遗传学研究，并签署了书面的知情同意书。

### 研究设置

这6项研究的队列筛选的详细内容、危险因素的评估和结局的确定已经做过报道(增刊附录，第二部分)。由于常见SNP被插补到所有队列中，我们对插补的质量较差的研究直接进行基因分型。基因分型、插补、治疗控制的详细方法已发表在增刊附录，第三部分。

### 基因分型

在每项研究，参与者都要进行MRI扫描，扫描的执行和结果的判读以一种不需要人口学特征或临床信息的标准化方式进行(增刊附录，第二部分)。MRI扫描上的梗死灶定义为在血管支配区内具有异常信号强度的区域，不伴有占位效应，梗死灶应≥3-4 mm。在所有的研究中都需要对梗死灶和扩大的血管周围间隙进行鉴别。根据有无MRI诊断的梗死灶将所有参与者分为两组。

### 研究人群

如果参与者具有基因分型、MRI扫描结果，且在完成磁共振检查（MRI诊断的隐匿性脑梗死）前无短暂性脑缺血发作或卒中病史，则被选中进行分
在 AGES Reykjavik 研究、ASPS 研究、FHS 研究和鹿特丹研究中，全部或几乎全部参与者均为欧洲白色人种，因此 ARIC 研究和 CHS 研究中的黑人参与者将不被纳入分析人群中。而且，CHS 研究未对基线时患有临床心血管疾病的参与者进行基因分型。在设计上，ASPS 研究未对 TIA 或卒中患者进行 MRI 扫描。同样，ASPS 研究和鹿特丹研究未对伴有痴呆的患者进行 MRI 扫描。表 1 列出了各项研究的参与者的例数和特点。

统计分析
每项研究都要拟合一个附加的基因模型，基因型剂量相关的趋势检验自由度为 1，较小的等位基因片段复制 0-2 倍，根据 MRI 有无梗死灶分为两组。应用 logistic 回归模型计算比值比 (OR) 及相应的 95% 可信区间 (95% CI)。首先对年龄和性别进行校正，以避免按因果途径对协变量进行调整。而且，在 ARIC 研究和 CHS 研究要对研究场所进行校正，在 FHS 研究要校正家族构成。为了探索潜在的机制，我们还需要在一个模型中校正最重要的关联——收缩压，在另外一个模型中校正是否存在高血压 (收缩压 ≥ 140 mmHg, 舒张压 ≥ 90 mmHg, 或已应用降压药物)。所有研究根据潜在的人口子结构进行筛选，可被忽略 (增刊附录，第四部分)。

荟萃分析
我们对来自 6 个研究的结果进行固定效应荟萃分析，对于 7 个队列 (加上鹿特丹研究 II) 应用反向变异数加权法。首先对每项研究进行质量控制、过滤、插补，然后把荟萃分析限制在所有研究共有 2 217 889 个常染色体的 SNP 中，较小的等位基因平均频率 >2%。荟萃分析 SNP 功能注释的详细内容可在增刊查询 (增刊附录，第五部分)。根据其他研究的建议 [6]，我们将全基因组显著性先验阈值定义为 5×10^-8。当 SNP 的 P 值范围为 5×10^-8 < P < 1×10^-5 时，高度提示存在相关性。本研究还对先前报道的与 MRI 诊断的隐匿性脑梗死有关的候选 SNP 或取代物进行了验证。

我们通过对 3C-Dijon 研究 [18, 19] 的 1822 例老年白人参与者和 ARIC 研究的 644 例黑人参与者进行
基因分型，试图复制本项研究发现的常见 SNP[10]。本研究在黑人样本中应用电脑模拟的复制技术从常见 SNP 中发现了 59 个 SNP（大小在 300 kb 以内）（增刊附录，第七部分）[10]。设定复制的阈值为单侧，$P=0.05$。

结果

在 9401 名参与者中，平均年龄 69.7 岁，女性 53.4%，1822 例（19.4%）存在至少一个磁共振诊断的梗死灶（见表 1）。经过荟萃分析，基因组膨胀因子 $\lambda=0.996$，提示没有概率值的显著性膨胀。图 1 为全基因组信号强度图（Manhattan）显示了 MRI 诊断的脑梗死的基因组位置及对应的概率值。从 p 终端终止，黑色实线为全基因组显著性预置的阈值，$P=5.0 \times 10^{-8}$；蓝色虚线为较宽松的阈值，也在文献中被应用，$P=5.0 \times 10^{-7}$；红色虚线为高度提示关联性的阈值，$P=1.0 \times 10^{-5}$。

图 1 全基因组信号强度图（Manhattan）显示了 MRI 诊断的脑梗死的基因组位置及对应的概率值。在每个染色体上（X 轴），结果从左向右标绘，到 p 终端终止。黑色实线为全基因组显著性预置的阈值，$P=5.0 \times 10^{-8}$；蓝色虚线为较宽松的阈值，也在文献中被应用，$P=5.0 \times 10^{-7}$；红色虚线为高度提示关联性的阈值，$P=1.0 \times 10^{-5}$。

相关性最 强的是染色体 20p12 中的 SNP rs2208454，位于 MACROD2（MACRO domain containing 2）基因的内含子 3 中，在 FLRT3（富含亮氨酸的纤维连接蛋白跨膜蛋白 3）基因的下游。MRI 诊断的脑梗死的 OR 值为 0.76（95% CI, 0.68–0.84; $P=4.64 \times 10^{-7}$）。校正后的收缩压（OR，0.76; 95% CI, 0.68–0.85）或高血压（OR，0.76; 95% CI, 0.68–0.84）对结果没有影响。图 2 显示了在 7 个队列中 rs2208454 的风险评估森林图。另外 22 个在 MACROD2 基因内含子 3 中的 SNP 也与 MRI 诊断的脑梗死存在相关性，$P<1.0 \times 10^{-5}$（表 2，增刊附录，表 I）。所有 SNP 与 rs2208454 存在连锁不平衡：对所有 SNP，$r^2>0.64$；对 17 个 SNP，$r^2>0.8$。对于这 22 个 SNP，1 个是 FLRT3（rs6110247）基因中的内含子，3 个是潜在的转录因子结合位点（rs6110247, rs743216 和 rs3789335）。图 3 显示了高发基因两侧 200 kb 的区域内的所有 SNP，及其概率值、重组率，还有这些区域的已知基因。

虽然在多数研究中，对 rs2208454 基因插补的质量较好（O/E[观察值 / 预期值] 比值 >0.95），但在 CHS 研究中质量较差（O/E 比值为 0.48）。因此，在 CHS 研究中，对 rs2208454 进行了基因分型。当荟萃分析中合并这些结果，不包括插补数据，OR 值同样是 0.76，但概率值变小为 $1.44 \times 10^{-7}$。

无论是在来自 3C-Dijon 研究的 1822 例白人选入者（平均年龄 72.5 岁，9.4% 存在至少一个 MRI 诊断的梗死灶，较小的等位基因频率为 22%,
Stroke  February 2010

表 2 与 MRI 诊断的隐匿性脑梗死相关性最强的 SNP 表型

<table>
<thead>
<tr>
<th>SNP</th>
<th>OR</th>
<th>95% CI</th>
<th>P值</th>
<th>功能</th>
<th>染色体位点</th>
<th>位基因</th>
<th>MAF</th>
<th>(95% CI)</th>
<th>P值</th>
<th>PAR</th>
<th>名称</th>
<th>距离</th>
<th>名称</th>
<th>距离</th>
<th>额外的 SNP，P&lt;10^-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2208454</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td>0.09</td>
<td>1.01</td>
<td>0.04</td>
<td>MACROD2</td>
<td>20:14213415</td>
<td>FLRT3</td>
<td>39.2</td>
<td>22 个内含子</td>
</tr>
<tr>
<td>rs1834018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td>0.06</td>
<td>0.76</td>
<td>0.84</td>
<td>MACROD2</td>
<td>16:56864743</td>
<td>CCDC113</td>
<td>1.32</td>
<td>18 个内含子</td>
</tr>
</tbody>
</table>
| rs2869036 |     |        |     |      |        |        | G   | 0.02   | 0.76 | 0.85 | MACROD2 | 15:76454627 | CRABP1    | 27.0 | 1 个内含子，2 个上游
| rs1471895 |     |        |     |      |        |        | G   | 0.06   | 1.44 | 1.23 | MACROD2 | 11:1297762 | GALNTL4   | 3.90e-06 | 1 个内含子    |
| rs4335430 |     |        |     |      |        |        | T   | 0.04   | 1.27 | 1.00 | MACROD2 | 1:215166854 | ESRRG     | 4.31e-06 | 1 个内含子    |
| rs17695069|     |        |     |      |        |        | G   | 0.11   | 1.34 | 1.52 | MACROD2 | 18:54801916 | ZNF532    | 4.54e-06 | 1 个内含子    |
| rs2284038 |     |        |     |      |        |        | G   | 0.36   | 1.20 | 1.11 | MACROD2 | 22:35965001 | RAC2     | 5.98e-06 | 1 个内含子    |
| rs12885474|     |        |     |      |        |        | A   | 0.12   | 0.75 | 0.66 | MACROD2 | 14:24451217 | STXB6     | 6.90e-06 | 1 个内含子    |
| rs11746929|     |        |     |      |        |        | A   | 0.27   | 0.81 | 0.73 | MACR0D2 | 5:149114067 | PPARC1B   | 8.35e-06 | 1 个内含子    |

此表列出了参考 SNP 编号 (rs)、功能和染色体 (chrom) 位置。OR 值、95% 可信区间和 P 值 (以 10 的幂表示) 均源自荟萃分析。每行仅列出与最低 P 值相关的 SNP 表型位点。最后一列显示的是在同一个位点上额外的 SNP 数目，某个 250 kb 以内的 SNP，这些也与表型相关，P<10^-5。关于这些额外的 SNP 的详细内容可在线查询，见表 1。等位基因的存在基于 NCBI build 36 的正链。较小的等位基因 (也叫编码的等位基因) 和较小的等位基因频率 (MAF) 基于荟萃分析实例中的等位基因频率。附录第八部分详细介绍了人群归因危险度 (PAR) 的计算方法。对于那些等位基因与 MRI 诊断的脑梗死呈负相关的 SNP (OR <1.0)，应用主要的等位基因作为风险等位基因计算 PAR。运用人类基因组织 (HUGO) 基因命名系统标志用于命名距离 SNP 最近的 2 个基因，与相关 SNP 的距离可从 5' 端 (开始) 到 3'端 (结束)，选择两者中最近的一个。对每个 SNP 结果的标准化差值以分布于 UCSC 基因组浏览器参考序列计划基因轨道 (hg18)，到基因的距离单位为千碱基对 (kb)。基于 NCBI build 36。

OR=0.99，95% CI 0.76-1.29，P=0.92)，还是在来自 ARIC 研究的 644 例黑人入选者 (平均年龄 61.5 岁，15.5% 存在至少一个 MRI 诊断的梗死灶，较小的等位基因频率为 6.5%，OR=1.26，95% CI 0.74-2.15，P=0.40)，都未能成功地得到 SNP。然而，另外 4 个 SNP (来自 rs2208454，在 MACROD2 基因内含有 3 个，大小为 50-151 kb) 在黑色人种中与 MRI 诊断的脑梗死存在显著的相关性 (rs7268327，P=0.045; rs1998237，P=0.006; rs4464346，P=0.0006; rs8116105，P=0.01)。在测试样本中，rs8116105 也与 MRI 诊断的脑梗死存在显著的相关性，但这种相关是反向的 (P=0.02; 增刊附录，表 II)。黑人与白人在这个区域的连锁不平衡模式存在着本质的差异 (图 4)。复制方法详见增刊附录，第七部分。

本研究的重复分析包含了既往有 TIA 或卒中的参与者 (表 1)。大体上，这些结果相似，都显示了较弱的相关性 (未列出数据)。我们也验证了先前报道过的候选 SNP 或其取代物与 MRI 诊断的脑梗死的相关性，但多因素分析结果显示均无统计学意义 (增刊附录，第九部分)。而且，我们也验证了最近的 CHARGE GWAS 荟萃分析中报道的与缺血性卒中相关的 Ninjurin-2 基因中的最常的 SNP，结果发现这些 SNP 与 MRI 诊断的梗死没有显著的相关性。最后，我们也探讨了 MACROD2 基因中的 SNP23(在最近的 CHARGE GWAS 荟萃分析中报道) 与缺血性卒中的相关性。经验证这些 SNP 与缺血性卒中没有显著的相关性。

讨论

这项关于 MRI 诊断的隐性脑梗死的 GWAS 数据荟萃分析包括了 9401 例伴有 TIA 或卒中病史的参与者，这些参与者来自 6 个基于社区的研究。与 MRI 诊断的脑梗死显著相关 (P=4.64×10^-7) 的 SNP...
为染色体 20p12 上的 rs2208454，位于 MACROD2 基因的内含子 3，FLRT3 基因控制区域的下游。较不常见的等位基因与较低的风险相关。22 个连锁不平衡的 SNP 与 MRI 诊断的脑梗死相关，P<1.0 × 10^{-5}。没有相关性达到我们预置的基因组显著性阈值，5.0 × 10^{-8}。
虽然，在黑人样本中，来自 rs2208454 基因的 4 个 SNP(大小在 200 kb 以内)与 MRI 诊断的脑梗死相关。但对 3C-Dijon 研究的 1822 例白人参与者和 ARIC 研究的 644 例黑人参与者进行的两个复制研究，没有发现 rs2208454 的相关性。最后，我们没有发现任何一个先前报道的 SNP 与 MRI 诊断的隐匿性脑梗死存在显著的相关性。

MACROD2 基因编码的蛋白质的功能仍未探明。此基因包含了大量结构域，在进化上，这些结构域在胎儿和成人的脑组织中进行储存和表达 [20, 21]。这些大的结构域与 ADP 核糖结合，可能是一个涉及很多程序的翻译后修饰，包括 DNA 修复、转录激活和抑制，及端粒和染色质的生物学性能 [22]。FLRT3 基因嵌套在 MACROD2 基因的内含子 3 中，用于编码纤维连接蛋白富含亮氨酸的跨膜蛋白质 3[23]。这种基因可在各种组织中表达，包括脑组织，并且在物种间很好的保存 [24]。它编码的蛋白质可协调同型的细胞粘附，并促进生成纤维细胞生长因子信号肽 [24]。这种信号肽可能参与血管和神经的生发 [25]。在动物实验中，FLRT3 基因可在神经细胞轴突损伤后促进轴突旁支形成 [26, 27]。

有趣的是，纵使 FLRT3 和 Ninjurin-2 均参与神经元损伤后应答反应的调节，然而最近 CHARGE 协作组发表的血性性卒中的 GWAS 研究结果显示，Ninjurin-2 的 SNP 与 MRI 诊断的隐匿性脑梗死无显著的相关性，并且本研究证实 MACROD2 的 SNP 与 MRI 诊断的血性性卒中不存在相关性。对这种差异的可能解释是：MRI 诊断的脑梗死中皮层下较小的梗死占有较大的比例，而血性性卒中存在更多的本质差异。需要进一步的研究，首先要明确 MRI 诊断的脑梗死，血性性卒中和缺血性卒中亚型的外军队列中重复验证这些相关性。

在公开的全基因组表达数量性遗传位点数据库中，MACROD2 的 23 个 SNP 与基因表达不存在显著的相关性，P<10-5 (增刊附录，第六部分)，但这个结果应该被谨慎的解释，因为不到一半的 SNP 可能存在连锁不平衡。这些研究中提供的任何一种基因分型排列，并且其表达数量性遗传位点可能是组织和损伤特异的。

本项荟萃分析具有一定的力度。它包括 6 项具有相似的 MRI 检查方案的大型队列研究。本项分析为了使人群分层风险降到最低，将研究人群限制为白色人种，并对基因分型进行了严格的质量控制。

MACROD2 中的 rs2208454 与 MRI 诊断的隐匿性脑梗死存在较好的相关性 (P=4.64×10-7)。未能成功地在一个独立的白人样本中复制常见 SNP，需要进一步扩大样本量来复制本研究的结果，并进一步研究其他推荐位点的关联性。在白人和黑人种群中对研究结果进行广泛复制可帮助达到精细定位的目的。明确相关基因在分子、临床和流行病学上的相互关系可帮助人们对新的角度理解隐匿性脑梗死的病理生理学和预防方法。