An Interactive Association of Common Sequence Variants in the Neuropeptide Y Gene With Susceptibility to Ischemic Stroke

Chaeyoung Lee, PhD; Minyoung Kong, MS

Background and Purpose—Knowledge of the genetic architecture of ischemic stroke has been quite limited. Most significant associations of candidate genes with ischemic stroke have been difficult to replicate. This might be because the associations were not strong enough for results to be consistent, and testing a mixture of heterogeneous ischemic strokes might lead to confounded genetic associations.

Methods—A preliminary association analysis with 28 sequence variants in 18 candidate genes (ACE, AGT, AGTR1, BDNF, CRP, F13B, LIF, MMP9, NPPA, NPY, PTGS2, SELP, SERPINE1, SREBF2, TFPI, THBD, VCAM1, and VEGF) revealed that NPY might be the most responsible for the susceptibility of ischemic stroke. Forty-five variants were discovered in the NPY gene by full sequencing, and 5 polymorphisms were selected based on their allele frequency and linkage disequilibrium estimates to conduct a thorough examination of their associations with ischemic stroke and its subtypes classified by TOAST. This study was conducted with 271 patients with ischemic stroke and 455 control subjects.

Results—In contrast to a slight significance for an allelic association with ischemic stroke, remarkable discrepancies between haplotype frequencies of control subjects and patients were found. Especially, TA and CC of the haplotypes composed of C4112T and A6411C in the NPY gene were associated with increased risk (\(P=1.8\times10^{-21}\), \(P=2.0\times10^{-11}\)). The interchanged haplotypes, TC and CA, were protective against the diseases (\(P=9.3\times10^{-12}\), \(P=6.0\times10^{-17}\)). The associations were also shown in major subtypes of ischemic stroke.

Conclusions—This remarkable haplotypic association suggested that the interaction between the 2 common sequence polymorphisms in NPY contributed to a great amount of phenotypic variability of ischemic stroke. (Stroke. 2007;38:2663-2669.)

Key Words: genetics ■ ischemia ■ stroke

A large number of studies have focused on the etiological and pathological mechanisms of ischemic stroke. However, knowledge on genetic architecture of ischemic stroke has been quite limited. Most candidate gene studies of ischemic stroke have had difficulty replicating the associations detected.1 The associations were not strong enough to be replicated, and the genetic effects were confounded, which was attributable to testing a mixture of heterogeneous ischemic strokes with various clinical manifestations and etiologies. The ischemic strokes with various etiologies suggested potentially different genetic backgrounds,2,3 and ignoring etiological differences might lead to conflicting results.4

In the current study, we conducted an extensive genetic association study for ischemic stroke, locating strong candidate genes to minimize false-positives and incorporating subtype data to avoid confounding effects.

Methods

Study Population
The study population was composed of patients with ischemic stroke treated at Hallym University Hospital from 2002 to 2005.5 Ischemic stroke was diagnosed by performing CT or MRI from patients with acute stroke within 7 days of onset. Among 327 patients, patients with tumor (n=3), chronic inflammatory diseases (n=12), diabetes mellitus (n=39), or a history of drug abuse (n=0) or alcohol intoxication (n=2) were excluded to avoid any possible confounded effects with them. A total of 271 patients with ischemic stroke were further categorized into its subtypes, large artery atherosclerosis (LAA, n=95), small vessel occlusion (n=110), cardioembolism (n=20), and the other strokes with rare or undetermined etiology using the Trial of Org 10172 in Acute Stroke Treatment (TOAST) classification.5 A subtype analysis was limited to LAA and small vessel occlusion because of small sample sizes of the other subtypes. Four hundred fifty-five subjects aged 55 or older who served as a control group were without a history of cerebral ischemic events and were randomly recruited from routine health checkups, including chest x-ray, gastroscopy, basic health checkup (blood test, urinalysis, liver function test, heart function test, and so on), optional cancer examinations, and a routine survey, before consulting at the same hospital. The exclusion criteria for the stroke patient group were also applied to the control group. Written informed consent was obtained from all subjects, and the study protocol was approved by the ethical committee.

Candidate Gene Selection
A preliminary association analysis with a large number of candidate genes was performed to identify a gene strongly associated with the
Table 1. Candidate Genes and Their Sequence Polymorphisms Used in the Preliminary Analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Location</th>
<th>Suggested Function</th>
<th>Polymorphism</th>
<th>Significance in Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angiotensin I converting enzyme</td>
<td>17q23.3</td>
<td>Atherosclerosis, renin–angiotensin system</td>
<td>289 bp Ins/Del (rs464699A)</td>
<td>CAD, IS</td>
</tr>
<tr>
<td>Angiotensinogen</td>
<td>1q42-43</td>
<td>Atherosclerosis, renin–angiotensin system</td>
<td>T235M (rs699)</td>
<td>MI, CHD</td>
</tr>
<tr>
<td>Angiotensin II receptor, type 1</td>
<td>3q21-25</td>
<td>Inflammation, renin–angiotensin system</td>
<td>A1166C (rs5186)</td>
<td>IS</td>
</tr>
<tr>
<td>Brain-derived neurotrophic factor</td>
<td>11p13</td>
<td>Neurogenesis</td>
<td>V66M (rs6265), C-270T</td>
<td>PD, AD</td>
</tr>
<tr>
<td>C-reactive protein, pentraxin-related</td>
<td>1q21-23</td>
<td>Inflammation</td>
<td>-286C/T/A (rs3091244)</td>
<td>CHD</td>
</tr>
<tr>
<td>Coagulation factor XIII</td>
<td>1q31-q32.1</td>
<td>Angiogenesis</td>
<td>His95Arg (rs6003)</td>
<td>VT</td>
</tr>
<tr>
<td>Leukemia inhibitory factor</td>
<td>22q12.2</td>
<td>Inflammation</td>
<td>C3640A (rs737812), T4524G (rs929271)</td>
<td>VT</td>
</tr>
<tr>
<td>Matrix metallopeptidase 9</td>
<td>20q11.2-13.1</td>
<td>Inflammation, atherosclerosis</td>
<td>C-1562T (rs3918242), Q279R (rs17576)</td>
<td>CAD</td>
</tr>
<tr>
<td>Natriuretic peptide precursor A</td>
<td>1p36.21</td>
<td>Inhibition of renin–angiotensin–aldosterone system, vasodilatation</td>
<td>G664A (rs3063)</td>
<td>HT</td>
</tr>
<tr>
<td>Neuropeptide Y</td>
<td>7p15.1</td>
<td>Cardiovascular homeostasis, vasoconstriction</td>
<td>C-1471T (rs16148), C4112T (rs16135)</td>
<td></td>
</tr>
<tr>
<td>Prostaglandin-endoperoxide synthase 2</td>
<td>1q25.2-q25.3</td>
<td>Inflammation, atherosclerosis</td>
<td>G-765C (rs20417)</td>
<td>MI</td>
</tr>
<tr>
<td>Selectin P</td>
<td>1q22-25</td>
<td>Inflammation</td>
<td>Val640Leu (rs6133)</td>
<td>IS</td>
</tr>
<tr>
<td>Serpin peptidase inhibitor, clade E</td>
<td>7q21.3-22</td>
<td>Inflammation</td>
<td>4G/5G (rs1798889)</td>
<td>MI, IS</td>
</tr>
<tr>
<td>Sterol regulatory element binding</td>
<td>22q13</td>
<td>Lipid homeostasis</td>
<td>G34995T (rs2289657), C6891T</td>
<td>VT</td>
</tr>
<tr>
<td>Transcription factor 2</td>
<td>2q32</td>
<td>Atherosclerosis</td>
<td>T-33C, C536T</td>
<td>VT, CHD</td>
</tr>
<tr>
<td>Thrombomodulin</td>
<td>20p12-cen</td>
<td>Inflammation</td>
<td>G-33A, A455V (rs17846140)</td>
<td>MI, CHD</td>
</tr>
<tr>
<td>Vascular cell adhesion molecule 1</td>
<td>1p32-p31</td>
<td>Inflammation</td>
<td>T-1594C (rs1041163), A22068 (rs3176879)</td>
<td>IS</td>
</tr>
<tr>
<td>Vascular endothelial growth factor</td>
<td>6p12</td>
<td>Neurogenesis, angiogenesis</td>
<td>G-1154A (rs1570369), C-7T (rs25649), C13553T (rs3025039)</td>
<td>VT, CHD, AD, MI, IS, PD</td>
</tr>
</tbody>
</table>

CAD indicates coronary artery disease, IS, ischemic stroke; MI, myocardial infarction; CHD, coronary heart disease, PD, Parkinson disease; AD, Alzheimer disease; VT, venous thrombosis; HT, hypertension; VD, vascular dementia.

Genotyping

Genomic DNA was isolated from peripheral blood cells using a commercially available kit from Qiagen. The sequence polymorphisms, except ACE I/D, were genotyped using the TaqMan polymerase chain reaction assay (Applied Biosystems). Reactions were carried out following the manufacturer’s protocol, and the products were analyzed using ABI PRISM 7900HT (Applied Biosystems). Polymorphism of ACE I/D was identified by polymerase chain reaction using the method described by Rigat et al. Genotyping was performed by laboratory personnel blind to case–control status of the samples.

Sequence Variant Discovery in the NPY Gene

We proceeded to a full-scale investigation of the NPY gene, which was selected as the most probable risk factor of ischemic stroke among candidate genes from our preliminary study. Approximately 9.5 kb covering the entire NPY gene and its promoter were sequenced with genomic DNA samples from 96 randomly selected healthy subjects recruited separately from control subjects. Primers were designed based on GenBank sequences (www.ncbi.nlm.nih.gov/Genkbank, accession number NT_0078191), and their sequences are available on our web site (www.hallym.ac.kr/~clee/NPY).

Statistical Analysis

Pairwise linkage disequilibrium was estimated by D’, and Hardy-Weinberg equilibrium was calculated using Haplovie 3.2. Haplotype frequencies were estimated by an expectation–maximization algorithm using the Arlequin program (http://lgb.unige.ch/arlequin). Odds ratios and their 95% CIs adjusted for gender, age, body mass index, hyperlipidemia, smoking, and hypertension were estimated by logistic regression analyses using SAS Release 9.1 (SAS Institute Inc.). The OR estimates were tested with and without adjustment for multiple testing. A Bonferroni correction was applied to conservatively identify a susceptible gene from the preliminary candidate gene study. A permutation test with 5000 iterations was performed to control false-positives in the association study of the NPY gene. This was because polymorphisms on the same gene were not independent, and the haplotypes sharing with the alleles of the loci were not independent as well.
Results

Clinical Characteristics of Patients and Control Subjects

Distributions of clinical characteristics in the groups used for the association study were compared in supplemental Table I, available online at http://stroke.ahajournals.org. No significant difference was observed between disease and control groups in terms of age, gender, hyperlipidemia, or smoking habit (P > 0.05). Hypertension was found more frequently in patient groups than in the control group (P < 0.05).

Preliminary Association Analysis With 18 Candidate Genes

A significant allelic, genotypic, and/or haplotypic association with ischemic strokes was shown in 4 of the 18 candidate genes (Table 2). The allelic association of C4112T in the NPY gene with LAA susceptibility was more significant than that of V66M in the BDNF gene with small vessel occlusion or that of T4524G in the LIF gene with cardioembolism. Furthermore, a strong association of haplotypes in the NPY gene was observed with ischemic stroke and its subtypes, and only this association was significant after the Bonferroni correction in the preliminary analysis (P < 0.05), which revealed that the NPY gene was the most probable as a risk factor of ischemic stroke.

Sequence Variant Discovery in the NPY Gene

Forty-five sequence polymorphisms, including 6 insertion/deletion loci, were discovered, of which 17 were first identified in this study (Figure). The allele frequencies of the all but 2 loci of 40 and 41 were in Hardy-Weinberg equilibrium (Figure B). For a subsequent association study, 3 additional SNPs (G-1484A, C5325T, and A6411C) were selected based on their minor allele frequency (Figure B) and linkage disequilibrium (D') between loci (Figure A).

Association Analysis With the NPY Gene

A significant allelic association with ischemic stroke was observed only in C4112T. Its T allele was associated with an increased risk of ischemic stroke (OR: 1.32; P < 0.05, supplemental Table II, available online at http://stroke.ahajournals.org). The subjects carrying the T allele were also susceptible to ischemic stroke in a dominant model (OR: 1.41; P < 0.05). Genotypes of C5325T were also associated with ischemic stroke (CC + CT versus TT, OR: 1.76; P < 0.05). When the analysis was extended to the subtypes of ischemic stroke, the association was observed only between alleles of C4112T and the incidence of LAA (OR: 1.43; P < 0.05; supplemental Table II).

A genetic association with ischemic stroke was further investigated by examining frequencies of the haplotypes formed by all the possible combinations at 2 to 5 loci. Many haplotypes were significantly associated with ischemic stroke, and these associations stood significant even by multiple testing that controlled false-positives (P < 0.05, Table 3). The haplotypes composed of the 2 loci, C4112T and A6411C, were more strongly associated with ischemic stroke than the other haplotypes. Their haplotype, TA or CC, was associated with an increased risk of ischemic stroke, whereas the interchanged haplotype, TC or CA, was with a decreased risk. Individuals with the haplotype TA or CC had a 5.7-fold increased risk of developing ischemic stroke compared with those with the haplotype TC or CA (95% CI: 1.43; P = 0.05).
The ORs for SNPs and their haplotypes from the logistic regression analysis using gender, age, body mass index, hyperlipidemia, smoking, and hypertension did not differ from those not adjusted for these factors ($P > 0.05$; data not shown).

**Discussion**

The NPY could be a potential candidate gene supported by its important role in central and peripheral neuronal regulation, cardiovascular control, and blood pressure homeostasis. To date, association studies using sequence variants in the NPY...
Table 3. Haplotypes in the NPY Gene Associated With Combined Ischemic Stroke

| Haplotype* | Frequency | OR 95% CI | P  | | | | L1 | L2 | L3 | L4 | L5 | Stroke | Control | H11002 | H11003 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 loci | | | | | | | | | | | | | | | | |
| G | C | 0.32 | 0.38 | 0.73 | 0.58–0.93 | 0.01 | NS | | | | | | | | | |
| C | C | 0.44 | 0.61 | 0.52 | 0.47–0.65 | 1.8×10⁻⁸ | 1.3×10⁻⁶ | | | | | | | | | |
| T | C | 0.40 | 0.24 | 2.00 | 1.56–2.56 | 1.9×10⁻⁸ | 1.5×10⁻⁶ | | | | | | | | | |
| T | C | 0.16 | 0.06 | 2.99 | 2.05–4.36 | 1.2×10⁻⁸ | 1.0×10⁻⁶ | | | | | | | | | |
| C | C | 0.69 | 0.61 | 1.35 | 1.07–1.72 | 0.011 | NS | | | | | | | | | |
| T | T | 0.15 | 0.24 | 0.44 | 0.47–0.78 | 5.1×10⁻⁴ | 0.042 | | | | | | | | | |
| T | T | 0.15 | 0.10 | 1.58 | 1.13–2.22 | 0.008 | NS | | | | | | | | | |
| C | C | 0.40 | 0.34 | 1.35 | 1.07–1.71 | 0.011 | NS | | | | | | | | | |
| T | C | 0.01 | 0.20 | 0.04 | 0.02–0.61 | 9.3×10⁻⁸ | 7.6×10⁻⁸ | | | | | | | | | |
| T | A | 0.39 | 0.14 | 0.14 | 0.02–0.61 | 9.3×10⁻⁸ | 7.6×10⁻⁸ | | | | | | | | | |

| Haplotype* | Frequency | OR 95% CI | P  | | | | L1 | L2 | L3 | L4 | L5 | Stroke | Control | H11002 | H11003 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 loci | | | | | | | | | | | | | | | | |
| G | C | 0.32 | 0.38 | 0.73 | 0.58–0.93 | 0.01 | NS | | | | | | | | | |
| T | A | 0.29 | 0.54 | 0.36 | 0.28–0.46 | 6.0×10⁻¹⁷ | 4.9×10⁻¹⁵ | | | | | | | | | |
| C | C | 0.15 | 0.07 | 2.01 | 1.41–2.66 | 1.1×10⁻⁴ | 0.009 | | | | | | | | | |
| C | C | 0.15 | 0.05 | 3.34 | 2.24–4.98 | 3.7×10⁻⁸ | 3.0×10⁻⁵ | | | | | | | | | |
| T | C | 0.72 | 0.55 | 1.30 | 1.18–1.43 | 0.003 | NS | | | | | | | | | |
| C | A | 0.12 | 0.18 | 0.72 | 0.50–1.03 | 0.34 | NS | | | | | | | | | |
| C | C | 0.13 | 0.20 | 0.71 | 0.50–1.04 | 0.34 | NS | | | | | | | | | |
| C | C | 0.39 | 0.24 | 1.91 | 1.50–2.43 | 1.4×10⁻⁷ | 1.2×10⁻⁵ | | | | | | | | | |
| C | A | 0.29 | 0.13 | 0.45 | 0.30–0.64 | 0.032 | NS | | | | | | | | | |
| C | C | 0.15 | 0.12 | 1.42 | 1.02–1.98 | 0.036 | NS | | | | | | | | | |
| C | C | 0.15 | 0.10 | 1.77 | 1.25–2.50 | 0.001 | NS | | | | | | | | | |
| C | C | 0.01 | 0.06 | 0.04 | 0.02–0.35 | 2.6×10⁻⁵ | 1.9×10⁻⁵ | | | | | | | | | |
| C | C | 0.01 | 0.06 | 0.04 | 0.02–0.35 | 2.6×10⁻⁵ | 1.9×10⁻⁵ | | | | | | | | | |

| Odds ratios and 95% CIs were adjusted for gender, age, body mass index, hyperlipidemia, smoking, and hypertension and only those with P<0.05 are listed. Each haplotype was compared with the other haplotypes composed of the same loci. P is the permuted P value with 5,000 iterations. NS indicates P>0.05. |

*1L1=G-1484A; L2=C-1471T; L3=A6411C; L4=L5=Δ6411C.

gene for complex traits have been conducted quite selectively. For instance, C-399T in promoter and Leu7Pro in exon 2 have been intensively studied after their respective associations with schizophrenia37 and cholesterol level38 were found. Only the Leu7Pro variant, an extremely rare or monomorphic locus in nonwhites such as blacks, Chinese, Japanese, and Koreans15,39 has been examined in the context of susceptibility for cardiovascular or cerebrovascular diseases.34,40–41 A significance was reported only in Swedish patients with hypertension with myocardial infarction and stroke.34

In the current study, powerful associations of NPY haplotypes with ischemic stroke were discovered, especially with C4112T and A6411C. The 2 loci could explain the genetic association between the NPY gene and ischemic stroke better than C-1471T and C4112T, which showed a considerable significance in the preliminary analysis. A cautious, but plausible, interpretation can be offered that the associations of the latter haplotypes might be derived as partial effects of the former haplotypes by linkage rather than as their own effects. Even any haplotypes composed of more than 2 loci, including C4112T and A6411C, were not as informative as those composed of the 2 loci in predicting a likelihood of susceptibility. For example, the haplotypes composed of the 2 loci and C-1471T also showed strong associations with the disease (Table 3). The associations were mainly attributable to the effects of the 2 loci, and the interaction effect with C-1471T was negligible.

Of the 4 haplotypes composed of C4112T and A6411C, TA and CA were associated with an increased risk of ischemic stroke, whereas the interchanged haplotypes, TC and AC, were found to be protective against the disease. The allelic contribution to genetic susceptibility in one locus depended absolutely on counter allele of the other locus, indicating a strong interaction between the loci.

The remarkable interaction effects were also demonstrated in major subtypes of ischemic stroke, LAA and small vessel occlusion (www.hallym.ac.kr/~clee/NPY.htm). The 4 haplotypes exerted consistent influence on the subtypes as predisposition or protection factors. The consistency suggested that the NPY gene might contribute to pathogenesis of heterogeneous ischemic strokes as more of a common denominator than as a subtype-specific factor. This may be in line with the roles of polypeptide NPY observed in previous studies in
which NPY-induced mitogenesis of vascular smooth muscle and affected both functions of vasoconstriction and vasodilatation. Its contrasting roles in vascular homeostasis might suggest a meaningful direction in understanding the pathophysiological mechanisms behind the association between NPY polymorphisms and ischemic stroke.

The haplotype association with ischemic stroke was not influenced by stroke risk factors analyzed in this study. Although hypertension was more prevalent in patients than in control subjects, it did not affect the genetic effects. Further studies with more expanded designs are in order to clarify other effects such as diabetes mellitus, chronic inflammatory diseases, and tumors on the haplotype association.

The current study provided the first evidence of the association between the NPY gene and susceptibility of ischemic stroke. The identified genetic variants conferring risk for the disease were commonly present SNPs. In addition, our results suggested an importance of haplotype analysis for complex traits by showing a strong interaction among SNPs and called for more attention on intronic SNPs that have been often neglected. Our findings should be replicated with a larger group of patients with ischemic stroke for practical applications. Also, further studies on the haplotype effects are warranted to elucidate their underlying mechanisms.

Acknowledgments

We appreciate 3 anonymous reviewers’ helpful comments and suggestions. We thank healthy participants, patients with ischemic stroke, and their family members whose contribution made this work possible. We also thank all the neurologists who recruited their patients for this study and our laboratory personnel who collected clinical data and participated in SNP discovery.

Source of Funding

This study was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health and Welfare, Republic of Korea (A0200007).

Disclosures

None.

References

33. Michel MC, Rascher W. Neuropeptide Y gene and ischemic stroke. 2669. Lee and Kong Neuropeptide Y Gene and Ischemic Stroke
An Interactive Association of Common Sequence Variants in the Neuropeptide Y Gene With Susceptibility to Ischemic Stroke
Chaeyoung Lee and Minyoung Kong

Stroke. 2007;38:2663-2669; originally published online August 16, 2007;
doi: 10.1161/STROKEAHA.107.482075
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2007 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/38/10/2663

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/