A Genome-Wide Study of Moyamoya-Type Cerebrovascular Disease in the Korean Population

Sung-Pil Joo, M.D.,1 Tae Sun Kim, M.D.,1 Il-Kwon Lee, PhD.,2 Joon-Tae Kim, M.D.,2 Man-Seok Park, M.D.,2 Ki-Hyun Cho, M.D.2

Departments of Neurosurgery,1 Neurology,2 Chonnam National University Hospital & Medical School, Gwangju, Korea
Genome Research Center for Hematopoietic Disease,1 Chonnam National Hwasun Hospital, Hwasun, Korea

Objective : Structural genetic variation, including copy-number variation (CNV), constitutes a substantial fraction of total genetic variability, and the importance of structural variants in modulating susceptibility is increasingly being recognized. CNV can change biological function and contribute to pathophysiological conditions of human disease. Its relationship with common, complex human disease in particular is not fully understood. Here, we searched the human genome to identify copy number variants that predispose to moyo-moya type cerebrovascular disease.

Methods : We retrospectively analyzed patients who had unilateral or bilateral steno-occlusive lesions at the cerebral artery from March, 2007, to September, 2009. For the 20 subjects, including patients with moyamoya type pathologies and three normal healthy controls, we divided the subjects into 4 groups : typical moyamoya (n=6), unilateral moyamoya (n=9), progression unilateral to typical moyamoya (n=2) and non-moyamoya (n=3). Fragmented DNA was hybridized on Human610Quad v1.0 DNA analysis BeadChips (Illumina). Data analysis was performed with GenomeStudio v2009.1, Genotyping 1.1.9, cnvPartition_v2.3.4 software. Overall call rates were more than 99.8%.

Results : In total, 1258 CNVs were identified across the whole genome. The average number of CNV was 45.55 per subject (CNV region was 45.4). The gain/loss of CNV was 52/249, having 4.7 fold higher frequencies in loss calls. The total CNV size was 904,657,868, and average size number variations can affect the expression of genes, alter the organization of chromatin, and/or influence the regulation of nuclotide polymorphisms (SNP). Research also reveals that copy number variations can affect the expression of genes, alter the organization of chromatin, and/or influence the regulation of genes in the vicinity.

Conclusion : Significant association between unilateral MMD and progression of unilateral to typical moyamoya was observed. The finding was confirmed again with clustering analysis. These data demonstrate that certain CNV associate with moyamoya-type cerebrovascular disease.

Key Words : Copy number variation (CNV) · Whole genome association study · Moyamoya disease.

INTRODUCTION

Typical moyamoya disease (MMD) is defined by specific angiographic findings of diffuse stenotic or occlusive lesions of the bilateral carotid fork and unique collateral vessels at the base of the brain. However, some atypical cases of MMD show unilateral lesions on angiography and a normal terminal portion of the contralateral internal carotid artery (ICA) or proximal middle cerebral artery (MCA), or even progression from unilateral to bilateral disease. Whether unilateral MMD is an early form of moyamoya disease remains controversial. Identifying predictive factors for the development of moyamoya in an at-risk population coupled with data supporting an effective treatment could lead to better outcomes for these patients through earlier diagnosis and treatment. Patients with unilateral moyamoya disease can be difficult to treat, particularly regarding predicting the likelihood of subsequent progression to the contralateral side. The speed of progression of the vasculopathy associated with moyamoya type vascular disease is extremely variable. Copy number variation (CNV) refers to an intermediate-scale genomic change in segments greater than 1.0 kilo base pairs but typically less than 5 mega bases in length. Typically CNV is defined as structural change since CNVs modify the structure of the genome. CNVs include both additional copies of sequence (duplications) and losses of genomic region (deletions). Evidence is accumulating that CNVs play important roles in human disease. Based on results, human CNV is now thought to affect more base pairs than other forms of mutation such as single nucleotide polymorphisms (SNP). Research also reveals that copy number variations can affect the expression of genes, alter the organization of chromatin, and/or influence the regulation of genes in the vicinity.

The purpose of the present study was to clarify the genetic characteristics of unilateral moyamoya disease compared to...
A Genome-Wide Study of Moyamoya-Type Cerebrovascular Disease in the Korean Population | SP Joo, et al.

MATERIALS AND METHODS

Study samples
We divided the patients into 4 groups on the basis of angiographic findings. The ‘unilateral moyamoya group’ (unimoya group, n=9) included patients with only unilateral abnormal vessels. The ‘typical moyamoya group’ (moyamoya group, n=6) included patients with bilateral steno-occlusive lesions of the ICA terminal portion and moyamoya vessels at the base of the brain. The ‘unilateral to bilateral moyamoya’ (n=2) comprised subjects with progression of unilateral lesion to bilateral lesions. The ‘non-moyamoya’ (healthy control, n=3) comprised subjects with steno-occlusive lesions of the ICA terminal portion or proximal MCA without moyamoya vessels (Table 1). We excluded patients with suspicious contralateral involvement at the time of diagnosis.

Using angiographic findings, we investigated the frequency of collateral vessels and the staging of both typical moyamoya and unilateral moyamoya disease. The staging was performed according to Suzuki and Kodama19, which was based on angiographic findings. Baseline patient characteristics (comparison of unilateral moyamoya and unilateral to bilateral moyamoya) are summarized in Table 2. Genomic DNAs were extracted from peripheral blood B-lymphocytes using QIAGEN DNA extraction kits according to manufacturer’s instruction. The quality and quantity of DNA were examined by NanoDrop / Picogreen assays and 1% agarose gel electrophoresis.

Copy number variation discovery
Approximately 200 ng of DNA was used to genotype each subject sample according to the manufacturer’s protocol (Illumina, San Diego, CA, USA). After accurate quantification, whole genome amplification, random fragmentation, purification, and hybridization, the specifically hybridized DNA was fluorescently labeled by allele specific primer extension and detected using a BeadArray scanner. Non-specifically hybridized fragments were removed by washing, and specifically hybridized DNAs were processed for the single base extension reaction and stained. Images were obtained according to standard Illumina scan procedures using Illumina BeadStation 500G Array scanner. Intensity files were processed by GenomeStudio GT module 1.1.9 with default analysis settings. Chromosomal aberrations are detected by comparing the normalized intensity of a subject sample and a reference sample using two modes of analysis. The first is a single-sample mode in which reference values are derived from canonical genotyping clusters (at 0, 0.5, and 1.0) created from clustering on normal reference samples. The second is a paired-sample mode in which direct intensity (R) comparisons between a subject sample and its paired reference sample are performed. The basis for detecting chromosomal aberrations are genomic plots of the log2 (Rsubject/Reference; log R

Table 1. Baseline patient data

<table>
<thead>
<tr>
<th>Group</th>
<th>No.</th>
<th>Sample group</th>
<th>ng/µL</th>
<th>Well</th>
<th>260/280</th>
<th>260/230</th>
<th>Call rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Unilateral MMD</td>
<td>292.02</td>
<td>A1</td>
<td>1.84</td>
<td>1.9</td>
<td>0.9984</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Unilateral MMD</td>
<td>265.84</td>
<td>B1</td>
<td>1.81</td>
<td>1.7</td>
<td>0.9985</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>Unilateral MMD</td>
<td>236.58</td>
<td>C1</td>
<td>1.86</td>
<td>1.91</td>
<td>0.9983</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>Unilateral MMD</td>
<td>216.9</td>
<td>D1</td>
<td>1.85</td>
<td>2.25</td>
<td>0.9983</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>Unilateral MMD</td>
<td>161.92</td>
<td>E1</td>
<td>1.85</td>
<td>1.7</td>
<td>0.9980</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>Unilateral MMD</td>
<td>235.7</td>
<td>F1</td>
<td>1.89</td>
<td>2.15</td>
<td>0.9982</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>Unilateral MMD</td>
<td>167.05</td>
<td>G1</td>
<td>1.85</td>
<td>1.78</td>
<td>0.9983</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>Unilateral MMD</td>
<td>210.3</td>
<td>H1</td>
<td>1.88</td>
<td>2.37</td>
<td>0.9981</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>Unilateral MMD</td>
<td>129.8</td>
<td>A3</td>
<td>1.88</td>
<td>2.38</td>
<td>0.9986</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>Typical MMD</td>
<td>313.46</td>
<td>B3</td>
<td>1.85</td>
<td>1.91</td>
<td>0.9982</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>Typical MMD</td>
<td>125.29</td>
<td>C3</td>
<td>1.87</td>
<td>1.7</td>
<td>0.9955</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>Typical MMD</td>
<td>71.48</td>
<td>D3</td>
<td>1.9</td>
<td>2.33</td>
<td>0.9981</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>Typical MMD</td>
<td>153.1</td>
<td>E3</td>
<td>1.87</td>
<td>2.34</td>
<td>0.9984</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>Typical MMD</td>
<td>286.6</td>
<td>F3</td>
<td>1.85</td>
<td>2.08</td>
<td>0.9986</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>Typical MMD</td>
<td>251.2</td>
<td>G3</td>
<td>1.86</td>
<td>2.06</td>
<td>0.9983</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>Unilateral=&gt;Typical MMD</td>
<td>257.13</td>
<td>H3</td>
<td>1.86</td>
<td>1.84</td>
<td>0.9983</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>Unilateral=&gt;Typical MMD</td>
<td>265.3</td>
<td>A5</td>
<td>1.86</td>
<td>2.06</td>
<td>0.9981</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>Control</td>
<td>142.10</td>
<td>B5</td>
<td>1.9</td>
<td>1.78</td>
<td>0.9984</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>Control</td>
<td>169.82</td>
<td>C5</td>
<td>1.87</td>
<td>1.7</td>
<td>0.9985</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>Control</td>
<td>282.71</td>
<td>D5</td>
<td>1.86</td>
<td>1.85</td>
<td>0.9984</td>
</tr>
</tbody>
</table>

MMD : moyamoya disease
ratio) and the allele frequency parameters, which are the known B allele frequencies of the three canonical clusters.

**Functional classification of genes**

To explore associated genes in terms of their molecular function and biologic pathway, Protein ANalysis Trough Evolutionary Relationships (PANTHER) analysis was performed. Proteins are classified into families and subfamilies of shared function, which are then categorized by molecular function and biological process ontology terms.

**Data processing and statistical analysis of SNPs**

Intensity files were processed by GenomeStudio GT module 1.1.9 with default analysis settings. Each SNP is analyzed independently to cluster and identify genotypes. Genotype calls are generated by comparing experimental data with those in the supplied cluster file. Calls are generally highly accurate (average call rate was 99.8%) and unambiguous for high quality samples. A false discovery rate was used for multiple testing corrections. *p*<0.05 was considered statistically significant.

**RESULTS**

**Distribution of CNVs/CNVRs and general characteristics**

We performed a genome-wide screen for CNVs by analysis of Illumina Human610 Quad v1.0 DNA analysis BeadChip (Illumina) for 17 patients with moyamoya disease and 3 controls. We examined CNVs for association with moyamoya diseases using linked genomic information with higher order functional information of genes deleted or gained in patients with moyamoya disease by searching the Gene Ontology (GO) and KEGG Kyoto Encyclopedia of Genes and Genomes pathway databases. The average call rate of samples (male 9, female 11) was 99.8% (Table 1) from 620091 markers in autosomal chromosomes (sex chromosomes and pseudo-autosomal markers were excluded). Selected were included in the imaging analysis if they did not depart from Hardy-Weinberg equilibrium. cnnPartition v2.3.4 was used for estimation of copy number and annotation regions. Genome-wide detection of genomic abnormalities in moyamoya disease using high-resolution single SNP array-based analysis revealed a high incidence of gains and losses. In total, 1258 CNVs were identified across the whole genome. The average number of CNV was 45.55 per genome (CNVR)=45.4. The median size of CNVs was 1095202 bp (range 569-5430796 bp). The gain/loss occurrence rate of CNV was 52/249, having a 4.7 fold higher frequencies in loss calls. Average copy neutral loss of heterozygosity was 30.5.

Total CNV size was 904,657,868 bp, and the average size was 993,038 bp. The largest portion of CNVs (613 calls) were 1 M-10 M in length. CNVRs were defined as more than 30% overlap by merging overlapping CNVs in two or more individuals. In total, 373 CNVRs were identified across the genome, with a median size of 1164624 bp (range 569-5509685 bp). The largest portion of these CNVRs (613 calls) were 1M-10M in length. 373 CNVRs encompassing 436.6 Mb accounted for -14.2% of human genome.

Novel CNVRs detected

A total of 373 CNVRs identified were compared with the 14478 CNVRs in the Database of Genomic Variants (DGV). If a CNVR overlapped a DGV entry, it was counted as a known CNVR. In total, 290 potentially novel CNVRs (77.7%) were identified, whereas 83 remaining CNVRs were known (22.3%). Unknown CNVRs could be Korean-specific CNVRs, indicating that previous CNV coverage of the human genome is incomplete and there is ethnic diversity in CNVRs. More detailed knowledge of CNVs in the human genome will be useful for genetic association studies.

**Differential distribution of CNVs between moyamoya subtypes**

There was a significant association between unilateral MMD and the allele frequency parameters, which are the known B allele frequencies of the three canonical clusters.
subjects) and progression of unilateral to typical moyamoya (2 subjects). The result was confirmed by hierarchical cluster analysis to identify subgroups of moyamoya with separate CNVRs profiles. Thirteen CNVRs (8 novel, 5 known CNVRs) encompassing 212 genes or Gene Locus (LOC) numbered unknown targets were common between the two. The most frequent CNVR was chr3:163699360-163712278, a known CNVR, with 2 homozygous losses and 7 losses.

**Candidate genes residing within the significantly associated regions**

Total of 140 genes with known function from CNVRs were further analyzed with the PANTHER GO classification tool. The 3 major biologic processes involved with those genes were metabolic process, cellular process, and cell communications (Fig. 1A). Classification by molecular function showed catabolic activity as well as binding and receptor activity (Fig. 1B). Of particular interest, contactin-5, collagen alpha-1 (XI) chain, activated CDC42 kinase 1, PAK-2p34, and phosphotidylinositol phosphatase PTPRQ gene are involved in angiogenesis. The functional implications of these 5 candidate genes in terms of vascular-wall homeostasis are largely unknown but warrant further genetic study using a larger sample size.

**DISCUSSION**

After the successful discoveries of genetic associations for common disorders using SNPs in genome-wide association scans, new efforts are ongoing to evaluate structural variations in disease, mainly in the form of CNV. Until recently, SNPs were thought to be the predominant form of genomic variation\(^1\,\,^2\). However, SNPs have been partially disappointing in strongly linking genetic variability with common disease\(^4\). Recent advancement in technology has facilitated a shift from locus-specific studies to genome-wide assessment of genetic variation. In 2004, two groups independently described the widespread presence of CNVs in the genomes of healthy people with no obvious genetic disorder\(^7\,\,^18\). With this accumulation of information, it now seems appropriate to review our current understanding of CNV and its significance in human phenotypic variation (including disease resistance and susceptibility) and to discuss future directions for studies in this field. Genetic analysis has the poten-

---

Fig. 1. PANTHER analysis. A: Classification by biologic process. B: Classification by molecular function. PANTHER: Protein ANalysis Through Evolutionary Relationships, GO: gene ontology.

**GO molecular function**

Total #genes: 140, Total #function hits: 158

**GO biological process**

Total #genes: 140, Total #process hits: 380
tial to uncover underlying pathogenic mechanisms of moyamoya disease that could enhance our understanding of vascular disease and identify novel targets for therapeutic intervention.

A number of genome-wide linkage studies have been performed in Japanese families with familial moyamoya that suggest linkage to 5 different chromosomal regions: 3p24.2-p26, 6q25, 8q23, 12p12, and 17q25. Penetrance is in part age dependent and influenced by genomic imprinting. Yamuchi et al. used microsatellite markers to interrogate chromosome 17 given the association between moyamoya disease and neurofibromatosis type 1 (17q11.2). Furthermore, Mineharu et al. performed separate analysis in 15 extended Japanese families using genome-wide parametric linkage analysis and identified a candidate locus at 17q25.3. These data show compelling evidence that a pathogenic gene is associated with familial moyamoya disease.

Another approach with significant potential to elucidate genetic factors associated with moyamoya disease is genome-wide association studies that involve the use of high-density SNP arrays. This approach was recently exemplified by an elegant study from Gunel’s group into the pathogenesis of intracranial aneurysms, and this pivotal effort resulted in the identification of common SNPs on chromosomes 2q, 8q, and 9p that show significant association with intracranial aneurysm. However, similar to moyamoya disease, genome-wide linkage studies of intracranial aneurysms in familial cases have been unable to identify a replicable loci, and candidate gene-based case-control association study have failed to produce replicable results.

Genome-wide association approaches are now being applied in moyamoya disease to uncover the underlying pathogenic mechanisms. Hallemeier et al. reported that collateral vessels more easily develop from the normal side to the involved side in patients with unilateral moyamoya disease than in those with bilateral moyamoya disease. Therefore, the hemodynamics of patients with unilateral moyamoya may be better, leading to fewer cerebrovascular complications.

One of the major concerns facing patients with unilateral moyamoya is how to address the non-affected side. Our data corroborate previous reports from our group and others that many patients with unilateral moyamoya may develop disease on the non-affected side in a very delayed fashion or not at all. The ability to accurately predict patients with unilateral moyamoya who are at greater risk for developing contralateral disease would be of substantial clinical advantage. An appreciation of the expected time course for progression, coupled with knowledge of specific risk factors associated with development of contralateral disease, would enable clinicians to more accurately design individualized follow-up strategies. Ultimately, this knowledge would enable increased numbers of patients to receive a diagnosis and be treated prior to experiencing a stroke. Future studies should also aim to develop genetic and serum biomarkers that can predict outcomes after the initial presentation. Biomarkers reflective of changes in the stenotic/occlusive intracranial environment after revascularization surgery will help clinicians observe disease evolution in patients with moyamoya disease. Although data on CNVs are still largely rudimentary, new developments in high-resolution scanning technology will likely facilitate the establishment of comprehensive CNV maps. It is unlikely that any one technology alone will allow thorough identification of all CNVs, so future work should focus on verifying primary results, integrating multiple data sources, and assigning population frequencies to these genomic variants.

A genome-wide screen for CNVs by analysis of Illumina Human610 Quad v1.0 DNA analysis BeadChip (Illumina) for 17 patients with moyamoya disease and 3 controls was carried out. DNA copy-number variation analysis using high-density SNP arrays reveals a molecular karyotype for individual patients and leads to the detection of novel chromosomal high-resolution SNP array-based analysis allowed us a high incidence of gains and losses in different moyamoya subtypes. In total, 1258 CNVs were identified across the whole genome. The gain/loss occurrence rate of CNV was 52/249, having a 4.7 fold higher frequencies in loss calls. Total CNV size was 904657868 bp, and the average size was 993038 bp. We also identified 290 unknown CNVRs that could be Korean-specific novel CNVRs, indicating that previous CNV coverage of the human genome is incomplete, particularly for different ethnic groups. Those novel CNVRs (77.7%) could be Korean-specific CNVRs, indicating that previous CNV coverage of the human genome is incomplete and there is ethnic diversity in CNVRs. Until recently, most of the reported CNVs have been identified in Caucasians, which may not be directly applicable to people of different ethnicities including Korean population. In this context CNVRs profiling in Korean will be useful resources for investigating the association between CNVs and disease phenotypes in East-Asian population. We also found an association of CNV’s between unilateral moyamoya and progression to typical moyamoya. 13 CNVRs were identified as DNA structural variations. In order to define associated genes in terms of their molecular function and biologic pathway, GO analysis was performed, and then categorized by molecular function and biological process ontology terms. Five candidate genes with angiogenesis GO terms were clustered in this region, which suggests that CNVRs are likely associated with blood vessel formation. This data seems consistent with previous observations that positive findings from one study often fail to replicate in other studies, but it could be due to small sample size of the study. Therefore, further studies on larger sample size, as well as functional validation will define their role in the pathogenesis and progression to typical moyamoya.

CONCLUSION

Moyamoya cerebrovascular disease is an important cause of stroke in children and young adults. Further investigations are needed to identify the underlying cause of moyamoya disease.
DNA copy-number variation analysis using high-density SNP arrays revealed a molecular karyotype for individual patients with moyamoya and led to the detection of novel chromosomal aberrations and allowed us a high incidence of gains and losses in different moyamoya subtypes. Many of those detectable lesions were found to be previously unidentified cryptic chromosomal aberrations. Total of 290 unknown CNVRs identified in this study could be Korean-specific novel CNVRs. In this context CNVRs profiling in Korean will be useful resources for investigating the association between CNVs and disease phenotypes in East-Asian population. Genome-wide association studies in moyamoya disease will not only help overcome limitations of previous genetic investigations in familial cases only but also be important for the development of novel therapies specific to moyamoya disease.

- Acknowledgements

This work was supported by a grant of the Chonnam National University hospital Research Institute of Clinical Medicine (Grant No : CRI 199063-1), and also supported by Research Institute of Medical Science of Chonnam National University (2008-CURIMS-DR003) and Special Research of Chonnam and also supported by Research Institute of Medical Science of Chonnam Na

References