Genome-Wide Linkage Scan of Common Stroke in Families From Northern Sweden

Sofie Nilsson-Ardnor, MD; Tomas Janunger, MSc; Per-Gunnar Wiklund, MD, PhD; Kurt Lackovic, PhD; Anna Karin Nilsson, MSc; Petter Lindgren, MStat; Stefan A. Escher, PhD; Birgitta Stegmayr, PhD; Kjell Asplund, MD, PhD; Dan Holmberg, PhD

Background and Purpose—Taking advantage of low genetic variations in northern Sweden, we performed a genome-wide linkage scan to investigate the susceptibility loci for common forms of stroke.

Methods—Fifty-six families, containing multiple cases of stroke and whose data had been previously used to replicate linkage to the phosphodiesterase 4D locus on chromosome 5q, were genotyped in a genome-wide scan. Fine mapping was performed, and subsequently 53 additional families from the same region were genotyped over the candidate regions.

Results—Linkage calculations were performed by using 3 different disease models, from a very broad (all stroke cases defined by World Health Organization MONICA criteria) to a narrower (ischemic stroke only) stroke phenotype. With all models, nonparametric multipoint linkage analysis yielded allele-sharing log of the odds (LOD) scores ≥1.2 at 9 locations: 1p34, 5q13, 7q35, 9q22, 9q34, 13q32, 14q32, 18p11, and 20q13. The highest allele-sharing LOD scores were obtained on chromosomes 5q (previously reported), 1p (LOD = 2.09), and 18p (LOD = 2.14). Fine mapping resulted in increased allele-sharing LOD scores for chromosome 5q (previously reported) and 9q22 (LOD = 1.56), but all others decreased. Combining these initial results with a subsequent analysis of 53 additional families, we obtained the highest allele-sharing LOD scores on chromosomes 5q, 13q, and 18p, although none reached the initial genome-wide allele-sharing LOD scores.

Conclusions—Genetic analysis of stroke in families from northern Sweden did not identify any new major stroke loci. This indicates that multiple minor susceptibility loci in addition to the previously known locus on chromosome 5 could contribute to the disease. (Stroke. 2007;38:34-40.)

Key Words: genetics ■ linkage ■ stroke

In Sweden, ≈30 000 stroke events occur annually. Studies on the genetic epidemiology of stroke as well as in animal models implicate genetic components that contribute to stroke risk. At present, only 1 whole-genome linkage study for common forms of stroke has been published, identifying a susceptibility locus on chromosome 5q12. In this region, the phosphodiesterase 4D gene was further investigated, and an association between genetic variants of phosphodiesterase 4D and ischemic stroke was later reported. In addition, a linkage study of myocardial infarction (MI) identified a region on chromosome 13q12–13, and in this region, the gene ALOX5AP, encoding 5-lipoxygenase–activating protein, was associated with a greater risk of MI and stroke. In northern Sweden, low genetic variation in the population has previously been taken advantage of for successful linkage studies of complex diseases. We have previously confirmed linkage to the phosphodiesterase 4D region in 56 extended pedigrees from northern Sweden. Here we report data for a whole-genome linkage scan of these pedigrees and fine mapping over 9 regions of interest that also includes an additional 53 families.

Subjects and Methods

Recruitment and Validation of Subjects

To identify probands for the study, we used a population-based stroke registry established in 1985 at the northern Sweden MONICA center. Familial cases of stroke were identified by questionnaires that were sent to all living patients in the registry who had been affected at <70 years of age between the years 1985 and 1996; these questionnaires asked for information on a family history of stroke. In northern Sweden, low genetic variation in the population has previously been taken advantage of for successful linkage studies of complex diseases. We have previously confirmed linkage to the phosphodiesterase 4D region in 56 extended pedigrees from northern Sweden. Here we report data for a whole-genome linkage scan of these pedigrees and fine mapping over 9 regions of interest that also includes an additional 53 families.

Stroke is available at http://www.strokeaha.org DOI: 10.1161/01.STR.0000251643.37454.16
and the questionnaire was thus answered by 79% of the surveyed individuals. A family history of stroke was reported by 654 individuals, and of those, 197 reported an affected sibling. The affected sibling was alive in 125 families, and in 16 of these, several affected siblings were still alive. The diagnosis for each affected family member was validated in the MONICA register or by hospital records. The World Health Organization MONICA criteria include subarachnoid hemorrhage (SAH) in its stroke definition but exclude transient ischemic attack. One hundred one families with at least 2 affected first-degree relatives available for sampling were finally identified. In the sampling procedure, all available affected family members, as well as available unaffected siblings and children of those affected, were included. Blood samples for DNA preparation were usually obtained by arrangement with the local general practitioner. After identification and collection of individual samples, families were submitted to genealogical studies with the use of public and private databases. Only families in which both parents of the proband had been born in northern Sweden (in either Norrbotten or Västerbotten, the 2 northernmost counties of Sweden) were included in the study. In the follow-up study, additional familial cases of stroke were identified by questionnaires sent to all living patients in the stroke registry who had been affected at \textit{H11021} years of age between the years 1997 and 2001 (1701 individuals). Also in this second recruitment, 79% answered the questionnaire. Fifty-nine families with at least 2 affected first-degree relatives and 4 families with second- or third-degree relatives available for sampling were identified. Medical records of the affected individuals in both recruitments were reviewed for the prevalence of risk factors by an experienced stroke physician, and ischemic subjects were subtype according to the TOAST classification system.\textsuperscript{13} Clinical and genealogical information of participants in the study was continuously updated by screening the registers for new stroke events among previously healthy siblings and by genealogical research. Each subject was given a unique identification number before genotyping. The participants in the study provided informed consent for access to medical records and donation of blood for genetic research. This study was approved by the research ethics committee of Umeå University and the data-handling procedures by the National Computer Data Inspection Board.

**Genotyping**

DNA was extracted from whole blood according to standard phenol purification methods. A total of 445 polymorphic microsatellite markers from the ABI prism linkage mapping set, v2.5 HD10 and HD5 (Applied Biosystems), were genotyped. In the second step of the study, additional markers from the ABI prism linkage mapping set, v2.5 HD5, were genotyped to investigate linkage regions observed in the initial genome-wide scan. Polymerase chain reaction conditions, genotype analysis, and examinations were performed as previously described.\textsuperscript{10} Genotypes inconsistent with mendelian laws were reexamined and corrected or deleted when possible.

**Statistical Analysis**

Linkage analysis was performed with a model-free, multipoint approach in Allegro\textsuperscript{14} with the functions previously described.\textsuperscript{10} We used the genetic map with interpolated genetic marker positions from David Duffy (http://www.qimr.edu.au/davidD/master_map.dat), in which the genetic positions are interpolated by locally weighted linear regression from the National Center for Biotechnology Information build 34.3 physical map positions and the Rutgers genetic map.\textsuperscript{15} Marker allele frequencies were estimated among all pedigree members according to the Merlin algorithm.\textsuperscript{16} To estimate empirical genome-wide probability values, we performed a simulation study of 1000 randomly created data sets under the null hypothesis of no linkage. The simulation, with the assumption of a genotyping success rate of 0.95, was performed in Allegro with the same allele frequencies, marker maps, and pedigrees as in the original 8.6-centimorgan (cM) genome-wide scan for disease model II.

**Results**

**Genome-Wide Scan**

In the initial family recruitment, we identified 101 families and collected 556 individual samples. Of the 101 identified families, 25 were excluded after genealogical and clinical validation or because of the lack of blood samples, with the criteria displayed in Figure 1. Of the remaining 76 families, the 56 families considered to provide the most information regarding identical-by-descent sharing were selected for genotyping in the genome-wide scan. These families included 129 affected individuals (including 11 cases of SAH) from a total of 376 individuals.
For statistical analysis, we selected 3 different disease models. Model I is broad, including all stroke cases captured by the World Health Organization MONICA definition of stroke, thus excluding transient ischemic attack but including SAH. Model II consisted of ischemic and intracerebral hemorrhagic stroke, and model III included ischemic stroke only. Further subdivision according to the TOAST classification was not considered for our linkage calculations because of the low number of subjects with complete classification data. In all models, the diagnosis was based on the first stroke event, with 1 exception: if the subject first experienced an SAH and later an ischemic stroke (4 cases), the second event was allowed as affected in both models II and III. The distribution of stroke subtypes within the families of model I is displayed in Table 1. To enable comparisons with previous studies that mainly investigated common stroke excluding SAH, we display in Table 2 the clinical characteristics and risk factor distribution for model II only. The characteristics of stroke cases in our study do not markedly differ from the general stroke population in northern Sweden.

Initial linkage analysis, with marker loci distributed at an average intermarker distance of 8.6 cM and application of the David Duffy genetic map with interpolated genetic map positions, revealed allele-sharing log of the odds (LOD) scores /H11022 at 9 locations: 1p34, 5q13, 7q35, 9q22, 9q34, 13q32, 14q32, 18p11, and 20q13 for all 3 disease models (Table 3 and Figure 2). The highest allele-sharing LOD scores were obtained on chromosomes 5q.
TABLE 3. Summary of Multipoint Allele-Sharing LOD Scores

<table>
<thead>
<tr>
<th>Cytogenetic Position</th>
<th>Model</th>
<th>GWS</th>
<th>Fine Mapping</th>
<th>Follow-Up</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LOD Score</td>
<td>56 fam.</td>
<td>53 fam.</td>
<td>109 fam.</td>
</tr>
<tr>
<td>1p34</td>
<td>I</td>
<td>2.08</td>
<td>1.32 &lt; 0.00</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>2.10</td>
<td>1.27 &lt; 0.00</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>1.25</td>
<td>1.35 &lt; 0.00</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>5q13</td>
<td>I</td>
<td>1.22</td>
<td>1.67 0.12</td>
<td>1.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>2.03</td>
<td>2.19 0.10</td>
<td>1.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>1.82</td>
<td>1.87 &lt; 0.00</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>7q35</td>
<td>I</td>
<td>1.20</td>
<td>1.05 &lt; 0.00</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>9q22</td>
<td>I</td>
<td>1.41</td>
<td>1.56 &lt; 0.00</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>1.43</td>
<td>1.45 &lt; 0.00</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>9q34</td>
<td>I</td>
<td>1.50</td>
<td>1.45 &lt; 0.00</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>1.23</td>
<td>1.09 &lt; 0.00</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>13q32</td>
<td>I</td>
<td>1.63</td>
<td>1.36 0.33</td>
<td>1.59</td>
<td></td>
</tr>
<tr>
<td>14q32</td>
<td>II</td>
<td>1.25</td>
<td>0.67 0.17</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>18p11</td>
<td>I</td>
<td>1.63</td>
<td>1.45 0.00</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>1.47</td>
<td>1.27 0.06</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>2.14</td>
<td>1.73 0.16</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>20q13</td>
<td>II</td>
<td>1.44</td>
<td>0.98 &lt; 0.00</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>1.64</td>
<td>1.14 &lt; 0.00</td>
<td>0.45</td>
<td></td>
</tr>
</tbody>
</table>

LOD scores >1.2 in any of the models in the genome wide linkage scan are displayed as well as fine mapping and follow-up LOD scores.

(Previously reported), 1p (LOD=2.10 at marker D1S255; model II), and 18p (LOD=2.14 at marker D18S59; model III). Simulation studies revealed that an allele-sharing LOD score value of 2.14 is expected to occur by chance once in every 5 genome-wide scans. These simulation studies have not taken into consideration that 3 different disease models were investigated. Such corrections would be difficult, however, because the models are not independent.

Fine Mapping

Next, the 9 chromosomal regions with allele-sharing LOD scores >1.2 were investigated further by genotyping 48 additional microsatellite markers, resulting in an average intermarker distance of 3.8 cM in these regions. Fine mapping resulted in increased allele-sharing LOD scores for chromosome 5q (previously reported) and 9q22 (LOD=1.56 at marker D9S287; model I). All other loci decreased, but 4 chromosomal regions still displayed an allele-sharing LOD score >1.2: 1p34 (LOD=1.27 at marker D1S233; model II), 9q34 (LOD score=1.45 at marker D9S1818; model I), 13q32 (LOD score=1.36 at marker D13S265; model I), and 18p11 (LOD score=1.73 at marker D18S59; model III; Table 3 and Figure 3A). On chromosome 9q22 and 9q34, linkage was detected in models I and II but not in the ischemic stroke—only model (model III). The strongest evidence for linkage to the chromosome 18 region was found for model III and to the chromosome 13 region for model I. The marker yielding the maximum allele-sharing LOD score on chromosome 18, D18S59, is the most telomeric marker genotyped on chromosome 18 and therefore needs to be interpreted with caution.

Follow-Up and Combined Analysis

In a follow-up investigation, we collected 63 additional families from the same region with 246 individual samples. Thirty families were excluded, mainly owing to a lack of samples at the time of genotyping, but also owing to genetic, logical and clinical validation (Figure 1) problems. The remaining 33 families from the second recruitment were added to the 20 ungenotyped families from the first recruitment, forming a follow-up cohort consisting of 129 affected individuals (including 15 cases of SAH) from a total of 242 individuals.

These 242 individuals were genotyped with the 48 fine-mapping microsatellite markers and 50 of the original genome-wide markers in all 9 regions of interest. As illustrated in Table 3, the additional 53 families did not contribute significantly to the linkage peaks identified in the initial 56 families. In the combined analysis including all 109 families, the highest allele-sharing LOD scores were obtained on chromosomes 5q (LOD=1.71, D5S1982; model II; Table 3), 13q (LOD=1.59 at marker D13S265; model I), and 18p (LOD=1.53 at marker D18S59; model III; Figure 3B).

Discussion

This article describes 1 of the few genome-wide scans for common forms of stroke. The initial linkage findings were refined in 2 steps. In the first fine-mapping step, we detected moderate linkage on chromosomes 9q, 13q, and 18p in addition to the previously described region on chromosome 5q. After inclusion of additional families in a second step, linkage was still detected in all of these regions apart from chromosome 9q.

The families under investigation were selected from the relatively genetically homogeneous region of northern Sweden. This was reinforced by genealogical studies and exclusion of families in which the parents of the proband had not been born in the region. We also enriched for families with the potential to provide maximum information regarding identical-by-descent sharing among affected individuals. This was obtained by selecting families with multiple affected family members who were available for genotyping and also families in which both an affected sib pair and unaffected relatives were available for genotyping. The families added in the second step were selected with the same genealogical criteria and displayed a similar stroke risk factor profile as the initial genome-wide scan, but they did not provide as much information regarding identical-by-descent sharing within the families as the first family set. This lack of power and decrease in the proportion of multiplex families could explain the lack of linkage in the separate analysis of the second family set.

Linkage calculations were performed with 3 different disease models, ranging from a very broad to a narrower stroke phenotype. Because most reports consider SAH a different disease entity, wherein the disease process involves structural proteins, proteases, and protein inhibitors in the extracellular matrix, the phenotype in model I, including SAH, might be too broad. However, in 23 of the 109 described families, SAH was clustered with other stroke diagnoses, and in several cases, the same patient had been affected by >1 type of stroke. This may indicate common
underlying susceptibility factors for all stroke types, similar to the way that the ALOX5AP gene has been reported to contribute to the susceptibility for both MI and ischemic stroke.\textsuperscript{7} The most significant finding for chromosome 13 was obtained by using a broad phenotype. One interesting gene in this region is guanylate cyclase 1, soluble, β-2, which plays an important role in the cardiovascular system as a receptor for nitric oxide. On chromosome 18, the ischemic stroke–only model provided the highest allele-sharing LOD score. In this region, adenylate cyclase–activating polypeptide 1, a regulator of vasodilation, has been suggested as a candidate gene for essential hypertension in a sib-pair study.\textsuperscript{18}

The use of families from relatively homogeneous populations to detect linkage in complex diseases has proven to be
a successful and efficient approach in genetic studies.\textsuperscript{5,7} Recruitment and study of stroke families are difficult, owing to both the disabling and fatal nature of the disease and the advanced age of the patients. This article also illustrates some of the difficulties in recruiting a sufficient number of stroke families for linkage studies, even though a large percentage of the population has a very positive attitude toward participating in genetic research.\textsuperscript{19} The recruitment and linkage study was facilitated by a high response rate to the questionnaires in this project, in combination with high-quality clinical registers with validated stroke diagnoses. Unfortunately, subphenotyping of ischemic subjects according to TOAST criteria\textsuperscript{13} could be obtained with certainty in only 41\% of cases.

Replication of previously published linkage data strengthens our current findings, but until now, only Gretarsdottir et al\textsuperscript{5} have reported a whole genome-wide scan for stroke. When comparing these 2 genome-wide scans, we noted that the only striking similarity is the chromosome 5 linkage finding, previously published.\textsuperscript{10} Helgadottir et al\textsuperscript{7} published a genome-wide scan for MI and detected linkage to chromosome 13; they also identified the \textit{ALOX5AP} gene, which showed an association with both MI and ischemic stroke. Our genome-wide scan detected a modest allele-sharing LOD score of 0.81 at marker D13S171 in the ischemic stroke--only model, in close proximity to the peak marker D13S289 from the MI study by Helgadottir et al.\textsuperscript{7} This might indicate a contribution of \textit{ALOX5AP} to the susceptibility for ischemic stroke in our population and will be investigated further.

In summary, we present a study design in a restricted number of well-characterized families with an accumulation of stroke cases from a relatively homogeneous population, with the possibility to detect linkage peaks of the same magnitude as large sib-pair studies in heterogeneous populations.\textsuperscript{20} Even though the applicability of linkage findings detected in homogeneous populations to other populations is uncertain, this method still appears to be an important approach to understand the genetics of complex diseases. Candidate genes in pathways identified in homogeneous populations can and should be tested for association in other populations. Nonsignificant linkage results in initial genome-wide scans of homogeneous populations have previously been used for the successful identification of candidate genes for complex diseases, with a strong association to the disease in more heterogeneous populations.\textsuperscript{21–23} Our genetic analysis of stroke in families from northern Sweden did not identify any new major stroke loci. This indicates that multiple minor susceptibility loci in addition to the previously known locus on chromosome 5 could contribute to the disease.

\textbf{Acknowledgments}

We thank the physicians and the participating probands and their families whose contribution made this study possible.

\textbf{Sources of Funding}

This work was supported by grants from the Kempe foundation, the Swedish Research Council, the Foundation of Strategic Research, and the King Gustaf V and Queen Victoria’s foundation.

\textbf{Disclosures}

None.

\textbf{References}


Genome-Wide Linkage Scan of Common Stroke in Families From Northern Sweden
Sofie Nilsson-Ardnor, Tomas Janunger, Per-Gunnar Wiklund, Kurt Lackovic, Anna Karin Nilsson, Petter Lindgren, Stefan A. Escher, Birgitta Stegmayr, Kjell Asplund and Dan Holmberg

*Stroke*. 2007;38:34-40; originally published online November 22, 2006; doi: 10.1161/01.STR.0000251643.37454.16
*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2006 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/38/1/34

**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/