Systemic lupus erythematosus (SLE) is an autoimmune disease with diverse and variable clinical manifestations and unknown etiology. Epidemiological and animal studies indicate that environmental and genetic factors are involved in the development of the disease. Several candidate gene loci (including the human leucocyte antigen (HLA) region, FcY receptors, and complement components) have been implicated through association studies, and multiple susceptibility loci have been detected in inbred mouse models of SLE. Until now, six groups have published genomewide scans with SLE as a phenotype in different ethnic groups. Recently, linkage to chromosome 2q37 (logarithm of odds (LOD) 4.24) in a Swedish population resulted in the identification of a new susceptibility gene PDCD1 in a large multinational study by Prokunina et al. The SLE associated allele of this immunoreceptor gene alters a binding site for the runt related transcription factor 1 (RUNX1), which is found in an intrinsic enhancer.

Stratification of pedigrees based on clinical manifestations has been used in recent studies that involved genomewide scans. The aim was to achieve genetically and clinically homogeneous sets of families and to increase the power to detect susceptibility genes for different subphenotypes of SLE. Altogether, 17 regions have been linked significantly to SLE with model based and non-parametric approaches; 11 of these in stratified studies. In addition, several other regions with suggestive linkage have been identified, but only some of those loci have been implicated in more than one study. The data suggest that multiple genes are involved in conferring susceptibility to SLE.

In our study, we conducted a nationwide and genomewide scan for SLE susceptibility loci in Finnish families multiply affected by SLE. The extensive hospital registration system in Finland allowed us to identify and recruit approximately 85% of all patients with SLE who needed hospital based treatment. Among these patients, we identified 35 multiplex families suitable for linkage mapping. We genotyped DNA samples from 73 patients with SLE and 96 healthy relatives with polymorphic microsatellite markers and analysed them with non-parametric linkage analysis. We obtained suggestive evidence for linkage (that is, non-parametric linkage scores that exceeded the threshold of one random occurrence per genome scan) for three regions in a region on chromosome 1q4 with a previous suggestive mapping result, a region on chromosome 6q25–q27 previously linked to insulin dependent diabetes and rheumatoid arthritis, and a new locus on chromosome 5p. In addition, the HLA region on chromosome 6p had positive linkage that reached the suggestive threshold after a marker gap was filled. Our genome scan identified regions for linkage disequilibrium mapping in a larger cohort of Finnish patients with SLE and control participants.

**Key points**

- This study aimed to identify susceptibility loci for systemic lupus erythematosus (SLE) in Finnish multiplex families.
- A genomewide scan with 417 polymorphic microsatellite markers was conducted in 35 Finnish families multiply affected by SLE. Simulation with model based and non-parametric approaches was used to establish thresholds for suggestive and significant linkage.
- Suggestive linkage was observed in three regions: 5p, 6q25-q27, and 14q21-q23. In addition, chromosome 6p (locus for human leucocyte antigen) exceeded the threshold after a marker gap was filled. In all four regions, non-parametric linkage scores increased as information provided by additional markers increased, but they did not reach the threshold for significant linkage.
- Linkage peaks in 6p and 14q add to evidence from previous studies that susceptibility genes for SLE are present in these genomic regions.

**PARTICIPANTS AND METHODS**

**Patient recruitment**

The patients were recruited as described previously. The study was approved by the local ethical committees according to applicable regulations. Patients treated at Kuopio and Helsinki University hospitals, patients registered by the Lupus Foundation of Finland, and patients who answered an advertisement published in patient bulletins were recruited in 1995. We identified patients who had been treated in the two university hospitals during 1992–1995 from the corresponding hospital registries and contacted all patients personally or by mail. A second recruitment phase started in autumn 1996. We contacted doctors (mainly rheumatologists) from 17 central hospitals in Finland and four other major hospitals that treated patients with SLE. A letter was sent to all patients with the clinical diagnosis of SLE who had been treated in these hospitals during 1993–1996. All patients were asked whether they had relatives or family members diagnosed with SLE or a connective tissue disease similar to SLE. We asked patients with a positive family history to participate and obtained informed consent. According to the prevalence of SLE in Finland, we succeeded in contacting roughly 1200 out of 1500 available patients personally (by phone or by mail) during the two phases of recruitment; this accounted for about 85% of all Finnish patients with SLE who needed...
hospital based treatment. Most of the families had only a single patient with SLE and thus could not be used for a linkage study.

**Family selection**

Fifty-three families who were affected multiply by SLE and thus were informative for genetic linkage were identified among all contacted patients with SLE. All patients from these families were interviewed by the same doctor (HJ) either personally or by telephone, and their case records from the hospitals were reviewed. All patients met the American College of Rheumatology criteria for the diagnosis of SLE. All available parents were recruited for phasing of chromosomes. If parents were not available, an unaffected sibling was sampled to allow reconstruction of parental genotypes. Of the identified families, 35 were informative for linkage mapping and hence were included in this mapping study.

**Polymerase chain reaction (PCR) and genotyping**

We amplified genomic DNA (20 ng) prepared from blood samples in 5 μl PCR assays with 0.33 μM fluorescently labelled primers and 0.2 units of AmpliTaq Gold DNA Polymerase (Applied Biosystems, Foster City, CA, USA). Before the PCR assays, we distributed the DNA to 384 well microtitre plates with a Hydra-96 microdispenser (Robbins Scientific, Sunnyvale, CA, USA), after which the solution of DNA was dried. We assembled the PCR with a Tecan Genesis 150/8 robotic sample processor (DNA Microarray Core Facility, Miami, FL, USA) and ran it in a Dual 384-Well GeneAmp PCR System 9700 (PE Biosystems, Foster City, CA, USA). After an initial heating step of 12 minutes at 95°C, we ran 30 cycles of PCR (10 cycles of 30 seconds at 95°C, 30 seconds at 55°C, and 30 cycles of 20 seconds at 72°C). We multiplexed the amplified fragments from PCR (13 markers in a pool on average) and separated them with capillary array electrophoresis (Molecular Dynamics MegaBACE 1000; Global Medical Instrumentation, Albertville, Minnesota, MN, USA). We further processed and analysed the capillary runs with Genetic Profiler software (version 1), which performs automatic sizing and allele calling. We reviewed all electropherograms manually before they were analysed. We used microsatellite markers from the linkage mapping set MD-10 (PE Biosystems, Foster City, CA, USA) in the primary genome scan. Of the 400 markers in the linkage mapping set, we were able to retrieve data for 388 markers. The average intermarker distance was 9.73 cM, and the data contained six gaps of 20–26 cM (the marker map is available at http://www.genome.helsinki.fi). Mendelian inheritance was confirmed in all families with Pedmanager software (Center for Genome Research, Cambridge, MA, USA) and PedCheck software (University of Pittsburgh, Pittsburgh, PA, USA). After we analysed the genome scan, we genotyped 29 additional markers to improve the inheritance information captured in regions of chromosomes 3p, 5p, 6p, 6q, 8, and 14q. We added the data to the genome scan data and reanalysed it.

**Data analysis and simulations**

As the mode of inheritance of SLE is not known, we carried out linkage analysis with non-parametric analysis with Genehunter software (version 2.0 for chromosomes 1–22 and version 1.3 for the X chromosome) (Fred Hutchinson Cancer Research Center, Seattle, WA, USA). Genehunter multipoint analyses the degree of identity by descent sharing among all affected pedigree members at each location of the genome. To get an estimate of the global p value, we performed allele dropping simulations under the null hypothesis of no linkage. The simulations consisted of 200 iterations, as described previously. In each iteration, we drew founder alleles from the estimated allele frequency distributions and introduced missing genotypes to exactly the same positions as in the real data. As a result, we obtained an empirical distribution for the non-parametric linkage scores, which corresponded to suggestive and significant linkage observed under the null hypothesis in the entire genome scan.

**RESULTS**

**Demographic, clinical, and laboratory features of families with SLE**

Seventy-three patients with SLE and 96 healthy relatives from 35 multiply affected families were identified. Their clinical features have been described previously. In 32/35 families, two family members were affected. The most common combination was sister and sister (16 families); this was followed by aunt and niece (eight families). Three families had three patients with SLE. In two of these families, the affected family members were first degree relatives; in the third family, a second degree relative (a cousin) and first degree relatives (mother and son) were involved. Seven (9.6%) of the 73 familial cases were men, which corresponded well with the sex distribution of SLE in the general population.

**Genome scan**

We used the marker set MD-10 to capture 50–60% of the maximum inheritance information throughout the genome. Loss of information was mainly caused by the pedigree structure (that is, missing parents). Table 1 shows the highest non-parametric linkage scores obtained and the corresponding information content (chromosome 3, chromosome 5, and chromosome 6).

We used simulations to establish the empirical thresholds for suggestive and significant linkage (that is, non-parametric linkage scores reached at random once per one or

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Marker</th>
<th>Non-parametric linkage score</th>
<th>p value</th>
<th>Information content</th>
</tr>
</thead>
<tbody>
<tr>
<td>3q</td>
<td>D3S1278</td>
<td>1.78</td>
<td>0.03</td>
<td>0.64</td>
</tr>
<tr>
<td>5p</td>
<td>D5S418</td>
<td>1.74</td>
<td>0.04</td>
<td>0.61</td>
</tr>
<tr>
<td>HLA region</td>
<td>D6S276</td>
<td>1.22</td>
<td>0.11</td>
<td>0.52</td>
</tr>
<tr>
<td>6q</td>
<td>D6S441</td>
<td>2.35</td>
<td>0.01</td>
<td>0.57</td>
</tr>
<tr>
<td>8</td>
<td>D8S260</td>
<td>0.62</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>14q</td>
<td>D14S288</td>
<td>1.51</td>
<td>0.07</td>
<td>0.55</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Marker</th>
<th>Non-parametric linkage score</th>
<th>p value</th>
<th>Information content</th>
</tr>
</thead>
<tbody>
<tr>
<td>3q</td>
<td>D3S1278</td>
<td>1.55</td>
<td>0.06</td>
<td>0.71</td>
</tr>
<tr>
<td>5p</td>
<td>D5S418</td>
<td>2.03</td>
<td>0.02</td>
<td>0.63</td>
</tr>
<tr>
<td>HLA region</td>
<td>D6S273</td>
<td>2.17</td>
<td>0.02</td>
<td>0.78</td>
</tr>
<tr>
<td>6q</td>
<td>D6S960</td>
<td>2.47</td>
<td>0.008</td>
<td>0.67</td>
</tr>
<tr>
<td>8</td>
<td>D8S260</td>
<td>0.48</td>
<td>0.31</td>
<td>0.46</td>
</tr>
<tr>
<td>14q</td>
<td>D14S587</td>
<td>2.20</td>
<td>0.02</td>
<td>0.62</td>
</tr>
</tbody>
</table>
20 genome scans, respectively\(^1\)). The best observed non-parametric linkage score under the null hypothesis of no linkage showed that non-parametric linkage = 1.7 is reached once per genome scan (suggestive linkage) and that the global p value = 0.05 corresponds to non-parametric linkage = 3 (significant linkage). Three loci exceeded the threshold for suggestive linkage; none was significant (table 1). To verify the highest non-parametric linkage scores and to cover gaps in the map, we genotyped more markers at regions with non-parametric scores >1.7 (chromosome 3q, chromosome 5p, and chromosome 6q) and regions with gaps >20 cM if the corresponding non-parametric linkage score was >1 (chromosome 6p (that is the HLA region), chromosome 8, and chromosome 14). In all regions with suggestive

---

**Figure 1**  Non-parametric multipoint linkage analyses in 35 families multiply affected by SLE. All chromosomes are shown. Chromosomes with peaks with non-parametric linkage $\geq 2.0$ shown in large panels (top row). Non-parametric linkage scores plotted with solid line (scale on left) and information content with dotted line (scale to right). In total, 417 markers were genotyped; average intermarker distance was 10 cM.
linkage, non-parametric linkage scores increased with additional markers (table 2). The most striking change in non-parametric linkage scores was seen in the HLA region, in which the non-parametric linkage score increased from 1.2 to 2.1. Figure 1 shows the overall non-parametric linkage scores, including all genotyped markers.

**DISCUSSION**

This was the first nationwide study in Finland to map genetic factors that confer susceptibility to SLE. We recruited >80% of patients with SLE in Finland, which has a population of 5 million. Our results suggest that no single major gene contributes to the risk of the disease in Finland, despite its population structure. Our genome scan showed three loci that exceeded the threshold of 1.7; this corresponded to the definition of suggestive linkage genomewide. Even after we increased the amount of information with additional markers, no locus reached the threshold for significant linkage. Ideally, the results should be verified in an independent dataset. Unfortunately, we cannot increase the dataset in Finland, because we already have sampled almost all available patients. These results can, however, be used to guide further association mapping with very high density marker maps, as suggested previously.

We identified two previously reported regions (the HLA region and chromosome 14q21–q23) and two novel regions on chromosome 5p and chromosome 6q as possibly linked to SLE. Linkage to HLA was in accordance with previous studies. Association between SLE and the class II and III genes in the human leucocyte antigen complex was seen in studies in the 1970s. Studies of patients of European descent consistently show an association between SLE and especially HLA DR2 and DR3. The HLA region also has been implicated in more recent linkage mapping studies. Gaffney and colleagues reported strong linkage to the HLA region; this was supported by results from Shai and colleagues and Lindqvist and colleagues in patients of Swedish pedigree.

Chromosome 14q21–q23 also has been implicated repeatedly in previous linkage studies, although no study alone has reported a significant linkage result. Evidence for linkage to chromosome 14q has been reported in three independent data sets: Gaffney and colleagues (LOD 2.81, p < 0.00016), Shai and colleagues (non-parametric linkage 2.02, p = 0.02) and Lindqvist and colleagues in patients of Swedish pedigree (LOD 1.15) but not in the combined Minnesota cohort or the Oklahoma dataset. The data support locus heterogeneity between the different datasets. Our study adds to the evidence that a susceptibility gene for SLE exists in chromosome 14q.

The biological relevance of the observed excess sharing of alleles in the novel regions on chromosomes 5p and 6q is more difficult to evaluate. Interestingly, the same region on chromosome 6q25–q27 has been implicated in other autoimmune diseases. Several studies support a locus that confers susceptibility to insulin dependent diabetes in this chromosomal region (IDDMS). Supportive evidence for the existence of a common autoimmune susceptibility locus in the region comes from a study by Myerscough and colleagues. They report evidence for linkage disequilibrium between rheumatoid arthritis and the markers D6S311 and D6S440. As the highest observed Z score in our study resided in the same region, chromosome 6q25–q27 may harbour a gene that influences the predisposition to autoimmunity.

Recently, Rioux and colleagues showed the power of linkage disequilibrium analysis in complex disease mapping. By genotyping polymorphic markers at 0.35 cM intervals in a region initially identified by linkage mapping, they successfully identified linkage disequilibrium and an ancestral haplotype that spanned 250 kb in patients with inflammatory bowel disease in an outbred Canadian population. A similar approach should be feasible among Finnish patients and controls in an expanded study. In a population characterised by local founder effects, linkage disequilibrium mapping in selected geographical areas might be a useful strategy for identifying relatively rare susceptibility alleles. We intend to use the power of linkage disequilibrium to identify common ancestral chromosomes among a larger cohort of patients with SLE and controls by high resolution mapping.

**ACKNOWLEDGEMENTS**

We thank all patients and family members who participated in this study and the many doctors who referred families and verified the diagnoses. We thank Ms Riitta Lehtinen, Ms Sirkka Ruohomaki, Ms Päivi Hantula, and Ms Anitta Hottinen for excellent technical assistance and Ms Anne Lehto for invaluable help with data management. We are indebted to Dr Tanja Laitinen for valuable discussions and comments.

**REFERENCES**

Linkage mapping of systemic lupus erythematosus (SLE) in Finnish families multiply affected by SLE

S Koskenmies, P Lahermo, H Julkunen, V Ollikainen, J Kere and E Widén

J Med Genet 2004 41: e2-e5
doi: 10.1136/jmg.2003.009977