A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families


*University of Minnesota Medical School, 312 Church Street SE, Minneapolis, MN 55455; and †Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157

Communicated by Arno G. Motulsky, University of Washington, Seattle, WA, October 16, 1998 (received for review September 14, 1998)

ABSTRACT Systemic lupus erythematosus (SLE) is an autoimmune multisystem inflammatory disease characterized by the production of pathogenic autoantibodies. Previous genetic studies have suggested associations with HLA Class II alleles, complement gene deficiencies, and Fe receptor polymorphisms; however, it is likely that other genes contribute to SLE susceptibility and pathogenesis. Here, we report the results of a genome-wide microsatellite marker screen in 105 SLE sib-pair families. By using multipoint nonparametric methods, the strongest evidence for linkage was found near the HLA locus (D6S257, log of odds (lod) = 3.90, P = 0.000011) and at three additional regions: 16q13 (D16S415, lod = 3.64, P = 0.000022), 14q21–23 (D14S276, lod = 2.81, P = 0.00016), and 2p12 (D20S186, lod = 2.62, P = 0.00025). Another nine regions (1p36, 1p13, 1q42, 2p15, 3q14–15, 3qter–q11, 4q28, 11p15, and 15q26) were identified with lod scores >1.00. These data support the hypothesis that multiple genes, including one in the HLA region, influence susceptibility to human SLE.

RESULTS AND DISCUSSION

Families were recruited by advertising for “Sisters with Lupus” (11), and, as a result, our sample was highly enriched for female patients (female: male 219:1), compared with the population estimate for lupus of ~90% female. Selected clinical and demographic data of the SLE patients in this study are shown in Table 1.

The sample studied in this report comprises the first 105 SLE sib-pair families collected (102 families with 2 SLE sibs, 2 families with 3 SLE sibs, and 1 family with 4 SLE sibs). All available parents (including three affected) were collected, and, in the absence of parents, an unaffected sibling was sampled to assist in reconstruction of parental genotypes. In three families an additional affected-first degree relative was

MATERIALS AND METHODS

Families. The recruitment of families for this study has been described (11). All patients met the 1997 American College of Rheumatology revised criteria for the diagnosis of SLE (7, 8).

Samples and Genotyping. Genomic DNA was isolated from peripheral blood cells by using standard conditions. Genotyping was performed by using an Applied Biosystems fluorescently labeled human linkage mapping set (version 1.0, panels 1–18; version 2.0, panels 19–27). PCR (32 cycles) was performed on an ABI 877 Catalyst robotic workstation [5 μl reactions; 8 ng of genomic DNA, 2.5 mM MgCl2, 0.2 mM dNTPs (Pharmacia), 1.65 pmol of 5’ and 3’ primers, 0.2 units of AmpliTaq Gold DNA polymerase (Perkin–Elmer) in 1× PCR Buffer II (Perkin–Elmer)].

RESULTS AND DISCUSSION

Families were recruited by advertising for “Sisters with Lupus” (11), and, as a result, our sample was highly enriched for female patients (female: male 219:1), compared with the population estimate for lupus of ~90% female. Selected clinical and demographic data of the SLE patients in this study are shown in Table 1.

The sample studied in this report comprises the first 105 SLE sib-pair families collected (102 families with 2 SLE sibs, 2 families with 3 SLE sibs, and 1 family with 4 SLE sibs). All available parents (including three affected) were collected, and, in the absence of parents, an unaffected sibling was sampled to assist in reconstruction of parental genotypes. In three families an additional affected-first degree relative was

MATERIALS AND METHODS

Families. The recruitment of families for this study has been described (11). All patients met the 1997 American College of Rheumatology revised criteria for the diagnosis of SLE (7, 8).

Samples and Genotyping. Genomic DNA was isolated from peripheral blood cells by using standard conditions. Genotyping was performed by using an Applied Biosystems fluorescently labeled human linkage mapping set (version 1.0, panels 1–18; version 2.0, panels 19–27). PCR (32 cycles) was performed on an ABI 877 Catalyst robotic workstation [5 μl reactions; 8 ng of genomic DNA, 2.5 mM MgCl2, 0.2 mM dNTPs (Pharmacia), 1.65 pmol of 5’ and 3’ primers, 0.2 units of AmpliTaq Gold DNA polymerase (Perkin–Elmer) in 1× PCR Buffer II (Perkin–Elmer)].

Pooled amplification products were electrophoresed through 5% polyacrylamide gels (FMC) for 2 hr at 3,000 V by using an ABI 377 DNA Sequencer. Semi-automated fragment sizing was performed by using GENESCAN 2.1 software (ABI) followed by allele calling with GENOTYPER 2.0 software (ABI). Each genotype was reviewed manually by two members of the research team to confirm the accuracy of allele calling. Poor performing markers (a total of 17) were excluded from the analysis. The overall data drop-out rate for the ~127,500 genotypes analyzed was <1.5%.

Data Analysis. Nonparametric multipoint analysis was performed with GENEHUNTER PLUS (12), a modified version of GENEHUNTER (13), by using the “all” statistic. Allele frequencies for the parameter file were generated from the unaffected parental genotypes for the cohort. Marker map positions were obtained from the sex-averaged maps compiled by J. Weber (Marshfield Clinic, Marshfield, WI) (www.marshmed.org/genetics/).

ABBREVIATIONS: sib, sibling; SLE, systemic lupus erythematosus; lod, logarithm of odds; MHC, major histocompatibility complex; cM, centimorgan.

†P.M.G. and G.M.K. contributed equally to this work.

‡Current address: University College, Erics Street, Dublin 7, Ireland.

To whom reprint requests should be addressed at: University of Minnesota Medical School, Box 334 Mayo, 312 Church Street SE, Minneapolis, MN 55455, e-mail: behre001@maroon.tc.umn.edu.
``potentially interesting,'' and applied the criteria of Lander 4.1, lod
mosome. We considered any region with a Zlr
calculated lod scores for each marker plotted for each chro-
Zlr, the test statistic generated by GENEHUNTER PLUS , and the
yses (data not shown).
these multipoint results were confirmed by single point anal-
tipoint linkage analysis was performed on the collected marker
itance was confirmed in all families, and nonparametric mul-
personalities included in the analysis. The final study cohort totaled 220 SLE

Table 1. Clinical and demographic features of 220 SLE patients

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis, ±SD</td>
<td>31 ± 11 years</td>
</tr>
<tr>
<td>Duration of disease, ±SD</td>
<td>12 ± 7 years</td>
</tr>
<tr>
<td>Sex, female:male</td>
<td>219:1</td>
</tr>
<tr>
<td>Ethnicity, %</td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>80</td>
</tr>
<tr>
<td>Hispanic</td>
<td>8</td>
</tr>
<tr>
<td>African-American</td>
<td>5</td>
</tr>
<tr>
<td>Asian</td>
<td>3</td>
</tr>
<tr>
<td>Mixed heritage</td>
<td>4</td>
</tr>
<tr>
<td>Laboratory clinical features*, %</td>
<td></td>
</tr>
<tr>
<td>ANA-positive</td>
<td>98</td>
</tr>
<tr>
<td>Anti-dsDNA-positive</td>
<td>46</td>
</tr>
<tr>
<td>Arthritis</td>
<td>85</td>
</tr>
<tr>
<td>Skin involvement</td>
<td>91</td>
</tr>
<tr>
<td>Pleuritis</td>
<td>53</td>
</tr>
<tr>
<td>Hematologic</td>
<td>47</td>
</tr>
<tr>
<td>Renal disease</td>
<td>30</td>
</tr>
<tr>
<td>CNS lupus</td>
<td>22</td>
</tr>
<tr>
<td>Pericarditis</td>
<td>18</td>
</tr>
<tr>
<td>More than one miscarriage</td>
<td>13</td>
</tr>
<tr>
<td>Medication history, %</td>
<td></td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>77</td>
</tr>
<tr>
<td>Antimalerials</td>
<td>65</td>
</tr>
<tr>
<td>Cytotoxic drugs</td>
<td>28</td>
</tr>
<tr>
<td>Intravenous steroids</td>
<td>19</td>
</tr>
</tbody>
</table>

*a Data represents the percentage of SLE patients having the indicated laboratory/clinical features and medication history at any time during the course of their disease.
*b Four individuals negative for ANA tested positive for anti-dsDNA antibodies and otherwise fulfilled criteria. These data are comparable with those described in other large series of SLE patients (31). ANA, antinuclear antibodies; dsDNA, double-stranded DNA; CNS, central nervous system.

included in the analysis. The final study cohort totaled 220 SLE patients and 155 unaffected parents or sibs. The racial com-
position of the 105 families studied was as follows: 84 Caucasian, 8 Hispanic, 6 African-American, 3 Asian, and 4 of mixed heritage.

Family members were genotyped with 341 highly polymorp-
phic markers across the 22 autosomes, at an average inter-
marker distance of 9.7 centimorgans (cM). Mendelian inher-
itance was confirmed in all families, and nonparametric mul-
tipoint linkage analysis was performed on the collected marker
data by using GENEHUNTER (13), with extensions to include
 calculation of appropriate logarithm of odds (lod) scores and
 support intervals (12).

Results of the multipoint analysis are shown in Fig. 1, with
Zp, the test statistic generated by GENEHUNTER PLUS, and the
 calculated lod scores for each marker plotted for each chro-
mosome. We considered any region with a Zp ≥ 2.0 as
“potentially interesting,” and applied the criteria of Lander
and Kruglyak (14) to further define regions of significant (Z ≥
4.1, lod ≥ 3.6, P = 0.000002) or suggestive (Z ≥ 3.2, lod ≥ 2.2,
P ≤ 0.0007) linkage. Overall, 25 of the 341 markers tested (7%)
gave Zp scores >2.1 and lod scores >1.0. Strikingly, 16 of the
25 positive markers clustered into 4 distinct genomic intervals
(Fig. 1 and Table 2). Two of these intervals contained a marker
that met criteria for significant linkage: 6p11-p21, mapping near the HLA region (D6S257, Zp = 4.24, lod = 3.90, P =
0.0000011), and 16q13 (D16S415, Zp = 4.09, lod = 3.64, P =
0.0000022). Two other regions fulfilled criteria for suggestive
linkage: 14q21-q23 (D14S276, Zp = 3.60, lod = 2.81, P =
0.000016) and 20p12 (D20S186, Zp = 3.48, lod = 2.62, P =
0.000025). Nine additional chromosomal regions were identi-
fied by single markers with Zp scores >2.1 and accompanying
lod scores ≥1.0 (Table 2). In all of the most suggestive areas,
these multipoint results were confirmed by single point analy-
yses (data not shown).

Because of sample size considerations, we stratified the data by ethnic group only for the 84 Caucasian families in the
overall sample of 105 families. The results of this analysis are
shown in Table 3. Lod scores dropped in 10 of the top 13
identified potential susceptibility intervals when non-
Caucasian families were eliminated from the analysis, sug-
gest ing that families of all ethnic groups contributed to the
evidence for genetic linkage in these regions. In contrast,
three intervals (4q28, 11p15, and 15q26) were characterized
by an improvement in lod scores when only the Caucasian
families were considered.

The strongest evidence for linkage was found at 6p11-p21,
with the best markers (D6S426 and D6S257) mapping just
centromeric to the HLA region (located at 6p21.3). There is
a long history of interest in the role of the major histocom-
patibility complex (MHC) in many autoimmune diseases,
including IDDM, familial psoriasis, multiple sclerosis, rheu-
matoid arthritis, and SLE. Genetic associations have been
shown for HLA-DR2 or HLA-DR3 alleles in SLE patients
from several ethnic groups (relative risks ranging from 2.0 to
3.4) (15). However, in large lupus-prone families, HLA
alleles do not necessarily segregate with disease (16). A
stronger association of HLA alleles (DO in particular) has
been demonstrated for specific autoantibody subsets in SLE
(17). Non-HLA genes within the MHC, notably the tumor
necrosis factor α gene and the complement components C2
and C4, also have been implicated in SLE (17). Further work
will be necessary to determine whether this 6p locus repre-
sents a polymorphic HLA locus or a linked gene within or
near HLA.

A recent comparison of the linkage results from 23
independent genome screens in various human and experi-
mental animal autoimmune or inflammatory diseases iden-
tified 18 “clusters” of non-MHC candidate human autoim-
mune loci (18). Of interest, two of the non-MHC intervals
identified in this screen of SLE families (16q13 and 11p15)
are located within clusters implicated in other human auto-
immune diseases. The 16q cluster is a large interval of ~35
cM that includes potential susceptibility loci for Crohn’s
disease (19), Blau syndrome (20), psoriasis (21), IDDM (22),
and asthma (23). Cluster 11p is a narrow interval (~2–3 cM)
identified in asthma (23), multiple sclerosis (24), and IDDM
(22). SLE is not known to associate strongly with any of
the diseases that define these clusters, but these data suggest
the possibility that there is a sharing of genetic predisposition to
multiple autoimmune diseases or, alternatively, that each of
these intervals may contain closely linked autoimmune
susceptibility genes.

The results of an independent genome scan of SLE families
performed by Moser et al. (25) at the Oklahoma Medical
Research Foundation support many of the findings of this
screen. Of the top 13 intervals identified in this sib-pair family
screen, 4 (20p12, 1p13, 1q42, and 4q28 with lod scores ranging
from 2.62 to 1.46) also were identified in the Oklahoma
Medical Research Foundation screen of mostly larger families
multiplex for SLE. The 20p12 region shows suggestive evi-
dence for linkage in this study (D20S186, Zp = 3.48, P =
0.000025, lod = 2.62) and was also one of the strongest loci
identified in the Oklahoma screen (25). This interval maps
centromeric to a recently identified susceptibility interval for
psoriasis (21). The locus in the 1q42 region, originally identi-
fied by Tsao et al. in a candidate approach in 52 SLE sib-pairs
(26), now has been identified in three independent SLE
screens [refs. 25 and 26 and this study (D1S235, Zp = 2.64, P =
0.0041, lod = 1.51)]. Despite these similarities, several of the
strong intervals identified in this screen (6p, 1q, and 16q) did
not show evidence for linkage in the Oklahoma Medical
Research Foundation study. Also, in our primarily Caucasian
sample, we find little evidence for linkage at 1q22–23. This
region showed evidence for significant linkage in the Okla-
homa Medical Research Foundation study and harbors the various Fc receptors, which are associated with renal disease in African-American SLE patients (27). These differences in the genome screen results may reflect "false positive" evidence...
for linkage at some loci (14) or may result from ethnic and/or genetic heterogeneity in the samples.

In summary, using an affected sib-pair family approach, we have identified four potential SLE loci with lod scores >2.6 and an additional nine loci with lod scores ranging from 1.00 to 1.68. The four most interesting regions (6p11-p21-HLA, 14q21–23, 16q13, and 20p12) show surprisingly strong evidence for linkage given the sample size of 114 affected sib pairs and a total of 127 affected relative pairs. Two of the non-MHC intervals identified in this screen (16q13 and 11p15) have been implicated in other human autoimmune diseases. These data support results from genetic studies in lupus-prone mouse strains suggesting that multiple genes are responsible for conferring susceptibility to SLE (28–30). It will now be important to confirm these findings in additional cohorts of sib-pair families and to initiate efforts to further narrow these intervals in preparation for gene identification.

We are grateful to all of the patients and their family members who participated in this study and the many physicians who referred families and verified diagnoses. Thanks to J. Harley and K. Moser for sharing data before publication. Supported by grants from the Minnesota chapters of the Arthritis Foundation and the Lupus Foundation of America, the Texas Gulf Coast Chapter of the Lupus Foundation of America, and National Institutes of Health Grants R01 AR 43271, SLE SCOR P50 AR45231, and KO8-AR01986. P.M.G. is supported by a Physician-Scientist Award from the University of Minnesota Department of Medicine.

Table 2. Summary of potential SLE susceptibility loci

| ZLR range | Interval | ZLR | p value | lod score | Positive markers | Map position, cM | IC
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;4.0</td>
<td>6p11-p21</td>
<td>3.69</td>
<td>0.00011</td>
<td>2.96</td>
<td>D6S426</td>
<td>60.44</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>16q13</td>
<td>3.81</td>
<td>0.000069</td>
<td>3.15</td>
<td>D16S3136</td>
<td>62.11</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>14q21-23</td>
<td>2.94</td>
<td>0.0016</td>
<td>1.87</td>
<td>D14S288</td>
<td>47.51</td>
<td>0.70</td>
</tr>
<tr>
<td>3.0–4.0</td>
<td>16q13</td>
<td>3.60</td>
<td>0.00016</td>
<td>2.81</td>
<td>D14S276</td>
<td>56.36</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>14q21-23</td>
<td>2.80</td>
<td>0.0025</td>
<td>1.71</td>
<td>D14S63</td>
<td>69.18</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>2.19</td>
<td>0.014</td>
<td>1.04</td>
<td>D6S462</td>
<td>99.01</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>2.0–3.0</td>
<td>2p15</td>
<td>2.78</td>
<td>0.0027</td>
<td>1.68</td>
<td>D2S337</td>
<td>80.69</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>1p13</td>
<td>2.66</td>
<td>0.0039</td>
<td>1.53</td>
<td>D1S252</td>
<td>150.27</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>1q42</td>
<td>2.64</td>
<td>0.0041</td>
<td>1.51</td>
<td>D1S235</td>
<td>254.64</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>4q28</td>
<td>2.60</td>
<td>0.0047</td>
<td>1.46</td>
<td>D4S424</td>
<td>144.46</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>3cent-q11</td>
<td>2.39</td>
<td>0.0084</td>
<td>1.24</td>
<td>D3S1271</td>
<td>117.76</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>11p15</td>
<td>2.34</td>
<td>0.0096</td>
<td>1.19</td>
<td>D11S922</td>
<td>2.11</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>2q21–33</td>
<td>2.27</td>
<td>0.012</td>
<td>1.12</td>
<td>D2S151</td>
<td>152.04</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>15q26</td>
<td>2.22</td>
<td>0.013</td>
<td>1.07</td>
<td>D15S127</td>
<td>86.81</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>1p36</td>
<td>2.15</td>
<td>0.016</td>
<td>1.00</td>
<td>D1S234</td>
<td>55.10</td>
<td>0.74</td>
</tr>
</tbody>
</table>

*Chromosome locations were determined from the maps available at www.gdb.org and http://cedar.genetics.soton.ac.uk/pub/.
†lod = $Z^2_{LR}/2\ln10$ (12).
‡Map positions obtained from the sex-averaged maps compiled by J. Weber (www.marshmed.org/genetics/).
§Information content (13) for the marker designated in each interval.

for linkage at some loci (14) or may result from ethnic and/or genetic heterogeneity in the samples.

In summary, using an affected sib-pair family approach, we have identified four potential SLE loci with lod scores >2.6 and an additional nine loci with lod scores ranging from 1.00 to 1.68. The four most interesting regions (6p11-p21-HLA, 14q21–23, 16q13, and 20p12) show surprisingly strong evidence for linkage given the sample size of 114 affected sib pairs and a total of 127 affected relative pairs. Two of the non-MHC intervals identified in this screen (16q13 and 11p15) have been implicated in other human autoimmune diseases. These data support results from genetic studies in lupus-prone mouse strains suggesting that multiple genes are responsible for conferring susceptibility to SLE (28–30). It will now be important to confirm these findings in additional cohorts of sib-pair families and to initiate efforts to further narrow these intervals in preparation for gene identification.

We are grateful to all of the patients and their family members who participated in this study and the many physicians who referred families and verified diagnoses. Thanks to J. Harley and K. Moser for sharing data before publication. Supported by grants from the Minnesota chapters of the Arthritis Foundation and the Lupus Foundation of America, the Texas Gulf Coast Chapter of the Lupus Foundation of America, and National Institutes of Health Grants R01 AR/ AI43271, SLE SCOR P50 AR45231, and KO8-AR01986. P.M.G. is supported by a Physician-Scientist Award from the University of Minnesota Department of Medicine.

Table 3. Comparison of lod scores obtained for Caucasian SLE sib-pair families versus entire family cohort

<table>
<thead>
<tr>
<th>Loci</th>
<th>Caucasian families*</th>
<th>All families †</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lod score ‡</td>
<td>lod score ‡</td>
<td></td>
</tr>
<tr>
<td>6p11-p21</td>
<td>2.50</td>
<td>3.90</td>
<td>−1.40</td>
</tr>
<tr>
<td>16q13</td>
<td>3.36</td>
<td>3.64</td>
<td>−0.28</td>
</tr>
<tr>
<td>14q21-23</td>
<td>2.56</td>
<td>2.81</td>
<td>−0.25</td>
</tr>
<tr>
<td>20p12</td>
<td>2.09</td>
<td>2.62</td>
<td>−0.53</td>
</tr>
<tr>
<td>2p15</td>
<td>1.50</td>
<td>1.68</td>
<td>−0.18</td>
</tr>
<tr>
<td>1p13</td>
<td>1.27</td>
<td>1.55</td>
<td>−0.26</td>
</tr>
<tr>
<td>1q42</td>
<td>0.46</td>
<td>1.51</td>
<td>−1.05</td>
</tr>
<tr>
<td>4q28</td>
<td>2.00</td>
<td>1.46</td>
<td>+0.54</td>
</tr>
<tr>
<td>3cent-q11</td>
<td>0.87</td>
<td>1.24</td>
<td>−0.37</td>
</tr>
<tr>
<td>11p15</td>
<td>1.41</td>
<td>1.19</td>
<td>+0.22</td>
</tr>
<tr>
<td>2q21–33</td>
<td>0.61</td>
<td>1.12</td>
<td>−0.51</td>
</tr>
<tr>
<td>15q26</td>
<td>2.09</td>
<td>1.07</td>
<td>+1.02</td>
</tr>
<tr>
<td>1p36</td>
<td>0.41</td>
<td>1.00</td>
<td>−0.59</td>
</tr>
</tbody>
</table>

*84 Caucasian families.
†105 total sib-pair families.
‡lod = $Z^2_{LR}/2\ln10$ (12).