

ARTICLE

# Haplotype associations define target regions for susceptibility loci in systemic lupus erythematosus

Sari Koskenmies<sup>1</sup>, Elisabeth Widén<sup>2</sup>, Päivi Onkamo<sup>3</sup>, Petteri Sevón<sup>3</sup>, Heikki Julkunen<sup>4</sup> and Juha Kere<sup>\*,1,2,5</sup>

<sup>1</sup>Department of Medical Genetics, University of Helsinki, Finland; <sup>2</sup>Finnish Genome Center, Helsinki, Finland; <sup>3</sup>Basic Research Unit, Department of Computer Science, Helsinki Institute for Information Technology, University of Helsinki, Finland; <sup>4</sup>Department of Internal Medicine, Peijas Hospital, Helsinki University Hospital, Finland; <sup>5</sup>Department of Biosciences at Novum and Clinical Research Centre, Karolinska Institutet, Stockholm, Sweden

Systemic lupus erythematosus (SLE) is an autoimmune disease characterized by diverse and variable clinical manifestations. The etiology of SLE is still unknown, but both environmental and genetic factors are involved. Recent genome-wide scans and candidate genes studies in different ethnic groups have already suggested susceptibility loci for SLE, but most of the genetic component remains unexplained. We have previously conducted a genome-wide scan in 35 Finnish families multiply affected with SLE. With 417 microsatellite markers, we detected suggestive linkage in regions on chromosomes 6q and 14q as well as HLA on 6p. The 14q locus has also been implicated in three previous genome scans on SLE, whereas a partially overlapping region on 6q was implicated in one previous study. In an effort to obtain additional evidence for susceptibility loci on 6q and 14q and in order to refine their positions, we performed fine mapping at 1 cM density across the suggestive regions of linkage. Our results show evidence for excess sharing of a haplotype on 14q and excess transmission of a haplotype on 6q. Our results are compatible with the idea of a founder effect for susceptibility genes in SLE in central eastern Finland and suggest a path to the isolation of the putative susceptibility genes.

*European Journal of Human Genetics* (2004) 12, 489–494. doi:10.1038/sj.ejhg.5201125

Published online 31 March 2004

**Keywords:** genome screen; autoimmune disease; complex disease

## Introduction

Systemic lupus erythematosus (SLE) is an inflammatory disease characterized by autoantibody production and tissue injury. The vast majority of patients are women in their childbearing age (sex ratio 9:1). Symptoms flare up and recede over time and can appear in a variety of forms, ranging from mild to very severe. SLE is a complex disease with an unknown etiology. Both environmental, such as certain drugs or viruses, and genetic factors are involved in the development of the disease.<sup>1,2</sup> The mode of inheritance

is unknown, but both familial aggregation and twin studies support a strong genetic component for SLE: the concordance ratio of affected monozygotic twins over dizygotic twins is 10-fold (>20 vs 2%) and the risk for a first-degree relative is 10–20 times higher compared to the general population.<sup>3,4</sup> Several candidate loci (including the HLA region, *Fcγ* receptors, complement components and recently programmed cell death 1 gene (*PDCDI*)) have been implicated through association studies and candidate susceptibility loci have been detected on inbred mouse models of SLE.<sup>5,6</sup>

Until present, six genome-wide scans using SLE as phenotype without stratification have been carried out on SLE in different ethnic groups using different analytical approaches.<sup>7–12</sup> Each genome scan revealed several suggestive loci, but only six have surpassed the significance threshold of LOD  $\geq 3.3$ . These loci are on chromosomes

\*Correspondence: Professor J Kere, Department of Biosciences at Novum, Karolinska Institutet, Huddinge 141 57, Sweden. Tel: +46 8 6089158; Fax: +46 8 7745538; E-mail: juha.kere@biosci.ki.se  
Received 23 June 2003; revised 25 September 2003; accepted 14 October 2003

1q22–q23 and 1q41–q42, 2q37, 4p16, 6p21–p11 and 16q13. The highest LOD score of 4.24 was found on 2q37, where recently the *PDCD 1* was found to associate with SLE in a large multinational study.<sup>6</sup> Moreover, Johanneson *et al*<sup>13</sup> have recently shown a significant linkage to 1q31 ( $Z = 3.79$ ) with a set of 87 multicaser families from Europe and Mexico, Colombia and the United States.

We have focused our genetic studies of SLE on the Finnish population characterized by multiple rural founder effects.<sup>14</sup> We reasoned that such a population structure might offer advantages for studying a relatively rare complex disease. Studies in multiple sclerosis, asthma, type II diabetes and obesity, as examples of complex diseases, have already revealed susceptibility loci in Finland, and SLE with its lower incidence might show even more discrete founder effects. We conducted a genome scan in 35 multiplex families with 388 microsatellite markers with subsequent fine mapping and found evidence for suggestive linkage in regions on 6q23–q27 (NPL 2.47,  $P = 0.008$ ) and 14q21–q23 (NPL = 2.2,  $P = 0.02$ ) as well as HLA on 6p (NPL = 2.17,  $P = 0.02$ ). The 6q23–q27 linkage region also associates with other autoimmune diseases such as diabetes and rheumatoid arthritis, and hence might suggest a candidate region for susceptibility in several autoimmune diseases.<sup>15–17</sup> Our genome scan also suggested a peak on 14q21–q23 (at D14S587), within 1 cM from the peak of Gaffney *et al*<sup>7</sup> (NPL = 2.81  $P < 0.00016$  at D14S276). Furthermore, this region has been reported in two other independent studies, and thus should be considered a confirmed susceptibility locus for SLE.<sup>9,12</sup>

To proceed toward the identification of the susceptibility genes and to test our initial hypotheses of founder effects, we conducted further fine mapping of two of the susceptibility regions on 6q and 14q. In addition to the 35 multiplex families used in the linkage study, we ascertained 31 simplex families from a restricted region in Central Eastern Finland (Savo province, Kuopio). Kuopio was included because an initial recruitment phase in Helsinki suggested that more patients than expected originated from the Kuopio region, where the prevalence of SLE appeared to be slightly elevated. In order to verify the result of our linkage regions, blood samples of 104 SLE patients and 158 healthy relatives were genotyped with 44 new markers (19 markers on 6q and 25 markers on 14q) and analyzed for haplotype associations. The results reveal short haplotype conservation between families, lending support to the founder hypothesis and suggesting a method of positional cloning of the susceptibility genes.

## Subjects and methods

### Patient recruitment

The recruitment of patients as well as their phenotypes have been described.<sup>18</sup> Briefly, about 1200 SLE patients from a total of 21 hospitals in Finland were asked

personally or by mail as to whether they had any relative or a family member with SLE or a connective tissue disease similar to SLE. Patients with a positive family history of SLE and patients with sporadic SLE attending the Helsinki and Kuopio University Hospitals were asked to participate in the study. Kuopio was included because an initial recruitment phase in Helsinki suggested that more patients than expected came from the Kuopio region. All patients included in this study met the American College of Rheumatology criteria for the diagnosis of SLE.<sup>19</sup> Comparing the estimated prevalence of SLE in Finland, 28/100 000, we were able to identify 80–85% of all Finnish SLE patients requiring hospital-based treatment.<sup>20</sup>

### Selection of multiplex families

A total of 53 families multiply affected by SLE were identified among all contacted SLE patients. Of these families, 35 families with 73 affected pedigree members and 96 healthy relatives were suitable for a linkage mapping study.<sup>18</sup> One patient per pedigree was selected at random for association analyses using a computer program designed for this purpose.

### Singleton trios

To facilitate linkage disequilibrium mapping, we recruited SLE patients from the same geographical area, the Savo region, in Central Eastern Finland. This region is also known for a previously characterized founder effect for the *MLH1* susceptibility gene for colon cancer.<sup>21</sup> Altogether, 38 patients and 68 relatives were collected. To allow the unambiguous reconstruction of haplotypes, available parents or alternatively spouses and offspring were collected. All SLE patients had at least one parent or grandparent from the Savo region. Population records were used to track possible distant relationships between families.

### Genotyping

PCR assays, with 20 ng of genomic DNA, were conducted as described previously.<sup>22</sup> Our linkage peaks were initially genotyped with an average density of 3.5 cM between markers. Those data were combined with the data from the present study for an overall association analysis. Combined, we have data for a total of 32 markers for the linkage peak on chromosome 14q and for 25 markers on chromosome 6q, yielding a final density of  $< 1$  cM between markers (Figure 1). All markers for fine mapping were selected from Marshfield genetic maps at <http://research.marshfieldclinic.org/genetics/> and the marker order was verified using the NCBI sequencing database (<http://www.ncbi.nlm.nih.gov>) and UCSC human genome assembly (<http://genome.ucsc.edu>). The marker order was in accordance with the deCode genetic linkage map.<sup>23</sup> For multiplex families, Mendel errors were checked using the Pedmanager and PedCheck software<sup>24</sup> and overall

genotype frequencies were verified for Hardy–Weinberg equilibrium.

### Association and linkage analyses

We combined all available genotypes for an overall association analysis. To find allele and haplotype associations and study their effects, we used two approaches: haplotype pattern mining (HPM) and transmission disequilibrium test (TDT). HPM is a data mining-based algorithmic approach to genetic association analysis, in which frequent patterns of haplotypes associated with a trait are sought.<sup>25,26</sup> HPM analyses case–control data using nontransmitted parental chromosomes as controls (ie, pseudo-controls). We extracted independent trios from the pedigrees using an in-house software tool designed for this purpose. In total, 181 disease-associated chromosomes and 175 control chromosomes were obtained from 89 trios. The following parameters were used for HPM: maximum length of the pattern, 10 markers; maximum number of gaps, 1; and minimum  $\chi^2$  for a pattern, 3. To compensate for variable marker densities and marker information contents, a total of 10 000 permutations, fashioned as described,<sup>25</sup> were run to obtain empirical *P*-values. Linkage analyses were performed with Genehunter2.<sup>27</sup>

## Results

### Patients

The fine mapping and association study included 262 individuals from 66 families. Families originating from Savo region comprised 38 patients and 68 healthy relatives. With the exception of three multiplex families, included already in the genome scan, the other 30 families were simplex trios mostly with mother, father and offspring. In 21 of 30 families with mother, father and child, the affected was either an offspring (in 13 families) or one of the parents (in eight families), available for haplotype reconstruction. In families with a missing family member (mostly spouse or father), another relative (sib or first-degree relative) was available for haplotyping. Consanguinity was studied from population records up to the level of the late 18th century (in most families up to the mid-19th century) and none was observed between the families. Relationships between pedigree members in 35 multiplex families have been described previously.<sup>22</sup>

### Haplotype association in 14q

A total of 32 markers were analyzed on chromosome 14q covering a region of 24 cM across the implicated region between markers D14S70 and D14S63 (Figure 1). There were three intervals with 3 cM between markers (the consecutive intervals D14S980–D14S274–D14S592–D14S63) and one with 2 cM (D14S587–D14S1064). The average polymorphism information content (PIC) value was 0.72. By linkage, the highest peak after fine mapping

was observed at D14S1055 (50.30 cM), NPL 2.22 ( $P=0.015$ ), 5 cM centromeric from our previous peak at D14S587 using a lower density of markers.<sup>22</sup>

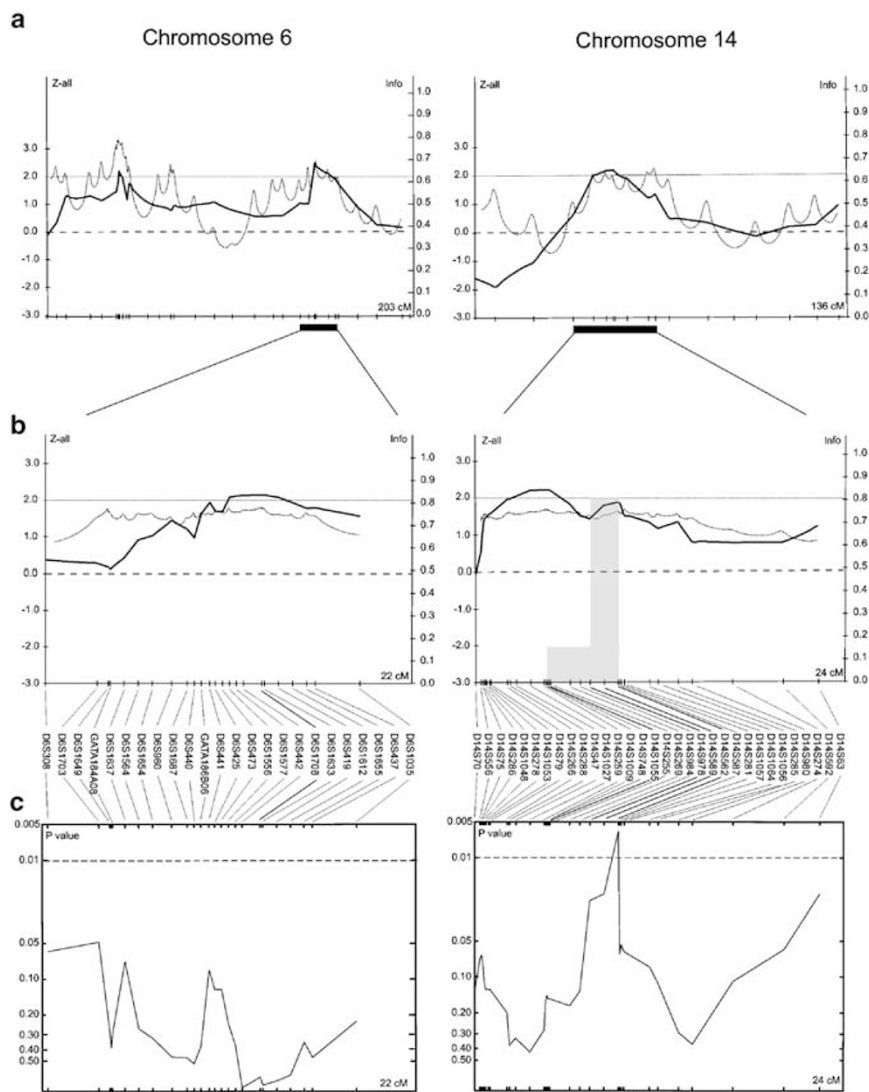
Trios from 65 families were analyzed with TDT and HPM using the randomization procedure to assess empirical significance. A significant association of a 2 cM long haplotype involving 10 chromosomes for the markers D14S978–D14S589–D14S562 at 53.2–55.2 cM (alleles 3-2-8) was identified (Figure 1). The pattern was present in 10 chromosomes from SLE patients but in none of the control chromosomes ( $\chi^2=9.95$ , empirical  $P=0.00598$  for marker D14S562, odds ratio not defined). The next best region of allele sharing was seen between markers D14S1009 and D14S748 (alleles 3-1) at 50.1–50.2 cM. Altogether, 15 patient chromosomes and two control chromosomes shared this haplotype, which, however, did not reach significance ( $\chi^2=9.99$ , empirical  $P=0.14$ , odds ratio = 13, 95% confidence interval 2.9–59).

Interestingly, both shared haplotypes 3-1 and 3-2-8 were found in the same chromosomes in three individuals, suggesting common origin. Indeed, the longest shared haplotype between two unrelated SLE patients covered 8 cM from D14S288 to D14S281 (Figure 2). Only five meioses were informative for TDT analyses for these haplotypes. The 3-2-8 haplotype was transmitted four times and nontransmitted one time ( $P=0.18$ ). For the 3-1 haplotype, there were six transmitted and two nontransmitted chromosomes ( $P=0.157$ ). The TDT thus remained inconclusive.

### Haplotype association in 6q

On chromosome 6q, the number of mapped markers was 25 across a 22 cM region from D6S308 to D6S1035. With the exception of two gaps with 4 and 3 cM between markers D6S308–D6S1703 and D6S437–D6S1035, respectively, all other markers had an intermarker distance of 1 cM or less, as listed in Figure 1. The mean marker information content (PIC) was 0.74. The highest NPL score observed in genome scan decreased slightly from 2.47 (marker D6S960 at 151.4 cM) to 2.27 ( $P=0.013$ ) and the peak marker was shifted 5 cM telomeric to D6S1708 at 157.8 cM. The NPL scores and the corresponding information content are shown in Figure 1.

Trios used for the HPM and TDT analyses were the same as for chromosome 14q analysis. By HPM analysis, some allele sharing was observed between markers D6SGA-TA184A08 (allele 1) and D6S1637 (allele 3) at 147–148 cM, 10 cM away from the peak of linkage. The haplotype GATA184A08–D6S1637 (alleles 1-3) was seen in 27 patient chromosomes and nine control chromosomes ( $\chi^2=8.017$ , empirical  $P=0.0719$  for marker D6S1637, odds ratio = 5.8, 95% confidence interval 2.6–13). The empirical association thus remained nonsignificant. However, in TDT analyses, the GATA184A08–D6S1637 haplotype 1-3



**Figure 1** (a) Linkage maps of chromosomes 6 and 14 from 35 Finnish families multiply affected by SLE, full chromosome views (Koskenmies *et al*, 2003). The distances between markers in linkage regions (black bars) are on average 3.5 cM. (b) Combined linkage results from 35 multiplex and 31 singleton (Savo) families after addition of fine-mapping markers, detailed maps. The intermarker distances are  $\leq 1$  cM. Gray bars show the shared associating haplotypes in 10 patients by HPM analysis. Detailed haplotypes are shown in Figure 2. (c) Association analyses using HPM. The dotted line depicts the empirical significance level of 0.01. The y-axis shows the *P*-value, expressed in logarithmic scale. The x-axis shows the genotyped region in cM (same scale as b).

was transmitted 15 times and not transmitted five times ( $\chi^2 = 5.00$ ,  $P = 0.025$ ).

#### Geographical distribution of the associated haplotypes

Our ascertainment scheme was intentionally biased to include an excess of families from Savo in central eastern Finland, in anticipation of a founder effect. Tracing back the birthplaces of parents or grandparents of the patients, we found that five out of 10 chromosomes (5/9 families) with 3-2-8 haplotype on chromosome 14 came from Savo. The population of this geographical region has previously

shown founder effects in studies with monogenic diseases.<sup>21,28</sup> When looking even more carefully, two of the Savo families from neighboring parishes shared a long haplotype of 8 cM, comprising both associated haplotypes 3-1 and 3-2-8, as described above. Two families with the 3-2-8 haplotype originate from western Finland. There are three families (two originating from Savo) where both the chromosome 14q haplotypes and also the 6q two-marker haplotype (3-1, 3-2-8 on chromosome 14 and 1-3 on chromosome 6) are found, and five families (three originating from Savo) with six patients, where the

D14S288 (47.37)	D14S47 (48.00)	D14S1027 (49.00)	D14S259 (50.00)	D14S1009 (50.10)	D14S748 (50.20)	D14S1055 (50.30)	D14S255 (50.40)	D1246269 (51.79)	D14S984 (52.50)	D14S978 (53.20)	D14S589 (54.20)	D14S562 (55.30)	D14S281 (55.40)
2	0	6	2	0	2	0	3	0	8	3	2	8	0
2	1	5	4	1	1	3	4	4	15	3	2	8	2
2	0	0	7	0	1	0	5	2	8	3	2	8	2
2	1	2	2	2	3	5	3	3	1	3	2	8	5
11	1	3	3	2	3	1	5	1	1	3	2	8	1
5	1	5	5	0	2	3	3	1	1	3	2	8	5
4	1	1	2	2	3	1	3	0	0	3	2	8	5
1	0	4	2	3	1	4	3	1	8	3	2	8	5
11	1	4	2	3	1	1	3	0	1	3	2	8	2
11	1	4	2	3	1	1	3	1	1	3	2	8	2

**Figure 2** Detailed analysis of independent SLE-associated haplotypes for marker loci on 14q from a set of patients of 35 multiplex and 31 singleton (Savo) Finnish families. The haplotype 3-2-8 is present in 10 patients and no controls and is transmitted four times and not transmitted once. The haplotype 3-1, present in 15 patients and two controls, is transmitted six times and not transmitted two times.

haplotype 3-2-8 or 3-1 on chromosome 14q and 1-3 on chromosome 6q are present. Considering the exclusion of close relationships between the families, such haplotype sharing is highly unexpected if not associated to the selection criterium (ie SLE), and the clustering of the associated haplotypes to eastern Finland may suggest a founder effect, possibly involving two loci on chromosomes 14q and 6q.

**Discussion**

The positional cloning of susceptibility genes in complex diseases proceeds from linkage mapping to association study under the hypothesis of shared ancestral mutations among the patients. In this study, we have verified our linkage study results with a dense map of markers and taken the first steps toward a genetic association study for two loci on chromosomes 14q and 6q. We designed our study to take advantage of the subisolate structure of Finland and targeted especially one such region in central eastern Finland. Remarkably, we found significantly excessive haplotype sharing between patients from neighboring parishes whose close relationships were, however, excluded at least back to the mid-19th century. Even more surprisingly, six patients from five unrelated families shared two loci in different chromosomes, an observation that is highly unlikely if not causally associated with the disease.

The length of the associated haplotypes is roughly in accordance with the population history. The chromosome 14q haplotype 3-2-8 spans 2cM and the chromosome 6q

haplotype 1-3 spans 1 cM. Such a long haplotype sharing is not expected unless caused by a relatively young founder effect. The area was settled in the early 16th century and similar founder effects have been noted there before.<sup>29,30</sup>

Our finding is compatible with previous studies on three independent data sets, but refines the search region considerably. Gaffney *et al*<sup>7</sup> found linkage of SLE to D14S276 (LOD 2.81,  $P < 0.00016$ ) using multipoint non-parametric methods with 105 sib pairs. Our conserved haplotype lies less than 1 cM away from the peak of Gaffney *et al*.<sup>7</sup> Moreover, Shai *et al*,<sup>9</sup> in a study of 80 multiplex families, and Lindqvist *et al*<sup>12</sup> with Swedish families have reported linkages to 14q with LOD scores of 2.02 ( $P = 0.02$ , D14S258) and 1.15 (D14S592), respectively. For chromosome 6q, linkage to SLE has previously been reported about 20 cM telomeric to our linkage peak.<sup>8</sup> Moreover, the same general region has been implicated in previous studies with insulin-dependent diabetes mellitus (6q25, IDDM5, between markers D6S476 and D6S473, LOD 4.5) and rheumatoid arthritis (D6S311 and D6S440,  $P = 0.016$  and 0.017, respectively).<sup>15,16,17</sup> The overtransmitted haplotype on 6q identified in our study coincides with the peak of Myerscough *et al*.<sup>17</sup>

No obvious candidate genes reside within the conserved haplotype regions in gene maps. It will be necessary to perform further fine mapping with tens of more tightly spaced SNP markers within and near the identified haplotypes on both 14q and 6q to confirm the suggested identity-by-descent of the haplotypes. Thereafter, positional cloning can proceed to verifying new transcripts and studying their polymorphisms and possible associations with SLE in this and other sample sets. We conclude that our fine-mapping results suggest potentially highly interesting regions for additional susceptibility genes in SLE.

**Acknowledgements**

We thank all the patients and their family members who participated in this study and the many physicians who referred families and verified the diagnosis. We thank Riitta Lehtinen, Sinikka Lindh, Sirkka Ruohomäki, Sirkku Ekström, Päivi Malinen and Teija Baas for expert help in the laboratory and sample collection, Anne Lehto for invaluable help with data management and Päivi Tikka and Päivi Lahermo for help with characterizing markers and genotyping. This study was supported by the Stockmann Foundation, Sigrid Jusélius Foundation, Helsinki University Central Hospital research funds, and Academy of Finland.

**References**

- Patel GK, Anstey AV: Rifampicin-induced lupus erythematosus. *Clin Exp Dermatol* 2001; 26: 260–262.
- James JA, Kaufman KM, Farris AD, Taylor-Albert E, Lehman TJ, Harley JB: An increased prevalence of Epstein–Barr virus infection in young patients suggests a possible etiology for systemic lupus erythematosus. *J Clin Invest* 1997; 100: 3019–3026.

- 3 Deapen D, Escalante A, Weinrib L *et al*: A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis Rheum* 1992; **35**: 311–318.
- 4 Vyse TJ, Todd JA: Genetic analysis of autoimmune disease. *Cell* 1996; **85**: 311–318.
- 5 Kelly JA, Moser KL, Harely JB: The genetics of systemic lupus erythematosus: putting the pieces together. *Genes Immun* 2002; **3** (Suppl 1): 71–85.
- 6 Prokunina L, Castillejo-Lopez C, Oberg F *et al*: A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet* 2002; **32**: 666–669.
- 7 Gaffney PM, Kearns GM, Shark KB *et al*: A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families. *Proc Natl Acad Sci USA* 1998; **95**: 14875–14879.
- 8 Moser KL, Neas BR, Salmon JE *et al*: Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African–American pedigrees. *Proc Natl Acad Sci USA* 1998; **95**: 14869–14874.
- 9 Shai R, Quismorio Jr FP, Li L *et al*: Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families. *Hum Mol Genet* 1999; **8**: 639–644.
- 10 Gaffney PM, Ortmann WA, Selby SA *et al*: Genome screening in human systemic lupus erythematosus: results from a second Minnesota cohort and combined analyses of 187 sib-pair families. *Am J Hum Genet* 2000; **66**: 547–556.
- 11 Gray-McGuire C, Moser KL, Gaffney PM *et al*: Genome scan of human systemic lupus erythematosus by regression modeling: evidence of linkage and epistasis at 4p16 15.2. *Am J Hum Genet* 2000; **67**: 1460–1469.
- 12 Lindqvist AK, Steinsson K, Johanneson B *et al*: A susceptibility locus for human systemic lupus erythematosus (hSLE1) on chromosome 2q. *J Autoimmun* 2000; **14**: 169–178.
- 13 Johanneson B, Lima G, Von Salome J, Alarcon-Segovia D, Alarcon-Riquelme ME: A major susceptibility locus for systemic lupus erythematosus maps to chromosome 1q31. *Am J Hum Genet* 2002; **71**: 1060–1071.
- 14 Kere J: Human population genetics: lessons from Finland. *Annu Rev Genom Hum Genet* 2001; **2**: 103–128.
- 15 Luo DF, Buzzetti R, Rotter JI *et al*: Confirmation of three susceptibility genes to insulin-dependent diabetes mellitus: IDDM4, IDDM5 and IDDM8. *Hum Mol Genet* 1996; **5**: 693–698.
- 16 Todd JA: Genetics of type 1 diabetes. *Pathol Biol (Paris)* 1997; **45**: 219–227.
- 17 Myerscough A, John S, Barrett JH, Ollier WE, Worthington J: Linkage of rheumatoid arthritis to insulin-dependent diabetes mellitus loci: evidence supporting a hypothesis for the existence of common autoimmune susceptibility loci. *Arthritis Rheum* 2000; **43**: 2771–2775.
- 18 Koskenmies S, Widen E, Kere J, Julkunen H: Familial systemic lupus erythematosus in Finland. *J Rheumatol* 2001; **28**: 758–760.
- 19 Tan EM, Cohen AS, Fries JF *et al*: The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982; **25**: 1271–1277.
- 20 Helve T: Prevalence and mortality rates of systemic lupus erythematosus and causes of death in SLE patients in Finland. *Scand J Rheumatol* 1985; **14**: 43–46.
- 21 Nyström-Lahti M, Sistonen P, Mecklin JP *et al*: Close linkage to chromosome 3p and conservation of ancestral founding haplotype in hereditary nonpolyposis colorectal cancer families. *Proc Natl Acad Sci USA* 1994; **94**: 6054–6058.
- 22 Koskenmies S, Lahermo P, Julkunen H, Ollikainen V, Kere J, Widén E: Linkage mapping of systemic lupus erythematosus (SLE) in Finnish multiplex families. *J Med Genet* 2004; **41**:e2.
- 23 Kong A, Gudbjartsson DF, Sainz J *et al*: A high-resolution recombination map of the human genome. *Nat Genet* 2002; **31**: 241–247.
- 24 O’Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
- 25 Toivonen HT, Onkamo P, Vasko K *et al*: Data mining applied to linkage disequilibrium mapping. *Am J Hum Genet* 2000; **67**: 133–145.
- 26 Sevon P, Ollikainen V, Onkamo P, Toivonen HT, Mannila H, Kere J: Mining associations between genetic markers, phenotypes and covariates. *Genet Epidemiol* 2001; **21** (Suppl 1): 588–593.
- 27 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.
- 28 Høglund P, Sistonen P, Norio R *et al*: Fine mapping of the congenital chloride diarrhea gene by linkage disequilibrium. *Am J Hum Genet* 1995; **57**: 95–102.
- 29 de la Chapelle A: Disease gene mapping in isolated human populations: the example of Finland. *J Med Genet* 1993; **30**: 857–865.
- 30 Lahermo P, Sajantila A, Sistonen P *et al*: The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. *Am J Hum Genet* 1996; **58**: 1309–1322.