

SHORT REPORT

Systemic lupus erythematosus (SLE) and chromosome 16: confirmation of linkage to 16q12–13 and evidence for genetic heterogeneity

Swapan K Nath^{*,1,2}, Bahram Namjou¹, David Hutchings¹, C. Phillip Garriott¹, Catherine Pongratz¹, Joel Guthridge¹ and Judith J James^{1,3}

¹Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA; ²Genetic Epidemiology Unit, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA; ³Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA

Systemic lupus erythematosus (SLE) is a multisystem autoimmune disease with significant morbidity and mortality, characterized by remarkable clinical variability with unknown etiology. Genetic contribution to the development of SLE is well established. Recently, we found evidence ($P < 0.004$) of linkage at 16p13 and 16q12–13 in a genome scan based on 37 Hispanic families. The main objective of this study is to replicate and confirm the linkage at these two genomic locations in two large independent replication data sets designated as, group-1 and group-2, consisting of 172 and 120 multiplex SLE families, respectively. We have found a significant evidence of linkage with high heterogeneity ($HLOD = 4.85$, $\alpha = 35\%$) at 16q12–13 in group-2. Other independent research groups also reported the SLE susceptibility at or close to 16q12–13 previously. Therefore, independent published reports, together with our initial linkage with Hispanics and followed by significant evidence from group-2, provide a strong and confirmed evidence for an SLE susceptibility locus at 16q12–13.

European Journal of Human Genetics (2004) 12, 668–672. doi:10.1038/sj.ejhg.5201209
Published online 28 April 2004

Keywords: systemic lupus erythematosus; genetic heterogeneity; linkage confirmation

Introduction

Systemic lupus erythematosus (SLE) [MIM 152700] is a multisystem autoimmune disease with high morbidity and mortality. The overall estimated prevalence in the United States is between 12 and 64 cases per 100 000 individuals.^{1,2} Significant gender differences (female:male = 9:1) are observed in prevalence, age at onset, premorbid conditions, clinical expression, course of illness, response to treatment, and morbid risk.^{1,2} Although the etio-pathogenesis of SLE remains largely unknown, the dysregulation of self-reactive B cells leading to immune complex activation and

complement consumption is well observed. The production of autoantibodies against nuclear components is the hallmark characteristic of SLE. The genetic basis for SLE is well established, although, various environmental factors, perhaps interacting with certain genes, also may play a significant role in development of SLE. The relative risk ratio to the sibs of an affected proband (λ_s) is about 20–40.³

Recently, in an effort to identify the novel susceptibility genes for SLE, we performed a genome-wide scan on 37 Hispanic families with multiple affected individuals.⁴ We have identified the evidence of linkage to SLE at two regions on chromosome 16 (16p13 and 16q12–13). While the evidence of linkage at 16q12–13 exceeded the recommended threshold for genome-wide suggestive linkage⁵ ($Z_{lr} = 3.06$, $LOD = 2.01$, $P = 0.001$ at 66.6 cM), the evidence of linkage at 16p13 was also nearly suggestive ($Z_{lr} = 2.68$, $LOD = 1.13$, $P = 0.004$ at 20.0 cM).

*Correspondence: Dr SK Nath, Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, Oklahoma 73104, USA. Tel: +1 405 271 7766; Fax: +1 405 271 4110; E-mail: swapan-nath@omrf.ouhsc.edu
Received 4 February 2004; revised 18 March 2004; accepted 19 March 2004

The main purpose of this study is to evaluate the linkage at these two genomic locations on chromosome 16 in two large independent data sets. Here we report the confirmation of linkage at 16q12–13 in one of the two ethnically diverse data sets. We also report the presence of linkage heterogeneity at the 16q12–13 susceptibility locus.

Materials and methods

All the SLE patients met the 1997 revised criteria for classification of SLE.^{6,7} Detailed ascertainment procedure for the SLE families is given elsewhere.^{4,8,9} Two independent replication samples consisting of 172 families (group-1) and 120 families (group-2) are used to verify our previously identified linkage effects. The group-1 consists of 110 European-American (EA) and 62 African-American (AA) families while group-2 consists of 82 EA and 38 AA families (Table 1). Although most of the families from group-1 are used in our previous genome scans, the families from group-2 are used for the first time for chromosome 16 replication study.

All the linkage analyses are performed using the families multiplex (at least two affected individuals) for SLE. Genomic DNA was isolated from peripheral blood cells, buccal cell swabs, mouth wash specimens, or EBV transformed cell lines using standard methods. A total of 11 microsatellite markers with an average marker spacing of 11 cM (range 8–15 cM) and a heterozygosity of 76% (range 56–94%) were typed for chromosome 16. We used the genetic map and intermarker distances from the Marshfield map.¹⁰ Positions of markers not in the Marshfield map were set by interpolation on the basis of physical distances. Major family and demographic characteristics were compared between these two groups. We assess the significance of the quantitative variables using one-way ANOVA and *post hoc* Tukey's test for multiple pair-wise comparisons between groups.

The evidence of linkage has been assessed by both a nonparametric, penetrance-independent, allele-sharing method, and a parametric method using the program GENEHUNTER-PLUS.^{11,12} GENEHUNTER-PLUS reports allele-sharing LOD scores and maximum Z scores for the

likelihood ratio (Zlr). Parametric linkage analysis is performed assuming both linkage homogeneity (LOD) and linkage heterogeneity (HLOD). Significance of LOD is determined by $\chi^2 = 4.6 \times \text{LOD}$ with one d.f. Since HLOD follows a complex statistical distribution, significance for the observed HLOD is first converted to a χ^2 , where $\chi^2 = 4.6 \times \text{HLOD}$. *P*-value (P_1) was then derived for χ^2 , using the χ^2 distribution with one d.f. The *P*-value for the HLOD score is then $0.5 \times [1 - (1 - P_1)(1 - P_1)]$.¹³ We have used the race-specific marker allele frequencies estimated from family founders when we analyzed the families with a specific racial background.

Results

The comparison of major family and demographic characteristics show that overall the four groups are similar (Table 1). We have assessed the significance of the quantitative variables using one-way ANOVA and *post hoc* Tukey's test for multiple pair-wise comparisons between groups. The mean age at onset for affected individuals does not vary significantly among four groups ($F = 1.54$, $P = 0.20$). Similar results have been found using the *post hoc* pair-wise comparison between the groups.

The results of multipoint nonparametric and parametric linkage analyses at 16q12–13 are given in Table 2. Virtually, no evidence of linkage was revealed by group-1. In contrast, moderate evidence of linkage was revealed in group-2 by the nonparametric analysis (Zlr = 2.45, $P = 0.007$). Once the evidence of linkage at 16q12–13 was found, we performed parametric multipoint analyses, employing simple dominant and recessive models. Greenberg *et al*¹⁴ have shown, as long as both dominant and recessive models are used, that simple genetic models are nearly as powerful to determine evidence of linkage as when the true parameters are known. While no evidence of linkage is found by the parametric analysis assuming linkage homogeneity (LOD = –19.83), a significant evidence of linkage is found by assuming genetic heterogeneity (HLOD = 4.85, $P = 4.5 \times 10^{-6}$) under a dominant model, and about 35% families are estimated to be linked with this locus. Although the analyses based on these two

Table 1 Selected clinical and demographic features of the affected SLE individuals

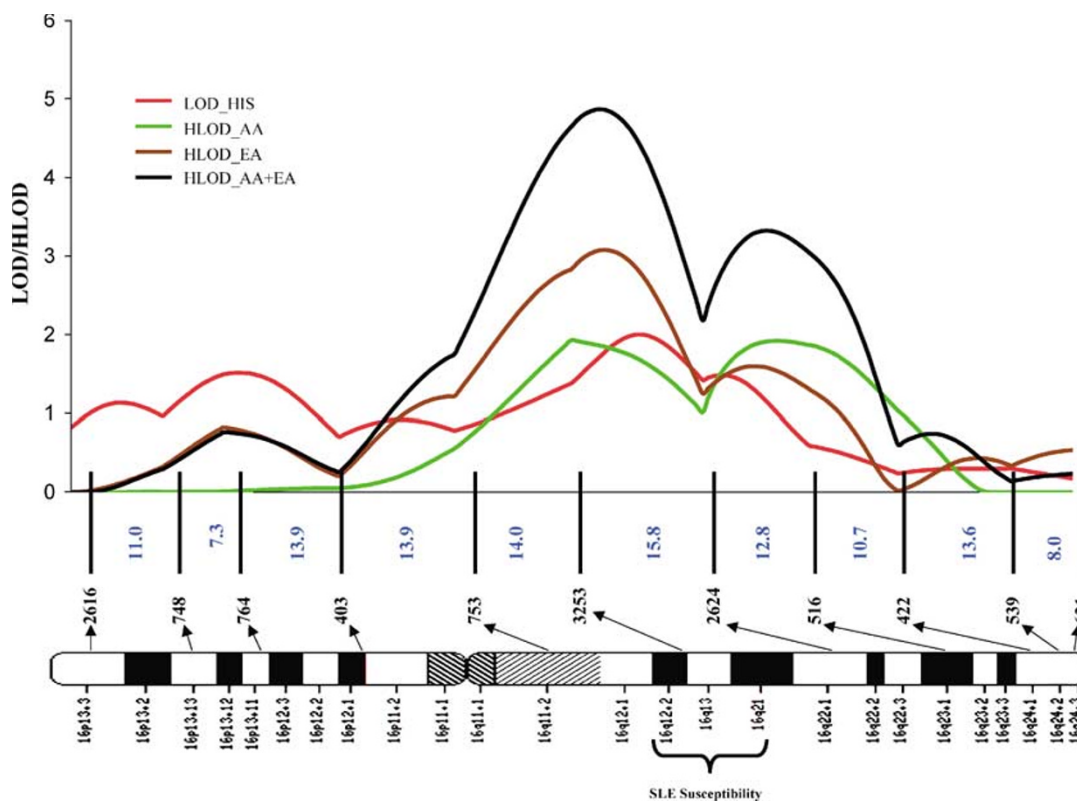
Characteristics	Group-1 ^a		Group-2	
	EA	AA	EA	AA
# Family	110	62	82	38
# SLE affected	250	144	172	78
Sex (male:female)	23:227	9:135	17:155	8:70
# SLE per family	2.3	2.3	2.1	2.0
Age at onset (years) ± SD	35.0 ± 12.2	34.3 ± 10.7	37.0 ± 14.6	34.1 ± 12.1

^aAA = African-American, EA = European-American.

Table 2 Multipoint parametric and nonparametric linkage analyses on chromosome 16 for group-1 and group-2. For comparison, the linkage results from Hispanic (Nath *et al*, 2004) families are also provided

Replication group	Ethnic group (n)	Nonparametric linkage				Parametric linkage			
		Zlr	LOD	P-value	Peak at (cM)	LOD (Model) ^a	HLOD	α (%)	Peak at (cM)
Nath <i>et al</i> (2004)	HIS (37)	3.06	2.01	0.001	68.0	-1.13 (Rec)	1.54	44	62.3
Group-1	EA (110)	1.24	0.33	0.107	32.2	-40.04 (Rec)	0.03	02	79.7
	AA (62)	0.75	0.12	0.237	11.0	-24.5 (Dom)	0.01	00	60.1
	EA+AA (172)	-0.66	0.09	0.255	67.2	-48.6 (Rec)	0.01	00	66.1
Group-2	EA (82)	1.72	0.64	0.043	66.4	-13.3 (Dom)	3.22	34	66.4
	AA (38)	2.05	0.91	0.020	60.1	-17.6 (Dom)	1.79	35	60.1
	EA+AA (120)	2.45	1.30	0.007	63.3	-19.83 (Dom)	4.85	35	64.8

^aDisease allele frequency and three corresponding genotypic penetrances for Recessive (Rec) and Dominant (Dom) models are: 0.1, 0.2, 0.0, 0.0 (Rec) and 0.001, 0.99, 0.99, 0.0 (Dom). The estimated prevalence under both models is 0.002, HLOD = LOD score assuming linkage heterogeneity, α = proportion of families linked, HIS = Hispanic, AA = African-American, EA = European-American.

**Figure 1** Multipoint linkage analysis in four groups of families (EA = European-American ($n = 82$), HIS = Hispanic ($n = 37$), AA = African-American ($n = 38$) and AA + EA = AA and EA combined ($n = 120$)). The marker names ('D16S' is removed from all the markers) and their relative positions and inter-marker distances (cM) are indicated by an arrow, vertical line and numbers.

methods identify two different peaks, but these peaks are very close to each other between markers D16S3253 and D16S2624. Similar patterns of peak location and linkage heterogeneity are found when linkage analyses are per-

formed separately in the individual group (Table 2 and Figure 1). The results of parametric linkage analysis under recessive models are significantly inferior to the dominant model (data not shown).

Discussion

Several genome screens have been performed to find the susceptibility genes for lupus. Interestingly and importantly, the evidence of linkage and association in and around 16q12–13 has previously been identified by at least four independent studies. First, the USC group¹⁵ found modest evidence ($P=0.017$) for linkage to marker D16S3136 at 62.1 cM in Mexican-American and EA families. Second, the Minnesota group¹⁶ reported significant evidence of SLE linkage at 16q12–13. They found maximum evidence of linkage (LOD=3.85, $P=1.3 \times 10^{-5}$) at 60.45 cM on marker D16S415 in mostly EA families. Third, the UCLA group¹⁷ confirmed the evidence of SLE linkage spanning between markers D16S753 and D16S757 at 16q11–q13 in mostly non-Caucasian families. They also found a positive epistatic interaction between this locus and the 1q23 SLE susceptibility locus. Fourth, a Chinese group¹⁸ found an association between D16S517 (16q12) and SLE at 58.46 cM after performing a family-based association study on the Chinese population. These four independent studies found maxima over a range of 9 cM (58 and 69 cM), which is a typical shift of linkage peak location for a complex disease linkage.

Our previously identified second linkage signal on chromosome 16 at 16p13 in Hispanics⁴ has not been replicated in any of these replication groups. Several reasons may contribute. First, a factor contributing to the lack of consistency across studies is locus heterogeneity, which can weaken or even eliminate evidence for linkage that is present only in a subset of families. This phenomenon is clearly exemplified from the linkage at 16q12–13 in which we did not find any evidence of linkage in any of the ethnic group from group-1. Although a significant evidence for linkage is found at 16q12–13 in group-2, it also showed tremendous genetic heterogeneity. Approximately, 65% genetic heterogeneity is estimated in group-2, which is consistent among the families and is irrespective of their ethnic origin. Second, current sample sizes would allow the reliable identification of only major genes that take up a large proportion of all the overall familial risk. In practice, this means that studies need 'luck', in addition to a high standard of data quality, to detect the loci with relatively small effect size. Third, this linkage might be a false positive linkage; therefore, it is not detected in the replication groups.

In conclusion, our results provide strong evidence that confirms the existence of a major SLE susceptibility gene on 16q12–13. In our data initially we have shown the evidence of linkage in Hispanics⁴ and in the present study, we show this SLE linkage in the families with AA, EA ethnic background. This susceptibility locus also demonstrates substantial genetic heterogeneity. Nevertheless, the finding of highly significant evidence in the present study and the

previously published results from four independent studies demonstrates the powerful evidence of linkage to this chromosomal region, and furthermore, this linkage is reproducible and 'consistent'. Several independent confirmations would justify further and more intensive approaches to identify the actual SLE susceptibility gene(s) at 16q12–13.

Acknowledgements

We thank the participating patients and family members, as well as their referring physicians. Some of the pedigrees were obtained from the Lupus Multiplex Registry and Repository (AR12253). We are grateful to Dr John B Harley for critical reading of the manuscript and for valuable advice during the course of this project. The study was supported by the NIH grant (AR048928) awarded to SKN.

References

- 1 Hochberg MC: The epidemiology of systemic lupus erythematosus; in Wallace DJ, Hahn BH (eds): *Dubois' lupus erythematosus*. Baltimore: Williams and Wilkins; 1997, pp 49–65.
- 2 Lawrence RC, Helmick CG, Arnett FC *et al*: Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* 1998; **41**: 778–799.
- 3 Wandstrat A, Wakeland E: The genetics of complex autoimmune diseases: non-MHC susceptibility genes. *Nat Immunol* 2001; **2**: 802–809.
- 4 Nath SK, Quintero-Del-Rio AI, Kilpatrick J, Feo L, Ballesteros M, Harley JB: Linkage at 12q24 with systemic lupus erythematosus (SLE) is established and confirmed in Hispanic and European-American families. *Am J Hum Genet* 2004; **74**: 73–82.
- 5 Lander ES, Kriglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
- 6 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.
- 7 Hochberg MC: Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997; **40**: 1725.
- 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* 1998; **95**: 14869–14874.
- 9 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.
- 10 Broman KW, Murray JC, Sheffield VC, White RL, Weber JL: Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 1998; **63**: 861–869.
- 11 Kriglyak 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.
- 12 Kong A, Cox NJ: Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 1997; **61**: 1179–1188.
- 13 Faraway JJ: Distribution of the admixture test for the detection of linkage under heterogeneity. *Genet Epidemiol* 1993; **10**: 75–83.
- 14 Greenberg DA, Abreu P, Hodge SE: The power to detect linkage in complex disease by means of simple LOD-score analyses. *Am J Hum Genet* 1998; **63**: 870–879.

- 15 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.
- 16 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.
- 17 Tsao B, Cantor RM, Grossman JM *et al*: Linkage and interaction of loci on 1q23 and 16q12 may contribute to susceptibility to systemic lupus erythematosus. *Arthritis Rheum* 2002; **46**: 2928–2936.
- 18 Feng XB, Shen N, Chen SL *et al*: Susceptibility gene of systemic lupus erythematosus in 16q12 in Chinese cohort. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2003; **20**: 27–30.