Am. J. Hum. Genet. 67:1460–1469, 2000

Genome Scan of Human Systemic Lupus Erythematosus by Regression Modeling: Evidence of Linkage and Epistasis at 4p16-15.2

C. Gray-McGuire,[1,5] K. L. Moser,[6] P. M. Gaffney,[6] J. Kelly,[1] H. Yu,[1] J. M. Olson,[5] C. M. Jedrey,[5] K. B. Jacobs,[5] R. P. Kimberly,[7] B. R. Neas,[2] S. S. Rich,[8] T. W. Behrens,[6] and J. B. Harley[1,3,4]

[1]Arthritis and Immunology Program, Oklahoma Medical Research Foundation, Departments of [2]Biostatistics and Epidemiology and [3]Medicine, University of Oklahoma Health Sciences Center, and [4]U.S. Department of Veterans Affairs Medical Center, Oklahoma City; [5]Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland; [6]Department of Medicine, University of Minnesota Medical School, Minneapolis; [7]Department of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham; and [8]Department of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC

Systemic lupus erythematosus (SLE) is a complex autoimmune disorder involving at least hormonal, environmental, and genetic factors. Familial aggregation, a 2%–3% sibling recurrence rate, monozygotic twin concordance >20%, association with several candidate genes, as well as the results of five genome scans support a genetic component. We present here the results of a genome scan of 126 pedigrees multiplex for SLE, including 469 sibling pairs (affected and unaffected) and 175 affected relative pairs. Using the revised multipoint Haseman-Elston regression technique for concordant and discordant sibling pairs and a conditional logistic regression technique for affected relative pairs, we identify a novel linkage to chromosome 4p16-15.2 ($P = .0003$ and LOD = 3.84) and present evidence of an epistatic interaction between chromosome 4p16-15.2 and chromosome 5p15 in our European American families. We confirm the evidence of linkage to chromosome 4p16-15.2 in European American families using data from an independent pedigree collection. In addition, our data support the published results of three independent studies for nine purportedly linked regions and agree with the previously published results from a subset of these data for three regions. In summary, results from two new analytical techniques establish and confirm linkage with SLE at 4p16-15.2, indicate epistasis between 4p16-15.2 and 5p15, and confirm other linkage effects with SLE that have been reported elsewhere.

Introduction

Recent advances in molecular genetics have resulted in increasingly more detailed physical and genetic maps of the human genome. As a result, genome scans using large numbers of highly polymorphic genetic microsatellite markers to study genetic diseases have become not only plausible but quite efficient (Lander and Botstein 1989; Reed et al. 1994). However, as results from genome scans for complex genetic disorders become available, the problem of establishing linkage and subsequent gene identification remains a significant challenge. Studies of schizophrenia (as reviewed by Tsuang et al. 1999), type I diabetes (as reviewed by Friday et al. 1999), and systemic lupus erythematosus (SLE [MIM 152700]) exemplify instances in which results vary substantially between studies, presumably because of heterogeneity of sample populations and differing methods of analysis.

SLE is a clinically heterogeneous autoimmune disorder that predominantly affects women (~90%) and is much more prevalent and severe in African Americans, compared with European Americans (Kaslow and Masi 1978). The etiology of SLE is complex, involving both environmental and genetic factors and, likely, a synergistic relationship between the two. High heritability (>66%) (Lawrence et al. 1987), increased concordance rates among monozygotic twins (25%–69%), compared with dizygotic twins and other full siblings (2%–3%) (Deapen et al. 1992; Reichlin et al. 1992), familial aggregation (Sestak et al. 1999), association with multiple candidate genes (as reviewed by Tan and Arnett 1998), and linkage analysis results from five different genome scans (Moser et al. 1998; Gaffney et al. 1998, 2000; Shai et al. 1999; Lindqvist et al. 2000) all support a genetic component. Results of cohort-comparison studies that include environmental triggers such as Epstein-Barr virus (James et al. 1997), human T cell lymphotropic virus (Brand et al. 1999), and hormonal and chemical exposure (Cooper et al. 1998) support a

Received July 17, 2000; accepted for publication October 19, 2000; electronically published November 9, 2000.

Address for correspondence and reprints: Dr. John B. Harley, Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, OK 73104. E-mail: john-harley@omrf.ouhsc.edu

0002-9297/2001/6706-0012$02.00

gene-environment interaction, and one cohort-comparison study supports a gene-gene interaction (Mehrian et al. 1998).

The five genome scans completed to date evaluated multicase SLE pedigrees collected by investigators in Oklahoma (Moser et al. 1998), Minnesota (scan I, Gaffney et al. 1998 and scan II, Gaffney et al. 2000), California (Shai et al. 1999), and Sweden (Lindqvist et al. 2000). Each used different analysis methods and populations. The Minnesota I (MNI), Minnesota II (MNII), and California (CA) scans used model-free analysis, primarily because their collections were ascertained on the basis of affected sibling pairs, whereas the Oklahoma (OK) and Sweden (SW) scans applied model-based techniques to collections composed primarily of extended multiplex families. The MNI, MNII, and SW collections were primarily European American (SW was Icelandic and Swedish), whereas the original OK collection consisted primarily of European American and African American pedigrees and the CA collection of European and Mexican American pedigrees.

After publication of the original OK SLE genome scan by Moser et al. (1998), the OK collection was expanded by 32 pedigrees and subsequently was reevaluated in the present study, using two newly developed model-free methods. The first method, a revision of the Haseman-Elston (HE) regression, evaluated genotypic and phenotypic information from concordant affected, concordant unaffected, and discordant sibling pairs (Elston et al. 2000). The other, a conditional logistic regression technique, used genetic information from affected-relative pairs (avuncular, grandparental, cousin, and sibling) (Olson 1999*b*). These methods allowed exploitation not only of the identical-by-descent (IBD) sharing information between affected pairs—siblings and other relatives (which are somewhat unique to our lupus pedigree collection)—but also of the sharing information between unaffected and discordant sibling pairs. These methods also freed the analysis of the limits to identification of linkages that are often imposed by an a priori specification of mode of inheritance.

The reevaluation of the expanded OK pedigree collection proved to be a successful application of two new analytical methods that resulted in identification of a novel linkage to 4p16-15.2, as well as confirmation of this linkage in an independent pedigree collection. Results of this scan identified an epistatic interaction between chromosome 4p16-15.2 and a candidate region at 5p15—the first such synergistic relationship to be shown in family studies of SLE. Finally, results replicated linkage signals at nine regions identified in independent studies and four regions identified in a subset of these data published elsewhere.

Subjects and Methods

Pedigrees

Pedigrees were enrolled in the study only after verification that at least two members met four or more of the American College of Rheumatology criteria for classification for SLE (Tan et al. 1982; Hochberg 1997) and that the relationship of the affecteds was potentially informative for linkage. After informed consent was obtained, blood samples and/or buccal swabs were collected, and genotypes were obtained for the members of 126 extended multiplex SLE pedigrees, including 744 individuals, 295 of which were classified as affected. Within these 126 pedigrees were a total of 469 sibling pairs useful for sibpair analysis, 127 African American sibling pairs, 297 European American sibling pairs, and 45 sib-pair pedigrees of Asian, Hispanic, Native American, or mixed origin (Other). A total of 175 affected full-sibling, half-sibling, avuncular, grandparental, and cousin pairs were included in the affected-relative pair analysis (table 1).

SLE is found predominantly in women, and, accordingly, our sample is enriched for female versus male affecteds (275:20). Likewise, SLE is more prevalent and severe in African Americans than in European Americans (Alarcon et al. 1999). Our sample, unlike the other collections mentioned above, is therefore enriched for African American affecteds (32% African American and 61% European American).

Data from the MNI and MNII pedigree collections were used to confirm the linkage result on chromosome 4p16-15.2 from the present study. These collections were 80% and 78% European American, respectively, and together contained 196 affected full-sibling pairs (153 European American, 17 African American, and 26 Other), 11 affected half-sibling pairs (6 European American, 4 African American, 1 Other), and 13 affected-relative pairs other than siblings (13 European American). The procedures for recruitment and genotyping of these 187 pedigrees (150 European American, 17 African American, and 20 Other) have been described elsewhere (Gaffney et al. 1998, 2000).

Genotyping

A total of 312 microsatellite markers were typed from the Version 8 Weber Screening Set, which has an average marker spacing of 11 cM. PCR was performed in 10.875-ml reaction volumes containing 5–15 ng of template DNA, 0.20 mM of M13 tailed primers (Research Genetics), 0.05 mM of IR40-labeled M13 primer (Li-Cor), 200 mM of each nucleotide, 10 mM TrisHCl (pH 8.3), 1.50 mM $MgCl_2$, and 0.375 U *Taq* DNA polymerase. Amplified fragments were detected using 6% polyacrylamide gels electrophoresed on automated Li-

Table 1

Composition of 126 Multiplex SLE Pedigrees, Separated by Ethnicity

Composition	Ethnicity (No. [%])			Total
	African American	European American	Other	
Total multiplex pedigrees	40 (31.7)	77 (61.1)	9 (7.2)	126
Total sib pairs (affected and unaffected)	127 (27.1)	297 (63.3)	45 (9.6)	469
Total nuclear families:	110 (34.7)	174 (54.9)	33 (10.4)	317
% with 1 or 2 sibs	71.8	69.1	75.7	
% with 3–5 sibs	26.4	29.8	21.3	
% with >5 sib	1.8	1.1	3.0	
Total ARPs:	52 (29.7)	112 (64.0)	11 (6.3)	175
% total full sibs	51.9	60.7	63.6	
% total half sibs	11.6	1.8	.0	
% total other pairs (avuncular,[a] grandparental, cousin)	36.5	37.5	36.4	

[a] Niece/aunt, niece/uncle, nephew/aunt, and nephew/uncle pairs.

Cor Model 4000 DNA sequencers. Gel images were collected using Base ImagIR software, version 4.0, and alleles were determined using the Gene ImagIR program, version 3.52. The Mammalian Genotyping Center performed initial genotyping of the screening markers for 29 pedigrees in Marshfield, WI (Center for Medical Genetics, Marshfield Medical Research Foundation Web site), using a fluorescent-based detection system.

Model-Free Linkage Analysis

Prior to any linkage analysis, sibling, half-sibling, and parent/child relationships were confirmed using statistics generated by RELTEST (Olson 1999*a*), a feature of the S.A.G.E. 4.0 package, version Beta 3 (1999). Multipoint linkage analysis was performed for all 22 autosomes by scanning 2-cM increments for linkage to the binary phenotype of interest (presence or absence of SLE). Multipoint IBD-sharing estimates were calculated for each of the 469 sib pairs. Then, the new HE regression method (Elston et al. 2000) was performed using the combined data set, as well as racial subsets of pedigrees (African American and European American). This method regresses the IBD-sharing values against the mean corrected cross product of the sib-pair trait differences as

$$E[(y_i - \bar{y})(y_j - \bar{y})] = \alpha + \beta\pi_{ij} ,$$

where y_i and y_j are the trait values for each sibling in a pair, $\bar{y}$ is the mean, α is the intercept, β is a parameter estimate, and π_{ij} is the IBD sharing value for the ijth pair. This ordinal measure of the dependent variable allows the phenotype to be treated as if it were continuous. As seen in the above equation, the correction of the dependent variable is dependent on a specified mean that is usually calculated from the sampled population. Because ascertainment bias may be introduced when pedigrees are collected on the basis of the presence of the desired phenotype and therefore may inflate the true population mean, we chose to analyze our data in two ways: (1) by defining $\bar{y}$ as the sample mean and (2) by defining $\bar{y}$ as the population prevalence of disease (~0). The results were similar; therefore, all results presented here were obtained using the correction factor ($\bar{y}$) set to the population prevalence of disease, to account (at least partially) for any ascertainment bias.

In addition, we modeled the covariance of all affected relative pairs as a function of marker allele–sharing IBD in the three groups defined above (All, African American, and European American), using the conditional logistic formulation implemented in the affected-relative pair (ARP) subtest (Olson 1999*b*) of S.A.G.E. 4.0 Beta 3 (1999). This method assesses linkage in affected relative pairs by constructing the following pair-specific likelihood ratio (LR) statistic:

$$\mathrm{LR} = \frac{\sum_{i=0,1,2} e^{\beta_i}\hat{f}_{ri}}{\sum_{i=0,1,2} e^{\beta_i} f_{ri}} ,$$

where e^{β_i} is equivalent to $e^{\log\ e^{\lambda_i}}$, λ_i is the relative risk to an individual who shares i alleles IBD with an affected relative, f_{ri} is the prior probability that a relative pair (type r) shares i alleles IBD, and $\hat{f}_{ri}$ is the estimated probability, conditional on the marker data, that a relative pair (type r) shares i alleles IBD. This LR is then easily transformed into a LOD score by summation of the base-10 logarithms of the above pair-specific likelihood ratios. In addition to information about linkage, the values of λ_i can also be used to infer mode of inheritance.

Lander and Kruglyak's criterion for establishing linkage in complex genetic-disease studies (Lander and Kruglyak 1995) was applied to all results obtained in the present analysis. Therefore, any result from the revised HE regression with $P \leq 2 \times 10^{-5}$ was considered

to have genomewide significance at the 5% level, as was any result from the conditional logistic regression with LOD $\geq$ 3.3. LOD scores and *P* values presented are not adjusted for genomewide significance; therefore, any region that meets this criterion is specifically noted as such.

To confirm or refute the evidence of linkage to chromosome 4p16-15.2, we applied the conditional logistic formulation method to the affected relative pairs in the combined MNI + MNII pedigree collections. We used genotypes from the same three microsatellite markers as those used in the genome scan presented here (D4S2366, D4S403, and D4S2639).

To address any possible duplication between the MNI + MNII and the present OK collections, we searched the genotype data at nine microsatellite markers and identified seven samples that were genetically identical. Upon examination of the pedigree structure, race, sex, and affection status of these samples, it was determined that five pedigrees were present in both data collections. These five families were dropped from the MNI + MNII collection prior to analysis.

Multiple regression modeling in the African American, European American, and All groups was conducted, using the revised HE regression method, to address possible interaction between regions that were suggestive of linkage. Three pairwise interactions were assessed for the European American subset, one for the African American subset, and three for the entire pedigree collection.

Finally, to better understand the genetic architecture of SLE, the locus-specific parameter estimates from the conditional logistic regression were evaluated for information on mode of inheritance (MOI). The parameter estimates were used to calculate the relative risk of monozygotic twin, offspring, and other affected relative pairs (λ_m, λ_o, and λ_r, respectively). Using these relative risks, we gained insight as to whether the model for that locus indicated a dominant or recessive MOI. That is, if λ_s ($\lambda_s = \frac{1}{4} + \frac{1}{2}\lambda_o + \frac{1}{4}\lambda_m$), the relative risk for a sibling, is approximately equal to λ_o, then it would appear that the majority of the affected sibling pairs share only one allele, and, therefore, the trait would appear dominant (Risch 1990). On the other hand, if λ_s were much larger than λ_o, then it would appear that a majority of affected sibling pairs share two alleles and the trait would therefore appear recessive (Risch 1990). Of course, these statements are true only when linkage is present in the region of interest.

Results

Scanning the genome for linkage using concordant and discordant sibpairs yielded 20 regions for which significance at the $P < .05$ level spanned, on average, 26 cM (minimum 14, maximum 42). There were 10 regions for which $.05 > P > .01$, 8 for which $.01 > P > .001$, and 2 for which $P < .001$ (table 2). The test for linkage among the ARPs identified 12 regions in which significance at a LOD of $\geq$1.5 spanned 6–26 cM, with a mean of 13 cM. Six regions were significant at LOD > 1.5, five at $2.0 <$ LOD < 3.0, and one at LOD > 3.0 (table 2). (Please note that all reported *P* values refer to results obtained using all possible sibling pairs [SIBPAL2], and all reported LOD scores refer to results obtained from ARPs.)

Evidence of Linkage at 4p16-15.2

The effect on chromosome 4p16-15.2 was among the 10 most significant results identified by the original OK genome scan (LOD = 2.18) (Moser et al. 1998), which used a subset of the same families represented here. The effect in this region is the most suggestive of linkage in the European American subset using both the SIBPAL2 ($P = .0003$, fig. 1) and ARP (LOD = 3.84, fig. 2) methods and is the most significant effect in the entire collection of 126 pedigrees when using ARPs (LOD = 3.44). Chromosome 4p16-15.2 is one of only six regions in the genome scans published to date that meets the Lander and Kruglyak (1995) threshold of genomewide significance (LOD $\geq$ 3.3 or $P \leq 2 \times 10^{-5}$) (1q22-24 [OK], 1q41 [OK] [Moser et al. 1999; Tsao et al. 1999], 2q37 [SW], 4p16-15.2 [OK], 6p21-11 [MNI + MNII], and 16q13 [MNI + MNII]).

Linkage to 4p16-15.2 is supported not only by the magnitude of the effect but also by the replication of this finding in an independent data collection. The combined MNI + MNII pedigree collection was analyzed using the conditional logistic, ARP method for the same microsatellite markers as those used in the present genome scan (D4S2366, D4S403, and D4S2639). Typing identical markers in both the OK and MNI + MNII collections allowed the direct comparison of results and avoided problems with confirmation that might arise because of marker variability and mislocalization. A LOD score of 1.5 was obtained with the European American ARPs from Minnesota. Although this result does not meet the Lander and Kruglyak criterion (1995), it is more than sufficient to be considered a significant confirmation according to the results of Xu et al. (1999), which, on the basis of analyses done using data simulated for the Genetic Analysis Workshop 11, state that a LOD $\geq$ 1.2 in any one confirmatory study is significant confirmation of a linkage effect.

Evidence of linkage to this region was not found in the African American populations of either the OK ($P = .99$, LOD = 0) or MNI + MNII (LOD = 0) data sets, which therefore suggests that linkage at 4p16-15.2 is specific to the European American subset of pedigrees. Certainly, this is not the first time that a potentially ra-

Table 2

Results of $P < .05$ or LOD > 1.5 for the OK Scan using SIBPAL2 and ARPs and Results from Other Lupus Genome Scans Significant at the Confirmatory Level of $P < .05$ or LOD > 1.2 and Located within 15 cM of the Microsatellite Used in the OK Scan

Region	Nearest Marker	Population	LOD Score; Model[a]	*P*			LOD Score Independent Study[b]
				SIBPAIR	SIBPAL2	ARP	
1q22-24[c]	D1S1679	African American	.323; DMix	.0055	.001	2.47	…
1q22-24[c]	D1S1679	All	.470; R50	.035	.009	2.75	…
2p13-11[d]	D2S1777	All	.015; R50	…	.019	…	…
2p13-q12	D2S1790	European American	.102; Rmix	…	.002	…	2.13 (SW)
2q31-34	D2S1391	All	.835; R50	.022	.012	…	…
2q35-37	D2S1363	All	.779; D50	…	.024	1.53	1.45 (MNII)
3p21-12	D3S1766	All	1.372; R100	…	.020	…	…
4p16-15.2[c]	D4S2366	European American	1.657; R100	.0032	.0003	3.84	1.50 (MNI + MNII),[e] 1.31(MNI + MNII)
4p16-15.2	D4S2366	All	1.385; R100	.0038	.002	3.44	…
4q31-33	D4S2368	European American	.809; D50	.05	.004	…	1.02 (MNII)
4q31-33	D4S2368	All	1.221; Rmix	.021	.018	…	…
5p15	D5S807	European American	.927; Dmix	…	.005	…	1.52 (SW)
5p15	D5S807	All	1.324; Dmix	…	.0005	1.75	…
6p24-23[d]	D6S2434	All	.628; D50	…	.006	1.54	…
6p24-23	D6S2434	African American	.601; D50	…	.009	2.06	1.58 (MNI + MNII)
6p22-21	D6S2439	African American	.155; D90	…	.011	1.70	1.54 (SW), 1.48(MNII)
6p12-q14	D6S1053	African American	1.186; R50	.007	…	2.36	3.9 (MNI)
7q36	D7S559	European American	.538; D50	…	.014	…	2.15 (MNII)
7q36	D7S559	All	.086; D90	…	.023	…	…
9p24-21	D9S925	African American	.113; Rmix	…	.004	2.08	…
9q21-32[d]	D9S910	African American	.200; R50	…	.012	…	…
12p12-11[c]	D12S1042	European American	1.946; D50	.014	…	1.79	…
12q24[c]	D12S395	European American	1.802; R50	.004	.009	2.04	…
12q24	D12S395	All	1.795; Rmix	.003	.006	…	…
17p13.3	D17S1298	European American	.690; R50	.03	…	1.57	…
17p11-q21[d]	D17S1299	European American	.864; R50	.011	.027	…	…
17p11-q21	D17S1299	All	.687; Rmix	.032	.009	…	…
17q21	D17S2180	African American	1.670; Rmix	.011	.015	…	…
21q21.1	D21S1437	African American	.707; D50	…	.027	…	1.58 (SW)
21q21.3-22[d]	D21S1440	African American	1.240; R50	.012	.010	…	…

[a] Models used in traditional LOD score analysis, outlined in Moser et al. (1998).

[b] Results for MNI, MNII, or MNI + MNII were found in the combined pedigree collection (African American and European American), with the exception of 4p16-15.2, which was specific to the European American families. Results for SW were found in the Swedish families, with the exception of 21q21.1, which was identified in the Icelandic families.

[c] Contain loci identified in the same racial group in Moser et al. (1998).

[d] Linkage signals not previously identified.

[e] Result obtained using analyses described herein.

cially specific linkage has been observed in the genetic analysis of SLE family data. Evidence of linkage to 1q22-24, for example, is the strongest in African Americans using both the present analysis techniques and the expanded data set ($P = .009$, LOD = 2.75), as well as using a traditional LOD-score method with either all 126 pedigrees (LOD = 3.97) or a subset of the present data (LOD = 3.45) (Moser et al. 1998). However, this effect is much less significant in the European Americans, regardless of method of analysis or size of data set. Racially specific linkages may account, at least in part, for the difficulty in replication of results between pedigree collections.

Using the parameter estimates produced by the conditional logistic regression technique, the mode of inheritance for the linkage at 4p16-15.2 was inferred for both the OK and the MNI + MNII data sets. The relative risk estimates (λ_s) for affected sibling pairs were 2.16 and 1.22 at the peak of the linkage signal for the OK and MN data sets, respectively. The relative risk estimates for affected parent–offspring pairs (λ_o) were 1.62 and 1.00, respectively. Because the sibling sharing estimates (λ_s) were approximately equal, in both pedigree collections, to that of the parent-offspring sharing estimate (λ_o), the linkage at 4p16-15.2 appears to follow a dominant rather than a recessive mode of inheritance.

Identification of Epistasis at 4p16-15.2

Choosing from those regions considered to be suggestive of linkage in the expanded OK pedigree collection, we identified six regions with which to perform

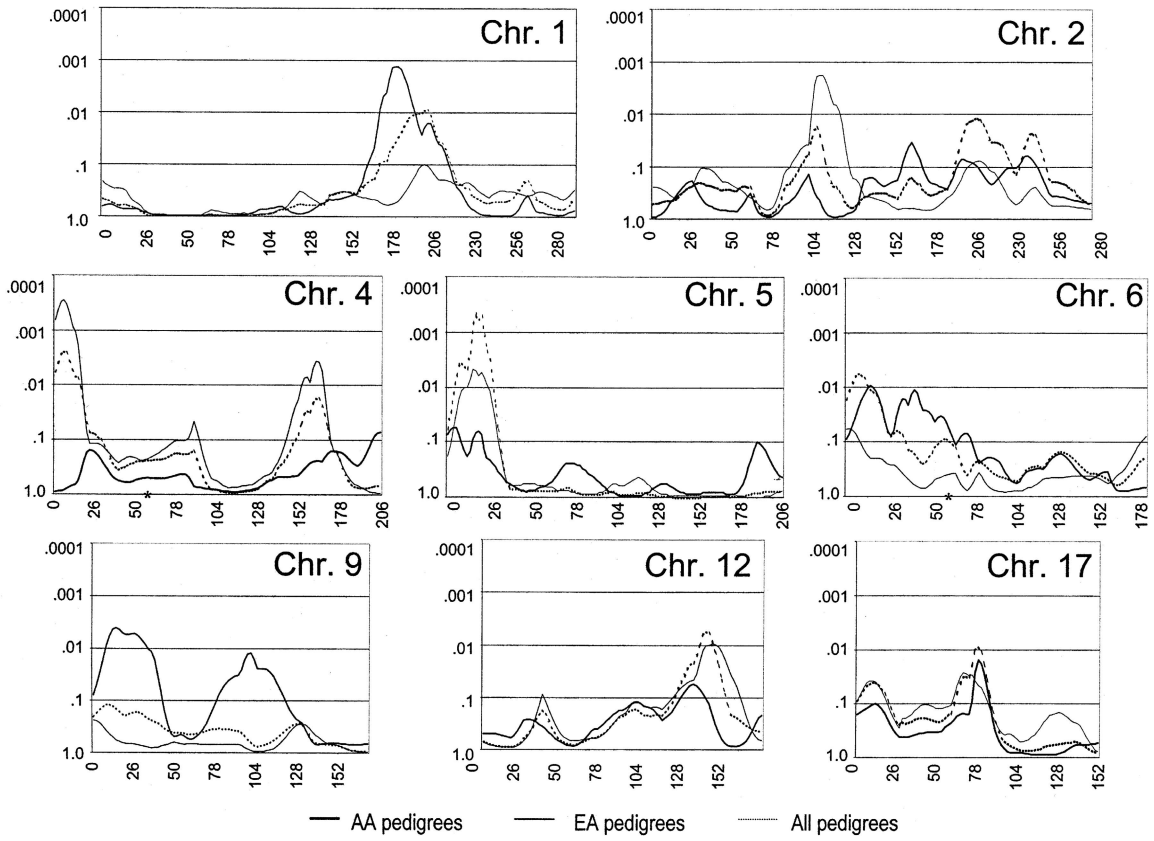

Figure 1 Chromosomes containing regions with SIBPAL2 significance <.01. Centimorgans are plotted on the *X*-axis and *P* values on the *Y*-axis. Asterisks(*) indicate locations of other linkage signals identified on chromosomes 4 and 6 (Gaffney et al. 1998; Lindqvist et al. 2000).

multiple-regression modeling (only results from the revised HE regression were considered, since it is with this method that we would assess interaction) (table 3). Inclusion in the multiple-regression modeling was based on both the magnitude ($P < .01$) of the result and support from previous studies (table 2). There were three regions chosen in the European American subset (4p16-15.2, 4p31-33, and 5p15), two in the African American subset (1q22-24 and 6p23-22), and three in the combined set (1q22-24, 5p15, and 6p24-23).

The European American subset model, including 4p16-15.2 ($P = .343$), 5p15 ($P = .591$), and their interaction (4p16-15.2 × 5p15), indicated a multiplicative relationship between the two regions. In addition to a significant *P* value for the interaction term in the model ($P = .038$), the parameter estimates (relative variance component measures) for the main effects in the presence of the interaction term also support an interaction. Because the relative variance components for the main effects (individual regions) are much less than that of the interaction (.0737, .000, and .3095, respectively), the global variance in the model can be attributed almost entirely to the interaction.

Identification of New Candidate Linkages

Three regions identified in this genome scan and significant at $P < .01$ had not been explicitly identified elsewhere. The effects on chromosome 12q24 in the European American subset ($P = .009$, LOD = 2.04) and on chromosome 17p11-q21 in all 126 pedigrees ($P = .009$) are unique to the present study and are not in the proximity of other signals that have been identified elsewhere. The effect on chromosome 6p24-23 (table 2, $P = .006$), found when analyzing all 126 pedigrees together, is within 10 cM of a region originally identified in the MNI scan (6p21-11) (Gaffney et al. 1998) (figs. 1 and 2). Although a definitive conclusion regarding the independence of the effects on chromosome 6p cannot yet be made, fine-mapping studies and multipoint analyses of this region should aid in making this determination.

Significant Effects Identified Elsewhere

Results from the present scan support linkage in 10 regions identified in independent studies elsewhere (MNI, MNII, and SW) (table 2). The effect on chro-

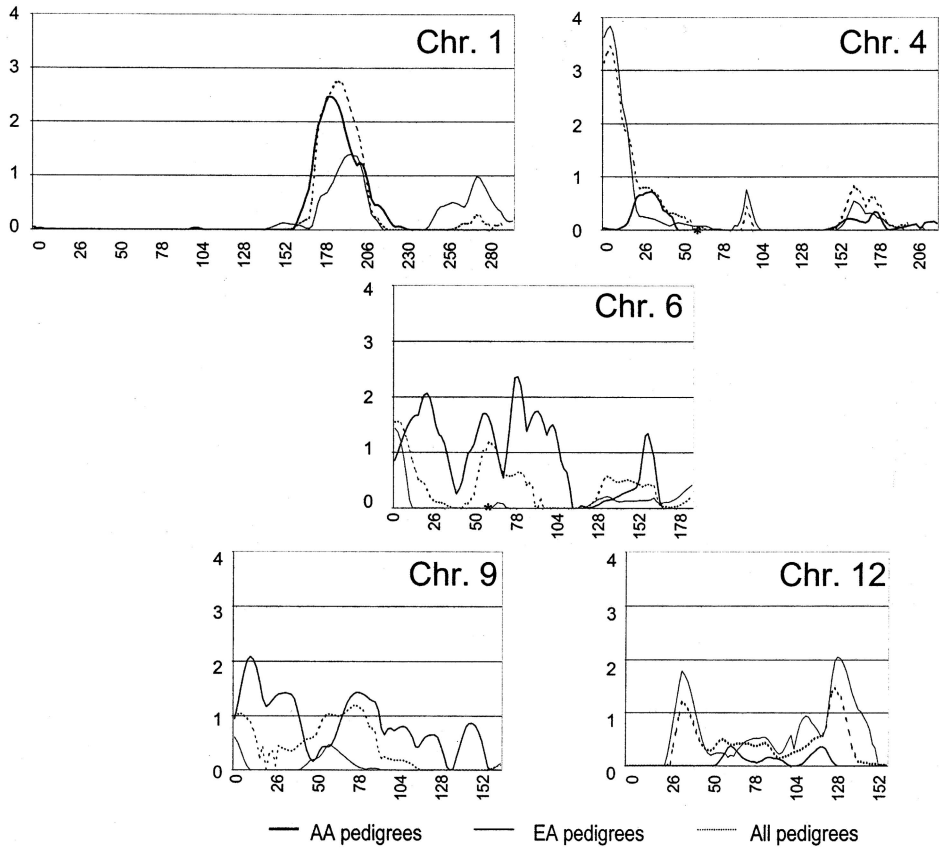

Figure 2 Chromosomes containing regions with ARP LOD scores >2.0. Centimorgans are plotted on the *X*-axis and LOD scores on the *Y*-axis. Asterisks (*) indicate locations of other linkage signals identified on chromosomes 4 and 6 (Gaffney et al. 1998; Lindqvist et al. 2000).

mosome 5p15, the most significant result in the combined data set for SIBPAL2 ($P = .0005$, fig. 1) and ARPs (LOD = 1.75, table 2) was identified independently in the SW scan (Lindqvist et al. 2000) (table 2). It is important to note that, although the effect on chromosome 5p15 is not the largest in magnitude for the combined data set, it remains the only effect that cannot be predominantly attributed to the effect of either the African American or European American subset. One region, 6p22.1-21, was identified in the present study, as well as in the SW (Lindqvist et al. 2000) and MNII (Gaffney et al. 2000) scans (table 2). Not suprisingly, this region encompasses HLA components such as C2, C4A, C4B, HLA-B, HLA-DR, and HLA-DQ, all of which have been associated with SLE in cohort-comparison studies (Schur et al. 1982).

The most suggestive linkage effect in the African American subset for each of the methods used in the present study was 1q22-24 ($P = .001$, fig. 1; LOD = 2.47, fig. 2). This region, like 4p16-15.2, was identified in the original OK scan (Moser et al. 1998). In fact, 1q22-24 overlaps a region containing several candidate genes, one of which, FcγRIIA, is the most significant single-point linkage result for the African American subset in the aforementioned scan (LOD = 3.37) (see Moser et al. 1998). The magnitude of the results obtained at 1q22-24 in the present analyses did not vary greatly with the addition of the FcγRIIA candidate gene into the multipoint analysis ($P = .004$, LOD = 2.36).

The identification of new linkage signals, as well as increased evidence for linkage to previously detected regions, could be attributable to the increase in sample size from the original OK data set to the current. However, a traditional LOD-score and sib-pair screen of the expanded collection (methods outlined in Moser et al. 1998) either failed to detect or detected at a lesser significance level all regions except those on chromosome 12 and 17, which are similar in significance across methods (table 2).

Discussion

The results of this study support the presence, in European American populations, of an SLE susceptibility locus on chromosome 4p16-15.2. The results further support the involvement of this susceptibility locus in a

Table 3

Results of Models Using One Marker (Single Regression) and Two Markers with Their Interaction (Multiple Regression) for the Concordant and Discordant Sibling Pairs in the Expanded OK Pedigree Collection

	Parameter Estimates		
Pedigree Subset and Region	Single Regression	Multiple Regression	*P*
European American:			
4p16-15.2	.2250	.1577	.1050
4q31-33	.1576	.0736	.3020
4p16-15.2 × 4q31-33		.1083	.3150
4p16-15.2	.2250	.0737	.3430
5p15	.1660	.0000	.5910
4p16-15.2 × 5p15		.3095	.0380
4q31-33	.1576	.1429	.1140
5p15	.1660	.1541	.0990
4q31-33 × 5p15		.0526	.4000
African American:			
1q22-24	.3149	.3008	.0440
6p24-23	.2650	.2454	.0780
1q22-24 × 6p23-22		.0000	.6060
All pedigrees:			
1q22-24	.1332	.1202	.0320
6p24-23	.1309	.1767	.0360
1q22-24 × 6p24-23		.0000	.7640
1q22-24	.1332	.1311	.0020
5p15	.1733	.1304	.0002
1q22-24 × 5p15		.0000	.9840
6p24-23	.1309	.0739	.2390
5p15	.1733	.1317	.1050
6p24-23 × 5p15		.1770	.3010

synergistic relationship with a potential susceptibility locus on chromosome 5p15 in an European American population. Finally, results support the continued application of these analytical methods to both new and previously collected pedigree data for the purposes of identifying previously undetected linkage signals and confirming strong but previously unreplicated results.

The results of two previous genome scans suggest linkage of SLE to chromosome 4. Gaffney et al. (2000) identified a moderate linkage signal at D4S424 (LOD = 1.50), and Lindqvist et al. (2000) presented evidence of linkage to D4S1627 (LOD = 3.20). However, both of these signals appear independent of the novel linkage mapped to 4p16-15.2 (figs. 1 and 2). D4S424 has been localized to 4q33, a band on the long arm of chromosome 4 and >150 cM from 4p16-15.2. D4S1627, although labeled in the SW scan as 4p15, has actually been localized to 4p13, the most telomeric band on the short arm of chromosome 4 and is >40 cM from 4p16-15.2. (Marker localization is according to the following standardized genetic databases: Center for Medical Genetics, Marshfield Medical Research Foundation; Cooperative Human Linkage Center; Genome Database; LDB, Genetic Location Database; and Whitehead Institute for Biomedical Research/MIT Center for Genome Research.)

The novel linkage identified at chromosome 4p16-15.2 is among the most convincing candidate linkages in lupus to date. Although the evidence (LOD = 3.84 and 1.5) is quite different for the OK and MNI + MNII collections respectively, it is typical for a confirmatory, targeted marker study (as opposed to a genome screen), to have a smaller effect, since no selection was made on the basis of the magnitude of the result. Accordingly, this effect both exceeds the threshold of LOD = 3.3 set by Lander and Kruglyak (1995) for the data set in which it was first identified, as well as exceeds the LOD = 1.2 criteria touted as necessary for confirmation by Xu et al. (1999). It is one of only six SLE candidate linkages that are significant at LOD = 3.3 (Lander and Kruglyak 1995) (1q22-24 [OK], 1q41 [OK, UCLA] [Tsao et al. 1999]), 2q37 (SW), 4p16-15.2 (OK), 6p21-11 (MNI + MNII), and 16q13 (MNI + MNII) and is one of only three regions to also meet the confirmation criterion set forth by Xu et al. (1999) (1q41[OK, UCLA, MNI + MNII], 4p16-15.2 (OK and MNI + MNII), and 6p21-11 (MNI + MNII, OK, SW).

Almost 70 genes are currently mapped to chromosome 4p16-15.2, 21 of which have either been associated with an autoimmune disease or appear to be potentially involved in autoimmune processes. This set includes several genes that are directly involved with B and T lymphocyte activity—namely, CD38, BST1, ZNF36, ZNF134, and ZNF136. Furthermore, a gene linked to lupus nephritis in an NZBxW murine model system (*sle6*) maps to the centromeric region of mouse chromosome 5 (Morel et al. 1999). This region is syntenic to human chromosome 4p16-15 (placement based on the standardized Mouse Genome Informatics database).

All complex genetic diseases potentially involve gene-gene interactions, at least to some extent. SLE is no exception. One such interaction between IL-10 and bcl-2 was proposed (Mehrian et al. 1998) on the basis of cohort comparison, but a gene-gene interaction based on family data has not yet been identified. However, we present the first evidence to date of a locus-locus interaction involved in lupus. Although a biological model underlying this interaction cannot be determined at this time, we propose, on the basis of the modeling techniques used, that the relationship between markers at 4p16-15.2 and 5p15 is epistatic rather than additive in nature.

The possibility that there are several potential SLE loci yet to be identified is only reinforced by the continued identification of new candidate linkages. For example, three regions that were not identified by the five preceding scans (Moser et al. 1998; Gaffney et al. 1998, 2000; Shai et al. 1999; Lindqvist et al. 2000)

were found to be significant at $P < .01$ (6p24-23, 12q24, and 17p11-q21) and $P < .05$ (2p13-11, 9p24-21, and 21q21.3-22) in the present analysis (table 2). Certainly, increased type I–error rates must be considered as the number of analyses increase. However, identification of new candidate regions such as those found in the present scan are perhaps a result of not only increased sample size but also improved analytical techniques that better maximize the information available from multiplex pedigrees.

In attempting to reconcile the various results from genetic-linkage studies of SLE, one must consider the clinical heterogeneity of the disease. SLE is highly variable among individual patients and, to some extent, between ethnic groups, since certain disease manifestations seem to cluster within racial populations, whereas others do not (Petri 1998). It is therefore not unreasonable to believe that there are genes specific to either a particular racial group or to a particular clinical manifestation. Work by S. Rao, J. M. Olson, C. Gray-McGuire, G. R. Bruner, J. Kelly, K. L. Moser, and J. B. Harley (unpublished data), using the expanded OK collection, assesses linkage to various clinical manifestations of SLE and compares the occurrence of those manifestations both between and within families. The results of this study show that there are indeed linkages to particular lupus manifestations, without regard to racial affiliation. This suggests that the inability to replicate a result within a particular racial subset does not necessarily refute linkage at that locus to lupus but, instead, could indicate a linkage that is specific to a phenotypic rather than a racial subset.

These results (table 2) also illustrate the dependence of many of the traditional linkage methods on prior knowledge of genetic mechanism. For example, 4p16-15.2, a region where linkage has now been established and confirmed, was not found at this magnitude, even in the same data set, using other methods. Certainly, it is no surprise that results differ by method applied, and, although there is a statistical penalty that should be considered when applying multiple methods to the same set of marker data, the results herein support the view that important linkages may be missed when analysis is restricted to a single method.

In conclusion, results of this genome scan support the involvement of a SLE susceptibility locus on chromosome 4p16-15.2 in an European American population. This result reaches a significance level attained by only one other lupus candidate linkage to date. Our results support an epistatic relationship between this susceptibility locus and a potential susceptibility locus on chromosome 5p15. Finally, although a discussion of the power of the methods used in this scan is beyond the scope of this paper, results do support the application of these and other new techniques to previously collected and recently expanded pedigree collections both to identify new candidate linkages and to confirm strong but unacknowledged linkages.

Acknowledgments

We are thankful for all the patients and family members who participated in this study, as well as the many referring physicians and the U.S. Department of Veterans Affairs. We are grateful for the contributions of Gail Bruner, Sharon Johnson, Debra Bacino, and Gerald Koelsch. This work was supported by grants from the National Institutes of Health (AR42460, AI24717, AR45231, AI31584, HG01577, and RR03655). The majority of this work was conducted at the Robert H. and Lynnie Spahn Center for Genotyping at the Oklahoma Medical Research Foundation. Some of the results in this report were obtained by use of S.A.G.E., which is supported by NCRR grant RR03655. Sixty-seven pedigrees (cohorts A and B) were obtained from the Lupus Multiplex Registry and Repository (AR-5-2221) (see http://omrf.ouhsc.edu/lupus).

Electronic-Database Information

The accession number and URLs for data in this article are as follows:

Center for Medical Genetics, Marshfield Medical Research Foundation, http://research.marshfieldclinic.org/genetics/
Cooperative Human Linkage Center, The, http://www.chlc.org/
Genome Database, The, http://gdbwww.gdb.org/
LDB, Genetic Location Database, http://cedar.genetics.soton.ac.uk/public_html/ldb.html
Mouse Genome Informatics, http://www.informatics.jax.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/omim (for systemic lupus erythematosus, MIM 152700)
Whitehead Institute for Biomedical Research/MIT Center for Genome Research, http://www.genome.wi.mit.edu/

References

Alarcon GS, Friedman AW, Straaton KV, Moulds JM, Lisse J, Bastian HM, McGwin G Jr, Bartolucci AA, Roseman JM, Reveille JD (1999) Systemic lupus erythematosus in three ethnic groups. III. A comparison of characteristics early in the natural history of the LUMINA cohort. LUpus in MInority populations: NAture vs. Nurture. Lupus 8:197–209

Brand A, Griffiths DJ, Herve C, Mallon E, Venables PJ (1999) Human retrovirus-5 in rheumatic disease. J Autoimmun 13: 149–154

Cooper GS, Dooley MA, Treadwell EL, St Clair EW, Parks CG (1998) Hormonal, environmental, and infectious risk factors for developing systemic lupus erythematosus. Arthritis Rheum 41:1714–1724

Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P, Walker A, Mack TM (1992) A revised es-

timate of twin concordance in systemic lupus erythematosus. Arthritis Rheum 35:311–318

Elston RC, Buxbaum S, Jacobs KB, Olson JM (2000) Haseman and Elston revisited. Genet Epidemiol 19:1–17

Friday RP, Trucco M, Pietropaolo M (1999) Genetics of type 1 diabetes mellitus. Diabetes Nutr Metab 12:3–26

Gaffney PM, Kearns GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML, Rohlf KE, Ockenden TC, Messner RP, King RA, Rich SS, Behrens TW (1998) A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families. Proc Natl Acad Sci 95:14875–14879

Gaffney PM, Ortmann WA, Selby SA, Shark KB, Ockenden TC, Rohlf KE, Walgrave NL, Boyum WP, Malmgren ML, Miller ME, Kearns GM, Messner RP, King RA, Rich SS, Behrens TW (2000) 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 66:547–556

Hochberg MC (1997) Updating the American College of Rheumatology revised criteria for systemic lupus erythematosus. Arthritis Rheum 40:1725

James JA, Kaufman KM, Farris AD, Taylor-Albert E, Lehman TJ, Harley JB (1997) An increased prevalence of Epstein-Barr virus infection in young patients suggests a possible etiology for systemic lupus erythematosus. J Clin Invest 100: 3019–3026

Kaslow RA, Masi AT (1978) Age, sex, and race effects on mortality from systemic lupus erythematosus in the United States. Arthritis Rheum 21:473–479

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Lawrence JS, Martins CL, Drake GL (1987) A family survey of lupus erythematosus 1 heritability. J Rheum 14:913–921

Lindqvist AKB, Steinsson K, Johanneson B, Kristjansdottir H, Arnasson A, Grondal G, Jonasson I, Magnusson V, Sturfelt G, Truedsson L, Svenungsson E, Lundberg I, Terwilliger JD, Gyllensten UB, Alarcon-Riquelme ME (2000) A susceptibility locus for human systemic lupus erythematosus (hSLE1) on chromosome 2q. J Autoimmunity 14:169–178

Mehrian R, Quismorio FP Jr, Strassmann G, Stimmlet MM, Horwitz DA, Kitridou RC, Gauderman WJ, Morrison J, Brautbar C, Jacob CO (1998) Synergistic effect between IL-10 and bcl-2 genotypes in determining susceptibility of systemic lupus erythematosus. Arthritis Rheum 41:596–602

Morel L, Tian X-H, Crocker BP, Wakeland EK (1999) Epistatic modifiers of autoimmunity in a murine model of lupus nephritis. Immunity 11:131–139

Moser KL, Gray-McGuire C, Kelly J, Asundi N, Yu H, Bruner GR, Mange M, Hogue R, Neas BR, Harley JB (1999) Confirmation of genetic linkage between human systemic lupus erythematosus and chromosome 1q41. Arthritis Rheum 42: 1902–1907

Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Asundi A, Bruner GR, Fox J, Kelly J, Henshall S, Bacino D, Dietz M, Hogue R, Koelsch G, Nightingale L, Shaver T, Abdou NI, Albert DA, Carson C, Petri M, Treadwell EL, James JA, Harley JB (1998) Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees. Proc Natl Acad Sci 95: 14869–14874

Olson JM (1999*a*) Relationship estimation by Markov-process models in a sib-pair linkage study. Am J Hum Genet 64: 1464–1472

——— (1999*b*) A general conditional-logistic model for affected-relative-pair linkage studies. Am J Hum Genet 65: 1760–1769

Petri M (1998) The effect of race on incidence and clinical course in systemic lupus erythematosus: the Hopkins lupus cohort. J Am Med Womens Assoc 53:9–12

Reed PW, Davies JL, Copeman JB, Bennett ST, Palmer SM, Pritchard LE, Gough SC, Kawaguchi Y, Cordell HJ, Balfour KM, Jenkins SC, Powell EE, Vignal A, Todd JA (1994) Chromosome-specific microsatellite sets for fluorescent-based, semi-automated genome mapping. Nat Genet 7:390–395

Reichlin M, Harley JB, Lockshin MD (1992) Serologic studies of monozygotic twins with systemic lupus erythematosus. Arthritis Rheum 35:457–464

Risch N (1990) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229–241

S.A.G.E. (1999) Statistical analysis for genetic epidemiology, release 4.0, Beta 3. Computer package, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH

Schur PH, Meyer I, Garavoy M, Carpenter CB (1982) Associations between systemic lupus erythematosus and the major histocompatibility complex: clinical and immunological considerations. Clin Immunol Immunopathol 24:263–275

Sestak AL, Timothy S, Moser KL, Neas BR, Harley JB (1999) Familial aggregation of lupus and autoimmunity in an unusual multiplex pedigree. J Rheum 26:1495–1499

Shai R, Quismorio FP, Li L, Kwon OJ, Morrison J, Wallace DJ, Neuwelt CM, Brautbar C, Gauderman WJ, Jacob CO (1999) Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families. Hum Mol Genet 8:639–644

Tan FK, Arnett FC (1998) The genetics of lupus. Curr Opin Rheum 10:399–408

Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfeild NF, Schaller JG, Talal N, Winchester RJ (1982) The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 25:1271–1277

Tsao BP, Cantor RM, Grossman JM, Shen N, Teophilov NT, Wallace DJ, Arnett FC, Hartung K, Goldstein R, Kalunian KC, Hahn BH, Rotter JI (1999) PARP alleles within the linked chromosomal region are associated with systemic lupus erythematosus. J Clin Invest 103:1135–1140

Tsuang MT, Stone WS, Faraone SV (1999) Schizophrenia: a review of genetic studies. Harv Rev Psychiatry 7:185–207

Xu J, Weisch DG, Eugene WT, Meyers DA (1999) Evaluation of replication studies, combined data analysis, and analytical methods in complex diseases. Genet Epidemiol 17:S773–S778