Research article

Genetic linkage and transmission disequilibrium of marker haplotypes at chromosome 1q41 in human systemic lupus erythematosus

Robert R Graham, Carl D Langefeld*, Patrick M Gaffney, Ward A Ortmann, Scott A Selby, Emily C Baechler, Katherine B Shark, Theresa C Ockenden, Kristine E Rohlf, Kathleen L Moser, William M Brown*, Sherine E Gabriel†, Ronald P Messner, Richard A King, Pavel Horak‡, James T Elder§, Philip E Stuart§, Steven S Rich* and Timothy W Behrens

Department of Medicine, Center for Immunology, University of Minnesota Medical School, Minneapolis, MN, USA
*Department of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, USA
†Department of Health Sciences Research and Division of Rheumatology, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA
‡Interni Klinica, Olumouc, Czech Republic
§Department of Dermatology, University of Michigan Medical School, Ann Arbor, MI, USA

Correspondence: Dr Timothy W Behrens, University of Minnesota Medical School, 6-126 BSBE Bldg, 312 Church Street SE, Minneapolis, MN 55455, USA. Tel: +1 612 625 4485; fax +1 612 625 2199; e-mail: behre001@umn.edu

Received: 6 March 2001
Revisions requested: 22 May 2001
Revisions received: 30 May 2001
Accepted: 19 June 2001
Published: 17 July 2001

Arthritis Res 2001, **3**:299–305

(Print ISSN 1465-9905; Online ISSN 1465-9913)

Abstract

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease characterized by the production of autoantibodies to a wide range of self-antigens. Recent genome screens have implicated numerous chromosomal regions as potential SLE susceptibility loci. Among these, the 1q41 locus is of particular interest, because evidence for linkage has been found in several independent SLE family collections. Additionally, the 1q41 locus appears to be syntenic with a susceptibility interval identified in the NZM2410 mouse model for SLE. Here, we report the results of genotyping of 11 microsatellite markers within the 1q41 region in 210 SLE sibpair and 122 SLE trio families. These data confirm the modest evidence for linkage at 1q41 in our family collection (LOD = 1.21 at marker D1S2616). Evidence for significant linkage disequilibrium in this interval was also found. Multiple markers in the region exhibit transmission disequilibrium, with the peak single marker multiallelic linkage disequilibrium noted at D1S490 (pedigree disequilibrium test [PDT] global *P* value = 0.0091). Two- and three-marker haplotypes from the 1q41 region similarly showed strong transmission distortion in the collection of 332 SLE families. The finding of linkage together with significant transmission disequilibrium provides strong evidence for a susceptibility locus at 1q41 in human SLE.

Keywords: 1q41, autoimmunity, linkage, systemic lupus erythematosus, transmission disequilibrium

Introduction

Systemic lupus erythematosus (SLE [MIM 152700] [1]) is an autoimmune disease characterized by the production of autoantibodies with specificity for a wide range of self-antigens. These antibodies cause disease directly by binding to target tissues (e.g. platelets and phospho-

ADPRT = adenosine diphosphate ribosyltransferase; cM = centiMorgan; C-TDT = combined S-TDT and TDT; ESRRG = estrogen-related receptor gamma; HLA = human leukocyte antigen; kb = kilobasepairs; LOD = logarithm of odds ratio; MIM = Mendelian Inheritance in Man (database; see [1]); PARP = poly ADP ribosyl polymerase; PCR = polymerase chain reaction; PDT = pedigree disequilibrium test; SLE = systemic lupus erythematosus; S-TDT = discordant-sib TDT; TDT = transmission disequilibrium test; T:NT = ratio of transmitted to nontransmitted alleles.

lipids), and indirectly by depositing immune complexes in vascular tissues, leading to organ damage [2]. The current evidence suggests that SLE is a complex genetic disease, with contributions from both environmental (e.g. ultraviolet light, viral infections) and genetic factors.

A number of recent genome-wide screens support the genetic hypothesis in human SLE, and have reported, in aggregate, 48 potential susceptibility loci [3–7] (reviewed [8]). Of these intervals, significant attention has been directed at the 1q41–42 chromosomal region. After the identification of the syntenic region in mouse (*Sle1*) as a susceptibility interval for murine SLE [9], a targeted study by Tsao *et al* [10] was the first to implicate this locus in human SLE (marker D1S229; $P < 0.005$). Subsequently, two independent SLE genome scans also found suggestive evidence for linkage at 1q41–42 [3–5]. Poly ADP ribosyl polymerase (PARP; also referred to as ADPRT [adenosine diphosphate ribosyltransferase]) was initially suggested as the relevant candidate gene in the region [11,12], but that finding was not reproduced in other family collections [13].

The work of Moser *et al* [14] provided the initial evidence that the susceptibility locus in this region might lie centromeric of PARP. In 127 multiplex SLE pedigrees, the best evidence for linkage was at D1S229, with the greatest extent of allele sharing in the white families at the D1S2616 marker. The data reported here confirm the evidence for genetic linkage at 1q41 in 210 SLE sibpair families (families with a pair of affected sibs) from the Minnesota collection [3,4]. Furthermore, we report new evidence for transmission disequilibrium of both single marker alleles and short marker haplotypes from the 1q41–42 interval in 122 trio families (families with a single affected offspring and both parents) and in the total dataset of 332 SLE families. These data suggest that a human SLE susceptibility locus is located centromeric to PARP near the D1S490 marker.

Materials and methods

SLE family collection

The collection of 187 affected sibpair families in Cohorts 1 and 2 of the Minnesota collection have been described in detail elsewhere [3,4]. An additional group of 24 sibpair families was collected and members were genotyped for the 11 markers in the 1q41–42 region. One hundred and twenty-two trio families were also collected. This study was approved by the Human Subjects Review Board at the Mayo Clinic and at the University of Minnesota. The clinical features of the sibpair and trio families are provided in the Supplementary material.

Samples and genotyping

Genomic DNA isolation and genotyping of families was performed as described [3,4]. The 12 markers from the 1q41–42 region originally typed by Moser *et al* [14] were typed in the Minnesota collection. Marker order (Fig. 1) was determined using the public databases described in the Supplementary material.

Data analysis

See Supplementary material for details of linkage analysis and transmission disequilibrium testing.

Results

Genetic linkage in the 1q41–42 region

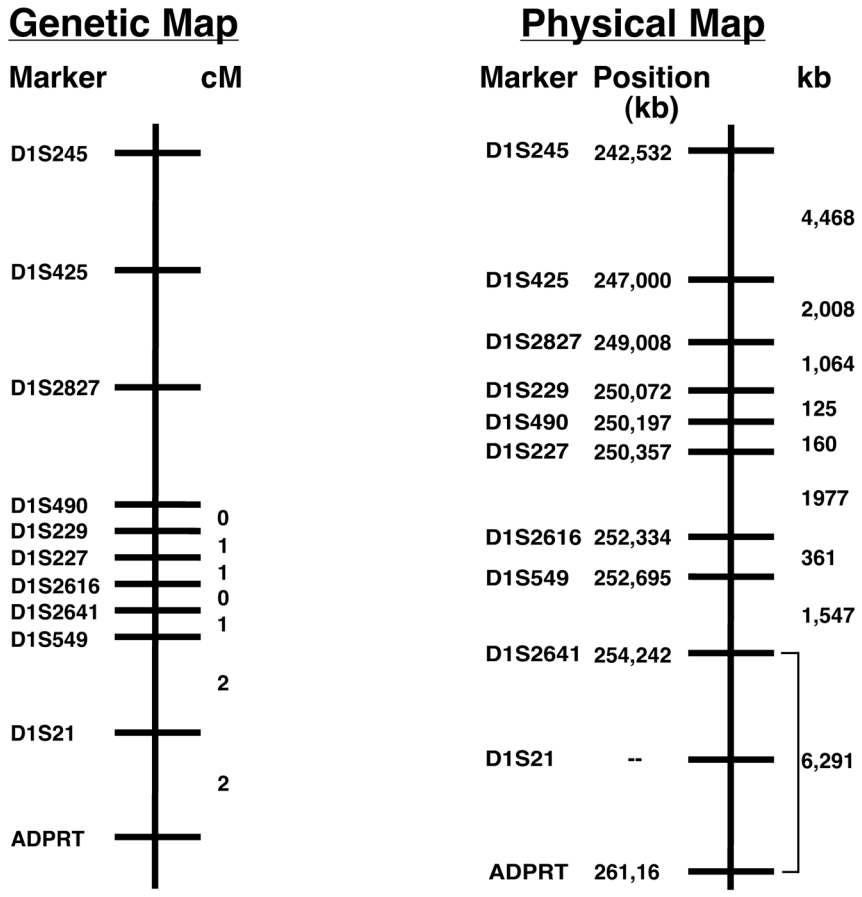
In the combined data from the original Cohorts 1 and 2 genome screens, the highest LOD score in the 1q41–42 region was 1.33 at marker D1S229 [4]. With the addition of five new fine-map markers and 24 new families, the best evidence for linkage was slightly reduced (LOD = 1.23) and shifted telomeric to marker D1S2616 (Fig. 2). This level of linkage support corresponds roughly to a nominal *P* value of 0.05. When only white families were considered, the peak multipoint LOD score remained at marker D1S2616 (LOD = 0.99; Fig. 2).

Transmission disequilibrium in the 1q41–42 region

The standard transmission disequilibrium test (TDT) [15], applied to the data for a single marker allele, failed to identify significant transmission disequilibrium (Fig. 3). Analyses with the C-TDT (the TDT combined with the discordant sib TDT [S-TDT]), which allows additional families containing discordant sibs to be evaluated for disequilibrium [16], were then performed. Among the 60 families with discordant siblings but missing parental information, several microsatellite alleles in the 1q41 region demonstrated significant evidence of transmission disequilibrium (Fig. 3). Markers D1S490-allele 6 ($P = 0.0044$), D1S229–allele 7 ($P = 0.0054$), and D1S227–allele 5 ($P = 0.0084$) showed the strongest evidence for disequilibrium using the C-TDT.

The TDT and C-TDT are limited by the fact that only one triad or discordant sibship is analyzed from a given family; thus these two tests examine only a subset of the data from general pedigrees. The pedigree disequilibrium test (PDT) improves upon these two tests by generating a measure of linkage disequilibrium for every discordant sibship and triad in a pedigree [17]. Because it examines the entire data set, the PDT can result in much higher power than the other two tests [17]. The version of the PDT used in this analysis has been modified to accept dyads (a single typed parent and an affected offspring). A dyad is informative only when the offspring has a heterozygous genotype different from that of the parent, thus avoiding the bias of using inferred phase assignments [18,19]. Using the PDT, three single marker alleles of D1S490 exhibited evidence of transmission disequilibrium (Fig. 3). Consistent with the C-TDT results, D1S490-allele 6 yielded the strongest PDT result ($P = 0.0071$).

Figure 1



Genetic and physical maps of the 1q32–42 region. The genetic map is based on the maps available at Marshfield, WI, USA (http://research.marshfieldclinic.org/genetics) and the Genome Database (http://www.gdb.org). The physical map was determined from the sequence data generated by the Human Genome Project (as described in the text), and is available from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). The physical distance between the markers was estimated from the available sequence data if the markers were not on the same contig, and the maps are not drawn to scale. The physical location of D1S213 between D1S2641 and ADPRT is not known.

Transmission disequilibrium of marker haplotypes in the 1q41–42 region

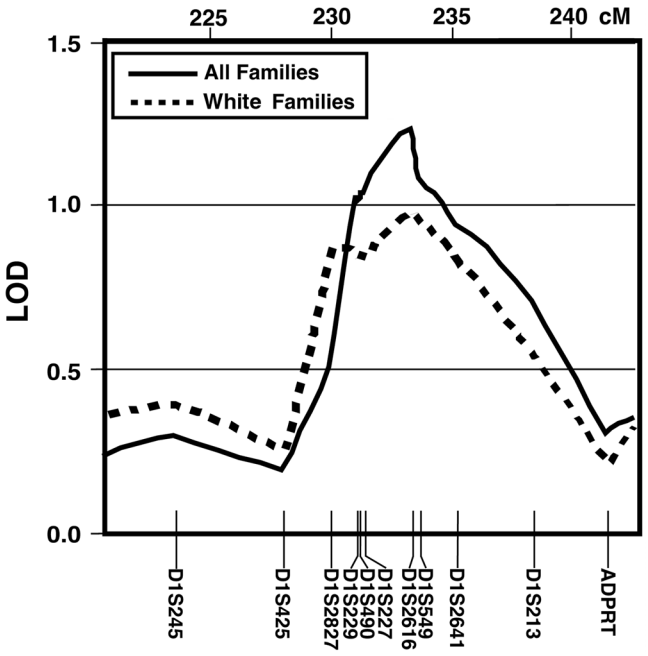
Two- and three-marker haplotypes from the 1q41–42 region were next tested for evidence of transmission disequilibrium (Fig. 3). Haplotype clusters were generated for 'exact match' haplotypes with more than 10 founders. Analysis of two- and three-marker haplotypes revealed that haplotypes from D1S490 to D1S549 had the greatest number of transmissions and yielded the most significant *P* values, with the majority of the significant haplotypes centered on D1S227. The two-marker haplotypes D1S425–allele 10/D1S2827–allele 11, D1S490–allele 7/D1S227–allele 5, and D1S2616–allele 6/D1S549–allele 12 were significant in the TDT, C-TDT, and PDT. The D1S490–allele 7/D1S227–allele 5 and D1S2616–allele 6/D1S549–allele 12 haplotypes had 63% (42:25) and 61% (52:33) transmitted:nontransmitted ratios (T:NT), respectively.

Using three-marker haplotypes, the strongest evidence for disequilibrium was in the region spanning the markers D1S229 to D1S2641. Notably, the D1S229–allele 7/D1S490–allele 7/D1S227–allele 5 (T:NT = 23:6, 79% T) and the D1S2616–allele 6/D1S549–allele 12/D1S2641–allele 10 (T:NT = 18:6, 75% T) haplotypes yielded biallelic PDT *P* values of < 0.008.

Global multiallelic disequilibrium statistic

The PDT test also calculates a measure of the linkage disequilibrium for all the alleles of a given marker in the form of a global multiallelic *P* value. Using this statistic, the 11 markers of 1q32–42 fine map were tested for global transmission disequilibrium (Fig. 4). Importantly, two markers in the 1q32–42 region reach the 0.05 level of significance in the data set from the 332 SLE families. D1S425 results in a global *P* value of 0.034, which

Figure 2



Linkage in the 1q41 region. GENEHUNTER PLUS (see reference [S4] of supplementary material) was used to perform multipoint nonparametric linkage analysis using the 11 markers shown in the total collection of 210 families with systemic lupus erythematosus and in a subset of 163 white families. Marker positions (cM) were estimated based on the available genetic and physical maps.

appeared to be largely driven by the contribution of families in the trio collection ($P = 0.0056$). D1S490 showed the most significant result using this test ($P = 0.0091$), with both the sibpair and trio collections contributing to the linkage disequilibrium observed. Significant results for this global disequilibrium statistic were not observed for any other markers in the 1q41–42 region, including ADPRT (PARP). Two- and three-marker haplotype analysis with the global PDT identified only one significant result. The D1S425–D1S2827 'window' yielded a PDT global *P* value of 0.032.

Discussion

The localization of genes in complex genetic diseases is a challenging proposition, given that these disorders are likely to show significant locus heterogeneity, genetic epistasis, and incomplete penetrance, as well as environmental effects. In SLE, all the available evidence points to a similarly complex genetic etiology, with six recent genetic linkage studies implicating as many as 48 genetic loci [3–7,20]. While some of the loci identified are likely to be false positives or relatively minor genes enriched in one population or another, the locus at 1q41 is unique in that it has provided a significant, though modest, linkage signal in three independent SLE populations [3-5,10,14]. Fine-mapping in the 1q41 region by Moser *et al* [14] also showed that the highest overall LOD score in 127 families of the Oklahoma collection was at D1S229, and the greatest degree of allele sharing in 78 white families was at D1S2616. D1S229 also showed the strongest evidence for linkage in the UCLA collection [10], and was the best marker in genome screens performed on the Minnesota family collection [3,4]. Thus, the evidence for linkage at 1q41 is reproducible in independent collections using the identical markers.

To detect association with SLE, we used the transmission disequilibrium test (TDT), as well as two additional tests, the C-TDT and the PDT. The PDT is the strongest test of association because of its ability to maximize the information extracted from complex pedigrees. Importantly, marker D1S490 in the 1q41 region was significant on the multiallelic PDT (Fig. 4). Supporting the global finding, the C-TDT and PDT tests also identified evidence for transmission disequilibrium with several alleles from D1S490 and nearby markers (Fig. 3). Marker D1S425 in the 1q32 region also demonstrated significant transmission disequilibrium, particularly in the trio collection (Fig. 4).

We sought to confirm and expand upon the results of the single-marker tests by examining haplotypes from the 1q32–41 region for the presence of transmission disequilibrium. Haplotype-based association methods may be more powerful than single-marker tests [21–23]. For example, the same microsatellite allele (i.e. same size microsatellite repeat) may be present on a number of haplotypes, some of which may not be associated with disease. Indeed, haplotype-based association methods employing a dense map of markers have been used to localize genetic effects to small segments of chromosomes [24–29]. The knowledge of marker order and intermarker distances in the 1q41 region allowed the generation of haplotypes with unambiguously determined 'phase' in our large collection of SLE families. The D1S425–D1S2827 two-marker haplotype window in the 1q32 region showed evidence for transmission disequilibrium via the global PDT, while no other windows showed evidence of association using this multiallelic test. However, several individual two- and three-marker haplotypes from the 1q32 and 1q41 region showed significant transmission disequilibrium (Fig. 3). These haplotype results should be viewed with some caution, since they are uncorrected for the multiple allele combinations tested, and the possibility of Type I errors may be increased.

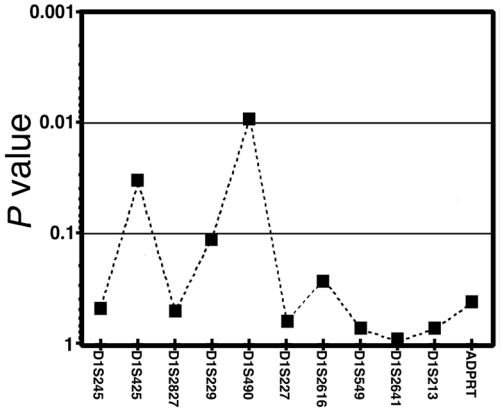
At present, the most interesting candidate gene in the region showing the strongest evidence for disequilibrium is the estrogen-related receptor gamma (ESRRG). This gene is found on the same contig as D1S490 (NT_004817) and is an orphan receptor within the steroid hormone receptor superfamily. It is expressed in lymphocytes and other

Figure 3

	D1S245	D1S425	D1S2827	D1S229	D1S490	D1S227	D1S2616	D1S549	D1S2641	D1S213	ADPRT	TDT		C-TDT	PDT
												T:NT	*P* value		
Single Marker	5											132:106	0.0919	0.0286	0.1770
		8										140:126	0.3906	0.0299	0.2230
		9										80:58	0.0611	0.0473	0.1420
				6								36:24	0.1213	0.0211	0.3700
				7								145:115	0.0628	0.0054	0.2220
					6							94:70	0.0609	0.0044	0.0071
					8							20:15	0.3981	0.0138	0.0462
					12							24:13	0.0706	0.0104	0.0096
						5						98:73	0.0559	0.0084	0.3060
							6					161:134	0.1160	0.0456	0.3090
Two-Marker Haplotypes	9	8										19:7	0.0186	0.0155	0.0593
		9	8									38:19	0.0119	0.0086	0.0790
		10	11									14:4	0.0184	0.0131	0.0295
				8	6							9:1	0.0114	0.0134	0.0578
				9	12							10:6	0.3173	0.2266	0.0455
					7	5						42:25	0.0378	0.0314	0.0451
						5	6					71:45	0.0158	0.0067	0.1020
						9	8					19:8	0.0343	0.0219	0.0814
							6	12				52:33	0.0393	0.0255	0.0159
							8	20				13:7	0.1797	0.1318	0.0272
								14	10			10:3	0.0522	0.0481	0.0455
Three-Marker Haplotypes	5	9	8									12:4	0.0455	0.0401	0.1320
				7	7	5						23:6	0.0016	0.0025	0.0043
					6	7	6					24:12	0.0455	0.0334	0.0387
					7	5	6					25:10	0.0112	0.0184	0.0690
						5	6	10				18:4	0.0028	0.0050	0.0628
							6	12	10			18:6	0.0143	0.0124	0.0075
									10	9	5	10:3	0.0522	0.0481	0.0588

Transmission disequilibrium of individual alleles and short marker haplotypes in the 1q32–42 region. Single-allele and two- or three-marker haplotype data sets were analyzed using the transmission disequilibrium test (TDT), the combined TDT and discordant-sib TDT (C-TDT), and the pedigree disequilibrium test (PDT). For the single-marker data, the entire collection of 332 families with systemic lupus erythematosus was used, while the two- and three-marker analyses were performed only on the 274 families with at least one typed parent. Only those marker alleles or haplotypes with at least one significant result are shown.

Figure 4



Pedigree Disequilibrium Test Global *P* value			
Marker	All Families	Sibpair Families	Trio Families
D1S245	0.4565	0.5961	0.4184
D1S425	**0.0345**	0.3637	**0.0056**
D1S2827	0.5096	0.8589	0.3352
D1S229	0.1061	0.0671	0.7693
D1S490	**0.0091**	**0.0425**	**0.0837**
D1S227	0.6291	0.2442	0.2367
D1S2616	0.2626	0.0992	0.7917
D1S549	0.7365	0.6735	0.3370
D1S2641	0.9587	0.2980	0.8911
D1S213	0.7308	0.9477	0.4312
ADPRT	0.4205	0.4184	0.4950

Global pedigree disequilibrium test at 1q32–42 in 332 families with systemic lupus erythematosus. The global (multiallelic) *P* values found using the pedigree disequilibrium test are graphed for each of the 11 markers that comprised the fine map. The 'All Families' data are graphed in the left panel.

tissues and is an interesting candidate, given the suspected role of sex hormones in the pathophysiology of both mouse and human lupus [30,31]. Other genes in the region include those for the cathelicidin antimicrobial peptide (near D1S2616), an innate microbial defense peptide expressed by the skin during inflammation; MARK (near D1S2641), a serine/threonine protein kinase; and at least three uncharacterized genes.

The data reported here provide some additional perspective on the initial reports that PARP might be the relevant gene in this locus. The finding of significant transmission distortion of marker alleles centromeric to PARP in the Minnesota collection suggests the possibility that the disequilibrium initially reported for PARP alleles may be due to more extensive disequilibrium – to include the PARP marker – in the families studied by Tsao and her colleagues compared with other populations and groups of families studied. A dense mapping of this interval by all the various groups and a pooling of data would help to resolve this question. It seems likely that, as in the HLA region [32], there will be a limited number of ancestral haplotypes in the 1q41 region, and that these ancestral haplotypes will be identifiable by typing a dense map of microsatellites. This should facilitate the identification of the responsible gene(s) in the region.

Conclusions

The data reported here confirm the evidence for genetic linkage at 1q41 in 210 SLE sibpair families from the Minnesota collection. Furthermore, we report new evidence for transmission disequilibrium of both single marker alleles and short marker haplotypes from the 1q41–42 interval in 122 trio families and in the total dataset of 332 SLE families. These data suggest that a human SLE susceptibility locus is located centromeric to PARP near the D1S490 marker.

Acknowledgements

We are grateful to the many patients and their families and physicians who have co-operated with this study. This work was supported by the General Clinical Research Center at the University of Minnesota and by grants from the Lupus Foundation of Minnesota, the Lupus Foundation of America – Texas Gulf Coast Chapter, and the National Institutes of Health (K08-AR0230105 [PMG], R01 AR/AI43271 [TWB], and SCOR in SLE P50 AR45231, project 1 [TWB]).

References

1. Online Mendelian Inheritance in Man (OMIM) – a catalog of human genes and genetic disorders. [http://www.ncbi.nlm.nih.gov/Omim/]
2. Rothfield N: **Systemic lupus erythematosus: clinical aspects and treatments.** In *Arthritis and Allied Conditions.* Edited by DJ McCarty. Philadelphia: Lea & Febiger; 1985:911-935.
3. Gaffney PM, Kearns GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML, Rohlf KE, Ockenden TC, Messner RP, King RA, Rich SS, Behrens TW: **A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families.** *Proc Natl Acad Sci U S A* 1998, **95**:14875-14879.
4. 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: **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.
5. Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Asundi N, 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: **Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees.** *Proc Natl Acad Sci U S A* 1998, **95**:14869-14874.
6. Lindqvist AK, 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: **A susceptibility locus for human systemic lupus erythematosus (hSLE1) on chromosome 2q.** *J Autoimmun* 2000, **14**:169-178.
7. Shai R, Quismorio FP, Jr., Li L, Kwon OJ, Morrison J, Wallace DJ, Neuwelt CM, Brautbar C, Gauderman WJ, Jacob CO: **Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families.** *Hum Mol Genet* 1999, **8**:639-644.
8. Gaffney PM, Moser KL, Graham RR, Behrens TW: **Recent advances in the genetics of systemic lupus erythematosus**. *Rheum Dis Clin N Am* 2001, in press.
9. Morel L, Mohan C, Yu Y, Croker BP, Tian N, Deng A, Wakeland EK**: Functional dissection of systemic lupus erythematosus using congenic mouse strains.** *J Immunol* 1997, **158**:6019-6028.
10. Tsao BP, Cantor RM, Kalunian KC, Chen CJ, Badsha H, Singh R, Wallace DJ, Kitridou RC, Chen SL, Shen N, Song YW, Isenberg DA, Yu CL, Hahn BH, Rotter JI: **Evidence for linkage of a candidate chromosome 1 region to human systemic lupus erythematosus.** *J Clin Invest* 1997, **99**:725-731.
11. Tan FK, Reveille JD, Arnett FC, Stivers DN, Tsao BP: **Poly(ADP)-ribose polymerase and susceptibility to systemic lupus erythematosus and primary antiphospholipid syndrome: comment on the article by Delrieu *et al.*** *Arthritis Rheum* 2000, **43**:1421-1423.
12. Tsao BP, Cantor RM, Grossman JM, Shen N, Teophilov NT, Wallace DJ, Arnett FC, Hartung K, Goldstein R, Kalunian KC, Hahn BH, Rotter JI: **PARP alleles within the linked chromosomal region are associated with systemic lupus erythematosus.** *J Clin Invest* 1999, **103**:1135-1140.
13. Criswell LA, Moser KL, Gaffney PM, Inda S, Ortmann WA, Lin D, Chen JJ, Li H, Gray-McGuire C, Neas BR, Rich SS, Harley JB, Behrens TW, Seldin MF: **PARP alleles and SLE: failure to confirm association with disease susceptibility.** J Clin Invest 2000, **105**:1501-1502.
14. Moser KL, Gray-McGuire C, Kelly J, Asundi N, Yu H, Bruner GR, Mange M, Hogue R, Neas BR, Harley JB: **Confirmation of genetic linkage between human systemic lupus erythematosus and chromosome 1q41**. *Arthritis Rheum* 1999, **42**:1902-1907.
15. Spielman RS, McGinnis RE, Ewens WJ: **Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM).** *Am J Hum Genet* 1993, **52**:506-516.
16. Spielman RS, Ewens WJ: **A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test.** *Am J Hum Genet* 1998, **62**:450-458.
17. Martin ER, Monks SA, Warren LL, Kaplan NL: **A test for linkage and association in general pedigrees: the pedigree disequilibrium test.** *Am J Hum Genet* 2000, **67**:146-154.
18. Dudbridge F, Koeleman BP, Todd JA, Clayton DG: **Unbiased application of the transmission/disequilibrium test to multilocus haplotypes.** *Am J Hum Genet* 2000, **66**:2009-2012.
19. Curtis D: **Use of siblings as controls in case-control association studies.** *Ann Hum Genet* 1997, **61**:319-333.
20. Gray-McGuire C, Moser KL, Gaffney PM, Kelly J, Yu H, Olson JM, Jedrey CM, Jacobs KB, Kimberly RP, Neas BR, Rich SS, Behrens TW, Harley JB: **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.
21. Collins A, Morton NE: **Mapping a disease locus by allelic association.** *Proc Natl Acad Sci U S A* 1998, **95**:1741-1745.
22. Terwilliger JD: **A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci.** *Am J Hum Genet* 1995, **56**:777-787.

23. Xiong M, Guo SW: **Fine-scale genetic mapping based on linkage disequilibrium: theory and applications.** *Am J Hum Genet* 1997, **60**:1513-1531.
24. Nair RP, Stuart P, Henseler T, Jenisch S, Chia NV, Westphal E, Schork NJ, Kim J, Lim HW, Christophers E, Voorhees JJ, Elder JT: **Localization of psoriasis-susceptibility locus PSORS1 to a 60-kb interval telomeric to HLA-C.** *Am J Hum Genet* 2000, **66**: 1833-1844.
25. Jenisch S, Henseler T, Nair RP, Guo SW, Westphal E, Stuart P, Kronke M, Voorhees JJ, Christophers E, Elder JT: **Linkage analysis of human leukocyte antigen (HLA) markers in familial psoriasis: strong disequilibrium effects provide evidence for a major determinant in the HLA-B/-C region.** *Am J Hum Genet* 1998, **63**:191-199.
26. Todd JA, Mijovic C, Fletcher J, Jenkins D, Bradwell AR, Barnett AH: **Identification of susceptibility loci for insulin-dependent diabetes mellitus by trans-racial gene mapping.** *Nature* 1989, **338**:587-589.
27. Degli-Esposti MA, Abraham LJ, McCann V, Spies T, Christiansen FT, Dawkins RL: **Ancestral haplotypes reveal the role of the central MHC in the immunogenetics of IDDM.** *Immunogenetics* 1992, **36**:345-356.
28. Degli-Esposti MA, Andreas A, Christiansen FT, Schalke B, Albert E, Dawkins RL**: An approach to the localization of the susceptibility genes for generalized myasthenia gravis by mapping recombinant ancestral haplotypes.** *Immunogenetics* 1992, **35**: 355-364.
29. Schmitt-Egenolf M, Eiermann TH, Boehncke WH, Stander M, Sterry W: **Familial juvenile onset psoriasis is associated with the human leukocyte antigen (HLA) class I side of the extended haplotype Cw6-B57- DRB1*0701-DQA1*0201-DQB1*0303: a population- and family-based study**. *J Invest Dermatol* 1996, **106**:711-714.
30. Evans MJ, MacLaughlin S, Marvin RD, Abdou NI: **Estrogen decreases *in vitro* apoptosis of peripheral blood mononuclear cells from women with normal menstrual cycles and decreases TNF-alpha production in SLE but not in normal cultures.** *Clin Immunol Immunopathol* **1997**, **82**:258-262.
31. Morel L, Croker BP, Blenman KR, Mohan C, Huang G, Gilkeson G, Wakeland EK**: Genetic reconstitution of systemic lupus erythematosus immunopathology with polycongenic murine strains.** *Proc Natl Acad Sci U S A* 2000, **97**:6670-6675.
32. Degli-Esposti MA, Leaver AL, Christiansen FT, Witt CS, Abraham LJ, Dawkins RL: **Ancestral haplotypes: conserved population MHC haplotypes.** *Hum Immunol* 1992, **34**:242-252.

Supplementary material

Supplementary materials and methods

Establishing marker order in the 1q41–42 region

The 12 markers from the 1q41–42 region originally typed by Moser *et al* [S1] were typed in the Minnesota collection (see Fig. 1 in the main text of this article). Genetic maps for the region were obtained from the Marshfield Clinic (Marshfield, WI, USA; http://research.marshfield-clinic.org/genetics) and the Genome Database (http://www.gdb.org). Since genetic maps have limited resolution in the case of closely linked markers, the available human genomic sequence from the region was analyzed to more accurately assess marker order and intermarker distances. The Human Genome Project Working Draft database (based on October 7, 2000 freeze) at the University of California, Santa Cruz (http://genome.ucsc.edu), was the primary source used to establish marker order. Three markers (D1S425, D1S229, and D1S213) could not be located in this database. The UniSTS database (http://www.ncbi.nlm.nih.gov/genome/sts/epcr.cgi) located D1S425 on contig NT_004656. According to Mapviewer tool on the National Center for Biotechnology Information site (http://www.ncbi.nlm.nih.gov/genome/guide/human), D1S425 lies 4 Mb telomeric of D1S245. BLAST (http://www.ncbi.nlm.nih.govb/BLAST/) analysis of the human genome sequence was used to locate D1S229 on the contig NT_004817, 125 kb centromeric of D1S490. D1S2860 was found to map to chromosome 3 and thus was omitted from the present analysis. Only 1 marker, D1S213, was not located on any assembled contigs or draft sequence.

Samples and genotyping

Panels of markers were optimized such that markers bearing the same fluorescent tag could be multiplexed in polymerase chain reaction (PCR). PCR (32 cycles) was performed on an ABI 877 Catalyst robotic workstation (5 µl reactions – 5 ng of genomic DNA, 2.5 mM $MgCl_2$, 0.2 mM dNTPs [Pharmacia, Piscataway, NJ, USA], 0.2 U Amplitaq Gold DNA polymerase [Perkin-Elmer, Foster City, CA, USA] in 1 × PCR Buffer II [Perkin-Elmer]). Individual primer-pair concentrations ranged from 0.31 to 3.30 pmol per reaction, based on the results of optimization runs. Pooled amplification products were electrophoresed through 5% polyacrylamide gels (FMC Bioproducts, Rockland, ME, USA) for 2 h at 3000 V on an ABI 377 DNA Sequencer. Semiautomated fragment sizing was performed using GENESCAN software (v. 2.1, ABI, Foster, CA, USA) followed by allele calling with GENOTYPER software (v. 2.5, ABI).

Data verification

Each genotype was reviewed manually by two members of the team to confirm the accuracy of allele calling. The genotype data from the 11 microsatellites used in the study were analyzed using PEDCHECK and RELATIVE to identify errors in Mendelian inheritance and to confirm familial relationships [S2,S3]. Additionally, the maximum-likelihood haplotypes for each pedigree (generated by GENEHUNTER PLUS [S4], a modified version of GENEHUNTER [S5]) were examined to identify multiple recombinants. Double recombinants were generally found to be genotyping errors and were corrected.

Linkage analysis

Multipoint nonparametric linkage analysis was performed using GENEHUNTER PLUS [S4]. We report the results based on the Sall scoring function, which emphasizes sharing the same allele identical-by-descent (IBD) among all affected family members [S6]. The reported logarithm of odds ratio (LOD) score is calculated as LOD = Zlr2/2ln10. In some instances, families of different ethnic groups were extracted from the master linkage file and analyzed separately. Allele frequencies used in the parameter files for each analysis were generated from the founder genotypes for the analyzed set of data (sibpair families, trio families, all families combined, and individual ethnic groups).

Transmission disequilibrium testing

The alleles from the 11-marker fine map were examined for the presence of transmission disequilibrium using the transmission disequilibrium test (TDT) [S7] and the discordant sibling TDT (S-TDT) test [S8]. Analysis was conducted using the combined TDT/S-TDT program version 1.1, set to accept one missing parent per trio [S8]. This program outputs separate results for the TDT, S-TDT and the combined TDT/S-TDT (referred to herein as C-TDT). In families with multiple affected individuals, only a single affected patient (the index case) was used for the TDT and C-TDT analyses.

Two- and three-marker haplotypes from the 1q32–41 region were tested for the presence of transmission disequilibrium. Maximum-likelihood haplotypes were generated using GENEHUNTER PLUS for the 274 families with at least one typed parent. Markers where phase could not be definitively determined were identified and labeled as missing data for the haplotype analysis. Founder haplotypes were then used to create two- and three-contiguous-marker exact-match clusters using a moving-window approach, as described by Nair *et al* [S9]. In order to prevent rare haplotypes from dominating the statistical tests and to better approximate the distributional assumptions of these tests, only those haplotypes with at least 10 founders were clustered. Haplotypes represented in fewer than 10 founders were pooled together. Every two- or three-marker 'window' was examined by moving the window across the 1q32–42 region one marker at a time from centromere to telomere. The resulting haplotype data were examined using the TDT and C-TDT with each haplotype considered as a single marker.

Additional information from the sibpair and multiplex families was extracted using the pedigree disequilibrium test (PDT) [S10]. The original form of the PDT incorporates information from every possible triad and discordant sibship in a given family, to yield an average measure of disequilibrium for the entire pedigree. In addition, we analyzed dyads (one parent and an affected offspring) when the offspring had a heterozygous genotype different from that of the parent [S11,S12]. The PDT was used to examine the single-marker and the two- and three-marker haplotype data.

Linkage analysis in families carrying 'risk' haplotypes

GENEHUNTER PLUS creates maximum-likelihood haplotypes when conducting multipoint linkage analysis. The GENEHUNTER PLUS haplotype output was used to identify families with founders carrying the seven three-marker haplotypes in the 1q41 region (D1S229–D1S2641) that had a TDT transmission:nontransmission ratio of at least 2:1 (66% transmission rate). Multipoint linkage analysis was conducted on the 'risk' and 'nonrisk' subsets as described above.

Supplementary results

Clinical features of systemic lupus erythematosus (SLE) in sibpair and trio families

The composition of the sibpair families studied in this report is summarized in Supplementary Table 1 (see below). Twenty-four additional families were included beyond the families that comprised the original Minnesota Cohorts 1 and 2 [S13,S14]. In addition, a collection of 122 trio families (one SLE patient with both parents) was assembled. Demographics and selected clinical characteristics of the affected individuals within the sibpair and trio collections are presented in Supplementary Table 2. The trio collection was highly enriched for white families (96% compared with 77% in the sibpair families). Compared with patients from the sibpair collection, trio individuals were more likely to be positive for anti-dsDNA antibody, and to have evidence of hematologic or renal disease. Lower percentages of trio patients had pleuritis, skin involvement, or CNS lupus. The differences in the clinical manifestations observed may reflect the greater ethnic diversity within the sibpair collection as well as potential differences in familial compared with sporadic SLE.

Linkage analysis in families carrying 'risk' haplotypes

The evidence for linkage was examined in the subset of families that carried the 'risk' haplotypes, defined as those families carrying three-marker haplotypes from D1S229 to D1S2641 that displayed at least a 2:1 transmission:nontransmission ratio (see Fig. 3 in the main text of this article). Families with the 1q41 risk haplotypes ($n = 88$) demonstrated a multipoint LOD score of 1.15 at D1S2616, while the remainder of the family collection ($n = 23$) showed a LOD score at D1S2616 of 0.30. Thus, SLE families carrying the 'risk' marker haplotypes with the strongest evidence for transmission disequilibrium were also the families that accounted for most of the evidence for linkage in the 1q41 region.

Localization of the 1q41 effect through haplotype analysis

The two most significant three-marker haplotypes identified in the moving-window analysis were D1S229–allele 7/D1S490–allele 7/D1S227–allele 5 and D1S2616–allele 6/D1S549–allele 12/D1S2641–allele 10. Interestingly, none of the 83 individuals who carried one of these two haplotypes contained both of these on a further extended haplotype. Furthermore, 33 of 55 patients with the 229/490/227 haplotype (60%) also bore allele 6 of D1S2616, while 14 of 28 individuals with the 2616/549/2641 haplotype also bore allele 5 of D1S227. This suggests the possibility that the major effect in this region is between markers D1S227 and D1S2616, an interval of approximately 0.96 megabasepairs. Alternatively, there might be two effects in the region, each carried on one of the haplotypes. The two-marker haplotype data provide some support for the former interpretation, since the

Supplementary Table 1

Composition of 210 Minnesota SLE sibpair and multiplex families

	Cohort 1	Cohort 2	New	Total
SLE families	104	82	24	210
Affected sibpairs	114	93	26	233
Affected relative pairs	127	111	35	273
Affected SLE individuals	220	179	53	452
Unaffected parents and sibs	155	101	41	297

Supplementary Table 2

Demographics and clinical features of 576 SLE patients

	Sibpair	Trio
Number of families	210	122
Number of affected individuals (sex, F:M)	440:10	123:0
Family ethnicity		
White	78	96
African-American	10	1
Hispanic	7	1
Asian	2	1
Mixed heritage	3	1
Laboratory/clinical features† (%)		
ANA positive	98	97
Anti-dsDNA positive	46	72***
Arthritis	85	84
Skin involvement	92	82**
Pleuritis	53	45*
Hematologic	48	72***
Renal disease	33	53***
CNS lupus	25	15**
Pericarditis	19	20

†Data represent the percentage of SLE patients having the indicated laboratory/clinical features at any time during the course of their disease. Differences in the clinical features between the trio and sibpair familiy collections were determined using generalized estimating equations to adjust for intrafamilial correlation within the sibpair families [S15]. ***$P < 0.0001$, **$P < 0.01$, *$P < 0.05$.
ANA = antinuclear antibodies; CNS = central nervous system

peak number of significant haplotype transmissions is found within the D1S227–D1S2616 interval, with the number of events falling off at surrounding windows. Although the current evidence for linkage disequilibrium at 1q41 extends over quite a broad interval, it is important to note that the marker density in the current study is still rather sparse. Efforts to further localize the gene in this region will require the identification and typing of additional microsatellite markers and single nucleotide polymorphisms (SNPs) in a large cohort of SLE cases and controls.

Supplementary references

S1. Moser KL, Gray-McGuire C, Kelly J, Asundi N, Yu H, Bruner GR, Mange M, Hogue R, Neas BR, Harley JB: **Confirmation of genetic linkage between human systemic lupus erythematosus and chromosome 1q41.** *Arthritis Rheum* 1999, **42**:1902-1907.

S2. Goring HH, Ott J: **Relationship estimation in affected sib pair analysis of late-onset diseases.** *Eur J Hum Genet* 1997, **5**:69-77.

S3. O'Connell JR, Weeks DE**: PedCheck: a program for identification of genotype incompatibilities in linkage analysis.** *Am J Hum Genet* 1998, **63**:259-266.

S4. Kong A, Cox NJ: **Allele-sharing models: LOD scores and accurate linkage tests.** *Am J Hum Genet* 1997, **61**:1179-1188.

S5. 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.

S6. Whittemore AS, Halpern J**: A class of tests for linkage using affected pedigree members.** ***Biometrics*** 1994, **50**:118-127.

S7. Spielman RS, McGinnis RE, Ewens WJ: **Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM).** *Am J Hum Genet* 1993, **52**:506-516.

S8. Spielman RS, Ewens WJ: **A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test.** *Am J Hum Genet* 1998, **62**:450-458.

S9. Nair RP, Stuart P, Henseler T, Jenisch S, Chia NV, Westphal E, Schork NJ, Kim J, Lim HW, Christophers E, Voorhees JJ, Elder JT: **Localization of psoriasis-susceptibility locus PSORS1 to a 60-kb interval telomeric to HLA-C.** *Am J Hum Genet* 2000, **66**: 1833-1844.

S10. Martin ER, Monks SA, Warren LL, Kaplan NL: **A test for linkage and association in general pedigrees: the pedigree disequilibrium test.** *Am J Hum Genet* 2000, **67**:146-154.

S11. Curtis D: **Use of siblings as controls in case-control association studies.** *Ann Hum Genet* 1997, **61**:319-333.

S12. Dudbridge F, Koeleman BP, Todd JA, Clayton DG: **Unbiased application of the transmission/disequilibrium test to multilocus haplotypes.** *Am J Hum Genet* 2000, **66**:2009-2012.

S13. 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: **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.

S14. Gaffney PM, Kearns GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML, Rohlf KE, Ockenden TC, Messner RP, King RA, Rich SS, Behrens TW: **A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families.** *Proc Natl Acad Sci U S A* 1998, **95**:14875-14879.

S15. Zeger SL, Liang KY: **Longitudinal data analysis for discrete and continuous outcomes.** *Biometrics* 1986, **42**:121-130.