D6S439 Microsatellite Identifies a New Susceptibility Region for Primary Sjögren's Syndrome

JUAN-MANUEL ANAYA, DORA RIVERA, LUIS G. PALACIO, MAURICIO ARCOS-BURGOS, and PAULA A. CORREA

ABSTRACT. ***Objective.*** To examine genetic variations in the region surrounding loci of the major histocompatibility complex, and to investigate the probable location of a new candidate region on the short arm of chromosome 6 predisposing to primary Sjögren's syndrome (SS).

Methods. We conducted an association study and positional candidate gene approach by microsatellite analysis. Five polymorphic microsatellite markers, D6S273, D6S439, D6S1645, D6S291, and DS61019, spanning the region 6p21.3, and establishing particular landmarks to discriminate between the human leukocyte antigen class II and tumor necrosis factor-α loci, were genotyped by polymerase chain reaction technique.

Results. A total of 64 patients with primary SS and 120 matched controls were examined. There was no genetic stratification among cases and controls. Genotype distribution analysis disclosed a significantly higher number of homozygotes for D6S439 locus in patients than in controls [odds ratio (OR): 3, 95% confidence interval (CI): 1.46-6.14, p = 0.004]. Confirmation of this homozygosity was established by the gene correlation intra-locus test (Fis value = +0.233, p = 0.0007). Allele D6S439*274 was associated to disease (OR: 3, 95% CI: 1.35-6.65, p = 0.006, p_c = 0.04). Among patients, no significant linkage disequilibrium (LD) value was found between the studied microsatellites and TAP, HLA-DRB1, or HLA-DQB1 loci. In controls, there was LD between D6S1645 and D6S291 loci.

Conclusion. Our results indicate that D6S439 microsatellite defines a new susceptibility region for primary SS, independent of LD with TAP and HLA DQ/DR. These findings might imply that a gene surrounding this location is causally related to the disease. (J Rheumatol 2003;30:2152–6)

Key Indexing Terms:
SJÖGREN'S SYNDROME BAK MAJOR HISTOCOMPATIBILITY COMPLEX
HOMOZYGOSITY D6S439 MICROSATELLITES

Sjögren's syndrome (SS) is a chronic autoimmune rheumatic and lymphoproliferative disease characterized by a progressive lymphocytic and plasma cell infiltration of the salivary and lachrymal glands, and the production of autoantibodies, mainly leading to dryness of the eyes (keratoconjunctivitis sicca), and mouth (xerostomia)[1]. SS may occur alone (primary) or in association (secondary) with almost any of the autoimmune rheumatic diseases, the most frequent being rheumatoid arthritis (RA).

From the Cellular Biology and Immunogenetic Unit, Corporación para Investigaciones Biológicas (CIB); the Rheumatology Unit, Clínica Universitaria Bolivariana, School of Medicine, Universidad Pontificia Bolivariana (UPB); and the Department of Biology, University of Antioquia (UA), Medellin, Colombia.

J-M. Anaya, MD, Associate Researcher, Cellular Biology and Immunogenetic Unit, CIB, and Professor of Medicine, School of Medicine, UPB; D. Rivera, BSc, Assistant Researcher; L.G. Palacio, PhD Student, Department of Biology, UA, and Instituto Neurológico de Antioquia, Medellin; M. Arcos-Burgos, MD, PhD, Chief, Population Genetics, Mutacarcinogenesis and Genetic Epidemiology Group, Department of Biology, and Professor, UA; P.A. Correa, MSc, Assistant Researcher, Cellular Biology and Immunogenetic Unit, CIB, and Assistant Professor, School of Medicine, UPB.

Address reprint requests to Dr. J-M. Anaya, Corporación para Investigaciones Biológicas, Cra. 72-A No 78-B-141, Medellín, Colombia. E-mail: janaya@cib.org.co

Submitted August 1, 2002; revision accepted March 23, 2003.

A role for genetic factors in the pathogenesis of primary SS has been suggested by a tendency toward familial aggregation and the presence of autoantibodies and other autoimmune diseases in family members[2,3]. The polymorphic major histocompatibility complex (MHC) genes are the best documented genetic region associated with the development of primary SS, most specifically human leukocyte antigen (HLA)-DR and -DQ alleles[4-6]. Stronger associations have been found between HLA class II alleles and autoantibodies (anti-Ro and anti-La) than to the disease itself[7-10]. This might be attributed to genetic variations in the region surrounding these loci. To further investigate this possibility and to examine if additional genes in the MHC region other than HLA-DRB1 and -DQB1 may contribute to susceptibility to primary SS, we undertook a positional candidate gene approach by microsatellite analysis, spanning the region 6p21.3.

Microsatellites are mostly found in non-coding regions and their polymorphisms rarely have functional consequences; however, they are used as tags to label particular genomic regions of a chromosome. By this approach, we identified a new candidate region on short arm of chromosome 6 associated with primary SS.

MATERIALS AND METHODS

Study population. Consecutive patients attending the Rheumatology Unit at the Corporación para Investigaciones Biológicas (CIB), and the Clinica Universitaria Bolivariana, in Medellin, were included. All patients met at least 4 of the European classification criteria for primary SS including a positive minor salivary gland biopsy[11]. Those patients with another rheumatologic autoimmune disease other than SS were excluded [i.e., RA, systemic lupus erythematosus (SLE), scleroderma], as well as patients with a previous history of lymphoma or chronic infectious disease.

Patient demographics and cumulative clinical and laboratory manifestations over the course of disease were obtained either by chart review or by interview and were collected in a standard data collection form. The clinical variables associated with SS were classified according to Oxholm, *et al*[12].

Persons without inflammatory or autoimmune disease, matched to patients by age (± 5 years), gender, geography, and socioeconomic status were included as controls. They were unrelated to the patients. All subjects were coded. The genotyping was blinded to case or control status. The code was broken when the genotyping was completed. This research, accomplished in accordance with Resolution No 008430 of 1993 from the Ministry of Health of the Republic of Colombia, was classified as research with minimal risk. The Ethics Committee of the CIB approved the present study.

Microsatellite typing. Genomic DNA was isolated from sodium dodecyl sulfate-lysed and proteinase-K treated peripheral blood cells by phenol-chloroform extraction. Five polymorphic microsatellite markers, D6S273, D6S439, D6S1645, D6S291, and D6S1019 were studied. These microsatellites span the region 6p21.3 with intervals of less than 5 centiMorgan (cM, unit of genetic distance equivalent to a 1% probability of recombination during meiosis; 1 cM corresponds approximately to 1000 kb) and thus establish particular landmarks to discriminate between the HLA class II and tumor necrosis factor (TNF) loci (Figure 1)[13].

Genotyping was performed by a non-radioactive method using polymerase chain reaction (PCR) in a total reaction volume of 20 ul that contained 50 ng of DNA, 5 pmol of each primer, 2 mM of dNTPs (Promega, Madison, WI, USA), and 2 U of Taq DNA polymerase (Promega). Single-step reaction was used for amplification. Initial denaturation (4 min at 95°C) was followed by 35 cycles of 95°C for 30 s, 60 s at 56°C/64°C, and 60 s at 72°C with a final extension at 72°C for 10 min. The primers used were: 5'CAACTTTTCTGTCAATCCA3' and 5'ACCAAACTTCAAATTTTCGG3' for D6S273; 5'GATGATTTAAGTTTCCTGTGGACC3' and 5'TTCAAGGACAGCCTCAGGG3' for D6S439; 5'CAGGAGAACCACTTGAACC3' and 5'CCCACTTAGCAGACAGAGAG3' for D6S1645; 5'CTCAGAGGATGCCATGTCTAAAATA3' and 5'GGGGATGACGAATTATTCACTAACT3' for D6S291; and 5'TCATGGGGTTGCAAAAGTAT3' and 5'CAGTGGCTGAACAGCATATG3' for D6S1019. The microsatellites' alleles were resolved on denaturing polyacrylamide gels comprising 6% acrylamide (19:1 with bis-acrylamide) (Promega), and 7 M urea (Promega). Products were silver stained using the GenePrint™ STR Systems (Promega). PCR products related to specific markers were identified by their size from gels. Data were analyzed using an internal size marker and external allelic ladders.

Statistical and genetic data analysis. Results are shown in averages ± standard deviation, and in frequencies. Differences among the case and the control marker gene frequencies were established by the z proportions test. In all cases the Bonferroni correction for multiple comparisons was applied. The expected Hardy-Weinberg (H-W) proportions were tested against the observed ones in order to detect significant departures from the H-W equilibrium. We followed the procedure described by Guo and Thompson[14], in a modified version of the Markov-chain random walk algorithm. To this effect 100,000 replicas in the Markov-chain were evaluated for a significance level of 0.05. To discriminate the directions of these H-W deviations, we performed subsequent tests to detect the presence of significant heterozygote excess or deficit.

As the presence of population stratification causes linkage disequilibrium overestimation, we tested the existence of population heterogeneity among the case and control sets, by using Wright's F statistics according to the non-biased method of Weir and Cockerham (θ value)[15]. Bootstrap tests involving 30,000 replications with a 95% confidence interval (CI) were performed to determine significant differences from 0. As the gametic phase was unknown, linkage disequilibrium between pairs of loci was tested with likelihood-ratio test, with empirical distribution obtained by a permutation procedure[14]. With this method and assuming linkage equilibrium, the likelihood of the data was computed by using the fact that, under this hypothesis, the haplotype frequencies obtained were the product of the allele frequencies. The likelihood of the data not assuming linkage equilibrium was obtained by applying the expectation-maximization algorithm to estimate haplotype frequencies[16]. The above analyses and estimations were performed using Genetic Data Analysis[17], Arlequin[18], and Genepop[19]. Finally, the association between different alleles and the presence of autoantibodies was evaluated by chi-square test. A p value below 0.05 was considered significant.

RESULTS

A total of 64 patients with primary SS (63 women) and 120 controls were genotyped. The mean age of the patients was not different from that of the controls (49 ± 13 vs 47 ± 12

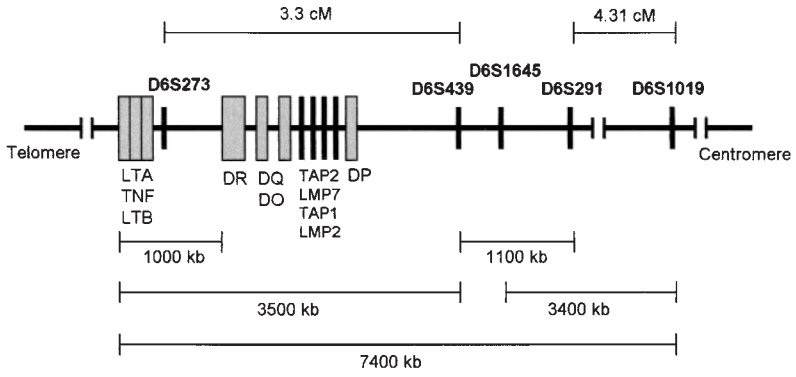

Figure 1. Disposition of microsatellite loci D6S273, DS6439, D6S1645, D6S291, and D6S1019 on short arm of chromosome 6 and their relationship with some MHC genes, as described in Santa Cruz Human Genome Project Working Draft[13]. cM: centiMorgan (unit of genetic distance equivalent to a 1% probability of recombination during meiosis), kb: kilobases. 1 cM corresponds approximately to 1000 kb.

yrs). The mean duration of primary SS was 6.9 ± 5.6 yrs. During the course of the disease, all patients presented exocrine involvement, 63% had arthralgia/arthritis, 23% cutaneous involvement, and 27% Raynaud's phenomenon. Antinuclear antibodies were detected in 55 (86%) patients, 38 (59%) patients had anti-Ro antibodies, and 20 (31%) patients had anti-La antibodies (by ELISA, INCSTAR, Stillwater, Minnesota, USA).

For cases and controls the gene frequencies estimated by the maximum likelihood method for each of the microsatellite loci are shown in Table 1. The Kolmogorov-Smirnov test to detect gene frequency distribution departures from the normal distribution was negative for each one of the 5 loci. The frequency of allele D6S439*274 was significantly different among the patients when compared with controls (Table 1). To avoid spurious associations due to stratification, we tested the existence of population heterogeneity among the case and control sets, as described[15]. There was no stratification between patients and control samples since the θ value (Fst) was not significantly different from 0 (Table 2). Thus, patients and controls were from a similar genetic background.

Genotype distribution analysis disclosed a significantly higher number of homozygotes for the D6S439 locus in patients with respect to controls [42% vs 19%, odds ratio (OR): 3, 95% CI: 1.46-6.14, p = 0.004] (Table 3). The Fis gene correlation intra-locus value for D6S439 indicated a strong deviation in the expected homozygote proportion, confirming this excess of homozygosity in patients (Table 4). No significant differences were observed in other loci concerning the number of homozygotes or heterozygotes.

No significant value was found when testing for the existence of linkage disequilibrium among the studied loci in the patients. However, in the control group linkage disequilibrium was observed between D6S291 and D6S1645 loci. The control population was in H-W disequilibrium for loci D6S273, D6S439, and D6S291, while patients were in H-W equilibrium for all the loci tested.

Table 2. Weir and Cockerham estimates of population subdivision. Theta-P, f, and F estimates represent the non-biased estimates of Wright classical Fis, Fit, and Fst subdivision coefficients, respectively; CI: confidence interval at 95% of the estimates after the bootstrap process of 30,000 replicas. There was no subdivision among the individual subpopulation level, the individual total population level (patients and controls all together), and the subpopulation total population level.

Bound	f	F	θ-P
Upper	0.03	0.042	0.01
Lower	–0.04	–0.04	0.002
Replicas, n	30,000	30,000	30,000
CI	95.0	95.0	95.0

No significant influence of the homozygosity at D6S439 on the presence of autoantibodies was observed. We did not find an association between allele D6S439*274 and the presence of autoantibodies or the clinical manifestations.

DISCUSSION

The strong homozygosity deviation at D6S439 microsatellite and the allelic imbalance at this locus indicate that this marker defines a candidate region for the acquisition of primary SS. These findings might imply that a gene(s) surrounding this location is causally related to the disease.

Primary SS is a multifactorial disease, phenotypically heterogeneous, and polygenic with genetic heterogeneity. In primary SS, the difficulty in obtaining fully informative pedigrees reflects its very complex genetic influence, with many contributing genes of low penetrance. In addition, a poor definition of disease subtypes, and the influence of environment on the disease, also complicate both defining and finding the most important genes associated with the pathogenic events leading to primary SS[20].

Table 1. Microsatellite allelic frequencies in patients with primary SS (pSS) and in controls.

D6S273			D6S439			D6S1645			D6S291			D6S1019		
Allele	pSS n = 128	Controls n = 240	Allele	pSS n = 106	Controls n = 230	Allele	pSS n = 126	Controls n = 240	Allele	pSS n = 128	Controls n = 240	Allele	pSS n = 124	Controls n = 240
126	0.023	0	274	0.142	0.052*	224	0.008	0.042	198	0.016	0.008	210	0.153	0.196
130	0.125	0.179	276	0.113	0.078	228	0.008	0.013	200	0.328	0.296	212	0.040	0.046
132	0.047	0.050	278	0.123	0.117	232	0.016	0.013	202	0.445	0.404	222	0.387	0.300
134	0.328	0.313	280	0.425	0.330	238	0.024	0.017	204	0.031	0.054	224	0.024	0.008
136	0.250	0.225	282	0.113	0.235	240	0.016	0.008	206	0.016	0.033	226	0.016	0.013
138	0.164	0.171	284	0.075	0.117	242	0.159	0.083	208	0.047	0.025	228	0.008	0.033
140	0.039	0.021	286	0	0.057	244	0.524	0.533	210	0.031	0.079	230	0.137	0.117
142	0.023	0.042	288	0.009	0.013	246	0.056	0.088	212	0.086	0.063	232	0.202	0.254
						248	0.063	0.063	214	0	0.004	234	0.024	0.013
						250	0.040	0.092	216	0	0.033	236	0	0.013
						252	0.079	0.046						

* OR: 3; 95% CI: 1.35–6.65; p = 0.006; p_c = 0.04.

Table 3. Microsatellites homozygosity distribution among patients with primary SS and controls.

Microsatellite	Group	Expected Homozygotes, n	Observed Homozygotes, n	Expected Heterozygotes, n	Observed Heterozygotes, n
D6S273	Patients	13.5	12	50.5	52
	Controls	25.3	21	94.7	99
D6S439	Patients	12.7	22*	40.3	31
	Controls	23.1	22*	91.9	93
D6S1645	Patients	19.5	14	43.5	49
	Controls	37.6	35	82.4	85
D6S291	Patients	20.0	15	44.0	49
	Controls	31.7	25	88.3	95
D6S1019	Patients	14.2	12	47.8	50
	Controls	24.8	36	95.2	84

* Homozygosity for D6S439 was observed in 42% of patients and in 19% of controls (OR: 3, 95% CI: 1.46–6.14, p = 0.004); no other significant differences were found.

Table 4. Hardy-Weinberg test in patients evaluating the heterozygote deficit as an alternative hypothesis. Fis values represent the gene correlation intra-locus. Fis values are significantly greater than 0 when there is a high correlation among the genes at one locus, meaning a deficit of heterozygotes (excess of homozygotes).

Locus	Fis	p
D6S273	–0.031	0.13
D6S439	+0.233	0.0007
D6S1645	–0.129	0.98
D6S291	–0.115	0.96
D6S1019	–0.047	0.06

Homozygosity deviation at locus D6S439 affecting the distribution of the genetic frequencies (Tables 1 and 3), implies that it could be a selection disturbing genotype distribution in patients with primary SS (either directional, or selection against heterozygotes). This explains, in part, the association between this locus and disease[21]. Susceptibility due to directional selection implies the participation of either environmental or chance events, which in the case of primary SS occur over time, since the disease is observed more frequently in the fifth decade of life. These interactions may be cumulative and do not all occur necessarily at the time when clinical manifestations appear.

The well known extensive polymorphism at the MHC loci is probably attributable to balancing selection caused by the superior performance of the immune system of individuals more heterozygous at these loci[22]. High heterozygosity suggests positive selection in favor of diversity. The opposite of balancing selection (i.e., selection against heterozygotes) will cause homozygosity and divergence among populations[21], as seemed to be the case in our patients for D6S439 microsatellite. Actually, we observed that in patients, most of the loci were in H-W equilibrium, suggesting that such loci were under disturbing forces including selection, mutation, or the continuation of random mating[23].

We observed that allele D6S439*274 was associated with disease (Table 1). The finding that D6S439*286 allele was absent in patients might be a reflection of the summative effect of other alleles at the locus. A positive association means that statistically, a non random distribution among marker polymorphisms and trait categories does exist. This biased distribution may or may not be caused by structural chromosomal linkage, since other genes, in linkage disequilibrium to the associated marker, may be causing the phenotype. This was not the case in our study since we did not observe linkage disequilibrium among microsatellites in patients. For this reason the Bonferroni correction was applied just for the independent alleles at D6S439 marker. Importantly, 50 of the patients from this study had been previously examined for TAP, HLA-DRB1, and -DQB1 polymorphism[10]. Taken together, analysis of these 50 patients showed no linkage disequilibrium between microsatellites and HLA or TAP loci ($p > 0.05$, after 10,000 replicas). Thus D6S439 appears to be an HLA independent susceptibility marker for primary SS (Figure 1). It is noteworthy that we did not observe D6S439 association with SLE or RA[24,25].

On the other hand, it is possible that the 2 groups could be stratified with respect to one or more unmeasured risk factors (either genetic or environmental) and this could be the cause of an association between disease and a marker allele. In our study, the distribution of the gene frequencies at the different loci fitted the normal distribution, as the Kolmogorov-Smirnov test did not show significant p values for departures from normality. Genetic stratification among cases and controls was not found (Table 2). In addition, caution was taken in matching cases and control samples, particularly in controlling variables such as age, sex, and socioeconomic status.

In conclusion, D6S439 tags a new susceptibility region for primary SS. The most likely gene related to the D6S439 chromosomal location appears to be *BAK-1,* which codes for the Bcl-2 antagonist/killer 1 protein[26]. Future replication and

transmission disequilibrium test studies might be useful in confirming our results, and facilitating structural mapping.

ACKNOWLEDGMENTS

We thank all the patients and participants of this study. We also thank the anonymous reviewers of this paper for their valuable comments.

REFERENCES

1. Jonsson R, Haga HJ, Gordon T. Sjögren's syndrome. In: Koopman W, editor. Arthritis and allied conditions. 14th ed. Philadelphia: Lippincott Williams & Wilkins; 2001:1736-59.
2. Reveille JD, Wilson RW, Provost TT, Bias WB, Arnett FC. Primary Sjögren's syndrome and other autoimmune diseases in families. Prevalence and immunogenetic studies in six kindreds. Ann Intern Med 1984;101:748-56.
3. Foster H, Stephenson A, Walker D, Cavanagh G, Kelly C, Griffiths I. Linkage studies of HLA and primary Sjogren's syndrome in multicase families. Arthritis Rheum 1993;36:473-84.
4. Fei HM, Kang H, Scharf S, Erlich H, Peebles C, Fox R. Specific HLA-DQA and HLA-DRB1 alleles confer susceptibility to Sjögren's syndrome and autoantibody production. J Clin Lab Anal 1991;5:382-91.
5. Reveille JD, Arnett FC. The immunogenetics of Sjögren's syndrome. Rheum Dis Clin North Am 1992;101:748-56.
6. Kang HI, Fei HM, Saito I, et al. Comparison of HLA class II genes in Caucasoid, Chinese, and Japanese patients with primary Sjögren's syndrome. J Immunol 1993;150:3615-23.
7. Reveille JD, Macleod MJ, Whittington K, Arnett FC. Specific amino acid residues in the second hypervariable region of HLA-DQA1 and DQB1 chain genes promote the Ro (SS-A)/La (SS-B) autoantibody responses. J Immunol 1991;146:3871-6.
8. Harley JB, Alexander EL, Bias WB, et al. Anti-Ro (SS-A) and anti-La (SS-B) in patients with Sjögren's syndrome. Arthritis Rheum 1986;29:196-206.
9. Bolstad AI, Wassmuth R, Haga HJ, Jonsson R. HLA markers and clinical characteristics in Caucasians with primary Sjögren's syndrome. J Rheumatol 2001;28:1554-62.
10. Anaya JM, Correa PA, Mantilla RD, Arcos-Burgos M. TAP, HLA-DQB1, and HLA-DRB1 polymorphism in Colombian patients with primary Sjögren's syndrome. Semin Arthritis Rheum 2002;31:396-405.
11. Vitali C, Moutsopoulos HM, Bombardieri S. The European community study group on diagnostic criteria for Sjögren's syndrome. Sensitivity and specificity of test for ocular and oral involvement in Sjögren's syndrome. Ann Rheum Dis 1994;53:637-47.
12. Oxholm P, Asmussen K, Axéll T, et al. Sjögren's syndrome: terminology. Clin Exp Rheumatol 1995;13:693-6.
13. UCSC Human Genome Project Working Draft, June 2002 assembly [cited June 10, 2003]. Available from: http://genome.cse.ucsc.edu
14. Guo S, Thompson E. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992;48:361-72.
15. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution 1984;38:1358-70.
16. Slatkin M, Excoffier L. Testing for linkage disequilibrium in genotypic data using the EM algorithm. Heredity 1996;76:377-83.
17. Lewis P O, Zaykin D. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0. 1997. [cited July 16, 2000] http://lewis.eeb.uconn.edu/lewishome/software.html
18. Schneider S, Kueffer JM, Roessli D, Excoffier L. Arlequin, Version 1.1. Available from: http://lgb.unige.ch/arlequin [cited July 29, 2003].
19. Raymond M, Rousset F. GENEPOP (version 3.2a): Population genetics software for exact test and ecumenicism. J Heredity 1995;86:248-9.
20. Correa, PA, Anaya, J-M. Inmunogenetica del síndrome de Sjögren. In: Anaya JM, Ramos M, García M, editors. Síndrome de Sjögren. Medellín: Corporación para Investigaciones Biológicas; 2001:104-12.
21. Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM. An introduction to genetic analysis, 5th edition. New York: W.H. Freeman; 1993:766-8.
22. Hughes AL, Yeager M. Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet 1998;32:415-35.
23. Weir BS. Disequilibrium. In: Weir BS, editor. Genetic data analysis II. Sunderland: Sinauer Associates; 1996:91-140.
24. Correa PA, Rivera D, Palacio LG, Arcos-Burgos M, Anaya JM. Microsatellite analysis refines a location for a susceptibility region to rheumatoid arthritis on 6p [abstract]. Arthritis Rheum 2002;46 Suppl:S98.
25. Correa PA, Rivera D, Arcos-Burgos M, Molina JF, Pinto LF, Anaya JM. Microsatellite analysis defines a susceptibility region to systemic lupus erythematosus close to TNF alpha locus [abstract]. Ann Rheum Dis 2003;62 Suppl 1:199.
26. Herber JA, Phillips S, Beck S, et al. Genomic structure and domain organization of the Bak gene. Gene 1998;211:87-94.