Am. J. Hum. Genet. 64:1406–1410, 1999

The Critical Region for Behçet Disease in the Human Major Histocompatibility Complex Is Reduced to a 46-kb Segment Centromeric of HLA-B, by Association Analysis Using Refined Microsatellite Mapping

Masao Ota,[1] Nobuhisa Mizuki,[3,4] Yoshihiko Katsuyama,[2] Gen Tamiya,[4] Takashi Shiina,[4] Akira Oka,[4] Hitoshi Ando,[5] Minoru Kimura,[4] Kaori Goto,[3,4] Shigeaki Ohno,[3] and Hidetoshi Inoko[4]

Departments of [1]Legal Medicine and [2]Pharmacy, Shinshu University School of Medicine, Matsumoto, [3]Department of Ophthalmology, Yokohama City University School of Medicine, Yokohama, and [4]Department of Genetic Information, Division of Molecular Life Science, Tokai University School of Medicine, and [5]Japanese Red Cross, Shonan Blood Center, Kanagawa, Japan

Summary

The HLA-B51 allele is known to be associated with Behçet disease. Recently, we found a higher risk for Behçet disease in the *MICA* gene, 46 kb centromeric of *HLA-B*, by investigation of GCT repetitive polymorphism within exon 5 of *MICA*. The pathogenic gene causing Behçet disease, however, has remained uncertain. Here, eight polymorphic microsatellite markers, distributed over a 900-kb region surrounding the *HLA-B* locus, were subjected to association analysis for Behçet disease. Statistical studies of associated alleles detected on each microsatellite locus showed that the pathogenic gene for Behçet disease is most likely found within a 46-kb segment between the *MICA* and *HLA-B* genes. The results of this mapping study, and the results of an earlier study of ours, suggest that *MICA* is a strong candidate gene for the development of Behçet disease.

Introduction

Behçet disease (MIM 109650) is characterized by four major symptoms: oral aphthous ulcers, skin lesions, ocular symptoms, and genital ulcerations. Occasionally, it is distinguished by inflammation of the tissues and of the organs, including those of the gastrointestinal tract, the central nervous system, the vascular system, the lungs, and the kidneys. This disease is defined as a refractory systemic inflammatory disorder and occurs sporadically—that is, most patients with Behçet disease are the only members of their families with the disease. The onset of Behçet disease is believed to be triggered by the involvement of some environmental factors in individuals with a particular genetic background, probably as a multifactorial disease. It has been well established that Behçet disease is associated with the HLA-B51 molecule that is relatively frequent, with a range of 45%–60% in many ethnic groups including Asian and Eurasian populations from Japan to the Middle East (Ohno et al. 1982). However, our previous study on polymorphism of the Tau-a microsatellite in the *TNF* gene region on the 200-kb centromeric side of the *HLA-B* gene suggested that the *HLA-B* gene may not be the primary locus responsible for Behçet disease and implicated some other pathogenic gene(s) located near the *HLA-B* gene (Mizuki et al. 1995). Further, we recently reported higher association of six GCT repetitions (A6 alleles) in a microsatellite polymorphism within exon 5 of the *MICA* (major histocompatibility complex [MHC] class I chain-related genes) gene than in HLA-B51 (Mizuki et al. 1997*a*). The *MICA* gene, spanning an 11-kb stretch of DNA, is located ~46 kb centromeric of the *HLA-B* gene. The function of a MICA protein has not been fully understood yet, but it has been suggested that it is recognized by certain T cells with variable-region Vδ1γδ T-cell receptors from intestinal epithelium (Groh et al. 1996, 1998). The characteristic symptoms of neutrophil hyperactivity and abnormality or hyperreactivity of γδ T cells in Behçet disease may be attributed to unusual expression and recognition of MICA molecules. Thus, having the *MICA* gene makes one highly susceptible to Behçet disease. However, only 76% of patients with Behçet disease possessed the A6 allele in *MICA*, so the strong association of *MICA-A6* with the disease may result from linkage disequilibrium with some unknown causative gene(s) around the *MICA–HLA-B* gene region. In this respect, fine mapping of the susceptible locus by means of more genetic markers in this region is needed.

Received September 9, 1998; accepted for publication March 1, 1999; electronically published April 5, 1999.

Address for correspondence and reprints: Prof. Hidetoshi Inoko, Department of Genetic Information, Division of Molecular Life Science, Tokai University School of Medicine, Bohseidai, Isehara, Kanagawa 259-1193, Japan. E-mail: hinoko@is.icc.u-tokai.ac.jp

0002-9297/99/6405-0021$02.00

In the present study, we did association analysis using eight microsatellite polymorphic markers identified around the *HLA-B* gene by our large-genome sequencing of the HLA class I region (Tamiya et al. 1998), to precisely map the pathogenic gene involved in the development of Behçet disease. To our knowledge, this is the first report of a genetic disease association study using microsatellite markers densely spread over a candidate gene region.

Subjects and Methods

Subjects

The test samples were obtained from 132 healthy controls and 74 Japanese patients with Behçet disease. The same patients reported elsewhere (Mizuki et al. 1997*a*) were studied. The disease was diagnosed according to the criteria proposed by the Japan Behçet Disease Research Committee at the Uveitis Clinic, Department of Ophthalmology, Yokohama City University School of Medicine or at the Department of Internal Medicine, Nanasawa Rehabilitation Hospital Comprehensive Stroke Center. Informed consent was obtained from all healthy donors and patients before collection of peripheral blood.

Genotyping for Microsatellite Alleles

To determine the number of repeat units of the microsatellite loci exhibiting polymorphisms, we synthesized unilateral primers by labeling at the 5′-ends with the fluorescent reagents, 6-FAM, HEX, and TET (PE Applied Biosystems). PCR primers and conditions for amplifying C1-2-A, *MICA*, C1-4-1, C1-2-5, C1-3-1, C2-4-4, and C3-2-11 microsatellites were the same as described elsewhere (Mizuki et al. 1997*a*; Ota et al. 1997; Tamiya et al. 1998). PCR primers and conditions for MIB alleles were comparable to those of Grimaldi et al. (1996). The PCR-amplified products were denatured for 5 min at 100°C, mixed with formamide-containing stop buffer, applied with a size standard marker of GS500 TAMRA (PE Applied Biosystems) to each lane, and run on a 4% polyacrylamide denaturing gel containing 8 M urea, in an automated DNA sequencer. Fragment sizes were determined automatically by means of the GeneScan software (PE Applied Biosystems).

Statistical Analysis

Gene (allele) frequencies were estimated by direct counting. The significance of the distribution of alleles between the patients with Behçet disease and healthy controls was tested by the χ^2 method with the continuity correction and Fisher's exact probability test (*P* test). *P* was corrected by multiplication by the number of mi-

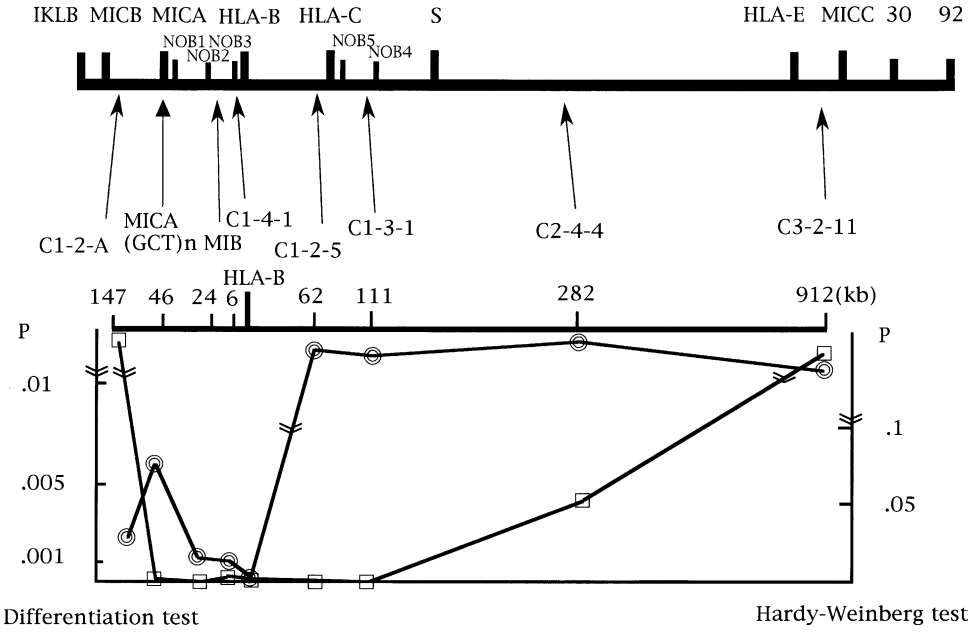

Figure 1 *P* values of the exact test of HWP and the genotypic differentiation test with the location of microsatellite markers used for association analyses of Behçet disease. These markers were scattered within the 1,300-kb segment from the IkBL to HLA-92 loci in the human MHC class I region. The locations of all markers used are displayed under the map, with the distance from each marker to the *HLA-B* gene, which is represented by kilobase pair (kb). Squares denote genotypic differentiation test; circles denote exact test of HWP.

Table 1

Statistically Significant Alleles Associated with Behçet Disease

Marker	No. of Alleles	Significant Allele	No. (%) of Patients	No. (%) of Controls	Odds Ratio (95% Confidence Interval)	χ^2	*P*	*Pc*
C1-2-A	13	182	22 (30)	17 (13)	2.86 (1.43–5.75)	8.77	.003	.04
MICA $(GCT)_n$	5	A6	56 (76)	62 (47)	3.51 (1.89–6.52)	15.97	.00006	.0003
MIB	12	336	7 (9)	40 (30)	.24 (.11–.55)	11.98	.0005	.006
		344	21 (28)	10 (8)	4.83 (2.23–10.47)	16.05	.00006	.00068
		348	33 (45)	16 (12)	5.84 (3.02–11.30)	27.58	.0000002	.000002
C1-4-1	6	217	65 (88)	87 (66)	3.74 (1.78–7.94)	11.79	.0006	.0036
C1-2-5	8	178	10 (14)	1 (1)	20.47 (4.48–93.42)	15.26	.00018[a]	.0033
		202	33 (45)	21 (16)	4.25 (2.26–8.01)	20.17	.0000071	.000057
C1-3-1	6	293	6 (8)	31 (23)	.28 (.12–.70)	7.61	.0058	.035
C2-4-4	11	255	28 (38)	22 (17)	3.04 (1.60–5.79)	11.56	.00067	.0074
C3-2-11	17	213	28 (38)	27 (20)	2.37 (1.27–4.43)	7.32	.0068	.116
HLA-B	25	B51	47 (64)	18 (14)	11.02 (5.82–20.87)	54.62	1.5×10^{-13}	3.7×10^{-12}

[a] Obtained by the Fisher's exact test (all other values were obtained by the χ^2 test).

crosatellite alleles observed in each locus (*Pc*). A level of *Pc* <.05 was accepted as statistically significant. The exact test of Hardy-Weinberg proportion (HWP) for multiple alleles and the genotypic differentiation test were performed by the Markov chain method within the GENEPOP software package (Guo and Thompson 1992; Raymond and Rousset 1995, 1997). The Markov chain method has the advantage of obtaining a complete enumeration for testing HWP in cases where the number of alleles is small and the sample size is small. A level of *P* <.1 or *P* <.01 was accepted as statistically significant for the Hardy-Weinberg equilibrium (HWE) or genotypic differentiation test, respectively (Guo and Thompson 1992; Raymond and Rousset 1995).

Results and Discussion

In the course of large-scale genome sequencing, we identified 533 microsatellite loci from dinucleotide to pentanucleotide repeats within the centromeric half of the HLA class I region spanning 1.3 Mb from the IkBL gene to the *HLA-92(L)* gene, which includes the *HLA-B, -C,* and *-E* genes (Mizuki et al. 1997*b*; Shiina et al. 1998; Tamiya et al. 1998). Of the 533 microsatellites, 27, including one (MIB) reported elsewhere (Grimaldi et al. 1996), were characterized and found to serve as an informative polymorphic genetic marker in a Japanese population (PIC: polymorphism content value = 0.69, average number of alleles = 9.6) (Tamiya et al. 1998). Here, eight markers were selected, by means of fluorescent-based automated fragment detection technique after PCR amplification, to detect repeat polymorphism in the Japanese patients with Behçet disease. These markers were densely distributed at the appropriate distance (C1-2-A: 147-kb centromeric, $[GCT]_n$ in exon 5 of *MICA*: 46 kb centromeric, MIB: 24 kb centromeric, C1-4-1: 6 kb centromeric, C1-2-5: 62 kb telomeric, C1-3-1: 111 kb telomeric, C2-4-4: 282 kb telomeric, C3-2-11: 767 kb telomeric) from the *HLA-B* locus (fig. 1). All eight markers indicated important genetic information for association analysis (heterozygosity/C1-2-A: 0.82, MICA-$[GCT]_n$: 0.82, MIB: 0.86, C1-4-1: 0.76, C1-2-5: 0.93, C1-3-1: 0.67, C2-4-4: 0.67, C3-2-11: 0.92). One or more alleles showing statistically significant differentiation in the patients, compared with the control group, were found at each locus (table 1). Among them, four were remarkably associated with the disease: A6 of MICA (χ^2 = 15.97, *Pc* = .0003), 344 (χ^2 = 16.05, *Pc* = .00068) and 348 (χ^2 = 27.58, *Pc* = .000002) of MIB, and 202 (χ^2 = 20.17, *Pc* = .000057) of C1-2-5. All the alleles in each microsatellite marker were named on the basis of the amplified fragment length, except for GCT repetitive polymorphism in the *MICA* gene (Mizuki et al. 1997*a*; Tamiya et al. 1998). The exact test of HWP was done, by means of the Markov chain method, at the eight microsatellite loci, as well as at the *HLA-B* locus in the patient and control groups (Guo and Thompson 1992). All nine markers tested, including *HLA-B,* were under the HWE (*P* > .4) in the healthy control group, whereas the four loci were significantly deviated from the HWE in the patient group (*P* < .1: C1-2-A, *MICA*, MIB, C1-4-1, *HLA-B*) (table 2). These results clearly suggest that the Behçet disease gene resides in close proximity to, probably centromeric of, *HLA-B,* because in the patient group the significant *P* values were obtained at only the four loci neighboring on the centromeric side of *HLA-B,* but not at the distant four loci located >62 kb telomeric of *HLA-B*. Genotypic differentiation concerned with the allelic distribution was also analyzed for the disease and the control groups. Generally, genotypic distribution should be identical within the same ethnic population. If genotypic differentiation is observed at a locus between the disease and control groups, allelic distribution at its locus should be

Table 2

Exact Test of HWP for Microsatellites and HLA-B Alleles

	P[a] (Standard Error)	
Marker	Controls	Patients
C1-2-A	.940 (.0011)	.034 (.0022)
MICA	.646 (.0007)	.099 (.0013)
MIB	.602 (.0019)	.027 (.0013)
C1-4-1	.451 (.0011)	.020 (.0011)
HLA-B	.841 (.0093)	.009 (.0023)
C1-2-5	.949 (.0007)	.423 (.0075)
C1-3-1	.721 (.0010)	.431 (.0001)
C2-4-4	.676 (.0033)	.749 (.0028)
C3-2-11	.62 (.0030)	.332 (.0090)

[a] In the exact P test by the Markov chain method, the dememorization period was 1,000 steps. P values <.1, regarded as statistically significant, are underlined.

under the influence of some genetic bias. The P values calculated by the Markov chain method were significantly low at the seven loci (*MICA,* MIB, C1-4-1, *HLA-B,* C1-2-5, C1-3-1, C2-4-4: $P < .01$) (table 3). *MICA,* MIB, *HLA-B,* and C1-2-5 showed the remarkably low $P < .00001$. These markers are embedded in the 108-kb interval between *MICA* and C1-2-5.

Taken together, these results suggest that the pathogenic gene for developing Behçet disease is most likely found within the 46-kb segment (*MICA–HLA-B*) on the centromeric side of the HLA-B locus, because only the four loci (*MICA,* MIB, C1-4-1, and *HLA-B*) on this segment provide significant evidence for association to the Behçet disease gene with the remarkably low P values in both the HWE and genotypic differentiation tests. These results are consistent with our previous findings that no allele in the *HLA-C* gene displays significant association with the disease when excluding *HLA-B51*-linked *HLA-C* alleles (Mizuki et al. 1996), that no expressed gene has been found between the *HLA-B* and *HLA-C* genes (Mizuki et al. 1997*b*; T. Shiina, N. Mizuki, K. Yabuki, S. Ohno, H. Inoko, unpublished data), and that there is a recombination hot spot just centromeric of the *MICA* gene (Mizuki et al. 1997*a*; Ota et al. 1997; Tamiya et al. 1998).

In our analysis elsewhere on the Tau-a microsatellite polymorphism located between the *TNF* and *MICA* genes (Mizuki et al. 1995), we suggested that the pathogenic gene responsible for the development of Behçet disease was not the *HLA-B* gene itself, but rather other genes between *TNF* and *HLA-B*. In fact, the alanine repeat polymorphism in the transmembrane region of the *MICA* gene localized between *TNF* and *HLA-B* was found to be significantly associated with the disease (Mizuki et al. 1997*a*). Further, we have reported the identification of three genes—*NOB1, NOB2,* and *NOB3* (new organization associated with *HLA-B*)—in this critical 46-kb segment between *MICA* and *HLA-B* (Mizuki et al. 1997*b*), but all of them have recently been established to be pseudogenes, with their expressed homologues on other chromosomes (K. Yabuki, N. Mizuki, T. Shiina, S. Ohno, H. Inoko, unpublished data). Collectively, the *MICA* gene is suggested to be a primary locus involved in the development of Behçet disease, in terms of its predicted function as an immunoregulatory molecule recognized by $\gamma\delta$ T cells (Groh et al. 1998) and its expressed tissues, such as epithelial cells and fibroblasts (Groh et al. 1996), where the inflammation of Behçet disease exclusively takes place. However, the possibility that *HLA-B* contributes to the pathogenesis of the disease as an additional or complementary risk factor cannot be excluded.

The present study also shows that association analysis using numerous microsatellite markers around a candidate gene region in a general population provides valuable information on focusing on a disease's critical region on the human genome. Association analysis is easier and more practical than methods in which affected family members are used. Association analysis may also prove effective in the analysis of sporadic, multifactorial, and common genetic diseases.

Acknowledgments

This work was supported by a grant from the Ministry of Health and Welfare of Japan.

Electronic-Database Information

Accession number and URL for data in this article are as follows:

Online Mendelian Inheritance in Man (OMIM), http://

Table 3

Genotypic Differentiation between the Control and Patient Groups

Marker	P[a] (Standard Error)
C1-2-A	.19450 (.00324)
MICA	.00001 (.00001)
MIB	.00000 (.00000)
C1-4-1	.00022 (.00005)
HLA-B5	.00000 (.00000)
C1-2-5	.00000 (.00000)
C1-3-1	.00011 (.00003)
C2-4-4	.00402 (.00029)
C3-2-11	.03727 (.00147)

Note.—In the genotypic differentiation test by the Markov chain method, the dememorization period was 1,000 steps.

[a] P values <.01, regarded as statistically significant, are underlined.

www.ncbi.nlm.nih.gov/Omim (for Behçet disease [MIM 109650]).

References

Grimaldi M-C, Clayton J, Pontarotti P, Cambon-Thomsen A, Crouau-Roy B (1996) New highly polymorphic microsatellite marker in linkage disequilibrium with HLA-B. Hum Immunol 51:89–94

Groh V, Bahram S, Bauer S, Herman A, Beauchamp M, Spies T (1996) Cell stress regulated human MHC class I gene expressed in gastrointestinal epithelium. Proc Natl Acad Sci USA 93:12445–12450

Groh V, Steinle A, Bauer S, Spies T (1998) Recognition of stress-induced MHC molecules by intestinal epithelial γδ T cells. Science 279:1737–1740

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372

Mizuki N, Ando H, Kimura M, Ohno S, Miyata S, Yamazaki M, Tashiro H, et al (1997*b*) Nucleotide sequence analysis of the HLA class I region spanning the 237-kb segment around the HLA-B and -C genes. Genomics 42:55–66

Mizuki N, Ohno S, Ando H, Kimura M, Ishihara M, Miyata S, Nakamura S, et al (1996) HLA-C genotyping in the patients with Behçet disease. Hum Immunol 50:47–53

Mizuki N, Ohno S, Sato T, Ishikawa M, Miyata S, Nakamura S, Naruse T, et al (1995) Microsatellite polymorphism located between the TNF and HLA-B genes in Behçet disease. Hum Immunol 43:129–135

Mizuki N, Ota M, Kimura M, Ohno S, Ando H, Katsuyama Y, Yamazaki M, et al (1997*a*) Triplet repeat polymorphism in the transmembrane region of the MICA gene: a strong association of six GCT repetitions with Behçet disease. Proc Natl Acad Sci USA 94:1298–1303

Ohno S, Ohguchi M, Hirose S, Matsuda H, Wakisaka A, Aizawa H (1982) Close association of HLA-Bw51 with Behçet disease. Arch Ophthalmol 100:1455–1458

Ota M, Katsuyama Y, Mizuki N, Ando H, Furihata K, Ohno S, Pivetti-Pezzi P, et al (1997) Trinucleotide repeat polymorphism within exon 5 of the MICA gene (MHC class I chain-related gene A): allele frequency data in the nine population groups Japanese, Northern Han, Hui, Uygur, Kazakhstan, Iranian, Saudi Arabian, Greek and Italian. Tissue Antigens 49:448–454

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49:1280–1283

Raymond M, Rousset F (1997) Genepop (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

ShiinaT, Tamiya G, Oka A, Yamagata T, Yamagata N, Kikkawa E, Goto K, et al (1998) Nucleotide sequencing analysis of the 146-kilobase segment around the IkNL and MICA genes at the centromeric end of the HLA class I region. Genomics 47:372–382

Tamiya G, Ota M, Katsuyama Y, Shiina T, Oka A, Makino S, Kimura M, et al (1998) Twenty-six new polymorphic microsatellite markers around the HLA-B, -C and -E loci in the human MHC class I region. Tissue Antigens 51:337–346