Title: Replication of a microsatellite genome-wide association study of Behcet disease in a Korean population

Yukihiro Horie¹, Akira Meguro², Nobuyoshi Kitaichi³,⁴, Eun Bong Lee⁵, Atsuhiro Kanda⁴, Kousuke Noda¹, Yeong Wook Song⁵, Kyung Sook Park⁶, Kenichi Namba¹, Masao Ota⁷, Hidetoshi Inoko⁸, Nobuhisa Mizuki², Susumu Ishida¹, and Shigeaki Ohno⁴

1. Department of Ophthalmology, Hokkaido University Graduate School of Medicine, Sapporo Japan
2. Department of Ophthalmology, Yokohama City University School of Medicine, Yokohama, Japan
3. Department of Ophthalmology, Health Sciences University of Hokkaido, Sapporo, Japan
4. Department of Ocular Inflammation and Immunology, Hokkaido University Graduate School of Medicine, Sapporo, Japan
5. Department of Internal Medicine and Graduate Program in Immunology, Medical Research Center, Seoul National University Hospital, Seoul, Korea
6. Department of Biology, College of Natural Sciences, Sungshin Women's University, Seoul, Korea
7. Department of Legal Medicine, Shinshu University School of Medicine, Matsumoto, Japan
8. Department of Molecular Life Science, Division of Molecular Medical Science and Molecular Medicine, Tokai University School of Medicine, Isehara, Japan

Address correspondence to: Nobuyoshi Kitaichi, MD, PhD
Department of Ocular Inflammation and Immunology, Hokkaido University Graduate School of Medicine, Sapporo, Kita-15, Nishi-7, Kita-ku, Sapporo 060-8638, Japan.
E-mail: nobukita@med.hokudai.ac.jp
Fax: +81-11-706-5948
Tel: +81-11-706-5944
Running head:

Microsatellite GWAS of Behcet disease

Key messages

- Replication study of the microsatellite GWAS was performed.
- *HLA-B*51 gene was strongly associated with Behcet disease in Korean population.
- One of the six microsatellite marker was significantly associated with Korean population.
Abstract

Objectives: Behcet disease is one of the major etiologies of uveitis causing blindness in Asian countries. A genome-wide association study identified 6 microsatellite markers as disease susceptibility loci for Japanese patients with Behcet disease. To confirm our recent results, these microsatellite markers were examined in Korean population as a replication study.

Methods: Study participants included 119 Behcet patients and 141 controls. All were enrolled in Korea. Association between the 6 reported microsatellite markers (D3S0186i, D6S0014i, D6S0032i, 536G12A, D12S0645i, and D22S0104i) and Behcet disease was analyzed. HLA-B was genotyped by sequence-based typing methods.

Results: A microsatellite marker located near the HLA-B region demonstrated significant association with Behcet disease ($P = 0.028$). The genotype and phenotype frequencies of the HLA-B*51 gene were significantly increased in patients (23.1% and 39.5%, respectively) compared with healthy controls (11.2% and 20.1%, respectively; $P < 0.001$).

Conclusions: Microsatellite analysis revealed that the HLA-B*51 gene was strongly associated with Behcet disease in a Korean population.

Key words: Microsatellite, GWAS, Behcet disease, replication study, HLA-B, Korean
**Introduction**

Behcet disease is a chronic, systemic, inflammatory disorder characterized by 4 major symptoms consisting of oral aphthous ulcers, genital ulcers, skin lesions, and recurrent ocular inflammation [1]. The disease is occasionally associated with inflammation in the vascular and/or central nervous systems and joints. Behcet disease is found predominantly between East Asia and the Mediterranean basin along the historic Silk Road; however, it is uncommon in the American continents, Oceania, and sub-Saharan Africa [2]. The distribution of intraocular inflammation and uveitis also differ in different regions of the world [3, 4].

In 1973, the first genetic factor was reported between Behcet disease and human leukocyte antigen (HL-A) 5 [5]. The nomenclature of HL-A5 was later changed to HLA-B5. Today, HLA-B5 is comprised of 3 subantigens: HLA-B51, -B52, and -B53. Susceptibility to Behcet disease is strongly associated with *HLA-B*51, as reported in different ethnic groups [2, 5, 6]. Populations with a high prevalence of *HLA-B*51 lie predominantly north of the equator, spanning Japan and Western Europe between 30° and 45° N [7]. The area is completely consistent with the countries where Behcet disease is common. Meanwhile, many other susceptibility genes, such as *HLA-A, -E, -F, and -G; tumor necrosis factor (TNF) -a; toll like receptor (TLR) 4; interleukin (IL)-1, -8, -10, -12, and -18; IL23R; and CD28* have been reported in relation to Behcet disease [8-13]. Most recently, reports from ourselves and others of genome-wide association studies (GWAS) using 500,000 single nucleotide polymorphism (SNP) microarrays have identified 2 new disease susceptibility loci for Behcet disease on Chr.1p31.3 and 1q32.1, aside from the HLA class I region, in Japanese, Turkish, Korean, UK Caucasian, and Greek populations [11, 12]. These results are compatible with the fact that approximately one-half of Behcet patients are *HLA-B*51-negative, and most *HLA-B*51 carriers never suffer from Behcet disease in their lifetime [2, 14].

However, there is little doubt but that the *HLA-B* region still demonstrates the strongest association with Behcet disease among all ethnic groups. A GWAS using microsatellite markers demonstrated that 6 (D3S0186i, D6S0014i, D6S0032i, 536G12A, D12S0645i, D22S0104i) of
23,465 markers differed significantly between patients and healthy subjects by using pooled DNA and individual typing methods in the Japanese population [8]. One of these significant 6 markers is located near the HLA-B region. Therefore, it is critically important to examine these indicated microsatellite markers among the people along the historic Silk Road, considering the prominent high prevalence of both the disease and the HLA-B*51 gene there. In the present study, these microsatellite markers were investigated in a Korean population as a replication study, and the association was also examined between clinical features and the gene frequencies.

Methods

Participants

In the present study, 119 Behcet patients and 141 healthy controls were enrolled. All the patients were from Seoul National University Hospital, Seoul, Korea and fulfilled the diagnostic criteria of the International Study Group for Behcet disease (ISGBD). Informed consent was obtained from all participating individuals. The ethical committee of Seoul National University Hospital approved the study. The procedures used conformed to the tenets of the Declaration of Helsinki. All of the patients and control subjects were Korean.

Genotyping

Genomic DNA was prepared from peripheral blood specimens using the QIAamp DNA Blood Mini Kit (Qiagen, Tokyo, Japan). Each polymerase chain reaction (PCR) product for the 6 MS markers (D3S0186i, D6S0014i, D6S0032i, 536G12A, D12S0645i, D22S0104i) was amplified by PCR reactions. Each MS marker was amplified using two primers: forward 5’- AGC TCT TCC TAA CTG ATA AGG AAG -3’ and reverse 5’- GTA AAG GTT GCT AGG TCC TGT T -3’ for D3S0186i, forward 5’- CCA TAT GCT AGA AAT TAT GGT ACT -3’ and reverse 5’- GTT TCA CTG TGT TGG CCA G -3’ for D6S0014i, forward 5’- TAA GTC TAA GAA TGT GAG ACC AAC -3’ and reverse 5’- GTA ATG CTG ATA ACG TTT ACT GTC -3’ for D6S0032i, forward 5’- GTG TGC
TTG TGT CTG TTA ATT G -3’ and reverse 5’- ACA CTA TAT TGT TAG CAA GTT ACT GAA C -3’ for 536G12A, forward 5’- GGC AGA GAC AGT GTC TTT CTC -3’ and reverse 5’- AGG TCA AGT GCA TGT TTG AC -3’ for D12S0645i, and forward 5’- TAA GGC TGA GTA GCA GTC TAC ATA -3’ and reverse 5’- TCA TTA AAG AAC TGG ATC CAT -3’ for D22S0104i. The reaction mixture was subjected to 5 min at 94°C; followed by 35 cycles of 1 min for denaturing at 94°C, 1 min for annealing at 57°C or 60°C, and 2 min for extension at 72°C; and 10 min for final elongation at 72°C using a PCR thermal cycler, GeneAmp System 9700 (Applied Biosystems, Foster City, CA, USA). Each forward primer was labeled at the 5’ end with 6-FAM, NED, or VIC to determine the number of MS repeats. Fragment length analysis was performed using an ABI3130 DNA sequencer (Applied Biosystems) and the number of MS repeats was estimated with GeneMapper v3.5 software using GS500(-250)Liz (Applied Biosystems) as a size marker.

Genotyping of the HLA-B gene was performed using PCR sequence-based typing methods, and the data were analyzed using Assign software. HLA-B typing by sequence-based typing methods was performed by using HLA typing kit (Allele SEQR HLA-B®, Abbott, Japan).

Statistical analysis
Allele frequencies were calculated by direct counting. The significance of the association was assessed using the chi-square test. The strength of all P-values was derived from a two-sided test; P < 0.05 was considered statistically significant.

Results
There was a significant difference between patients and healthy controls in one of 6 microsatellite markers. The allele frequency of “308” of the microsatellite marker D6S0032i was statistically higher among Behcet patients (14.3%) compared with healthy subjects (8.2%) [P = 0.028, odds ratio (OR) = 1.86, 95% confidence interval (CI) = 1.06-8.66] (Table 1).
Genotype and phenotype frequencies at the *HLA-B* locus were examined, and 23 antigens were identified in 119 patients and 141 controls. The genotype frequency of *HLA-B*<sup>51</sup> was significantly higher among patients (23.1%) than among healthy controls (11.2%) \( P = 0.0003, \text{OR} = 2.39, \text{95% CI} = 1.48-9.23 \) [in particular, *HLA-B*<sup>5101</sup> \( P = 0.0006, \text{OR} = 2.31, \text{95% CI} = 1.42-9.12 \)] (Table 1). The phenotype frequency of *HLA-B*<sup>51</sup> was significantly higher among patients (39.5%) than among healthy controls (20.1%) \( P = 0.00064, \text{OR} = 2.59, \text{95% CI} = 1.49-9.74 \) \( \text{(Table 1)} \).

**Discussion**

To confirm recent GWAS studies in Japan [8], it is important to repeat population specific analysis and other population data analysis. In the present study, we determined the reproducibility of the recent Japanese GWAS results, and successfully found the new disease susceptibility gene by using microsatellite markers in a Korean population. Microsatellite polymorphisms have been used in the genetic association study because they are highly polymorphic, relatively simple, and inexpensive [15]. Recent technologies have reduced the necessary time and cost to genotype SNPs [16]. However, because of their lower variability, SNPs may be considered as carrying relatively less information at each locus. Furthermore, the microsatellite mutation rate was estimated to be \( 10^{-2} - 10^{-5} \) each generation; more frequent than that of SNPs \( 2.5 \times 10^{-8} \) [17]. That is to say, although microsatellite analysis has the great advantages of higher polymorphism and sensitivity, it also carries the disadvantage of relatively high genotyping error rates due to its highly polymorphic nature [18]. Therefore, it is necessary to replicate the obtained susceptible genes in other ethnic groups by other means. This is why we have organized the present study.

The marker D6S0032i, located at 1.1 Mb telomeric of *HLA-B* within the HLA class I region, showed statistical significance in Korean Behcet patients. However, there was no significant difference between patients and controls for the other 5 markers located near the *ROBO1* (roundabout, axon guidance receptor, homolog 1), *HLA-L and B, PPIL4* (peptidylprolyl isomerase—
like 4), \textit{SOX5} (sex determining region Y–box 5), and \textit{IGLV1-40 (IGL variable 1-40)} genes. The previous Turkish GWAS using fewer than 400 microsatellite markers reported chromosomes 12p12–13 and 6p22–24 as the susceptible loci for Behcet disease [19]. However, our recent GWAS in Japanese people using 23,465 microsatellite markers disagreed with that Turkish study.

In this type of study, however, there will be sources of bias that may distort the results. In this study, we have identified the following possibilities:

1. Classification bias: We adopted International Study Group (ISG) Criteria in this study, however, Japanese GWAS study adopted Japan Behcet Disease Research Committee’s Criteria. The use of different standardised criteria may have led to misclassification when comparing the frequencies of systemic features.

2. Ascertainment bias: All samples in the current work were collected by rheumatologists, however in the Japanese GWAS, patients were mainly recruited at eye clinics. The patients with ocular lesions was significantly fewer in this study (37.8%) than recent Japanese GWAS study (89.3%) \((P =0.0000067)\) [9]. It is possible that some patients with quite mild ocular inflammation might be misidentified as no ocular manifestations at rheumatology clinics when they did not consult ophthalmologists.

There are 3.61% of genetic differences between Japanese and Koreans [20]. Additional replication studies in other ethnic groups may be required for better understanding the mechanisms, etiology, and characteristics of Behcet disease.

In conclusion, we performed a microsatellite analysis in a population of Korean Behcet patients as a replication study, and demonstrated significant association with the marker located near the \textit{HLA-B} region. The \textit{HLA-B*51} gene is significantly associated with patients with ocular manifestations.
Acknowledgements

The authors thank clinical staffs, patients, and their families for their participation in and dedication to the project.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Statement

This work was supported by grants-in-aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the Ministry of Health, Labour and Welfare, Japan.
References


Table 1. Genotype frequencies of the D6S0032i, and genotyped and phenotyped HLA-B allele frequencies in the patients and healthy controls of the Korean population

<table>
<thead>
<tr>
<th>D6S0032i</th>
<th>Patients 2n=238</th>
<th>%</th>
<th>Controls 2n=280</th>
<th>%</th>
<th>Odds (95%CI)</th>
<th>$\chi^2$</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td>304</td>
<td>7</td>
<td>2.9</td>
<td>10</td>
<td>3.6</td>
<td>0.82 (0.31-8.80)</td>
<td>0.16</td>
<td>0.69</td>
</tr>
<tr>
<td>308</td>
<td>34</td>
<td>14.3</td>
<td>23</td>
<td>8.2</td>
<td>1.86 (1.06-8.66)</td>
<td>4.84</td>
<td>0.028</td>
</tr>
<tr>
<td>313</td>
<td>109</td>
<td>45.8</td>
<td>163</td>
<td>58.2</td>
<td>0.61 (0.43-7.64)</td>
<td>7.95</td>
<td>0.0048</td>
</tr>
<tr>
<td>321</td>
<td>88</td>
<td>37</td>
<td>84</td>
<td>30</td>
<td>1.37 (0.95-7.56)</td>
<td>2.82</td>
<td>0.093</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Patients 2n=238</th>
<th>%</th>
<th>Controls 2n=278</th>
<th>%</th>
<th>Odds (95%CI)</th>
<th>$\chi^2$</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td>B*51</td>
<td>55</td>
<td>23.1</td>
<td>31</td>
<td>11.2</td>
<td>2.39 (1.48-9.23)</td>
<td>13.2</td>
<td>0.00028</td>
</tr>
<tr>
<td>B*5101</td>
<td>52</td>
<td>21.8</td>
<td>30</td>
<td>10.8</td>
<td>2.31 (1.42-9.12)</td>
<td>11.73</td>
<td>0.00062</td>
</tr>
<tr>
<td>B*5102</td>
<td>3</td>
<td>1.3</td>
<td>1</td>
<td>0.4</td>
<td>3.54 (0.22-10.0)</td>
<td>1.35</td>
<td>0.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>n=119</th>
<th>%</th>
<th>n=139</th>
<th>%</th>
<th>Odds (95%CI)</th>
<th>$\chi^2$</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td>B*51</td>
<td>47</td>
<td>39.5</td>
<td>28</td>
<td>20.1</td>
<td>2.59 (1.49-9.74)</td>
<td>11.64</td>
<td>0.00064</td>
</tr>
</tbody>
</table>

* Significant p values are in bold