

## ORIGINAL ARTICLE

# Targeted resequencing of candidate genes reveals novel variants associated with severe Behçet's uveitis

Sang Jin Kim<sup>1,2,8</sup>, Seungbok Lee<sup>3,4,8</sup>, Changho Park<sup>4</sup>, Jeong-Sun Seo<sup>3,4,5,6</sup>, Jong-Il Kim<sup>3,4,5,7</sup> and Hyeong Gon Yu<sup>1</sup>

Behçet's disease (BD) is a chronic systemic inflammatory disorder characterized by four major manifestations: recurrent uveitis, oral and genital ulcers and skin lesions. To identify some pathogenic variants associated with severe Behçet's uveitis, we used targeted and massively parallel sequencing methods to explore the genetic diversity of target regions. A solution-based target enrichment kit was designed to capture whole-exonic regions of 132 candidate genes. Using a multiplexing strategy, 32 samples from patients with a severe type of Behçet's uveitis were sequenced with a Genome Analyzer IIx. We compared the frequency of each variant with that of 59 normal Korean controls, and selected five rare and eight common single-nucleotide variants as the candidates for a replication study. The selected variants were genotyped in 61 cases and 320 controls and, as a result, two rare and seven common variants showed significant associations with severe Behçet's uveitis ( $P < 0.05$ ). Some of these, including rs199955684 in *KIR3DL3*, rs1801133 in *MTHFR*, rs1051790 in *MICA* and rs1051456 in *KIR2DL4*, were predicted to be damaging by either the PolyPhen-2 or SIFT prediction program. Variants on *FCGR3A* (rs396991) and *ICAM1* (rs5498) have been previously reported as susceptibility loci of this disease, and those on *IFNAR1*, *MTFHR* and *MICA* also replicated the previous reports at the gene level. The *KIR3DL3* and *KIR2DL4* genes are novel susceptibility genes that have not been reported in association with BD. In conclusion, this study showed that target enrichment and next-generation sequencing technologies can provide valuable information on the genetic predisposition for Behçet's uveitis.

*Experimental & Molecular Medicine* (2013) 45, e49; doi:10.1038/emm.2013.101; published online 18 October 2013

**Keywords:** Behçet's disease; next-generation sequencing; uveitis

## INTRODUCTION

Behçet's disease (BD) is a systemic inflammatory disease characterized by recurrent oral aphthous ulcers, genital ulcers, skin lesions and ocular involvement.<sup>1,2</sup> Although uncommon, BD may cause serious manifestation including systemic vasculitis, central nervous system involvement and sight-threatening uveitis.<sup>1,2</sup>

The etiology of BD is not understood fully, but some environmental factors are thought to trigger the immunopathogenesis of BD in people with a genetic predisposition. Innate immunity is involved in inflammation along with the

adaptive immune system, most importantly T cells.<sup>1,2</sup> *HLA-B51* is the most extensively investigated genetic risk factor for BD.<sup>1-4</sup> A strong association of *HLA-B51* with BD and similarities in the geographic distribution of BD and the susceptibility alleles have been reported.<sup>5</sup> In addition to *HLA-B51*, many genetic variations have been reported to be associated with BD, including *HLA-A26*, *MICA*, *TNF*, *TLR4* and others.<sup>6-8</sup> Recent genome-wide association studies have revealed an association at the *IL10* and *IL23R-IL12RB2* loci.<sup>9,10</sup>

Although many genome-wide association studies based on single-nucleotide polymorphism (SNP) genotyping have

<sup>1</sup>Department of Ophthalmology, Seoul National University College of Medicine, Seoul, South Korea; <sup>2</sup>Department of Ophthalmology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea; <sup>3</sup>Genomic Medicine Institute (GMI), Medical Research Center, Seoul National University, Seoul, South Korea; <sup>4</sup>Department of Biomedical Sciences, Seoul National University Graduate School, Seoul, South Korea; <sup>5</sup>Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Seoul, South Korea; <sup>6</sup>Macrogen, Seoul, South Korea and <sup>7</sup>Psoma Therapeutics, Seoul, South Korea

<sup>8</sup>These authors contributed equally to this work.

Correspondence: Professor J-I Kim, Genomic Medicine Institute, Medical Research Center, Seoul National University College of Medicine, 101 Daehang-ro, Jongno-gu, Seoul 110-799, South Korea.

E-mail: jongil@snu.ac.kr

or Professor HG Yu, Department of Ophthalmology, Seoul National University College of Medicine, 101 Daehang-ro, Jongno-gu, Seoul 110-744, South Korea.

E-mail: hgonyu@snu.ac.kr

Received 17 July 2013; accepted 29 July 2013

suggested the existence of several significant loci in complex diseases including BD, only a small fraction of the heritable variations have been identified. One possible explanation is that many different rare or pathogenic variants, which are difficult to find using SNP microarray experiments, contribute substantially to the genetic susceptibilities of these diseases. Therefore, it has been suggested that sequencing of candidate genes would be an efficient method to investigate the contribution of previously unrecognized rare or pathogenic variants to the target disorder.

Whole-exome sequencing or targeted sequencing with massively parallel or next-generation sequencing (NGS) technology has shown promising results in several diseases, mostly those with Mendelian inheritance such as Duchenne and Becker muscular dystrophy.<sup>11–14</sup> Exome or candidate gene sequencing studies of complex diseases have also been undertaken.<sup>15,16</sup> Herein, to investigate the genetic predisposition for BD with severe uveitis, we used a targeted and massively parallel sequencing method to explore the genetic diversity of exonic regions in 132 candidate genes. We also performed a replication experiment using pathogenic candidates selected from the targeted sequencing.

## MATERIALS AND METHODS

### Study subjects

The Institutional Review Board of the Seoul National University Hospital approved the study protocol. The enrolment of participants and blood collection were performed in accordance with the Declaration of Helsinki. Informed consent was obtained from all enrolled patients and healthy subjects. A total of 61 unrelated Korean individuals with BD were enrolled in this study. The diagnosis of BD was made by the criteria of the Behçet's Research Committee of Japan, and patients with either complete or incomplete types of BD were included.<sup>17</sup> Complete ophthalmological examinations including slit-lamp biomicroscopic examination, dilated fundus examination and fluorescein angiography were conducted for all patients. All patients showed severe chronic recurrent vision-threatening uveitis in both eyes and were followed up for more than 1 year at the uveitis clinic, Department of Ophthalmology, Seoul National University Hospital. For the replication study, samples were obtained from 320 healthy Korean subjects (controls) aged  $\geq 50$  years with no evidence of rheumatological disease or eye disease except senile cataract. All controls were unrelated to each other or to the patients.

### Genomic DNA extraction

After the subjects provided informed consent, genomic DNA was extracted from peripheral blood leucocytes using FlexiGene DNA kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Sample concentrations were measured by using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

### Targeted exonic sequencing of candidate genes

We selected 132 candidate genes that had been reported or suspected to play a role in the pathogenesis of Behçet's uveitis (Supplementary Table S1). In brief, the 132 genes included cytokines such as *IL6*, *IL10*, *IFNG*, *TNF* and so on; cytokine receptors such as *IFNGR1*; chemokines such as *CCL2*; other immune-related genes such as

*TLR4*, *NOD2* and so on; and previously reported BD-associated genes such as *F5*.

For targeted sequencing, 32 patient samples were used. Based on the RefSeq gene set, exonic regions of these genes were captured using a SureSelect Target Enrichment System kit (Agilent Technology, Palo Alto, CA, USA) and sequenced through a Genome Analyzer IIx (Illumina, San Diego, CA, USA). The detailed procedures were described in our previous report<sup>14</sup> and the only difference was that eight samples were applied to each capture kit.

The generated reads were aligned to National Center for Biotechnology Information (NCBI) human reference build 37 using the GSNAP (Genomic Short-read Nucleotide Alignment Program) alignment tool.<sup>18</sup> For further analyses, we used only the bases that were covered by  $\geq 8$  uniquely aligned reads with  $\geq 20$  mean *Q*-scores. Single-nucleotide variants (SNVs) and short insertions/deletions (indels) were called according to the algorithm in our previous study.<sup>19</sup>

### Selection of candidate variants

We compared the frequency of each variant with that of the control samples. The controls for this discovery step were the samples from 59 normal Koreans comprising eight whole-genome sequencing and 51 whole-exome sequencing samples.<sup>16,20</sup> For a relevant comparison, we used only variant positions covered well in  $\geq 20$  controls. In addition, we selected further pathogenic candidates among variants based on the gene annotation information: nonsynonymous SNVs, coding sequence indels and splicing site variants. The Reference Sequence (RefSeq) gene database was used for the gene annotation.

The comparison was conducted in two different ways. First, we focused on novel rare variants with allele frequencies  $< 0.01$  in the controls that were detected at the same time in  $\geq 2$  case subjects. Novelty was determined by comparison with dbSNP release 132. Second, the genotype frequency of each variant was compared in several ways between the case and control groups to screen for common candidates among the variants. We applied both dominant and recessive models for comparison, and we also compared the allele frequencies between the two groups. Fisher's exact test was used for all comparisons. A *P*-value of  $< 0.05$  was considered as the standard for candidate selection.

The selected rare and common candidates were checked manually in terms of base coverage depth. We excluded from further experiments those with  $> 2$  alleles or with a coverage depth ratio between the reference and variant alleles that deviated far from 0.5 throughout the samples. Finally, linkage disequilibrium (LD) was estimated among the final candidates; if strong LD was found ( $r^2 > 0.8$ ), we selected one among the related candidates for the further replication study.

### Replication study on candidate variants using additional case and control subjects

We genotyped the 13 candidate variants in all 61 case samples, including 31 samples that were sequenced in the first step, and 320 control samples from healthy Korean individuals using the TaqMan SNP genotyping assay (Applied Biosystems, Foster City, CA, USA) (Supplementary Table S2). Some missing genotypes from this experiment were checked through PCR and a subsequent Sanger sequencing method (Supplementary Table S3).

As we did when examining the discovery set, we also applied genotype frequency comparisons to each candidate in both the dominant and recessive models, along with allele frequency

comparison. The only difference was that the  $\chi^2$  test was used for the analyses of the common candidates. Odds ratios were calculated, and a *P*-value of <0.05 was considered significant.

Two computational algorithms, PolyPhen-2 and SIFT (Scale-Invariant Feature Transform), were used to predict the functional impacts of nonsynonymous SNVs in this study.<sup>21,22</sup> PolyPhen-2 distinguishes variants with drastic effects from other neutral variants, and the results for each variant were classified into two subcategories: damaging and benign. The output of SIFT showed a normalized probability score. Positions with normalized probabilities of <0.05 were predicted to be damaging, and those with normalized probabilities of  $\geq 0.05$  were predicted to be tolerated.

## RESULTS

### Characteristics of patients

The demographic and clinical characteristics of the patients with BD are shown in Table 1. All patients had a complete or incomplete type of BD with severe uveitis.

### Summary of targeted sequencing in patient samples

The sequencing summary of 32 patient samples is shown in Supplementary Table S4. The mean coverage depth was  $73.68 \times$  (median  $52.03 \times$ ), and the mean percentage of  $\geq 8 \times$  covered bases in the bait region was 71.77% (median 88.09%). Supplementary Table S5 summarizes the SNV and indel counts of each patient sample. As shown in both tables, five samples were very poorly covered by sequencing (<10%), and the numbers of variants called were markedly lower than those of other samples.

### Pathogenic candidate variants

Figure 1 shows the overall flow of this study, including the processes for candidate selection and the replication experiment. In total, we called 3139 SNVs and 269 indels from the targeted sequencing. Among them, 1489 SNVs and 79 indels

were well covered in  $\geq 20$  control samples, and 301 SNVs and 5 indels were identified as possibly pathogenic at the same time by the criteria described in the MATERIALS AND METHODS. We narrowed the pathogenic candidates further to 5 rare SNVs and 12 common SNVs by comparing the genotype frequencies between the case and control groups. Among these, 13 SNVs were finally chosen for the replication study with additional samples, and the detailed information of them is shown in Supplementary Table S6. Their allele frequencies were similar with those from the 1000 Genomes project, except three rare and two common variants without information.

### Replication study with additional study subjects

We genotyped the candidate variants in 61 cases and 320 controls using TaqMan SNP genotyping assay. The case subjects included 31 patient samples that had been sequenced at the discovery step and 99.0% of genotypes were shown to be concordant between the two genotyping methods.

As a result, two rare SNVs and seven common SNVs were shown to have significant associations with BD (Table 2). Their possible effects were consistent with those of the discovery step, which were predicted to be damaging for all candidates. Among the rare candidates, one nonsynonymous SNV in *KIR3DL3* (rs199955684) and one in *IFNAR1* had significantly higher frequencies in the BD group than in the controls. The substitution of threonine to isoleucine in *KIR3DL3* (T395I) was predicted to affect protein function by both PolyPhen-2 and SIFT. Other three rare SNVs also showed the *P*-values near 0.05. Among the common candidates, seven SNVs in *MTHFR* (rs1801133), *FCGR3A* (rs396991), *MICA* (rs1051790, rs61736348, rs41546114), *ICAM1* (rs5498) and *KIR2DL4* (rs1051456) had significantly higher allele frequencies in the BD group than in the controls. In case of three SNVs (rs1051790, rs41546114 and rs1051456), the recessive model showed more significant results than the dominant model, and the highest odds ratio (21.42) was found for rs41546114. The substitutions of alanine to valine in *MTHFR* (A222V) and leucine to valine in *MICA* (L145V) were also predicted to be damaging by both PolyPhen-2 and SIFT, and the P209A change of *KIR2DL4* was predicted to be damaging by PolyPhen-2. The variants on *FCGR3A* and *ICAM1* have been previously reported as susceptibility loci of BD and its ocular involvement.<sup>23–26</sup>

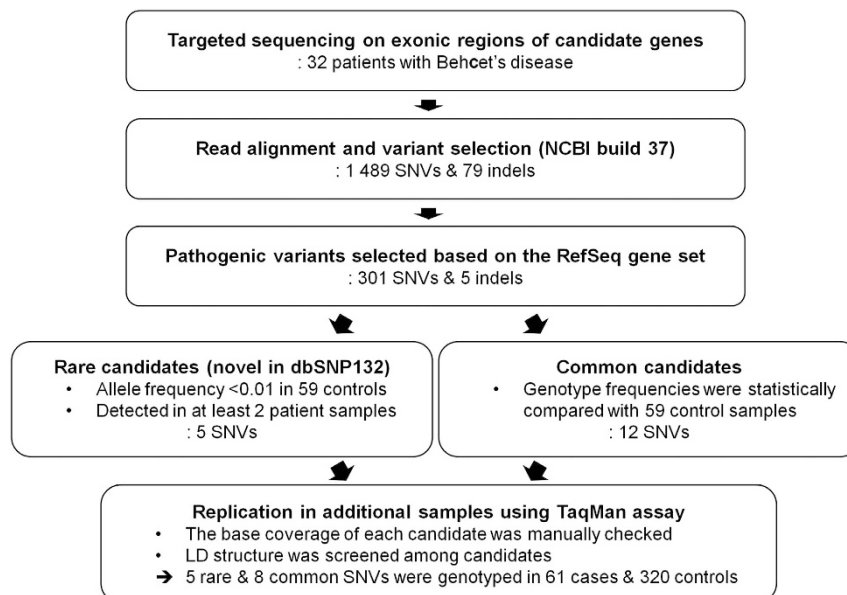
Figure 2 shows the distribution the 13 pathogenic candidates in patient samples. Rare SNVs seem to be widely distributed throughout the samples with few overlaps. In case of common candidates, rs1051790 and rs41546114 in *MICA* had moderate LD between them ( $r^2 = 0.56$ ), whereas others seem to be randomly distributed in case samples.

## DISCUSSION

To identify the genetic predisposition for BD, we used target enrichment and NGS methods to investigate the genetic diversity of exonic regions in 132 candidate genes, which was followed by a replication study. Among the 1489 SNVs

**Table 1 Demographic and clinical characteristics of the study subjects**

	Patients with Behçet's disease		Controls for replication study
	For targeted sequencing	For replication study	
Number	32	61	320
Age (years)	38.4 ± 10.4	39.3 ± 9.7	65.6 ± 9.1
Male (%)	23 (71.9)	42 (68.9)	130 (40.6)
<i>Behçet's disease</i>			—
Complete (%)	23 (71.9)	42 (68.9)	
Incomplete (%)	9 (28.1)	19 (31.2)	
<i>Uveitis (%)</i>	32 (100)	61 (100)	—
Panuveitis (%)	29 (90.6)	52 (85.3)	
Posterior uveitis (%)	3 (9.4)	9 (14.8)	
Oral ulcers (%)	27 (84.4)	59 (96.7)	—
Skin lesions (%)	31 (96.9)	54 (88.5)	—
Genital ulcers (%)	24 (75.0)	43 (70.5)	—



**Figure 1** Overall scheme of the experiment. This flow chart shows briefly all steps in the experiment of this study, including the processes of candidate selection and the additional replication experiment. indel, insertion/deletion; LD, linkage disequilibrium; SNV, single-nucleotide variant.

**Table 2** List of 13 nonsynonymous SNVs genotyped in 61 patients with Behçet's uveitis and 320 normal controls

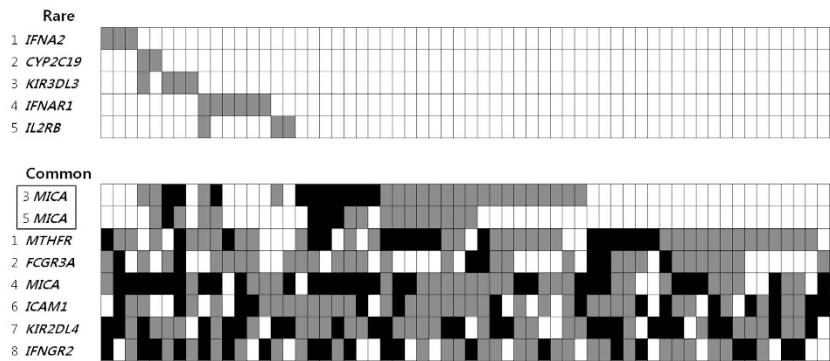
Chr	Position	Gene	Allele	dbSNP ID (build 137)	Dominant		Recessive		Allele frequency		SIFT	PolyPhen-2
					P-value	OR	P-value	OR	P-value	OR		
<i>Rare</i>												
9	21384971	<i>IFNA2</i>	C/T	rs146352658	0.087	4.05	—	—	0.088	3.97	Tolerated	Benign
10	96540403	<i>CYP2C19</i>	C/A	Novel	0.069	10.71	—	—	0.069	10.55	Damaging	Benign
19	55247514	<i>KIR3DL3</i>	C/T	rs199955684	<b>0.026</b>	5.49	—	—	<b>0.027</b>	5.34	Damaging	Damaging
21	34726042	<i>IFNAR1</i>	A/T	Novel	<b>0.001</b>	11.42	—	—	<b>0.001</b>	10.88	Tolerated	Benign
22	37524600	<i>IL2RB</i>	C/T	rs188869923	0.056	5.41	—	—	0.057	5.3	Tolerated	Benign
<i>Common</i>												
1	11856378	<i>MTHFR</i>	C/T	rs1801133	0.072	1.94	0.082	1.61	<b>0.025</b>	1.59	Damaging	Damaging
1	161514542	<i>FCGR3A</i>	G/T	rs396991	0.127	1.6	0.078	2.43	<b>0.047</b>	1.56	Tolerated	Benign
6	31378956	<i>MICA</i>	C/G	rs1051790	<b>0.041</b>	1.84	<b>0.022</b>	2.47	<b>0.005</b>	1.87	Damaging	Damaging
6	31379817	<i>MICA</i>	T/C	rs61736348	<b>0.012</b>	4.56	<b>0.016</b>	1.55	<b>0.002</b>	2.06	Tolerated	Benign
6	31382831	<i>MICA</i>	C/T	rs41546114	<b>0.002</b>	2.97	<b>0.001</b>	21.42	< <b>0.001</b>	3.06	—	—
19	10395683	<i>ICAM1</i>	A/G	rs5498	<b>0.025</b>	2.23	0.156	1.53	<b>0.018</b>	1.63	Tolerated	Benign
19	55317669	<i>KIR2DL4</i>	C/G	rs1051456	0.189	1.81	<b>0.027</b>	1.55	<b>0.022</b>	1.63	Tolerated	Damaging
21	34787312	<i>IFNGR2</i>	A/G	rs9808753	0.793	0.87	0.395	1.24	0.777	1.08	—	Benign

Abbreviations: OR, odds ratio; SIFT, Scale-Invariant Feature Transform; SNV, single-nucleotide variant. Significant *P*-values are highlighted in bold.

and 79 indels identified in target regions, 2 rare and 7 common SNVs showed significant associations with BD. In addition, we found that three pathogenic variants of *KIR3DL3* (T395I), *MTHFR* (A222V, rs1801133) and *MICA* (L145V, rs1051790) were predicted to be damaging by the two computational algorithms, PolyPhen-2 and SIFT.

We tried to find some functional candidates for the BD pathogenesis. The strategies for selecting functional candidates

in this study were as follows. First, we focused only on whole-exonic regions of 132 genes, most of which had been reported as susceptibility genes for BD. Second, we selected possible pathogenic variants with changes in the amino acid sequences of the corresponding proteins. For this strategy, we chose to apply the NGS technology because it could fully cover the exonic regions of several genes at the base level. However, when using the NGS method, successful analysis requires even



**Figure 2** Distribution of the 13 pathogenic candidates throughout the replication cases. Upper and lower plots show which kinds of rare and common pathogenic candidates were found for each sample, respectively. Grey squares indicate heterozygous variants and black squares indicate homozygous variants. The numbers assigned to each variant correspond to the order of variants listed in Table 2. Common 3 (rs1051790) and 5 (rs41546114) variants have moderate linkage disequilibrium with each other ( $r^2 > 0.5$ ).

and sufficient sequencing coverage throughout study samples. Unfortunately, five samples were scarcely covered, and some other samples also showed fluctuating coverage patterns (Supplementary Tables S4 and S5). Although this can be explained by many factors such as poor sample quality, technical problems and experimental errors, the most acceptable explanation would be the application of too many samples to one capture kit. Our previous study, which showed uniform coverage patterns throughout study samples, used the same experimental procedures as in this study; the only difference was that we applied four samples to one capture kit in our previous study and eight in this study.<sup>14</sup> In addition, if more sequencing throughput had been acquired, we would have expected more sufficient coverage even if the coverage patterns had fluctuated.

To overcome this uneven coverage in the discovery set, we first selected samples that satisfied our base quality criteria (bases covered by  $\geq 8$  uniquely aligned reads with  $\geq 20$  mean Q-scores) for each variant position. The statistical comparison was conducted next using only the selected samples for the position. In this way, variants that have too many missing genotypes and therefore high *P*-values would be excluded, and only variants with significant differences between two groups would remain as candidates. For the replication study, we genotyped all available cases again using TaqMan probes to check the accuracy of sequencing variant calls and to make up missing genotypes for more powerful analyses.

One rare SNV in *KIR3DL3* (rs199955684) and one common SNV in *KIR2DL4* (rs1051456) showed significant associations with BD in our study. Killer cell immunoglobulin-like receptors (KIRs) are members of the immunoglobulin superfamily that are expressed by natural killer (NK) cells and some T-cell subsets.<sup>27</sup> *KIR3DL3* expression is limited to CD56<sup>bright</sup> NK cells, which have a functional role in the innate immune response as the primary source of NK cell-derived immunoregulatory cytokines.<sup>27,28</sup> In BD, the number of NK cells increases in peripheral blood and aqueous humor.<sup>29,30</sup> NK cells may play an important role in controlling inflammation

in BD.<sup>31,32</sup> The *KIR3DL3* and *KIR2DL4* genes have rarely been suggested as susceptibility genes of BD in previous studies. The role of KIRs in BD has not been established, but the variations in these genes may affect the functions of NK cells and the pathogenesis of BD.

The *MTHFR* gene encodes methylene tetrahydrofolate reductase that uses folate to metabolize and thereby remove homocysteine.<sup>33</sup> The A222V polymorphism (rs1801133) in this gene is associated with reduced enzyme activity and therefore causes impaired remethylation of homocysteine to methionine along with subsequent hyperhomocysteinemia.<sup>33</sup> Homocysteine is suggested to be a risk factor for the hypercoagulability and thrombotic complications observed in BD patients.<sup>34,35</sup> The significant association of *MTHFR* with BD in our study is consistent with association between this variation (rs1801133) and an increased risk of ocular involvement in BD reported by Ozkul *et al.*<sup>36</sup>

In this study, the frequency of the *MICA*\*A5.1 allele (rs1051790, *MICA*\*00801) was significantly higher in patients with BD. *MICA* is located within the human major histocompatibility complex, 46 kb centromeric to the *human leukocyte antigen B (HLA-B)* gene.<sup>37</sup> A strong LD was observed between several alleles of *MICA* including *MICA*\*A6 and *HLA-B51* in both the patients with BD and normal controls.<sup>38,39</sup> Thus, the significant association of *MICA*\*A5.1 allele with BD might be caused by LD with *HLA-B51*, which is known to show a strong association with BD. However, because the LD between *MICA*\*A5.1 and *HLA-B51* has not been investigated, and HLA typing was not done in our study, the exact role of *MICA*\*A5.1 in BD remains unclear.

In addition to the variants predicted to be damaging at the prediction tools, there exist some more SNVs that showed the possible associations with BD in this study. One rare and novel variant in *IFNAR1* (N465Y) showed the lowest *P*-value in the dominant model among candidates. The role of interferon- $\alpha$  (IFN- $\alpha$ ) in BD pathogenesis was suggested in a previous study and IFN- $\alpha$  uses the receptor dimer of *IFNAR1* and *IFNAR2*.<sup>40</sup>

Among common SNVs, rs396991 in *FCGR3A* and rs5498 in *ICAM1* also have reached the significance level in this study, and both of these variants have been previously reported to be associated with BD in several studies.<sup>23–26</sup>

In this study, we successfully replicated previous studies and identified several novel susceptibility variants of target genes using the NGS technology. However, there are some limitations to our study such as the insufficient sequencing coverage, a small sample size of cases and inclusion of only one ethnicity; thus, our findings should be interpreted with caution. In addition, BD is a multifactorial disease caused by several genetic and environmental factors. Even though several reports have shown the successful application of NGS methods to discovery of genetic determinants of complex traits, further studies with larger sample sizes and inclusion of different ethnicities are needed to confirm the results of this study.<sup>15,16</sup>

In conclusion, target enrichment and massively parallel sequencing technologies have provided valuable information on the genetic predisposition for BD; in particular, by focusing on pathogenic candidates that alter the corresponding transcripts. In addition to the genes or variants that replicated previous studies, newly detected genetic variants, such as rs199955684 in *KIR3DL3*, should be investigated further with a larger sample size from diverse populations.

## ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (no. 2011-0030738).

- Deuter CM, Kotter I, Wallace GR, Murray PI, Stubiger N, Zierhut M. Behçet's disease: ocular effects and treatment. *Prog Retin Eye Res* 2008; **27**: 111–136.
- Direskeneli H. Behçet's disease: infectious aetiology, new autoantigens, and HLA-B51. *Ann Rheum Dis* 2001; **60**: 996–1002.
- Ohno S, Ohguchi M, Hirose S, Matsuda H, Wakisaka A, Aizawa M. Close association of HLA-Bw51 with Behçet's disease. *Arch Ophthalmol* 1982; **100**: 1455–1458.
- Piga M, Mathieu A. Genetic susceptibility to Behçet's disease: role of genes belonging to the MHC region. *Rheumatology (Oxford)* 2011; **50**: 299–310.
- Verity DH, Marr JE, Ohno S, Wallace GR, Stanford MR. Behçet's disease, the Silk Road and HLA-B51: historical and geographical perspectives. *Tissue Antigens* 1999; **54**: 213–220.
- Durrani K, Papaliodis GN. The genetics of Adamantiades-Behçet's disease. *Semin Ophthalmol* 2008; **23**: 73–79.
- Horie Y, Meguro A, Ota M, Kitaichi N, Katsuyama Y, Takemoto Y *et al*. Association of TLR4 polymorphisms with Behçet's disease in a Korean population. *Rheumatology (Oxford)* 2009; **48**: 638–642.
- Meguro A, Inoko H, Ota M, Katsuyama Y, Oka A, Okada E *et al*. Genetics of Behçet disease inside and outside the MHC. *Ann Rheum Dis* 2010; **69**: 747–754.
- Mizuki N, Meguro A, Ota M, Ohno S, Shiota T, Kawagoe T *et al*. Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet* 2010; **42**: 703–706.
- Remmers EF, Cosan F, Kirino Y, Ombrello MJ, Abaci N, Satorius C *et al*. Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nat Genet* 2010; **42**: 698–702.
- Bilguvar K, Ozturk AK, Louvi A, Kwan KY, Choi M, Tatli B *et al*. Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. *Nature* 2010; **467**: 207–210.
- Chou LS, Liu CS, Boese B, Zhang X, Mao R. DNA sequence capture and enrichment by microarray followed by next-generation sequencing for targeted resequencing: neurofibromatosis type 1 gene as a model. *Clin Chem* 2010; **56**: 62–72.
- Shearer AE, DeLuca AP, Hildebrand MS, Taylor KR, Gurrola J 2nd, Scherer S *et al*. Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. *Proc Natl Acad Sci USA* 2010; **107**: 21104–21109.
- Lim BC, Lee S, Shin JY, Kim JI, Hwang H, Kim KJ *et al*. Genetic diagnosis of Duchenne and Becker muscular dystrophy using next-generation sequencing technology: comprehensive mutational search in a single platform. *J Med Genet* 2011; **48**: 731–736.
- Johansen CT, Wang J, McIntyre AD, Martins RA, Ban MR, Lanktree MB *et al*. Excess of rare variants in non-genome-wide association study candidate genes in patients with hypertriglyceridemia. *Circ Cardiovasc Genet* 2012; **5**: 66–72.
- Lee S, Paik SH, Kim HJ, Ryu HH, Cha S, Jo SJ *et al*. Exomic sequencing of immune-related genes reveals novel candidate variants associated with alopecia universalis. *PLoS One* 2013; **8**: e53613.
- Suzuki Kurokawa M, Suzuki N. Behçet's disease. *Clin Exp Med* 2004; **4**: 10–20.
- Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 2010; **26**: 873–881.
- Kim JI, Ju YS, Park H, Kim S, Lee S, Yi JH *et al*. A highly annotated whole-genome sequence of a Korean individual. *Nature* 2009; **460**: 1011–1015.
- Ju YS, Kim JI, Kim S, Hong D, Park H, Shin JY *et al*. Extensive genomic and transcriptional diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean individuals. *Nat Genet* 2011; **43**: 745–752.
- Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009; **4**: 1073–1081.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P *et al*. A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248–249.
- Aksu K, Kitapcioglu G, Keser G, Berdeli A, Karabulut G, Kobak S *et al*. FcγR1a, IIIa and IIIb gene polymorphisms in Behçet's disease: do they have any clinical implications? *Clin Exp Rheumatol* 2008; **26** (4 Suppl 50), S77–S83.
- Ben Dhifallah I, Karray EF, Sassi F, Hamzaoui K. Intercellular adhesion molecule 1 K469E gene polymorphism is associated with presence of skin lesions in Tunisian Behçet's disease patients. *Tissue Antigens* 2010; **75**: 74–78.
- Kim EH, Mok JW, Bang D, Lee ES, Lee S, Park KS. The ICAM1469\*E is associated with susceptibility to ocular lesions and vasculitis in Korean patients with Behçet's disease. *Adv Exp Med Biol* 2003; **528**: 235–236.
- Saadat M. Association between K469E polymorphism of intracellular adhesion molecules-1 ICAM-1 and Behçet's disease. *Saudi Med J* 2006; **27**: 1934.
- Trundle AE, Hiby SE, Chang C, Sharkey AM, Santourlidis S, Uhrberg M *et al*. Molecular characterization of KIR3DL3. *Immunogenetics* 2006; **57**: 904–916.
- Cooper MA, Fehniger TA, Turner SC, Chen KS, Ghaheri BA, Ghayur T *et al*. Human natural killer cells: a unique innate immunoregulatory role for the CD56(bright) subset. *Blood* 2001; **97**: 3146–3151.
- Suzuki Y, Hoshi K, Matsuda T, Mizushima Y. Increased peripheral blood gamma delta+ T cells and natural killer cells in Behçet's disease. *J Rheumatol* 1992; **19**: 588–592.
- Yu HG, Lee DS, Seo JM, Ahn JK, Yu YS, Lee WJ *et al*. The number of CD8+ T cells and NKT cells increases in the aqueous humor of patients with Behçet's uveitis. *Clin Exp Immunol* 2004; **137**: 437–443.
- Ahn JK, Chung H, Lee DS, Yu YS, Yu HG. CD8brightCD56+ T cells are cytotoxic effectors in patients with active Behçet's uveitis. *J Immunol* 2005; **175**: 6133–6142.
- Yamaguchi Y, Takahashi H, Satoh T, Okazaki Y, Mizuki N, Takahashi K *et al*. Natural killer cells control a T-helper 1 response in patients with Behçet's disease. *Arthritis Res Ther* 2010; **12**: R80.
- Ueland PM, Hustad S, Schneede J, Refsum H, Vollset SE. Biological and clinical implications of the MTHFR C677T polymorphism. *Trends Pharmacol Sci* 2001; **22**: 195–201.
- Cahill M, Karabatzaki M, Meleady R, Refsum H, Ueland P, Shields D *et al*. Raised plasma homocysteine as a risk factor for retinal vascular occlusive disease. *Br J Ophthalmol* 2000; **84**: 154–157.

- 35 Er H, Evereklioglu C, Cumurcu T, Turkoz Y, Ozerol E, Sahin K *et al*. Serum homocysteine level is increased and correlated with endothelin-1 and nitric oxide in Behçet's disease. *Br J Ophthalmol* 2002; **86**: 653–657.
- 36 Ozkul Y, Evereklioglu C, Borlu M, Taheri S, Calis M, Dundar M *et al*. 5,10-Methylenetetrahydrofolate reductase C677T gene polymorphism in Behçet's patients with or without ocular involvement. *Br J Ophthalmol* 2005; **89**: 1634–1637.
- 37 Mizuki N, Ota M, Kimura M, Ohno S, Ando H, Katsuyama Y *et al*. 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* 1997; **94**: 1298–1303.
- 38 Yabuki K, Mizuki N, Ota M, Katsuyama Y, Palimeris G, Stavropoulos C *et al*. Association of MICA gene and HLA-B\*5101 with Behçet's disease in Greece. *Invest Ophthalmol Vis Sci* 1999; **40**: 1921–1926.
- 39 Park SH, Park KS, Seo YI, Min DJ, Kim WU, Kim TG *et al*. Association of MICA polymorphism with HLA-B51 and disease severity in Korean patients with Behçet's disease. *J Korean Med Sci* 2002; **17**: 366–370.
- 40 Liu X, Yang P, Wang C, Li F, Kijlstra A. IFN-alpha blocks IL-17 production by peripheral blood mononuclear cells in Behçet's disease. *Rheumatology (Oxford)* 2011; **50**: 293–298.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Supplementary Information accompanies the paper on Experimental & Molecular Medicine website (<http://www.nature.com/emm>)