Microarray and whole-exome sequencing analysis of familial Behçet’s disease patients

Daisuke Okuzaki¹,²,*, Kazuyuki Yoshizaki³,⁴,*, Toshio Tanaka⁴, Toru Hirano⁶, Kohshiro Fukushima², Takanori Washio⁶ & Hiroshi Nojima¹,²

Behçet’s disease (BD), a chronic systemic inflammatory disorder, is characterized by recurrent oral and genital mucous ulcers, uveitis, and skin lesions. We performed DNA microarray analysis of peripheral blood mononuclear cell (PBMC) mRNA from 41 Japanese BD patients and revealed elevated levels of interleukin (IL) 23 receptor (IL23R) mRNA in many BD patients. DNA sequencing around a SNV (Rs12119179) tightly linked to BD revealed an elevated frequency of the C genotype, consistent with a previous report that IL23R is a susceptibility locus for BD. Notably, four of these BD patients are members of familial BD; a whole-exome sequencing (WES) of these BD patients identified 19 novel single-nucleotide variations (SNVs) specific to these patients. They include heterozygous SNVs in the genes encoding IL-1 receptor-associated kinase 4 (IRAK4), nucleotide-binding oligomerization domain (NOD)-like receptor family pyrin domain-containing 14 (NRP14) and melanoma antigen-encoding gene E2 (MAGEE2); IRAK4 harbors a missense mutation, whereas NRP14 and MAGEE2 harbor nonsense mutations. These SNVs may serve as genetic markers that characterize BD.

Behçet’s disease (BD) is a chronic systemic inflammatory disorder characterized by four major manifestations: recurrent oral and genital mucous ulcers, uveitis, and skin lesions. Because BD is more prevalent in certain geographical regions, with the highest incidence in countries along the ancient silk route spanning from Mediterranean countries to the Middle East and Japan, putative genetic variants and specific environmental factors are considered to be important for its etiology. Across multiple ethnicities, many BD patients have the HLA-B*51 variant of the human leukocyte antigen (HLA) class I allele. Because many healthy people also have this HLA-B51 variant, other genetic variations and/or environmental factors must be required to fully explain the etiology of BD.

Genome-wide association studies (GWAS) carried out in Turkey and Japan demonstrated that HLA-B*51, HLA-A*26, IL10, and IL23R–IL12RB2 are susceptibility loci for BD, and also revealed an interaction between HLA-B*51 and ERAP1. GWAS in a Chinese cohort confirmed the association between IL10 polymorphisms and BD. In addition, single-nucleotide polymorphism (SNP) mapping of MHC with logistic regression revealed that the HLA-B/MICA region and the region between HLA-F and HLA-A are independently associated with BD. A GWAS in a Korean population identified a novel BD-associated locus encompassing the gene encoding GTPases of immunity-associated protein (GIMAP), which plays a role in peripheral T-cell function. However, the GIMAP association did not replicate in European BD patients. High copy-number variations (CNVs) of complement component C4A confer a risk for BD, and serum C4 protein level is significantly elevated in Chinese BD patients. Non-synonymous variants identified by deep exonic resequencing confirmed the association of...
IL23R and TLR4 with BD, suggesting the involvement of innate immune and bacterial sensing mechanisms in BD pathogenesis.12

In this study, we sought to identify genes whose mRNA levels were commonly up- or down-regulated in peripheral blood mononuclear cells (PBMCs) of BD patients, because such genes may be directly linked to the pathogenesis of BD. We identified several genes, including IL23R, that were commonly up-regulated in all BD patients examined. Moreover, we identified a rare case of familial BD over three generations, and performed genome-wide exome analysis to identify genes specific to these patients. Based on our results, we propose a novel model for the pathogenesis of BD.

Results

DNA microarray analysis of PBMCs from BD patients. To determine whether gene expression profiles of many BD patients share common abnormalities relative to those of healthy volunteers (HVs), we performed genome-wide complementary DNA microarray analyses using an Agilent Hu44K array. RNA samples were isolated from PBMCs of 41 individual BD patients (32 females and nine males) and 17 HVs (Supplementary Fig. S1). Because we wanted to obtain putative gene markers that distinguish BD from healthy controls even during its inactive phase, BD patients were examined regardless of symptoms or disease activity/inactivity. When we arranged the top 50 up-regulated genes according to decreasing fold-change values (Fig. 1a) or the bottom 50 down-regulated genes according to increasing fold-change values (Fig. 1b), we found several genes were up-regulated (>3.0-fold, including PCDH18, BBS3, and PP2A), or down-regulated (<20-fold, including RPS4Y1, DDX3Y, and RPS4Y2) in many BD patients.

However, several down-regulated genes tended to be up-regulated when only male patients were considered (green arrows in Fig. 1b), suggesting that some up- or down-regulated genes in the microarray data were gender-specific, which could have hampered the comprehensive analysis of the data. Moreover, the number of female patients (32) was much larger than that of male patients (nine). To prevent other important genes from being overlooked, we also arranged the data according to decreasing or increasing fold change values for male patients (Supplementary Fig. S2) and female patients (Supplementary Fig. S3). These analyses identified other markedly up-regulated genes, including BAAT, CYSLTR1, IL23R, NMU and GUCYTF (Supplementary Fig. S2a, S3a), or down-regulated genes, including IFI27, OLFM4, NS4A2 and PIGC (Supplementary Fig. S2b, S3b). IL23R, in particular, is important because SNPs in the intergenic region between IL23R and IL12RB2 is associated with BD.13-15 The putative functions of these genes in the pathogenesis of BD are discussed in Supplementary Results. It is remarkable that many of the identified genes are unknown; this suggests that much remains to be done to determine the genetic basis of BD.

DNA sequencing at the IL23R–IL12RB2 locus for familial BD patients. Notably, IL23R mRNA was up-regulated in 39 of 41 BD patients (Fig. S2a). Previous GWAS identified associations of BD with SNPs in the intergenic region between IL23R and IL12RB2.12,13,15 IL23R is also a susceptibility locus for a number of inflammatory and immune-linked diseases, including inflammatory bowel disease,13 psoriasis,14 psoriatic arthritis,15 ankylosing spondylitis,16 acute anterior uveitis,17 Vogt–Koyanagi–Harada syndrome,18 and idiopathic achalasia.19 These findings suggest that up-regulation of IL23R is involved in the pathogenesis of BD.

To determine whether up-regulation of IL23R mRNA levels was due to the SNV at a susceptibility locus for BD,4,5 we determined the DNA sequence around the variant nucleotide (Rs12119179). BD patients had a variety of allele combinations (Supplementary Fig. S4a), including the A/A, A/C, or C/C genotypes (Fig. 2a). All familial BD patients had the C/C genotype, whereas HB1 and SB2 had the A/C genotype in HB1 and SB2 (Fig. 2b). IL23R mRNA levels (Fig. S2a) were almost identical between the A/A, A/C, and C/C genotypes (Fig. 2c). The C/C genotype was more frequent in the BD patients (red bars in Fig. 2d) than in the healthy Japanese population (black bars in Fig. 2d) (http://www.ncbi.nlm.nih.gov/SNP/snp_viewTable.cgi?pop=1411). Notably, mRNA levels of IL12RB2 and IL10 were unaltered (Supplementary Fig. S5). These results support a previous report that this SNV (Rs12119179) is related to BD4,5, but also indicates that it is not sufficient to cause BD symptoms.

WES analysis of familial BD patients. Interestingly, two daughters (BD20 and BD47) and a granddaughter (BD50) of patient BD26 also suffered from BD; the husband of BD47 (HB1) was healthy, and their two sons (SB1 and SB2) were too young to diagnose for BD at the time of the study (Fig. 3a). Because cases of familial BD are rare, we performed WES analysis to determine whether these patients possess any common SNVs in protein-coding regions. On average, we generated 6.9Gb of sequence (68,383,039 reads) per sample to a mean depth of 134×. Approximately 99.3% of reads passed quality control and were used for mapping. Ultimately, an average of 5.46 Gb of high-quality sequence (56,741,144 reads) per sample were used for SNV/indel calling, yielding a mean gross overall coverage of 106× of the target. Among these sequences, an average of 3.58 Gb of sequence mapped to the exome target. The mean capture efficiency varied across the targets: 95.5% of the target had ≥4× coverage, 90.5% had ≥10× coverage, and 82.6% had ≥20× coverage. We identified 224,391, 234,927, 233,108, and 216,409 high-quality SNVs in each of the four samples, of which 11,405 passed the filter criteria. In parallel, we predicted 15,693, 15,102, 16,280, and 16,188 small indels in these samples, of which 431 passed the filter criteria.

Because known SNVs are present in the healthy population, we searched for novel (unknown) SNVs relative to the whole human genome sequence in the data bank. In one of the alleles of the familial BD patients (BD20, BD26, BD47, and BD50), we identified 26 putative novel SNVs, including one nonsense SNV in NLRP14 and twenty missense SNVs (Fig. 3b,c); no novel SNVs occurred in both alleles. Sanger DNA sequencing of these chromosomal regions confirmed that these heterozygous SNVs were indeed present in these BD patients, but not in HB1 (yellow vertical arrow). Some of these SNVs were either present or absent in SB1 and SB2; their physiological significance remains elusive due to the lack of clinical data at this stage.
Although NLRP14 expression has been reported to be testis-specific\textsuperscript{20}, NLRP14 mRNA is also detected in other tissues (http://www.genecards.org/cgi-bin/carddisp.pl?gene=NLRP14). Moreover, the family of cytosolic tripartite NLR receptors to which NLRP14 belongs plays a key role in innate immunity. The truncated NLRP14 protein encoded by the variant allele in the familial BD patients includes only a portion of the N-terminal PYD domain (Fig. 3b), which was revealed to be important in a crystallographic study\textsuperscript{21}. Thus, NLRP14 could be involved in the pathogenesis of BD in these patients, provided that this truncated protein is not easily degraded in living cells. In addition, a missense mutation of IL-1 receptor-associated kinase 4 (IRAK4) (Fig. 3b) was shared by these BD patients. IRAK4 plays a pivotal role in signaling in the innate immune system\textsuperscript{22}; therefore, this variant may also be implicated in the pathogenesis of BD. We will examine this possibility in future work.

To determine whether other BD patients also possess these SNVs, we selected five other BD patients and subjected their genomic DNA to Sanger sequencing. Most of these SNVs were not present in other BD patients,
suggesting that they were specific to members of this family and not related to BD pathogenesis (Fig. 3c). One exception was CDK5RAP2 (pink font in Fig. 3c): BD11 had an SNV (C/A) in one of her CDK5RAP2 copies. However, the other BD patients did not possess this SNV (Supplementary Fig. S4b); thus BD11 is an exceptional case.

MAGEE2 SNVs in BD patients. Another notable gene was MAGEE2, which was classified as a novel SNV when we obtained our WES data; we performed the SNV analysis of MAGEE2 because it harbored a nonsense mutation in all BD patients but not in HB1 (Fig. 4a). Although updated genomic data show that MAGEE2 is no more a novel SNV, its variety among BD patients is notable. Indeed, Sanger DNA sequencing revealed that five other BD patients, including BD1, had the A/A genotype (Fig. 4b–i), whereas HB1 had the C/C genotype (Fig. 4b–iii). We examined the rest of the BD patients (Fig. S5B) and found that 89.1% of BD patients had the A/A genotype (Fig. 4b–i), whereas five patients (BD16, BD18, BD22, BD27, and BD38) had the A/C genotype (Supplementary Fig. S4c). The frequency of the A/C genotype in BD patients (10.9%) was higher than that in the Japanese population overall (2.3%) (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=rs1343879). Notably, the nonsense SNV (A genotype) encodes a truncated protein, whereas the C genotype encodes the

Figure 2. DNA sequences at the IL23R–IL12RB2 locus of our BD patients. (a) DNA sequence ladder around the SNV (Rs12119179) for BD3, BD2, and BD5. Arrows indicate signal peaks for A/A, A/C, and C/C genotypes, respectively. (b) Genotypes of the Rs12119179 SNV for the BD family members. (c) Comparison of IL23R mRNA levels (see Fig. 1A) for BD patients with A/A, A/C, and C/C genotypes. (d) Frequencies of A/A, A/C, and C/C genotypes of SNV (Rs12119179) in our BD patients (red bars) or the healthy Japanese population (black bars). Frequencies of A and C populations calculated from A/A, A/C, and C/C genotypes are shown.
full-size protein (Supplementary Fig. S4d). These results suggest that the \textit{MAGEE2} SNV is involved in the pathogenesis of BD.

**Discussion**

Here, we report the results of DNA microarray analyses of PBMCs from 41 Japanese BD patients. On the basis of our findings, we propose that up-regulation of \textit{PCDH18}, \textit{BAAT}, \textit{CysLTR1}, and \textit{IL23R} in PBMCs may serve as diagnostic and etiologic markers for BD (Fig. 1, S2, S3). We also performed whole-exome analysis of four familial BD patients (Fig. 2a), revealing 26 putative novel SNVs in one of the chromosomes of these patients. Notably, 19 of the novel SNVs were not present in the healthy subject HB1 (Fig. 2b,c), suggesting that these SNVs were related to BD pathogenesis in these patients. An SNV in \textit{MAGEE2} (Fig. 4), which was present in all BD patients but not in...
HB1, may generate a truncated protein; likewise, the nonsense SNV in \textit{NLRP14} encodes a truncated form of the protein (Fig. 2b) that contains only a portion of an important N-terminal domain\textsuperscript{21}. Notably, the related protein NLRP3 regulates innate immunity as a component of the inflammasome, which is responsible for inflammatory processes through activation of caspase-1 and maturation of the inflammatory cytokines pro-interleukin-1\textbeta (pro-IL-1\textbeta) and pro-IL-18\textsuperscript{23}. Indeed, a recombinant, non-glycosylated form of the human interleukin-1 receptor antagonist (IL-1Ra), Anakinra, is therapeutically effective in some BD patients\textsuperscript{24}. Because IRAK4 also controls innate immunity\textsuperscript{25}, the mutations found in both NLRP14 and IRAK4 suggests that at least a subset of BD is an auto-inflammatory syndrome caused by abnormalities in innate immunity. It remains elusive if unrelated BD patients have different mutations in the same gene; this will be examined in our future work.

Microarray analysis revealed up-regulated \textit{IL23R} mRNA levels in BD patients (Fig. S2a). Because \textit{IL23R} is in a susceptibility locus for BD\textsuperscript{4,5}, we determined the DNA sequence around the SNV (Rs12119179) at this locus. The frequency of the C/C genotype was higher in BD patients than in healthy Japanese populations (Fig. 3). Although this result supports the relevance of this SNV to BD, it also demonstrates that the presence of this variant alone is not sufficient to cause BD symptoms. Neutrophil invasion into sites of localized inflammation (e.g., erythema nodosum, folliculitis, and genital ulcer) in the absence of infection is a notorious feature of BD. The elevation of \textit{IL23R} mRNA in BD patients suggests that enhanced Th17 activation resulting from IL23 stimulation may be responsible for the observed neutrophil activation. Based on these considerations and the results presented in this study, we propose a new model for the pathogenesis of BD involving \textit{IL23R}, \textit{NLRP14}, \textit{IRAK4} and \textit{MAGEE2}.

\textbf{Figure 4.} \textit{MAGEE2} genotype and mRNA levels in BD patients. (a) \textit{MAGEE2} on chromosome X contained a known SNV (A/C) in BD patients (+) but not in HB1 (−). (b) DNA sequence ladders around the \textit{MAGEE2} SNV region for BD1, BD16, and HB1. Arrows indicate signal peaks for the A/A, A/C, and C/C genotype, respectively. (c) Percentage (%) of A/A, A/C, and C/C genotypes of \textit{MAGEE2} in our BD patients (red bars) and the healthy Japanese population (black bars). Percentage of A and C populations calculated from A/A, A/C, and C/C genotypes are shown. (d) A and C genotypes generate truncated (green bar) or full-size (blue bar) \textit{MAGEE2} protein, respectively.
Methods

Human subjects and ethical considerations. Forty-three BD patients were enrolled at Osaka University Hospital (BD1–41 plus the mother and daughter of BD20) between 2001 and 2009; the genders and ages of BD patients are provided in Figure S1. The study was reviewed and approved by the Research Ethics Committee of Osaka University, and written informed consent was obtained from all participants. The experimental methods involving patients were carried out in accordance with the approved guidelines. The diagnosis of BD was established according to standard criteria proposed by the Japan BD Research Committee. Serum samples were obtained from patients regardless of their symptoms and level of disease activity/inactivity.

Target selection and sequencing. Exome sequencing was conducted on four DNA samples from patients with familial BD: patient BD47 and her aunt (BD20), grandmother (BD26), and mother (BD50). Genomic DNA was extracted from PBMC using the PAXgene Blood DNA Kit (QIAGEN), sheared into 150–200 bp fragments, and used to make a library for multiplexed paired-end sequencing (Illumina). The resultant library was hybridized to biotinylated cRNA oligonucleotide baits from the SureSelect Human All Exon 50 Mb kit (Agilent Technologies) for exome capture. Targeted sequences were purified using magnetic beads, amplified, and sequenced on an Illumina HiSeq2000 platform in paired-end 101 bp configuration. The raw sequence data were submitted to the NCBI SRA database under accession No. SRP059981 (NCBI BioProject PRJNA288379).

References
1. Hatemi, G., Seyahi, E., Fresko, I., Talarico, R. & Hamuryudan, V. Behçet's syndrome: a critical digest of the 2013–2014 literature. Clin. Exp. Rheumatol. 32, S112–122 (2014).
2. Gül, A. Genetics of Behçet's disease: lessons learned from genomewide association studies. Curr. Opin. Rheumatol. 26, 56–63 (2014).
3. Ombrello et al. Behçet disease-associated MHC class I residues implicate antigen binding and regulation of cell-mediated cytotoxicity. Proc. Natl. Acad. Sci. USA 111, 8867–8872 (2014).
4. Remmers et al. Genome-wide association study identifies variants in the MHC class I, II, and IL23R-IL12RB2 regions associated with Behçet's disease. Nat. Genet. 42, 698–702 (2010).
5. Mizuki et al. Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. Nat. Genet. 42, 703–706 (2010).
6. Kirino et al. Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. Nat. Genet. 45, 202–207 (2013).
7. Kuznarov et al. Behçet's disease in HLA-B*51 negative Germans and Turks shows association with HLA-Bw4-80L. Arthritis. Res. Ther. 16, R16 (2014).
8. Wu et al. IL10 polymorphisms associated with Behçet's disease in Chinese Han. Hum. Immunol. 75, 271–276 (2014).
9. Lee et al. Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behçet's disease. Ann. Rheum. Dis. 72, 1510–1516 (2013).
10. Ortiz-Fernández et al. GIMAP and Behçet disease: no association in the European population. Ann. Rheum. Dis. 73, 1433–1434 (2014).
11. Hou et al. Copy number variations of complement component C4 are associated with Behçet's disease but not with ankylosing spondylitis associated with acute anterior uveitis. Arthritis. Rheum. 65, 2963–2970 (2013).
12. Kirino et al. Targeted resequencing implicates the familial Mediterranean fever gene MEFV and the toll-like receptor 4 gene TLR4 in Behçet disease. Proc. Natl. Acad. Sci. USA 110, 8134–8139 (2013).
13. Duerr et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science. 314, 1461–1463 (2006).
14. Cargill et al. A large-scale genetic association study confirms IL1B and leads to the identification of IL23R as psoriasis-risk genes. Am. J. Hum. Genet. 80, 273–290 (2007).
15. Liu et al. A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. PLoS Genet. 4, e1000041; doi:10.1371/journal.pgen.1000041 (2008).
16. Hirschfeld et al. Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. N. Engl. J. Med. 360, 2544–2555 (2009).
17. Robinson et al. Genetic dissection of acute anterior uveitis reveals similarities and differences in associations observed with ankylosing spondylitis. Arthritis. Rheumatol. 67, 140–51 (2015).
18. Hou et al. Genome-wide association analysis of Vogt-Koyanagi-Harada syndrome identifies two new susceptibility loci at 1p31.2 and 10q21.3. Eur. Respir. J. 46, 1107–1111 (2015).
19. Latino et al. Impact of genetic polymorphisms on the pathogenesis of idiopathic achalasia: Association with IL33 gene variant. Hum. Immunol. 75, 364–369 (2014).
20. Westerveld et al. Mutations in the testis-specific NALP14 gene in men suffering from spermatogenic failure. Hum. Reprod. 21, 3178–3184 (2006).
21. Ebl, C., Hessenberger, M., Wenger, J. & Brandstetter, H. Structures of the NLRP14 pyrin domain reveal a conformational switch mechanism regulating its molecular interactions. Acta Crystallogr. D. Biol. Crystallogr. 70, 2007–2018 (2014).
22. Suzuki, N., Suzuki, S. & Yeh, W.C. IRAK-4 as the central TIR signaling mediator in innate immunity. Trends Immunol. 23, 503–506 (2002).
23. Elliott, E. I. & Sutterwala, F. S. Initiation and perpetuation of NLRP3 inflammasome activation and assembly. Immunol. Rev. 265, 35–52 (2015).
24. Cantarini et al. Effectiveness and tuberculosis-related safety profile of interleukin-1 blocking agents in the management of Behçet's disease. Autoimmun. Rev. 14, 1–9 (2015).
25. Netea, M. G., Wijmenga, C. & O'Neill, L. A. Genetic variation in Toll-like receptors and disease susceptibility. Nat. Immunol. 13, 535–42 (2012).

Acknowledgements

We thank the patients and healthy volunteers who participated in this study. We also thank Ms. Kana Ooi, Ms. Wang Mian, and Mr. Katsuma Matsuura for technical assistance with cDNA microarray analysis, and Dr. Patrick Hughes and Ms. Nicola Edwards of Bioedit Ltd. for their critical reading of the manuscript. This work was primarily supported by a grant-in-aid from the Health Science Research grant from the Ministry of Health and Welfare of Japan. This work was also supported in part by grants-in-aid to HN from the Bio-Medical Cluster Project In Saito; Innovation Plaza Osaka of the Japan Science and Technology Agency (JST); Scientific Research on Priority Areas “Applied Genomics”; Scientific Research (S), Exploratory Research; and the Science and...
Technology Incubation Program in Advanced Regions from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (to HN).

Author Contributions
H.N. contributed to the conception and design of this study and drafted the manuscript. K.Y., T.T. and T.H. collected patients’ blood and performed relevant clinical research. D.O. performed DNA microarray analysis. T.W. performed WES. D.O., K.F. and H.N. performed PCR and DNA sequencing. All authors read and approved the final manuscript.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Okuzaki, D. et al. Microarray and whole-exome sequencing analysis of familial Behçet's disease patients. Sci. Rep. 6, 19456; doi: 10.1038/srep19456 (2016).