Dense genotyping of immune-related loci implicates host responses to microbial exposure in Behçet's disease susceptibility

Masaki Takeuchi1,2, Nobuhisa Mizuki3, Akira Meguro2, Michael J Ombrello3, Yohei Kirino4, Colleen Satorius1, Julie Le1, Mary Blake5, Burak Erer1, Tatsukata Kawagoe2, Duran Ustek6, Ilkınr Tugal-Tutkun7, Emire Seyahi8, Yılmaz Ozyazgan9, Inês Sousa10,11, Feraydoun Davatchi12, Vânia Francisco10,11, Farhad Shahram12, Bahar Sadeghi Abdollahi12, Abdolhadi Nadjii12, Niloofar Mojarad Shafiei12, Fahmida Ghaderibarmi12, Shigeaki Ohno13, Atsuhsa Ueda4, Yoshiaki Ishigatsubo4, Massimo Gadina5, Sofia A Oliveira10,11, Ahmet Gül14, Daniel L Kastner1,15 & Elaine F Remmers1,15

We analyzed 1,900 Turkish Behçet's disease cases and 1,779 controls genotyped by the Immunochip. The most significantly associated SNP was rs1050502, a tag SNP for HLA-B*51. In the Turkish discovery set, we identified three new risk loci, IL1A–IL1B, IRF8, and CEBPB–PTPN1, with genomewide significance (P < 5 × 10−8) by direct genotyping and ADO–EGR2 by imputation. We replicated the ADO–EGR2, IRF8, and CEBPB–PTPN1 loci by genotyping 969 Iranian cases and 826 controls. Imputed data in 608 Japanese cases and 737 controls further replicated ADO–EGR2 and IRF8, and meta-analysis additionally identified RIPK2 and LACCT. The disease-associated allele of rs4402765, the lead marker at IL1A–IL1B, was associated with both decreased IL-1α and increased IL-1β production. ABO non-secretor genotypes for two ancestry-specific FUT2 SNPs showed strong disease association (P = 5.89 × 10−15). Our findings extend the list of susceptibility genes shared with Crohn’s disease and lesy and implicate mucosal factors and the innate immune response to microbial exposure in Behçet’s disease susceptibility.

Behçet's disease is a systemic vasculitis that manifests with oral ulcers, uveitis, skin inflammation, genital ulcers, and inflammation in other organs.1–2. Behçet's disease is relatively common in modern-day countries located along the ancient Silk Route.3. The geographical distribution of Behçet's disease and lack of consistency with expected patterns of Mendelian inheritance, despite a high sibling risk ratio (λs = 11.4–52.5) (ref. 4), suggest that multiple genetic and environmental factors contribute to disease susceptibility. Although genetic studies have identified multiple susceptibility loci4–14, these genetic factors do not fully explain the apparent disease heritability.

To further clarify the genetic etiology of Behçet's disease, we genotyped 2,014 Turkish cases and 1,826 population controls using the Immunochip15. After quality control filtering, 130,647 autosomal markers genotyped in 1,900 cases and 1,799 controls were subjected to association tests. Association analysis showed the strongest disease association within the major histocompatibility complex (MHC) region (Fig. 1). Consistent with our previous findings using genome-wide association study (GWAS) genotype data16, imputation of the MHC region and classical human leukocyte antigens (HLAs) showed the strongest association for HLA-B*51 (P = 5.67 × 10−99; Table 1 and Supplementary Fig. 1) among all markers. Other HLA alleles also showed significant association with Behçet's disease (Supplementary Table 1), and regression analysis among MHC class I and II molecules confirmed independent disease-protective association for HLA-A*03 and disease susceptibility association for HLA-B*15 (Supplementary Table 2). Association analysis for SNPs showed the strongest disease association at rs105052 (P = 9.99 × 10−99; Table 1), a synonymous

1. Inflammatory Disease Section, National Human Genome Research Institute, US National Institutes of Health, Bethesda, Maryland, USA. 2. Department of Ophthalmology and Visual Science, Yokohama City University Graduate School of Medicine, Yokohama, Japan. 3. Translational Genetics and Genomics Unit, National Institute of Arthritis and Musculoskeletal and Skin Diseases, US National Institutes of Health, Bethesda, Maryland, USA. 4. Department of Stem Cell and Immune Regulation, Yokohama City University Graduate School of Medicine, Yokohama, Japan. 5. Translational Immunology Section, National Institute of Arthritis and Musculoskeletal and Skin Diseases, US National Institutes of Health, Bethesda, Maryland, USA. 6. Department of Genetics, Institute for Experimental Medicine, Istanbul University, Istanbul, Turkey. 7. Department of Ophthalmology, Istanbul University, Istanbul, Turkey. 8. Department of Rheumatology, Cerrahpaşa Faculty of Medicine, Istanbul University, Istanbul, Turkey. 9. Department of Ophthalmology, Cerrahpaşa Faculty of Medicine, Istanbul University, Istanbul, Turkey. 10. Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal. 11. Instituto Gulbenkian de Ciência, Oeiras, Portugal. 12. Rheumatology Research Center, Tehran University of Medical Sciences, Tehran, Iran. 13. Department of Ophthalmology, Hokkaido University Graduate School of Medicine, Sapporo, Japan. 14. Department of Internal Medicine, Division of Rheumatology, Istanbul University, Istanbul, Turkey. 15. These authors jointly directed this work. Correspondence should be addressed to D.L.K. (kastnerd@mail.nih.gov) or E.F.R. (elaine.remmers@nih.gov).

Received 12 September 2016; accepted 12 January 2017; published online 6 February 2017; doi:10.1038/ng.3786
variant for isoleucine at position 47 of the HLA-B molecule and a tag SNP for HLA-B*51 ($r^2 = 1$).

In a previous Immunochip study of Behçet’s disease, Hughes et al.\textsuperscript{17} reported that rs116799036, a SNP in the HLA-B–MICA intergenic region, was more strongly associated with disease than HLA-B*51 and also reported three additional independent disease susceptibility markers in the MHC region, rs12552170, rs114854070, and HLA-Cw*16:02. In our study, the association $P$ value for rs116799036 was 1.3 × 10$^{-5}$-fold higher (less significant) than the $P$ value for HLA-B*51, and regression analysis identified significant residual association of HLA-B*51 after conditioning on each of the other markers (Table 1). In contrast, the associations of the other markers were completely abrogated by conditioning on HLA-B*51 (Table 1 and Supplementary Fig. 1), and the associations of these markers were strongly correlated with their linkage disequilibrium (LD) with HLA-B*51 (Supplementary Fig. 2).

Outside of the MHC region, three new loci, IL1A–IL1B (rs3783550, $P$ value corrected for genomic inflation ($P_{GC} = 2.12 \times 10^{-8}$), IRF8 (rs11117433, $P_{GC} = 2.73 \times 10^{-8}$), and CEBPB–PTPN1 (rs913678, $P_{GC} = 1.96 \times 10^{-9}$), and three previously reported loci, IL10, CCR1, and IL12A, displayed genome-wide significant associations with Behçet’s disease based on Immunochip genotyping (Fig. 1, Table 2, and Supplementary Fig. 3). We also replicated the disease association of markers in other loci previously reported for Behçet’s disease (IL23R–IL12RB2, ERAP1, KLRC4, and FUT2) but did not find evidence supporting an association for TNFAIP3 reported in a Han Chinese population\textsuperscript{10} or IRKL–CNTN5 recently reported in a Spanish population\textsuperscript{14} (Supplementary Table 3).

Four significantly associated loci (IL1A–IL1B, IL12A, IRF8, and CEBPB–PTPN1) were not identified by our previous GWAS. Two of these loci (IL12A and IRF8) had no markers on the GWAS array in strong LD with the Immunochip lead SNP, and the associations were thus driven by the increased coverage on the Immunochip. The greater power provided by the larger sample size here also contributed to identifying disease associations at all four loci (Supplementary Table 4).

For 12 of the 20 new loci with association $P < 5 \times 10^{-5}$ (Table 2 and Supplementary Table 5), imputation identified a more significant association, including rs4402765 in the IL1A–IL1B locus ($P_{GC} = 3.85 \times 10^{-9}$) and a genome-wide significant association at rs7075773 in the ADO–EGR2 locus ($P_{GC} = 2.96 \times 10^{-8}$; Supplementary Fig. 3 and Supplementary Table 6). Conditional regression analysis of the regions that exceeded genome-wide significance identified an independent contribution to disease susceptibility at rs7203487 in the IRF8 locus after conditioning on the lead SNP, rs11117433 (Supplementary Table 7).

For replication, the 21 lead SNPs genotyped by the Immunochip in the 20 new loci with association $P < 5 \times 10^{-5}$ in the Turkish population were genotyped in the Iranian population comprising 969 cases and 826 controls. Four of these loci (ADO–EGR2, LACC1, IRF8, and CEBPB–PTPN1) replicated in the Iranian population (Supplementary Table 8). In a meta-analysis of the Turkish and Iranian populations, ADO–EGR2, IRF8, and CEBPB–PTPN1 exceeded genome-wide significance (Table 3).

We also evaluated the newly associated markers in the Japanese population, where possible, using imputed data from the previous Japanese GWAS in 608 cases and 737 controls, and we replicated association of rs9316059 in LACC1 (Supplementary Table 9), which also exceeded genome-wide significance in a meta-analysis with Turkish data (Supplementary Table 10). Furthermore, disease association of rs2121033 in the LACC1 locus, which is in strong LD with rs9316059 ($r^2 = 0.99$), was observed in all three populations, and a combined

### Table 1: Numeric association testing and conditional regression analysis for HLA-B*51, rs1050502, and four markers for which independent disease associations for Behçet’s disease were reported in a previous Immunochip study

| Marker       | A1/A2 | OR   | 95% CI          | $P$         | $P_{Conditional}$^a |
|--------------|-------|------|-----------------|-------------|---------------------|
| HLA-B*51     | +/−   | 3.26 | 2.89–3.68       | $5.67 \times 10^{-18}$ | –                   |
| rs1050502    | T/C   | 3.25 | 2.88–3.66       | $9.99 \times 10^{-19}$ | –                   |
| rs116799036  | AG    | 3.13 | 2.78–3.52       | $7.36 \times 10^{-17}$ | 0.29                |
| rs12552170   | AG    | 2.42 | 2.13–2.76       | $2.07 \times 10^{-15}$ | 0.036               |
| HLA-Cw*16:02 | +/−   | 2.93 | 2.33–3.68       | $6.90 \times 10^{-13}$ | 0.003               |
| rs114854070  | AG    | 1.32 | 1.18–1.47       | $5.07 \times 10^{-7}$  | 0.22                |

A1, minor allele; A2, major allele; OR odds ratio; CI, confidence interval.

*aThe residual $P$ value after conditioning on each marker determined using imputed data from 3,535 samples from Turkey with no missing genotype data for the six markers included in the regression analysis (1,840 cases, 1,735 controls). $P < 1 \times 10^{-5}$ was considered to indicate independent association in the conditional regression analysis. \textsuperscript{b}Present or absent. \textsuperscript{c}Reported in Hughes et al.\textsuperscript{17}.
**Table 2** Genome-wide significant associations with Behçet's disease of markers genotyped on the Immunochip in 1,900 cases and 1,779 controls from Turkey

| Marker | Nearest gene(s) | Chr. | hg19 position (bp) | A1/A2 | Allele freq. | OR    | 95% CI    | P      | P_{OC} |
|--------|----------------|------|-------------------|-------|--------------|-------|-----------|--------|--------|
|        |                |      |                   |       | Cases        | Controls |           |        |        |
| New loci |                |      |                   |       |              |         |           |        |        |
| rs3783550 | IL1A–IL1B     | 2    | 113,532,885       | G/T   | 0.360        | 0.298   | 1.33      | 1.20–1.46 | 1.29 × 10^{-8} | 2.12 × 10^{-8} |
| rs11117433 | IRF8          | 16   | 8,601,956         | C/G   | 0.074        | 0.113   | 0.63      | 0.54–0.74 | 1.67 × 10^{-7} | 2.73 × 10^{-7} |
| rs913678 | CEBPB–PTPN1   | 20   | 48,955,424        | C/T   | 0.474        | 0.404   | 1.33      | 1.21–1.46 | 1.10 × 10^{-9} | 1.96 × 10^{-9} |

Previously reported loci

| Marker | Chr. | hg19 position (bp) | A1/A2 | Allele freq. | OR    | 95% CI    | P      | P_{OC} |
|--------|------|-------------------|-------|--------------|-------|-----------|--------|--------|
| rs1518110 | IL10   | 1    | 206,944,861       | A/C   | 0.368        | 0.302   | 1.34      | 1.22–1.48 | 2.63 × 10^{-9} | 4.55 × 10^{-9} |
| rs7616215 | CCR1   | 3    | 46,205,686        | C/T   | 0.270        | 0.340   | 0.72      | 0.65–0.79 | 4.94 × 10^{-11} | 9.60 × 10^{-11} |
| rs17753641 | IL1A   | 3    | 159,647,674       | G/A   | 0.073        | 0.040   | 1.90      | 1.54–2.34 | 8.11 × 10^{-10} | 1.45 × 10^{-9} |

*Bold indicates genome-wide significance. Meta-analysis was performed for populations in which association for the variant exceeded the replication threshold (Online Methods). A1, minor allele in Turks; A2, major allele in Turks; OR, odds ratio; CI, confidence interval; P, inconsistency index; P_{het}, P for heterogeneity; TUR, Turkish; IRN, Iranian; JPN, Japanese.

**Table 3** Results of association tests and meta-analysis of new susceptibility loci identified in this study

| Marker (locus) | A1/A2 | Population | OR | 95% CI | P      | P_{OC} |
|----------------|-------|------------|----|--------|--------|--------|
| rs3783550 (IL1A–IL1B) | G/T | TUR | 1.33 | 1.20–1.46 | 1.29 × 10^{-8} |        |
|                |      | IRN | 1.13 | 0.98–1.31 | 0.098   |        |
|                |      | JPN | 1.11 | 0.93–1.33 | 0.24    |        |
| rs2230801 (RIPK2) | C/T | TUR | 1.43 | 1.22–1.68 | 9.60 × 10^{-6} |        |
|                |      | IRN | 1.11 | 0.84–1.46 | 0.47    |        |
|                |      | JPN | 3.41 | 1.80–6.47 | 6.39 × 10^{-5} |        |
| rs224127 (ADO–EGR2) | A/G | TUR | 1.26 | 1.15–1.39 | 1.56 × 10^{-6} |        |
|                |      | IRN | 1.18 | 1.03–1.35 | 0.017   |        |
|                |      | JPN | 1.30 | 1.11–1.51 | 0.0011  |        |
| rs1509966 (ADO–EGR2) | A/G | TUR | 0.80 | 0.73–0.87 | 1.47 × 10^{-6} |        |
|                |      | IRN | 0.79 | 0.69–0.90 | 5.09 × 10^{-4} |        |
|                |      | JPN | 0.91 | 0.77–1.07 | 0.24    |        |
| rs2121033 (LACC1) | G/C | TUR | 0.79 | 0.71–0.87 | 8.88 × 10^{-6} |        |
|                |      | IRN | 0.78 | 0.67–0.91 | 0.0012  |        |
|                |      | JPN | 0.69 | 0.58–0.83 | 4.68 × 10^{-5} |        |
| rs7203487 (IRF8) | C/T | TUR | 1.38 | 1.21–1.57 | 1.10 × 10^{-6} |        |
|                |      | IRN | 1.42 | 1.17–1.72 | 4.13 × 10^{-4} |        |
| rs142105922 (IRF8) | AAT/– | TUR | 0.63 | 0.52–0.77 | 5.58 × 10^{-6} |        |
|                |      | IRN | 0.68 | 0.51–0.91 | 0.0088  |        |
|                |      | JPN | 0.59 | 0.43–0.82 | 0.0013  |        |
| rs11117433 (IRF8) | C/G | TUR | 0.63 | 0.54–0.74 | 1.67 × 10^{-6} |        |
|                |      | IRN | 0.75 | 0.58–0.96 | 0.023   |        |
| rs913678 (CEBPB–PTPN1) | C/T | TUR | 1.33 | 1.21–1.46 | 1.10 × 10^{-9} |        |
|                |      | IRN | 1.29 | 1.13–1.48 | 1.59 × 10^{-4} |        |
| rs1041884 (IRF8) | C/T | TUR | 1.32 | 1.22–1.42 | 9.44 × 10^{-13} | 1.43 × 10^{-12} |

**Letters**

meta-analysis found highly significant association for this marker ($P_{OC} = 3.54 \times 10^{-11}$, Table 3).

With imputed data available from both the Turkish and Japanese populations, we were able to perform a more comprehensive analysis of the new suggestive regions, allowing identification of associations with alternate markers other than the lead ones identified in the Turkish discovery collection. Meta-analysis of all markers with suggestive evidence of association ($P < 5 \times 10^{-5}$) in the Turkish population that were also available in the Japanese data set identified four loci with genome-wide significant association, including the rs2121033 marker in the LACC1 locus described above. New markers in moderate LD with the Immunochip lead marker were identified in ADO–EGR2 and IRF8 (Table 3 and Supplementary Table 11), and conditional analysis on the Turkish genotypes suggested that these new markers...
were not independent of the lead SNPs (Supplementary Table 12). The combined Turkish and Japanese meta-analysis also identified a new genome-wide significant association at rs2230801 ($P_{GC} = 6.57 \times 10^{-9}$), a missense variant of RIPK2 (p.Ile259Thr) (Table 3 and Supplementary Table 11). In the Iranian collection, genotypes for the four markers that achieved genome-wide significance in the Turkish and Japanese meta-analysis demonstrated at least nominal evidence of association ($P < 0.05$), except for the rs2230801 variant in RIPK2 (Supplementary Table 13), for which the power to replicate was low ($0.32$). All the newly identified Behçet’s disease susceptibility markers are located in putative functional regions and have predicted functional effects, either altering protein structure or the expression of nearby genes (Supplementary Table 14).

Although replication of associations in the $IL1A$–$IL1B$ region did not reach statistical significance in the Iranian or Japanese samples, there was a trend for association, with a higher frequency of the risk allele in cases (Supplementary Tables 8 and 9). An expression quantitative trait locus (eQTL) study showed that the lead SNP associated with Behçet’s disease risk, rs4402765, was also the SNP most significantly associated with $IL1A$ gene expression in lymphoblastoid cells and showed that the disease risk allele was associated with reduced gene expression (Fig. 2a, b, Supplementary Fig. 3, and Supplementary Table 14). Consistent with the effect of this marker on $IL1A$ gene expression, we also found that the amount of IL-1α protein in culture supernatants from healthy donor peripheral blood mononuclear cells (PBMCs) stimulated with zymosan was reduced in homozygotes for the rs4402765 risk allele as compared to homozygotes for the protective allele (Fig. 2c). Although the published eQTL study did not show significant association of rs4402765 genotype with $IL1B$ gene expression, we found that the concentration of IL-1β protein was significantly higher in the cell culture supernatants of homozygotes for the risk allele than it was in supernatants from homozygotes for the protective allele (Fig. 2d).

IL-1α is highly expressed in the epidermis and has an important role in skin barrier functions against pathogens18. IL-1α is also required for effective host defense against disseminated candidiasis19. Our findings suggest that genetically encoded reductions in IL-1α expression may contribute to susceptibility to Behçet’s disease by weakening host response to and defense against invading pathogens. Disease susceptibility may also be increased by the risk allele’s effect on IL-1β production in response to microbial pathogens. IL-1β levels are elevated in individuals with Behçet’s disease20,21. Recently, blockade of IL-1 or IL-1β in patients with Behçet’s disease has been reported to be effective22,23. Our study raises the intriguing possibility that the decreased barrier function of IL-1α combines with the increased inflammatory response of IL-1β to increase Behçet’s disease risk.

$FUT2$ was recently reported to confer Behçet’s disease susceptibility in a meta-analysis of Iranian and Turkish GWAS data12. We have expanded this analysis to examine functionally relevant homoyzogous genotypes in a large sample size from three populations. $FUT2$ encodes α(1,2)-fucosyltransferase, which synthesizes secreted H antigen—the precursor of the ABO histo-blood group antigens—in body fluids and the intestinal mucosa24. The rs601338[A] allele (Turks and Iranians) and rs1047781[T] allele (Japanese) are ancestry-specific $FUT2$ non-secretor mutations (encoding p.Trp143Ter and p.Ile129Phe, respectively), for which homoyzogosity leads to an ABO non-secretor phenotype24. We found significant associations of rs601338 with disease in Turks ($P = 6.51 \times 10^{-10}$) and in Iranians ($P = 1.65 \times 10^{-5}$), and we also found significant association of rs1047781 with disease in Japanese ($P = 6.50 \times 10^{-4}$; Supplementary Table 15). These non-secretor genotypes are also associated with Crohn’s disease risk25,26 and with gut microbiome composition27,28. The non-secretor phenotype has also been associated with increased predisposition or resistance to different infectious agents29–31. Meta-analysis of the two common $FUT2$ non-secretor genotypes in Turks, Iranians, and Japanese yielded highly significant association ($P = 5.89 \times 10^{-15}$; Supplementary Table 15), providing evidence that ABO non-secretion, particularly at mucosal surfaces, increases the risk for Behçet’s disease and implicating the microbial–host interface in disease pathogenesis.

Our study has increased the number of susceptibility loci shared by Behçet’s disease and inflammatory bowel disease (IBD), which have many clinical features in common, to 11, adding 4 loci ($ADO$–$EGR2$, LACC1, IRF8, and CEBPB–PTPN1), thus indicating substantial genetic similarity between Behçet’s disease and IBD (Supplementary Tables 16 and 17) (refs. 25,32,33). In addition, markers within the reported susceptibility loci for IBD demonstrated greater enrichment for associations with Behçet’s disease than expected by chance (Supplementary Fig. 4). Comparing among subgroups of IBD, our new findings emphasize higher genetic similarity of Behçet’s disease with Crohn’s disease ($ADO$–$EGR2$, RIPK2, LACC1, and IRF8) than with ulcerative colitis ($ADO$–$EGR2$ and CEBPB–PTPN1) (Supplementary Table 17) (refs. 25,33).

We also identified substantial overlap in the susceptibility loci for Behçet’s disease and leprosy caused by infection with Mycobacterium leprae, adding three (RIPK2, ADO–EGR2, and LACC1) of the now four shared loci (Supplementary Table 17) (refs. 34–36). The reported
suspicious loci for leprosy also showed relative enrichment for associations with Behçet’s disease (Supplementary Fig. 4).

Interestingly, the minor allele of the LACCI lead SNP, rs2121033, confers protection for Behçet’s disease but is in high LD with a common coding variant, rs3764147 (p.Ile254Val; \( r^2 = 0.93 \)), that increases risk for IBD and Crohn’s disease (Supplementary Table 17). Furthermore, a rare mutation altering a residue in the laccase domain of the LACCI protein, p.Cys284Arg, cosegregates with Mendelian systemic juvenile idiopathic arthritis and Crohn’s disease in consanguineous families\(^{37,38}\). A recent study reported that the p.Ile254Val variant leads to impaired protein function, and Lact\(\alpha\)−/− mice produce decreased IL-1β in response to lipopolysaccharide (LPS) treatment, consistent with a role for IL-1β in Behçet’s disease pathogenesis\(^{39}\).

The minor allele of rs913678 (C) in the CEBPB−PTPN1 locus also showed opposite directions of effect in Behçet’s disease (risk) and IBD (protective) (Supplementary Table 17). This allele is associated with decreased gene expression (Supplementary Table 14), and Cebpb−/− mice show increased susceptibility to pathogens\(^{40,41}\). These opposite effects suggest that different mechanisms involving these loci increase disease risk for Behçet’s disease as compared with IBD.

A limitation of the Immunochip approach is that only selected genetic regions were explored in this study. In addition, the Immunochip has a potential problem in genotype calling accuracy because it is a custom array. To avoid miscalling, we applied strict protocols for quality control and cluster file preparation in the Turkish population (Online Methods). The peak genotyped markers for each disease-associated locus in the Turkish population showed robust clustering (Supplementary Fig. 5).

This Immunochip study with the largest discovery collection thus far and with two additional populations in the replication phase provides robust evidence that \( HLA-B^*51 \) in the \( HLA-B \) locus is the primary genetic source of disease risk and identifies multiple new susceptibility loci for Behçet’s disease. Genes in these loci contribute to the elucidation of disease pathogenesis by identifying disease-associated pathways, including pathways involved in host defense, inflammation, and immune response (Supplementary Table 18). Although a connection has not yet been proven, pathogenic infections have been proposed as an important environmental factor contributing to both the development and exacerbation of Behçet’s disease\(^{42}\). These pathways help to establish a link between genetic factors and environmental factors, such as microbial exposures, that together contribute to disease susceptibility. Our current findings implicate genetic determinants of mucosal barrier function and the host response to pathogens in Behçet’s disease susceptibility and draw important parallels and distinctions with respect to other immune-related diseases.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS
This research was supported by the Intramural Research Programs of the National Human Genome Research Institute and the National Institute of Arthritis and Musculoskeletal and Skin Diseases. We thank all the patients, the healthy controls, and medical staff for their enthusiastic support during this research study. M.T. is supported by a Fellowship for Japanese Biomedical and Behavioral Researchers at the NIH from the Japan Society for the Promotion of Science Research and a grant from the Japan Foundation for Applied Enzymology. Y.K. is supported by grants from the Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (grant 26713036), the Kanue Foundation for the Promotion of Medical Science, the Takeda Science Foundation, the SENSHIN Medical Research Foundation, and the Yokohama Foundation for Advancement of Medical Science. This research was also supported by the Portuguese Fundação para a Ciência e a Tecnologia (grant CMUP/ERJ/PEP/0028/2013, fellowship SFRH/BPD/70008/2010 to I.S., and an Investigator–FCT contract to S.A.O.) and the Research Committee of the Tehran University of Medical Sciences (grant 137/714). We thank A.E. Wilson for insightful comments on this manuscript.

AUTHOR CONTRIBUTIONS
M.T., M.O., A.G., D.L.K., and E.F.R. designed the study. M.T., A.M., M.J.O., M.B., M.G., A.D.L.K., and E.F.R. carried out the analysis. M.T., N.M., A.M., M.J.O., Y.K., C.S., J.L., M.B., B.E., T.K., D.U., I.T.-T., E.S., Y.O., I.S., F.D., V.F., F.S., B.S.A., A.N., N.M.S., F.G., S.O., A.U., Y.I., M.G., S.A.O., A.G., D.L.K., and E.F.R. procured samples and generated data. M.T., M.O., M.B., M.G., A.D.L.K., and E.F.R. wrote the manuscript. All authors read and approved the final version of the manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

URLs. GWAS catalog, http://www.genome.gov/gwastudies/; IMPUTE2, https://mathgen.stats.ox.ac.uk/impute/impute_v2_v2.html; SNP2HLA, https://www.broadinstitute.org/mpg/snp2hla/; CatTS, http://csq.sph.umich.edu/abecasis/CatTS/; META, https://mathgen.stats.ox.ac.uk/genetics_software/meta/meta.html; Genevar, https://www.sanger.ac.uk/resources/software/genevar/; Blood eQTL browser, http://genenetwork.nl/bloodeqtlbrowser/; haploReg, http://www.broadinstitute.org/mammals/haploreg/; PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/; GTEx, http://www.gtexportal.org/ImmunBase, https://www.immunobase.org/; DAVID, https://david.ncifcrf.gov/.

METHODS
Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

REFERENCES
1. Sakane, T., Takeino, M., Suzuki, N. & Inaba, G. Behçet’s disease. N. Engl. J. Med. 341, 1284–1291 (1999).
2. Hatemi, G. et al. EULAR recommendations for the management of Behçet disease. Ann. Rheum. Dis. 67, 1656–1662 (2008).
3. Verity, D.H., Marr, J.E., Ohno, S., Wallace, G.R. & Stanford, M.R. Behçet’s disease, the Silk Road and HLA-B51: historical and geographical perspectives. Tissue Antigens 54, 213–220 (1999).
4. Gull, A. & Ohno, S. HLA-B*51 and Behçet disease. Curr. Opin. Immunol. 20, 37–43 (2012).
5. Ono, S., Aoki, K., Sugiuira, S., Nakayama, E. & Itakura, K. HLA-A5 and Behçet disease. Lancet 2, 1383–1384 (1973).
6. de Menthon, M., Lavalley, M.P., Maldini, C., Guillemin, L. & Mahr, A. HLA-B51/BS and the risk of Behçet’s disease: a systematic review and meta-analysis of case-control genetic association studies. Arthritis Rheum. 61, 1287–1296 (2009).
7. Remmers, E.F. et al. Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R–IL12RB2 regions associated with Behçet disease. Nat. Genet. 42, 698–702 (2010).
8. Muñiz, N. et al. Genome-wide association studies identify IL23R–IL12RB2 and IL10 as Behçet disease susceptibility loci. Nat. Genet. 42, 703–706 (2010).
9. Kirino, Y. 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).
10. Li, H. et al. TNFAIP3 gene polymorphisms confer risk for Behçet’s disease in a Chinese Han population. Hum. Genet. 132, 293–300 (2013).
11. Kirino, Y. et al. Targeted resequencing implicates the familial Mediterranean fever gene MEVF and the toll-like receptor 4 gene TLR4 in Behçet disease. Proc. Natl. Acad. Sci. USA 110, 8134–8139 (2013).
12. Xavier, J.M. et al. FUT2: filling the gap between genes and environment in Behçet’s disease? Ann. Rheum. Dis. 74, 618–624 (2015).
13. Kpoden, J.H., Mart, J.E., Huitinga, J. et al. Genenome-wide association study in an admixed case series reveals IL12A as a new candidate in Behçet disease. PLoS One 10, e0119085 (2015).
14. Ortiz-Fernández, L. et al. Genetic analysis with the Immunochip platform in Behçet disease. Identification of residues associated in the HLA class I region and new susceptibility loci. PLoS One 11, e0161305 (2016).
15. Cortes, A. & Brown, M.A. Promise and pitfalls of the Immunochip. Arthritis Res. Ther. 13, 101 (2011).
16. Ombrello, M.J. 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).
17. Hughes, T. et al. Identification of multiple independent susceptibility loci in the HLA region in Behçet’s disease. Nat. Genet. 45, 319–324 (2013).
18. Barland, C.O. et al. Inflammation-induced interleukin-1α stimulation improves barrier homeostasis in aged murine epidermis. J. Invest. Dermatol. 122, 330–336 (2004).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.
19. Vonk, A.G. et al. Endogenous interleukin (IL)-1α and IL-1β are crucial for host defense against disseminated candidiasis. J. Infect. Dis. 193, 1419–1426 (2006).

20. Liang, L. et al. IL-1β triggered by peptidoglycan and lipopolysaccharide through TLR2/4 and RDS–NLRP3 inflammasome–dependent pathways is involved in ocular Behçet’s disease. Invest. Ophthalmol. Vis. Sci. 54, 402–414 (2013).

21. Kim, E.H., Park, M.J., Park, S. & Lee, E.S. Increased expression of the NLRP3 inflammasome components in patients with Behçet’s disease. J. Inflamm. (Lond.) 12, 41 (2015).

22. Case, F. et al. Biological treatments in Behçet’s disease: beyond anti-TNF therapy. Mediators Inflamm. 2014, 107421 (2014).

23. Gül, A. et al. Interleukin-1β-regulating antibody XOMA 052 (gevokizumab) in the treatment of acute exacerbations of resistant uveitis of Behçet’s disease: an open-label pilot study. Ann. Rheum. Dis. 71, 563–566 (2012).

24. Ferrer-Admetlla, A. et al. A natural history of FUT2 polymorphism in humans. Mol. Biol. Evol. 26, 1993–2003 (2009).

25. Franke, A. et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn’s disease susceptibility loci. Nat. Genet. 47, 267–271 (2015).

26. Liu, H. et al. Discovery of six new susceptibility loci and analysis of pleiotropic effects in leprosy. Nat. Genet. 47, 1831–1832 (2014).

27. Wacklin, P. et al. Secretor genotype (FUT2 gene) is strongly associated with the composition of Bifidobacteria in the human intestine. PLoS One 6, e20113 (2011).

28. Rausch, P. et al. Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 (Secretor) genotype. Proc. Natl. Acad. Sci. USA 108, 19030–19035 (2011).

29. Sheinfeld, J., Schaeffer, A.J., Cordon-Cardo, C., Rogatko, A. & Fair, W.R. Association of the Lewis blood-group phenotype with recurrent urinary tract infections in women. N. Engl. J. Med. 320, 773–777 (1989).

30. Marionneau, S., Ainaud, F., Bovin, N.V., Le Pendu, J. & Ruvoën-Clouet, N. Influence of the combined ABO, FUT2, and FUT3 polymorphism on susceptibility to Norwalk virus attachment. J. Infect. Dis. 192, 1071–1077 (2005).

31. Rupp, C. et al. FUT2 genotype is a risk factor for dominant stenosis and biliary candida infections in primary sclerosing cholangitis. Aliment. Pharmacol. Ther. 39, 873–882 (2014).

32. Hatemi, I. et al. Frequency of pathergy phenomenon and other features of Behçet’s syndrome among patients with inflammatory bowel disease. Clin. Exp. Rheumatol. 26 (Suppl. 50), S91–S95 (2008).

33. Jostins, L. et al. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491, 119–124 (2012).

34. Zhang, F.R. et al. Genomewide association study of leprosy. N. Engl. J. Med. 361, 2609–2618 (2009).

35. Liu, H. et al. Discovery of six new susceptibility loci and analysis of pleiotropic effects in leprosy. Nat. Genet. 47, 267–271 (2015).

36. Sales-Marques, C. et al. NOD2 and CCDC122–LACC1 genes are associated with leprosy susceptibility in Brazilians. Hum. Genet. 133, 1525–1532 (2014).

37. Wakil, S.M. et al. Association of a mutation in LACC1 with a monogenic form of systemic juvenile idiopathic arthritis. Arthritis Rheumatol. 67, 288–295 (2015).

38. Patel, N. et al. Study of Mendelian forms of Crohn’s disease in Saudi Arabia reveals novel risk loci and alleles. Gut 63, 1831–1832 (2014).

39. Cader, M.Z. et al. C13orf31 (FAMIN) is a central regulator of immunometabolic function. Nat. Immunol. 17, 1046–1056 (2016).

40. Sprechani, I. et al. Lymphoproliferative disorder and imbalanced T-helper response in C/EBPβ-deficient mice. EMBO J. 14, 1932–1941 (1995).

41. Tanaka, T. et al. Targeted disruption of the NF-IL6 gene discloses its essential role in bacteria killing and tumor cytototoxicity by macrophages. Cell 80, 353–361 (1995).

42. Zierhut, M. et al. Immunology and functional genomics of Behçet’s disease. Cell. Mol. Life Sci. 60, 1903–1922 (2003).
ONLINE METHODS

Subjects. We studied 2,014 Behçet’s disease cases and 1,826 genetically matched controls composed of the discovery and replication cohorts in previous GWAS and imputation studies2,5. We also included 969 Iranian cases and 826 controls recruited in a previous study42 and 608 Japanese cases and 737 controls recruited in a previous Japanese GWAS43 for replication. All Turkish and Iranian individuals affected with Behçet’s disease were diagnosed according to the International Study Group criteria for Behçet’s disease44 (Supplementary Table 19). All Japanese individuals affected with Behçet’s disease were diagnosed according to the Japanese Behçet’s disease criteria44 (Supplementary Table 20). Characteristics of each population are shown in Supplementary Table 21. All study participants provided written informed consent, and the study was approved by the ethics committee of each investigative institution.

Genotyping. We genotyped 2,014 Behçet’s disease cases and 1,826 healthy controls from the Turkish population on an Illumina Select HD custom genotyping array (Illuminomach) according to Illumina’s protocols. All samples were genotyped at the National Institutes of Health (Bethesda, Maryland, USA). Genotypes were called by GenCall using GRCh Build 37/hg19 mapping.

Data quality control. The cluster file was generated from samples with initial call rate >0.986 by the Illumina GenomeStudio GenTrain2.0 algorithm. Samples were excluded for call rate <0.85. After recalculating, markers were excluded for call frequency <0.95 and GenTrain score <0.5. After data cleaning, the data of 3,737 samples across 185,548 markers were exported to Golden Helix SVS 8.3.3 software. Markers on chromosomes X and Y were excluded. For further quality control, samples were excluded for call rate <0.95 and markers were excluded for call rate <0.95, minor allele frequency (MAF) <0.01, and deviation from Hardy–Weinberg equilibrium \( (P < 1 \times 10^{-5}) \).

A set of 38,256 LD-pruned markers with \( r^2 < 0.5 \) was used to estimate identity by descent. For each pair or trio of individuals with pi-hat >0.18, the sample with the higher or highest call rate was included. Principal-component analysis was used to estimate population stratification (Supplementary Fig. 6).

A set of LD-pruned markers remaining after removal of the MHC region and long-range LD regions45 was used for estimation of genomic inflation factors, \( \lambda_{GC} \) and \( \lambda_{0,00} \) (Supplementary Fig. 7). After quality control, a total of 1,900 samples, 1,779 controls, and 130,647 markers were included in the association analysis. This sample size provides greater than 83.6% power to detect a disease with 0.4% prevalence46.

Analysis of the MHC region. Immunochip genotyping data in the MHC region were imputed to type classical HLA alleles by SNP2HLA using the reference data collected by the Type 1 Diabetes Genetic Consortium47. Additional SNP genotypes from this region were also imputed by IMPUTE2 (ref. 48) after phasing by SHAPEIT49. For quality control, markers with MAF <0.01 and Hardy–Weinberg equilibrium \( P < 1 \times 10^{-5} \) in controls were excluded. The concordance rate per allele in 2,186 samples for which HLA-B*51 was directly typed was 98.6%.

Statistical association tests. Single-marker associations were evaluated by basic allele tests comparing the allele frequencies in cases and controls using Golden Helix SVS 8.3.3 software. The correlation/trend test was performed, and \( P < 5 \times 10^{-8} \) was considered to be the threshold for genome-wide significance. We also evaluated \( P \) values corrected by genomic inflation of the Turkish population (\( P_{GC} \)). Disease associations with markers reported in the previous studies were also evaluated. Conditional logistic regression analysis was performed to identify independently associated markers. After conditioning on a lead marker in each genome-wide significant locus, an additional marker was considered to be independently associated with Behçet’s disease when \( P < 5 \times 10^{-5} \). Statistical power for the original GWAS collection, the Immunochip study, and the replication cohorts was calculated with CaTS50.

Additional genotyping and imputation. From the Immunochip association analysis, we selected the lead marker(s) with \( P < 5 \times 10^{-5} \) from 20 new loci (3 loci from Table 2 and 17 loci from Supplementary Table 5). Markers located within about \( \pm 100 \) kb of these lead markers were selected for imputation. Because the \( IL1A–IL1B \) and \( PTPN1 \) loci were sparsely genotyped by the Immunochip, fine-mapping was performed for these loci before imputation with iPLEX assays (TOF MS, Agena) using the same Turkish samples. The Tagger SNP selection tool from HapMap was used to select SNPs with the intent of obtaining 100% coverage of the HapMap phase 3 SNPs with MAF greater than 1% in the CEU HapMap population with pairwise \( r^2 > 0.8 \). Although already tagged, additional SNPs with \( r^2 > 0.8 \) with the most significantly associated SNP of the region were also included. After combining the fine-mapping and Immunochip data, we imputed these loci by IMPUTE2. The same loci were also imputed from Japanese GWAS data for the replication study. The 1000 Genomes Project Phase 1 integrated data set51 was used as the reference panel for imputation. Markers with info score >0.8 and genotypes with probability >0.9 were included in analyses. For quality control, markers with MAF <0.01 and Hardy–Weinberg equilibrium \( P < 1 \times 10^{-5} \) in controls were excluded.

Replication. Disease associations for susceptibility markers previously reported for 11 loci outside of the MHC region were analyzed for genotyped and imputed markers. \( P < 0.0045 \) (0.05 corrected for 11 loci) was considered to indicate replication. Lead SNPs genotyped by Immunochip in the Turkish population for each new suggestive locus with \( P < 5 \times 10^{-5} \) were selected for genotyping in Iranian individuals by iPLEX assays (TOF MS, Agena) for replication. Imputed Japanese GWAS data were also used for replication in the Japanese population. Meta-analysis in multiple populations was performed using META48. For a comprehensive assessment of the new suggestive loci in imputed Turkish Immunochip and Japanese GWAS data, all the available markers with suggestive association in the Turkish population \( (P < 5 \times 10^{-5}) \) were analyzed \( (n = 215) \). The \( P \) value threshold for replication in the Iranian and Japanese cohorts was corrected for the number of independent markers \( (n = 37) \) after LD pruning to \( r^2 < 0.8 \) \((P < 0.0014)\). The \( P \) value of heterogeneity and \( I^2 \) were calculated to evaluate heterogeneity between populations. \( P_{\text{BLM}} < 0.05 \) and \( I^2 > 0.5 \) were considered to be significant.

Association analysis for homozygous FUT2 non-secretor alleles. Turkish Immunochip data from the FUT2 locus were used to impute regional variants by the same methods as other loci to obtain genotyping data for rs601338, the common FUT2 non-secretor SNP. Genotyping of the common Asian non-secretor allele at rs1047781 was performed in the Japanese population \( (594 \text{ cases and } 692 \text{ controls}) \) using the Taqman \( 5' \)-exonuclease assay with validated Taqman primer–probe sets (Applied Biosystems). The probe fluorescence signal was detected using the StepOnePlus Real-Time PCR System (Applied Biosystems) following the manufacturer’s protocol. Genotype data for rs601338 from the previous study42 were used for the meta-analysis of homozygous non-secretor genotypes among the three populations.

Annotation. To develop mechanistic hypotheses, we investigated chromatin states, conservation, and regulatory motifs altered by SNPs identified in this study using HaploReg \( (v4.1) \) (ref. 52). The functional effects of nonsynonymous coding variants were predicted by PolyPhen-2 (ref. 53). eQTL data were extracted from Genevar44,55, the Blood eQTL browser56, and GTEx57 to investigate the association between a disease susceptibility SNP and a target gene. \( P \)-value significance thresholds were applied as described in the original reports45–57.

Cytokine assays in PBMCs. Whole-blood samples from healthy controls were collected in sodium heparin tubes. PBMCs were purified by Ficoll (Ficoll-Paque PLUS, GE Healthcare) using Leucosep tubes (Greiner Bio-One) by gradient centrifugation. Cells were then washed twice with PBS (Life Technologies) \( (400 \text{ g for } 10 \text{ min at room temperature followed by } 250 \text{ g for } 12 \text{ min at } 4°C) \) and once with RPMI-1640 medium (Life Technologies) with FBS (300 g for 5 min at \( 4°C)\). The washed PBMCs were plated in triplicate in a 96-well plate \( (2 \times 10^5 \text{ cells/ml}) \) in RPMI-1640 medium with FBS. Cells were left untreated or were stimulated with zymosan \( (10 \text{ mg/ml}) \) at \( 37°C \) in \( 5\% \text{ CO}_2 \) for 24 h. The supernatants of cultured PBMCs were collected after centrifugation and stored at \( –80°C \). Cytokine concentrations were detected using the Affymetrix eBioScience Human Simplex kits for IL-1α and IL-1β and a Bio-Rad Bio-Plex 200 Luminox system according to the manufacturer’s instructions. We used 50 μl of PBMC culture supernatant for the immunoassays. The data were analyzed for statistical significance using the two-tailed Mann–Whitney test \( (P < 0.05) \).
Behçet’s disease susceptibility loci and overlap with other diseases. Susceptibility loci and lead SNPs that overlapped Behçet’s disease susceptibility loci in other immune-related diseases and leprosy were extracted from ImmunoBase and the National Human Genome Research Institute (NHGRI) GWAS catalog. If the Behçet’s disease lead SNP or a SNP in strong LD ($r^2 > 0.8$) was associated with the other disease, concordance of allelic effect was ascertained. In diseases with susceptibility loci with no disease-associated markers in strong LD with the Behçet’s disease marker, concordant effects of disease-associated alleles were identified by eQTL effects, if available. LD data from the CEU population of the 1000 Genomes Project was applied. eQTL databases (Genevar54,55, Blood eQTL browser56, GTEx57, and summary data from 12 studies available in HaploReg v4) were used to extract gene expression data. Permutation tests were performed to evaluate the number of disease risk loci shared by Behçet’s disease and IBD (Crohn’s disease and ulcerative colitis) or leprosy by a random selection of 1,000,000 sets of the same number of susceptibility loci for each disease from the RefSeq gene list (22,345 genes). Associations with Behçet’s disease in this study for markers located in susceptibility loci for IBD or leprosy from previous studies were plotted to evaluate whether they were enriched for associations with Behçet’s disease.

Pathway analysis. The pathways in which susceptibility genes are involved were analyzed in the Gene Ontology (GO) database by DAVID (v6.7) (ref. 58). We used 21 susceptibility genes (including 9 newly associated genes) from this study in the pathway analysis. $P$ values were corrected by Benjamini’s method. $P_{\text{corrected}} < 0.05$ was considered to be significant.

Data availability. Statistical summary data for all markers genotyped with the Immunochip for which the discovery cohort showed genome-wide significant or suggestive association ($P < 5 \times 10^{-8}$) are provided in a Supplementary Data Set.