



SCIENTIFIC REPORTS



OPEN

Behçet's disease and genetic interactions between HLA-B*51 and variants in genes of autoinflammatory syndromes

Sergio Burillo-Sanz¹ , Marco-Antonio Montes-Cano¹, José-Raúl García-Lozano¹, Israel Olivas-Martínez¹, Norberto Ortego-Centeno², Francisco-José García-Hernández³, Gerard Espinosa⁴, Genaro Graña-Gil⁵, Juan Sánchez-Bursón⁶, María Rosa Juliá⁷, Roser Solans⁸, Ricardo Blanco⁹, Ana-Celia Barnosi-Marín¹⁰, Ricardo Gómez de la Torre¹¹, Patricia Fanlo¹², Mónica Rodríguez-Carballeira¹³, Luis Rodríguez-Rodríguez¹⁴ , Teresa Camps¹⁵, Santos Castañeda¹⁶, Juan-Jose Alegre-Sancho¹⁷, Javier Martín¹⁸ & María Francisca González-Escribano¹

Behçet's disease (BD) is an immune-mediated systemic disorder with a well-established genetic base. In a previous study, using a next generation sequencing approach, we found many rare variants and some functional polymorphisms in genes related to autoinflammatory syndromes (AID): *CECR1*, *MEFV*, *MVK*, *NLRP3*, *NOD2*, *PSTPIP1* and *TNFRSF1A* in our BD cohort. Our strategy did not allow us to establish either number of patients with variants, proportion of individuals accumulating them or relationship with other genetic factors. With the goal to answer these questions, the individual samples were sequenced. Additionally, three functional polymorphisms: *NLRP3* p.Gln703Lys, *NOD2* p.Arg702Trp and p.Val955Ile were genotyped using TaqMan assays. A total of 98 patients (27.6%) carried at least one rare variant and 13 of them (3.7%) accumulated two or three. Functional regression model analysis suggests epistatic interaction between B51 and *MEFV* ($P = 0.003$). A suggestive protective association of the minor allele of *NOD2* p.Arg702Trp ($P = 0.01$) was found in both, B51 positive and negative individuals. Therefore, a high percentage of patients with BD have rare variants in AID genes. Our results suggest that the association of *MEFV* with BD could be modulated by the HLA molecules; whereas the protective effect of *NOD2* p.Arg702Trp would be independent of HLA.

¹Department of Immunology, Hospital Universitario Virgen del Rocío (IBiS, CSIC, US), Sevilla, 41013, Spain.

²Department of Internal Medicine, Hospital Clínico San Cecilio, Granada, 18003, Spain. ³Department of Internal Medicine, Hospital Universitario Virgen del Rocío, Sevilla, 41003, Spain. ⁴Department Autoimmune Diseases, Hospital Clinic Universitari, Barcelona, 08036, Spain. ⁵Department of Rheumatology, Complejo Hospitalario Universitario, A Coruña, 15006, Spain. ⁶Department of Rheumatology, Hospital Universitario de Valme, Sevilla, 41014, Spain. ⁷Department of Immunology, Hospital Universitari Son Espases, Palma de Mallorca, Illes Balears, 07120, Spain. ⁸Department of Internal Medicine, Autoimmune Systemic Diseases Unit, Hospital Vall d'Hebron, Universidad Autonoma de Barcelona, Barcelona, 08035, Spain. ⁹Department of Rheumatology, Hospital Universitario Marqués de Valdecilla, Santander, 39008, Spain. ¹⁰Department of Internal Medicine, Complejo Hospitalario Torrecárdenas, Almería, 04009, Spain. ¹¹Department of Internal Medicine, Hospital Universitario Central de Asturias, Asturias, 33011, Spain. ¹²Department of Internal Medicine, Hospital Virgen del Camino, Pamplona, 31008, Spain. ¹³Department of Internal Medicine, Hospital Universitari Mútua Terrassa, Terrassa, 08221, Spain. ¹⁴Department of Rheumatology, Hospital Clínico San Carlos, Madrid, 28040, Spain. ¹⁵Department of Internal Medicine, Hospital Regional Universitario, Málaga, 29010, Spain. ¹⁶Department of Rheumatology, Hospital de la Princesa, IIS-Princesa, Madrid, 28006, Spain. ¹⁷Department of Rheumatology, Hospital Universitario Doctor Peset, Valencia, 46017, Spain. ¹⁸Instituto de Parasitología y Biomedicina "López-Neyra", CSIC, PTS, Granada, 18016, Spain. Correspondence and requests for materials should be addressed to M.F.G.-E. (email: mariaf.gonzalez.sspa@juntadeandalucia.es)

Behçet's disease (BD) is a complex and systemic inflammatory syndrome; from a clinical point of view, it is a vasculitis that presents heterogeneous phenotypes. Genital and oral ulcers are the first and most common manifestations, although bilateral uveitis and cutaneous involvement are also present with a relatively high frequency, and there are occasional manifestations in other organs or systems, such as the nervous system^{1–3}. Gender distribution is variable among countries, but unlike other immune-mediated diseases there is not a predominance of females among affected subjects⁴. BD is, in general, a rare disease although its prevalence is higher along the old “Silk Road”, especially in Turkey (up to 370 cases per 100000 population)⁴. The pathogenic mechanisms of BD remain unknown, although there are accumulated evidences of genetic predisposition and data suggesting that an aberrant immune response to certain infectious agents and some environmental factors may trigger the disease⁵. The first described genetic factor associated with the disease was HLA-B*51⁶, which is present in up to 60% of patients. More recently, genome wide association studies (GWAS) reinforce HLA-B*51 as the most important genetic factor in susceptibility to this disease and, in addition, they support the contribution of other genes related with the immune system, such as *IL23R*, *IL10*, *STAT4*, *CCR1-CCR3*, *KLRC4*, *ERAP1*, *TNFAIP3* and *FUT2*⁷.

There is a considerable clinical overlap of BD with autoinflammatory syndromes (AID) which are characterised by a misregulation of the inflammasome proteins that lead to an increased release of IL-1b. AID are genetically conditioned diseases, caused by aberrant proteins encoded by pathogenic variants of genes related with both the inflammasome itself (for example, NLRP3) as well as with other sensing proteins involved in pathogen- and damage-associated molecular patterns (PAMPs and DAMPs) or in the signalling pathways⁸. Heat shock proteins (HSP), are a group of intracellular proteins with scavenger roles found from bacteria to mammals. Human NF-M is homolog of bacterial HSP-65, and BD patients' sera show crossreactivity with these two proteins, suggesting that bacterial HSP could be recognised by TLRs and act as a trigger antigen for BD development^{9,10}. Based on the above, it is plausible that variants in genes involved in AID could also be related to BD. In this sense, in a previous study, which included *CECR1*, *MEFV*, *MVK*, *NLRP3*, *NOD2*, *PSTPIP1* and *TNFRSF1A*, we reported that many rare variants in these AID related genes are found in BD patients¹¹.

Gene-gene interactions are assumed to be a common factor in the genetic components of human diseases¹², therefore, it is interesting to establish possible relationships among the different factors involved in the susceptibility to diseases. For instance, interaction between HLA and other genes associated with diseases has been demonstrated in BD and other immuno-mediated syndromes related to HLA class I, such as ankylosing spondylitis and psoriasis^{7,13,14}. Our previous study with AID related genes was performed using next generation sequencing (NGS) with a pooled strategy which did not allow us to assign each variant to an individual. The aim of the present study was to determine the number of patients carrying variants, the proportion of individuals that accumulate variants and the relationship with other genetic factors, such as HLA-B*51.

Materials and Methods

Subjects. The study population consisted of 355 patients diagnosed with BD according the 1990 International Study Group classification criteria for this disease^{15,16}. All the patients had Spanish European ancestry¹⁶, 43.7% were male and 36% were B51 positive. The clinical features of this cohort have been previously published¹¹. The study was approved by the ethical committees of all centers and hospitals involved (Hospital Universitario Virgen del Rocío, Hospital Clínico San Cecilio, Hospital Universitari Clínic, Complejo Hospitalario Universitario A Coruña, Hospital Universitario de Valme, Hospital Universitari Son Espases, Hospital Vall d'Hebron, Hospital Universitario Marqués de Valdecilla, Complejo Hospitalario Torrecárdenas, Hospital Universitario Central de Asturias, Hospital Virgen del Camino, Hospital Universitari Mútua Terrassa, Hospital Clínico San Carlos, Hospital Regional Universitario de Málaga, Hospital de la Princesa, Hospital Universitario Doctor Peset, Instituto de Parasitología y Biomedicina López-Neyra). The healthy control population in the rare variants' analysis, consisted on the whole genome sequences of 107 individuals from the Iberian population of Spain (IBS population), included in the phase 3 of the 1000 Genomes Project; whereas in polymorphisms analysis 363 in-house healthy controls were used.

AID Related Genes Genotyping. In a previously published study with the same cohort, 54 rare variants (minor allele frequency, MAF < 0.05 in IBS population) were found in the 7 genes included in the study by using an NGS approach of pooled samples¹¹. In the present study, in order to identify those patients carrying each variant, all the samples included in the pools which contained each rare variant were individually studied. 51 of the 54 rare variants were sequenced by Sanger and the remaining three (NLRP3 Val198Met: rs121908147, C__64676050_10, NOD2 Asn289Ser: rs5743271, C__26935007_10 and TNFRSF1A Arg121Gln: rs4149584, C__12029224_20) were TaqMan genotyped. In addition, to determine whether close variants (in the same exon) are in the same (*cis*) or in different (*trans*) chromosome, a Hexaprimer Amplification Refractory Mutation System PCR (H-ARMS PCR) was used¹⁷. The pairs of variants studied were: *MEFV* p.Leu110Pro and p.Glu148Gln, using the primers MEFV-H_ARMS_110_148-F1, R1, F2, R2, F3 and R3 and *NOD2* p.Arg311Trp and p.Arg703Cys with the primers NOD2-H_ARMS_311_703-F1, F2, R1 and R2 (Supplementary Table 1). In the case of *MEFV* H-ARMS, the following thermal-cycler conditions were used: 3 min at 95 °C followed by 35 cycles of: 20 s at 95 °C, 30 s at 65 °C and 45 s at 72 °C. In this case, an informative band of 145 bp is present only when both variants are in *trans* (Supplementary Fig. 1A). Regarding *NOD2*, a modified H-ARMS was used, since the 2 variants are more distant, a two-PCR reactions system was chosen (one with the pair of primers F1 and R1 and another with the pair F2 and R2). The thermal-cycler conditions were: 20 s at 95 °C, 30 s at 70 °C and 90 s at 72 °C. In this case, an informative band of 1228 bp in one of the PCR tubes is present only when both variants are in *trans* (Supplementary Fig. 1B).

Regarding the three functional polymorphisms found in the previous study, NLRP3 p.Gln703Lys, NOD2 p.Arg702Trp and p.Val955Ile were individually genotyped in all the samples included in those pools in which the variants had been detected. The following TaqMan assays were used: NLRP3 p.Gln703Lys (rs35829419,

C__25648615_10), NOD2 p.Arg702Trp (rs2066844, C__11717468_20) and p.Val955Ile (rs5743291, C__25651076_10). Real-Time PCR assays were performed using TaqPath™ ProAmp™ Master Mix (Thermo Fisher Scientific, Waltham, MA), in a 7500 Fast Real-Time PCR System, according to manufacturer's recommendations. No statistically significant differences in the allelic frequencies of the polymorphisms were observed between NGS and TaqMan assay results. The remaining 4 polymorphisms (MAF > 0.05 in IBS population) found in our previous study were excluded from the present study because they do not have a functional relevance.

Regarding the controls, genotypes of rare variants were retrieved from the IBS population of the 1000 Genomes Project database and they were imported to PLINK software for genotype counting and association testing. In addition, in the cases of the functional polymorphisms NLRP3 p.Gln703Lys, NOD2 p.Arg702Trp and p.Val955Ile, 363 in-house individuals were genotyped using the same TaqMan assays as in patients and used as local ethnically matched controls.

HLA-B Low Resolution Genotyping. The locus HLA-B was genotyped in the 343 BD patients and the 363 local controls by using a standard low resolution (first set of two digits) PCR-SSOP Luminex method (LABType SSO One Lambda Inc., Canoga Park, CA). The remaining 12 patients could not be genotyped in HLA because of lack of material or its poor quality. Regarding the IBS controls, HLA-B genotype of IBS population of the 1000 genomes project was called at the same level of resolution as the BD patients using a pipeline optimised for OptiType software¹⁸. Briefly, chr6:29910247-31325022 hg19 region was retrieved from the mapped exome BAM files and converted to FASTQ with bamtools; the readings were aligned to all the available HLA allele sequences of IMGT/HLA with razerS3 and these alignments were analysed by OptiType software for HLA class I type calling. Data were visually inspected for proper coverage of HLA class I exons. The allelic frequency of HLA-B*51 by this method in IBS population (N = 107) was 0.09, which is not significantly different from that obtained in the general Spanish population with a standard HLA typing method (0.07)¹⁹.

Statistical analysis. To analyse the relationship between HLA-B*51 and rare variants in the AID genes, BD patients were divided into two groups: B51 positive (assigned as affected phenotype) and B51 negative (assigned as non-affected phenotype) and the regions studied were those defined by the field refGene.name2. SKAT test implemented by Variant Tools software^{20,21} was used to analyse the distribution of rare variants in each of the genes included in the study in both patient groups. Additionally, to detect interactions between rare variants of each pair of AID genes as well as between each AID genes and HLA-B*51, a recently developed functional regression model was used with the R package FRGEpistasis²². This method has been used for quantitative traits²³, but it is also suitable for the binary ones. Briefly, the cases/controls were encoded as 1/0, the AID genes variants were codified as 0 (wild-type) or 1 (variant) and also the B51 status (1 positive, 0 negative). P-values were adjusted (P_{adj}) according to Bonferroni's (considering 8 tests: 7 AID genes and B51), therefore P-values below 0.0063 were considered statistically significant. Type I error was estimated using a simulation with data retrieved from the 1000 genomes database (Supplementary Table 2). In addition, FRG analysis was also done in our BD patients and IBS controls with randomly assigned the phenotypes (affected and unaffected) (Supplementary Table 3). Q-Q plots of the p-values obtained in FRG with both, the real and the random data are displayed in Supplementary Fig. 2.

In regards to the disease association study of the 3 functional polymorphisms: NLRP3 p.Gln703Lys, NOD2 p.Arg702Trp and p.Val955Ile, each sample was encoded as 0, 1 and 2 according to the number of copies of the minor allele of each polymorphism carried by the individual. Allelic and genotypic distributions in dominant (for minor allele) and allelic models in BD patients and controls were compared with X2 in two by two contingency Tables with 1-degree freedom using PLINK software and P-values were corrected by Bonferroni's adjustment considering 12 tests, P values < 0.004 were considered statistically significant and the Odds Ratios (ORs) were calculated with *fmsb* package for R. Finally, to analyse interaction between B51 and these functional polymorphisms, conditional logistic regression models were constructed by categorizing all the individuals according to the presence/absence of the B51 (dominant model) and the genotype of each polymorphism (dominant model for the minor allele) and they were analysed with Epi Info 2002. Additionally, the R package FRGEpistasis²² was also used to detect interaction of these polymorphisms with B51.

To quantify linkage disequilibrium (LD) when it was relevant, the R values were calculated in the patients and local controls by using PLINK, whereas in the EUR population, LD between variants was consulted in HaploReg v.4.1²⁴. A descriptive circular layout figure, performed with circlize R package, was used to graphically represent the rare variants that are found together in same patient²⁵.

Ethical use of human participants statement. All participants signed a written informed consent to publish identified information prior to their enrolment in the study. All methods were performed after obtaining the signed informed consent from patients, and in accordance to our institution ethical committee: Hospital Universitario Virgen del Rocio's ethical committee, and with the Helsinki Declaration of 1975, as revised in 1983.

Results

The individual genotyping of the samples revealed that 98 patients out of the total of 355 in our cohort (27.6%) carried at least one of the total of 54 rare variants found in the genes included in the study. From these 98 patients, 82 harboured only one rare variant, 13 had two and 3 had three, in the same or different genes. All the patients with one variant had a single copy, except two homozygous (one NOD2 p.Gly908Arg and the other TNFRSF1A p.Arg121Gln). In regard to the 13 patients carrying two variants, in 3 cases the variants were located in the same gene and in 10 cases they were in different genes. Regarding the 3 patients with 3 variants, 2 of them had the combined allele MEFV p.Arg408Gln/p.Pro369Ser and one variant in a different gene, whereas the remaining patient had each variant in a different gene. Figure 1 illustrates connections between variants and Table 1 displays the list of combinations found in our cohort. In order to know whether variants in the same gene were in *cis* or

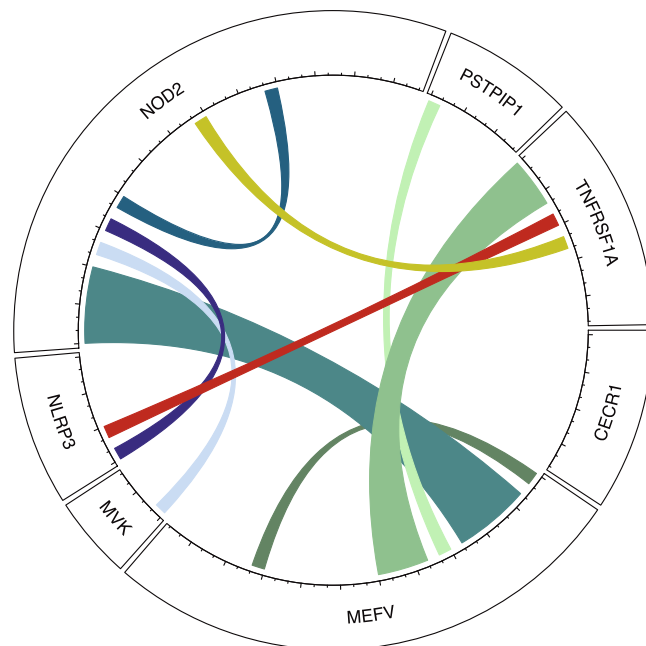


Figure 1. Descriptive circular layout showing coexistence of rare variants in patients. Length of sectors corresponds to number of variants; width of lines corresponds to the number of connections, each tick in inner circle axis represents one variant in one patient.

Type of combination	Variant combination	Nr. Patients
Homozygous	NOD2 Gly908Arg ^a	1
	TNFRSF1A Arg121Gln	1
Variants in the same gene	MEFV Arg408Gln, Pro369Ser ^b	1
	MEFV Glu148Gln, Leu110Pro ^c	1
	NOD2 Arg703Cys, Arg311Trp ^d	1
Variants in different genes	MEFV Glu148Gln, PSTPIP1 Val122Ile	1
	MEFV Ile591Thr, NOD2 Gly908Arg	1
	MEFV Ile591Thr, TNFRSF1A Arg121Gln	1
	MEFV Met694Val, NOD2 Leu349Phe	2
	MEFV Lys695Arg, NOD2 Ala755Val	1
	MEFV Ala744Ser, TNFRSF1A Arg121Gln	1
	MVK Val377Ile, NOD2 Gly908Arg ^a	1
	NLRP3 Val198Met, TNFRSF1A Arg312Lys	1
	NLRP3 Arg488Lys, NOD2 Asn289Ser	1
	MEFV Glu148Gln, TNFRSF1A Arg121Gln, NOD2 Leu248Arg	1
	MEFV Arg408Gln, Pro369Ser, TNFRSF1A Arg121Gln ^b	1
	MEFV Arg408Gln, Pro369Ser, NOD2 Asn289Ser ^b	1

Table 1. List of rare variants found in combination in the seven AID genes studied in our cohort of BD patients. ^aPatient homozygous NOD2 Gly908Arg is the same as the one with the variant MVK Val377Ile. ^bMEFV Arg408Gln and Pro369Ser are in a strong LD in the European control population ($r^2 = 1$). A total of 3 patients have these two variants in our cohort. ^cMEFV Leu110Pro / Glu148Gln LD has not been found in control population, but this combination has been reported as a complex allele in FMF and according to the PCR-H-ARMS performed in this patient these variants are in *cis*. ^dAccording to the results of the PCR-H-ARMS performed in this patient these variants are in *trans*.

in *trans*, PCR-H-ARMS was performed in the patients with MEFV p.Glu148Gln/p.Leu110Pro and with NOD2 p.Arg703Cys/p.Arg311Trp. In the case of MEFV, the variants were in *cis* whereas in NOD2 they were in *trans* (Supplementary Fig. 1A,B). No additional study was made in the 3 individuals with the combination MEFV p.Arg408Gln/p.Pro369Ser since these variants have been reported as being in a strong LD in the European control population ($r^2 = 1$), therefore *cis* configuration was assumed.

The results of the SKAT test suggested non-significant differences in the distribution of rare variants between B51 positive and B51 negative patients ($p > 0.05$, Table 2). Nevertheless, functional regression model analysis,

Gene	SKAT P-value ^a (N = 336)	FRG P-value ^b (N = 443)
<i>CECR1</i>	0.77	0.51
<i>MEFV</i>	0.75	0.003
<i>MVK</i>	1.0	0.50
<i>NLRP3</i>	0.73	0.19
<i>NOD2</i>	0.99	0.77
<i>PSTPIP1</i>	0.82	0.40
<i>TNFRSF1A</i>	0.46	0.85

Table 2. Analysis of the distribution of rare variants in seven genes related to AID in B51 positive and negative BD patients using two different approaches. ^aB51 positive (N = 147) vs. B51 negative (N = 189) patients were compared using SKAT. ^bBD patients (N = 336) vs. IBS 1000 Genomes controls (N = 107) were compared using FRG analysis for epistasis. Only the results of the epistasis analysis between B51 and the AID genes are displayed; although interactions between pairs of AID genes were also analysed, no interactions were detected (p-values > 0.05 in all the cases). Bonferroni's adjusted significance level is p-value < 0.0063.

		Patients		Healthy Controls		P	OR (95% CI)
		BD N = 355	Local N = 363	IBS N = 107	Overall ^a		
NLRP3 Gln703Lys	A/A	0 (0%)	0 (0%)	3 (2.8%)	3 (0.6%)		
	C/A	32 (9.0%)	32 (8.8%)	10 (9.3%)	42 (8.9%)		
	C/C	322 (91.0%)	331 (91.2%)	94 (87.8%)	425 (90.4%)		
	A	32 (0.05)	32 (0.04)	16 (0.08)	48 (0.05)	0.58	0.88 (0.56–1.39)
	C	676 (0.95)	694 (0.96)	198 (0.92)	892 (0.95)		
NOD2 Val955Ile	A/A	8 (2.3%)	2 (0.6%)	4 (3.7%)	6 (1.3%)		
	A/G	51 (14.4%)	58 (16%)	20 (18.7%)	78 (16.6%)		
	G/G	296 (83.4%)	303 (83.4%)	83 (77.6%)	386 (82.1%)		
	A	67 (0.09)	62 (0.09)	28 (0.13)	90 (0.1)	0.92	0.98 (0.71–1.37)
	G	643 (0.91)	664 (0.91)	186 (0.87)	850 (0.9)		
NOD2 Arg702Trp	T/T	0 (0%)	1 (0.3%)	0 (0.0%)	1 (0.2%)		
	C/T	18 (5.0%)	33 (9.1%)	12 (11.2%)	45 (9.6%)		
	C/C	336 (95.0%)	329 (90.6%)	95 (88.8%)	424		
					(90.2%)		
	T	18 (0.03)	35 (0.05)	12 (0.06)	47 (0.05)	0.011	0.496 (0.29–0.86)
	C	690 (0.97)	691 (0.95)	202 (0.94)	893 (0.95)		

Table 3. Distribution of frequencies of NLRP3 and NOD2 functional polymorphisms in BD cases and healthy controls. ^aNo differences in allelic frequencies between local and IBS 1000 genomes controls were found, therefore, overall control data were used to calculate p-values and ORs. Bonferroni's adjusted significance level considered significant is P-values < 0.004.

suggests epistatic interaction between HLA-B*51 and MEFV (p < 0.05, Table 2) in BD. No evidences of interaction between pairs of AID related genes were detected.

Table 3 displays the frequencies of the three functional polymorphisms in BD patients and healthy controls. A suggestive association was found in the case of NOD2 p.Arg702Trp (p = 0.011 in the allelic model and p = 0.014 in the dominant model: TT + CT vs. CC) with a protective effect of the minor allele. For the two remaining polymorphisms, no-significant association with BD was found (p > 0.05). No LD was detected between the two NOD2 polymorphisms included (R < 0.1 in all cases). With respect to the relationship of these polymorphism with B51, conditional regression analysis suggests that the protective effect of NOD2 p.Arg702Trp is independent of the HLA-B51, because the OR is decreased in both, B51 positive and negative groups (Table 4). The functional regression model analysis does not suggest epistatic interaction between B51 and these polymorphisms (p > 0.05 in all the cases).

Discussion

In the present study, a substantial percentage (about 30%) of the BD patients was found to have some rare variant in at least one out the seven AID related genes that were included. The proteins encoded by these AID genes are related with the IL-1b release, mainly through sensing bacterial or viral components and activating the NLRP3 inflammasome and, therefore, it would be reasonable to expect some degree of interaction among them. In this way, a relatively high ratio of our BD patients carried variants in different AID genes, suggesting that interplay between variants might exist. Hypothetically, these variants in AID genes could modify the association of B51 with BD or vice versa. It is necessary to note that, genetic interaction is difficult to detect because multilocus

HLA B51	NOD2 Arg702Trp	Patients	Controls	P	OR (95% CI)
+	+	10 (3.0)	5 (1.4)	0.04	3.11 (1.04–9.23)
+	–	136 (40.6)	46 (12.7)	<10 ^{–5}	4.59 (3.13–6.73)
–	+	8 (2.4)	27 (7.5)	0.06	0.46 (0.20–1.03)
–	–	181 (54.0)	284 (78.4)	<10 ^{–5}	1.0

Table 4. Conditional logistic regression model of HLA-B51 and NOD2 p.Arg702Trp.

genotype combinations for gene-gene interaction increase exponentially and require a larger sample size as well as more computation burden. In addition, most of the methods developed to detect gene-gene interactions have been designed for common variants and it is difficult to apply them to data with rare variants because the low frequency of these variants leads to a lack of statistical power^{26,27}. In spite of the use of a method designed for rare variants, no-interactions between AID genes were detected, although this result could be attributable to the low sample size, since only 13 patients accumulate variants in different genes. In addition, it is necessary to take into account that, from a biological point of view, gene-gene interaction is the phenomenon in which the effects of one gene on a trait (for example a disease) are modified by another gene or several other genes; whereas statistical epistasis is described as a deviation from additivity in a linear statistical model²⁸. In order to assess interaction between rare variants of AID genes and HLA-B*51, two different approaches were used; in the first, patients B51 positive and negative were compared with SKAT whereas in the second, patients and controls were compared using FRG. The only discrepancy observed between both tests was in the case of *MEFV* in which, both tests have apparently contradictory results. According to SKAT, no differences in the distribution of rare variants were detected between B51 positive and negative individuals whereas our results with FRG suggest interaction between HLA-B*51 and *MEFV*. The estimated type I error in our simulation is even lower for FRG than for SKAT therefore, discrepancy could be based in the characteristics of each test, SKAT was designed to study association of rare variants in complex diseases by grouping variants and checking the association of each entire gene, whereas FRG was specifically designed to study gene-gene interactions with rare variants. Contradictory results in the association of *MEFV* with BD have been reported^{29–32}, in fact, in our previous study in AID genes, *MEFV* was found associated only with SKAT but not with the rest of the statistical tests used¹¹. To our acknowledgement, no previous work on interaction between HLA and *MEFV* has been published for BD. Our results must be interpreted with caution, nevertheless, there is some evidence suggesting that involvement of *MEFV* in autoimmune diseases could be associated with HLA molecules^{33–35} and conversely, differential effects of HLA class I molecules in Japanese patients with familial Mediterranean fever (FMF), even in those with high-penetrance *MEFV* mutations, have been identified³⁶. *MEFV* protein interacts with cellular components of inflammasome whose activity is augmented by the pathogenic mutations³⁷. Therefore, it is reasonable to consider that FMF-associated *MEFV* mutations enhance the cellular responses in the inflammatory diseases which are also associated with certain HLA polymorphisms. In other words, mutations and polymorphisms in the *MEFV* gene are associated with subclinical inflammation therefore, it is possible that the presence of *MEFV* mutations increase the baseline level of inflammation, and the presence of a certain HLA molecule (such as HLA-B*51 in BD) modulates the disease. Although the same type of interaction could be hypothesized for any of the other AID genes studied, no more interactions were detected.

Regarding the polymorphisms in *NLRP3* and *NOD2*, only a protective association of *NOD2* p.Arg702Trp was found. The loss of function variants of *NOD2* have been reported as possibly protective in BD^{30,38}. For this reason, in our previous study, the frequencies of *NOD2* p.Asn289Ser, p.Arg702Trp and p.Gly908Arg in our patient cohort were compared with those of the 1000 genomes IBS control population¹¹. Although no significant differences were found, a trend was detected for p.Arg702Trp ($p = 0.09$) that in the present study reaches statistical significance by increasing the number of controls by including local ethnically matched individuals. Therefore, the present study suggests a protective effect of this *NOD2* variant at least in Caucasian patients. This effect seems to be independent of HLA-B*51.

In conclusion, a high percentage of patients with BD have one or several variants in AID genes. Additionally, although our data require replication, they suggest that the association of *MEFV* with BD could be modulated by the HLA molecules; whereas the protective effect of the loss of function of *NOD2* p.Arg702Trp polymorphism would be independent of HLA.

References

- Alpsoy, E. *et al.* Clinical features and natural course of Behçet's disease in 661 cases: A multicentre study. *British Journal of Dermatology* **157**, 901–906, <https://doi.org/10.1111/j.1365-2133.2007.08116.x> (2007).
- Kayıkçıoğlu, M. *et al.* Endothelial functions in Behçet's disease. *Rheumatology International* **26**, 304–308, <https://doi.org/10.1007/s00296-005-0590-1> (2006).
- Mendes, D. *et al.* Behçet's disease—a contemporary review. *Journal of autoimmunity* **32**, 178–188, <https://doi.org/10.1016/j.jaut.2009.02.011> (2009).
- Sakane, T., Takeno, M., Suzuki, N. & Inaba, G. Behçet's Disease. *New England Journal of Medicine* **341**, 1284–1291, <https://doi.org/10.1056/NEJM199910213411707> (1999).
- Lehner, T. The role of heat shock protein, microbial and autoimmune agents in the aetiology of Behçet's disease. *International reviews of immunology* **14**, 21–32, <https://doi.org/10.3109/08830189709116842> (1997).
- Yazici, H., Chamberlain, M. A., Schreuder, I., D'Amaro, J. & Muftuoglu, M. HLA antigens in Behçet's disease: a reappraisal by a comparative study of Turkish and British patients. *Annals of the rheumatic diseases* **39**, 344–348 (1980).

7. Kirino, Y. *et al.* Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nature Genetics* **45**, 202–207, <https://doi.org/10.1038/ng.2520> (2013).
8. Manthiram, K., Zhou, Q., Aksentjevich, I. & Kastner, D. L. The monogenic autoinflammatory diseases define new pathways in human innate immunity and inflammation. *Nature Immunology* **18**, 832–842, <https://doi.org/10.1038/ni.3777> (2017).
9. Lule, S. *et al.* Behcet Disease serum is immunoreactive to neurofilament medium which share common epitopes to bacterial HSP-65, a putative trigger. *J Autoimmun* **84**, 87–96, <https://doi.org/10.1016/j.jaut.2017.08.002> (2017).
10. Hasan, H. W. *et al.* Experimental mucosal induction of uveitis with the 60-kDa heat shock protein-derived peptide 336-351. *Eur. J. Immunol* **28**, 2444–2455, [https://doi.org/10.1002/\(SICI\)1521-4141\(199808\)28:08%3C2444::AID-IMMU2444%3E3.0.CO;2-N](https://doi.org/10.1002/(SICI)1521-4141(199808)28:08%3C2444::AID-IMMU2444%3E3.0.CO;2-N) (1998).
11. Burillo-Sanz, S. *et al.* Mutational profile of rare variants in inflammasome-related genes in Behçet disease: A Next Generation Sequencing approach. *Scientific Reports* **7**, 8453–8453, <https://doi.org/10.1038/s41598-017-09164-7> (2017).
12. Moore, J. H. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Human heredity* **56**, 73–82, 73735 (2003).
13. Evans, D. M. *et al.* Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* **43**, 761–767, <https://doi.org/10.1038/ng.873> (2011).
14. Strange, A. *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet* **42**, 985–990, <https://doi.org/10.1038/ng.694> (2010).
15. International Study Group for Behet's, D. Criteria for diagnosis of Behçet's disease. International Study Group for Behçet's Disease. *Lancet* **335**, 1078–1080, [https://doi.org/10.1016/0140-6736\(90\)92643-V](https://doi.org/10.1016/0140-6736(90)92643-V) (1990).
16. 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–e0161305, <https://doi.org/10.1371/journal.pone.0161305> (2016).
17. Piccioli, P. *et al.* Hexaprimer amplification refractory mutation system PCR for simultaneous single-tube genotyping of 2 close polymorphisms. *Clin Chem* **54**, 227–229, <https://doi.org/10.1373/clinchem.2007.095703> (2008).
18. Szolek, A. *et al.* OptiType: Precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316, <https://doi.org/10.1093/bioinformatics/btu548> (2014).
19. Montes-Cano, M. *et al.* HLA and non-HLA genes in Behçet's disease: a multicentric study in the Spanish population. *Arthritis Research & Therapy* **15**, R145–R145, <https://doi.org/10.1186/ar4328> (2013).
20. Wang, G. T. *et al.* Variant association tools for quality control and analysis of large-scale sequence and genotyping array data. *American journal of human genetics* **94**, 770–783, <https://doi.org/10.1016/j.ajhg.2014.04.004> (2014).
21. Wu, M. C. *et al.* Rare-variant association testing for sequencing data with the sequence kernel association test. *American journal of human genetics* **89**, 82–93, <https://doi.org/10.1016/j.ajhg.2011.05.029> (2011).
22. Zhang, F., Boerwinkle, E. & Xiong, M. Epistasis analysis for quantitative traits by functional regression model. *Genome Res* **24**, 989–998, <https://doi.org/10.1101/gr.161760.113> (2014).
23. Zhang, F., Xie, D., Liang, M. & Xiong, M. Functional Regression Models for Epistasis Analysis of Multiple Quantitative Traits. *PLoS Genet* **12**, e1005965, <https://doi.org/10.1371/journal.pgen.1005965> (2016).
24. Ward, L. D. & Kellis, M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Research* **40**, D930–D934, <https://doi.org/10.1093/nar/gkr917> (2012).
25. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812, <https://doi.org/10.1093/bioinformatics/btu393> (2014).
26. Sham, P. C. & Purcell, S. M. Statistical power and significance testing in large-scale genetic studies. *Nature reviews. Genetics* **15**, 335–346, <https://doi.org/10.1038/nrg3706> (2014).
27. Steen, K. V. Travelling the world of gene-gene interactions. *Briefings in bioinformatics* **13**, 1–19, <https://doi.org/10.1093/bib/bbr012> (2012).
28. Fisher, R. The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb* **52**, 399–433 (1918).
29. Baruch, Y. *et al.* MEFV, TNFRSF1A and CARD15 mutation analysis in Behçet's disease. *Clinical and experimental rheumatology* **29**, S24–27 (2011).
30. Kirino, Y. *et al.* Targeted resequencing implicates the familial Mediterranean fever gene MEFV and the toll-like receptor 4 gene TLR4 in Behcet disease. *Proceedings of the National Academy of Sciences* **110**, 8134–8139, <https://doi.org/10.1073/pnas.1306352110> (2013).
31. Tasliyurt, T., Yigit, S., Rustemoglu, A., Gul, U. & Ates, O. In *Gene* Vol. 530 100–103 (2013).
32. Wu, Z. *et al.* Association between MEFV Mutations M694V and M680I and Behçet's Disease: A Meta-Analysis. *PLOS ONE* **10**, e0132704, <https://doi.org/10.1371/journal.pone.0132704> (2015).
33. Cheraoui, B. *et al.* Boundaries between familial Mediterranean fever and juvenile spondyloarthritis: Analysis of three French retrospective cohorts. *Joint, bone, spine: revue du rhumatisme*, <https://doi.org/10.1016/j.jbspin.2018.01.010> (2018).
34. Cosan, F. *et al.* Association of familial Mediterranean fever-related MEFV variations with ankylosing spondylitis. *Arthritis Rheum* **62**, 3232–3236, <https://doi.org/10.1002/art.27683> (2010).
35. Wakil, S. M. *et al.* Association of a mutation in LACC1 with a monogenic form of systemic juvenile idiopathic arthritis. *Arthritis & Rheumatology (Hoboken, N.J.)* **67**, 288–295, <https://doi.org/10.1002/art.38877> (2015).
36. Yasunami, M. *et al.* Identification of Disease-Promoting HLA Class I and Protective Class II Modifiers in Japanese Patients with Familial Mediterranean Fever. *PLoS One* **10**, e0125938, <https://doi.org/10.1371/journal.pone.0125938> (2015).
37. Chae, J. J., Aksentjevich, I. & Kastner, D. L. Advances in the understanding of familial Mediterranean fever and possibilities for targeted therapy. *Br J Haematol* **146**, 467–478, <https://doi.org/10.1111/j.1365-2141.2009.07733.x> (2009).
38. Kappen, J. H. *et al.* Low prevalence of NOD2 SNPs in Behçet's disease suggests protective association in Caucasians. *Rheumatology (Oxford, England)* **48**, 1375–1377, <https://doi.org/10.1093/rheumatology/kep292> (2009).

Acknowledgements

This work was supported by Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III (PI16/01373), Fondos FEDER and Plan Andaluz de Investigación (CTS-0197).

Author Contributions

S.B.S.: Sanger-sequencing, polymorphisms genotyping and H-ARMS PCR performing, data analysis, manuscript preparation and submission. M.F.G.E.: Experimental design, data analysis, manuscript preparation. M.A.M.C.: Experimental support, data analysis, manuscript preparation. J.R.G.L.: Experimental support, manuscript preparation. I.O.M.: Experimental procedures. N.O.C.: Samples and clinical data provider. F.G.H.: Samples and clinical data provider. G.E.: Samples and clinical data provider. G.G.G.: Samples and clinical data provider. J.S.B.: Samples and clinical data provider. M.J.: Samples and clinical data provider. R.S.: Samples and clinical data provider. R.B.: Samples and clinical data provider. A.B.M.: Samples and clinical data provider. R.G.T.: Samples and clinical data provider. P.F.: Samples and clinical data provider. M.R.C.: Samples and clinical data provider. L.R.R.: Samples and clinical data provider. T.C.: Samples and clinical data provider. S.C.: Samples and clinical data provider. J.A.S.: Samples and clinical data provider. J.M.: Samples and clinical data provider.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-39113-5>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019