RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL ARTICLE

Non-classical human leucocyte antigens in ankylosing spondylitis: possible association with HLA-E and HLA-F

Margarida Rodrigues Santos,[1] Ana Rita Couto,[1] Iris Foroni,[1] Bruno Filipe Bettencourt,[1] Zhixiu Li,[2] Raquel Meneses,[1] Lawrie Wheeler,[2] Joaquim Pereira,[1] Fernando Pimentel-Santos,[3] João Eurico Fonseca,[4] Helena Alves,[5] António Martinho,[6] Manuela Lima,[7] Matthew A Brown,[2] Jácome Bruges-Armas[1,3]

To cite: Santos MR, Couto AR, Foroni I, *et al*. Non-classical human leucocyte antigens in ankylosing spondylitis: possible association with HLA-E and HLA-F. *RMD Open* Published Online First: [please include Day Month Year]. doi:10.1136/rmdopen-2018-000677

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/rmdopen-2018-000677).

Received 2 March 2018
Revised 18 May 2018
Accepted 10 June 2018

For numbered affiliations see end of article.

Correspondence to
Margarida Rodrigues Santos; margarida.mp.santos@azores.gov.pt

ABSTRACT

Objectives Ankylosing spondylitis (AS) is the most prevalent form of spondyloarthritis, with a known genetic association with the HLA-B27 molecule. The aim of this study was to assess the contribution of the HLA-G, HLA-E and HLA-F to AS susceptibility/protection in Portuguese patients with HLA-B27 AS and HLA-B27 unaffected controls.

Methods High-resolution typing of *HLA-G*, *HLA-E* and *HLA-F* was performed in 228 patients with HLA-B27 AS and 244 HLA-B27 unaffected controls. Allelic, genotypic and haplotypic frequencies were compared between cohorts. To replicate the results, single nucleotide polymorphisms (SNPs) in *HLA-E* and *HLA-F* genes were typed in Australian cohorts. For further confirmation, a group of European-descent patients with AS and unaffected controls were genotyped for Major Histocompatibility Complex SNPs using the Illumina microarray.

Results In the Portuguese population, no significant differences were found in *HLA-G*. For HLA-E, a significant difference was detected for the genotype *HLA-E*01:01:01/01:03:01* (p=0.009; pc=0.009; OR=0.51), with a protection effect. For HLA-F, significant differences were detected in the allele *HLA-F*01:01:02* (p=0.0049; pc=0.0098; OR=0.60) and corresponding SNP rs2075682 (p=0.0004; pc=0.0008; OR=0.53), suggesting protection and in the genotype *HLA-F*01:01:01/01:03:01* (p=0.011; pc=0.043; OR=2.00), suggesting a susceptibility effect. Three *G-E-F* haplotypes with significant differences were detected but occur in a very small number of individuals. The only significant differences detected in the replication studies were for *HLA-E* rs1059510 in the Australians and for *HLA-F* rs1736924 in the European-descent cohorts.

Conclusion Our results reveal suggestive AS protective and susceptibility effects from both *HLA-E* and *HLA-F* loci, however with population differences. To our knowledge, this is the first study showing association of *HLA-F* with AS.

Key messages

What is already known about this subject?

► Ankylosing spondylitis (AS) is strongly associated with HLA-B27, with only 20% of the overall genetic risk. Previous studies raised the possibility that AS susceptibility could be related to other MHC genes.

What does this study add?

► To our knowledge, this is the first study showing potential association of HLA-F with AS. AS protective and susceptibility effects from both HLA-E and HLA-F loci were detected, although with population differences.

How might this impact on clinical practice?

► The better knowledge of AS genetics will help to unveil the molecular mechanisms of the disease and contribute to the design of novel therapeutics for disease treatment.

INTRODUCTION

Spondyloarthropathies (MIM: 106300; SpAs) are the second most common group of chronic inflammatory rheumatic disorders among the adult Caucasian population.[1–3] Ankylosing spondylitis (AS) is the most prevalent form of SpA, affecting 0.3%–0.5% of Europeans, with a worldwide prevalence of 0.1%–1.4%.[3 4]

Despite the poor understanding of this pathology, it has long been known that susceptibility to AS is strongly associated with the Major Histocompatibility Complex (MHC) and in particular with the HLA-B27 molecule. HLA-B27 confers the greatest known AS risk as it is found in over 90% of patients with AS of European ancestry, but only ≈8% of healthy controls.[5] However, studies indicate that *HLA-B27* is only responsible for ≈20% of the overall genetic risk, suggesting a contribution of other genes to disease susceptibility.[4 6] Previous studies raised the possibility that AS susceptibility could be related to other MHC genes.[7–9]

The MHC is a large cluster of 128 genes and 96 pseudogenes, located on the short arm of chromosome 6, many of which with important roles in the immune system. Traditionally, the MHC is divided into classes containing groups of genes with related functions; the MHC class I and II genes encode for human leucocyte antigens (HLA), proteins that are displayed on the cell surface. In humans, MHC class I molecules comprise the classical (class Ia) HLA-A, HLA-B and HLA-C, and the non-classical (class Ib) HLA-E, HLA-F, HLA-G and HLA-H (HFE) molecules.[10] Both categories are similar in their mechanisms of peptide binding and presentation and in the induced T-cell responses.[11] The most remarkable feature of MHC class Ib molecules is their highly conserved nature[12] exhibiting very low levels of allelic polymorphism.[13] Understanding of the different roles played by class Ib molecules is rapidly increasing, with known roles including pathogen recognition, virus-induced immunopathology, tumour immune surveillance and regulation of autoimmunity.[14]

HLA-G biological features include restricted tissue expression, the presence of membrane-bound and soluble isoforms, generated by alternative splicing, limited protein variability, unique molecular structure, with a reduced cytoplasmic tail and modulation of the immune response.[15] HLA-E is the best-characterised MHC class Ib molecule. With a low level of polymorphism,[10] HLA-E molecules bind peptides derived from the signal sequences of classical MHC I molecules and present these to NK cells bearing CD94/NKG2 receptors.[16] However, it has become apparent that HLA-E molecules can bind a wider selection of peptides than canonical MHC I leader sequence-derived peptides and that these peptides can be presented to CD8 T cells.[17–20]

The *HLA-F* gene was first identified in 1990[21 22] being, so far, the least characterised non-classical class I molecule[23 24] and neither its native structure nor function is well known.[25] Several studies confirmed HLA-F protein expression in a number of diverse tissues and cell lines, including bladder, skin and liver cell lines, but no surface expression was detected in the majority of them, except for activated lymphocytes.[24] Just like other class Ib molecules, HLA-F restrictive tissue expression suggests specialised functions and tight transcriptional control of the gene[26] with unique potential regulator motifs.[21]

The aim of this study was to assess the possible contribution of non-classical *HLA-G, HLA-E* and *HLA-F* to AS susceptibility/protection in cohorts of patients with HLA-B27-positive AS and unaffected controls, independently of the HLA-B27 effect.

METHODS

Subjects

Three sets of patients with HLA-B27-positive AS and unaffected controls were used in this study. The Portuguese set was composed by individuals of Portuguese ancestry, recruited from the Azores and mainland Portugal. Patients with AS (n=228) were diagnosed using the modified New York Criteria.[27] The control group included 244 HLA-B27-positive unaffected subjects older than 35 years. One hundred individuals of each group were randomly selected for typing HLA-G 3′ UTR 14 bp indel. In order to replicate the results obtained in the Portuguese population, a group of single nucleotide polymorphisms (SNPs), in *HLA-E* and *HLA-F* genes, were genotyped in 222 Australian patients with HLA-B27-positive AS and 618 HLA-B27-positive unaffected controls. The third set were previously reported European-descent patients with AS and unaffected controls genotyped for MHC SNPs using the Illumina microarray.[7] The control group frequencies were in Hardy-Weinberg equilibrium (HWE) for all the studied SNPs.

Genotyping

Genomic DNA was extracted from peripheral blood cells according to standard procedures. The amplification of specific fragments of *HLA-G*, *HLA-E* and *HLA-F* was performed with optimised protocols. For the Portuguese cohorts, a published protocol was used for *HLA-G* typing.[28] For the amplification of *HLA-G* 3′ UTR 14 bp indel fragments, primers were designed using Primer3 software.[29] PCR primers for *HLA-E* and *HLA-F* were designed, covering all the already known polymorphic regions in either exonic or intronic regions, with Primer3 software.[29] Primer design for *HLA-E* and *HLA-F* fragments was based on sequences from IMGT/HLA Database (http://www.ebi.ac.uk/ipd/imgt/hla/align.html). Primer sequences and PCR conditions are available on request.

Sequencing was performed using an ABI 3130xl Genetic Analyzer with Big Dye Terminator V.1.1 and V.3.1 (Applied Biosystems). *HLA-E* and *HLA-G* genotype assignment was performed using Assign SBT 3.5+ software (Conexio Genomics, Fremantle, Australia). *HLA-F* allele assignment was performed manually, using the library compiled from the current ImMunoGeneTics database.[30 31] *HLA-G* 3′ UTR 14 bp indel genotyping was performed by fragment analysis method using an ABI Prism 310 Genetic Analyzer (Applied Biosystems) with 310 GeneScan V.3.1.2 software. Allele designations are according to the WHO Nomenclature Committee.[30]

Australian cohorts were typed using two different approaches; *HLA-E* and *HLA-F* SNPs rs2075682, rs17875379, rs2076183, rs1059510 and rs1264457 were typed with the previously described sequence-based typing protocol. The remaining SNPs were genotyped using Taqman SNP genotyping assays (Applied Biosystems, Foster City, California, USA) performed according to the manufacturer's instructions. Genotyping reactions were carried out with an ABI 7500 Fast Real-Time PCR System. Allele calls were obtained by the AB software V.2.0.5, by the analysis of allelic discrimination plots.

Statistical analysis

Arlequin software V.3.5.1.3[32] was used to test for Hardy-Weinberg equilibrium, perform haplotype inference and measure linkage disequilibrium (LD) (D′ and r^2). LD measuring for SNPs was performed using SNPClip from LDlink (https://analysistools.nci.nih.gov/LDlink/). Genotype and allele frequencies were directly enumerated. Inferred haplotypes were assessed for missingness. Haplotypes with >10% missingness were excluded. SNP association analysis was performed using Cochrane-Armitage test of trend as implemented in PLINK V.1.07 (http://pngu.mgh.harvard.edu/purcell/plink/). The Bonferroni correction was obtained using the formula ($p_{corrected}=1-(1\ p)^n$), where the number (n), used for correction, was the number of comparisons one or more of which shows a significant result (http://www.dorak.info/hla/stat.html). The correction procedure for SNPs included the removal of SNPs in LD ($r^2\approx1$), where the number (n) to use for correction was the remaining number of SNPs with significant differences between groups. ORs with a 95% CI were calculated for the minor alleles of the genotyped SNPs. OR >1 indicates a susceptibility allele and OR <1 indicates a protective allele. Haplotype association analysis was performed using Fisher's exact test.

Genotypes from an Immunochip dataset were used to impute *HLA-B27* using SNP2HLA,[33] and *HLA-F* using 1000G reference haplotypes as previously described.[7]

Preventive fraction was calculated using the following formula: [A/(A+B)]−[C/(C+D)], where A is the number of individuals positive for the SNP in patient cohort, B is the number of individuals positive for the SNP in control cohort, C is the number of individuals negative for the SNP in patient cohort and D is the number of individuals negative for the SNP in control cohort.

RESULTS

HLA-G

The *HLA-G* analysis was based on exons 2, 3 and 4 polymorphisms, allowing a four-digit high-resolution typing. The allelic and genotypic frequencies of patients with HLA-B27 AS and unaffected controls are shown in table 1. A total of five different alleles were identified; the most representative allele, with a frequency of 82% in both patients with AS and unaffected controls, was

Table 1 *HLA-G* allelic and genotypic frequencies in patients with HLA-B27 AS and HLA-B27 unaffected controls

		HLA-B27 Portuguese patients with AS	HLA-B27 Portuguese unaffected controls	
Alleles		**FAs**	**Fc**	**P values**
	*G*01:01*	0.82	0.82	1.00
	*G*01:03*	0.05	0.06	0.33
	*G*01:04*	0.07	0.08	0.90
	*G*01:05N*	0.01	0.00	0.36
	*G*01:06*	0.05	0.04	0.27
Genotypes		**FAs**	**Fc**	**P values**
	*G*01:01/G*01:01*	0.68	0.66	0.77
	*G*01:01/G*01:03*	0.07	0.13	0.07
	*G*01:01/G*01:04*	0.11	0.13	0.58
	*G*01:01/G*01:05N*	0.00	0.00	0.50
	*G*01:01/G*01:06*	0.10	0.06	0.17
	*G*01:03/G*01:03*	0.00	0.00	0.48
	*G*01:03/G*01:04*	0.01	0.00	0.11
	*G*01:04/G*01:04*	0.00	0.00	0.50
	*G*01:04/G*01:06*	0.01	0.01	0.68
	*G*01:05 N / G*01:05N*	0.00	0.00	0.48
	Alleles	**FAs**	**Fc**	**P values**
14 bp Indel	+14 bp (In)	0.34	0.38	0.46
	−14 bp (Del)	0.66	0.62	0.46
Genotypes		**FAs**	**Fc**	**P values**
	Homoz −14 bp (Del) (254)	0.39	0.41	0.89
	Homoz +14 bp (In) (268)	0.15	0.09	0.27
	Heteroz	0.46	0.50	0.67

AS, ankylosing spondylitis; FAs, frequency in patient group; Fc, frequency in control group; P values, Fisher's exact test.

*G*01:01*. Ten different genotypes were identified. The most prevalent genotype was clearly *G*01:01/G*01:01*. In this locus, no significant differences were observed between patients with AS and unaffected controls.

HLA-E

The *HLA-E* analysis was based on the known intronic and exonic polymorphisms, allowing six-digit high-resolution typing. Only three alleles were identified; the most prevalent, in patients with AS and unaffected controls, was *E*01:01:01*, followed by *E*01:03:02*. Six genotypes were identified, being the most prevalent *E*01:01:01/E*01:01:01* and *E*01:01:01/E*01:03:02* (table 2).

No significant differences were detected in the *HLA-E* alleles between patients with AS and unaffected controls. However, a significant difference between groups was detected for the genotype *E*01:01:01/E*01:03:01*, suggesting a protective effect for AS. This genotype had a frequency of 0.12 in patients with AS, almost doubling in unaffected controls (0.21) with a preventive fraction of 16%. Significant differences were also detected in rs1059510, where carriage of the genotype TT more than doubled in unaffected controls; however, it was not significant after Bonferroni correction, and rs1264457 where genotype AG is augmented in the control group suggesting a protective effect.

HLA-F

The *HLA-F* analysis was based on 18 SNP differences, in both intronic and exonic areas, allowing a six-digit high-resolution typing. Fourteen presented significant differences between patients with AS and unaffected controls and seven maintained significance after the Bonferroni correction (table 3).

Five alleles were identified; the most prevalent allele was *F*01:01:01*. Fourteen genotypes were identified and *F*01:01:01/F*01:01:01* was the most prevalent. Significant differences were detected in the alleles *F*01:01:02* and *F*01:03:01*, respectively. The first allele is increased in unaffected controls with a protective fraction of 12%, suggesting a putative protective effect and the second is increased in patients with AS with an attributable risk of 10%. The significance of allele *F*01:03:01* was not maintained after Bonferroni correction (table 3).

Four genotypes with significant differences were detected: *F*01:01:01/F*01:03:01* and *F*01:01:03/F*01:01:03* with 17% and 33% of attributable risk, respectively, which suggests a susceptibility effect. On the other hand, a protective effect is suggested by two other genotypes *F*01:01:02/F*01:01:03* and *F*01:01:02/F*01:03:01* with preventive fractions of 29% and 25%, respectively. After Bonferroni correction, genotype *F*01:01:01/F*01:03:01* was the only that maintained the significance.

Significant differences were detected in variants rs2075682 and rs1736924, but only the first maintained significance after Bonferroni correction (table 3).

Haplotypes

A total of 28 different *HLA-G*, *HLA-E* and *HLA-F* haplotypes were inferred for patients with AS and 26 for controls. The most frequent haplotypes, displaying the same frequency in patients with AS and unaffected controls, were *G*01:01/E*01:01:01/F*01:01:01* (25%), *G*01:01:01/E*01:03:02/F*01:01:01* (20%) and *G*01:01/E*01:01:01/F*01:01:02* (11%). In addition, three low-frequency extended haplotypes were inferred. Haplotype *G*01:03/E*01:01:01/F*01:01:02* (p=0.0003; pc=0.0008), with frequencies 0.012 and 0.053 in patients with AS and unaffected controls, respectively, integrates the allele *F*01:01:02*. On the other hand, the haplotype *G*01:01/E*01:03:01/F*01:03:01* (p=0.0002; pc=0.0005) with frequency of 0.043 and 0.004 in patients with AS and unaffected controls, respectively, integrates the allele *F*01:03:01*. Another haplotype, *G*01:01/E*01:03:01/F*01:01:01* (p=0.000003; pc=0.00001), showed significant differences between patients with AS and unaffected controls (0.015 and 0.074, respectively); although significant, these haplotypes were present in only a very small number of individuals.

Replication study

In order to replicate the obtained results, 10 SNPs (two in *HLA-E* and eight in *HLA-F* loci) were typed in cohorts of Australian patients with HLA-B27-positive AS and unaffected controls. The results are displayed in tables 2 and 3. Significant differences were detected in HLA-E, for rs1059510 genotypes CC and CT.

For further confirmation, an Immunochip-typed dataset with 4466 patients with AS of European ancestry (Australian and UK samples) and 9753 UK unaffected controls was studied, with *HLA-B*27* status imputed using SNP2HLA, and genotypes in the *HLA* region imputed using 1000G reference haplotypes. The results of the Immunochip dataset imputation for 18 SNPs (15 in *HLA-F* and 3 in *HLA-E*) showed four SNPs with significant differences (rs1736924, rs1632953, rs1736926 and rs1736925). However, as these were in LD, three of them were removed (rs1632953, rs1736926 and rs1736925) and only the SNP rs1736924 was maintained for further analysis (table 4). Association was performed by logistical regression using PLINK V.1.90b3.36. The p value of conditioning on *HLA-B*27* obtained for rs1736924 is in the same order of magnitude of the one obtained for the Portuguese population (table 4).

DISCUSSION

The *HLA-G*, *HLA-E* and *HLA-F* genes encode for molecules involved in regulation of autoimmune disease.[34]

The presence of HLA-G molecules in both membrane-bound and soluble forms has been associated with tolerogenic functions against innate and adaptative immune system.[35] Studies have indicated an immunoregulatory role of *HLA-G* wider than maintenance of tolerance on fetal–placental interface, describing the expression

Table 2 *HLA-E* allelic and genotypic frequencies and MAF in patients with HLA-B27 AS and HLA-B27 unaffected controls

		HLA-B27 Portuguese patients with AS	HLA-B27 Portuguese unaffected controls				
	Alleles	**FAs**	**Fc**	**P values**	**pCA**	**Bonferroni**	**OR (95% CI)**
	*E*01:01:01*	0.65	0.62	0.38			1.14 (0.87 to 1.48)
	*E*01:03:01*	0.10	0.12	0.41			0.82 (0.55 to 1.24)
	*E*01:03:02*	0.25	0.26	0.76			0.95 (0.70 to 1.27)
	Genotypes						
	*E*01:01:01/E*01:01:01*	0.41	0.33	0.08			1.41 (0.97 to 2.07)
	*E*01:01:01/E*01:03:01*	0.12	0.21	0.01		0.01	0.51 (0.31 to 0.85)
	*E*01:01:01/E*01:03:02*	0.35	0.37	0.85			0.95 (0.65 to 1.39)
	*E*01:03:01/E*01:03:01*	0.00	0.00	0.48			NA
	*E*01:03:01/E*01:03:02*	0.08	0.03	0.06			2.40 (1.02 to 5.69)
	*E*01:03:02/E*01:03:02*	0.04	0.06	0.21			0.56 (0.23 to 1.35)
SNP	**Alleles**	**AS MAF**	**C MAF**	**P values**	**pCA**	**Bonferroni**	**OR (95% CI)**
rs1059510	T/C	0.24	0.26	0.10	0.08		0.90 (0.67 to 1.20)
rs1264457	G/A	0.36	0.38	0.65	0.64		0.91 (0.70 to 1.18)
SNP	**Genotypes**	**FAs**	**Fc**	**P values**	**pCA**	**Bonferroni**	**OR (95% CI)**
rs1059510	CC	0.54	0.54	1.00			0.98 (0.69 to 1.41)
	CT	0.44	0.39	0.31			1.22 (0.85 to 1.75)
	TT	0.02	0.07	0.03		0.06	0.35 (0.13 to 0.89)
rs1264457	AA	0.41	0.34	0.09			1.38 (0.96 to 2.00)
	AG	0.46	0.57	0.02		0.045	0.65 (0.45 to 0.93)
	GG	0.12	0.09	0.31			1.39 (0.78 to 2.46)
		HLA-B27 Australian patients with AS	**HLA-B27 Australian unaffected controls**				
SNP	**Alleles**	**AS MAF**	**C MAF**	**P values**	**pCA**	**Bonferroni**	**OR (95% CI)**
rs1059510	T/C	0.44	0.39	0.10	0.08		1.21 (0.97 to 1.51)
rs1264457	G/A	0.45	0.44	0.65	0.64		1.06 (0.84 to 1.32)
	Genotypes	**FAs**	**Fc**	**P values**	**pCA**	**Bonferroni**	**OR (95% CI)**
rs1059510	CC	0.27	0.36	0.02		0.03	0.65 (0.47 to 0.92)
	CT	0.58	0.49	0.02		0.04	1.45 (1.06 to 1.97)
	TT	0.15	0.15	1.00			1 (0.65 to 1.55)
rs1264457	AA	0.32	0.31	0.80			1.04 (0.75 to 1.47)
	AG	0.44	0.49	0.26			0.83 (0.61 to 1.14)
	GG	0.23	0.20	0.27			1.24 (0.85 to 1.81)

AS, ankylosing spondylitis; AS MAF, minor allele frequency in AS patient group; Bonferroni, Bonferroni correction; C MAF, minor allele frequency in control group; FAs, patient group frequency; Fc, control group frequency; NA, not applicable; P values, Fisher's exact test; pCA, Cochrane-Armitage test of trend; SNP, single nucleotide polymorphism.

of *HLA-G* in several inflammatory myopathies, atopic dermatitis and cutaneous psoriasis.[36] Taking into consideration this inflammatory disease association, one could hypothesise a putative role in AS. The presence of 14 bp sequence insertion in HLA-G 3′ UTR has been associated with significantly reduced mRNA levels and lower soluble HLA-G in healthy subjects' serum.[15] In this study, no significant differences were identified between patients with HLA-B27 AS and HLA-B27 unaffected controls.

The existence of two major *HLA-E* alleles with similar frequency in most populations, probably resulting from balancing selection, indicates that there may be a functional difference between the alleles.[12 13] They differ at only one amino acid position (107), where an arginine

Table 3 *HLA-F* allelic and genotypic frequencies in patients with HLA-B27 AS and HLA-B27 controls

		HLA-B27 Portuguese patients with AS	HLA-B27 Portuguese unaffected controls				
	Alleles	**FAs**	**Fc**	**P values**	**pCA**	**Bonferroni**	**OR (95% CI)**
	*F*01:01:01*	0.54	0.52	0.59			1.08 (0.83 to 1.41)
	*F*01:01:02*	0.14	0.21	0.005		0.01	0.60 (0.42 to 0.86)
	*F*01:01:03*	0.13	0.13	0.92			1.02 (0.69 to 1.52)
	*F*01:03:01*	0.18	0.13	0.04		0.08	1.48 (1.03 to 2.13)
	*F*01:04*	0.01	0.01	1.00			0.85 (0.19 to 3.82)
	Genotypes	**FAs**	**Fc**	**P values**		**Bonferroni**	**OR (95% CI)**
	*F*01:01:01/F*01:01:01*	0.30	0.30	1.00			1.01 (0.67 to 1.51)
	*F*01:01:01/F*01:01:02*	0.14	0.20	0.08			0.63 (0.38 to 1.05)
	*F*01:01:01/F*01:01:03*	0.15	0.14	0.79			1.12 (0.66 to 1.90)
	*F*01:01:01/F*01:03:01*	0.19	0.11	0.01		0.04	2.01 (1.17 to 3.44)
	*F*01:01:01/F*01:04*	0.005	0.004	1.00			1.14 (0.07 to 18.25)
	*F*01:01:02/F*01:01:02*	0.05	0.04	0.65			1.27 (0.51 to 3.20)
	*F*01:01:02/F*01:01:03*	0.01	0.06	0.02		0.08	0.25 (0.07 to 0.89)
	*F*01:01:02/F*01:03:01*	0.03	0.08	0.01		0.04	0.32 (0.13 to 0.82)
	*F*01:01:02/F*01:04*	0.00	0.004	1.00			0
	*F*01:01:03/F*01:01:03*	0.05	0.01	0.03		0.11	4.33 (1.19 to 15.75)
	*F*01:01:03/F*01:03:01*	0.02	0.03	0.59			0.70 (0.23 to 2.18)
	*F*01:01:03/F*01:04*	0.0000	0.004	1.00			0
	*F*01:03:01/F*01:03:01*	0.04	0.02	0.24			2.32 (0.69 to 7.81)
	*F*01:03:01/F*01:04*	0.01	0.004	0.60			2.28 (0.21 to 25.33)
SNP	**Alleles**	**AS MAF**	**C MAF**	**P values**	**pCA**	**Bonferroni**	**OR (95% CI)**
rs2075682	T/A	0.12	0.21		0.0004	0.0008	0.53 (0.37 to 0.75)
rs17875380	A/C	0.007	0.01		0.55		0.65 (0.15 to 2.72)
rs1736924	C/T	0.18	0.13		0.03	0.059	1.48 (1.03 to 2.13)
rs1736922	T/C	0.15	0.14		0.79		1.05 (0.73 to 1.52
		HLA-B27 Australian patients with AS	**HLA-B27 Australian unaffected controls**				
SNP	**Alleles**	**AS MAF**	**C MAF**	**P values**	**pCA**		**OR (95% CI)**
rs2075682	T/A	0.11	0.14	0.10	0.11		0.75 (0.54 to 1.06)
rs17875380	A/C	0.007	0.008	1.00	0.78		0.83 (0.23 to 3.04)
rs1736924	C/T	0.17	0.17	1.00	0.99		1.00 (0.75 to 1.34)
rs1736922	T/C	0.18	0.21	0.27	0.25		0.85 (0.65 to 1.13)

AS, ankylosing spondylitis; AAS MAF, minor allele frequency in AS patient group; Bonferroni, Bonferroni corrected p value; C MAF, minor allele frequency in control group; FAs, frequency in patient group; Fc, frequency in control group; P value, Fisher's exact test; pCA, Cochrane-Armitage test of trend; SNP, single nucleotide polymorphism.

E^R (*E*01:01*) is replaced by a glycine E^G (*E*01:03*).[13] Both variants are indistinguishable in their structure and peptide binding features, although HLA-E^G homozygous cells seem to express higher levels of HLA-E at the cell surface.[12 13] A previous study investigating HLA-B27 extended haplotypes in Sardinia revealed, in patients with AS, a significantly higher frequency of E^R.[37 38] On the other hand, and for the same polymorphism, a markedly increased prevalence of heterozygosity among HLA-B27-positive unaffected controls was found, suggesting a protective role of E^G in AS.[37]

The results of our investigation showed a highly significant difference in the genotype *E*01:01:01/E*01:03:01* between patients with AS and unaffected controls. This genotype is twice as frequent in HLA-B27 unaffected controls when compared with patients with HLA-B27 AS. The biological meaning of this result, suggesting a putative protective effect with a protective fraction of 16%, is still

Table 4 Summary of results obtained with Immunochip dataset imputation

Chr	*HLA-F*	A1	OR	P (condition on *HLA-B*27*)
6	rs2075682	T	0.92	0.260
6	rs17875380	A	1.22	0.229
6	rs146403415 merged into rs1736924	C	0.88	0.025
6	rs1736922	T	0.92	0.086
Chr	***HLA-E***	**A1**	**OR**	**P (condition on *HLA-B*27*)**
6	rs114942539 merged into rs1059510	T	1.059	0.155
6	rs115492845 merged into rs1264457	G	1.06	0.137
6	rs188968394 merged into rs17875370	A	NA	NA

NA, not applicable.
Chr: Chromosome

unclear since no significant differences were observed in allelic frequencies between both groups. Further studies are necessary to clarify this fact. Furthermore, in this study, the rs1264457 genotype AG is augmented in the control group, with results similar to those previously obtained by Paladini *et al*,[38] but not replicated in the Australian/British cohort. Significant differences in rs1059510 genotype TT were detected but not maintained after Bonferroni correction. In contrast, and showing population discrepancies for the same SNP, in the Australian cohorts the significant differences were detected in genotypes CC and CT, both maintained after Bonferroni correction, the first with an OR suggesting protection and the other with an OR suggesting susceptibility. Data obtained in a 2015 study also suggest that *HLA-E* polymorphisms influence rheumatoid arthritis (RA) susceptibility and modulate the clinical outcome of anti-tumour necrosis factor treatment in female RA cases.[39]

The HLA-F role in the immune system regulation has been gradually unveiled and the wide-range immunoregulatory capacity is becoming clear, despite the limited understanding of the underlying structural and biochemical properties that govern its function.[40] It is known that HLA-F is expressed at the surface of activated lymphocytes; high level of HLA-F surface expression was observed in activated B, T and NK cells, with the exception of Treg,[24] but it was generally thought that HLA-F does not present antigen and instead may function as an empty, open conformer (OC) that heterodimerises with other MHC-I OC molecules.[25 41] However, a recent study establishes that HLA-F can exist both as an OC and peptide-presenting MHC molecule, with distinct NK cell receptor binding partners.[40] Expressed during an inflammatory response as OC, both HLA-F and MHC-I OC have been implicated in a novel pathway for uptake of extracellular antigen for cross-presentation.[42] HLA-F binds most allelic forms of MHC-I OC. This physical interaction likely stabilises the otherwise unstable MHC-I OC, contributing to their facility as ligands for a specific subset of killer Ig-like receptors (KIRs).[43] The coincident upregulation of KIRs with the upregulation of the HLA-F-HLA I heavy chain complex suggests an immunoregulatory role of HLA-F in inflammatory response.[44] Both KIR3DS1[45 46] and KIR3DL2[47] were previously associated with increased susceptibility to AS. Interestingly, HLA-F OCs were established as ligands of KIR3DS1, and it was also demonstrated that cell-context-dependent expression of HLA-F may explain the widespread influence of KIR3DS1 in human diseases, including AS.[45 48] Specifically, KIR3DL2 interacts with HLA-B*27 expressed as a homodimer without peptide.[49] Considering the strong association of HLA-B*27 with AS and the capacity of KIR3DL2 to detect HLA-F and MHC-I OC expressed as a homodimer raises the possibility that KIR3DL2 recognition of HLA-B*27 may represent an aberrant encounter that does not typically occur under resting conditions, leading to immune dysregulation. This possibility may relate to the expression of KIR3DL2 on Th17 CD4 T cells and their apparent increase in responsiveness in patients with AS.[43 47] In our study, two *HLA-F* alleles and four genotypes displayed significant differences in the Portuguese cohorts. However, it should be taken into account the low frequencies of some of these genotypes. It was possible to group the SNPs with significant differences into each one of the alleles (*F*01:01:02* and *F*01:03:01*). From this group of SNPs, two with potential interest were highlighted, given that they are located in *HLA-F* gene regions that may influence in one case expression levels and in the other case molecule conformation.

Variant rs2075682, allele *F*01:01:02*, with a highly significant difference between patients with AS and unaffected controls, is positioned in the SXY module X box of *HLA-F* promoter. It is the only allele with a thymine in that position, being T the ancestral allele. It is possible that this modification, in a promoter zone that controls gene transactivation, influences *HLA-F* expression, conferring a protective effect on unaffected HLA-B27-positive individuals.

HLA-F SXY module shows strong sequence homology to those of the classical MHC class I genes; it is composed by S, X and Y boxes and is bound by RFX, CREB/ATF and NFY protein complex, implicated in CIITA activity. Besides other induction pathways, *HLA-F* is induced by Class II Transactivator (CIITA), which is in agreement with the SXY module sequence conservation.[26]

Interestingly, the other variant (rs1736924) with a significant difference in both the Portuguese and Immunochip-typed Australian/British populations, although not significant enough to survive Bonferroni correction in the Portuguese population (pc=0.06), is part of the genomic sequence that is translated into the α3 domain

of HLA-F protein. HLA α3 domain residue conservation is greater than α1 and α2, with virtually no regions of particular sequence variability. Seventy-eight per cent of the residues are totally conserved and 7% are conservatively replaced and similar percentages are found for the HLA-F.[50] In rs1736924, a serine at residue 272 of α3 is substituted by a proline. Serine has an uncharged polar side chain and proline, with a non-polar side chain unique among the standard 20 amino acids, can sometimes substitute for other small amino acids, although it does not often substitute well.[51] Taking into account previous studies, two hypotheses could be proposed to explain the possible effect of rs1736924 substitution: (1) conformational changes in the molecule could transform the flexible loop that clamps the CD8 homodimer to the α3 domain, compromising the binding of CD8, consequent HLA stabilisation and recognition by TCR.[52] (2) The presence of a proline in position 272 may influence the binding of β_2m to the α chain, stimulating the formation of HLA-F dimers. Although interesting, the statistical significance of our findings is not sufficiently strong to support our hypothesis, requiring further investigation.

Regarding the association discrepancies seen in the populations in this study, differences in the association of HLA and non-HLA genes with AS have already been reported between Portuguese and other Caucasian populations. An *ERAP1* haplotype, conferring protective effect to AS in HLA-B27-positive individuals was described, for the first time, with this same cohort.[53] An association study between *IL23* and *ERAP1* genes and AS in the Portuguese population reported that no association was established with *IL23R* rs11209032 in Portuguese or Spanish populations, which was strongly associated with AS in other Caucasian populations. In addition, the study did not demonstrate any protective effect against AS for the Arg381Gln SNP (rs11209026) in the *IL23R* gene showing deviations between the Portuguese and other Caucasian populations.[54]

This is the first study in which the non-classical *HLA-G, HLA-E* and *HLA-F* genes are meticulously investigated for possible association with AS, independently of the HLA-B27 risk contribution. The results are promising, revealing protective and susceptibility effects from both *HLA-E* and *HLA-F* loci. To our knowledge, this is the first study showing potential association of *HLA-F* with AS.

Author affiliations
[1]Serviço Especializado de Epidemiologia e Biologia Molecular, Hospital de Santo Espirito da Ilha Terceira, EPER, Angra do Heroismo, Portugal
[2]Institute of Health and Biomedical Innovation, Translational Research Institute, Queensland University of Technology, Brisbane, Australia
[3]CEDOC, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa, Portugal
[4]IMM, Universidade de Lisboa, Lisboa, Portugal
[5]Centro de Histocompatibilidade do Norte, Instituto Português do Sangue e da Transplantação, Porto, Portugal
[6]Centro de Sangue e Transplantação de Coimbra, Instituto Português do Sangue e da Transplantação, Coimbra, Portugal
[7]Faculdade de Ciências e Tecnologia, Universidade dos Açores, Ponta Delgada, Portugal

Acknowledgements We wish to acknowledge all the individuals who accepted to participate in this study. We would also like to acknowledge Dr Maria José Peixoto, Dr Bruna Parreira and Dr Vanessa Faustino for organising, optimising and sequencing some of the samples.

Contributors MRS, MAB and JB-A conceived the study. FPS, JEF and JB-A provided clinical support to patient follow-up. MRS, IF, ZL, RM, LW and JP performed experiments. BFB, MRS, ZL, LW and IF performed statistical analysis. BFB, MRS, ZL, LW and MAB interpreted data. ARC, AM, HA, ML, MAB and JB-A gave critical revision of the manuscript for important intellectual content. All authors have read and accepted the submitted version of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Comissão de Ética do Hospital de Santo Espírito da Ilha Terceira.

Provenance and peer review Not commissioned; externally peer reviewed.

Data statement No additional data are available.

REFERENCES

1. Breban M. Genetics of spondyloarthritis. *Best Pract Res Clin Rheumatol* 2006;20:593–9.
2. Brown MA, Pile KD, Kennedy LG, *et al*. A genome-wide screen for susceptibility loci in ankylosing spondylitis. *Arthritis Rheum* 1998;41:588–95.
3. Bruges-Armas J, Lima C, Peixoto MJ, *et al*. Prevalence of spondyloarthritis in Terceira, Azores: a population based study. *Ann Rheum Dis* 2002;61:551–3.
4. Thomas GP, Brown MA. Genomics of ankylosing spondylitis. *Discov Med* 2010;10:263–71.
5. Reveille JD. Major histocompatibility genes and ankylosing spondylitis. *Best Pract Res Clin Rheumatol* 2006;20:601–9.
6. Ellinghaus D, Jostins L, Spain SL, *et al*. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat Genet* 2016;48:510–8.
7. Cortes A, Pulit SL, Leo PJ, *et al*. Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. *Nat Commun* 2015;6:7146.
8. Bettencourt BF FI, Couto AR, Lima M. Beyond HLA-B*27, clinical and molecular advances in ankylosing spondylitis. *Genetics in Ankylosing Spondylitis*: InTech, 2012:105–34.
9. Kim K, Bang SY, Lee S, *et al*. An HLA-C amino-acid variant in addition to HLA-B*27 confers risk for ankylosing spondylitis in the Korean population. *Arthritis Res Ther* 2015;17:342.
10. Pietra G, Romagnani C, Moretta L, *et al*. HLA-E and HLA-E-bound peptides: recognition by subsets of NK and T cells. *Curr Pharm Des* 2009;15:3336–44.
11. Rodgers JR, Cook RG. MHC class Ib molecules bridge innate and acquired immunity. *Nat Rev Immunol* 2005;5:459–71.
12. van Hall T, Oliveira CC, Joosten SA, *et al*. The other Janus face of Qa-1 and HLA-E: diverse peptide repertoires in times of stress. *Microbes Infect* 2010;12:910–8.
13. Strong RK, Holmes MA, Li P, *et al*. HLA-E allelic variants. Correlating differential expression, peptide affinities, crystal structures, and thermal stabilities. *J Biol Chem* 2003;278:5082–90.
14. Hofstetter AR, Sullivan LC, Lukacher AE, *et al*. Diverse roles of non-diverse molecules: MHC class Ib molecules in host defense and control of autoimmunity. *Curr Opin Immunol* 2011;23:104–10.
15. Donadi EA, Castelli EC, Arnaiz-Villena A, *et al*. Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. *Cell Mol Life Sci* 2011;68:369–95.
16. Braud VM, Allan DS, O'Callaghan CA, *et al*. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature* 1998;391:795–9.
17. Hansen SG, Wu HL, Burwitz BJ, *et al*. Broadly targeted CD8+ T cell responses restricted by major histocompatibility complex E. *Science* 2016;351:714–20.
18. Ulbrecht M, Modrow S, Srivastava R, *et al*. Interaction of HLA-E with peptides and the peptide transporter in vitro: implications for its function in antigen presentation. *J Immunol* 1998;160:4375–85.

19. Lampen MH, Hassan C, Sluijter M, *et al*. Alternative peptide repertoire of HLA-E reveals a binding motif that is strikingly similar to HLA-A2. *Mol Immunol* 2013;53:126–31.
20. Lampen MH, Verweij MC, Querido B, *et al*. CD8+ T cell responses against TAP-inhibited cells are readily detected in the human population. *J Immunol* 2010;185:6508–17.
21. Geraghty DE, Wei XH, Orr HT, *et al*. Human leukocyte antigen F (HLA-F). An expressed HLA gene composed of a class I coding sequence linked to a novel transcribed repetitive element. *J Exp Med* 1990;171:1–18.
22. Ishitani A, Sageshima N, Hatake K. The involvement of HLA-E and -F in pregnancy. *J Reprod Immunol* 2006;69:101–13.
23. Boyle LH, Gillingham AK, Munro S, *et al*. Selective export of HLA-F by its cytoplasmic tail. *J Immunol* 2006;176:6464–72.
24. Lee N, Ishitani A, Geraghty DE. HLA-F is a surface marker on activated lymphocytes. *Eur J Immunol* 2010;40:2308–18.
25. Goodridge JP, Burian A, Lee N, *et al*. HLA-F complex without peptide binds to MHC class I protein in the open conformer form. *J Immunol* 2010;184:6199–208.
26. Gobin SJ, van den Elsen PJ. Transcriptional regulation of the MHC class Ib genes HLA-E, HLA-F, and HLA-G. *Hum Immunol* 2000;61:1102–7.
27. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361–8.
28. Abstracts of the 18th European histocompatability conference, 8–11 May 2004, Sofia, Bulgaria. *Genes Immun* 2004;5(Suppl 1):S1–79.
29. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000;132:365–86.
30. Robinson J, Mistry K, McWilliam H, *et al*. The IMGT/HLA database. *Nucleic Acids Res* 2011;39(Database issue):D1171–6.
31. Robinson J, Halliwell JA, McWilliam H, *et al*. The IMGT/HLA database. *Nucleic Acids Res* 2013;41(Database issue):D1222–27.
32. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 2005;1:47–50.
33. Jia X, Han B, Onengut-Gumuscu S, *et al*. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One* 2013;8:e64683.
34. Kim JJ, Hong SJ, Hong YM, *et al*. Genetic variants in the HLA-G region are associated with Kawasaki disease. *Hum Immunol* 2008;69:867–71.
35. Baricordi OR, Stignani M, Melchiorri L, *et al*. HLA-G and inflammatory diseases. *Inflamm Allergy Drug Targets* 2008;7:67–74.
36. Brenol CV, Veit TD, Chies JA, *et al*. The role of the HLA-G gene and molecule on the clinical expression of rheumatologic diseases. *Rev Bras Reumatol* 2012;52:82–91.
37. Cascino I, Paladini F, Belfiore F, *et al*. Identification of previously unrecognized predisposing factors for ankylosing spondylitis from analysis of HLA-B27 extended haplotypes in Sardinia. *Arthritis Rheum* 2007;56:2640–51.
38. Paladini F, Belfiore F, Cocco E, *et al*. HLA-E gene polymorphism associates with ankylosing spondylitis in Sardinia. *Arthritis Res Ther* 2009;11:R171.
39. Iwaszko M, Świerkot J, Kolossa K, *et al*. Polymorphisms within the human leucocyte antigen-E gene and their associations with susceptibility to rheumatoid arthritis as well as clinical outcome of anti-tumour necrosis factor therapy. *Clin Exp Immunol* 2015;182:270–7.
40. Dulberger CL, McMurtrey CP, Hölzemer A, *et al*. Human leukocyte antigen F presents peptides and regulates immunity through interactions with NK cell receptors. *Immunity* 2017;46:1018–29.
41. Wainwright SD, Biro PA, Holmes CH. HLA-F is a predominantly empty, intracellular, TAP-associated MHC class Ib protein with a restricted expression pattern. *J Immunol* 2000;164:319–28.
42. Goodridge JP, Lee N, Burian A, *et al*. HLA-F and MHC-I open conformers cooperate in a MHC-I antigen cross-presentation pathway. *J Immunol* 2013;191:1567–77.
43. Goodridge JP, Burian A, Lee N, *et al*. HLA-F and MHC class I open conformers are ligands for NK cell Ig-like receptors. *J Immunol* 2013;191:3553–62.
44. Foroni I, Couto AR, Bettencourt BF. HLA-E, HLA-F and HLA-G—the non-classical side of the MHC cluster. *HLA and Associated Important Diseases*: IntechOpen, 2014.
45. Díaz-Peña R, Vidal-Castiñeira JR, Mulero J, *et al*. Activating killer immunoglobulin-like receptors genes are associated with increased susceptibility to ankylosing spondylitis. *Clin Exp Immunol* 2015;180:201–6.
46. Zuo HN, Wang ZL, Cui DR, *et al*. Genetic variations in the KIR gene family may contribute to susceptibility to ankylosing spondylitis: a meta-analysis. *Mol Biol Rep* 2014;41:5311–9.
47. Bowness P, Ridley A, Shaw J, *et al*. Th17 cells expressing KIR3DL2+ and responsive to HLA-B27 homodimers are increased in ankylosing spondylitis. *J Immunol* 2011;186:2672–80.
48. Garcia-Beltran WF, Hölzemer A, Martrus G, *et al*. Open conformers of HLA-F are high-affinity ligands of the activating NK-cell receptor KIR3DS1. *Nat Immunol* 2016;17:1067–74.
49. Kollnberger S, Chan A, Sun MY, *et al*. Interaction of HLA-B27 homodimers with KIR3DL1 and KIR3DL2, unlike HLA-B27 heterotrimers, is independent of the sequence of bound peptide. *Eur J Immunol* 2007;37:1313–22.
50. Tysoe-Calnon VA, Grundy JE, Perkins SJ. Molecular comparisons of the beta 2-microglobulin-binding site in class I major-histocompatibility-complex alpha-chains and proteins of related sequences. *Biochem J* 1991;277(Pt 2):359–69.
51. Barnes MR. *Bioinformatics for geneticists: a bioinformatics primer for the analysis of genetic data*. 2nd edn. Chichester: Wiley, 2007.
52. Gao GF, Tormo J, Gerth UC, *et al*. Crystal structure of the complex between human CD8alpha(alpha) and HLA-A2. *Nature* 1997;387:630–4.
53. Bettencourt BF, Rocha FL, Alves H, *et al*. Protective effect of an ERAP1 haplotype in ankylosing spondylitis: investigating non-MHC genes in HLA-B27-positive individuals. *Rheumatology* 2013;52:2168–76.
54. Pimentel-Santos FM, Ligeiro D, Matos M, *et al*. Association of IL23R and ERAP1 genes with ankylosing spondylitis in a Portuguese population. *Clin Exp Rheumatol* 2009;27:800–6.