Functionally distinct *ERAP1* allotype combinations distinguish individuals with Ankylosing Spondylitis

Emma Reeves[a], Alexandra Colebatch-Bourn[b], Tim Elliott[a,c], Christopher J. Edwards[b,c,1], and Edward James[a,c,1]

[a]Cancer Sciences Unit, Faculty of Medicine and [b]National Institute for Health Research Wellcome Trust Clinical Research Facility, University Hospital Southampton National Health Service Foundation Trust, Southampton General Hospital, Southampton SO16 6YD, United Kingdom; and [c]Institute for Life Sciences, University of Southampton, Southampton SO17 1BJ, United Kingdom

Edited by Peter Cresswell, Yale University School of Medicine, New Haven, CT, and approved November 5, 2014 (received for review May 15, 2014)

For more than 40 y, expression of HLA-B27 has been strongly associated with the chronic inflammatory disease Ankylosing Spondylitis (AS); however, the mechanisms underlying this association are still unknown. Single nucleotide polymorphisms within the aminopeptidase endoplasmic reticulum aminopeptidase 1 (*ERAP1*), which is essential for trimming peptides before they are presented to T cells by major histocompatibility complex (MHC) class I molecules, have been linked with disease. We show that *ERAP1* is a highly polymorphic molecule comprising allotypes of single nucleotide polymorphisms. The prevalence of specific *ERAP1* allotypes is different between AS cases and controls. Both chromosomal copies of *ERAP1* are codominantly expressed, and analysis of allotype pairs provided clear stratification of individuals with AS versus controls. Functional analyses demonstrated that ERAP1 allotype pairs seen in AS cases were poor at generating optimal peptide ligands for binding to murine H-2K^b and -D^b and the AS-associated HLA-B*2705. We therefore provide strong evidence that polymorphic ERAP1 alters protein function predisposing an individual to AS via its influence on the antigen processing pathway.

ERAP1 | Ankylosing Spondylitis | HLA-B27 | antigen processing | antigen presentation

Ankylosing Spondylitis (AS) is a chronic inflammatory rheumatic disease that causes peripheral joint inflammation, enthesitis, and typical lesions in the spine that may lead to fusion. Many factors are thought to contribute to disease susceptibility, with the primary genetic risk factor being the expression of human leukocyte antigen (HLA)-B27, found in 95% of AS-affected individuals. This strong association between HLA-B27 and AS ($P > 10^{-200}$) has been known for 40 y; however, the precise mechanisms underlying this association are still not known. Recent genome-wide association studies (GWAS) have revealed single nucleotide polymorphisms (SNPs) in two genes associated with AS, endoplasmic reticulum aminopeptidase 1 (*ERAP1*) (*ARTS1*), and *IL-23R* (1). Subsequent studies have confirmed this association to the extent that *ERAP1* has the second strongest association after HLA-B27 ($P > 10^{-27}$) (2, 3).

The ER resident aminopeptidase, ERAP1, performs a key final step within the major histocompatibility complex class I (MHC I) antigen processing pathway, trimming N-terminal residues to generate peptides that are an optimal length for loading onto MHC I (4, 5). This trimming is critical for the generation of many antigenic epitopes in vivo and influences the generation of the antigenic peptide repertoire (6–8). Because the functions of ERAP1 and HLA-B27 intersect in the ER at the final stages of the antigen processing pathway, it is reasonable to postulate that the disease mechanism or pathogenicity arise from defects/variations in the function of molecules that lie on this pathway. Interestingly, ERAP1 is only associated with AS in individuals expressing HLA-B27 further supporting this theory (2).

The contribution of individual SNPs to AS risk has been examined in many studies (reviewed in ref. 9). However, the ability to apportion disease risk to each SNP is often difficult because of strong linkage disequilibrium between polymorphisms of the same gene. Despite this, studies have identified the combination of the major allele of rs30187 (coding for R528) and the minor allele at rs10050860 (coding for N575) being protective against AS (a fourfold reduction in AS risk) (2). In addition, haplotypes K528/D575/R725, K528/D575/E730, R528/D575/R725/E730, and K528/D575/R725/Q730 have been shown to be associated with AS (10–12). The in vitro assessment of single amino acid ERAP1 variants corresponding to the AS-associated SNPs revealed reduced trimming activity for R528, Q725, and E730 but not N575 (2, 13, 14). Additional *ERAP1* polymorphisms may contribute to function and/or disease association because fine mapping SNP analysis of *ERAP1* identified the presence of additional SNPs in AS patients (3). These types of analyses, however, do not allow for the effects of SNPs that occur in combinations/haplotypes to be determined and are of particular importance because, in a recent study, we showed that naturally occurring *ERAP1* molecules are polymorphic, existing as haplotypes, hereafter referred to as allotypes, which consist of multiple combinations of AS-associated SNPs, and that they have functional differences with different amino acid specificities (15). In addition, natural polymorphic ERAP1 molecules have been shown to alter the repertoire of peptides presented by HLA-B27 (16). Indeed, the combined effects of two AS-associated SNPs, K528R and D575N, reveals a hierarchy of activity; K528/N575 have the greatest and R528/D575 the lowest with the relative activity of K528 and R528 dependent on residue 575 (17). These findings show that the SNPs are cis acting and suggest the role of *ERAP1* SNPs in AS association works at a level more complex

Significance

The immune system performs surveillance to identify infected or cancerous cells through recognition of small protein fragments called antigenic peptides on their surface. To do this, the peptides must be cut to a specific length by an enzyme called endoplasmic reticulum aminopeptidase 1 (ERAP1). Variation in this enzyme has recently been linked to the inflammatory rheumatic disease Ankylosing Spondylitis (AS). We have found that ERAP1 is highly polymorphic in humans and that specific combinations of ERAP1 are found in people with AS. These disease-associated combinations have a reduced ability to generate peptides for presentation at the cell surface by MHC class I molecules, including HLA-B27. Understanding this finding may allow easier identification of individuals with AS and allow stratification into prognostic groups.

Author contributions: E.R., T.E., C.J.E., and E.J. designed research; E.R., A.C.-B., and E.J. performed research; E.R., A.C.-B., T.E., C.J.E., and E.J. analyzed data; and E.R., T.E., C.J.E., and E.J. wrote the paper.

Conflict of interest statement: A patent relating to the work has been filed.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. KM357870–KM357891).

[1]To whom correspondence may be addressed. Email: eddjames@soton.ac.uk or cedwards@soton.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1408882111/-/DCSupplemental.

than individual SNPs. Therefore, it is important to identify the degree of polymorphism and allotype heterogeneity within *ERAP1*, whether particular allotypes associate with AS, and the role of ERAP1 molecules in AS pathogenesis.

Here, we characterize the full-length coding sequence of individual *ERAP1* allotypes from a cohort of AS cases and non-AS controls, revealing 13 different allotypes, and propose a standardized nomenclature reflecting the highly polymorphic nature of *ERAP1*. These *ERAP1* allotypes revealed a level of disease distinction with the frequency of those identified in non-AS controls, significantly different from those observed in AS cases. Furthermore, characterization of allotype pairs revealed stratification of AS cases and control groups because no combination was shared between groups. Functional analysis of these pairs also segregated case and control groups, with those found in AS cases being predominantly dysfunctional. Importantly, ERAP1 allotype pairs from AS cases are unable to restore MHC I levels to those observed from non-AS controls, indicating an inability to generate optimal HLA-B*2705 ligands and, therefore, significantly different peptide repertoires were presented at the cell surface.

Results

ERAP1 Is a Highly Polymorphic Molecule. We have shown that ERAP1 exists as distinct allotypes within individuals, with the majority of allotypes consisting of at least two AS-associated polymorphisms (15). Given the association of *ERAP1* SNPs with AS, we therefore wanted to investigate whether particular *ERAP1* allotypes were associated with AS. To this end, we isolated the full-length coding sequence of *ERAP1* from AS cases and controls. Using molecular cloning, we sequenced *ERAP1* genes from a cohort of 17 clinically characterized cases and 19 control samples assembled from age- and sex-matched cases of noninflammatory rheumatic illnesses (osteoarthritis, osteoporosis), non-AS inflammatory conditions (rheumatoid arthritis and systemic lupus erythematosus), and healthy volunteers. Samples were tissue typed, confirming that all AS cases were HLA-B*2705 positive (Table S1). Analysis of the full-length *ERAP1* coding sequence revealed 13 distinct allotypes based on amino acid sequences. The allotypes were found to contain multiple polymorphisms, which included the five SNPs previously shown to be associated with AS (Table 1). Further investigation revealed a number of conservative nucleotide variations, which, although not changing protein sequence, further delineated *ERAP1* molecules (Table S2). Because *ERAP1* is highly polymorphic (22 different sequences: 13 allotypes identified from 36 individuals), we undertook to standardize the *ERAP1* allotype sequence nomenclature to allow unique identification of *ERAP1* allotypes. To this end, we established the nomenclature ERAP1*000:00:00, where the first group of three digits identifies *ERAP1* molecules with coding amino acid differences defining the distinct allotypes. The second group of digits denotes variation within allotypes that represent conservative nucleotide changes. The final group of digits discriminates molecules within allotypes that have variation in intronic and/or untranslated regions (5′ and 3′ UTR; not examined in this study). We applied this standardizing nomenclature to the *ERAP1* allotypes we identified from our cohort and listed the amino acid positions where variation between allotypes was most frequent (Table 1). We identified 15 differences throughout the coding sequence, comprising six previously described nonsynonymous polymorphisms at amino acid positions 127, 349, 528, 575, 725, and 730; six that have not been reported previously at positions 82, 102, 115, 581, 737, and 752; and three others at previously described positions but encoding different amino acids (F199C, L727P, and M874T). The greatest extent of amino acid variation was between allotypes *001 and *002, which have 13 differences throughout the coding sequence with other allotypes incorporated varying combinations of the 15 differences (Table 1). We also identified three allotypes with additional diversity in conservative nucleotides, the greatest being for allotype *001, where seven subtypes were identified, perhaps reflecting its high frequency in the population (Table 1 and Table S2). To further assess the relationship between *ERAP1* allotypes, we performed phylogenetic analysis of the identified nucleotide and amino acid sequences. The resultant unrooted phylogenetic trees reveal two major branches defined by differences at six positions (82, 102, 115, 199, 581, 737) (Fig. 1). We have described functional variation among *ERAP1* encoded by nine different allotypes that broadly fell into three functional groups: "normal," "hypo," and "hyper" trimmers (15). When this trimming function is superimposed on the phylogenetic tree of amino acid sequences, we found some evidence of clustering of functionally similar allotypes (Fig. 1). Intriguingly, the hyperactive *006 and *007 are closely related to the hypoactive *005 and normal *008 allotypes, only varying at one or two positions. The hyperactive allotypes contain a Q725 polymorphism, indicating that it is important in the acquisition of a hyperactive trimming phenotype.

ERAP1 Allotypes Distinguish AS Case Samples from Matched Controls. We next determined the *ERAP1* allotypes expressed in AS cases (n = 34) and controls (n = 38; Table 1). Some allotypes were found to be more prevalent in controls (*002 and *011), whereas others were more prevalent in cases (*001 and *005). Previous assessment of the trimming function of these ERAP1 molecules showed that allotype *002 trimmed peptide precursors efficiently, whereas allotype *001 was hypoactive (ref. 15; where they were referred to as WT and 5SNP, respectively). Analysis of the second most-frequent case allotype, *005, showed that the trimming function depended on the amino acid substrate and was hypoactive for some substrates but not others (ref. 15; referred to as K528R and below). Therefore, there was a possible

Table 1. Identity and frequency of ERAP1 allotypes in the populations studied

ERAP1 allotypes	Frequency		Amino acid at indicated position														
	Controls (n = 38) n (%)	Cases (n = 34) n (%)	82	102	115	127	199	349	528	575	581	725	727	730	737	752	874
*001	8 (21)	15 (44)	V	I	L	P	F	V	R	N	L	Q	L	E	V	R	M
*002	17 (44.5)	1 (3)	I	L	P	R	S	M	K	D	S	R	L	Q	A	R	V
*003	1 (2.5)	0	I	L	P	R	S	M	K	D	S	R	L	Q	A	G	V
*004	0	1 (6)	I	L	P	R	S	M	K	D	S	R	A	Q	A	R	V
*005	4 (11)	10 (29)	I	L	P	R	S	M	R	D	S	R	L	Q	A	R	V
*006	0	2 (6)	I	L	P	R	S	M	K	D	S	Q	L	E	A	R	V
*007	0	2 (6)	I	L	P	R	S	M	R	D	S	Q	L	Q	A	R	V
*008	1 (2.5)	0	I	L	P	R	S	M	R	D	S	R	L	E	A	R	V
*009	0	1 (3)	V	I	L	P	F	V	R	D	S	R	L	Q	A	R	V
*010	2 (5)	0	I	L	P	R	S	V	K	N	S	Q	L	Q	A	R	V
*011	4 (11)	0	V	I	L	R	F	M	R	D	L	R	L	E	V	R	V
*012	0	1 (3)	I	L	P	R	S	V	K	D	S	R	L	Q	A	R	V
*013	1 (2.5)	0	I	L	P	P	S	M	K	D	S	R	A	Q	A	R	V

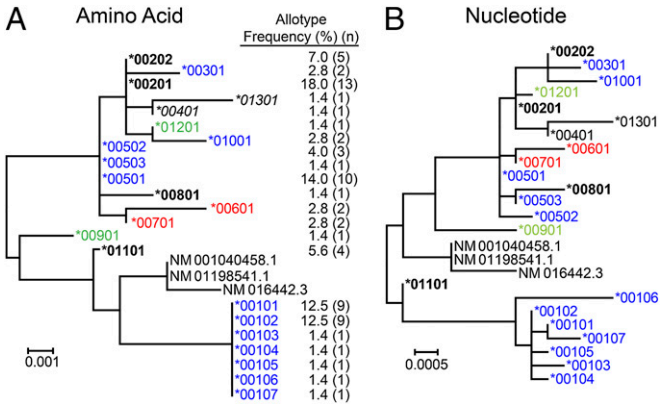

Fig. 1. Phylogenetic analysis of ERAP1 allotypes. *ERAP1* amino acid (*A*) and nucleotide (*B*) sequences were used to generate unrooted maximum likelihood phylogenetic trees. The frequency of each allotype in the amino acid tree is indicated. The overall trimming function of each allotype is also indicated; hyperactive trimmers are in red, hypoactive trimmers in blue, intermediate trimmers in green, and efficient trimmers in bold type. Allotypes in italics have not been assessed.

association between allotype and disease; however, this association was not clearly evident at the level of ERAP1 function.

Because both chromosomal copies of ERAP1 are codominantly expressed, we next determined the combinations of allotype in our AS cohort and control group. Interestingly, the majority of samples were heterozygous for *ERAP1* (32/36) and, strikingly, no allotype pair observed in cases was also seen in control samples (Table 2). For example, the *001 allotype, the most prevalent in AS cases, was not found in combination with *002 in cases, although this allotype pair was present in approximately a third (37%) of those identified in controls. Furthermore, the *002 allotype was observed in most of the controls (15/19), but in only one case (1/17) indicating that AS cases could be distinguished from controls based on their *ERAP1* allotype combination.

Importance of Combined Allotypes in the Case Cohort: AS Patient ERAP1 Allotype Pairs Reveal an Overall Reduced Trimming Function. With the *ERAP1* allotype pairs showing clear differences between AS cases and controls, we investigated whether the combined trimming functions of codominantly expressed *ERAP1* molecules were also different. We chose to measure the trimming function of ERAP1 in situ in the antigen processing pathway of living cells by using a well-characterized assay, which we have used (15), to measure function of ERAP1 allotypes and allotype pairs. An N-terminally extended precursor (AIVMK-SIINFEHL or X5-SHL8) was transfected into Erap1-deficient cells along with ERAP1 and assessed for generation of SHL8 complexed with H-2K^b MHC I by activation of SHL8-specific CD8$^+$ T cells. Trimming in Erap1-deficient cells was reduced by 90% compared with normal cells but could be restored by transfecting allotype *002 (Fig. 2*A*). ERAP1*002, containing the active site GAMEN motif mutation (E320A), was nonfunctional and was used as a negative control throughout the study (Fig. 2*A*). We reconstituted Erap1-deficient cells with pairs of allotypes corresponding to those combinations identified from individuals and confirmed equivalent expression by Western blot (Fig. S1). The ability of AS case ERAP1 combinations to generate SHL8 from X5-SHL8 was significantly reduced in most instances (Fig. 2 *B–D* and Fig. S2). This observation was in stark contrast to control allotype combinations where the predominant trimming function was similar to homozygous *002 allotypes (Fig. 2 *A*, *C*, and *D* and Fig. S2). The difference in ability to trim peptide precursor is most evident when comparing responses observed between control and AS case allotype pairs, where AS group trimming function was ~50% of that of the controls (Fig. 2*C*). Thus, discrimination between case and control populations was seen at the level of function only when the combined function of both codominantly expressed ERAP1 allotypes were analyzed. Interestingly, when *001 or *005 are paired with a *002 allotype (as occurs in some controls), the trimming function was good (Fig. 2*D*). However, when both *001 and *005 allotypes are combined (as in AS cases), the trimming function was poor (Fig. 2 *A* and *D*). The observed restoration of a normal trimming function when *001 or *005 are coexpressed with *002 is therefore consistent with a simple loss-of-function phenotype for *001 and *005 (Fig. 2*D* and Fig. S2). The majority of allotype pairs from AS cases consisted of two allotypes with poor trimming activities (Table 2, Fig. 2*D*, and Fig. S2). Where hypoactive allotypes appeared in the control group, they were typically paired with a normal functioning allotype such as the relatively frequent pairing of *001 with *002. Normal functioning allotypes that appeared in the AS case cohort were typically paired with allotypes that in combination demonstrated poor trimming capacity; for example, *002 paired with *006 (Fig. 2*D* and Fig. S2). This trimming phenotype is consistent with the *006 allotype being hyperactive and, thus, exerting a dominant negative trimming function.

Affect of ERAP1 Allotype Combinations on Peptide Repertoire and MHC I Expression. We have shown that while X5-SHL8 is an informative index substrate for broadly classifying ERAP1 function, fine substrate specificity is also observed among ERAP1 variants (15). To determine whether the observed trimming effects of X5-SHL8 was a fair representation of more global trimming function, we assessed the ability of ERAP1 pairs to restore cell surface expression of H-2D^b and -K^b in Erap1-deficient cells to normal levels; Erap1-deficient cells have a 20–30% reduction in MHC I (6), which was restored to normal levels following *002 transfection (Fig. 2*E*). We measured the ability of allotype pairs to restore MHC I expression and plotted the results as a direct comparison with the effect of the *002 transfectants. All allotype pairs found in the control group were able to restore cell surface MHC I levels (Fig. 2*E*). Conversely, most disease-associated pairs were unable to restore MHC I levels (Fig. 2*E*; >50% reduction). We noted one exception in AS cases, the combination of *003 and *012, which induced almost complete restoration. This finding suggests that ERAP1 allotypes in the majority of AS cases are unable to generate stabilizing ligands for H-2D^b/K^b.

HLA-B*2705 is the most prevalent HLA-B27 subtype associated with AS and was expressed by all AS patients in our cohort. We therefore investigated the effect of ERAP1 allotype pairs on HLA-B*2705 cell surface expression. Erap1-deficient cells were transfected with HLA-B*2705, human β_2M, and combinations of ERAP1 allotypes found in patient or control groups and the

Table 2. Identity and frequency of ERAP1 allotype combinations in the populations studied

	Frequency	
Allotype combination	Controls (n = 19), n (%)	Case (n = 17), n (%)
*001+*002	7 (37)	0
*002+*005	3 (16)	0
*002+*002	2 (11)	0
*002+*011	2 (11)	0
*010+*011	2 (11)	0
*001+*008	1 (5)	0
*002+*003	1 (5)	0
*005+*013	1 (5)	0
*001+*005	0	9 (53)
*001+*001	0	2 (12)
*001+*007	0	2 (12)
*002+*006	0	1 (6)
*004+*006	0	1 (6)
*005+*009	0	1 (6)
*003+*012	0	1 (6)

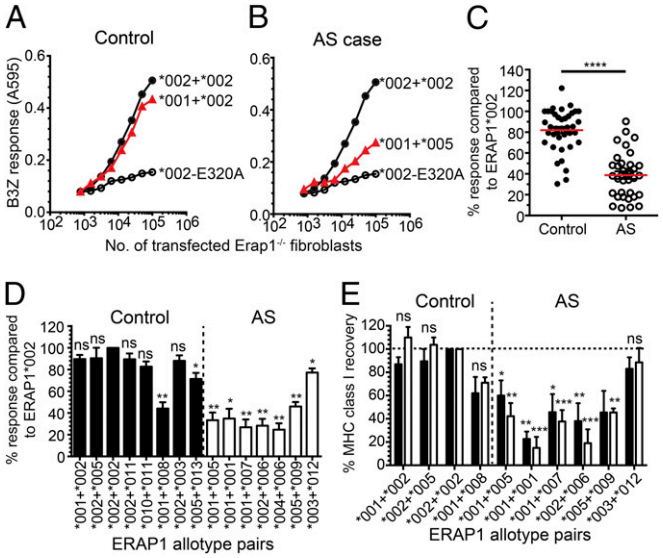

Fig. 2. ERAP1 allotype pairs isolated from AS cases have impaired trimming capacity. Erap1-deficient cells were transfected with ERAP1 allotypes corresponding to individual allotype pairs identified in cases and controls and X5-SHL8 and assessed for trimming using B3Z. (*A* and *B*) Representative line graphs showing trimming of the most common allotype pairs from controls (*A*) or cases (*B*) as indicated. The positive (*002) and negative (*002-E320A) control allotypes were also transfected. Data are representative of four experiments. (*C* and *D*) The relative maximum B3Z response of observed allotype pairs identified in control or case groups with each symbol representing a single allotype pair transfection (*C*) or individual allotype pairs compared with *002 allotype pair are shown (*D*). Data are pooled from four experimental repeats ±SEM (****$P < 0.0001$; **$P < 0.01$; *$P < 0.05$; ns, not significant). (*E*) The ability of ERAP1 allotype pairs to restore MHC I levels in Erap1-deficient cells was assessed. Erap1-deficient cells were transfected with allotype pairs as above (*002+*011, *010+*011, *002+*003, *005+*013, and *004+*006 not done), and the levels of H-2K^b (filled bars) or H-2D^b (open bars) were assessed and compared with *002. Results show data pooled from three experiments ±SEM (***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; ns, not significant). The dashed line represents 100% restoration of MHC I levels.

expression of HLA-B*2705 was measured by flow cytometry. Control ERAP1 allotype pairs showed a significant increase in HLA-B*2705 levels compared with allotype pairs found in AS cases (28% versus 2%; Fig. 3 *A* and *B*). Examination of the effect of each identified ERAP1 allotype pair on HLA-B*2705 levels revealed that most of the control ERAP1 pairs (5/7) increased cell surface expression by 20–40% (Fig. 3*C*) with the remaining two ERAP1 combinations (*002+*003 and *005+*013) increasing expression by <10% compared with the nonfunctional *002-E320A (Fig. 3*C*). By contrast, all of the ERAP1 pairs from AS individuals showed little difference in B*2705 expression compared with *002-E320A, with some even decreasing HLA-B*2705 levels (Fig. 3 *B* and *C*). These data suggest that, similar to the results obtained for H-2D^b and H-K^b, the ERAP1 pairs found in AS cases generate fewer peptides capable of stabilizing B*2705 compared with combinations found in controls.

To further investigate the effect of ERAP1 combinations on HLA-B*2705 cell surface levels, we used an ERAP1 KO 293T human cell line. This cell line was created by using the CRISPR/Cas9 system (18) to target ERAP1 and introduce a double-stranded nick, which, following repair, introduced frame-shift mutations resulting in premature stops in both copies of ERAP1. These ERAP1 KO 293T cells do not produce any detectable ERAP1 protein and fail to trim X5-SHL8 precursor when transfected (Fig. S3). ERAP1 KO 293T cells expressing HLA-B*2705 (Fig. S4) transfected with ERAP1 pairs found in controls showed a significant increase in HLA-B*2705 compared with ERAP1 pairs from AS cases (15% versus 1%; Fig. 3 *D* and *E*). Individual examination of ERAP1 pairs revealed that all those identified in controls increased HLA-B*2705 levels by 10–20% (Fig. 3*F*). By contrast, only three of the seven AS case ERAP1 pairs identified increased HLA-B*2705 cell surface expression, and even then did so with reduced efficiency (<5%), with the other combinations leading to decreased HLA-B*2705 expression (up to −5%; Fig. 3*F*). This observation further confirmed that AS case ERAP1 pairs generate fewer HLA-B*2705 stabilizing peptide ligands. It is therefore likely that the repertoire of peptides presented by HLA-B*2705 at the cell surface is significantly different between cases and controls.

Discussion

In this study, we have shown that *ERAP1* is highly polymorphic with 13 distinct allotypes assembled from at least 15 nonsynonymous nucleotide variants identified from 36 genomes. The polymorphic nature of *ERAP1* has been reported (3, 15), with both studies also identifying novel polymorphisms. Indeed, there are greater than 70 nonsynonymous polymorphisms in *ERAP1*, although the majority is found in only a small fraction of samples (dbSNP database). Our analysis of the complete coding sequence revealed a further nine polymorphic variants, three having been previously observed coding for different amino acids (199, 727, and 874; dbSNP database). Interestingly, phylogenetic analysis revealed six of the novel variants (82, 102, 115, 199, 581, 737) that formed the basis for the main branch point of *ERAP1* (Fig. 1). In almost all allotypes identified (71/72), the six variants were coinherited, forming a backbone, suggestive of an early evolutionary branching based on these variants. Many studies have shown the linkage between *ERAP1* SNPs and disease risk in different populations (reviewed in ref. 9); the strongest linkage found at residues 528 and 730 (3), observed in allotypes from

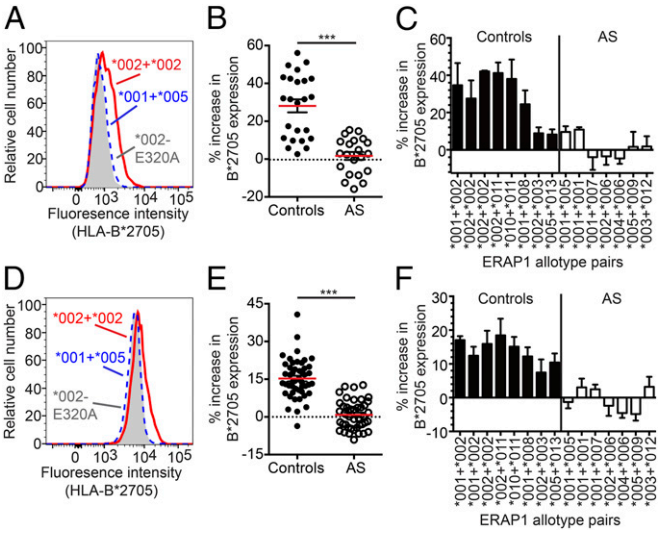

Fig. 3. AS case ERAP1 allotype pairs fail to increase HLA-B*2705 cell surface expression. Flow cytometry analysis of HLA-B*2705 cell surface expression by Erap1-deficient fibroblasts (*A*–*C*) or ERAP1 KO 293T cells (*D*–*F*) transfected with each of the 15 ERAP1 allotype pairs identified from control and AS case groups compared with *002-E320A. (*A* and *D*) Representative histograms showing HLA-B*2705 expression after transfection of Erap1-deficient fibroblasts (*A*) and ERAP1 KO 293T cells (*D*) with an example of allotype pairs from control and AS case groups. (*B* and *E*) Comparison of HLA-B*2705 cell surface expression in Erap1-deficient fibroblasts (*B*) or ERAP1 KO 293T cells (*E*) transfected with allotype pairs from control and AS case groups. Each symbol represents a single allotype pair transfection from three (*B*) or five (*E*) independent experiments; ***$P < 0.0001$. (*C* and *F*) The effect of individual ERAP1 allotype pairs from control and AS cases on HLA-B*2705 cell surface expression following transfection into Erap1-deficient fibroblasts (*C*) or ERAP1 KO 293T cells (*F*). Data are pooled from three (*C*) or five (*F*) independent experiments ±SEM.

both case and control groups. Previous studies have indicated an association between K528 and AS susceptibility, whereas we found an association with R528 (controls = 17/21 (R/K), cases = 28/6; $P = 0.002$); the reason for this difference is unclear. However, previous studies have noted an association with single SNP, whereas our study unequivocally identifies individual ERAP1 allotypes showing that the context in which SNPs appear within an allotype is important in determining disease association and ERAP1 function. Homozygosity at 528 was also associated with AS susceptibility (controls = 4/19, cases = 17/17; $P = 8 \times 10^{-6}$), consistent with a poor trimming phenotype being associated with disease (R528 shows a 30% reduction in trimming compared with K528) (15), and our observation that poor trimming R528-containing allotypes are rescued in heterozygotes when the accompanying allotype has normal function. Haplotypes linked with AS containing these polymorphisms have also been imputed from SNP analysis (refs. 10–12; K528/D575/R725, K528/D575/E730, R528/D575/R725/E730, and K528/D575/R725/Q730). We identified one of these (K528/D575/E730), or ERAP1*006 unambiguously in our full-length sequencing approach. This allotype was only present in AS cases, albeit at a low frequency (Tables 1 and 2). Despite the association of single SNPs with AS, the demonstration that *ERAP1* SNPs assemble into allotypes and that SNPs affect ERAP1 function in cis (15–17), in the context of these allotypes, shows the limited expediency of simple SNP analysis as a disease marker. Even allotype analysis was of limited value, with the highest precision being achieved only when both allotypes present in a single individual's genome were identified unambiguously. Moreover, this stratification was rationalized at the functional level because AS-associated allotype pairs encoded ERAP1 with dysfunctional trimming activity arising from the pairing of two hypofunctional allotypes or a normal function and a hyperfunctional allotype indicating that, in general, the overall trimming phenotype of two allotypes is summative which may be achieved through independent activity either in competition or in sequence, or via modulation resulting from a physical interaction between the two ERAPs, similar to the alteration of function observed when ERAP1 heterodimerizes with ERAP2 (19). Interestingly, one allotype combination (*003+*012) showed restoration of H-2K^b/D^b, but not B27 expression (Figs. 2 and 3) which may reflect a combination where the substrate specificity of each allotype is such that the peptidome thus generated contains abundant epitopes for H-2 but not HLA-B27. In support of this idea, the trimming phenotype of *012 is substrate specific and similar to that observed for X-SHL8 substrates (*012 = MV in ref. 15), indicating an overtrimming phenotype for B27.

The mechanism by which differences in ERAP1 primary structure contribute to the differences in function we observe is not clear. Four of the six (82, 102, 115, and 199) together with residue 127 are located in domain I away from the active site. Interestingly, the end of the S1 specificity pocket in domain II borders residues 181 and 183 in domain I and, therefore, the observed polymorphisms may affect the formation of the catalytic site (14, 20). In addition, residue 127 may affect conformational transition from open to closed states as proposed (21). Similarly, the AS-associated residue 528 is likely to affect the ability to adopt a correct catalytic conformation because it lies in a region of the molecule that has been proposed to articulate the conformational change. Residue 581 is situated in a β-strand in domain III and, similarly to residue 575 (closely located as part of a loop), may affect flexibility of domain III (17). Residue 349 is close to the active site and, therefore, may affect trimming. By contrast, residues 727 and 737 form part of an α-helix also containing the AS-associated residues 725 and 730 in domain IV. These residues are located within the substrate binding cavity, which may interact with the C terminus of peptide substrates as part of the "regulatory" domain and alter the binding and/or trimming specificity of ERAP1.

Although it is not known why ERAP1 is so polymorphic, the identification of an ERAP1 trimming-resistant HIV gag epitope (22) and targeting of ERAP1 by human cytomegalovirus (23) indicates selective pressure from infectious agents/pathogens similar to, but to a lesser extent than, that observed for HLA (MHC). It will be interesting to observe whether particular ERAP1 allotypes are associated with better protection to some pathogens. One consequence of increased genetic diversity in ERAP1 could be that the evolution of allotypes that confer better protection to a particular pathogen may, when expressed in individuals of particular HLA types such as B*2705 and B*5701, predispose these individuals to autoimmune disease (1, 24).

How allotypic variation of ERAP1 function could contribute to AS pathogenesis remains a matter for speculation, but our study strongly supports its impact on the biochemistry and antigen presentation of HLA-B27 via changes in the peptidome. Based on our findings, we propose a model that links the relative activity of ERAP1 variants to disease via their effect on biochemical features of HLA-B27 previously implicated in AS (Fig. 4). HLA-B27 has a propensity to form heavy chain homodimers ($B27_2$) in the ER and nonclassical forms, including $B27_2$, at the cell surface as a result of limited peptide supply, impaired peptide selection, or high B27 expression (25, 26). $B27_2$ have thus been implicated in the pathogenesis of AS through two different mechanisms: either the induction of the unfolded protein response (UPR) in the ER (27), or activation of innate and/or Th17 cells through KIR3DL2 engagement at the cell surface (28). Our data unify these mechanisms based on an understanding of ERAP1 function. The inability of ERAP1 variants with high (hyper) or low (hypo) trimming activity may lead to the restricted supply of optimal peptides (Figs. 2 and 3). Differences in trimming properties may predispose an individual to the formation of ER $B27_2$ and UPR. In addition, the generation of suboptimal HLA-B27 ligands created by aberrant ERAP1 activity that bind with sufficient affinity to pass intracellular quality control (16, 29) may dissociate rapidly at the cell surface, leading to increased aberrant forms of B27. These mechanisms are not necessarily mutually exclusive, nor do they preclude other possible mechanisms such as the ability of different ERAP1 variants to generate specific arthritogenic peptides.

Finally, the ERAP1 homolog ERAP2 has been linked with AS and a change in trimming function (30, 31). Despite there not being a mouse ortholog of ERAP2, we found little difference in the phenotype of HLA-B27 between murine Erap1$^{-/-}$ cells and human ERAP1 KO 293T cells, suggesting that any effect ERAP2 has on peptide generation is small, and supporting the idea that ERAP1 is the main component of peptide trimming in cells (30). Nevertheless, it remains to be determined whether ERAP2

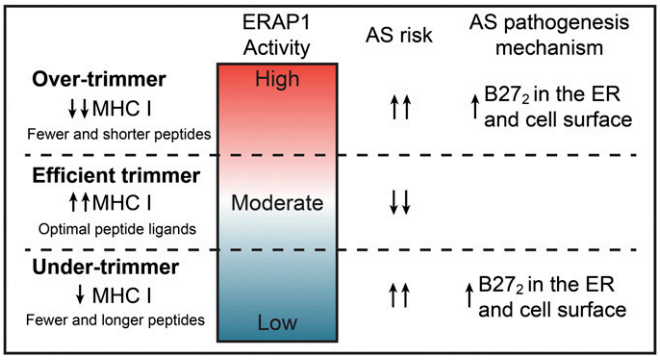

Fig. 4. Model representing the link between the ERAP1 trimming activity of an allotype pair and disease. ERAP1 allotype pairs from individuals have a broad spectrum of trimming activity. Those with trimming activities toward the extreme ends of this spectrum have a greater risk of developing AS. This increased risk is manifested in two different ways: (*i*) Aberrant ERAP1 activity results in increased misfolded and HLA-B*2705 homodimers in the ER inducing the unfolded protein response. (*ii*) Aberrant ERAP1 activity generates unstable peptide ligands and results in increased cell surface HLA-B*2705 homodimers activating NK and/or Th17 cells.

molecules in 293T cells are a low activity variant, which can be verified by using the ERAP1 KO 293T cells.

In conclusion, this study provides a framework for understanding how ERAP1 function could impact on disease pathogenesis and how the distinct allotype combinations in AS cases may serve as biomarkers for disease stratification and a target for treatment.

Materials and Methods

AS Case and Control Samples. All samples were obtained from patients and controls recruited with informed consent according to the protocol approved by the National Research Ethics Services and Southampton Research Ethics Committee. Diagnosis of AS was confirmed by using the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axial spondyloarthritis and the modified New York criteria (Table S1).

ERAP1 Isolation. *ERAP1* was amplified from cDNA as described (15). The PCR amplicon was cloned into vectors pcDNA3.1, pcDNA3.1V5/His (Life Technologies), or a modified pcDNA3.1HA and sequenced to identify individual allotypes. *ERAP1* allotype sequences were aligned, and phylogenetic analysis were performed by using MEGA 5 software.

Cell Lines, T-Cell Activation, and MHC I Recovery Assays. Erap1-deficient cells were transfected with both ERAP1 allotypes, recreating the combinations present within individuals and ES-AIVMK-SHL8 (X5-SHL8) minigene construct (15) by using FuGENE 6 (Promega). Presentation of trimmed SHL8 and activation of B3Z T-cell hybridoma was assessed as described (15). For murine MHC I recovery, 48 h after transfection Erap1-deficient cells were stained for H-2K^b (Y3) and H-2D^b (B22.249). For HLA-B*2705 expression, either Erap1-deficient or ERAP1 KO 293T-B27 (generated by using the CRISPR/Cas9 system; ref. 18) expressing cells were transfected with ERAP1 allotype pairs as above and stained for HLA-B*2705 expression (ME1). Cells were analyzed by flow cytometry with a BD FACS Canto II. The percentage of MHC class I recovery was calculated thus: [mean fluorescence intensity (MFI) of ERAP1 combination – MFI *002-E320A]/(MFI *002 – MFI *002-E320A) × 100. Percent change in HLA-B*2705 expression by ERAP1 allotypes was calculated relative to cells transfected with the nonfunctional *002-E320A (*SI Materials and Methods* and Fig. S4).

Immunoblots. Expression of ERAP1 was determined by immunoblot as described (15). Immunoblots were probed with anti-human ARTS1 (R&D Systems), anti-V5 (Life technologies), anti-HA (Abcam), or anti-glyceraldehyde 3-phosphate dehydrogenase, -GAPDH (Abcam) antibodies followed by HRP-conjugated secondary antibody and SuperSignal West Pico or Femto chemiluminescent substrate (Thermo Scientific). ERAP1 protein expression was quantified from immunoblots by using ImageJ software and compared with GAPDH.

Statistical Analysis. One-way ANOVA with Dunnett's post hoc test was performed for analysis of differences between multiple groups and control (GraphPad prism).

ACKNOWLEDGMENTS. We acknowledge the technical expertise of Nasia Kontouli, clinical support from Helen Platten, and ethics submission support from Pamela Freeman. This study was supported by the Southampton National Institute for Health Research Wellcome Trust Clinical Research Facility. This work was supported by Cancer Research UK Project Grant C25722/A8410 (to E.J. and T.E.), Cancer Research UK Programme Grant C7056/A11946 (to T.E.), and a Medical Research Council Studentship (to E.R.).

1. Burton PR, et al.; Wellcome Trust Case Control Consortium; Australo-Anglo-American Spondylitis Consortium (TASC); Biologics in RA Genetics and Genomics Study Syndicate (BRAGGS) Steering Committee; Breast Cancer Susceptibility Collaboration (UK) (2007) Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 39(11):1329–1337.
2. Evans DM, et al.; Spondyloarthritis Research Consortium of Canada (SPARCC); Australo-Anglo-American Spondyloarthritis Consortium (TASC); Wellcome Trust Case Control Consortium 2 (WTCCC2) (2011) Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 43(8):761–767.
3. Harvey D, et al. (2009) Investigating the genetic association between ERAP1 and ankylosing spondylitis. *Hum Mol Genet* 18(21):4204–4212.
4. Serwold T, Gonzalez F, Kim J, Jacob R, Shastri N (2002) ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature* 419(6906):480–483.
5. York IA, et al. (2002) The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. *Nat Immunol* 3(12):1177–1184.
6. Hammer GE, Gonzalez F, Champsaur M, Cado D, Shastri N (2006) The aminopeptidase ERAAP shapes the peptide repertoire displayed by major histocompatibility complex class I molecules. *Nat Immunol* 7(1):103–112.
7. Yan J, et al. (2006) In vivo role of ER-associated peptidase activity in tailoring peptides for presentation by MHC class Ia and class Ib molecules. *J Exp Med* 203(3):647–659.
8. York IA, Brehm MA, Zendzian S, Towne CF, Rock KL (2006) Endoplasmic reticulum aminopeptidase 1 (ERAP1) trims MHC class I-presented peptides in vivo and plays an important role in immunodominance. *Proc Natl Acad Sci USA* 103(24):9202–9207.
9. Keidel S, Chen L, Pointon J, Wordsworth P (2013) ERAP1 and ankylosing spondylitis. *Curr Opin Immunol* 25(1):97–102.
10. Choi CB, et al. (2010) ARTS1 polymorphisms are associated with ankylosing spondylitis in Koreans. *Ann Rheum Dis* 69(3):582–584.
11. Kadi A, et al. (2013) Investigating the genetic association between ERAP1 and spondyloarthritis. *Ann Rheum Dis* 72(4):608–613.
12. Maksymowych WP, et al. (2009) Association of a specific ERAP1/ARTS1 haplotype with disease susceptibility in ankylosing spondylitis. *Arthritis Rheum* 60(5):1317–1323.
13. Evnouchidou I, et al. (2011) Cutting Edge: Coding single nucleotide polymorphisms of endoplasmic reticulum aminopeptidase 1 can affect antigenic peptide generation in vitro by influencing basic enzymatic properties of the enzyme. *J Immunol* 186(4):1909–1913.
14. Kochan G, et al. (2011) Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc Natl Acad Sci USA* 108(19):7745–7750.
15. Reeves E, Edwards CJ, Elliott T, James E (2013) Naturally occurring ERAP1 haplotypes encode functionally distinct alleles with fine substrate specificity. *J Immunol* 191(1):35–43.
16. García-Medel N, et al. (2012) Functional interaction of the ankylosing spondylitis-associated endoplasmic reticulum aminopeptidase 1 polymorphism and HLA-B27 in vivo. *Mol Cell Proteomics* 11(11):1416–1429.
17. Martín-Esteban A, Gómez-Molina P, Sanz-Bravo A, López de Castro JA (2014) Combined effects of ankylosing spondylitis-associated ERAP1 polymorphisms outside the catalytic and peptide-binding sites on the processing of natural HLA-B27 ligands. *J Biol Chem* 289(7):3978–3990.
18. Cong L, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–823.
19. Saveanu L, et al. (2005) Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nat Immunol* 6(7):689–697.
20. Nguyen TT, et al. (2011) Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1. *Nat Struct Mol Biol* 18(5):604–613.
21. Alvarez-Navarro C, López de Castro JA (2014) ERAP1 structure, function and pathogenetic role in ankylosing spondylitis and other MHC-associated diseases. *Mol Immunol* 57(1):12–21.
22. Draenert R, et al. (2004) Immune selection for altered antigen processing leads to cytotoxic T lymphocyte escape in chronic HIV-1 infection. *J Exp Med* 199(7):905–915.
23. Kim S, et al. (2011) Human cytomegalovirus microRNA miR-US4-1 inhibits CD8(+) T cell responses by targeting the aminopeptidase ERAP1. *Nat Immunol* 12(10):984–991.
24. Kirino Y, et al. (2013) Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nat Genet* 45(2):202–207.
25. Bird LA, et al. (2003) Lymphoblastoid cells express HLA-B27 homodimers both intracellularly and at the cell surface following endosomal recycling. *Eur J Immunol* 33(3):748–759.
26. McHugh K, et al. (2014) Expression of aberrant HLA-B27 molecules is dependent on B27 dosage and peptide supply. *Ann Rheum Dis* 73(4):763–770.
27. Turner MJ, et al. (2005) HLA-B27 misfolding in transgenic rats is associated with activation of the unfolded protein response. *J Immunol* 175(4):2438–2448.
28. Bowness P, et al. (2011) Th17 cells expressing KIR3DL2+ and responsive to HLA-B27 homodimers are increased in ankylosing spondylitis. *J Immunol* 186(4):2672–2680.
29. Chen L, et al. (2014) Critical role of endoplasmic reticulum aminopeptidase 1 in determining the length and sequence of peptides bound and presented by HLA-B27. *Arthritis Rheum (Munch)* 66(2):284–294.
30. Evnouchidou I, et al. (2012) A common single nucleotide polymorphism in endoplasmic reticulum aminopeptidase 2 induces a specificity switch that leads to altered antigen processing. *J Immunol* 189(5):2383–2392.
31. Tsui FW, et al. (2010) Association of an ERAP1 ERAP2 haplotype with familial ankylosing spondylitis. *Ann Rheum Dis* 69(4):733–736.