HLA associations in inflammatory arthritis: emerging mechanisms and clinical implications

Robert Busch[1], Simon Kollnberger[2] and Elizabeth D. Mellins[3]

[1] Department of Life Sciences, University of Roehampton, Whitelands College, Holybourne Avenue, London, United Kingdom

[2] School of Medicine, Cardiff University, UHW Main Building, Heath Park, Cardiff, United Kingdom

[3] Department of Pediatrics, Program in Immunology, Stanford University Medical Center, Stanford, CA, United States

*emails: robert.busch@roehampton.ac.uk; mellins@stanford.edu

ORCID IDs:

0000-0002-4132-0751 (RB)

0000-0002-6904-3397 (SK)

0000-0003-2577-139X (EDM)

Abstract (149 words)

Our understanding of the mechanisms underlying HLA associations with inflammatory arthritis continues to evolve. Disease associations have been refined, and interactions of HLA genotype with other genes and environmental risk factors in determining disease risk have been identified. This Review provides basic information on the genetics and molecular function of HLA molecules, as well as general features of HLA associations with disease. We summarise evidence for various peptide-dependent and peptide-independent mechanisms by which HLA alleles might contribute to the pathogenesis of three types of inflammatory arthritis: rheumatoid arthritis, spondyloarthritis and systemic juvenile idiopathic arthritis. Also discussed are HLA allelic associations that shed light on the genetic heterogeneity of inflammatory arthritides and on the relationships between adult and pediatric forms of arthritis. Clinical implications range from improved diagnosis and outcome prediction to the possibility of using HLA associations in developing personalized strategies for the treatment and prevention of these diseases.

[ED: Some general editorial comments are replied to in the marginal comments. For replies to reviewers, please see the accompanying file.]

[H1] Introduction

Highly variable human leukocyte antigen (HLA) genes were first discovered in the 1950s; they were recognized as important effectors of transplant rejection and immune responses to specific antigens, following pioneering studies on their murine homologs in the H2 gene region (**Box 1** and refs. therein). Across species, this gene region is referred to as the major histocompatibility complex (MHC). The identification of HLA genes led to their intense investigation as inherited risk factors for conditions linked with autoimmunity, including rheumatic diseases. Among the earliest findings were the associations of HLA-B27 serotype (related to *HLA-B*27* genotype, see **Box 1** for nomenclature) with ankylosing spondylitis (AS)[1,2] and other spondyloarthritides (SpA), a group of arthritides with characteristic involvement of the spine, including psoriatic arthritis, inflammatory bowel disease (IBD) with SpA, and reactive arthritis as well as AS[3,4](**Table 1**). Similarly, early studies identified associations of HLA-DR4 serotype (related to *HLA-DRB1*04* genotype) and other HLA-DR alleles (which contain similar sequences at positions 70-74 of the β1 subunit, the 'shared epitope' or SE)[5,6] with rheumatoid arthritis (RA; **Table 1**). These associations have been replicated extensively and refined using genome-wide association study (GWAS) data (reviewed in [7,89]). HLA associations with many other rheumatic diseases have been identified, including, for example, systemic sclerosis[10], inflammatory myositis[11], and systemic lupus erythematosus (SLE)[12]. In general, different diseases are associated with different alleles of different HLA genes, although some alleles are associated with more than one disease.

The current paradigm for the pathogenesis of HLA-associated diseases holds that both genetic predisposition and environmental triggers contribute to disease development. HLA genes typically make the largest individual genetic contribution to inherited disease susceptibility, whether judged by the degree of confidence in the disease associations (extremely low *P* values) or by the magnitude of their effects on disease risk (high odds ratios)[13,14]. However, these genes are thought to account for substantially less than half of the full genetic burden[15,16], the remainder being attributed to polygenic influences and, as yet, incompletely mapped[17-19]. Disease occurrence among genetically susceptible individuals (penetrance) is typically low, reflecting risk modification by hormones, environmental factors (for example, infection, microbiota, lifestyle and diet), and stochastic effects on the formation of antigen receptor repertoires (such as in RA[20]). Moreover, the importance of any one genetic

contribution depends on the presence of other genetic factors (gene–gene interaction, also known as epistasis) or environmental factors (gene–environment interaction).

In this Review, we discuss HLA associations with diseases primarily characterized by inflammation-driven arthritis, such as RA, SpA and their pediatric counterparts. We also touch on other diseases in which inflammation-driven arthritis occurs, such as systemic sclerosis, SLE, dermatomyositis and systemic juvenile idiopathic arthritis (sJIA). We begin by describing the various HLA genes and functions of HLA molecules and summarize general features of HLA–disease associations. We present current hypotheses on the function of HLA risk alleles in disease etiology, focusing on RA and SpA, in which mechanistic understanding is arguably the most advanced, and sJIA, in which an HLA association has recently been identified. Finally, we consider the clinical implications of HLA associations for diagnosis, prognosis, definition of disease subtypes, therapy and prevention.

[H1] HLA genes and molecules

The genes coding for HLA proteins are located in a region on chromosome 6, interspersed between other genes involved in immune recognition and regulation (**Box 1; Fig. 1**). HLA proteins have peptide-binding domains and are expressed at the cell surface, allowing peptides to be presented for inspection by T cells via clonally diverse αβ T cell receptors (TCRs) (**Fig. 2**). Class I and class II HLA proteins are distinguishable, based on genomic location, expression, structure and function (**Fig. 1; Fig. 2; Box 2**). Class I HLA molecules primarily present peptides generated in the cytoplasm to cytolytic $CD8^+$ T cells[21] (**Fig. 2a; Box 2**). Class II HLA molecules bind peptides derived from endocytic processing, often after uptake from outside the cell, and present them to $CD4^+$ T cells, which produce cytokines to regulate other immune cells (**Fig. 2b; Box 2**). The recognition of HLA-bound peptides by TCRs drives thymic selection and other tolerance mechanisms, as well as antigen-specific T-cell responses to infection and, in autoimmunity, to self proteins.

Pathogen-driven selection and diversification of HLA genes during evolution has resulted in multiple HLA class I and class II genes, and multiple alleles at any one gene (see **Box 1**), such that most individuals are heterozygous at most HLA genes[22,23]. Much of the amino acid variation in HLA class I and class II alleles affects the shape and charge distribution within the peptide binding groove. These variations, in turn, control peptide side-chain preferences at important positions in the peptide

('anchor residues'), yielding distinct repertoires of self-peptides and non-self-peptides that bind to each allele[24]. Databases of HLA class I-binding peptides and class II-binding peptides are available (for example, the Immune Epitope Database and Analysis Resource)[25], as are algorithms that predict peptide-binding preferences for HLA alleles[26,27] (epitope prediction software for MHC class I and class II molecules).

[H1] HLA associations with disease

Many investigations of HLA disease associations have been case-control studies, which compare the frequencies of HLA alleles between groups of healthy individuals and patients with a specific condition (for HLA typing methods, see **Box 1**). In many early studies, HLA association signals were broad and difficult to pinpoint, because genetic recombination events in the HLA gene region rarely occur outside of a small number of recombination hot spots[22,28]. Thus, clustered polymorphisms on the same chromosome are co-inherited as haplotypes, spanning a number of adjacent HLA genes (linkage disequilibrium)[22,28]. For example, particular HLA-DRB1 alleles are consistently accompanied by particular HLA-DQ alleles (which are closely linked as shown in **Fig. 1**), and some extended haplotypes span much of the HLA region[29]. Distinguishing the disease-related alleles within these haplotypes has been a big challenge in the field.

The advent of high-density SNP arrays, which compare hundreds of thousands of single-nucleotide polymorphisms (SNPs) between thousands of patients and controls, has helped to overcome this problem[9,30,31]. Systematic procedures can resolve broad association signals and pinpoint the associated polymorphic sites at one or a few HLA gene loci[19,30,32]. These fine-mapped associations reflect polymorphisms at particular amino acid positions of HLA proteins and often modulate peptide interactions, although other interactions can also be affected. Many of these re-defined associations occur over diverse ethnic groups, updating an earlier impression that HLA associations tend to differ between ethnic groups[7,33]. In some instances, the associated sites are promoter polymorphisms that regulate expression levels of HLA class II molecules (for example, promoters of *HLA-DRB1*, *DQA1* and *DQB1* in systemic lupus erythematosus, or of *HLA-DPB1 in* systemic sclerosis)[34,35] or the expression of

HLA region-encoded accessory proteins, which influence peptide loading (for example, in *HLA-DOA* in RA and in *TAP2* in systemic sclerosis; see **Figs. 1** and **2** and **Box 2**)[35-37].

HLA–disease associations can have gene dosage effects. Odds ratios are higher for homozogous than for heterozygous genotypes of risk alleles (for example, in associations of SE-containing *DRB1* alleles with both RA[38] and rheumatoid factor (RF)-positive, polyarticular juvenile idiopathic arthritis (JIA)[39], and in at least some studies of *B*27* associations with AS[40]). In addition, in JIA, having an increased number of risk alleles has been correlated with a lower average age of disease onset[41].

HLA alleles associated with rheumatic diseases commonly confer increased risk of disease. Some HLA alleles, however, are less frequent in patients than in healthy individuals, implying that these alleles confer protection, either because they are less pathogenic than typical HLA alleles, or because they mediate an anti-inflammatory effect. For example, combined genetic, epidemiological and immunological evidence suggests a protective effect of *HLA-DRB1*13* in RA (**Table 1**), systemic lupus erythematosus and systemic sclerosis (reviewed in [42]).

[H1] Mechanisms of HLA disease associations

Different rheumatic diseases are associated with different HLA genes, suggesting that there is no unifying mechanism underpinning these associations. Conversely, as some HLA gene associations are shared for different diseases, some mechanistic features might also overlap. This idea is exemplified by the SpAs, all of which are associated with *HLA-B*27*(reviewed in [4]). Indeed, the association of *HLA-B*27* with disease in adults, adolescents and children with SpA suggests similar underlying mechanisms across the age spectrum[31]. Likewise, the expression of SE-containing *DRB1* alleles in both adults with RF-positive RA and pediatric patients with RF-positive polyarticular JIA suggests overlapping mechanisms[43]. In this section, we summarize the current data and concepts concerning the mechanisms of HLA association with RA, SpA, and systemic JIA (sJIA).

[H2] Rheumatoid arthritis and the HLA-DRB1 alleles

RA is characterized by symmetrical, typically erosive polyarthritis affecting distal joints. Anti-citrullinated protein (ACPA) and RF autoantibodies contribute to diagnosis, and their presence is associated with worse outcomes (as discussed further below)[44].

Twin studies indicate that this disease has a heritability of 53-68%, with little difference in heritability between ACPA-positive and ACPA-negative RA or between ethnic groups[45-47]. For ACPA-positive RA, the contribution of SE-containing *HLA-DRB1* alleles explains ~20% of the genetic risk, but it accounts for only a small percentage of the genetic risk of ACPA-negative RA[46]. HLA gene associations with ACPA-negative RA are less extensively characterised, but associations with *HLA-DRB1*03* and *HLA-B*08* have been identified[48].

[H3] Refinement of HLA associations with RA risk or protection

The original shared epitope (SE) hypothesis identified a sequence motif (Q[K/R]RAA in single-letter code) spanning residues 70-74 of the HLA-DR β chain as the distinguishing feature of the RA-associated HLA-DRB1 alleles known at the time (*HLA-DRB1*01:01*, **04:01*, and **04:04* in current nomenclature; cf. **Table 1**)[6]. Structural studies showed that the SE influenced the side-chain preferences of a pocket (P4) in the peptide binding groove of HLA-DR and included both TCR and peptide contact residues in the β-chain region flanking the bound peptide[49] (**Fig. 3a, *left***). An informative exception was a variant HLA-DR4 allele, *HLA-DRB1*04:02*[6], which is not associated with increased risk of RA[50-52]. *DRB1*04:02* encodes the sequence 70-DERAA-74 in place of the SE sequence, reversing the charge at the P4 pocket[6]. This finding also suggested that *HLA-DRB1*, rather than a closely linked gene within the *DRB1*04* haplotype, is the principal susceptibility locus[50]. Later studies identified additional SE-containing alleles (*DRB1*04:05, *04:08, *14:02, *10:01* and related others; **Table 1**) and revealed them to carry different levels of risk of RA, with *HLA-DRB1*04:01* having the strongest association in patients with European ancestry[38]. In 2012, GWAS data identified a minimal set of single-nucleotide polymorphisms that explained the HLA association with RA[32]. This analysis confirmed the importance of residues K71 and A74 in the traditional SE of HLA-DRB1 in disease susceptibility and identified an even stronger contribution from residue 11 in the P4 pocket in the floor of the peptide binding site (**Fig. 3a, *left***). These three residues explained all other association signals in the DRB1 gene, including residue 13, which is in linkage disequilibrium with residue 11[32].

Independent RA associations with single amino acid polymorphisms in *HLA-B* and *HLA-DPB1* were also identified [32]. Similar patterns were found in an Asian RA cohort[33].

Similarly, studies of DRB1 variation have considered whether protection against RA accrues from the presence of specific DRB1 alleles, such as *DRB1*13* [42,52](**Table 1**) or *DRB1*04:02* [50], or, rather, with the presence of particular sequence motifs, such as 70-DERAA-74 in place of the SE, or the presence of single amino acid polymorphisms, such as I67 or D70 [53,54]. This is not fully resolved.

Interestingly, several family studies show an increased risk of developing RA in patients who lack DRB1 SE alleles themselves, but whose mothers have SE+ DR alleles – an influence of non-inherited maternal alleles (NIMA) (reviewed in[55,56]). Not all such studies confirm the NIMA effect, which is thought to relate to the ability of maternal cells to enter the fetal circulation and shape immune system development. The mechanistic implications remain unclear.

[H3] Gene–environment interactions involving the shared epitope

Epidemiologic studies of ACPA-positive RA have provided striking evidence for a gene–environment interaction: two environmental factors, smoking and gingivitis, increase the risk of RA, but only in carriers of SE-containing HLA-DRB1 alleles[57-61]. Both smoking and gingivitis induce protein citrullination, a post-translational modification of positively charged arginine residues to citrulline, which is catalyzed by peptidylarginine deiminases (PAD) (**Fig. 3b**)[62]. Smoking-induced inflammation upregulates host PAD enzymes in the lungs[57]. *Porphyromonas gingivalis*, a bacterium associated with gingivitis, has its own PAD enzyme; another gingivitis-associated pathogen, *Aggregatibacter actinomycetemcomitans*, produces a toxin that induces hypercitrullination of proteins in neutrophils and promotes their extracellular release[60]. This evidence suggests a model in which the SE confers risk of autoreactivity directed against citrullinated self-proteins, which are generated at sites of neutrophilic inflammation. The participating self-reactive lymphocytes and autoantibodies may then attack the joint.

[H3] Neoantigen hypothesis

An attractive hypothesis is that, in RA, citrullination can make some self-antigens appear to be non-self, which can lead to the loss of self-tolerance (known as the neoantigen hypothesis). This hypothesis

might explain the association of ACPA-positive RA with the SE of HLA-DRB1, as citrullination differently affects the binding of some peptides to HLA-DR proteins with and without the SE (**Fig. 3a, *right***). Studies of peptides derived from foreign antigens unrelated to RA provided evidence that the P4 pocket of the HLA-DR0401 protein accommodates neutral, polar or negative side chains, but is far less accommodating to positively-charged arginine residues[63-65]. This suggested that self peptides with arginine residues at position P4 (the residue that binds the P4 pocket) may not be presented under non-inflammatory conditions, when PAD is unavailable, and thus fail to induce tolerance. Citrullination of such arginyl side chains, which neutralizes their positive charges, occurs in inflammatory states when PAD is released by neutrophils[57,60]. At least some of the resulting P4-citrullinated self peptides acquire the ability to bind HLA-DR0401 molecules[66,67]. These citrullinated self peptides are then available to stimulate effector/memory CD4+ T helper cells, which are expanded in the circulation of patients with RA, compared to healthy individuals, whereas regulatory T cells cells are depleted[67] (**Fig. 3a, *right***). This process probably enables B cells with surface immunoglobulin specific for citrullinated self-antigens to recruit T cell help and produce ACPAs. Crystallographic studies and binding assays evaluating the interaction of HLA-DR0401 protein with several P4-citrullinated self peptides confirm gains in binding following citrullination[67,68].

Studies with other shared epitope-containing HLA-DR variants and more peptides have shown, however, that the amount of binding gained from P4-citrullination is variable, and in several cases no gain is apparent, suggesting that the magnitude of the effect is context-dependent[68,69]. Citrullination at other anchor or non-anchor residues usually has no effect on (or weakens) binding[68,70]. Importantly, to support a neoantigen hypothesis, it is not necessary to postulate that all P4-citrullination events result in improved binding to shared epitope-containing HLA-DR alleles – the loss of tolerance to a small number of neoantigens might be enough to explain an increased disease risk.

Similar considerations may answer the reverse question of why other HLA-DRB1 alleles have no association with ACPA-positive RA (**Supp. Fig. 1**). The RA-nonassociated HLA-DR4 variant protein, DR0402 (with its *DRB1*04:02*-encoded β-chain harbouring the 70-DERAA-74 sequence in place of the SE), binds P4-citrullinated peptides, but unlike HLA-DR0401, this variant binds unmodified (P4-arginine-containing) precursors almost as well, suggesting that the parent peptides are available to induce tolerance (**Supp. Fig. 1a**)[67]. The binding (or lack of binding) of self-peptides to other HLA-DR

alleles is presumably unaffected by citrullination of arginine residues (**Supp. Fig. 1b**), as apart from their different charges, arginine and citrulline are very similar in structure. Citrullination might also affect other aspects of antigen presentation in RA (for example, by modifying the proteolytic processing of self-proteins), but there is no known reason for these other aspects to select for SE-containing HLA-DRB1 alleles[67,71].

Another mechanism has been proposed[72] to explain the protective effects of *HLA-DRB1*13* in ACPA-positive RA, although confirmation is still needed. *HLA-DRB1*13* encodes the DERAA sequence, as does *HLA-DRB1*04:02*, and the DR13 protein has been reported to be broken down into peptides with a DERAA core, which bind to HLA-DQ [72]. These peptides induce T cell tolerance against homologous microbial peptides and citrullinated vinculin peptides, to which ACPA-positive patients with RA would otherwise mount pathogenic T cell responses.

The discovery of anti-carbamylated protein (anti-CarP) autoantibodies in a subset of patients with RA has raised new questions about the mechanisms by which post-translational modifications break self tolerance. Carbamylation is a post-translational modification of lysine by cyanate, which is produced from thiocyanate by neutrophil myeloperoxidase during oxidative stress (**Fig. 3b**). Anti-CarP autoantibodies and ACPAs are not usually co-expressed in the same patients and do not seem to cross-react with each other[73]. Anti-CarP antibodies are not associated with the shared epitope or with smoking[74]. The effect of lysine carbamylation on the presentation of self peptides by SE+ DR alleles has not been characterised, but when examined using foreign peptides, the SE does not seem to discriminate strongly against lysine at P4 [63]. As a result, the SE may not confer an allele-specific mechanism for breaking tolerance through gain of binding after carbamylation. This would explain the lack of SE association with anti-CarP autoantibodies. However, the mechanism by which carbamylation breaks self-tolerance remains unresolved.

The neoantigen model is attractive because it accounts for the selective presence of ACPAs (but not anti-CarP autoantibodies) in patients with RA who carry shared epitope-containing HLA-DRB1 alleles; these alleles discriminate between certain non-binding parental peptides (containing arginine at P4) and post-translationally modified peptides (containing citrulline at P4). Notably, analogous mechanisms have been identified in celiac disease and in type 1 diabetes, wherein tissue transglutaminase creates neoantigens that bind to HLA-DQ risk alleles[75,76].

[H3] Calreticulin signalling hypothesis

Another mechanistic hypothesis to explain the association of SE-containing *DRB1* alleles with RA proposes that the SE region of the HLA-DR β chain provides a ligand for cell signalling that other *HLA-DRB1* alleles lack, by interaction with calreticulin (**Supp. Fig. 2a).** Calreticulin was discovered as a calcium-binding protein that chaperones nascent glycoproteins in the ER; its translocation to the plasma membrane of damaged cells activates phagocytosis[77]. SE-containing peptides or intact SE-containing HLA-DR molecules bind calreticulin[78], which stimulates proinflammatory changes, such as nitric oxide production, in antigen-presenting cells (APCs), in a CD91-dependent fashion[79]. The same interaction has been proposed to explain how, in $CD8^{-}CD11c^{+}$ dendritic cells, SE-containing peptides trigger the production of IL-6 and IL-23, which activate T_H17 cells, whereas in $CD8^{+}CD11c^{+}$ dendritic cells the same stimulus inhibits the production of indoleamine 2,3 dioxygenase (IDO), an enzyme important in regulatory T-cell activation[80]. A cyclised SE-containing peptide has been shown to bind calreticulin, enhance osteoclast differentiation and activation *in vitro*, and exacerbate bone erosion in murine arthritis[81]; these processes could contribute to erosive disease in RA patients with SE-containing DRB1 alleles[82]. Intriguingly, citrullination of calreticulin increases its affinity for the shared epitope, providing a link between the shared epitope and environmental factors known to increase the risk of erosive RA[83]. However, these SE-linked innate immune effects do not account for specific autoimmunity against citrullinated self antigens.

[H3] Low CLIP affinity hypothesis

Another mechanistic hypothesis derives from the observation that SE-containing HLA-DR molecules have a low affinity for class II-associated invariant chain peptides (CLIP), the nested set of invariant chain (Ii) fragments that occupy the peptide binding groove prior to the exchange for endosomal peptides during HLA class II–peptide complex assembly (**Fig. 2b**)[84]. For most HLA class II alleles, efficient peptide exchange requires catalysis of peptide release by the accessory molecule HLA-DM (**Fig. 2b**), but alleles with a lower affinity for CLIP, such as HLA-DR0401, can release CLIP spontaneously and might load peptides without the participation of HLA-DM[85] (**Supp. Fig. 2b**). As HLA-DM selects for a stably bound peptide repertoire, HLA-DR protein variants with low CLIP affinity

may acquire less stably bound peptides; indeed, these proteins have been shown to undergo increased peptide exchange at the cell surface[85]. This may influence loss of self-tolerance; for example, at sites of inflammation, extracellular proteases generate peptides (neoepitopes) that can be presented by HLA molecules owing to peptide exchange at the surface of APCs. The spontaneous release of CLIP and subsequent loading of weakly-bound peptides are also associated with a faster turnover of HLA class II molecules, leading to lower surface expression, which could impair central and peripheral tolerance mechanisms[86].

Whether low CLIP affinity is relevant to the contribution of the DRB1 SE to RA susceptibility remains unclear, but some evidence implicates this mechanism in other HLA associations. A *HLA-DPB1* polymorphism (Gly84) impairs the folding of the CLIP region of Ii into the peptide-binding groove [87]. This impairment enables endosomal peptide loading of HLA-DP without HLA-DM editing (**Supp. Fig. 2b**) [88], as well as premature loading with TAP-transported peptides in the endoplasmic reticulum (ER)[89]. Intriguingly, this polymorphism has been associated with susceptibility to ACPA-positive RA (and other diseases) in Japanese individuals (**Table 1**)[37]. Moreover, in the nonobese diabetic mouse model of type 1 diabetes, mutagenesis of CLIP to strengthen its binding to another low CLIP-affinity class II allele protects against disease[90].

[H3] Influence of the shared epitope on microbiota

The intestinal microbiota are known to affect immune regulation and autoimmunity, and HLA genes might, in turn, influence the composition of the intestinal microbiota. For example, compared to healthy individuals, the abundance of *Prevotella copri* was increased in the intestines of patients with new-onset RA, but not in patients with established RA. Even though the abundance of *P. copri* was influenced by HLA genes, this bacterium was most abundant in patients without SE-containing DRB1 alleles[91], arguing against a mechanism whereby the SE-containing risk alleles select arthritogenic microbiota. In HLA-DR4-transgenic mice, expression of shared epitope-containing alleles is associated with increased intestinal permeability, complex alterations in microbial composition, and the intestinal expression of type 17 cytokines[92], but the relevance of these findings to RA remains unclear.

[H2] Spondyloarthritis and HLA-B*27

The SpAs share associations of varying strength (greatest for AS) with *HLA-B*27:05* and other *HLA-B*27* subtypes, of which over 100 have been identified, though only the more prevalent subtypes have been examined for disease associations (**Table 1**)[93]. SpA characteristically affects the spine and sacroiliac joints; other clinical features include extra-articular enthesitis, peripheral arthritis (particularly in the lower limbs) and osteitis[94]. The expression of *HLA-B*27:05* in rats and mice causes joint pathology with some features of SpA[95]. Disease manifestations in *HLA-B*27* transgenic rats include destructive arthritis, enthesitis, spondylitis and ankylosis, which resembles human disease more closely than the disease in *HLA-B*27* transgenic mice[95]. Data from an early twin study indicated that although AS risk is mostly controlled by genetic rather than environmental variation, *HLA-B*27* genotype contributes less than half of the heritable component[96]. More recently, researchers reported that *HLA-B*27* contributes 20.1% of the heritability of AS, and identified other genes that could explain 4.3% of the heritability, although the majority of the genetic heritability remained unexplained[19]. A GWAS identified additional AS-associated genes[14], including the gene that encodes endoplasmic reticulum aminopeptidase 1 (ERAP1), which mediates N-terminal trimming of class I-bound peptides in the ER during HLA class I peptide loading[14]. Alleles of ERAP1 that confer reduced enzymatic activity against certain peptide substrates have been reported to be protective against AS[97-99].

Notably, polymorphisms in *ERAP1* are only linked to AS in patients who express *HLA-B*27* [36] or *HLA-B*40* [30] (another AS risk allele; **Table 1**) and not in patients who express other *HLA-B* alleles. This epistatic (gene/gene) interaction suggests that ERAP1 contributes to disease susceptibility by exacerbating the pathogenic role of HLA-B*27. Another epistatic interaction results in a greater risk for developing AS in patients heterozygous for two HLA class I risk alleles (HLA-B*27 and HLA-B*40:01, the latter encoding the Bw60 serotype) than would be expected from the additive effect of the two alleles on disease risk [100,101].

Additional AS-associated variants at genes such as the IL23 receptor, IL1 receptor type II and T cell transcription factors increase the propensity for pro-inflammatory signalling by the T_H17-promoting cytokines, IL-1 and IL-23[14]. These findings are consistent with data from immunological studies and clinical trials implicating Th17 and other IL-17-secreting cells in SpA[102].

In SpA, inflammation of the sacroiliac and intervertebral joints is associated with both erosive bone changes and new bone formation at the entheses. HLA-B27 does not directly affect bone formation in

transgenic mice or osteoblast differentiation by mouse or human bone precursor cells *in vitro* [80]; rather, epigenetic and gene expression analyses suggest that new bone formation occurs in SpA owing to HLA-B27-driven inflammation[103,104]. The mechanistic hypotheses proposed to account for these findings are depicted in **Fig. 4a** and explained in the remainder of this section.

[H3] Arthritogenic peptide hypothesis

A possible source of HLA-B27-driven inflammation in SpA is the presentation of pathogen-derived, 'arthritogenic' peptides (that is, exogenous peptides that are sufficiently similar to self-antigens to incite autoimmunity) to $CD8^+$ T cells and the resulting selection of high affinity autoreactive T cells (similar to the high-affinity T-cell selection that occurs during repeated antigen presentation in chronic infections[105]). T cells that recognize HLA-B27-bound candidate arthritogenic peptides from intracellular bacteria have been identified in patients with reactive arthritis[106]. A link between peptide presentation and disease could result from differences in features of peptides bound to disease-associated versus non-associated HLA-B27 alleles[107-110], of which presentation by the former could lead to the selection of autoreactive T cells. However, no qualitative differences between peptides that can bind these allelic groups have been identified[111]. Nonetheless, consistent with an antigen-driven process, oligoclonal T cell expansions are present in the blood and at sites of inflammation in patients with SpA[112-114]. Interestingly, joint-infiltrating $CD8^+$ T cells that, unusually, produce IL-17 have been detected in patients with PsA[115]. Indeed, arthritis risk in psoriasis patients has been linked to *HLA-B*27*, specifically glutamic acid at position 45[116] and asparagine at position 97[117] , both of which are located in the peptide-binding groove and likely to influence peptide presentation (**Fig. 4b**)[118]. However, in transgenic rats with high copy numbers of the HLA-B27 transgene, $CD8^+$ T cells are not required for the development of arthritis and intestinal disease [119]. This finding, together with the low frequencies of self-reactive or cross-reactive $CD8^+$ T cells in patients with other SpAs, such as AS, and the apparent function of T_H17 cell-derived cytokines in SpA (see above[102]), suggests that HLA-B27-restricted $CD8^+$ T cell involvement cannot fully account for SpA. However, failure to control the activity of high avidity autoreactive CD8+T cells could contribute to pathogenesis in some types of SpA, such as reactive arthritis and PsA.

[H3] Cell surface open conformations of HLA-B27

Another mechanistic hypothesis to explain the association of *HLA-B*27* alleles with SpAs is based on the fact that a free cysteine (Cys67; **Fig. 4b**) (and perhaps other cysteines)[120] in HLA-B27 heavy chains promotes the formation of peptide-free, open conformations of HLA-B27 heavy chains (including dimers and possibly other forms), which do not include β_2-microglobulin. Such open conformations are expressed at the cell surface and participate in pathogenic immune signalling (**Fig. 4a**)[121-123]. Cell surface HLA-B27 open conformations are generated from unstable peptide–B27 complexes during HLA class I recycling[121]. Cys67-dependent HLA-B27 homodimer formation has been detected on the cell surface of transfected B cell lines and SpA patient and *HLA-B*27* transgenic rat monocytes[121,122,124,125]. Furthermore, in addition to HLA-B27 homodimers, HLA-B27 could also dimerize or oligomerize with other cell surface molecules, such as HLA-F, which binds to disassembling HLA-class I molecules[121,126]. Unique molecular features of HLA-B27, including Cys67, could promote the formation of free class I heavy chains; during class I recycling, cell surface HLA-B27 open conformations are generated from unstable peptide–B27 complexes[121,125]. The expression of such conformations is increased on cells in the inflamed intestine and joints of patients with SpA [124,127,128]. Additionally, in a *HLA-B*27* transgenic rat model of arthritis, the expression of open conformations of HLA-B27 on immune cells increases with disease progression[125].

The enzymes ERAP1 and ERAP2 trim peptides and control the supply of appropriately-sized peptides for class I binding, thus controlling the rate of formation of cell surface peptide–HLA-B27 complexes (**Fig. 2a**) and the resulting HLA-B27 open conformations (**Fig. 4a**). The evidence linking specific *ERAP1* alleles with differences in the levels of HLA-B27 heavy chain open conformations is controversial. In one report, SpA-associated *ERAP1* variants were associated with an increased production of HLA-B27 open conformations[129]. Protective *ERAP1* variants were associated with reduced cell surface expression of class I heavy chains on cell lines and SpA patients' cells[129]. In this study, gene silencing or pharmacologic inhibition of ERAP1 reduced the expression of HLA-B27 open conformations, reduced their binding to killer cell immunoglobulin-like immune receptors (KIR), and reduced T cell-mediated production of IL-17 in both *in vitro* and *ex vivo* experiments[129]. By contrast, in other studies, *ERAP1* gene silencing and the expression of protective *ERAP1* variants increased the cell surface expression of HLA-B27 open conformations and intracellular formation of HLA-B27 heavy

chains[130,131]. These observations highlight the difficulties with extrapolating data from cell lines overexpressing HLA-B27 to disease. As ERAP1 affects the supply of peptides to HLA-B27, differences in the expression of *ERAP1* and *HLA-B*27* genes in patients and experimental systems could contribute to these contradictory observations.

Unlike peptide–HLA-B27 complexes, which interact with the TCR, HLA-B27 open conformations bind strongly to the inhibitory killer cell receptor, KIR3DL2, and leukocyte immunoglobulin-like receptor (LILR) family members, LILRB2 and LILRB5 [132,133]. HLA-B2705 protein forms more open conformations that bind to KIR and LILR than the non-disease-associated molecule, HLA-B2709[134]. KIR3DL2 is expressed on natural killer cells, γδ T cells and activated $CD4^+$ and $CD8^+$ αβ T cells[135,136]. KIR3DL2 also binds to other class I heavy chains, including HLA-F, which has also been linked to SpA[137,138]. However, KIR3DL2 binds more strongly to open conformations of HLA-B27 than to other HLA class I heavy chains[139]. LILRB2 and LILRB5 receptors that can bind HLA-B27 open conformations are expressed on monocytes and osteoclasts[140-143]. Thus, the binding of HLA-B27 open conformations to immune receptors probably affects the function of a diverse array of immune cells in SpA.

The binding of KIR3DL2 to open conformations of HLA-B27 promotes the production of IL-17A by T_H17 cells *in vitro*[144]. HLA-B27-positive patients with SpA have increased numbers of KIR3DL2-positive T_H17 cells and NK cells in their peripheral blood, peripheral joints and inflamed intestinal sites[136,145]. KIR3DL2-positive T_H17 cells account for the majority of IL-23 receptor-expressing T cells in patients with SpA[146]. These cells also express markers of activated T cells and other markers that suggest migration from the intestine, and they show evidence of oligoclonal expansion[144,146]. Inhibiting the binding of KIR3DL2 to HLA-B27 open conformations on transfected cells and on cells that naturally express HLA-B27 both inhibits survival of activated T cells from patients with SpA and healthy controls *in vitro* and limits IL-17 production by patient T cells *ex vivo* [144,147]. The binding of open conformations of HLA-B27 to KIR3DL2 on NK cells inhibits their production of IFNγ, potentially reducing an important brake on IL-17 production[124,139,147]. Because KIR3DL2 is an inhibitory receptor, KIR3DL2–HLA-B27 interactions would be expected to reduce activation signals in KIR3DL2-expressing cells. T cell survival and differentiation depends on a balance between activating and inhibitory signals. Diminished T cell signalling, in T_H17 cell differentiation cultures and in a mouse model of SpA, promotes the production of IL-17 and causes arthritis, in large part, by enhancing the

survival of autoreactive T cells and increasing the ratio of T_H17 to T_H1 cells[148,149]. Thus, the propensity of HLA-B27 to form open conformations seems to be linked to increased IL-17 production and an increased T_H17–T_H1 ratio, which have important functions in driving SpA pathogenesis. Hence, open B27 conformations might synergize with other factors promoting T_H17 activity or limiting T_H1 activity. The precise contribution of KIR3DL2 and LILR interactions with HLA-B27 open conformations to the differentiation of T_H17 cells, osteoclasts and other immune subsets in SpA warrants further investigation.

[H3] HLA-B27 misfolding, the unfolded protein response and autophagy

HLA-B27-mediated activation of the unfolded protein response (UPR) to ER stress, and its putative contribution to arthritis, have been extensively reviewed elsewhere[150]. The free Cys67 and other molecular properties of HLA-B27 heavy chains promote HLA-B27 misfolding in the ER[150]. Inefficient intracellular disposal of misfolded molecules causes ER stress and promotes the production of inflammatory mediators through the induction of the UPR[150,151]. Misfolded proteins in the ER are disposed of by ER-associated degradation (ERAD). The slow assembly of peptide–HLA-B27 complexes predisposes the cell to increased levels of ER-associated degradation (ERAD), the induction of an UPR and autophagy, especially when HLA-B27 production is upregulated during inflammation[150]. Disease-associated HLA-B27 subtypes oligomerize, accumulate in the ER more readily, and are more susceptible to ERAD than non-associated class HLA-B27 subtypes[152-154]. Increased sequestration of unfolded HLA-B27 by the ER chaperone BiP (immunoglobulin binding protein) increases the expression of UPR-related genes and pro-inflammatory immune genes, such as IL-23[155]. The UPR also promotes the production of IL-1α and IFNβ, which increase and inhibit T_H17 cell differentiation, respectively. Autophagy, induced as a consequence of a UPR, also promotes IL-23 production.

In *HLA-B*27* transgenic rats, which express very high levels of HLA-B27, misfolded HLA-B27 drives the UPR in macrophages[156] and is associated with increased osteoclast formation[157]. However, there are conflicting reports concerning whether the UPR is upregulated in mononuclear cells and macrophages from the synovial tissue of patients with SpA[153,158,159]. While an early study showed upregulation of BiP in mononuclear cells from peripheral joints of AS patients[133], a later study showed activated macrophages from patients with SpA produced more IL-23 than those from healthy individuals, but

without any notable upregulation of the UPR[160]. Other studies showed that while autophagy was increased without an ongoing UPR in the intestinal tissue of patients with SpA, autophagy was not upregulated in their peripheral blood or joints[161,162]. Disease-associated polymorphisms in *ERAP1* are not correlated with ER stress in patients with AS[163]. Overall, further work is needed to distinguish the direct effects of HLA-B27 on ER stress and the UPR from effects occurring downstream of inflammation.

[H3] HLA-B27 and the microbiome

HLA-B27 biological effects could influence SpA pathogenesis by shaping the microbiome. HLA-B27 promotes the survival of Gram-negative intracellular bacteria in transfected monocyte cell lines and in HeLa epitheloid cell lines expressing HLA-B27[164,165]. *In vivo*, any possible direct effects of HLA-B27 on the survival of microbes that promote arthritis are difficult to distinguish from the effects of an ineffective immune response. In *HLA-B*27* transgenic rats, antibiotic administration inhibits osteoclast formation and attenuates arthritis[166]. *HLA-B*27* transgenic rats have an increased susceptibility to Gram-negative intracellular bacterial infections, which could trigger development of SpA[167,168]. Perturbations in the intestinal microbiome in rodent and human SpA could be because of imbalances in type 1 and type 17 cytokine production[169,170]. A recent study in *HLA-B*27* transgenic rats emphasizes the relationship between the host genetic background and associated changes in multiple microbes in establishing an inflammatory environment that promotes arthritis[171]. This area is currently under intense investigation in human studies.

[H2] sJIA and HLA-DR

In many inflammatory arthritides, the mechanisms underpinning HLA disease associations are less well-understood than in RA and SpA, especially where structural studies have not provided mechanistic clues, as exemplified by sJIA. GWAS data from patients of European ancestry identified a sJIA association with *HLA-DRB1*11:04* and *HLA-DRB1*11:01*[172]. The researchers pinpointed a residue in the HLA-DR β-chain, glutamic acid at position 68, that confers increased risk of sJIA; this residue does not interact directly with bound peptide. sJIA has many features of innate immune dysfunction, including systemic inflammation responsive to IL-1 and IL-6 inhibition[173], and the absence of RF and

ACPA autoantibodies (as well as the absence of anti-nuclear antibodies in most patients)[174]. However, some evidence supports a contribution from CD4+ T cells to established sJIA[175-178], and a subset of patients with sJIA develop chronic erosive arthritis without continued systemic features[179]. Thus, one model is that the initial phase of sJIA is dominated by innate immune activation and a subsequent phase is T cell dominant[180]. In this model, IL-1 driven development of T_H17 cells is important for progression to the second (arthritic) phase[181]. In line with this model, the *HLA-DRB1*11* association could reflect HLA-DR-restricted T cell responses in this second phase[182]. Alternatively, as cell surface HLA-DR binds superantigens, which then engage and activate T cells that express particular TCR Vβ families[183], a HLA-DR–superantigen interaction could be preferentially promoted by the HLA-DRB1*11 molecule[184]. Another possibility is that HLA-DR molecules could regulate innate immune responses in sJIA. For example, intracellular HLA-DR has been reported to influence Toll-like receptor signal strength in monocytes and dendritic cells[185]. An increased understanding of sJIA immunobiology is required to refine the mechanistic hypotheses on the association of *HLA-DRB1*11* with this disease.

[H1] Clinical implications

[H2] Diagnosis and prognosis

For diagnosis, even the strongest HLA associations are generally insufficient alone, because most people with risk alleles never develop the associated disease. However, strong HLA disease associations can sometimes contribute to a diagnosis. For example, identifying the presence of disease-associated *HLA-B*27* alleles, and thus a risk of SpA, can aid the diagnostic evaluation of back pain[186]. Another potential clinical application of HLA typing is early identification of individuals at high arthritis risk among patients with psoriasis[187188]. One analysis identified glutamic acid at position 45 (e.g., in the peptide binding groove of *HLA-B*07, B*08, B*27*, *B*38* and *B*39*; **Fig. 4b**) as the critical risk factor associated with PsA in psoriasis patients[116]. A second analysis, with age of psoriasis onset considered as a covariate, found that the primary risk for PsA in psoriasis patients derives from asparagine (as found in HLA-B*27) or serine (in B*07 and B*08) residues at amino acid position 97 of HLA-B (**Fig. 4b**)[117]. The discrepancy between the two studies may reflect differences in the clinical subtypes between PsA cohorts in the two studies, as there is evidence that HLA-B allelic associations

may differ with PsA clinical phenotype[189]. Although mapping associations to amino acid residues is shedding light on PsA pathogenesis, in practice, screening of psoriasis patients for arthritis risk, if proven useful, is likely to employ simple typing for HLA-B*27 in the near term[117].

The scope for using HLA genetic data in diagnosis and prognosis might increase as patient genomic data become more widely available. Importantly, HLA allele associations are also being used as part of quantitative polygenic risk scores, which are scores based on weighted contributions of alleles at multiple gene loci. By including variation in multiple genes (HLA and non-HLA related genes; for example, >100 loci for RA[190]) and incorporating the effects of gene–gene interactions[191,192], polygenic risk scores can improve the prediction of inherited susceptibility for complex genetic traits, such as rheumatic diseases. Their clinical use is generally not imminent, but holds promise for the future[193].

[H2] Disease subtypes

Phenotypic subtypes of patients with the same clinical diagnosis can have distinct HLA associations, often with distinct autoantibody profiles. As discussed above, the most severe, erosive form of RA is associated both with HLA-DRB1 alleles that contain the SE[33,58] and with seropositivity for RF, ACPA or anti-cyclic citrullinated peptide (anti-CCP) autoantibodies, although whether ACPAs are merely markers or active drivers of disease severity is unclear[194]. Among patients with RA, carriers of the DRB1 SE are also more responsive to DMARDs and some biologic therapeutics than non-carriers[194,195]. Interestingly, the protective effect of the *HLA-DRB1*13* allele seems to be confined to ACPA-positive RA[196]. HLA-DR alleles are not always the primary risk factors for particular subgroup associations; for example, the *HLA-DQB1*03:01* is associated with bronchiectatic airway disease and emphysema in RA, whereas *HLA-DQB1*03:02* confers resistance to these conditions[197,198].

Additional examples of autoantibody-defined patient subsets with different associated HLA alleles include dermatomyositis and systemic sclerosis. In dermatomyositis, carriage of anti-Mi-2 (a nuclear antigen) antibodies is associated with *HLA-DRB1*07:01,* whereas carriage of anti-TIF1-γ (transcription intermediary factor 1) antibodies is associated with *HLA-DQB1*02*[199]. In systemic sclerosis, the two main serological patient subsets, defined by carriage of anti-centromere antibodies (anti-CA) or anti-topoisomerase antibodies (anti-Topo I), also are distinguished by genetic differences at HLA class II

loci[35,200]: for example, *DRB1*15:02* (anti-CA) and *DRB1*16:02* (anti-Topo I) in the Han Chinese population[200].

HLA allelic associations can also support classification schemes built primarily on clinical phenotype. For example, In JIA, HLA class II allelic associations support the International League of Associations for Rheumatology (ILAR) classification of JIA subtypes[201]. Specifically, oligoarticular JIA (≤ 4 joints), RF-positive polyarticular JIA (≥ 5 joints), and systemic JIA are linked to different HLA alleles. Susceptibility alleles for oligoarticular JIA are *HLA-DRB1*08*, *HLA-DRB1*11* and *HLA-DPB1*02:01*[202,203], for RF+ polyarticular JIA are DRB1 SE alleles[204], and for systemic JIA, *HLA-DRB1*11:01* and *HLA-DRB1*11:04*[172]. The association of some HLA class II alleles with particular phenotypes in JIA provides support for the existence of particular clinical subtypes of JIA[31]. In addition, the overlap in HLA associations of oligoarticular JIA and RF-negative polyarticular JIA supports many analyses that have combined these two subgroups, which share early onset, female predominance and a high prevalence of antinuclear autoantibodies[205]. In 2017, HLA alleles were found to be shared between JIA subtypes and clinically similar types of adult inflammatory arthritis[31]. Notably, the HLA associations were stronger in the paediatric subtypes compared with their adult counterparts, consistent with a greater degree of heritability. Genetically and immunologically similar diseases in adults and children ultimately might be re-classified as the same disease[206]. HLA alleles will be instrumental in such classifications.

[H2] Therapy

No hypothesis linking particular HLA alleles to autoimmune or inflammatory mechanisms has been proven conclusively. However, in SE-positive RA and *HLA-B*27*-positive AS, therapeutic strategies based on mechanistic models might ultimately provide strong support for specific hypotheses (**Fig 3**, **Fig 4, Supp. Fig. 2**). For example, the model invoking shared epitope-containing allele-specific presentation of post-translationally modified peptides can be tested by developing inhibitors of PAD4[207]. Alternatively, a critical contribution of T cell responses to citrullinated peptides in RA pathogenesis would be supported if RA disease activity is curbed by therapeutic vaccination or a T cell therapy focused on these neoantigens. Encouraging preliminary findings have been reported for immunotherapy with tolerogenic dendritic cells presenting citrullinated peptides in patients with RA[208].

In the HLA-B27 example, targeting T_H17 cell-promoting cytokines or IL-17 has proven therapeutic efficacy in AS[209,210], regardless of how HLA-B27 boosts T_H17 cell immunity, whereas targeting IL-23, specifically induced via the HLA-B27 misfolding–UPR pathway in the ER (**Fig. 4**), has been ineffective[211]. If blocking KIR3DL2 or open conformations of HLA-B27 were efficacious in patients[212,213], such a finding would strongly support a pathogenic function for this interaction. Moreover, although the exact function of ERAP1 in AS is unclear, the fact that ERAP1 variants with reduced activity are protective suggests that enzyme inhibitors could be beneficial[214,215]. Indeed, preliminary observations indicate that ERAP1 inhibitors limit the production of IL-17 by patient Th17 cells *ex vivo*[129].

Furthermore, the shared HLA associations between adult and childhood RF-positive arthritis or in SpA provide a rationale for testing treatments in children that have efficacy in adults. The use of drugs targeting the IL-17 axis, currently in use in adults[209,210] will probably become part of the armamentarium for HLA-B27-associated SpA in children and adolescents[216]. Any therapeutic benefit from such an approach would support the notion of shared disease mechanisms in children and adults.

[H2] Prevention

Studies have begun to elucidate how genetic and environmental factors combine to affect disease risk. Smoking[57-59] and gingivitis[60,61] are environmental risk factors for RA, but only in individuals who carry SE-containing HLA-DRB1 alleles. In principle, this knowledge enables the identification of at-risk individuals so that preventative approaches, such as smoking cessation and oral hygiene, can be personalized. Such genetically-targeted behavior modification could become an increasingly viable strategy as the cost of genotyping declines. Polygenic risk prediction models that incorporate environmental risk factors are being evaluated as an approach to tailor disease prevention strategies[217]. The prospect of reducing the possibility of chronic rheumatic disease might motivate individuals who are genetically at risk of disease to adhere to preventative measures.

[H1] Conclusions

Rheumatology has been at the forefront of understanding the genetic basis of complex autoimmune conditions. Accumulating data support the idea that HLA molecules contribute to disease through a

number of mechanisms, including the presentation of pathogenic neoantigens in RA and the production of pathogenic abnormal cell-surface HLA-B27 conformations in AS. The therapeutic promise of such mechanistic insights remains to be fully realised. Meanwhile, genetic data are increasingly becoming available to rheumatologists, patients and individuals at risk. Such information might aid diagnosis, disease stratification, and personalised treatment strategies.

References

1 Brewerton, D. A. *et al.* Ankylosing spondylitis and HL-A 27. *Lancet* **1**, 904-907 (1973).

2 Schlosstein, L., Terasaki, P. I., Bluestone, R. & Pearson, C. M. High association of an HL-A antigen, W27, with ankylosing spondylitis. *N Engl J Med* **288**, 704-706, doi:10.1056/NEJM197304052881403 (1973).

3 Brewerton, D. A. *et al.* Reiter's disease and HL-A 27. *Lancet* **302**, 996-998 (1973).

4 Bowness, P. HLA-B27. *Annu Rev Immunol* **33**, 29-48, doi:10.1146/annurev-immunol-032414-112110 (2015).

5 Stastny, P. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. *N Engl J Med* **298**, 869-871, doi:10.1056/NEJM197804202981602 (1978).

6 Gregersen, P. K., Silver, J. & Winchester, R. J. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum* **30**, 1205-1213 (1987).

7 Newton, J. L., Harney, S. M., Wordsworth, B. P. & Brown, M. A. A review of the MHC genetics of rheumatoid arthritis. *Genes Immun* **5**, 151-157, doi:10.1038/sj.gene.6364045 (2004).

8 Viatte, S., Plant, D. & Raychaudhuri, S. Genetics and epigenetics of rheumatoid arthritis. *Nat Rev Rheumatol* **9**, 141-153, doi:10.1038/nrrheum.2012.237 (2013).

9 Brown, M. A., Kenna, T. & Wordsworth, B. P. Genetics of ankylosing spondylitis--insights into pathogenesis. *Nat Rev Rheumatol* **12**, 81-91, doi:10.1038/nrrheum.2015.133 (2016).

10 Chairta, P., Nicolaou, P. & Christodoulou, K. Genomic and genetic studies of systemic sclerosis: A systematic review. *Hum Immunol* **78**, 153-165, doi:10.1016/j.humimm.2016.10.017 (2017).

11 Rothwell, S. *et al.* Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. *Ann Rheum Dis* **75**, 1558-1566, doi:10.1136/annrheumdis-2015-208119 (2016).

12 Kim, K. *et al.* The HLA-DRbeta1 amino acid positions 11-13-26 explain the majority of SLE-MHC associations. *Nat Commun* **5**, 5902, doi:10.1038/ncomms6902 (2014).

13 Wellcome Trust Case Control, C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-678, doi:10.1038/nature05911 (2007).

14 Australo-Anglo-American Spondyloarthritis, C. *et al.* Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* **42**, 123-127, doi:10.1038/ng.513 (2010).

15 Deighton, C. M., Walker, D. J., Griffiths, I. D. & Roberts, D. F. The contribution of HLA to rheumatoid arthritis. *Clin Genet* **36**, 178-182 (1989).

16 Rubin, L. A. *et al.* Investigating the genetic basis for ankylosing spondylitis. Linkage studies with the major histocompatibility complex region. *Arthritis Rheum* **37**, 1212-1220 (1994).

17 Diogo, D. *et al.* Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. *Am J Hum Genet* **92**, 15-27, doi:10.1016/j.ajhg.2012.11.012 (2013).

18 Stahl, E. A. *et al.* Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet* **44**, 483-489 (2012).

19 Cortes, A. *et al.* Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nature genetics* **45**, 730-738, doi:10.1038/ng.2667 (2013).

20 Angelotti, F. *et al.* One year in review 2017: pathogenesis of rheumatoid arthritis. *Clin Exp Rheumatol* **35**, 368-378 (2017).
21 Yewdell, J. W. & Bennink, J. R. The binary logic of antigen processing and presentation to T cells. *Cell* **62**, 203-206 (1990).
22 Trowsdale, J. & Knight, J. C. Major histocompatibility complex genomics and human disease. *Annu Rev Genomics Hum Genet* **14**, 301-323, doi:10.1146/annurev-genom-091212-153455 (2013).
23 Spottiswoode, C. N. & Busch, R. Vive la difference! Self/non-self recognition and the evolution of signatures of identity in arms races with parasites. *Phil. Trans. R. Soc. B* **374**, 20180206 (2019).
24 Rammensee, H. G., Friede, T. & Stevanoviic, S. MHC ligands and peptide motifs: first listing. *Immunogenetics* **41**, 178-228 (1995).
25 Vita, R. *et al.* The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* **43**, D405-412, doi:10.1093/nar/gku938 (2015).
26 Jensen, K. K. *et al.* Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* **154**, 394-406, doi:10.1111/imm.12889 (2018).
27 Karosiene, E., Lundegaard, C., Lund, O. & Nielsen, M. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* **64**, 177-186, doi:10.1007/s00251-011-0579-8 (2012).
28 Beck, S. & Trowsdale, J. The human major histocompatability complex: lessons from the DNA sequence. *Annu Rev Genomics Hum Genet* **1**, 117-137, doi:10.1146/annurev.genom.1.1.117 (2000).
29 Stewart, C. A. *et al.* Complete MHC haplotype sequencing for common disease gene mapping. *Genome Res* **14**, 1176-1187, doi:10.1101/gr.2188104 (2004).
30 Cortes, A. *et al.* Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. *Nat Commun* **6**, 7146, doi:10.1038/ncomms8146 (2015).
31 Hinks, A. *et al.* Fine-mapping the MHC locus in juvenile idiopathic arthritis (JIA) reveals genetic heterogeneity corresponding to distinct adult inflammatory arthritic diseases. *Ann Rheum Dis* **76**, 765-772 (2017).
32 Raychaudhuri, S. *et al.* Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet* **44**, 291-296 (2012).
33 Okada, Y. *et al.* Risk for ACPA-positive rheumatoid arthritis is driven by shared HLA amino acid polymorphisms in Asian and European populations. *Hum Mol Genet* **23**, 6916-6926 (2014).
34 Raj, P. *et al.* Regulatory polymorphisms modulate the expression of HLA class II molecules and promote autoimmunity. *Elife* **5**, doi:10.7554/eLife.12089 (2016).
35 Mayes, M. D. *et al.* Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. *Am J Hum Genet* **94**, 47-61, doi:10.1016/j.ajhg.2013.12.002 (2014).
36 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, doi:10.1038/ng.873 (2011).
37 Okada, Y. *et al.* Contribution of a Non-classical HLA Gene, HLA-DOA, to the Risk of Rheumatoid Arthritis. *Am J Hum Genet* **99**, 366-374, doi:10.1016/j.ajhg.2016.06.019 (2016).
38 Morgan, A. W. *et al.* The shared epitope hypothesis in rheumatoid arthritis: evaluation of alternative classification criteria in a large UK Caucasian cohort. *Arthritis Rheum* **58**, 1275-1283, doi:10.1002/art.23432 (2008).
39 Nepom, B. S. *et al.* Specific HLA-DR4-associated histocompatibility molecules characterize patients with seropositive juvenile rheumatoid arthritis. *J Clin Invest* **74**, 287-291, doi:10.1172/JCI111413 (1984).

40 Jaakkola, E. *et al.* Finnish HLA studies confirm the increased risk conferred by HLA-B27 homozygosity in ankylosing spondylitis. *Ann Rheum Dis* **65**, 775-780, doi:10.1136/ard.2005.041103 (2006).
41 Murray, K. J. *et al.* Age-specific effects of juvenile rheumatoid arthritis-associated HLA alleles. *Arthritis Rheum* **42**, 1843-1853, doi:10.1002/1529-0131(199909)42:9<1843::AID-ANR8>3.0.CO;2-M (1999).
42 Furukawa, H. *et al.* The role of common protective alleles HLA-DRB1*13 among systemic autoimmune diseases. *Genes Immun* **18**, 1-7, doi:10.1038/gene.2016.40 (2017).
43 Hinks, A. *et al.* Brief Report: The Genetic Profile of Rheumatoid Factor-Positive Polyarticular Juvenile Idiopathic Arthritis Resembles That of Adult Rheumatoid Arthritis. *Arthritis Rheumatol* **70**, 957-962, doi:10.1002/art.40443 (2018).
44 Aletaha, D. *et al.* 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* **62**, 2569-2581, doi:10.1002/art.27584 (2010).
45 MacGregor, A. J. *et al.* Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum* **43**, 30-37, doi:10.1002/1529-0131(200001)43:1<30::AID-ANR5>3.0.CO;2-B (2000).
46 van der Woude, D. *et al.* Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. *Arthritis Rheum* **60**, 916-923, doi:10.1002/art.24385 (2009).
47 Terao, C. *et al.* A twin study of rheumatoid arthritis in the Japanese population. *Mod Rheumatol* **26**, 685-689, doi:10.3109/14397595.2015.1135856 (2016).
48 Han, B. *et al.* Fine mapping seronegative and seropositive rheumatoid arthritis to shared and distinct HLA alleles by adjusting for the effects of heterogeneity. *Am J Hum Genet* **94**, 522-532, doi:10.1016/j.ajhg.2014.02.013 (2014).
49 Dessen, A., Lawrence, C. M., Cupo, S., Zaller, D. M. & Wiley, D. C. X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II. *Immunity* **7**, 473-481 (1997).
50 Wordsworth, B. P. *et al.* HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. *Proc Natl Acad Sci U S A* **86**, 10049-10053 (1989).
51 Gao, X. J. *et al.* A variant of HLA-DR4 determines susceptibility to rheumatoid arthritis in a subset of Israeli Jews. *Arthritis Rheum* **34**, 547-551 (1991).
52 van der Woude, D. *et al.* Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in four European populations. *Arthritis Rheum* **62**, 1236-1245, doi:10.1002/art.27366 (2010).
53 van der Helm-van Mil, A. H. *et al.* An independent role of protective HLA class II alleles in rheumatoid arthritis severity and susceptibility. *Arthritis Rheum* **52**, 2637-2644, doi:10.1002/art.21272 (2005).
54 Mackie, S. L. *et al.* A spectrum of susceptibility to rheumatoid arthritis within HLA-DRB1: stratification by autoantibody status in a large UK population. *Genes Immun* **13**, 120-128, doi:10.1038/gene.2011.60 (2012).
55 Harney, S., Newton, J., Milicic, A., Brown, M. A. & Wordsworth, B. P. Non-inherited maternal HLA alleles are associated with rheumatoid arthritis. *Rheumatology (Oxford)* **42**, 171-174 (2003).
56 Hsieh, H. J. *et al.* Using the maternal-fetal genotype incompatibility test to assess non-inherited maternal HLA-DRB1 antigen coding alleles as rheumatoid arthritis risk factors. *BMC Proc* **1 Suppl 1**, S124 (2007).

57 Sparks, J. A. & Karlson, E. W. The Roles of Cigarette Smoking and the Lung in the Transitions Between Phases of Preclinical Rheumatoid Arthritis. *Curr Rheumatol Rep* **18**, 15, doi:10.1007/s11926-016-0563-2 (2016).
58 Murphy, D., Mattey, D. & Hutchinson, D. Anti-citrullinated protein antibody positive rheumatoid arthritis is primarily determined by rheumatoid factor titre and the shared epitope rather than smoking per se. *PLoS One* **12**, e0180655, doi:10.1371/journal.pone.0180655 (2017).
59 Kim, K. *et al.* Interactions between amino acid-defined major histocompatibility complex class II variants and smoking in seropositive rheumatoid arthritis. *Arthritis Rheumatol* **67**, 2611-2623 (2015).
60 Sakkas, L. I., Daoussis, D., Liossis, S. N. & Bogdanos, D. P. The Infectious Basis of ACPA-Positive Rheumatoid Arthritis. *Front Microbiol* **8**, 1853, doi:10.3389/fmicb.2017.01853 (2017).
61 Lundberg, K., Wegner, N., Yucel-Lindberg, T. & Venables, P. J. Periodontitis in RA-the citrullinated enolase connection. *Nat Rev Rheumatol* **6**, 727-730, doi:10.1038/nrrheum.2010.139 (2010).
62 Wang, S. & Wang, Y. Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis. *Biochim Biophys Acta* **1829**, 1126-1135, doi:10.1016/j.bbagrm.2013.07.003 (2013).
63 Rothbard, J. B. *et al.* Reversal of HLA restriction by a point mutation in an antigenic peptide. *Int Immunol* **1**, 487-495 (1989).
64 Busch, R. *The interaction of peptides with human class II histocompatibility antigens* PhD thesis, University of London, (1992).
65 Sette, A. *et al.* HLA DR4w4-binding motifs illustrate the biochemical basis of degeneracy and specificity in peptide-DR interactions. *J Immunol* **151**, 3163-3170 (1993).
66 Hill, J. A. *et al.* Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J Immunol* **171**, 538-541 (2003).
67 Scally, S. W. *et al.* A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med* **210**, 2569-2582, doi:10.1084/jem.20131241 (2013).
68 Ting, Y. T. *et al.* The interplay between citrullination and HLA-DRB1 polymorphism in shaping peptide binding hierarchies in Rheumatoid Arthritis. *J Biol Chem* (2018).
69 James, E. A. *et al.* HLA-DR1001 presents "altered-self" peptides derived from joint-associated proteins by accepting citrulline in three of its binding pockets. *Arthritis Rheum* **62**, 2909-2918 (2010).
70 Sidney, J. *et al.* Citrullination only infrequently impacts peptide binding to HLA class II MHC. *PLoS One* **12**, e0177140 (2017).
71 Dendrou, C. A., Petersen, J., Rossjohn, J. & Fugger, L. HLA variation and disease. *Nat Rev Immunol*, doi:10.1038/nri.2017.143 (2018).
72 van Heemst, J. *et al.* Crossreactivity to vinculin and microbes provides a molecular basis for HLA-based protection against rheumatoid arthritis. *Nat Commun* **6**, 6681, doi:10.1038/ncomms7681 (2015).
73 Shi, J. *et al.* Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proc Natl Acad Sci U S A* **108**, 17372-17377, doi:10.1073/pnas.1114465108 (2011).
74 Jiang, X. *et al.* Anti-CarP antibodies in two large cohorts of patients with rheumatoid arthritis and their relationship to genetic risk factors, cigarette smoking and other autoantibodies. *Ann Rheum Dis* **73**, 1761-1768, doi:10.1136/annrheumdis-2013-205109 (2014).

75 Busch, R. *et al.* On the perils of poor editing: regulation of peptide loading by HLA-DQ and H2-A molecules associated with celiac disease and type 1 diabetes. *Expert Rev Mol Med* **14**, e15, doi:10.1017/erm.2012.9 (2012).
76 Roep, B. O., Kracht, M. J., van Lummel, M. & Zaldumbide, A. A roadmap of the generation of neoantigens as targets of the immune system in type 1 diabetes. *Curr Opin Immunol* **43**, 67-73, doi:10.1016/j.coi.2016.09.007 (2016).
77 Lu, Y. C., Weng, W. C. & Lee, H. Functional roles of calreticulin in cancer biology. *Biomed Res Int* **2015**, 526524, doi:10.1155/2015/526524 (2015).
78 Ling, S., Cheng, A., Pumpens, P., Michalak, M. & Holoshitz, J. Identification of the rheumatoid arthritis shared epitope binding site on calreticulin. *PLoS One* **5**, e11703, doi:10.1371/journal.pone.0011703 (2010).
79 Ling, S., Pi, X. & Holoshitz, J. The rheumatoid arthritis shared epitope triggers innate immune signaling via cell surface calreticulin. *J Immunol* **179**, 6359-6367 (2007).
80 De Almeida, D. E. *et al.* Immune dysregulation by the rheumatoid arthritis shared epitope. *J Immunol* **185**, 1927-1934, doi:10.4049/jimmunol.0904002 (2010).
81 Fu, J. *et al.* A small shared epitope-mimetic compound potently accelerates osteoclast-mediated bone damage in autoimmune arthritis. *J Immunol* **191**, 2096-2103, doi:10.4049/jimmunol.1203231 (2013).
82 Holoshitz, J. *et al.* An HLA-DRB1-coded signal transduction ligand facilitates inflammatory arthritis: a new mechanism of autoimmunity. *J Immunol* **190**, 48-57, doi:10.4049/jimmunol.1202150 (2013).
83 Ling, S., Cline, E. N., Haug, T. S., Fox, D. A. & Holoshitz, J. Citrullinated calreticulin potentiates rheumatoid arthritis shared epitope signaling. *Arthritis Rheum* **65**, 618-626, doi:10.1002/art.37814 (2013).
84 Patil, N. S. *et al.* Rheumatoid arthritis (RA)-associated HLA-DR alleles form less stable complexes with class II-associated invariant chain peptide than non-RA-associated HLA-DR alleles. *J Immunol* **167**, 7157-7168 (2001).
85 Busch, R. *et al.* Achieving stability through editing and chaperoning: regulation of MHC class II peptide binding and expression. *Immunol Rev* **207**, 242-260 (2005).
86 Rinderknecht, C. H. *et al.* DM influences the abundance of major histocompatibility complex class II alleles with low affinity for class II-associated invariant chain peptides via multiple mechanisms. *Immunology* **131**, 18-32, doi:10.1111/j.1365-2567.2010.03282.x (2010).
87 Anczurowski, M. & Hirano, N. Mechanisms of HLA-DP Antigen Processing and Presentation Revisited. *Trends Immunol* **39**, 960-964, doi:10.1016/j.it.2018.10.008 (2018).
88 Anczurowski, M. *et al.* Mechanisms underlying the lack of endogenous processing and CLIP-mediated binding of the invariant chain by HLA-DP(84Gly). *Sci Rep* **8**, 4804, doi:10.1038/s41598-018-22931-4 (2018).
89 Yamashita, Y. *et al.* HLA-DP(84Gly) constitutively presents endogenous peptides generated by the class I antigen processing pathway. *Nat Commun* **8**, 15244, doi:10.1038/ncomms15244 (2017).
90 Ito, Y. *et al.* Rapid CLIP dissociation from MHC II promotes an unusual antigen presentation pathway in autoimmunity. *J Exp Med*, doi:10.1084/jem.20180300 (2018).
91 Scher, J. U. *et al.* Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. *Elife* **2**, e01202, doi:10.7554/eLife.01202 (2013).
92 Gomez, A. *et al.* Loss of sex and age driven differences in the gut microbiome characterize arthritis-susceptible 0401 mice but not arthritis-resistant 0402 mice. *PLoS One* **7**, e36095, doi:10.1371/journal.pone.0036095 (2012).
93 Khan, M. A. Polymorphism of HLA-B27: 105 subtypes currently known. *Curr Rheumatol Rep* **15**, 362, doi:10.1007/s11926-013-0362-y (2013).

94 Sieper, J. & van der Heijde, D. Review: Nonradiographic axial spondyloarthritis: new definition of an old disease? *Arthritis and rheumatism* **65**, 543-551, doi:10.1002/art.37803 (2013).
95 Braem, K. & Lories, R. J. Insights into the pathophysiology of ankylosing spondylitis: contributions from animal models. *Joint Bone Spine* **79**, 243-248, doi:10.1016/j.jbspin.2011.09.008 (2012).
96 Brown, M. A. *et al.* Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum* **40**, 1823-1828, doi:10.1002/1529-0131(199710)40:10<1823::AID-ART15>3.0.CO;2-1 (1997).
97 Evnouchidou, I. *et al.* 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**, 1909-1913, doi:10.4049/jimmunol.1003337 (2011).
98 Kochan, G. *et al.* Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc Natl Acad Sci U S A* **108**, 7745-7750, doi:10.1073/pnas.1101262108 (2011).
99 Chen, L. *et al.* Critical role of endoplasmic reticulum aminopeptidase 1 in determining the length and sequence of peptides bound and presented by HLA-B27. *Arthritis Rheumatol* **66**, 284-294, doi:10.1002/art.38249 (2014).
100 Wei, J. C. *et al.* Interaction between HLA-B60 and HLA-B27 as a Better Predictor of Ankylosing Spondylitis in a Taiwanese Population. *PLoS One* **10**, e0137189, doi:10.1371/journal.pone.0137189 (2015).
101 van Gaalen, F. A. *et al.* Epistasis between two HLA antigens defines a subset of individuals at a very high risk for ankylosing spondylitis. *Ann Rheum Dis* **72**, 974-978, doi:10.1136/annrheumdis-2012-201774 (2013).
102 Gaston, J. S. H. & Jadon, D. R. Th17 cell responses in spondyloarthritis. *Best Pract Res Clin Rheumatol* **31**, 777-796, doi:10.1016/j.berh.2018.07.010 (2017).
103 Neerinckx, B., Kollnberger, S., Shaw, J. & Lories, R. No evidence for a direct role of HLA-B27 in pathological bone formation in axial SpA. *RMD open* **3**, e000451, doi:10.1136/rmdopen-2017-000451 (2017).
104 Li, Z. *et al.* Epigenetic and gene expression analysis of ankylosing spondylitis-associated loci implicate immune cells and the gut in the disease pathogenesis. *Genes and immunity*, doi:10.1038/gene.2017.11 (2017).
105 Crux, N. B. & Elahi, S. Human Leukocyte Antigen (HLA) and Immune Regulation: How Do Classical and Non-Classical HLA Alleles Modulate Immune Response to Human Immunodeficiency Virus and Hepatitis C Virus Infections? *Frontiers in immunology* **8**, 832, doi:10.3389/fimmu.2017.00832 (2017).
106 Appel, H. *et al.* Use of HLA-B27 tetramers to identify low-frequency antigen-specific T cells in Chlamydia-triggered reactive arthritis. *Arthritis research & therapy* **6**, R521-534, doi:10.1186/ar1221 (2004).
107 Fiorillo, M. T. *et al.* Allele-dependent similarity between viral and self-peptide presentation by HLA-B27 subtypes. *The Journal of biological chemistry* **280**, 2962-2971, doi:10.1074/jbc.M410807200 (2005).
108 Narzi, D. *et al.* Dynamical characterization of two differentially disease associated MHC class I proteins in complex with viral and self-peptides. *Journal of molecular biology* **415**, 429-442, doi:10.1016/j.jmb.2011.11.021 (2012).
109 Frauendorf, E., von Goessel, H., May, E. & Marker-Hermann, E. HLA-B27-restricted T cells from patients with ankylosing spondylitis recognize peptides from B*2705 that are similar to bacteria-derived peptides. *Clinical and experimental immunology* **134**, 351-359 (2003).
110 Hulsmeyer, M. *et al.* Dual, HLA-B27 subtype-dependent conformation of a self-peptide. *J Exp Med* **199**, 271-281 (2004).

111 Schittenhelm, R. B., Sian, T. C., Wilmann, P. G., Dudek, N. L. & Purcell, A. W. Revisiting the arthritogenic peptide theory: quantitative not qualitative changes in the peptide repertoire of HLA-B27 allotypes. *Arthritis & rheumatology* **67**, 702-713, doi:10.1002/art.38963 (2015).
112 Faham, M. *et al.* Discovery of T Cell Receptor beta Motifs Specific to HLA-B27-Positive Ankylosing Spondylitis by Deep Repertoire Sequence Analysis. *Arthritis & rheumatology* **69**, 774-784, doi:10.1002/art.40028 (2017).
113 Duchmann, R., Lambert, C., May, E., Hohler, T. & Marker-Hermann, E. CD4+ and CD8+ clonal T cell expansions indicate a role of antigens in ankylosing spondylitis; a study in HLA-B27+ monozygotic twins. *Clinical and experimental immunology* **123**, 315-322 (2001).
114 Dulphy, N. *et al.* Common intra-articular T cell expansions in patients with reactive arthritis: identical beta-chain junctional sequences and cytotoxicity toward HLA-B27. *J Immunol* **162**, 3830-3839 (1999).
115 Menon, B. *et al.* Interleukin-17+CD8+ T cells are enriched in the joints of patients with psoriatic arthritis and correlate with disease activity and joint damage progression. *Arthritis Rheumatol* **66**, 1272-1281, doi:10.1002/art.38376 (2014).
116 Okada, Y. *et al.* Fine mapping major histocompatibility complex associations in psoriasis and its clinical subtypes. *Am J Hum Genet* **95**, 162-172, doi:10.1016/j.ajhg.2014.07.002 (2014).
117 Bowes, J. *et al.* Cross-phenotype association mapping of the MHC identifies genetic variants that differentiate psoriatic arthritis from psoriasis. *Ann Rheum Dis* **76**, 1774-1779, doi:10.1136/annrheumdis-2017-211414 (2017).
118 Buxton, S. E., Benjamin, R. J., Clayberger, C., Parham, P. & Krensky, A. M. Anchoring pockets in human histocompatibility complex leukocyte antigen (HLA) class I molecules: analysis of the conserved B ("45") pocket of HLA-B27. *J Exp Med* **175**, 809-820 (1992).
119 Taurog, J. D. *et al.* Spondylarthritis in HLA-B27/human beta2-microglobulin-transgenic rats is not prevented by lack of CD8. *Arthritis and rheumatism* **60**, 1977-1984, doi:10.1002/art.24599 (2009).
120 Lenart, I. *et al.* The MHC Class I heavy chain structurally conserved cysteines 101 and 164 participate in HLA-B27 dimer formation. *Antioxid Redox Signal* **16**, 33-43, doi:10.1089/ars.2010.3693 (2012).
121 Bird, L. A. *et al.* Lymphoblastoid cells express HLA-B27 homodimers both intracellularly and at the cell surface following endosomal recycling. *European journal of immunology* **33**, 748-759, doi:10.1002/eji.200323678 (2003).
122 Kollnberger, S. *et al.* Cell-surface expression and immune receptor recognition of HLA-B27 homodimers. *Arthritis and rheumatism* **46**, 2972-2982, doi:10.1002/art.10605 (2002).
123 MacLean, L. *et al.* Sulphydryl reactivity of the HLA-B27 epitope: accessibility of the free cysteine studied by flow cytometry. *Annals of the rheumatic diseases* **51**, 456-460 (1992).
124 Payeli, S. K. *et al.* Inhibiting HLA-B27 homodimer-driven immune cell inflammation in spondylarthritis. *Arthritis and rheumatism* **64**, 3139-3149, doi:10.1002/art.34538 (2012).
125 McHugh, K. *et al.* Expression of aberrant HLA-B27 molecules is dependent on B27 dosage and peptide supply. *Annals of the rheumatic diseases* **73**, 763-770, doi:10.1136/annrheumdis-2012-203080 (2014).
126 Goodridge, J. P. *et al.* HLA-F and MHC-I open conformers cooperate in a MHC-I antigen cross-presentation pathway. *J Immunol* **191**, 1567-1577, doi:10.4049/jimmunol.1300080 (2013).
127 Raine, T. *et al.* Consistent patterns of expression of HLA class I free heavy chains in healthy individuals and raised expression in spondyloarthropathy patients point to physiological and pathological roles. *Rheumatology (Oxford)* **45**, 1338-1344 (2006).

128 Tsai, W. C. *et al.* Free HLA class I heavy chain-carrying monocytes--a potential role in the pathogenesis of spondyloarthropathies. *J Rheumatol* **29**, 966-972 (2002).
129 Chen, L. *et al.* Silencing or inhibition of endoplasmic reticulum aminopeptidase 1 (ERAP1) suppresses free heavy chain expression and Th17 responses in ankylosing spondylitis. *Annals of the rheumatic diseases* **75**, 916-923, doi:10.1136/annrheumdis-2014-206996 (2016).
130 Haroon, N., Tsui, F. W., Uchanska-Ziegler, B., Ziegler, A. & Inman, R. D. Endoplasmic reticulum aminopeptidase 1 (ERAP1) exhibits functionally significant interaction with HLA-B27 and relates to subtype specificity in ankylosing spondylitis. *Annals of the rheumatic diseases* **71**, 589-595, doi:10.1136/annrheumdis-2011-200347 (2012).
131 Tran, T. M., Hong, S., Edwan, J. H. & Colbert, R. A. ERAP1 reduces accumulation of aberrant and disulfide-linked forms of HLA-B27 on the cell surface. *Molecular immunology* **74**, 10-17, doi:10.1016/j.molimm.2016.04.002 (2016).
132 Giles, J. *et al.* HLA-B27 homodimers and free H chains are stronger ligands for leukocyte Ig-like receptor B2 than classical HLA class I. *J Immunol* **188**, 6184-6193, doi:10.4049/jimmunol.1102711 (2012).
133 Zhang, Z. *et al.* The Leukocyte Immunoglobulin-Like Receptor Family Member LILRB5 Binds to HLA-Class I Heavy Chains. *PloS one* **10**, e0129063, doi:10.1371/journal.pone.0129063 (2015).
134 Cauli, A. *et al.* The arthritis-associated HLA-B*27:05 allele forms more cell surface B27 dimer and free heavy chain ligands for KIR3DL2 than HLA-B*27:09. *Rheumatology*, doi:10.1093/rheumatology/ket219 (2013).
135 Gaur, P., Misra, R. & Aggarwal, A. Natural killer cell and gamma delta T cell alterations in enthesitis related arthritis category of juvenile idiopathic arthritis. *Clinical immunology* **161**, 163-169, doi:10.1016/j.clim.2015.07.012 (2015).
136 Chan, A. T., Kollnberger, S. D., Wedderburn, L. R. & Bowness, P. Expansion and enhanced survival of natural killer cells expressing the killer immunoglobulin-like receptor KIR3DL2 in spondylarthritis. *Arthritis and rheumatism* **52**, 3586-3595, doi:10.1002/art.21395 (2005).
137 Goodridge, J. P., Burian, A., Lee, N. & Geraghty, D. E. HLA-F and MHC class I open conformers are ligands for NK cell Ig-like receptors. *J Immunol* **191**, 3553-3562, doi:10.4049/jimmunol.1300081 (2013).
138 Santos, M. R. *et al.* Non-classical human leucocyte antigens in ankylosing spondylitis: possible association with HLA-E and HLA-F. *RMD open* **4**, e000677, doi:10.1136/rmdopen-2018-000677 (2018).
139 Hatano, H. *et al.* The D0 Ig-like domain plays a central role in the stronger binding of KIR3DL2 to B27 free H chain dimers. *J Immunol* **194**, 1591-1601, doi:10.4049/jimmunol.1402214 (2015).
140 Borges, L., Hsu, M. L., Fanger, N., Kubin, M. & Cosman, D. A family of human lymphoid and myeloid Ig-like receptors, some of which bind to MHC class I molecules. *J Immunol* **159**, 5192-5196 (1997).
141 Hogan, L. E., Jones, D. C. & Allen, R. L. Expression of the innate immune receptor LILRB5 on monocytes is associated with mycobacteria exposure. *Sci Rep* **6**, 21780, doi:10.1038/srep21780 (2016).
142 Mori, Y. *et al.* Inhibitory immunoglobulin-like receptors LILRB and PIR-B negatively regulate osteoclast development. *J Immunol* **181**, 4742-4751 (2008).
143 Jones, D. C. *et al.* Alternative mRNA splicing creates transcripts encoding soluble proteins from most LILR genes. *European journal of immunology* **39**, 3195-3206, doi:10.1002/eji.200839080 (2009).
144 Ridley, A. *et al.* Activation-Induced Killer Cell Immunoglobulin-like Receptor 3DL2 Binding to HLA-B27 Licenses Pathogenic T Cell Differentiation in Spondyloarthritis. *Arthritis & rheumatology* **68**, 901-914, doi:10.1002/art.39515 (2016).

145 Rysnik, O. *et al.* Non-conventional forms of HLA-B27 are expressed in spondyloarthritis joints and gut tissue. *Journal of autoimmunity* **70**, 12-21, doi:10.1016/j.jaut.2016.03.009 (2016).
146 Bowness, P. *et al.* Th17 cells expressing KIR3DL2+ and responsive to HLA-B27 homodimers are increased in ankylosing spondylitis. *J Immunol* **186**, 2672-2680, doi:10.4049/jimmunol.1002653 (2011).
147 Wong-Baeza, I. *et al.* KIR3DL2 Binds to HLA-B27 Dimers and Free H Chains More Strongly than Other HLA Class I and Promotes the Expansion of T Cells in Ankylosing Spondylitis. *J Immunol* **190**, 3216-3224, doi:10.4049/jimmunol.1202926 (2013).
148 Purvis, H. A. *et al.* Low-strength T-cell activation promotes Th17 responses. *Blood* **116**, 4829-4837, doi:10.1182/blood-2010-03-272153 (2010).
149 Ruutu, M. *et al.* beta-glucan triggers spondylarthritis and Crohn's disease-like ileitis in SKG mice. *Arthritis and rheumatism* **64**, 2211-2222, doi:10.1002/art.34423 (2012).
150 Navid, F. & Colbert, R. A. Causes and consequences of endoplasmic reticulum stress in rheumatic disease. *Nature reviews. Rheumatology* **13**, 25-40, doi:10.1038/nrrheum.2016.192 (2017).
151 Dangoria, N. S. *et al.* HLA-B27 misfolding is associated with aberrant intermolecular disulfide bond formation (dimerization) in the endoplasmic reticulum. *J Biol Chem* **277**, 23459-23468 (2002).
152 Guiliano, D. B. *et al.* Polymorphisms in the F Pocket of HLA-B27 Subtypes Strongly Affect Assembly, Chaperone Interactions, and Heavy-Chain Misfolding. *Arthritis & rheumatology* **69**, 610-621, doi:10.1002/art.39948 (2017).
153 Dong, W. *et al.* Upregulation of 78-kDa glucose-regulated protein in macrophages in peripheral joints of active ankylosing spondylitis. *Scandinavian journal of rheumatology* **37**, 427-434, doi:10.1080/03009740802213310 (2008).
154 Jeanty, C. *et al.* HLA-B27 subtype oligomerization and intracellular accumulation patterns correlate with predisposition to spondyloarthritis. *Arthritis & rheumatology* **66**, 2113-2123, doi:10.1002/art.38644 (2014).
155 Goodall, J. C. *et al.* Endoplasmic reticulum stress-induced transcription factor, CHOP, is crucial for dendritic cell IL-23 expression. *Proc Natl Acad Sci U S A* **107**, 17698-17703, doi:10.1073/pnas.1011736107 (2010).
156 DeLay, M. L. *et al.* HLA-B27 misfolding and the unfolded protein response augment interleukin-23 production and are associated with Th17 activation in transgenic rats. *Arthritis Rheum* **60**, 2633-2643 (2009).
157 Layh-Schmitt, G., Yang, E. Y., Kwon, G. & Colbert, R. A. HLA-B27 alters the response to tumor necrosis factor alpha and promotes osteoclastogenesis in bone marrow monocytes from HLA-B27-transgenic rats. *Arthritis and rheumatism* **65**, 2123-2131, doi:10.1002/art.38001 (2013).
158 Neerinckx, B., Carter, S. & Lories, R. J. No evidence for a critical role of the unfolded protein response in synovium and blood of patients with ankylosing spondylitis. *Annals of the rheumatic diseases* **73**, 629-630, doi:10.1136/annrheumdis-2013-204170 (2014).
159 Gu, J. *et al.* Clues to pathogenesis of spondyloarthropathy derived from synovial fluid mononuclear cell gene expression profiles. *J Rheumatol* **29**, 2159-2164 (2002).
160 Zeng, L., Lindstrom, M. J. & Smith, J. A. Ankylosing spondylitis macrophage production of higher levels of interleukin-23 in response to lipopolysaccharide without induction of a significant unfolded protein response. *Arthritis and rheumatism* **63**, 3807-3817, doi:10.1002/art.30593 (2011).
161 Ciccia, F. *et al.* Evidence that autophagy, but not the unfolded protein response, regulates the expression of IL-23 in the gut of patients with ankylosing spondylitis and subclinical gut inflammation. *Annals of the rheumatic diseases* **73**, 1566-1574, doi:10.1136/annrheumdis-2012-202925 (2014).

162 Neerinckx, B., Carter, S. & Lories, R. IL-23 expression and activation of autophagy in synovium and PBMCs of HLA-B27 positive patients with ankylosing spondylitis. Response to: 'Evidence that autophagy, but not the unfolded protein response, regulates the expression of IL-23 in the gut of patients with ankylosing spondylitis and subclinical gut inflammation' by Ciccia et al. *Annals of the rheumatic diseases* **73**, e68, doi:10.1136/annrheumdis-2014-206277 (2014).
163 Kenna, T. J. *et al.* Disease-associated polymorphisms in ERAP1 do not alter endoplasmic reticulum stress in patients with ankylosing spondylitis. *Genes and immunity* **16**, 35-42, doi:10.1038/gene.2014.62 (2015).
164 Penttinen, M. A. *et al.* Enhanced intracellular replication of Salmonella enteritidis in HLA-B27-expressing human monocytic cells: dependency on glutamic acid at position 45 in the B pocket of HLA-B27. *Arthritis and rheumatism* **50**, 2255-2263, doi:10.1002/art.20336 (2004).
165 Antoniou, A. N. *et al.* Salmonella exploits HLA-B27 and host unfolded protein responses to promote intracellular replication. *Annals of the rheumatic diseases*, doi:10.1136/annrheumdis-2018-213532 (2018).
166 Ansalone, C., Utriainen, L., Milling, S. & Goodyear, C. S. Role of Gut Inflammation in Altering the Monocyte Compartment and Its Osteoclastogenic Potential in HLA-B27-Transgenic Rats. *Arthritis & rheumatology* **69**, 1807-1815, doi:10.1002/art.40154 (2017).
167 Manasson, J. *et al.* Gut Microbiota Perturbations in Reactive Arthritis and Post-Infectious Spondyloarthritis. *Arthritis & rheumatology*, doi:10.1002/art.40359 (2017).
168 Yu, D. & Kuipers, J. G. Role of bacteria and HLA-B27 in the pathogenesis of reactive arthritis. *Rheum Dis Clin North Am* **29**, 21-36, v-vi (2003).
169 Lin, P. *et al.* HLA-B27 and human beta2-microglobulin affect the gut microbiota of transgenic rats. *PloS one* **9**, e105684, doi:10.1371/journal.pone.0105684 (2014).
170 Breban, M. *et al.* Faecal microbiota study reveals specific dysbiosis in spondyloarthritis. *Annals of the rheumatic diseases* **76**, 1614-1622, doi:10.1136/annrheumdis-2016-211064 (2017).
171 Gill, T., Asquith, M., Brooks, S. R., Rosenbaum, J. T. & Colbert, R. A. Effects of HLA-B27 on Gut Microbiota in Experimental Spondyloarthritis Implicate an Ecological Model of Dysbiosis. *Arthritis & rheumatology* **70**, 555-565, doi:10.1002/art.40405 (2018).
172 Ombrello, M. J. *et al.* HLA-DRB1*11 and variants of the MHC class II locus are strong risk factors for systemic juvenile idiopathic arthritis. *Proc Natl Acad Sci U S A* **112**, 15970-15975, doi:10.1073/pnas.1520779112 (2015).
173 Pardeo, M., Bracaglia, C. & De Benedetti, F. Systemic juvenile idiopathic arthritis: New insights into pathogenesis and cytokine directed therapies. *Best Pract Res Clin Rheumatol* **31**, 505-516, doi:10.1016/j.berh.2018.02.002 (2017).
174 Mahmud, S. A. & Binstadt, B. A. Autoantibodies in the Pathogenesis, Diagnosis, and Prognosis of Juvenile Idiopathic Arthritis. *Front Immunol* **9**, 3168, doi:10.3389/fimmu.2018.03168 (2018).
175 Ruperto, N. *et al.* Abatacept in children with juvenile idiopathic arthritis: a randomised, double-blind, placebo-controlled withdrawal trial. *Lancet* **372**, 383-391, doi:10.1016/S0140-6736(08)60998-8 (2008).
176 Omoyinmi, E. *et al.* Th1 and Th17 cell subpopulations are enriched in the peripheral blood of patients with systemic juvenile idiopathic arthritis. *Rheumatology (Oxford)* **51**, 1881-1886, doi:10.1093/rheumatology/kes162 (2012).
177 Raziuddin, S., Bahabri, S., Al-Dalaan, A., Siraj, A. K. & Al-Sedairy, S. A mixed Th1/Th2 cell cytokine response predominates in systemic onset juvenile rheumatoid arthritis: immunoregulatory IL-10 function. *Clin Immunol Immunopathol* **86**, 192-198 (1998).

178 Hugle, B. *et al.* Transcription factor motif enrichment in whole transcriptome analysis identifies STAT4 and BCL6 as the most prominent binding motif in systemic juvenile idiopathic arthritis. *Arthritis Res Ther* **20**, 98, doi:10.1186/s13075-018-1603-2 (2018).
179 Janow, G. *et al.* The Systemic Juvenile Idiopathic Arthritis Cohort of the Childhood Arthritis and Rheumatology Research Alliance Registry: 2010-2013. *J Rheumatol* **43**, 1755-1762, doi:10.3899/jrheum.150997 (2016).
180 Nigrovic, P. A. Review: is there a window of opportunity for treatment of systemic juvenile idiopathic arthritis? *Arthritis Rheumatol* **66**, 1405-1413, doi:10.1002/art.38615 (2014).
181 Zielinski, C. E. *et al.* Pathogen-induced human TH17 cells produce IFN-gamma or IL-10 and are regulated by IL-1beta. *Nature* **484**, 514-518, doi:10.1038/nature10957 (2012).
182 Nigrovic, P. A. Autoinflammation and autoimmunity in systemic juvenile idiopathic arthritis. *Proc Natl Acad Sci U S A* **112**, 15785-15786, doi:10.1073/pnas.1521837113 (2015).
183 Kissner, T. L. *et al.* Activation of MyD88 signaling upon staphylococcal enterotoxin binding to MHC class II molecules. *PLoS One* **6**, e15985, doi:10.1371/journal.pone.0015985 (2011).
184 Scholl, P. R. *et al.* Effect of isotypes and allelic polymorphism on the binding of staphylococcal exotoxins to MHC class II molecules. *J Immunol* **144**, 226-230 (1990).
185 Liu, X. *et al.* Intracellular MHC class II molecules promote TLR-triggered innate immune responses by maintaining activation of the kinase Btk. *Nat Immunol* **12**, 416-424, doi:10.1038/ni.2015 (2011).
186 Poddubnyy, D. *et al.* Development of an ASAS-endorsed recommendation for the early referral of patients with a suspicion of axial spondyloarthritis. *Ann Rheum Dis* **74**, 1483-1487, doi:10.1136/annrheumdis-2014-207151 (2015).
187 Winchester, R. *et al.* HLA associations reveal genetic heterogeneity in psoriatic arthritis and in the psoriasis phenotype. *Arthritis Rheum* **64**, 1134-1144, doi:10.1002/art.33415 (2012).
188 Queiro, R., Morante, I., Cabezas, I. & Acasuso, B. HLA-B27 and psoriatic disease: a modern view of an old relationship. *Rheumatology (Oxford)* **55**, 221-229, doi:10.1093/rheumatology/kev296 (2016).
189 FitzGerald, O., Haroon, M., Giles, J. T. & Winchester, R. Concepts of pathogenesis in psoriatic arthritis: genotype determines clinical phenotype. *Arthritis Res Ther* **17**, 115, doi:10.1186/s13075-015-0640-3 (2015).
190 Okada, Y., Eyre, S., Suzuki, A., Kochi, Y. & Yamamoto, K. Genetics of rheumatoid arthritis: 2018 status. *Ann Rheum Dis*, doi:10.1136/annrheumdis-2018-213678 (2018).
191 Wei, W. H. *et al.* Major histocompatibility complex harbors widespread genotypic variability of non-additive risk of rheumatoid arthritis including epistasis. *Sci Rep* **6**, 25014 (2016).
192 Wei, W. H., Loh, C. Y., Worthington, J. & Eyre, S. Immunochip Analyses of Epistasis in Rheumatoid Arthritis Confirm Multiple Interactions within MHC and Suggest Novel Non-MHC Epistatic Signals. *J Rheumatol* **43**, 839-845 (2016).
193 Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet* **19**, 581-590, doi:10.1038/s41576-018-0018-x (2018).
194 Alivernini, S. *et al.* Is ACPA positivity the main driver for rheumatoid arthritis treatment? Pros and cons. *Autoimmun Rev* **16**, 1096-1102, doi:10.1016/j.autrev.2017.09.002 (2017).
195 Choi, S. T. & Lee, K. H. Clinical management of seronegative and seropositive rheumatoid arthritis: A comparative study. *PLoS One* **13**, e0195550, doi:10.1371/journal.pone.0195550 (2018).
196 Lundstrom, E. *et al.* Opposing effects of HLA-DRB1*13 alleles on the risk of developing anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-

negative rheumatoid arthritis. *Arthritis Rheum* **60**, 924-930, doi:10.1002/art.24410 (2009).
197 Oka, S. *et al.* Association of human leukocyte antigen alleles with chronic lung diseases in rheumatoid arthritis. *Rheumatology (Oxford)* **55**, 1301-1307 (2016).
198 Hillarby, M. C. *et al.* HLA associations in subjects with rheumatoid arthritis and bronchiectasis but not with other pulmonary complications of rheumatoid disease. *Br J Rheumatol* **32**, 794-797 (1993).
199 Parkes, J. E. *et al.* Genetic background may contribute to the latitude-dependent prevalence of dermatomyositis and anti-TIF1-gamma autoantibodies in adult patients with myositis. *Arthritis Res Ther* **20**, 117, doi:10.1186/s13075-018-1617-9 (2018).
200 He, D. *et al.* Association of the HLA-DRB1 with scleroderma in Chinese population. *PLoS One* **9**, e106939, doi:10.1371/journal.pone.0106939 (2014).
201 Petty, R. E. *et al.* International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* **31**, 390-392 (2004).
202 Thomson, W. *et al.* Juvenile idiopathic arthritis classified by the ILAR criteria: HLA associations in UK patients. *Rheumatology (Oxford)* **41**, 1183-1189 (2002).
203 Hollenbach, J. A. *et al.* Juvenile idiopathic arthritis and HLA class I and class II interactions and age-at-onset effects. *Arthritis Rheum* **62**, 1781-1791, doi:10.1002/art.27424 (2010).
204 Prahalad, S. *et al.* Hierarchy of risk of childhood-onset rheumatoid arthritis conferred by HLA-DRB1 alleles encoding the shared epitope. *Arthritis Rheum* **64**, 925-930, doi:10.1002/art.33376 (2012).
205 Ravelli, A. *et al.* Patients with antinuclear antibody-positive juvenile idiopathic arthritis constitute a homogeneous subgroup irrespective of the course of joint disease. *Arthritis Rheum* **52**, 826-832, doi:10.1002/art.20945 (2005).
206 Nigrovic, P. A., Raychaudhuri, S. & Thompson, S. D. Review: Genetics and the Classification of Arthritis in Adults and Children. *Arthritis Rheumatol* **70**, 7-17, doi:10.1002/art.40350 (2018).
207 Willis, V. C. *et al.* Protein arginine deiminase 4 inhibition is sufficient for the amelioration of collagen-induced arthritis. *Clin Exp Immunol* **188**, 263-274, doi:10.1111/cei.12932 (2017).
208 Benham, H. *et al.* Citrullinated peptide dendritic cell immunotherapy in HLA risk genotype-positive rheumatoid arthritis patients. *Sci Transl Med* **7**, 290ra287, doi:10.1126/scitranslmed.aaa9301 (2015).
209 Baker, K. F. & Isaacs, J. D. Novel therapies for immune-mediated inflammatory diseases: What can we learn from their use in rheumatoid arthritis, spondyloarthritis, systemic lupus erythematosus, psoriasis, Crohn's disease and ulcerative colitis? *Ann Rheum Dis* **77**, 175-187, doi:10.1136/annrheumdis-2017-211555 (2018).
210 Baeten, D. *et al.* Secukinumab, an Interleukin-17A Inhibitor, in Ankylosing Spondylitis. *N Engl J Med* **373**, 2534-2548, doi:10.1056/NEJMoa1505066 (2015).
211 Baeten, D. *et al.* Risankizumab, an IL-23 inhibitor, for ankylosing spondylitis: results of a randomised, double-blind, placebo-controlled, proof-of-concept, dose-finding phase 2 study. *Ann Rheum Dis* **77**, 1295-1302, doi:10.1136/annrheumdis-2018-213328 (2018).
212 Marroquin Belaunzaran, O. *et al.* HLA-B27-Homodimer-Specific Antibody Modulates the Expansion of Pro-Inflammatory T-Cells in HLA-B27 Transgenic Rats. *PloS one* **10**, e0130811, doi:10.1371/journal.pone.0130811 (2015).
213 Schmitt, C., Marie-Cardine, A. & Bensussan, A. Therapeutic Antibodies to KIR3DL2 and Other Target Antigens on Cutaneous T-Cell Lymphomas. *Front Immunol* **8**, 1010, doi:10.3389/fimmu.2017.01010 (2017).
214 Keidel, S., Chen, L., Pointon, J. & Wordsworth, P. ERAP1 and ankylosing spondylitis. *Curr Opin Immunol* **25**, 97-102, doi:10.1016/j.coi.2012.11.002 (2013).

215 Roberts, A. R. *et al.* ERAP1 association with ankylosing spondylitis is attributable to common genotypes rather than rare haplotype combinations. *Proc Natl Acad Sci U S A* **114**, 558-561, doi:10.1073/pnas.1618856114 (2017).
216 Weiss, P. F. & Colbert, R. A. Juvenile Spondyloarthritis: A Distinct Form of Juvenile Arthritis. *Pediatr Clin North Am* **65**, 675-690, doi:10.1016/j.pcl.2018.03.006 (2018).
217 Chatterjee, N., Shi, J. & Garcia-Closas, M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet* **17**, 392-406, doi:10.1038/nrg.2016.27 (2016).
218 Snell, G. D. & Higgins, G. F. Alleles at the histocompatibility-2 locus in the mouse as determined by tumor transplantation. *Genetics* **36**, 306-310 (1951).
219 consortium, M. s. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* **401**, 921-923, doi:10.1038/44853 (1999).
220 Marsh, S. G. *et al.* Nomenclature for factors of the HLA system, 2010. *Tissue Antigens* **75**, 291-455, doi:10.1111/j.1399-0039.2010.01466.x (2010).
221 Robinson, J. *et al.* The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* **43**, D423-431, doi:10.1093/nar/gku1161 (2015).
222 Fasano, R. M. *et al.* Genotyping Applications for Transplantation and Transfusion Management: The Emory Experience. *Arch Pathol Lab Med* **141**, 329-340, doi:10.5858/arpa.2016-0277-SA (2017).
223 Jeanmougin, M., Noirel, J., Coulonges, C. & Zagury, J. F. HLA-check: evaluating HLA data from SNP information. *BMC Bioinformatics* **18**, 334, doi:10.1186/s12859-017-1746-1 (2017).
224 Matzaraki, V., Kumar, V., Wijmenga, C. & Zhernakova, A. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome Biol* **18**, 76, doi:10.1186/s13059-017-1207-1 (2017).
225 Neefjes, J., Jongsma, M. L., Paul, P. & Bakke, O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* **11**, 823-836, doi:10.1038/nri3084 (2011).
226 Bjorkman, P. J. *et al.* Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**, 506-512, doi:10.1038/329506a0 (1987).
227 Bjorkman, P. J. *et al.* The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329**, 512-518, doi:10.1038/329512a0 (1987).
228 Brown, J. H. *et al.* Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33-39, doi:10.1038/364033a0 (1993).
229 Long, E. O., Kim, H. S., Liu, D., Peterson, M. E. & Rajagopalan, S. Controlling natural killer cell responses: integration of signals for activation and inhibition. *Annu Rev Immunol* **31**, 227-258, doi:10.1146/annurev-immunol-020711-075005 (2013).
230 Munz, C. Autophagy proteins in antigen processing for presentation on MHC molecules. *Immunol Rev* **272**, 17-27, doi:10.1111/imr.12422 (2016).
231 Engelhard, V. H. The contributions of mass spectrometry to understanding of immune recognition by T lymphocytes. *Int J Mass Spectrom* **259**, 32-39, doi:10.1016/j.ijms.2006.08.009 (2007).
232 Lukacs-Kornek, V. & Turley, S. J. Self-antigen presentation by dendritic cells and lymphoid stroma and its implications for autoimmunity. *Curr Opin Immunol* **23**, 138-145, doi:10.1016/j.coi.2010.11.012 (2011).
233 Andersson, G. Evolution of the human HLA-DR region. *Front Biosci* **3**, d739-745 (1998).
234 Blum, J. S., Wearsch, P. A. & Cresswell, P. Pathways of antigen processing. *Annu Rev Immunol* **31**, 443-473, doi:10.1146/annurev-immunol-032712-095910 (2013).
235 Neerincx, A. & Boyle, L. H. Properties of the tapasin homologue TAPBPR. *Curr Opin Immunol* **46**, 97-102, doi:10.1016/j.coi.2017.04.008 (2017).

236 Turunen, S. *et al.* Rheumatoid arthritis antigens homocitrulline and citrulline are generated by local myeloperoxidase and peptidyl arginine deiminases 2, 3 and 4 in rheumatoid nodule and synovial tissue. *Arthritis Res Ther* **18**, 239, doi:10.1186/s13075-016-1140-9 (2016).
237 Stewart-Jones, G. B. *et al.* Crystal structures and KIR3DL1 recognition of three immunodominant viral peptides complexed to HLA-B*2705. *Eur J Immunol* **35**, 341-351, doi:10.1002/eji.200425724 (2005).
238 Huizinga, T. W. *et al.* Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. *Arthritis Rheum* **52**, 3433-3438, doi:10.1002/art.21385 (2005).
239 Furuya, T. *et al.* Differential association of HLA-DRB1 alleles in Japanese patients with early rheumatoid arthritis in relationship to autoantibodies to cyclic citrullinated peptide. *Clin Exp Rheumatol* **25**, 219-224 (2007).
240 Shimane, K. *et al.* An association analysis of HLA-DRB1 with systemic lupus erythematosus and rheumatoid arthritis in a Japanese population: effects of *09:01 allele on disease phenotypes. *Rheumatology (Oxford)* **52**, 1172-1182, doi:10.1093/rheumatology/kes427 (2013).
241 Scally, S. W. *et al.* Molecular basis for increased susceptibility of Indigenous North Americans to seropositive rheumatoid arthritis. *Ann Rheum Dis* **76**, 1915-1923, doi:10.1136/annrheumdis-2017-211300 (2017).
242 Oka, S. *et al.* Protective effect of the HLA-DRB1*13:02 allele in Japanese rheumatoid arthritis patients. *PLoS One* **9**, e99453, doi:10.1371/journal.pone.0099453 (2014).
243 El Mouraghi, I. *et al.* Polymorphisms of HLA-A, -B, -Cw and DRB1 antigens in Moroccan patients with ankylosing spondylitis and a comparison of clinical features with frequencies of HLA-B*27. *Tissue Antigens* **85**, 108-116, doi:10.1111/tan.12515 (2015).
244 Reveille, J. D. *et al.* HLA class I and II alleles in susceptibility to ankylosing spondylitis. *Ann Rheum Dis* **78**, 66-73, doi:10.1136/annrheumdis-2018-213779 (2019).
245 Liu, Y. *et al.* Predominant association of HLA-B*2704 with ankylosing spondylitis in Chinese Han patients. *Tissue Antigens* **75**, 61-64, doi:10.1111/j.1399-0039.2009.01379.x (2010).

Acknowledgements

E.D.M was funded by the NIH and the Daylight Foundation. R.B. is a past recipient of a Senior Research Fellowship and Research Progression Award (18543 and 20648 respectively) from Arthritis Research UK.

Author Contributions

The authors contributed equally to all aspects of the article.

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.

Supplementary information

Supplementary information is available for this paper at https://doi.org/10.1038/xxxxxxx

Related links

The IPD-IMGT/HLA database: www.ebi.ac.uk/ipd/imgt/hla/

HLA System nomenclature: hla.alleles.org/nomenclature/naming.html

Immune Epitope Database and Analysis Resource: www.iedb.org

Epitope prediction software: www.cbs.dtu.dk/services/NetMHCII-2.3

Epitope prediction software for MHC (including HLA) class II:

www.cbs.dtu.dk/services/NetMHCIIpan-3.2

Epitope prediction software for MHC (including HLA) class I: www.cbs.dtu.dk/services/NetMHCcons/

Genome-wide SNP database (humans and other species):

https://www.ncbi.nlm.nih.gov/projects/SNP/

HLA-check tool: omictools.com/hla-check-tool

George D. Snell's Nobel lecture: www.nobelprize.org/prizes/medicine/1980/snell/lecture/

Key points

- The HLA gene complex is one of the best-studied regions of the human genome, and technical advances continue to deepen our understanding of its associations with inflammatory arthritis.
- HLA genotypes corroborate distinct clinical and immunological subtypes between patients with the same clinical diagnosis or within the juvenile idiopathic arthritis umbrella.
- HLA associations are helping to clarify mechanistic overlap between pediatric and adult forms of arthritis.
- In some rheumatic conditions, potential mechanisms to explain HLA associations are emerging: shared epitope-dependent neoantigen presentation in ACPA-positive rheumatoid arthritis and HLA-B27 open conformations stimulating innate immune receptors in spondyloarthritis, among others.
- HLA associations can provide adjuncts to diagnosis and prognosis of rheumatic conditions, and novel therapeutic and preventative approaches might develop from further mechanistic studies of these associations.

Box 1. HLA gene polymorphism, nomenclature and typing techniques

The mouse MHC, called H2, was recognized in the 1940s as a gene region that controls tumour transplant rejection[218] (see George D. Snell's Nobel Lecture). Its human homolog, the HLA complex on human chromosome 6p21.3, was first identified in the late 1950s using serological techniques[22,71]. Cloning of individual HLA genes and sequencing of HLA proteins culminated in the sequencing of the entire HLA gene region in 1999, ahead of the rest of the human genome[22,219]. The HLA region contains a dense cluster of immune-related genes (**Fig. 1**), of which the genes coding for classical transplantation antigens, i.e., the peptide-binding HLA class I and class II glycoproteins, exhibit the greatest allelic polymorphism. As a result, most individuals are heterozygous at most HLA class I and class II gene loci, and paternally-inherited and maternally-inherited allelic variants at these loci are co-dominantly expressed. Thousands of alleles at several classical HLA class I and class II genes have been sequenced, and any two alleles typically differ at multiple nucleotide positions.

There is a well-established nomenclature for HLA genes[220] (see HLA System nomenclature for a detailed explanation and an annotated example). Letters (*A*, *B* and *C* for class I and *DP*, *DQ* and *DR* for class II) designate each locus; for class II genes, *A* and *B* suffixes (thus, *DRA*, *DRB*, etc.) are used to designate genes coding for the α and β chain polypeptide, respectively. Additional loci encoding alternate HLA-DR β-chains, which are present in some haplotypes, are numbered (e.g., *DRB4*). Locus names are followed by an asterisk and several pairs of digits, separated by colons, which identify the following types of allelic variation, in order: the major variants correlating with HLA serotypes; all

other coding variation; synonymous nucleotide substitutions in coding regions; noncoding differences including promoter polymorphisms. In practice, genotyping information may be incomplete, in which case unknown variation is omitted (thus, *B*27:05* may have many non-coding subtypes, but this genotype states only the coding polymorphisms). A well-curated database of HLA gene, cDNA and protein sequences is the IPD-IMGT/HLA database[221].

HLA gene-encoded protein variants are commonly named without intervening asterisks and colons; we follow this convention here. Thus, the *HLA-DRB1*04:01* allele codes for the β-chain polypeptide of the HLA-DR0401 class II protein, and the *HLA-B*27:05* allele codes for the heavy chain of the HLA-B2705 class I protein. Any associated non-polymorphic chains, such as the HLA-DRα or class I-associated β_2-microglobulin polypeptides are not specified: thus, for example, "DR0401 protein" refers to the heterodimer formed by the two polypeptide chains encoded by *DRB1*04:01* and (non-polymorphic) *DRA*01:01* genes.

In early studies, HLA types were primarily characterized using serotyping with panels of alloantisera, which contain antibodies specific for alleles that differ from that of the serum donor. These serologically distinct groups of alleles are still reflected in the 2-digit first number of the current allele nomenclature. For example, serotyping can distinguish B27 from B53 antigens using alloantisera, correlating with *B*27* and *B*53* genotypes, but is unable to distinguish the differences between *B27:05* and *B27:03* gene variants. For clinical purposes, HLA typing is now performed by a combination of molecular and immunological techniques[222]. When microarrays are employed for genome-wide analysis of single-nucleotide polymorphisms (SNPs), HLA gene region SNPs together with data in reference panels can be used to infer a reasonably accurate HLA genotype by a process called imputation[223] (for example, using the HLA-check tool, a software tool for imputation of HLA genotype from SNP data, and the genome-wide SNP database). Complete nucleotide sequencing is the gold standard for HLA genotyping; the technology continues to evolve[224].

Box 2: Function of MHC molecules

The major immune function of MHC molecules is to enable T cells to detect host cells that have been exposed to infectious pathogens[225]. In such host cells, foreign proteins are degraded (processed) and the resultant peptides bind to MHC proteins, which are then displayed (presented) on the plasma

membrane, where they can activate T cells (**Fig. 2**). The T cell antigen receptor (TCR) interacts with parts of the peptide and parts of the MHC molecule (known as 'dual recognition')[226-228].

MHC class I

MHC class I molecules (HLA-A, HLA-B and HLA-C in humans) acquire peptides derived from cytosolic protein degradation by the proteasome, transported by TAP transporters, and trimmed by ERAP (**Fig. 2a**). The TAP1 and TAP2 subunits of the peptide transporter, and PSMB8 and PSMB9, two cytokine-inducible proteasome subunits, are encoded in the HLA class II region (**Fig. 1**). In cells infected with cytosolic pathogens, pathogen-derived peptides enter the class I pathway and are presented to $CD8^+$ T cells, which then kill the infected cells. In addition, MHC class I–self-peptide complexes on healthy host cells engage inhibitory receptors on natural killer (NK) cells, delivering signals that protect the healthy cells from NK cell-mediated attack[229]. Diminished MHC class I expression on stressed or infected cells removes this inhibition, making them more susceptible to NK cell-mediated cell death.

MHC class II

MHC class II molecules (HLA-DR, HLA-DQ and HLA-DP in humans) are constitutively expressed on professional antigen presenting cells (APCs; such as monocytes/macrophages, dendritic cells, and B lymphocytes), which take up pathogens or their products. MHC class II molecules acquire peptides derived from endocytic processing of these exogenous antigens (**Fig. 2b**) and present them to $CD4^+$ T cells, which then activate or regulate effector responses through cell-surface interactions or cytokine release[21]. Endogenous peptides present in endosomes (for example, derived from autophagy of cytosolic proteins) can also become bound to MHC class II molecules[230]. MHC class II $\alpha\beta$ dimers are synthesised in the ER, associate with invariant chain (Ii), and are delivered to endosomes, where Ii is degraded by proteases. The subsequent release of residual class II-associated invariant chain peptides (CLIP) from the peptide-binding groove by HLA-DM creates MHC class II molecules competent for peptide loading (**Fig. 2b**). HLA-DM and its inhibitor, HLA-DO, are also encoded in the MHC class II region (**Fig. 1**).

Tolerance induction

Importantly, MHC proteins are occupied by a diverse cargo of self-peptides ($\geq 10^4$ distinct species), arising from the physiological turnover of host proteins[24,225,231]. Presentation of these peptides to T cells by APCs in the thymus, including thymic epithelial cells, and at extra-thymic sites in the absence

of inflammation, induces T cell tolerance[232], whether by T cell elimination (clonal deletion), anergy, or differentiation into regulatory T cells.

Legend to Figures

Fig 1: Organization of the HLA gene region.

The HLA gene region is shown with each bar representing a gene (simplified and not to scale). At the centromeric end (bottom) are 3 pairs of expressed genes coding for the α and β chains of classical class II molecules: *HLA-DPA1, HLA-DPB1; HLA-DQA1, HLA -DQB1;* and *HLA-DRA, HLA-DRB1*. *HLA-DRB3*, *HLA-DRB4*, or *HLA-DRB5* (*DRBx* in the Figure) encode alternate expressed HLA-DR β chains in many haplotypes[22,233]. In the telomeric class I region (top), the *HLA-A, HLA-C,* and *HLA-B* genes code for classical class I heavy chains (also known as class Ia genes), which assemble with β_2-microglobulin (encoded on chromosome 15). Classical class I and II genes display extraordinary allelic variation, except for HLA-DRA[22]. Interspersed among the class II loci are genes that regulate antigen presentation, such as *TAP1* and *TAP2* an *PSMB8* and *PSMB9*, which function in the class I pathway, as well as genes for *HLA-DM* and *HLA-DO*, in the class II pathway[85,225,234](explained in **Box 2**). The HLA class I region contains genes encoding 'non-classical' class Ib heavy chains, which are not involved in peptide presentation to $CD8^+$ $TCR\alpha\beta^+$ T cells[22]. The class III region encodes non-polymorphic immune molecules that are not directly involved in antigen presentation (such as complement components and TNF).

Fig. 2: MHC class I and class II proteins: maturation and function.

Distinct accessory molecules enable MHC class I and class II molecules to acquire peptides from different locations[21,225,234].

a) MHC class I molecules acquire peptides from proteasome-mediated degradation of cytosolic proteins. These peptides are imported into the ER by the transporter associated with antigen presentation (TAP1/TAP2), trimmed by ER aminopeptidases (ERAP), and loaded onto nascent class I heavy chain/β2-microglobulin complexes associated with TAP, tapasin, and other components (not shown). Tapasin ensures selection of a stable class I-bound peptide repertoire by promoting peptide exchange. In the ER and Golgi apparatus, class I molecules may further associate with the tapasin-related protein, TAPBPR, with consequences that are as yet unclear[235]. After transport to the cell surface, the MHC class I molecules present bound peptides to $CD8^+$ T cells.

b) Nascent MHC class II molecules bind to the invariant chain (Ii) polypeptide in the endoplasmic reticulum (ER). After transport to endosomes, Ii is trimmed, leaving class II-associated Ii peptides (CLIP) in the peptide-binding groove. HLA-DM promotes the exchange of CLIP for peptides generated in the endocytic pathway, but HLA-DO can also bind to and inhibit HLA-DM. After transport to the cell surface, MHC class II molecules present bound peptides to $CD4^+$ T cells.

Fig. 3 Neoantigen formation and presentation in RA

a) *Left*, Structure of the HLA-DR0401 peptide-binding site (PDB entry 4MCZ, top view)[67], complexed with citrullinated vimentin peptide. Side chains of citrulline at the P4 position of the peptide are shown in light red; three β-chain residues (V11, K71, A74) reported by Raychaudhuri *et al.*[32] to explain RA association, in green; and the other residues in the 70-74 β-chain region (the SE) in gold. *Right*, Neoantigen model. SE-containing HLA-DR variants present modified peptides with electrically neutral citrulline residues at P4, but they fail to present the unmodified parent peptides with positively-charged arginine at P4. Consequences for tolerance are explained in the text. In patients with a SE-containing HLA-DR allele, tolerogenic vaccination with citrullinated peptides might normal immune regulation (red arrow).

b) Neoantigen formation by post-translational modification. *Left,* Peptidylarginine deiminases (PAD) mediate calcium-dependent citrullination of arginine residues on polypeptide chains (left). *Right*, in the presence of cyanate (generated by oxidation of thiocyanate by neutrophil myeloperoxidase (MPO) at sites of inflammation[236]), lysine residues are carbamylated to homocitrulline, which is identical to citrulline except for an extra methylene group[73]. Inhibition of neutrophil enzymes (red arrows) might disrupt the presentation of neoantigens that perpetuate autoimmunity.

Fig. 4. Mechanisms of *HLA-B*27* association with SpA

a) *Left,* Misfolding of HLA-B27 in the endoplasmic reticulum (ER) leads to disposal by ER-associated degradation (ERAD) in the cytosol, and to ER stress, which synergizes with pro-inflammatory signals to favour IL-23 secretion, which, in turn, encourages pathogenic $CD4^+$ T helper 17 (T_H17) cell differentiation. *Middle,* Correctly-folded, cell surface HLA-B27–peptide complexes present

arthritogenic peptides and activate CD8$^+$ T cells. *Right*, Endosomal turnover of folded HLA-B27 complexes leads to the production of disulfide-bonded HLA-B27 heavy-chain dimers and other open conformations. The open conformations return to the plasma membrane and provide aberrant signals, e.g., *via* the NK cell receptor, KIR3DL2, to CD4$^+$ Th17 cells. IL-17 and IL-23 signalling contributes to osteoclast development. Therapeutic targets (red arrows) might include IL-17 and IL-23 and their signalling pathways, KIR3DL2, and cell surface HLA-B27 open conformations.

b) Structure of the peptide binding site of HLA-B2705 (PDB entry 2BST, top view)[237]. The cysteine at position 67 is implicated in the generation of open conformation heavy chain dimers. Residues 45 and 97 (glutamate and asparagine, respectively, in B*27:05) are located within the binding groove and influence peptide binding[118].

Table 1: Representative studies of HLA class I or class II susceptibility loci in RA and AS[a]

HLA molecule	Notable associated alleles or haplotypes	Type	Odds Ratio (95% CI)	Population	Ref.
ACPA-positive rheumatoid arthritis					
HLA-DRB1	Shared epitope (SE) alleles[b] +/+ (homozygous) +/- (heterozygous)	Risk alleles	11.79 (6.58-21.13) 4.37 (2.88–6.65)	White (Dutch)	238
	*HLA-DRB1*01:01*	SE Risk allele	1.38 (1.28-1.50)	White (European)	32[c]
	*HLA-DRB1*04:01*	SE Risk allele	4.14 (3.86-4.44)	White (European)	32
	*HLA-DRB1*04:04*	SE Risk allele	3.17 (2.83-3.54)	White (European)	32
	*HLA-DRB1*04:05*	SE Risk allele	2.31 (1.77-3.01)	White (European)	32
	*HLA-DRB1*04:05*	SE Risk allele	3.4 (2.0-5.7)	Japanese	239
	*HLA-DRB1*04:05*	SE Risk allele	3.93 (3.09-5.05)	Korean	32
	*HLA-DRB1*04:08*	SE Risk allele	5.48 (4.11-7.30)	White (European)	32
	*HLA-DRB1*09:01*	Risk allele[d]	1.43 (1.11-1.85)	Japanese	240
	*HLA-DRB1*10:01*	SE Risk allele	2.53 (2.04-3.14)	White (European)	32
	*HLA-DRB1*14:02*	SE Risk allele	2.38 (1.38-4.12)	Indigenous North American	241
	*HLA-DRB1*13:01*	Protective allele	0.24 (0.09-0.59)	White (European)	52
	*HLA-DRB1*13:02*	Protective allele	0.42 (0.31-0.58)	Japanese	242
Ankylosing spondylitis					
HLA-B27	*HLA-B*27*	Risk allele	27.5 (10.8-72.1)	Moroccan	243
	*HLA-B*27*	Risk allele	81.4 (49.8-134.3)	White (Dutch)	101
	*HLA-B*27*	Risk allele	87.7 (66.8-115.0)	White (US)	244
	*HLA-B*27*	Risk allele	120.8 (83.3-204.8)	Taiwanese	100
	*HLA-B*27:02*	Risk allele	43.4 (29.8-63.2)	White (European descent)	30
	*HLA-B*27:04*	Risk allele	86.8 [na][e]	Chinese (Han)	245
	*HLA-B*27:05*	Risk allele	62.4 (56.9-68.4)	White (European descent)	30
HLA-B40	*HLA-B*40.01*[f]	Risk allele	1.8 (1.2-2.8)	White (Dutch)	101
	*HLA-B*40:01*[f]	Risk allele	1.79 (1.29-2.49)	Taiwanese	100
HLA-B47	*HLA-B*47:01*	Risk allele	2.35 (1.43–3.86)	White (European descent)	30
HLA-B7	*HLAB*07:02*	Protective allele	0.82 (0.74–0.91)	White (European descent)	30
HLA-B57	*HLAB*57:01*	Protective allele	0.75 (0.61–0.92)	White (European descent)	30

[a] This table focuses on HLA-DRB1 alleles for RA and HLA-B alleles for AS. Both diseases have been shown to have associations with other HLA loci[32,244].

[b] Shared epitope alleles used in ref 238 analysis: *DRB1*01:01, *01:02, *01:04, *04:01, *04:04, *04:05, *04:08, *04:13, *04:16, *10:01, *14:02.*

[c] All data from reference 32 in this Table are extracted from Supplementary Table 3 of that paper.
[d] OR shown is based on analysis of *HLA-DRB1* alleles in all RA compared to healthy controls. Comparing ACPA+ to ACPA- RA patients, *HLA-DRB1*09:01* OR=2.02 (1.48-2.75)[240]. This allele has the sequence 70-RRRAE-74 in the SE region.
[e] B*2704 is more strongly associated with AS than B*2705 (OR = 2.5 (1.4–4.2) in Han Chinese[245].
[f] Meta-analysis of B*40:01 OR 2.2 (1.8-2.8). Epistatic effect: B27-positive/B40:01-neg a has relative risk (RR) of 69 (40-111), B27-positive/B40:01-positive has RR of 342 (147-708)[101]. Results in a Taiwanese cohort are similar[100].
na, confidence interval not available

Glossary terms

- **Epistasis**: the ability of one gene to influence the effect of another gene on a phenotype (such as the risk of developing a rheumatologic condition); also known as gene –gene interactions.
- **Haplotypes:** a constellation of allelic variants at closely-linked loci, which are preferentially or exclusively inherited together owing to to linkage disequilibrium
- **Linkage disequilibrium:** a lack of recombination within a stretch of DNA (relative to recombination frequencies expected on the basis of the length of DNA involved), such that allelic variants contained within that stretch are systematically co-inherited as a haplotype.
- **Neoantigen:** An antigenic peptide that arises from the post-translational modification of a self-protein, thus appearing foreign or new; the generation of such peptides is a proposed mechanism for breaking self-tolerance.
- **Polygenic risk score:** a composite score based on weighted contributions of a large number of allelic variants across the genome to provide greater stratification of risk, especially at the tail ends of the risk distribution.
- **Polymorphism:** variability, within a population, of a gene or a group of genes. In the HLA region, several types of polymorphism can be distinguished, including polygeny (the existence of multiple gene loci coding for polypeptides with similar functions), the presence of alternate HLA-DRB genes in some HLA-DRB1 haplotypes, and allelic variation.