ABSTRACT
Both ankylosing spondylitis (AS) and rheumatoid arthritis (RA) are common, highly heritable conditions, the pathogenesis of which are incompletely understood. Gene-mapping studies in both conditions have over the last couple of years made major breakthroughs in identifying the mechanisms by which these diseases occur. Considering RA, there is an over-representation of genes involved in TNF signalling and the NFκB pathway that have been shown to influence the disease risk. There is also considerable sharing of susceptibility genes between RA and other autoimmune diseases such as systemic lupus erythematosus, type 1 diabetes, autoimmune thyroid disease and celiac disease, with thus far little overlap with AS. In AS, genes involved in response to IL12/IL23, and in endoplasmic reticulum peptide presentation, have been identified, but a full genomewide association study has not yet been reported.

Introduction
Rheumatoid arthritis (RA) and ankylosing spondylitis (AS) are the prototypic seropositive and seronegative inflammatory arthropathies respectively. Whilst the clinical distinction between these common conditions has long been recognised, apart from differences in their HLA-associations, until recently little has been known about the differences in their genetic associations. The genomewide association study era has demonstrated that there are major differences between the genes involved in these conditions, with little overlap thus far identified. This review outlines our current knowledge of the genetics of each condition, and highlights the major differences that underlie them.

Rheumatoid arthritis
Genetic factors clearly play a major role in determining individual risk of developing RA. The disease runs strongly in families, and is highly heritable. The heritability of RA is about 60% (i.e. 60% of the risk of developing RA is due to genetic factors) (1). The likelihood of siblings of an RA case also developing the disease is 2-6-times greater than the likelihood of developing RA in the general community, with a higher recurrence risk ratio observed in families with more severe cases (2, 3). This suggests that genes affecting susceptibility to RA also influence its severity.

Major histocompatibility complex genes and RA
The main gene causing RA, HLA-DRB1, has been known for more than 25 years (4). This gene accounts for about 30% of the risk of developing RA (5), and in the last couple of years rapid progress has been made identifying the further genes involved in susceptibility to the disease.

RA is associated with HLA-DRB1 alleles that encode a shared sequence of amino-acids termed the ‘shared epitope’ (SE; consisting of 70QRRAA74, 70RRRAA74 or 70QKRAA74) which comprise residues 70-74 in the third hypervariable region (HVR3) of the DRβ1 chain (6). The alleles carrying this nucleotide sequence are DRB1*0401, *0404, *0405, *0408, *0101, *0102, *1402, *09 and *1001. These alleles have differential strength of association with RA, with the highest risk associated with DRB1*0401 or *0404 homozygotes or compound heterozygotes. The absolute risk of developing RA if carrying a SE encoding allele is 1 in 35 for *0401, 1 in 20 for *0404, *0405, *0408, *0101, *0102, *1402, *09 and *1001. These alleles have differential strength of association with RA, with the highest risk associated with DRB1*0401 or *0404 homozygotes or compound heterozygotes. The absolute risk of developing RA if carrying a SE encoding allele is 1 in 35 for *0401, 1 in 20 for *0404 and 1 in 80 for *0101 (7). The association of HLA-DRB1 is restricted to anti-CCP positive RA, and there is a strong interaction involved with cigarette smoking, since SE-carriers that smoke are much more likely to develop RA than non-smokers (Fig. 1) (8).
It is likely that further MHC genes are involved in RA-susceptibility, and that they either interact differently with, or are carried on different haplotypes by, different SE-carrying alleles (9). Preliminary studies using dense SNP maps have identified other regions likely to harbour these genes (10). Recently the MHC Class III gene, *AIF1*, has been associated with RA and T1DM, and over-expressed in peripheral blood and in synovium of RA cases (11, 12). Further genes are also likely to be involved, but large studies controlling for the complex patterns of MHC linkage disequilibrium will be required to identify these.

**Protein tyrosine phosphatase 22 (PTPN22)**

The gene *PTPN22* was found to be associated with RA having originally been identified as a susceptibility gene for type 1 diabetes mellitus (T1DM) (13). It encodes a 110-kD cytoplasmic protein tyrosine phosphatase (Lyp) that is thought to function as a down-regulator of T-cell receptor dependent responses through interaction with a negative regulatory kinase, Csk. This gene is also associated with a variety of other autoimmune diseases including systemic lupus erythematosus, Addison’s disease and Hashimoto’s thyroiditis, in addition to RA (14) and T1DM, helping to explain the tendency for these conditions to segregate together. *PTPN22* is not associated with RA in Asians, because the key associated variant (rs2476601, R620W) is not polymorphic in Asian populations (15).

**Peptidyl-arginine deiminase 4 (PADI4)**

The gene *PADI4* has been comprehensively demonstrated to be associated with RA in Japanese and Korean populations (16), but not in Caucasians. The enzyme PADI4 citrullinates peptides, antibodies against which (anti-CCP antibodies) are highly specific for RA. The discovery of the association of *PADI4* with RA strongly suggests that anti-CCP antibodies are important in disease pathogenesis, and not merely an immunological epiphenomenon.

**NFKB pathway genes**

In the past 24 months several new genes have been identified and confirmed as being associated with RA that have helped define the likely immunopathogenesis of the disease. A common feature of the RA-associated genes *TRAF1*, *TNFAIP3*, *PRKCQ*, *TNFRSF14* and *CD40* is that they are all involved in TNF/NFκB signaling pathways. A region on chromosome 9q33 containing the genes *C5* and *TRA1F* has been robustly associated with RA (17), and fine-mapping studies make it increasingly likely that *TRA1F* (TNF receptor-associated factor 1) is the key associated gene. The TRAF1 protein is involved in signal transduction from TNF receptor family members, including *CD40* and *TNFRSF14*, which have also recently been associated with RA (18). Neither the *TRA1F/C5* nor *TNFRSF14* loci have yet been convincingly associated with other autoimmune diseases. **CD40** polymorphisms have recently been associated with multiple sclerosis. A region on chromosome 6q23, in addition to being associated with RA (17, 19), has also been associated with SLE, T1DM, and celiac disease. The region contains two known genes (*TNFAIP3* and *OLIG3*), as well as a pseudogene *PTPN11*. At this stage the evidence most strongly supports *TNFAIP3* as being the key associated gene in the region.

**T-cell differentiation/activation genes**

Genes influencing T-cell differentiation/activation are also over-represented in those that have been associated with RA to date. **STAT4** is significantly associated...
with RA (23), and also with other organ-specific autoimmune diseases including T1DM and SLE. STAT4 is a transcription factor particularly involved in driving CD4 lymphocyte differentiation into Th1 lymphocytes, and stimulating interferon-γ secretion. STAT4 knockout mice develop less severe collagen-induced arthritis, and suppression of STAT4 levels using antisense oligonucleotides reduces collagen-induced arthritis activity, even when applied after the onset of arthritis (24).

CTLA-4 is a protein involved in co-stimulation regulation of T-cell activation. Polymorphisms of CTLA-4 are associated with autoimmune thyroid disease, T1DM, and celiac disease, and RA (25). PRKCQ in addition to being involved in NFκB signaling has widespread effects on T cell differentiation and survival. These findings confirm that abnormalities of T cell differentiation and function are likely involved in RA causation, and are not just bystander effects of the RA disease process.

Other regions
In some genetic regions associated with RA, the key associated gene is quite unclear. Association has been reported and confirmed with the chromosome 4q27 region encoding both the IL2 and IL21 genes (26). IL2RB has been convincingly associated with RA, so IL2 would seem a likely candidate, but IL21 is also a strongly proinflammatory cytokine. It may prove that both are involved.

A region on chromosome 12q13 has been associated with RA (27) and also recently with T1DM and multiple sclerosis. The key associated gene here is quite uncertain. In multiple sclerosis and type 1 diabetes though, there is suggestive evidence that the gene encoding the vitamin D activating enzyme, 1-alpha 25-hydroxyvitamin D hydroxylase, is involved (reviewed in (28)). Vitamin D has multiple effects on immunological function, and if this gene is confirmed as being the polymorphic gene responsible for the association at this locus, it would represent a highly tractable therapeutic target. In RA, the strongly associated SNP lies in KIF5A (kinesin family member 5A), a functionally unlikely candidate for RA, but the association signal includes the gene PIP4K2C, encoding phosphatidylinositol 4 phosphate 5 kinase type II gamma. The latter has recently been identified as an autoantigen in RA, raising potential mechanisms by which it may influence RA susceptibility (29).

Conclusion
The list of genes definitely associated with RA is likely to continue to increase rapidly as sample sizes increase, and different populations are studied. However the genes robustly identified to date place T-cell activation, and NFκB signaling at the centre of the process leading to the development of RA. These findings are already providing valuable foundations from which hypothesis driven research can proceed, and identified potential novel therapeutic targets.

Ankylosing spondylitis
Familial aggregation of ankylosing spondylitis (AS) has been recognized for many years and the association with HLA-B27, first reported in 1973 (30, 31), constitutes the strongest genetic association observed with any human immune disease. Siblings of a case have a 50-fold risk of developing the condition compared to the general population and disease concordance is 63% in identical as compared to 13% in non-identical twins (32, 33). The latter study also estimated that 97% of the population variance can be explained by genetic effects and environmental influences are likely to be ubiquitous and therefore non-contributory to population variance. Genetic factors also appear to play an important role in influencing disease severity as affected sibling pairs have closer scores for disability, pain, and radiological damage than expected (34-36). An important role for non-B27 genes is highlighted by the increased risk of disease in B27 positive relatives of AS patients compared to B27 positive individuals in the general population as well as the higher concordance rate of disease in monzygotic twins (63%) compared to B27 positive dizygotic twin pairs (23%) (33, 37).

Major histocompatibility complex genes
Although the dominant gene accounting for the association at the major histocompatibility complex (MHC) is B27, there appears to be a role for other HLA-B and non-HLA-B MHC genes.

HLA-B27 and B27 subtypes.
The risk for AS in B27 positive individuals is in the order of 2-5% and 90-95% of Caucasians are B27 positive. The strength of this association is less evident among certain ethnic groups and in those with concomitant psoriasis or inflammatory bowel disease where only 60-80% carry B27.

The Anthony Nolan Trust database reported 58 subtypes of B27 as of April 2009 but for most, disease association with AS is unclear since the subtypes have been reported in only a few unaffected individuals. AS has been reported with the following subtypes: B*2702, 2703, 2704, 2705, 2706, 2707, 2708, 2709, 2710, 2714, 2715, 2719 and 2730 (38). Studies of B27 subtypes have been very helpful in clarifying the following aspects into the mechanism of the B27 association with AS. First, because B27 subtypes associated with disease occur on different HLA haplotypes, this draws attention to B27 as the principle gene associated with disease and not some adjacent B27 linked MHC gene. Second, analysis of the amino acid sequence of associated and non-associated B27 subtypes, has shown that the primary differences are observed in the B27 antigen binding groove and studies have shown that this results in differences in the range of peptides presented by B27 to cytotoxic T-cells (39). B*2706 and B*2709 are less strongly associated with disease and differ from the disease associated B2705 subtype at amino acid position 116 (both subtypes) and position 114 (B*2704) in the antigen binding pocket of B27 that binds the C-terminal amino acid of the antigenic peptide so that peptides with arginine or tyrosine at the C-terminus are not bound. There is presently sufficient epidemiological data to support the association between B*2704, B*2705, B*2702 and B*2707 with AS. It was originally thought that
B*2706 and B*2709 were protective but there are now case reports reporting disease in carriers of B*2709 (40, 41) and one family study has shown that B2704/2706 compound heterozygotes still develop AS (13). It is more likely that these subtypes have a weaker association with AS than B*2705. Although the mechanistic basis for the differential subtype associations with disease is still unclear, it is unlikely that this is on the basis of a shared epitope as has been established in RA.

**Non-B27 HLA genes**

Population studies and examination of multiplex families have implicated HLA-B60 in risk for disease in B27 positive Caucasians (42). It has also been implicated in B27 negative Chinese with AS (43), while an increased risk has been associated with B39 in B27 negative Japanese cases (44). Interestingly B39 shares the same amino acids that make up the B-pocket of the antigen binding groove in B27 and is capable of binding the same peptides. The mechanism of the B60 association is unclear and it may reflect an HLA haplotype that carries other genes associated with disease. B*1403 has been associated with disease and also shares a similar sequence to B27 in the antigen binding groove (45).

Evaluation of associations with other MHC genes is extremely challenging due to the complexity associated with extreme variation at different MHC loci and the extensive linkage disequilibrium observed between loci. One study that compared B27 matched case and control haplotypes showed a strong association with HLA-DRB1 regardless of the B27 status of the chromosome (46). Case control analysis matching for B27 identified 2 haplotypes (B27 positive/DRB107 positive and B27 negative/DRB103 positive) that were associated with susceptibility to disease. This study also implicated genes in the class 3 region that include the C4A/B components of the complement pathway, CREBL1, involved in the unfolded protein stress response in the endoplasmic reticulum, and CYP21A2, a member of the cytochrome P450 family. Identification of the true disease associated gene in this region will require large data-sets with greater marker densities.

**Non-major histocompatibility complex genes**

The component of the increased risk in siblings attributable to non-MHC loci is approximately equivalent to the entire genetic component of insulin dependent diabetes and greater than that for rheumatoid arthritis. Recent work has identified several novel genetic associations in non-MHC regions.

**Interleukin-1 (IL-1) gene cluster**

Association of members of the IL-1 gene cluster on chromosome 2q13 has been reported in several studies in Caucasian and Asian populations (47-53) although consistency has been lacking in identifying the principle gene associated with disease. The IL-1 gene cluster is a 360 kilobase region containing 9 genes with sequence homology either to the pro-inflammatory IL-1 agonists, IL-1alpha (IL-1A) and IL-1beta (IL-1B), or the anti-inflammatory IL-1 antagonist IL-1RN, encoding the protein IL-1RA. A recent meta-analysis of 2,675 AS cases and 2,592 ethnically matched controls from 12 discrete cohorts in 10 countries demonstrated that the strongest association was observed with 3 single nucleotide polymorphisms in the IL-1A gene with no significant heterogeneity of effects between centers (51). The population attributable risk fraction was estimated at 4-6%. Odds ratios for the association with disease were quite modest for each polymorphism (≈1.2) and similar to that reported for associated genes in RA. This study was important in also emphasizing the required sample size to detect effects of the size observed with IL-1A and is a likely explanation for the variability of findings in prior studies of IL-1 gene cluster members and AS.

**Endoplasmic reticulum aminopeptidase-1 (ERAP1)**

A recent genome-wide association scan conducted in 1,000 patients with AS and 1,500 control subjects from the Welcome Trust Case Control Consortium (WTCC) and the Australo-Anglo American Spondylitis Consortium (TASC) using 14,500 non-synonymous coding single nucleotide polymorphisms demonstrated associations with SNPs located in the gene encoding an endoplasmic reticulum amino peptidase, ERAP1, on chromosome 5 and interleukin 23 receptor gene on chromosome-1 (54). ERAP1 has been implicated in 2 biologic functions: 1. N-terminus trimming of peptides that have been transported into the endoplasmic reticulum where they bind to Class 1 HLA molecules prior to cell surface presentation to T-cells (55); 2. trimming of surface expressed cytokine receptors and specifically the tumor necrosis factor receptor 1 (TNFR1) (56), interleukin 6 receptor alpha (57), and the type II interleukin 1 receptor (58). A replication study in 3 Canadian case control cohorts has confirmed this association and also demonstrated a specific ERAP1 haplotype that was strongly associated with disease in all 3 cohorts (59).

**Interleukin-23 receptor (IL23R)**

An association with IL23R has been demonstrated in several cohorts in addition to the WTCC and TASC cohorts (60-62). It has also been demonstrated to be associated with both inflammatory bowel disease and psoriasis (63, 64). IL23 is an important cytokine and together with IL-1 and IL-6 regulates the differentiation of T-cells into a distinct phenotype that expresses interleukin 17 (65). IL17 induces pro-inflammatory cytokines in several cell types within the joint and has been shown to act synergistically with tumor necrosis factor alpha (66). It is elevated in both synovial fluid and serum of patients with spondyloarthropathy (67). Therapeutic strategies targeting IL23 have been successful in psoriatic arthritis but have yet to be assessed in AS (68). A recent report has also shown that over expression of IL23 is a pivotal feature of subclinical gut inflammation in AS (69). An important role for IL23 in directly activating the innate intestinal immune system is an additional property distinct from its role in triggering the proliferation of TH17 cells (70).
Conclusion
Although we now have quite long lists of genes identified which are involved in RA and AS, thus far there appears little overlap between factors affecting susceptibility to the two diseases. As predicted by the patterns of familial aggregation of these diseases, RA shares many susceptibility genes with autoimmune diseases such as TIDM, celiac disease, autoimmune thyroid disease and even systemic lupus erythematosus. A feature of the genetic determinants of RA is an over-representation of genes involved in NFκB pathways. In contrast, in AS, the genes identified to date have been involved in Th17 lymphocyte differentiation, and likely, peptide presentation by HLA Class I. Many more genes remain to be identified for each condition, but even with only a moderate proportion of the genetic susceptibility to each disease defined, these novel insights have validated the genetics research programs that have contributed to these achievements.

References

43. WEI JC, TSAI WC, LIN HS, TSAI CY, CHOU CT: HLA-B60 and B61 are strongly associated with ankylosing spondylitis in HLA-B27-negative Taiwan Chinese patients. Rheumatology (Oxford) 2004; 43: 839-42.


