Genetics of ankylosing spondylitis

M.A. Brown, B.P. Wordsworth, J.D. Reveille

Matthew A. Brown, MD, Wellcome Trust Centre for Human Genetics, Spondyloarthritis and Bone Disease Research Group, Oxford; B. Paul Wordsworth, Roosevelt Drive, Headington, United Kingdom; John D. Reveille, Department of Medicine, University of Texas Health Science Center at Houston, USA.

MAB is supported by the Arthritis Research Campaign (UK).

Please address correspondence to:
Matthew A. Brown, Wellcome Trust Centre for Human Genetics, Spondyloarthritis and Bone Disease Research Group, Roosevelt Drive, Headington, OX3 7BN, United Kingdom.
E-mail: mbrown@well.ox.ac.uk

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2002.

Key words: Ankylosing spondylitis, spondyloarthritis, genetics, linkage, association, HLA.

ABSTRACT
Ankylosing spondylitis is a common inflammatory rheumatic disease. Both susceptibility to and clinical manifestations of the disease are highly heritable. Although some genes, notably HLA-B27, have been implicated in susceptibility to the disease, the genetic of the condition are complex and many more genes involved in the condition await discovery.

Introduction
Ankylosing spondylitis (AS) is the second most common autoimmune rheumatic disease and despite major clues as to its causes such as the relationship with reactive arthritis and the association with HLA-B27, its aetiology remains very poorly understood. A major genetic predisposition was suggested by the strong familial predisposition to the disease, and this was confirmed in the early 1970’s by the discovery of the association with HLA-B27. HLA-B27 is probably the major gene involved in susceptibility to AS, but operates in combination with other genes to determine the majority of susceptibility to the disease, and has little role in determining disease severity. Gene mapping studies have implicated non-B27 genes both within the MHC and elsewhere as being involved in susceptibility to the disease, but only very preliminary studies of genes involved in disease severity have been undertaken.

Genetic epidemiology of AS
The high degree of familial recurrence of AS has been long recognised, but until formal heritability studies in twins were performed, the extent to which this was due to shared environmental or genetic factors was unknown. From twin recurrence risk studies the heritability of susceptibility to AS has been estimated at 97% (95% confidence interval 92-99%) (1). The concordance rate for B27-positive DZ twin pairs (23%) is still considerably below that of MZ twin pairs (63%), indicating a large genetic component not linked to the MHC (1).

Disease severity determinants in AS
The genetic control of disease severity has also been investigated in twins. All measures of disease activity and functional impairment were more similar in MZ than DZ twins, although the number of concordant twin pairs studied was quite small (1). Nahal and colleagues examined familiality of disease phenotypes, and noted increased clustering of arthritis and uveitis, but not late age of onset, or peripheral enthesitis (2). Familiality of pain and disability indices, and pelvic and spinal radiographic changes have also been reported (3), but neither study differentiated shared environmental factors from genetic factors. More recently a complex segregation analysis has demonstrated a high degree of genetic control of disease severity assessed by the Bath AS Disease Activity index (BASDAI) and functional impairment assessed by the Bath AS Functional Index (BASFI) (4). Heritability of these measures was estimated at 51% and 68% respectively; modelling studies indicated that a monogenic model fitted the observed data most closely, with polygenic and environmental models being rejected. As all the affected relative pairs in this study were B27-positive, a gene other than B27 must be involved, supporting the hypothesis that the genes controlling the process of ankylosis are different from those controlling susceptibility to AS.

Family recurrence of AS
The number of genes involved in susceptibility to AS has been estimated by recurrence risk modelling. In diseases with a significant genetic component, the recurrence risk of a disease in relatives of cases falls with increasing distance of relatedness to the proband and the rate of reduction is determined by both the number of genes involved and
their mechanism of interaction (5). This data was used to test different genetic models in comparison with the observed data, allowing an estimate of the likely number of genes involved and their pattern of interaction (6). The best fitting model was a five-locus model with multiplicative interaction between loci. This suggests that the major genes involved in susceptibility to AS are likely to be of at least moderate magnitude, an encouraging thought for those involved in gene-mapping studies.

Recurrence risk in families has been shown to differ according to the gender of the proband, with an increased risk apparent in the relatives of younger, female probands (7, 8). This is consistent with a polygenic contribution to AS. If women have a higher threshold for developing the disease (as suggested by the gender imbalance in AS), then affected women may have more susceptibility genes than men, increasing the risk to their offspring (9). Similarly, if early onset of AS reflects the presence of greater genetic susceptibility, there should be a correspondingly increased recurrence risk in the offspring of younger probands. The exclusion of significant susceptibility loci on the X-chromosome indicates that the gender bias in risk of disease and in recurrence risk in relatives of AS cases is not explained by X-chromosome encoded factors (10).

An increased risk of developing AS in first-born children and in the children of younger mothers has been reported (11). The increased risk of AS with low birth-order was not replicated in a study of multicase-spondyloarthropathy families (12), or British AS families (13), and no association between maternal age and risk of AS in children was observed in another British study (14). Potential explanations for a birth-order effect include the development of maternal anti-B27 antibodies in B27-negative mothers of B27-positive offspring, or increased risk of early childhood infection in high birth-order children, possibly shaping the immune system to protect children from developing AS (11). The latter suggestion has been raised as an explanation for the paucity of AS amongst B27-positive rural west-Africans (15).

**Gene-mapping studies in AS**

The gene-mapping methods available to researchers to identify genes in any disease can be divided into linkage disequilibrium (‘association’) approaches and linkage studies, and in turn these can either be candidate gene or whole genome studies [reviewed in (16, 17)].

**Linkage mapping**

Genetic linkage refers to the cosegregation of alleles at two separate genetic loci in families. Linkage mapping is used to create low-resolution maps of the likely chromosomal location of disease-associated genes. It is feasible to screen the entire human genome for genes of moderate effect by linkage with 300-400 polymorphic markers. However, linkage studies are inefficient at fine localisation (‘fine mapping’) of susceptibility loci in polygenic diseases (18).

**Linkage disequilibrium mapping**

Linkage disequilibrium (LD) refers to the inheritance of markers together on a chromosome strand in a population more often than would be expected by chance, creating a particular haplotype. A variety of approaches to linkage disequilibrium mapping exist including case-control (population association) studies and within-family association studies using parental or sibling data as the control group. The extent of linkage disequilibrium varies considerably throughout the genome and in different populations. Genetically homogenous populations have wider areas of linkage disequilibrium, increasing the power to detect it and possibly thereby to map genes. Across the genome linkage disequilibrium varies widely and islands of increased linkage disequilibrium occur which can extend hundreds of kilobases (19-21). There is some suggestion that the underlying DNA structure affects recombination rates and therefore the extent of linkage disequilibrium. GC-poor regions have been shown to have more extensive LD than GC-rich regions, and it has been suggested that in GC-poor regions one marker every 100 kilobases may be sufficient, whereas in GC-rich regions a density of one marker per 10 kilobases may be required (22). Efforts are underway to define these islands to reduce the density of markers that would be required for LD mapping.

Hitherto the paucity of success in identifying genes in complex human genetics has encouraged scepticism about the likelihood of success of current approaches (23,24). Two recent successes have demonstrated that the method can be successful; linkage disequilibrium mapping has identified the NOD2 gene as being involved in Crohn’s disease (25), and the calpain-10 gene in type 2 diabetes mellitus (26). The association of variants of the NOD2 gene with Crohn’s disease is of particular interest to the genetics of spondyloarthropathies, given the clinical association between the conditions. NOD2 mutations have also been demonstrated to cause a monogenic disease, Blau Syndrome (Online Mendelian Inheritance in Man reference 186580), characterised by granulomatous small joint arthritis, skin rash, and non-granulomatosus uveitis (27). NOD2 is expressed in monocytes, and promotes apoptosis through nuclear factor kappa-B activation. It has structural similarity to a class of disease-resistance genes in plants that induce localised cell death at the site of pathogen invasion and is thought to function as a cytosolic receptor for pathogenic components derived from bacteria. NOD2 mutations are not involved in susceptibility either to primary AS or AS complicating inflammatory bowel disease (28). Whether NOD2 mutations affect susceptibility to B27-associated uveitis is unknown, but given the association with uveitis in Blau Syndrome is an interesting hypothesis.

**MHC genes and AS**

**HLA-B27**

The association of B27 with AS remains among the strongest immunogenetic of any disease, only bettered by the association of the DRB1*1501-DQB1*0602 haplotype with narcolepsy. Extremely strong linkage of the MHC with AS has been demonstrated in several studies (29, 30), including a
maximum LOD score of 15.6 in a recently whole genome screen (31). Yet only a small proportion of subjects carrying B27 develop the disease (32-34). The world-wide association of B27 with AS and B27-transgenic animal models of spondyloarthropathies, strongly suggest that B27 itself rather than a nearby gene which is involved, but this remains unproven formally in humans. Population studies have shown that different disease-associated B27-subtypes carry different MICA/MICB or HLA-C alleles (35-37). This indicates that the gene involved lies between MICB and HLA-C, a distance of 131 kilobases which contains several predicted genes.

To date, up to 24 subtypes of HLA-B27 have been reported. HLA-B*2722 is no longer regarded as a legitimate subtype-the sequence described from a Filipino of unknown disease status (38, 39). The subtypes are presumed to have evolved from a “parent” HLA-B27 allele-HLA-B*2705 (Fig. 1), and evolved along at least three separate paths: one, in Caucasians and Africans, involving substitutions in the first domain (encoded by exon 2 of the HLA-B27 gene); the second, predominantly in Asians, involved one same substitution in the first domain and a variable number of substitutions in similar positions in the second domain (encoded by exon 3); and the third group, predominantly in Caucasians, which involved only substitutions in the second domain. Finally, there are B*2713, which differs from B*2705 at another location than the outermost two domains, and B*2718, which appears to have evolved separately.

Studies of the association of these allelic variants have suggested differences in susceptibility to AS particularly with B*2706 (now split into B*2706, B*2711 and B*2722) and B*2709, which are considered as being less disease-associated, and B*2704, which may confer greater disease susceptibility that other B27 subtypes (40) pointing to the B27 gene itself being critical in susceptibility to AS. However, these observations can potentially still be explained by differences in linkage disequilibrium with another nearby gene. Certainly B27 subtype differences play no role in western European populations in determining those B27-positive individuals who will develop AS (41).

**Non-B27 MHC genes**

There is strong evidence that B27 is not the only MHC gene involved in susceptibility to AS. As indicated above, linkage studies will not have sufficient resolution to differentiate effects of HLA-B27 from those of other MHC genes. Within-family linkage disequilibrium approaches also must cope with the strong linkage disequilibrium that occurs across the MHC complicates population genetic approaches to identifying the genes involved. Interethnic and functional studies are probably going to be required to identify the individual genes involved.

**HLA-B60**

Robinson and colleagues first reported the association of B60 with B27-positive AS (42), a finding that has been replicated in many but not all studies in

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>αα</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>B*2703</td>
<td>1</td>
<td>African</td>
</tr>
<tr>
<td>B*2717</td>
<td>1</td>
<td>Unc</td>
</tr>
<tr>
<td>B*2701</td>
<td>3</td>
<td>Mexico</td>
</tr>
<tr>
<td>B*2702</td>
<td>3</td>
<td>Caus.</td>
</tr>
<tr>
<td>B*2716</td>
<td>3</td>
<td>Caus.</td>
</tr>
<tr>
<td>B*2708</td>
<td>4</td>
<td>Caus.</td>
</tr>
<tr>
<td>B*2723</td>
<td>7</td>
<td>Caus.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>αα</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>B*2704</td>
<td>1</td>
<td>Asian</td>
</tr>
<tr>
<td>B*2715</td>
<td>1</td>
<td>Asian</td>
</tr>
<tr>
<td>B*2725</td>
<td>1</td>
<td>Unc.</td>
</tr>
<tr>
<td>B*2721</td>
<td>1</td>
<td>Caus.</td>
</tr>
<tr>
<td>B*2711</td>
<td>1</td>
<td>Asian</td>
</tr>
<tr>
<td>B*2720</td>
<td>1</td>
<td>Asian</td>
</tr>
<tr>
<td>B*2724</td>
<td>1</td>
<td>Unc.</td>
</tr>
</tbody>
</table>

Fig. 1. Alleles of HLA-B27. Possible origins. Number of amino acid substitutions (0 amino acid). (Courtesy of Prof. M.A. Khan).
both Caucasian and Asian populations. Small studies of B27-negative AS have demonstrated an association of B60 in these cases, indicating that the gene involved must be either B60 or be in tight LD with it (1, 41, 43).

**HLA-DRB1**

Association studies of HLA-B27-DRB1 haplotypes in AS have shown by both case-control and within-family methods that the B27-DR1 haplotype is associated with AS in British (1, 44) and French studies (45). Further evidence of a second MHC locus in addition to B27 comes from studies of inflammatory bowel disease (IBD) and AS complicating IBD. A rare subtype of DR1, HLA-DRB1*0103, is associated with IBD (46), increased severity of IBD, and with arthritis complicating IBD (47, 48). It has recently been demonstrated in patients with IBD and AS that DRB1*0103 is strongly associated with the two diseases (48), and that this association is independent of B27 (49). These findings indicate that it is extremely likely that a further MHC gene is involved in susceptibility to AS, particularly when complicated by IBD. Association of HLA-DR4 with spondyloarthritides has recently been reported, but the analysis in this paper employs multiple cases from each family (50, 51). Thus the findings represent a combination of linkage and association, and it is not possible to be sure whether DR4 itself is truly associated with AS, or if this finding is due to linkage of the MHC with AS and linkage disequilibrium of HLA-B27 and -DR4 (52). An association between peripheral arthritis complicating AS and DR4 has been reported (53) but not replicated by other studies. The rare DRB1*0408-DQB1*0301 haplotype was reported to be associated with reactive arthritis in Finns (54), but this was not confirmed in another Finnish study (55), perhaps because the latter studied appropriately used B27-positive controls. Associations between DR7 and the presence of peripheral arthritis in AS has been reported by several groups (56-58). This is consistent with reported associations between DR7 and psoriatic arthritis (57, 59, 60), which has a higher frequency of peripheral arthritis than observed in AS. Disease progression in psoriatic arthritis is associated with the B27-DR7 haplotype (61). In a study of British Caucasian AS patients, DR7 was associated with a younger age of onset of disease, but not with other disease manifestations (44). DR8 is associated with iritis in Japanese adults with AS (62). This finding has been replicated in a study of Norwegian adult AS cases, where it was also found to be associated with a younger age of disease onset (63). A DR8 bearing haplotype (HLA-A2-B46-DR8) was associated with psoriatic arthritis in Japanese subjects, but DR8 itself was not associated with disease (64). DR8 homozygosity was associated with susceptibility to AS in British Caucasians, but no effect of DR8 on susceptibility to iritis was noted (44).

**TNFa**

The TNF-308 polymorphism is associated with AS in some European populations (65-67). Several small negative studies of TNF promoter polymorphisms have been reported, but these lack adequate power to detect the relatively small effects being investigated. The TNF-308.2 allele when found on B27-bearing haplotypes has been demonstrated to be associated with higher TNF production, perhaps explaining its protective effect (68). However a large negative study in British Caucasians demonstrated no association between AS and TNFa promoter polymorphisms, but confirmed the TNF-308.2 association in a southern German AS cohort, suggesting that TNF-308 may simply lie on a disease-associated haplotype rather than being involved in disease-susceptibility itself (67). Further studies of B27-TNF haplotypes will be required to determine this.

**Other class III genes**

Other genes that have been investigated within the MHC include MICA, TAP, LMP2 and LMP7, HSP70 and complements. Studies of TAP, LMP2 and LMP7 have yielded conflicting and at best weakly positive results, and further studies will be required to determine if these findings represent true association. Several studies of MICA have demonstrated no B27-independent association of MICA alleles and AS (36, 37, 69-71), making it quite unlikely that this gene has an independent role in AS. HSP70 –1, HSP-2 and HSP-hom lie between the C2 and TNFa polymorphisms (72). The study used general population controls rather than B27-matched controls, and for the HSP70-2 study, the control genotypes were not in Hardy-Weinberg equilibrium and are therefore unlikely to reflect the truly healthy population genotype frequencies. Two previous studies have found no association between HSP70 genotypes and AS (73, 74).

These studies all suggest that further MHC genes may be involved in aspects of the AS-phenotype, but at this point studies are not sufficiently definitive to directly implicate individual genes. Further large studies taking into account the complexity of linkage disequilibrium within the MHC will be required to define the actual gene or genes involved.

**Non-MHC genes and AS**

Whole genome scans aiming to broadly define the loci involved in AS are underway in several countries. Two whole genome studies in British Caucasians have been published (31, 75), involving a total of 188 families with 255 affected sibling pairs (31). These screens provide strong evidence as to the loci encoding the non-MHC genetic susceptibility to AS. Regions on chromosomes 1, 2, 6, 9, 10, 16 and 19 were identified with at least moderate linkage to the disease. The strongest linkage observed outside of the MHC is on chromosome 16q, where maximum linkage was observed at 101cM from the p–telomere (LOD = 4.7), equivalent to a genome-wide significance level of < 0.005 (76). Both screens showed significant support for this locus, with screen 1 achieving LOD = 4.1 at 106cM and screen 2 LOD = 1.2 at 99cM,
making it quite unlikely that either represents a chance finding. Independent support of this finding has been reported in a preliminary report on a further genome-wide screen (77). The region of linkage is very broad, with the 3-LOD confidence interval extends from 84-114cM, and contains numerous potential candidate genes. Further refinement of this interval by high density association/linkage disequilibrium mapping will be required to identify the actual genes involved. The magnitude of the genetic effect observed in affected sibling pair linkage screens is measured by the statistic I, which is the ratio of the observed / expected number of pairs sharing 0 alleles identical by descent. The magnitude of the chromosome 16q locus is L = 1.8 (95% CI 1.3-2.4), equivalent to 13% or 2.2% of the recurrence risk ratio for polygenic multiplicative or additive models respectively. This is roughly equivalent to the magnitude of the genetic effect of HLA-DRB1 in rheumatoid arthritis. Loci of this magnitude are quite likely not to be detected in future genome-wide screens for purely statistical reasons (78), and it is to be expected that many if not most true-positive findings in genome-wide scans will not be replicated in future scans. Even the MHC region has not been linked with AS in all studies (79). Researchers must not be discouraged by this problem in the pursuit of the non-B27 genetic component of AS.

Genome-wide screens are also underway in families collected in France (77), North America and Canada (80), and family collections in other centres. Over the next 5 years, over 1000 sibling pairs are likely to be screened, providing researchers with an excellent map of genetic susceptibility to AS. Positional candidate gene studies have now identified two non-MHC genes which are likely to be involved in AS susceptibility. Firstly, the cytochrome P450 2D6 gene (‘debrisoquine hydroxylase’, CYP2D6) has been implicated by case-control studies in two separate populations (81, 82), with moderate support from both within-family association methods (p = 0.01) and linkage studies (LOD = 1.0) (31, 82). The strength of association by case-control methods was extremely strong (p = 0.0007) but the relative risk of 2.1 only moderate and it is unlikely that this gene has a major role to play in AS-susceptibility. It is postulated that altered metabolism of a natural toxin or potential immunogen by the CYP2D6 gene may increase susceptibility to AS.

More recently, case-control association studies have been used to investigate the interleukin 1 complex which lies in a region on chromosome 2 linked to AS in genome-wide studies (31). Two studies have now reported over-representation of the IL1-receptor antagonist VNTR allele 2 in AS cases (83,84), and one negative study has been reported (85) which was too small to identify anything other than a large effect. This allele is reported to be associated with increased IL-1RA production, which would be expected to inhibit the proinflammatory cytokine, IL-1. Whether IL-1RA levels or production are elevated in AS is unknown. The IL-1 complex contains genes encoding many other important players in the immune system and it may be that the IL-1RA findings reflect linkage disequilibrium with another nearby gene. The ank/ank mouse develops ectopic calcium hydroxyapatite crystal deposition, leading to vertebral fusion resembling AS and chondrocalcinosis (86). This mouse model has been investigated in the past due to its similarity with human AS and it has been demonstrated that B27 carriage and immune suppression have no effect on the ank/ank phenotype (87, 88). This indicates that the progress of ankylosis in the mouse is not mediated by classical immunological pathways. The genetic defect in the ank/ank mouse has recently been identified (89). The defective gene is thought to be a membrane pyrophosphate transporter; dysfunction of the gene causes elevation of intracellular inorganic pyrophosphate (PPI) and reduction in extracellular PPI (89). PPI is a major inhibitor of calcium hydroxyapatite crystal deposition, hence extracellular deficiency of this metabolite may promote calcification. Studies are required to determine if polymorphism in this gene has any effect on either susceptibility to AS, or as seems more likely, the rate of ankylosis.

**Conclusion**

In conclusion, rapid progress has been made over the past 5 years in defining the genetic epidemiology of AS. The location of some of the genes encoding the non-MHC genetic risk for AS have been identified, and the studies commenced which should identify these genes. The next decade will certainly be an exciting period for research into the genetics of AS, with great potential to determine at least some of the underlying pathogenic mechanism in this poorly understood disease.

**References**

12. **SAID-NAHAL R, MICELI-RICHARD C, DOU-
Genetics of ankylosing spondylitis / M.A. Brown et al.


80. MCGARRY F, NEILLY J, ANDERSON N, STURROCK R, FIELD M: A polymorphism within the interleukin 1 receptor antagonist (IL-1Ra) gene is associated with ankylosing spondylitis. *Rheumatol Int* 2001; 40: 1359-64.


