

Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis

M A Brown, S H Laval, S Brophy, A Calin

Abstract

Objectives—It has long been suspected that susceptibility to ankylosing spondylitis (AS) is influenced by genes lying distant to the major histocompatibility complex. This study compares genetic models of AS to assess the most likely mode of inheritance, using recurrence risk ratios in relatives of affected subjects. **Methods**—Recurrence risk ratios in different degrees of relatives were determined using published data from studies specifically designed to address the question. The methods of Risch were used to determine the expected recurrence risk ratios in different degrees of relatives, assuming equal first degree relative recurrence risk between models. Goodness of fit was determined by χ^2 comparison of the expected number of affected subjects with the observed number, given equal numbers of each type of relative studied.

Results—The recurrence risks in different degrees of relatives were: monozygotic (MZ) twins 63% (17/27), first degree relatives 8.2% (441/5390), second degree relatives 1.0% (8/834), and third degree relatives 0.7% (7/997). Parent-child recurrence risk (7.9%, 37/466) was not significantly different from the sibling recurrence risk (8.2%, 404/4924), excluding a significant dominance genetic component to susceptibility. Poor fitting models included single gene, genetic heterogeneity, additive, two locus multiplicative, and one locus and residual polygenes ($\chi^2 > 32$ (two degrees of freedom), $p < 10^{-6}$ for all models). The best fitting model studied was a five locus model with multiplicative interaction between loci ($\chi^2 = 1.4$ (two degrees of freedom), $p = 0.5$). Oligogenic multiplicative models were the best fitting over a range of population prevalences and first degree recurrence risk rates.

Conclusions—This study suggests that of the genetic models tested, the most likely model operating in AS is an oligogenic model with predominantly multiplicative interaction between loci.

(*Ann Rheum Dis* 2000;59:883–886)

The importance of genes encoded within the major histocompatibility complex (MHC) in the aetiology of ankylosing spondylitis (AS) has been well established by association^{1,2} and linkage studies.^{3,4} Few families have been reported in which AS has occurred in the absence of HLA-B27 (B27) or in which AS

segregates separately from B27, suggesting that the inheritance of AS depends on the presence of B27.

There are considerable data suggesting that B27, though essential, is not sufficient to explain the genetic epidemiology of AS. Only a small proportion of B27 positive subjects develop AS, suggesting the presence of other susceptibility factors. Twin studies demonstrating high heritability of susceptibility to AS (97%) indicate that these factors are primarily genetic.⁵ The concordance rate for B27 positive dizygotic twin pairs (24%) is considerably lower than for MZ twins (63%), suggesting the presence of other genes influencing disease susceptibility. Gene mapping studies have implicated other genes, both within the MHC (for example, HLA-B60 and HLA-DR1) and distantly encoded (CYP2D6), which are likely to have small effects on disease susceptibility.^{6,7} Linkage studies in affected sibling pair families have also identified several non-MHC regions likely to contain further susceptibility genes.³

The recurrence risk in relatives of affected subjects (the proportion of relatives of an affected subject who develop the disease themselves) is determined by the number of genes involved, the magnitude of their individual effects, and the manner of their interaction. Risch has reported a non-parametric approach to studying different genetic models to assess likely modes of inheritance in genetic diseases employing familial recurrence risk ratios.⁸ In this study Risch's approach has been used to investigate the likely genetic models operating in AS.

Methods

Recurrence risk rates in different degrees of relatives of white patients with AS were determined using the results of all studies published in peer reviewed journals. Studies were included if they had either complete ascertainment of a population or proband ascertainment with careful control against ascertainment bias. Case reports or studies without specific measures to avoid ascertainment bias were not included. Table 1 outlines the studies meeting these criteria and their findings.

The population prevalence of AS was set at 0.1% (95% confidence interval 0.03 to 0.24%) from the studies of van der Linden and colleagues.⁹ This study was used because it was determined using screening methods similar in sensitivity to those employed by the recurrence risk studies. Alteration in the population prevalence directly influences the recurrence risk ratio (recurrence risk/population prevalence),

Spondyloarthritis
Research Group,
Wellcome Trust Centre
for Human Genetics,
Roosevelt Drive,
Headington, OX3 7BN,
UK

M A Brown
S H Laval

Royal National
Hospital for
Rheumatic Diseases,
Upper Borough Walls,
Bath, BA1 1RL, UK
S Brophy
A Calin

Correspondence to:
Dr Brown
mbrown@well.ox.ac.uk

Accepted for publication
10 April 2000

Table 1 Recurrence risk of ankylosing spondylitis in different degrees of probands of affected subjects

Study	Relationship	MZ twins	Siblings	Parent-child	First degree	Second degree	Third degree
de Blécourt ¹⁸	Number affected	—	24	5	29	8	7
	Number studied	—	401	191	592	834	997
	Recurrence risk (%)	—	6.0	2.6	4.9	1.0	0.7
Brown ⁵	Number affected	6	4	—	4	—	—
	Number studied	8	32	—	32	—	—
	Recurrence risk (%)	75	12.5	—	12.5	—	—
Calin ¹¹	Number affected	—	363	32	363	—	—
	Number studied	—	4298	275	4298	—	—
	Recurrence risk (%)	—	8.4	11.6	8.4	—	—
Carter ¹⁹	Number affected	—	10	—	10	—	—
	Number studied	—	169	—	169	—	—
	Recurrence risk (%)	—	5.9	—	5.9	—	—
Emery ²⁰	Number affected	1	0	—	0	—	—
	Number studied	2	4	—	4	—	—
	Recurrence risk (%)	50	0	—	0	—	—
Jarvinen ²¹	Number affected	3	3	—	3	—	—
	Number studied	6	20	—	20	—	—
	Recurrence risk (%)	50	15	—	15	—	—
Moesman ²⁰	Number affected	7	—	—	—	—	—
	Number studied	11	—	—	—	—	—
	Recurrence risk (%)	63.6	—	—	—	—	—
Total	Number affected	17	404	37	441	8	7
	Number studied	27	4924	466	5390	834	997
	Recurrence risk (%)	63.0	8.2	7.9	8.2	1.0	0.7
	(95% CI)	(42 to 81)	(7.4 to 9.0)	(5.6 to 10.7)	(7.4 to 8.9)	(0.2 to 1.7)	(0.1 to 1.4)
	Recurrence risk ratio	630	82	79	82	10	7

and different studies have reported different population prevalences. A recent study reporting a population prevalence of 1.0% used magnetic resonance imaging (MRI) scanning to screen for cases, a considerably more sensitive method than has been used previously in studies either of population prevalence or recurrence risk.¹⁰ It would be expected that the use of MRI scanning, or a similarly sensitive screening modality, in studies of recurrence rates would observe higher recurrence rates than those previously reported. Therefore models using prevalences in the range 0.1%–0.5% were tested against the observed data, but not as high as 1.0% as no comparable recurrence risk data are available.

Risch defines the recurrence risk ratio in relatives of type R (λ_R) as λ_R =recurrence risk in relatives of type R/population prevalence. In single gene models, $\lambda_R - 1$ falls by half with each increase in distance of relationship to the proband, such that:

$$\lambda_1 - 1 = 2(\lambda_2 - 1) = 4(\lambda_3 - 1).$$

Where there is no dominance variance component to disease susceptibility (see below), this relation can be extended to MZ twins, such that:

$$\lambda_{MZ} - 1 = 2(\lambda_1 - 1).$$

This relation is also true for polygenic models with additive interaction between loci, and approximately for genetic heterogeneity models (where the same phenotype may result from different susceptibility genes).

In polygenic multiplicative models, in which there is multiplicative interaction between large numbers of loci, the recurrence risk ratio falls by the square root with each increase in distance of relationship, such that:

$$\begin{aligned} \lambda_{MZ} &= \lambda_1^2 \\ \lambda_2 &= \lambda_1^{0.5} \\ \lambda_3 &= \lambda_2^{0.5} = \lambda_1^{0.25} \end{aligned}$$

Dominance variance is the component of genetic susceptibility due to interaction between alleles of a particular locus. Whereas there is a dominance variance component to the genetic correlation between siblings, this is

not the case between parents and children, who can only share one copy of any gene identical by descent. No significant difference was noted between parent-child and sibling recurrence risks (7.9% v 8.2%, $\chi^2=0$, odds ratio = 1.0, 95% confidence interval 0.7 to 1.4), indicating that any dominance variance component in AS must be small. In the largest study examining both types of first degree relatives, the parent-child rate was actually greater than the sibling recurrence rate, a finding not consistent with any genetic model.¹¹ The dominance and additive variance components to susceptibility can be calculated from the parent-child and sibling recurrence risk ratios assuming a single locus model, the model most affected by misspecification of dominance variance. The dominance variance component is calculated to be only 4% of the total variance. The estimated MZ recurrence risk ratio for a single locus model assuming dominance variance is 130, not significantly different from the figure estimated assuming no dominance variance (125). Therefore for all further calculations it was assumed that there is no significant dominance variance component in AS.

When a combination of these equations is used, any variety of genetic model can be studied. The models studied were single locus/additive/genetic heterogeneity, polygenic multiplicative, two locus multiplicative, one locus and residual polygenic multiplicative component, and five loci with multiplicative interaction and residual polygenic multiplicative component. More complex models of genetic interaction may exist but are not considered in this study. The models considered broadly reflect the likely models operating, and greater definition of the models, though increasing the precision of the estimates would not add greatly to the findings presented.

For the two locus multiplicative and one locus with residual polygenes models, one locus was assumed to be HLA, which in AS has a recurrence risk value (λ_{HLA}) of 3.6.¹² Oligogenic models using λ_{HLA} of 3.6, and 2–10 loci

Table 2 Observed and expected recurrence risk ratios for different genetic models assuming no dominance variance component and a population prevalence of 0.13%

	Observed	Single locus	Polygenic multiplicative	Two locus multiplicative	HLA, residual polygenes	Five locus multiplicative
MZ twins	630	163	6694	282	3202	780
First degree relatives	82	82	82	82	82	82
Second degree relatives	10	41	9	27	11	15
Third degree relatives	7	21	3	11	4	5

with equal λ and a residual polygenic component were fitted against the observed data. The best fitting oligogenic models had no residual polygenic component, and the best fitting such model had five genes, consisting of HLA and four genes each with $\lambda=2.2$. The first degree relative recurrence risk ratio was fixed at the observed figure of 82 for all models.

Goodness of fit of models was compared with the observed data by χ^2 analysis comparing the expected and observed number of affected MZ, second and third degree relatives for each model. For each model, the expected number of affected subjects for each degree of relationship was calculated from the predicted recurrence risk ratio and the number of subjects with that degree of relationship actually studied.

There is a clear bias in AS between the sexes, with men more commonly affected than women. Most available recurrence risk data are not specific for one sex, and therefore the effects of the sex of the subject were not considered in this study. It is assumed that the inheritance of AS in men and women follows similar genetic models, an assumption supported by the available recurrence risk data.¹¹

Results

Table 2 shows the expected recurrence risk ratios for each genetic model. The results clearly demonstrate that single gene, polygenic multiplicative, one locus (HLA) with residual polygenes models, and the two locus multiplicative model match the observed data poorly ($\chi^2=129, 90, 32, 36$; $p<10^{-6}, <10^{-6}, <10^{-6}, <10^{-6}$) respectively. The expected recurrence risk ratios for MZ twins under the polygenic multiplicative and one locus (HLA) are extremely high, but these models accurately estimate the recurrence risk ratio in more distant relatives. The single gene and two locus multiplicative models underestimate the MZ recurrence risk ratio and greatly overestimate the second degree relative recurrence risk ratios.

The best fitting model is an oligogenic multiplicative model with a small residual polygenic effect. The five locus with residual polygenes is not significantly different from the observed data ($\chi^2=1.4, p=0.5$), and estimates the recurrence risk in MZ twins, second and third degree relatives quite closely. Models with 3–9 genes fitted the observed data ($p<0.05$), but three gene models can be excluded as the magnitude of the genes involved ($\lambda=4.8$) would be larger than the MHC ($\lambda=3.6$). The presence of genes of this magnitude has been excluded by whole genome linkage studies, which clearly indicate that the MHC is the major susceptibility locus in AS.³

These findings were largely unaffected by the population prevalence chosen. Increasing the population prevalence alters the recurrence risk ratios for all types of relatives, but the models fitting the data were similar for a range of population prevalences. Oligogenic multiplicative models of 2–6 loci in addition to the MHC fit the observed recurrence risk rate data for population prevalences from 0.1 to 0.2% (data not shown). With a population prevalence of 0.5%, non-oligogenic models (monogenic, HLA and polygenic residual genes, and the polygenic model) were rejected, but the ability to differentiate between oligogenic models with different numbers of genes was poor, and models with 20 genes fitted the data similarly to those with only two (data not shown). The first degree relative recurrence risk rate varies between 0% and 25% in various studies, but in studies including more than 100 relatives, the rate is between 4.9% and 8.4%. Oligogenic multiplicative models are still favoured where lower first degree relative recurrence risk rates are chosen, but models with a larger number of genes are permitted. For example, if the true first degree relative recurrence risk is 50, then only oligogenic multiplicative models with two to 14 loci fit the observed data. Conversely, where higher first degree relative recurrence risk rates are employed, models with fewer genes are permitted. Overall, however, the finding that oligogenic multiplicative models are the best fitting is sustained over a range of population prevalence and first degree recurrence risk rates.

Discussion

This study suggests that the likely genetic model in AS is an oligogenic model with multiplicative interaction between loci. The number of loci involved and their individual magnitude is of great importance in determining the feasibility of mapping such genes. Genes of magnitude $\lambda \leq 1.3$ require extremely large sample sizes to be identified by either linkage studies in affected sibling pair families or by linkage disequilibrium mapping.¹³ The finding that single gene and polygenic models (with or without a contribution from the MHC region) were excluded by this study is therefore of great relevance. The data predict the existence of genes lying distant to the MHC of sufficient magnitude to be tractable to positional cloning.

The precise number of genes involved cannot be accurately modelled by such studies, particularly where the population prevalence increases relative to the recurrence rate, as the power to discriminate between multilocus models is small, with all such models giving

similar predictions. However, a model with only five genes models the observed data quite closely, and models with three to nine genes were consistent with the observed data. Models with 10 or more genes interacting multiplicatively fail to model accurately the observed recurrence risks, significantly overestimating the MZ twin recurrence risk ratio and underestimating the third degree relative recurrence risk ratio (data not shown). Oligogenic multiplicative models with two to four loci model the observed data closely, but only if the recurrence risk for each locus is fairly high (non-HLA loci $\lambda \geq 2.8$). Five locus models estimated the recurrence risk well with non-MHC loci of magnitude $\lambda = 2.2$. Genes of this magnitude are sufficiently large to be identified by current linkage-mapping techniques.

Although there are data supporting the autosomal dominant role of B27 in AS,^{14,15} the finding that the sibling recurrence risk ratio is not significantly greater than the parent-child recurrence risk ratio suggests that there is not a great dominance variance component to susceptibility to AS. This might be explained if there was a substantial non-MHC genetic component to AS, or, alternatively, if the MHC acted codominantly. If the latter situation is the case, then either B27 does not act dominantly, or there are other MHC susceptibility genes either operating in trans relative to B27, or acting recessively. HLA-B60 is one gene that in white populations acts in trans with B27 to increase susceptibility to AS,^{2,16} but there is increasingly strong evidence of other MHC susceptibility genes in AS, which may explain the apparent lack of a strong dominance genetic effect in this disease.¹⁷

Although this study supports the presence of non-MHC genetic susceptibility in this common rheumatic condition, the ultimate proof of this theory will be the isolation of the genes involved. Perhaps the most important conclusion from the modelling presented here is that the genes involved are likely to be of sufficient magnitude to be identifiable by current gene mapping techniques.

This study was funded by the Arthritis Research Campaign. Steven Laval is funded by the Oliver Bird Fund, Nuffield Foundation; Sinead Brophy by the Colonel WW Pilkington Charitable Trust; and Dr Andrei Calin by the Royal National Hospital for Rheumatic Diseases NHS Trust.

- 1 Brewerton DA, Hart FD, Nicholls A, Caffrey M, James DC, Sturrock RD. Ankylosing spondylitis and HL-A27. *Lancet* 1973;ii:904-7.
- 2 Brown MA, Pile KD, Kennedy LG, Calin A, Darke C, Bell J, et al. HLA class I associations of ankylosing spondylitis in the white population in the United Kingdom. *Ann Rheum Dis* 1996;55:268-70.
- 3 Brown MA, Pile KD, Kennedy LG, Campbell D, Andrew L, March R, et al. A genome-wide screen for susceptibility loci in ankylosing spondylitis. *Arthritis Rheum* 1998;41:588-95.
- 4 Rubin LA, Amos CI, Wade JA, Martin JR, Bale SJ, Little AH, et al. Investigating the genetic basis for ankylosing spondylitis. Linkage studies with the major histocompatibility complex region. *Arthritis Rheum* 1994;37:1212-20.
- 5 Brown MA, Kennedy LG, MacGregor AJ, Darke C, Duncan E, Shatford JL, et al. Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum* 1997;40:1823-8.
- 6 Beyeler C, Armstrong M, Bird HA, Idle JR, Daly AK. Relationship between genotype for the cytochrome P450 CYP2D6 and susceptibility to ankylosing spondylitis and rheumatoid arthritis. *Ann Rheum Dis* 1996;55:66-8.
- 7 Brown MA, Campbell S, Pile KD, Calin A, Ebringer A, Weeks DE, et al. Polymorphisms of the CYP2D6 gene increase susceptibility to ankylosing spondylitis [abstract]. *Arthritis Rheum* 1998;41:S355.
- 8 Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 1990;46:222-8.
- 9 van der Linden S, Valkenburg H, Cats A. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals: a family and population study. *Br J Rheumatol* 1983;22:18-19.
- 10 Braun J, Bollow M, Remlinger G, Eggens U, Rudwaleit M, Distler A, et al. Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. *Arthritis Rheum* 1998;41:58-67.
- 11 Calin A, Brophy S, Blake D. Impact of sex on inheritance of ankylosing spondylitis: a cohort study. *Lancet* 1999;354:1687-90.
- 12 Laval S, Timms A, Bradbury L, Edwards S, Rubin L, Simionovitch K, et al. Whole genome screening in ankylosing spondylitis confirms presence of major non-MHC susceptibility loci on chromosomes 7, 10, 16 and 19 [abstract]. *Rheumatology* 2000;suppl 1:83.
- 13 Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516-17.
- 14 Suarez AM, Russell AS. B27 homozygosity and ankylosing spondylitis. *J Rheumatol* 1987;14:302-4.
- 15 Arnett FC Jr, Schacter BZ, Hochberg MC, Hsu SH, Bias WB. Homozygosity for HLA-B27. Impact on rheumatic disease expression in two families. *Arthritis Rheum* 1977;20:797-804.
- 16 Robinson WP, van der Linden S, Khan MA, Rentsch HU, Cats A, Russell A, et al. HLA-Bw60 increases susceptibility to ankylosing spondylitis in HLA-B27+ patients. *Arthritis Rheum* 1989;32:1135-41.
- 17 Hohler T, Schaper T, Schneider PM, Meyer zum Buschenfelde KH, Marker-Hermann E. Association of different tumor necrosis factor alpha promoter allele frequencies with ankylosing spondylitis in HLA-B27 positive individuals. *Arthritis Rheum* 1998;41:1489-92.
- 18 de Blécourt J, Polman A, de Blécourt-Meindersma T. Hereditary factors in rheumatoid arthritis and ankylosing spondylitis. *Ann Rheum Dis* 1961;20:215-23.
- 19 Carter N, Williamson L, Kennedy L, Brown M, Wordsworth B. Sibling recurrence risk in ankylosing spondylitis. *Rheumatology (Oxford)* 2000;39:445.
- 20 Lawrence J. *Rheumatism in populations*. London: Heinemann, 1977:282-324.
- 21 Jarvinen P. Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis Rheum* 1995;38:381-3.