Linkage of rheumatoid arthritis to the candidate gene NRAMP1 on 2q35

Marie-Anne Shaw, David Clayton, Sara E Atkinson, Hazel Williams, Nancy Miller, Dean Sibthorpe, Jenefer M Blackwell

Abstract
The macrophage resistance gene NRAMP1 regulates priming/activation of macrophages for enhanced TNFα, IL1β, and MHC class II expression. Since all of these functions are of potential importance in the induction or maintenance of both autoimmunity disease, samples from the Arthritis and Rheumatism Council’s repository of multicae rheumatoid arthritis families were typed for a dinucleotide repeat in the NRAMP1 promoter region and four other 2q34 (TNP1) or 2q35 (IL8R, VIL1, DES) marker genes. Identity by descent (IBD) sib pair analysis using a three locus haplotype NRAMP1-IL8RB-VIL1, or NRAMP1 alone, provided preliminary evidence (maximum lod score = 1.01, p = 0.024) for a gene in this region contributing to susceptibility to rheumatoid arthritis, or Candida for NRAMP1 as the disease susceptibility gene was supported by a significant bias (p=0.048) towards transmission of the NRAMP1 promoter region allele 3 in affected offspring. (J Med Genet 1996;33:672–677)

Key words: linkage; rheumatoid arthritis; NRAMP1.

Two questions of fundamental importance in rheumatoid arthritis are (1) which genes confer disease susceptibility, and (2) does an infectious agent trigger disease? Epidemiological data suggest a two locus multiplicative model for genetic susceptibility, to which HLA-DR has been estimated to contribute 37%. The major gene(s) accounting for the remaining 63% has yet to be identified.

A plausible candidate for non-MHC gene regulation of early induction of rheumatoid arthritis is the human homologue, NRAMP1, for the murine macrophage resistance gene Ity/Lsh/Beg (NRamp1). Ity/Lsh/Beg was identified as a single gene controlling receptive susceptibility to the intracellular pathogens Salmonella typhimurium, Leishmania donovani, and Mycobacterium bovis BCG, and was mapped to the distal end of mouse chromosome 1. A candidate for Ity/Lsh/Beg was identified by positional cloning, a full length cDNA sequence obtained, and its candidacy confirmed by functional analysis of transfected macrophage cell lines. The gene has been renamed Nrampl, the Natural resistance associated macrophage protein. Through a region of conserved synteny on chromosome 2q35, human NRAMP1 has been sequenced and detailed physical and genetic mapping of the region undertaken. This work also showed that the IL8 receptor gene cluster, IL8RA, IL8RB, and the pseudogene IL8RBP, lies in close proximity (~130 kb distal) to the NRAMP1 gene.

Functional studies indicate that Nram1 regulates, and acts early in the pathway to, macrophage priming/activation for antimicrobial activity. The gene therefore has many pleiotropic effects on macrophage function, including differential upregulation of the early response gene KC, a neutrophil chemotactant belonging to the C-X-C family of small cytokines, as well as TNFα, IL1β, inducible nitric oxide synthase, and MHC class II expression, leading to higher levels in resistant macrophages. All of these functions are of potential importance in the induction or maintenance of both autoimmune disease. Hence, an animal resistant to intracellular pathogens may be susceptible to autoimmune disease. A number of agents may trigger Nrampl regulated macrophage responses, including interferon γ, bacterial lipopolysaccharide (LPS), and mycobacterial lipoarabinomannnan (LAM). A role for NRAMP1 in rheumatoid arthritis would therefore also be consistent with evidence for a bacterial/mycobacterial aetiology, either as a trigger for autoimmune or as an agent for the perpetuation of disease.

To test the hypothesis that NRAMP1 influences rheumatoid arthritis, samples from the Arthritis and Rheumatism Council (ARC) Repository were typed for a dinucleotide repeat in the NRAMP1 promoter region and for polymorphisms at four other 2q34/35 genes, and the data analysed using an affected sib pair maximum likelihood identity by descent (IBD) method of linkage analysis and transmission disequilibrium testing.

Material and methods
Families from the ARC National Repository have been described in detail by Worthington et al. and Hay et al. including diagnostic criteria and family structures. In this study, 61 nuclear families from the repository were analysed, described in Worthington et al. as families 001-046, 051-066, 075-109, 129-182, 196, 198-201, 205-210, 237, 238, 309, and 319. Five of these families had three affected offspring and one family had four affected offspring, giving a total of 76 affected sib pairs (although a weighting factor was applied in sib
Table 1  Pairwise lod scores (sex averaged) for linkage between pairs of genes on 2q33-q35 from 61 pedigrees analysed using LINKAGE

<table>
<thead>
<tr>
<th>Recombination fraction</th>
<th>0.001</th>
<th>0.01</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNP1-NRAMP1</td>
<td>-96.78</td>
<td>-2.26</td>
<td>-0.32</td>
<td>0.95</td>
<td>1.34</td>
<td>1.29</td>
<td>0.81</td>
</tr>
<tr>
<td>NRAMP1-IL8R</td>
<td>3.87</td>
<td>3.87</td>
<td>3.90</td>
<td>3.48</td>
<td>3.02</td>
<td>2.05</td>
<td>1.09</td>
</tr>
<tr>
<td>IL8R-VIL1</td>
<td>4.04</td>
<td>4.03</td>
<td>3.95</td>
<td>3.87</td>
<td>2.90</td>
<td>1.82</td>
<td>0.91</td>
</tr>
<tr>
<td>VIL1-DES</td>
<td>-94.55</td>
<td>3.03</td>
<td>3.94</td>
<td>4.16</td>
<td>3.81</td>
<td>2.71</td>
<td>1.51</td>
</tr>
</tbody>
</table>

Pair analyses for families with multiple sibs, reducing the effective number of sib pairs contributing to the analysis to 68, see below). Nineteen sib pairs had both parents typed, 10 sib pairs had one parent typed, and 32 had no parental DNA available for typing. The total number of DNA samples available for typing for 2q34/35 genetic polymorphisms was 318, 129 affected and 189 unaffected.

Samples from all family members were included in typing for a repeat polymorphism immediately 5' of position -594 bp in the NRAMP1 promoter region.11 Four variants occur within a possible enhancer element containing the Z-DNA forming dinucleotide repeat originally sequenced as t(gg)10. Alternates 1, 2, 3 and 3 vary in repeat number (11, 10, 9, 8, respectively) for the larger array of gt repeats. Allele 4 carries the sequence t(gg)10t(gg)10. All samples were also included in typing for a new restriction fragment length polymorphism (RFLP) in IL8R,12 and for RFLPs in genes proximal (TNP1, 2q34) and distal (VIL1 and DES, 2q35) to NRAMP1 and IL8R.13 The gene order is known to be TNP1-NRAMP1-IL8RA-IL8RB-VIL1-DES. The majority of samples provided a scorable genotype. Linkage between these genes was analysed by LIPED,17 using previously reported,12-15 gene frequencies: TNP1 A1=0.366 A2=0.634; IL8R A1=0.314 A2=0.686; VIL1 A1=0.672 A2=0.328; and DES A1=0.659 A2=0.341. NRAMP1 allele frequencies were those calculated by SPLINK (see below) in this analysis: A1=0.001 A2=0.999.

Linkage between markers and disease susceptibility was analysed by the affected sib pair method, using the maximum likelihood identity by descent (IBD) method pioneered by Risch11 and later extended by Holmans14 and Holmans and Clayton15 to allow for uncertainty of IBD assignment owing to unknown phase of marker haplotypes and to missing or incomplete parental genotype data. The statistical method involves the use of maximum likelihood to estimate the IBD sharing probabilities and haplotype frequencies simultaneously, and is implemented in the computer program SPLINK (available from the second author).

In the IBD analyses, affected sib trios were dealt with by treating each trio as the three independent pairwise comparisons, but giving each comparison a weight of only 2/3 in the analysis.22 Similarly the six possible pairwise comparisons of the affected 4-tuple were each given a weight of 1/2. This is a conservative strategy (P Holmans, D Clayton, unpublished observation). For each marker loci examined, the transmission of individual alleles from heterozygous parents to affected offspring was examined using the transmission disequilibrium test.23,24

Results

Use of a three locus haplotype

Maximum lod scores for linkage between NRAMP1-IL8R and IL8R-VIL1 occurred at a recombination fraction of 0 (table 1), indicating that there were no recombinants between these markers in the 61 rheumatoid arthritis families. A haplotype for these three loci was therefore used in the sib pair analyses, increasing the heterozygosity from 0.371 (NRAMP1 alone) to 0.766. Zero recombination between these markers is consistent with the physical NRAMP1-VIL1 interval of 155 kb and IL8R-VIL1 interval of 10-30 kb.16

Sib pair analysis

The results of the IBD analyses for all markers are summarised in table 2. IBD sharing probabilities (0:1:2 alleles IBD) for the three locus haplotype, estimated using the maximum likelihood method,14 were 0.118:0.500:0.382 giving a likelihood ratio chi-squared test statistic of 4.63. This provides a maximum lod score (MLS) of 1.01 which corresponds asymptotically to statistical significance at the p=0.024 level. The contributions which individual families make to these probabilities are represented diagrammatically in fig 1. Affected sib pairs whose IBD status is known exactly plot at the vertices of the triangles. The plotting positions of pairs calculated using maximum likelihoods reflect the relative probabilities of their assignments. In the left hand figure, the plotting positions for the probabilistic IBD assignments assume the null hypothesis (no linkage, 0.25:0.5:0.25). This plot represents

Table 2  Summary of IBD sib pair analysis using SPLINK. After weighting, the 76 affected sib pairs equated to 65-68 independent sib pairs contributing to the analysis for different genes. Effective sample size is calculated by SPLINK as the equivalent number of fully informative sib pairs. \( \chi^2 \) = likelihood ratio chi-squared test statistic. MLS = maximum lod score, \( p = \) probability, PIC = polymorphic information content

<table>
<thead>
<tr>
<th>Estimated IBD sharing probabilities</th>
<th>0-ibd</th>
<th>1-ibd</th>
<th>2-ibd</th>
<th>Effective sample size</th>
<th>( \chi^2 )</th>
<th>MLS</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNP1</td>
<td>0.250</td>
<td>0.500</td>
<td>0.250</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NRAMP1</td>
<td>0.137</td>
<td>0.420</td>
<td>0.453</td>
<td>13</td>
<td>4.64</td>
<td>1.01</td>
<td>0.024</td>
</tr>
<tr>
<td>NRAMP1-IL8R-VIL1</td>
<td>0.118</td>
<td>0.500</td>
<td>0.382</td>
<td>27</td>
<td>4.63</td>
<td>1.01</td>
<td>0.024</td>
</tr>
<tr>
<td>DES</td>
<td>0.182</td>
<td>0.500</td>
<td>0.318</td>
<td>19</td>
<td>0.90</td>
<td>0.20</td>
<td>0.171</td>
</tr>
</tbody>
</table>

NRAMP1: heterozygosity = 0.371, PIC score = 0.302.
NRAMP1-IL8R-VIL1: heterozygosity = 0.766, PIC score = 0.740.
TNP1: heterozygosity = 0.487, PIC score = 0.368.
DES: heterozygosity = 0.437, PIC score = 0.342.
null hypothesis

Maximum likelihood

Figure 1 Contributions of individual families to IBD haplotype sharing probabilities. An affected sib pair whose IBD status is known exactly is plotted at the vertices of the triangles. The plotting position of other pairs depends on the relative probabilities of their assignments. All 59 families informative for the three locus haplotype contribute to the maximum lod score. Open circles indicate single sib pairs while closed circles indicate the contribution of more than one sib pair. In the latter case the number of pairs is shown alongside the plotting point.

Table 3 NRAMP1-IL8R-VIL1 haplotype frequencies derived from 59 families/66 sib pair comparisons using SPLINK

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1-1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2-1-1</td>
<td>0.010</td>
</tr>
<tr>
<td>3-1-1</td>
<td>0.133</td>
</tr>
<tr>
<td>1-2-1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2-2-1</td>
<td>0.134</td>
</tr>
<tr>
<td>3-2-1</td>
<td>0.406</td>
</tr>
</tbody>
</table>

The evidence on which the significance test is based. The right hand figure uses the assignment probabilities at the final maximum likelihood estimate of the sharing probabilities. Summation of these assignments provides the maximum likelihood estimates. None of the 10 fully informative families showed 0-IBD for the NRAMP1-IL8R-VIL1 haplotype. The positions for the 49 families obtained using probability assignments depict the clear balance of probability away from the null hypothesis. In this maximum likelihood analysis, the marker haplotype frequencies are estimated internally. This guards against errors which may follow from misspecification by taking haplotype frequencies from inappropriate populations. The fact that these frequencies are estimates rather than known constants is allowed for in the likelihood ratio tests. The final estimates of the NRAMP1-IL8R-VIL1 haplotype frequencies are shown in Table 3.

Table 4 Transmission disequilibrium testing

<table>
<thead>
<tr>
<th>Allele</th>
<th>No of transmissions from heterozygous parents</th>
<th>No of families</th>
<th>No of transmissions for each allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNP1</td>
<td>54</td>
<td>38</td>
<td>1</td>
</tr>
<tr>
<td>NRAMP1</td>
<td>31</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>IL8R</td>
<td>21</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>VIL1</td>
<td>38</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>DES</td>
<td>52</td>
<td>35</td>
<td>1</td>
</tr>
</tbody>
</table>

When the IBD analysis was repeated for NRAMP1 alone (table 2), the maximum lod estimates of the IBD sharing probabilities were (0.127:0.420:0.453), again giving a likelihood ratio chi-squared test statistic of 4.64 and a MLS of 1.01. Hence, even without the additional polymorphic information content which the haplotype provides, a positive association between NRAMP1 and susceptibility to rheumatoid arthritis is observed. In contrast, no linkage between rheumatoid arthritis and the markers either proximal (TNP1) or distal (DES) to the three locus haplotype (table 2) was observed, even though the polymorphic information content for these markers was higher than for the NRAMP1 promoter region polymorphism.

TRANSMISSION TESTING

While NRAMP1 is a candidate locus, and linkage has been established to 2q35, it is possible that another gene in the region of the three locus haplotype is responsible for susceptibility. To address this issue, transmission of alleles from heterozygous parents was examined for the five 2q34/35 genes using the same set of families (table 4). Since problems may arise in transmission testing where parental genotype information is incomplete, only matings where parental genotypes were conclusively assigned were used in the analysis. In the majority of cases this meant families where both parents were available for typing. For NRAMP1, two families were included where the heterozygous parental genotype was deduced on the basis of genotypes of offspring from the extended family, that is, unaffected as well as affected offspring were available. In both cases, there was only one possible parental genotype which could be assigned, and this could be done without reference to the affected person being tested for transmission disequilibrium. The only significant deviation from the expected 1:1 transmission of the two alleles to affected sibs under the null hypothesis was obtained for NRAMP1, where allele 3 was
transmitted to affected children in preference to allele 2 (chi-squared=3.9, p=0.048). In testing for transmission disequilibrium, multiple independent tests have been carried out. Hence, it could be argued that a multiple testing correction factor should be applied. Multiplying our p values by the number of tests performed (2 alleles x 5 loci = 10) would clearly remove all significance from the NRAMP1 TDT allelic association in this small dataset. However, as Coyman et al26 argue, if we had tested only the loci showing independent evidence of linkage (NRAMP1), a correction factor of only 2 (= 2 alleles, 1 locus) could be applied. Since we could also argue that a one tailed test could be applied in comparing alleles 2 versus 3, that is, that we would expect a one way skew towards allele 3 in affected offspring on functional grounds (S Searle, J M Blackwell, unpublished data, see Discussion), our original p value for NRAMP1 TDT could be halved. Multiplying by 2 would then retain the TDT for NRAMP1 within the 5% significance level.

### Discussion

Results of the IBD sib pair analysis presented here provide evidence (MLS=1.01, p=0.024) for the presence of a gene contributing to susceptibility to rheumatoid arthritis in the region of the NRAMP1-IL8R-VIL1 haplotype on human chromosome 2q35. Since this region contained an a priori candidate gene (NRAMP1) for disease susceptibility on functional grounds, there was no multiple testing problem. We have performed a single test of the null hypothesis of no linkage to 2q35, with a resulting pointwise nominal significance level31 of 2.4%. Since this region was not identified as part of a genome wide search, the more stringent criteria for genome wide significance levels suggested by Lander and Kruglyak31 do not apply. Candidacy for NRAMP1 as the disease susceptibility gene within 2q35 was supported by (1) IBD analysis for NRAMP1 alone providing an equivalent MLS of 1.01, and (2) a significant bias (p=0.048) towards transmission of the NRAMP1 promoter region allele 3 in affected offspring. The latter observation suggests that the NRAMP1 promoter region polymorphism itself may be the disease associated genetic defect in these families. Although a role for IL8R genes cannot be discounted, NRAMP1 itself does provide the more plausible functional candidate for non-MHC control of susceptibility to rheumatoid arthritis. Its pleiotropic effects5-7 include roles in cytokine production, in particular TNFα and IL1β, and in MHC class II antigen presentation to CD4+ T cells. There is evidence that macrophages are directly involved in the inflammatory process in rheumatoid arthritis, and that TNFα plays a major role in autoimmune disease. IL1β and TNFα are both expressed by synovial tissue macrophages in rheumatoid arthritis.32,33 GM-CSF is also produced in large quantities by activated macrophages in the synovium, and acts to induce IL-1β production and enhance class II expression.34 Hence, there is good evidence that the activation phenotypes associated with NRAMP1 regulated macrophage activation could contribute directly to the inflammatory response and pathology associated with rheumatoid arthritis. Results of continuing research in our laboratory (S Searle, J M Blackwell, unpublished data) analysing promoter function in a reporter gene system indicate that allele 3 of the NRAMP1 promoter region polymorphism drives NRAMP1 expression more efficiently than allele 2. This would be consistent with overexpression of the NRAMP1 gene, and hence hyperactivation of macrophages, contributing to disease susceptibility. Although additional functional polymorphisms/mutations in NRAMP1 may be found, there is good evidence that this promoter region polymorphism may represent the functional 2q35 polymorphism contributing to disease susceptibility in families analysed here.

In previous studies, polymorphic HLA haplotypes have been shown to contribute to disease susceptibility in rheumatoid arthritis, either through the direct role of MHC class II molecules in presenting antigen to autoim-

<table>
<thead>
<tr>
<th>NRAMP1-IL8R-VIL1 (%)</th>
<th>HLA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-IBD 1.8</td>
<td>2.3</td>
</tr>
<tr>
<td>1-IBD 3.9</td>
<td>3.5</td>
</tr>
<tr>
<td>2-IBD 6.0</td>
<td>6.2</td>
</tr>
</tbody>
</table>

### Results of the IBD sib pair analysis

Table 5  Risk of rheumatoid arthritis in sib sharing 0, 1, or 2 NRAMP-IL8R-VIL1 2q35 or HLA haplotypes IBD with an affected sib
immune T cells or through linkage disequilibrium with polymorphic elements regulating TNFα production.14 The possible interaction between NRAMP1 and MHC class II regulation of disease susceptibility is of particular interest. One hypothesis for polygenic control in rheumatoid arthritis is that non-MHC genes determine susceptibility to disease per se and act early in the induction phase of autoimmunity, whereas HLA determines disease severity through modulation of T cell responses to particular antigens.15,16 This parallels observations in the mouse where Nramp1 influences the early macrophage response to infection, and hence susceptibility to disease per se, while polymorphism in class II molecules at the mouse H-2 locus influences disease severity.17 For control of infection, the best possible interaction between genes occurs when an Nramp1 resistant mouse carries an H-2 curing haplotype, leading to enhanced antigen presentation to protective interferon γ generating CD4+ T helper 1 cells. In autoimmune disease, this might represent the worst possible interaction. T cells, particularly CD4+ cells, contribute to pathology in rheumatoid arthritis. The involvement of cells recognising antigen in association with HLA class II molecules fits well with the association of rheumatoid arthritis with particular DR types18 and also points to a significant role for antigen presenting cells such as macrophages. Further analysis of the interaction between DR and NRAMP1 in determining RA susceptibility will be of particular interest.

NRAMP1 as a candidate for susceptibility to rheumatoid arthritis is also interesting in relation to a bacterial/mycobacterial aetiology for the disease. The gene was first identified for its role in controlling resistance to intramacrophage pathogens. Many arthritic diseases have been shown to have intracellular pathogens as triggers, and the hypothesis that rheumatoid arthritis is a slow bacterial infection with autoimmunity a secondary consequence has not been disproved.19 In the case of rheumatoid arthritis, no self antigen has been shown to be associated with the initiation of disease. Mycobacteria contain inducers of TNF-α, the most potent of which is LAM,20 while organisms associated with reactive arthritis have LPS which can be identified in affected joints.21 Both LAM and LPS are important triggers of Nramp1 regulated macrophage activation.22 Support for a role of mycobacteria in arthritis has come from the use of Lewis rats which are susceptible to adjuvant induced arthritis and respond particularly to the mycobacterial 65 kDa hsp. Indeed, arthritis is inducible by a single T cell clone recognising the non-conserved 180-188 region of mycobacterial hsp65.23 T cells isolated from the synovial fluid of patients have been found to give a localised response to both conserved and non-conserved mycobacterial hsp65 epitopes.24 The reactivity against non-conserved epitopes of hsp65 cannot be accounted for by cross reactions with the human homologue. Again, the interaction between NRAMP1 and MHC class II molecules in directing the T cell response to bacterial/mycobacterial antigens may be of particular importance in initiation and maintenance of disease.

The primary function of NRAMP1 remains unknown. While its molecular structure is compatible with a transport function, the presence of an SH3 binding domain at the N-terminus of the molecule points to a role in signal transduction.25 In the murine system, the final effector mechanism for antimicrobial activity relies on induction of nitric oxide synthase and generation of nitric oxide.26 This, or any one or combination of the multiple pleiotropic effects of the gene, would be sufficient to account for a role for NRAMP1 in regulating autoimmune disease. In this study, only 47% of families had one (16%) or both (31%) parents of affected sibs available for study, reducing the power of our analysis in terms of equivalence to fully informative affected sib pairs (table 2). This is a general problem in late onset diseases. Our results are therefore preliminary, and require additional information in additional datasets before the real contribution of NRAMP1 to disease susceptibility in rheumatoid arthritis is known. Functional studies analysing NRAMP1 expression in synovial tissue macrophages would also be of interest, and further studies regarding activation phenotypes in macrophages from people bearing different NRAMP1 genotypes. These continuing studies of NRAMP1 in rheumatoid arthritis patients and families provide an exciting basis to further genetic and functional characterisation of disease susceptibility.

This work was supported by grants from the Arthritis and Rheumatism Council and the Wellcome Trust. We would like to thank Dr Bill Ollier, ARC Epidemiology Research Unit, Manchester University, for his help in making material available from the ARCC Repository.

Linkage of rheumatoid arthritis to NRAMP on 2q35


