Chapter 5

Single Nucleotide Polymorphisms (SNPs) in the C5 gene associate with susceptibility to RA.

Abstract

Rheumatoid arthritis (RA) is the most common inflammatory arthritis and is a major cause of disability. Early theories on the pathogenesis of RA focused on autoantibodies and immune complexes. T cell-mediated antigen-specific responses, T cell-independent cytokine networks, and aggressive tumor-like behaviour of rheumatoid synovium have also been implicated. More recently, the contribution of autoantibodies has returned to the forefront. The complement network was initially implicated in human RA, indirectly, by the co-localization of C3 fragments with immune complexes in joint tissue, and by the demonstration that complement activity, as well as early-acting components (C2, C4), is routinely depressed in synovial fluid of patients. More recently, more direct evidence for complement activation in arthritic joints has been reported. Mice deficient for C5 are resistant to serum induced arthritis and anti-C5 monoclonal antibody treatment prevents arthritis in mice. Therefore, it is tempting to speculate that the complement system also plays an important role in disease pathogenesis in human.

Here we show that genetic polymorphisms located in the C5 gene locus associate with RA, as a common haplotype consisting of 3 SNPs located at positions rs25681, rs17611 and rs2416808 is significantly overrepresented in the RA-population. These data indicate that this haplotype in the C5 region confers susceptibility to RA.
**Introduction**

The complement system plays an important role in the immune system, providing a highly effective means for the destruction of invading micro-organisms and for immune complex elimination\(^1,2\). Three pathways of complement activation have been described, the classical pathway, the alternative pathway and the lectin (i.e. mannan-binding lectin and ficolins) pathway. All lead to cleavage of C3, into C3a and C3b. Binding of C3b enables a better clearance of pathogens and immune complexes as well as the generation of the lytic membrane attack complex, C5b-9. C5, the fifth component of the complement system is a 190,000 M\(_r\) glycoprotein consisting of 1666 aa in two disulfide-linked polypeptide chains, C5\(\alpha\) and C5\(\beta\). When activated by the C5 convertases, C5 is cleaved into C5a and C5b. C5a is the most potent complement-derived proinflammatory peptide\(^3\). It serves not only as a chemoattractant but also initiates multiple defence mechanisms in leukocytes. Through its multiple binding sites, C5b initiates and directs the assembly of the membrane attack complex.

C5 is encoded by a single gene copy on chromosome 9q34.1 which contains 41 exons that span a genomic region of 79 kb. The gene encodes the open reading frame for C5\(\alpha\) (exon 1-16), as well as for C5\(\beta\) (exons 16-41). Both are transcribed as a single 6.0 kb mRNA and translated into a pre-C5 protein (in a \(\beta\)-\(\alpha\) orientation). These single chain precursors are processed into the mature form of C5 by the removal of the four internal residues to yield a two-chain protein\(^4\).

Through activation and interaction with respective receptors on various cells, complement is closely linked to the adaptive immune response. A proper functioning complement system is also required for physiological tissue regeneration and repair. Complement defects increase the susceptibility to infection and are frequently associated with autoimmune disorders\(^5\) such
as, for example rheumatoid arthritis (RA). RA is a chronic inflammatory disease that attacks peripheral joints, leading to their erosion and destabilisation. Although the causes for RA are still elusive, it is generally accepted that autoimmune reactions towards joints are of critical importance.

K/BxN T cell receptor transgenic mice are a model of inflammatory arthritis, similar to rheumatoid arthritis. In this arthritis model, both T and B cells are required for disease initiation, but arthritogenic immunoglobulins (Igs) alone can induce arthritis in lymphocyte-deficient recipients. These observations indicate that these antibodies are primarily driving the disease. Subsequently, it was shown that Fc receptors as well as the complement network are crucial to disease induction and progression. Likewise, mice lacking FcRγ were not susceptible to arthritis induction upon injection of collagen or adjuvant; and lack of the inhibitory receptor FcγRIIB was found to exacerbate collagen-induced arthritis (CIA) in susceptible mouse strains or permit its induction in normally resistant strains.

Whether FcγR are also involved in RA is not known, but FcγRIII is detected on synovial intima in normal and arthritic human joints and on invading macrophages in the latter. Moreover, an FcγRIII gene polymorphism has been correlated with human RA susceptibility indicating that also in humans FcR are involved in disease pathogenesis.

That the complement network plays a role in human RA was for the first time detected, indirectly, by the colocalization of C3 fragments with immune complexes in joint tissue, and by the demonstration that complement activity, as well as early-acting components (C2, C4), is routinely depressed in synovial fluid of patients. More recently, more direct evidence of complement activation in arthritic joints has been reported. In murine models of RA, especially CIA, C5 deficiency has frequently been correlated with disease resistance, although this has not always been the case. Thus, there is evidence implicating both FcRs and the complement network in the aetiology of RA.
As mice deficient for C5 are unsusceptible to K/BxN serum induced arthritis$^{17}$ and anti-C5 monoclonal antibody treatment prevents arthritis in mice$^6$, we wished to address the question whether C5 would also be implicated in RA by determining whether certain genetic variance in the C5 locus is associated with RA.
Materials and Methods

Patients

The distribution of C5 alleles was determined in 289 consecutive RA patients who attended the out-patient clinic of the Leiden University Hospital in 1994\textsuperscript{18}.

A standard diagnostic evaluation was performed at the first visit. This consisted of a patient history, physical and laboratory examinations, and radiographs of the hands and feet\textsuperscript{19}. Baseline laboratory examination included an enzyme-linked immunosorbent assay (ELISA) for IgM-RF as previously described\textsuperscript{20} and an anti-CCP-2 antibody ELISA (Immunoscan RA Mark 2; Euro-Diagnostica, Arnhem, The Netherlands) which was performed according to the manufactures instructions (sensitivity 74\%, specificity 97-99\%). Patients included in this study all had a definite diagnosis of RA according to the American College of Rheumatology (formerly, the American Rheumatism Association) 1987 revised criteria\textsuperscript{21} 1 year after inclusion in the study.

SNP typing

Single Nucleotide Polymorphism (SNP) typing in the C5 gene were performed by polymerase chain reaction using specific primers (information on the primer sequence and thermal cycling conditions will be provided on request) and by PCR-restriction fragment length polymorphism analysis. rs25681 and rs17611 were genotyped using time-of-light mass spectrometry based Sequenom Mass Array Platform (Sequenom, San Diego, CA, USA).

Quality Control

Data for each SNP were reviewed independently to verify their quality. We considered each SNP to be validated and correctly genotyped for all individuals only if the genotype data met
a series of strict criteria including minimum signal intensity specifications and unambiguous genotype reading. Replication of the genotyping of 10% of the samples for each SNP was performed randomly and no inconsistencies were detected.

**Statistical analysis**

All analysis was performed using SPSS and Statcalc (Centers for Disease Control and Prevention, Atlanta, GA). Haplotype frequencies were inferred using Haplovie\textsuperscript{22}. P values less than or equal to 0.05 were considered significant.
Results

Polymorphisms in the complement factor 5 gene

The human C5 molecule lies on the long arm of chromosome 9, at loci 9q34.1. It contains 97938 bases, running from position 120794170 up to position 120892108, according to the NCBI database, and codes for 1676 amino acids. The gene consists of 41 exons, containing 79 kb. A large number of SNPs are present in the gene, comprising the introns, the exons and the untranslated regions (UTR’s) that span the gene. Many SNPs of the C5 gene are already registered in the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP). 3 SNPs spanning the C5 gene were chosen, to investigate their association with RA (figure 1).

**C5 gene**

![C5 gene diagram](image)

**figure 1**: the position of the three genotyped SNPs (rs25681, rs17611 and rs2416808) on the C5 gene.
Table 1: Characteristics of 289 of patients with arthritis of the Leiden University Hospital

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at onset (mean – SD)</td>
<td>49.9 (14.0)</td>
</tr>
<tr>
<td>Female</td>
<td>61%</td>
</tr>
<tr>
<td>Rheumatoid factor positive</td>
<td>89%</td>
</tr>
<tr>
<td>SE positive</td>
<td>73%</td>
</tr>
<tr>
<td>DRB1.04 positive</td>
<td>22%</td>
</tr>
<tr>
<td>Erosions of hand and feet</td>
<td>90%</td>
</tr>
<tr>
<td>Mean disease duration (years)</td>
<td>13.5 (9.9)</td>
</tr>
</tbody>
</table>

Association of SNPs with rheumatoid arthritis

289 RA patients and 250 controls were genotyped for all 3 SNPs (Table 1). There was no significant deviation from Hardy-Weinberg equilibrium in the cases or in the controls (Table 2).

A Chi-square trend test was first performed to determine whether there was any difference in genotype distribution between cases and controls, followed by an allelic and recessive model analysis.

Table 2: Case control analysis of C5

<table>
<thead>
<tr>
<th>SNP</th>
<th>Alleles</th>
<th>Cases</th>
<th>Controls</th>
<th>Allele Frequency</th>
<th>OR</th>
<th>P</th>
<th>OR</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1/2</td>
<td>11 12 22</td>
<td>Total</td>
<td>Cases 11 12 22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2416808</td>
<td>G/A</td>
<td>59 122 62</td>
<td>243</td>
<td>33 127 79 223</td>
<td>0.494</td>
<td>1.44 2.00</td>
<td>0.0049</td>
<td>0.0034</td>
</tr>
<tr>
<td>rs17611</td>
<td>G/A</td>
<td>90 117 44</td>
<td>251</td>
<td>71 125 51 247</td>
<td>0.592</td>
<td>1.23 1.39</td>
<td>0.1034</td>
<td>0.089</td>
</tr>
<tr>
<td>rs25681</td>
<td>C/T</td>
<td>90 117 44</td>
<td>251</td>
<td>72 124 51 247</td>
<td>0.592</td>
<td>1.22 1.36</td>
<td>0.1177</td>
<td>0.11</td>
</tr>
</tbody>
</table>

SNP rs2416808 (G) is significantly associated with RA whereas SNPs rs17611 (G) and rs256181 (C) show a trend. The alleles are overrepresented in the RA population and point towards a recessive effect as seen in the significance level (p<0.05) for the recessive model analysis.
Linkage disequilibrium (LD) analysis reveals that the three SNPs are linked (table 3) with SNPs rs17611 and rs25681 being in complete LD (D’=1, R^2=0.99)

Table 3: Linkage disequilibrium analysis

<table>
<thead>
<tr>
<th>SNP1</th>
<th>SNP2</th>
<th>D’</th>
<th>R^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2416808</td>
<td>rs17611</td>
<td>0.734</td>
<td>0.35</td>
</tr>
<tr>
<td>rs2416808</td>
<td>rs25681</td>
<td>0.733</td>
<td>0.347</td>
</tr>
<tr>
<td>rs17611</td>
<td>rs25681</td>
<td>1.000</td>
<td>0.996</td>
</tr>
</tbody>
</table>

Since the level of LD is relative high (as shown by figure 2) it indicates that the effect seen might be attributed to a haplotype effect.

**figure 2:** Linkage disequilibrium level.

**Haplotype effect**

To investigate this hypothesis, haplotype frequencies were inferred for cases and controls. The 3 SNPs (rs25681, rs17611 and rs2416808) were used to determine the haplotype frequencies. As seen from Table 4, haplotype GGC is the most frequent in the population.
The region shows a high level of linkage disequilibrium explaining the prevalence of only 4 haplotypes out of the $2^3$ possible haplotypes. The cut off value for reliable haplotype frequency estimation was 2%. The 4 major haplotypes account for 99-100 % of the total population. Haplotype GGC is significantly overrepresented in the RA population (OR=1.26 P=0.0028), indicating that haplotype GGC confers susceptibility to RA.

**Table 4: haplotype structure and frequency in C5**

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>Frequency</th>
<th>Case/Control Ratio</th>
<th>Case/Control Frequency</th>
<th>ChiSq</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGC</td>
<td>0.401</td>
<td>234:292, 179:327</td>
<td>0.446, 0.354</td>
<td>8.964</td>
<td>0.0028</td>
</tr>
<tr>
<td>AAT</td>
<td>0.383</td>
<td>190:336, 205:301</td>
<td>0.362, 0.405</td>
<td>2.054</td>
<td>0.1518</td>
</tr>
<tr>
<td>AGC</td>
<td>0.163</td>
<td>76:450, 93:413</td>
<td>0.144, 0.184</td>
<td>3.043</td>
<td>0.0811</td>
</tr>
<tr>
<td>GAT</td>
<td>0.052</td>
<td>26:500, 28:478</td>
<td>0.049, 0.055</td>
<td>0.18</td>
<td>0.6714</td>
</tr>
</tbody>
</table>
**Discussion**

Previously, it has been shown that mice deficient for C5 are resistant to K/BxN serum induced arthritis\(^{17}\). Moreover, anti-C5 monoclonal antibody treatment prevents arthritis in mice\(^6\), indicating that C5 is crucially involved in the process that contributes to arthritis. In this study, we investigated the contribution of the C5 gene in human rheumatoid arthritis.

We show that in the human C5 locus, SNPs rs25681 (C), rs17611 (G) and rs2416808 (G) are significantly associated with the susceptibility of RA. The alleles are overrepresented in the RA population and point towards a dominant effect. The association appeared to be a haplotype effect. Using the 3 SNPs (rs 25681, rs17611 and rs2416808), haplotype CGG is the most frequent in the population, and it is significantly overrepresented in the RA population. However, to verify these findings replication in other RA populations is required. Likewise, a more extensive analysis should be performed by incorporating more SNPs that are also localized outside the C5 gene in order to determine whether these associations are, indeed, contributed by C5 or are a consequence of genes that are located in the vicinity of the C5 gene. Moreover, such an analysis will be of importance to identify the causal SNP that is underlying the observed associations.

In conclusion, we provided evidence for association of a particular haplotype of the C5 gene with development of RA. Further studies should elucidate the precise nature of this association, thereby contributing to the further understanding of RA.
References


