ETHNIC DIVERSITY OF CLASS III GENES IN AUTOIMMUNE DISEASE

Joann M. Moulds

Department of Microbiology and Immunology, MCP Hahnemann University School of Medicine

TABLE OF CONTENTS

1. Abstract
2. Fourth component of complement (C4)
   2.1. Variability in C4 gene size
   2.2. Polymorphism of the C4A and C4B genes
   2.3. Chido/Rodgers blood group polymorphism
   2.4. C4 genes and disease
      2.4.1. Systemic lupus erythematosus
      2.4.2. Other autoimmune diseases
3. Second component of complement (C2)
4. Factor B (Bf)
5. Tumor necrosis factor (TNF)
6. Conclusions
7. Acknowledgments
8. References

1. ABSTRACT

One may wonder why the genes for the class III region are even situated in the human Major Histocompatibility Locus (MHC). However, upon closer inspection we find that the genes for the complement components, tumor necrosis factor (TNF) and others may play an important role in host immune defenses. Thus, this region on chromosome six may be more appropriately thought of as an immune response area. Accordingly, there is a high degree of polymorphism in many genes within the MHC, the most notable being the HLA genes. In the class III region C2, factor B (Bf) and C4 are polymorphic in as many populations studied to date. The purpose of this review will be to briefly describe the class III region, identify the genetic polymorphism found in various ethnic groups and define their roles in autoimmune diseases.

2. FOURTH COMPONENT OF COMPLEMENT (C4)

2.1. Variability in C4 Gene Size

The class III region of the MHC contains numerous genes, many having unknown function and includes: the opposite strand gene (OSG), second component of complement (C2), fourth component of complement (C4A & C4B), 21-hydroxylase (CYP21A & CYB21B), RP1 & RP2, RD, Factor B (Bf), heat shock protein (Hsp70), tumor necrosis factor (TNF), tenascin (TNXA & TXNB), several HLA-B associated transcripts (BAT) genes and ~20 genes yet to be named (1). This highly variable region is characterized by polymorphisms, variation in gene size and variable gene number. The genes for C2-Bf-C4A-C4B-CY21B cover an area of 120 kb and traditionally have been considered as a group or “complotype”. More recently, the genes RP and TNX have been shown to be present between Bf and C4 and form a genetic module termed RCCX (1). Other genes mapped to this region including RD, SKI2W and DOM3Z, but their functions are unknown at this time (2). For a more detailed description of the Class III MHC genes, the reader is referred to the review by Yu (1).

Plasma C4 exists as two functionally distinct isotypes ie. C4A and C4B, whose genes exhibit a size difference. Although most C4A genes are of the “long” variety (21 kb), the C4B genes can either be long or “short” (14.6 kb). The added length of the long genes is due to a retroviral gene insertion (HERV-K) that adds an additional 6.36 kb. C4A genes usually have this insertion within intron nine while the C4B genes may or may not carry the insert. Of note, the HERV-K insert is lost in the classical C4A gene deletion that removes 18 kb of both the C4A and C4B genes and which can be identified by a 6.4 bp Taq I RFLP (3). The majority of studies determining the size and number of C4 genes have been done in Caucasians where the 6.4 kb fragment is a marker for the C4 gene deletion. Although the presence of short C4 genes at both loci is rare in Caucasians, this gene structure has been found in high frequency among the Kaingang South American Indians (4).

The C4 genes can occur in variable copy number. In the more common bimodular genotype, C4A occupies the 5’ locus while C4B usually is at the 3’ locus. Using RFLP, gene duplications were found in the Kaingang Indians which were not observed on protein allotyping thus
Ethnic Diversity of Class III

Table 1. C4 subtypes based on PCR typing of Rg and Ch genes

<table>
<thead>
<tr>
<th>C4A</th>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ch-5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ch-4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rg-3</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Rg-1</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Ch+5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Ch+4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Ch+6</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ch+1</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
</tbody>
</table>

C4B: 01 02 03 04 05 06 07 08

| Ch-5 | +  | +  | +  | +  | 0  | 0  | 0  | 0  |
| Ch+4 | +  | +  | +  | +  | +  | +  | 0  | 0  |
| Ch+6 | +  | +  | +  | +  | +  | +  | +  | +  |
| Ch+1 | +  | +  | 0  | 0  | +  | +  | +  | +  |
| Ch+5 | 0  | 0  | 0  | 0  | +  | +  | +  | 0  |
| Ch+4 | 0  | 0  | 0  | 0  | +  | 0  | +  | 0  |
| Ch+6 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| Rg-3 | 0  | 0  | 0  | 0  | +  | 0  | +  | 0  |
| Rg+1 | 0  | 0  | +  | +  | 0  | 0  | +  | +  |

Modified from Schneider et al. (7)

Table 2. Chido/Rodgers blood group phenotypes in two ethnic groups

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Caucasians</th>
<th>Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rg: 1.2</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>Rg: 1.2</td>
<td>2.5%</td>
<td>0</td>
</tr>
<tr>
<td>Rg: -1.2</td>
<td>2.5%</td>
<td>0</td>
</tr>
<tr>
<td>Ch: 1.2,3</td>
<td>88%</td>
<td>75%</td>
</tr>
<tr>
<td>Ch: 1.2,3</td>
<td>5%</td>
<td>24%</td>
</tr>
<tr>
<td>Ch: 1.2,3</td>
<td>3%</td>
<td>0</td>
</tr>
<tr>
<td>Ch: -1.2,3</td>
<td>4%</td>
<td>1%</td>
</tr>
</tbody>
</table>

suggested that the extra genes are not expressed. Duplication of the C4A locus has also been found in Australian Aborigines (5).

2.2. Polymorphism of the C4A and C4B Genes

Genetic typing for allelic variants of C4 is routinely carried out using high voltage agarose gel electrophoresis often in conjunction with functional (hemolytic overlays) or immunological (Western blot) methods. The C4B alleles are more hemolytically active than C4A alleles since their reactive thioester bond preferentially binds to hydroxyl groups that are found in abundance on the red cell membrane. Most C4A alleles code for the Rodger (Rg) blood group antigens while C4B alleles code for Chido (Ch) antigens. Thus, using these methods almost 50 variants of the C4 genes have been defined (6). Schneider et al. (7) have described a genomic PCR-based method to type for the Chido/Rodgers antigens and have proposed a new nomenclature based on a combination of the above methods (Table 1).

Because of the association of C4 null genes, C4*Q0, with autoimmune diseases (see below), many population studies of C4 allotypes have been reported. Caucasians are probably the largest group studied but many other races have been C4 phenotyped. In all cohorts C4A*3 and C4B*1 are the two most common alleles at the A and B loci, respectively. C4 allotypes have been useful in tracing populations; e.g., there is a close similarity of complement haplotypes between the British and the immigrants to Tasmania (8).

Studies of Africans and African derived populations have found increased frequencies of some unusual alleles including: C4A*91 and C4B*92 when compared to other races. But even among one ethnic group the actual frequency may differ based on location. For instance, the C4A*91, C4B*1 haplotype had a frequency of 11.8% in Khoi but only 2% in Xhosas, South Africa (9). When the C4A*91 was paired with the C4B*Q0 allele (found in the extended haplotype HLA-B42, C4A*12, C4A*91, C4B*Q0, DR18) the frequency rose to 13%.

In studies of Mexican-Americans residing in Texas (10), an increased frequency of C4B*3 was noted as well as a decrease of the C4B null gene. The relative paucity of C4B*Q0 in Mexicans may reflect the Spanish influx as this population has a low frequency (6%) of C4B null genes.(11) In Southern Brazil, the occurrence of C4B*Q0 is also very low in the Kaingang but high in the Guarani Amerindians. In this same group an unusual C4A*3 variant was found in which there was a gene duplication at the C4A locus (C4A*0301, 0304) (Ref. 4).

In comparing allele frequencies between European Caucasians and Northern Indians, significant differences were noted by Kramer et al. (12) They found a decrease of C4A*Q0 (0.08) which was more like Hungarian Gypsies (0.02) than Hungarian Caucasians (0.17), thus supporting the theory that Hungarian Gypsies immigrated into Northern India. However, other studies of both Muslims and Hindus have found a high frequency of C4A*Q0, a finding that correlates well with the high frequency of SLE in Southern Asia (13).

2.3. Chido/Rodgers Blood Group Polymorphism

Human antibodies to epitopes found on C4 were first described in the 1960’s from patients which had been multiply transfused or from multiparous females. These sera eventually defined 2 Rodgers and 6 Chido blood group antigens as well as the WH antigen. The majority of the blood group frequency studies have been performed in Caucasians but there a few serological reports from other groups. We have found (unpublished data) a high incidence of C4 genes in African-Americans which demonstrate the “reversed antigenicity” phenomenon, i.e. these C4A alleles carry Ch1 in place of Rg1. Interestingly, the duplicated C4A*3 gene found by Weg-Remers et al. (4) had Rg1, Ch1 and Ch3. Almost 100% of Japanese (Table 2) have the Rg: 1,2 phenotype while the Chido phenotype is split among two major types, i.e. Ch: 1,2,3 (75%) and Ch: 1,2,3 (24%). Because of the association between C4A null genes and SLE (see below), it has been reported that the Rodgers negative phenotype occurs in ~10% of lupus patients vs. 2% of healthy controls (14).

2.4. C4 Genes and Disease

2.4.1. Systemic Lupus Erythematosus

Genes within the MHC are believed to be involved in the pathogenesis of SLE, however, it is difficult to determine which of these genes is most important because of strong linkage disequilibrium. The most common haplotype in Caucasian patients of Western European descent is HLA A1, B8, DR3, Bf*S, C2*C, C4A*Q0, C4B*1. The importance of C4A*Q0 in SLE has
been proven by studies of many ethnic groups including: French-Canadians, Greeks, African-Americans, Japanese, Chinese, etc (15-17). An increased frequency of C4A*Q0 was found in the absence of HLA-DR3 suggesting that the C4 genes themselves were the genetic risk factors. However, not all studies have come to this same conclusion. In a study of three ethnic groups in the USA, Reveille et al. (10) did not find C4A*Q0 to be a risk factor in either African- or Mexican Americans. The strongest associations between SLE and C4 appear to be the gene deletion (18,19) that removes most of the C4A gene, the CYP21A gene and part of the C4B gene as well as the HERV-K insert. If this insert is reverse transcribed it may act as an “intrinsic” DNA vaccine. Recently, high levels of another retrovirus, HRV-5, have been reported in both SLE and RA patients suggesting an important role for retroviruses in autoimmune diseases (20). However, C4A*Q0 is not always found as the genetic risk gene for SLE. In studies of Mexican-Americans, African-Americans, Chinese and Spaniards the C4B*Q0 gene was increased in frequency among SLE patients (11,17). An increased frequency of the deleted C4B genotype has been reported among SLE patients having diffuse proliferative glomerulonephritis again suggesting that a gene deletion may be the important risk factor.

2.4.2. Other autoimmune diseases

Although numerous studies have proven the importance of C4 genes in SLE, a similar association has been made with other autoimmune diseases. Both C4A*Q0 (21) and C4B*Q0 (22,23) have been reported as markers of disease susceptibility for systemic sclerosis (scleroderma) among Northern Europeans. Franciotti (24) found an increased frequency of C4A nulls in Italian patients having relapsing-remitting multiple sclerosis. This association was equally split between deleted C4A genes and non-expressed alleles suggesting that lack of C4A protein itself may be important in disease pathogenesis. In the Japanese, C4B*Q0 is associated with diffuse scleroderma as well as anti-topoisomerase antibody (25).

The Japanese have searched for MHC associations among patients diagnosed with Sjogren’s Syndrome. Moriuchu et al. (26) initially reported an association of C4A*Q0 with disease in 10 of 28 patients which appeared to be separate from the HLA-DRw53 association. Further studies indicated that the C4A null was not due to the typical 18 kb deletion but the exact mechanism remained unknown (27).

Other autoimmune diseases have also been studied in the Japanese in order to delineate the roles of both class II and III MHC genes. In a study of rheumatoid arthritis (RA), Takeuchi et al. (28) an increased frequency of C4A*Q0 and C4B*5 in patients. However, there was a strong association of the C4B*5 allele with the haplotype HLA-Bw54, Bw59, DR 4.1, DQw4. Several papers from the United Kingdom have reported an association between C4B null alleles and a complication of RA known as Felty’s Syndrome (29-31).

In addition to the associations of DR4 and C4 null alleles with RA, both loci are also reported as susceptibility genes for insulin-dependent diabetes mellitus (IDDM). Again the association appears to be with the gene deletion at least in Caucasians (32,33). Other autoimmune disease associated with deleted C4 genes include autoimmune hepatitis in French-Canadians and juvenile dermatomyositis in the U.K (34,35). Interestingly, the C4 genes were not primary genetic risk factors in adults (either white or black) having myositis (36). In Mexican-Americans the FC31 complotype has been associated seronegative spondylarthropathies (37).

3. SECOND COMPONENT OF COMPLEMENT (C2)

The structural gene for C2 lies between Bf and TNF in the class III MHC region. Although genetic polymorphism occurs, it is less common than C4 having ~10 variants. The C2C variant is the most common variant being found in 97% of most populations (13). Studies of C2 in various ethnic groups have shown that frequencies in Whites, Blacks, Asians, etc. are similar for the C2*C and C2*B alleles. However, in the Japanese several rare alleles have been identified. In Caucasians the C2*B allele is associated with HLA-Bw15, Cw3 as well as Bw40. In Japanese the rare C2*AT allele is associated with Bw62 and the C2*B with HLA-Bw61.

C2 deficiency is the most common complement deficiency occurring in ~1.5% of the population. Like total C4 deficiency, total C2 deficiency is strongly associated with SLE and in Caucasians is in linkage disequilibrium with HLA-A10 (A25), B18, DR2, BF*S. In a Brazilian study, C2 deficiency was found in 6.6% of SLE patients compared to 2.2% of controls (38). Two types of C2 deficiency have been described. The molecular mechanism for the first appears to be a 28 bp deletion in exon 6 causing the loss of a splice site, resulting in a frameshift and creation of a stop codon. Type II deficiency is caused by nonsense mutations that block secretion of C2. Low plasma C2 levels have been found in several autoimmune disease such as sclerosis and IDDM.

4. FACTOR B (BF)

Factor B allotyping by agarose gel electrophoresis has identified 3 common phenotypes (F, FS, and S) and ~15 rare variants. However, using iso-electric focusing or PCR-RFLP of codon 7 two subtypes (FA and FB) can be identified (39). The frequencies of BF alleles is highly variable; for instance, BF*S is very common in Europeans while BF*F is found in African derived populations (13). In general, the frequency of BF*F in the Indian subcontinent is intermediate between the two other races but there are considerable caste differences. The strong linkage disequilibrium with other MHC genes has hampered studies of BF and autoimmune disease.

5. TUMOR NECROSIS FACTOR (TNF)

Tumor necrosis factor is a cytokine produced as one of the first responses to infection but it may also play a role in the pathogenesis of many autoimmune disorders. There appears to be a genetic component to the amount of
TNF that is produced and this often correlates with the HLA-DR genotype. The TNF gene exhibits a promoter polymorphism at –308 and –238. At –308 the two alleles are TNF1 (G) and TNF2 (A). There is strong linkage disequilibrium between the TNF*30K allele and HLA-DR 3 and individuals with this haplotype often have higher levels of TNF. In studies of Taiwanese (40) and USA Blacks (41) the TNF2 allele was associated with an increased risk for SLE, however, in South African Blacks there was a reduction of this allele (42).

In close proximity to the TNF (TNFα) gene is the TNFβ locus also referred to as lymphotoxin-α (LTo) as well as several microsatellite markers. These too have been investigated in relation to HLA and disease expression in SLE (43). Among Caucasians the data suggest that TNF microsatellites are not independent markers for SLE (44,45). Although the TNFa2, b2 type is increased in SLE it most often is part of the HLA-B8, SC01, DR3 haplotype (45,46). Kim et al. (47) have investigated TNFβ genes and found that homozygosity for TNFB*2 was strongly associated with lupus nephritis among Koreans.

The role of TNF in rheumatoid arthritis has been well documented and there are now new therapeutic agents available for treatment based on these studies. However, the role of TNF genes is difficult to interpret separate from their linked HLA haplotypes and may in fact show some interaction. In studies of Caucasians the HLA-DRB1*0401 was associated with the microsatellite marker TNFa6 and c1 (48). Recently, Mu et al. found that the most severe RA occurred among patients having TNFα11 and DR alleles bearing the “shared epitope” (48). A Japanese report has implicated genes between TNF and HLA-B in Bechet’s disease (49).

6. CONCLUSIONS

More and more genes are being identified within the Class III region of the MHC. Many of these genes have been and continue to be associated with various autoimmune diseases. However, clearly there are many populations and autoimmune diseases to be studied before we fully understand the role of these genes in autoimmunity.

7. ACKNOWLEDGMENTS

This work was supported by a Hulda Irene Duggins Arthritis Investigator Award from the Arthritis Foundation and a NIH grant (R01 AI 42367).

8. REFERENCES


13. Ad'Hiah A. H & S. S. Papiha: Complement components C2, C3, and C4 (C4A and C4B) and BF


**Key Words:** C4, C2, TNF, Complement, MHC, Review

**Send correspondence to:** Joann M. Moulds Ph. D., Department of Microbiology and Immunology, MCP Hahnemann University School of Medicine, 2900 Queen Lane, Philadelphia, PA 19129-1096, Tel: 215-991-8379, Fax: 215-848-2271, E-mail: jmm56@drexel.edu