THE ASSOCIATION OF MHC GENES WITH AUTISM

Anthony R. Torres¹, Alma Maciulis², Dennis Odell³

¹Center for Persons with Disabilities and Biology Dept., Utah State University, Logan, Utah, ²Center for Persons with Disabilities, Utah State University, Logan, Utah, ³Center for Persons with Disabilities and Biology Dept., Utah State University, Logan, Utah

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Human Complement C4 in Autism
4. Conclusions
5. Acknowledgments
6. References

1. ABSTRACT

Several immune abnormalities have been noted in autistic subjects. These associations have been extended to the Major Histocompatibility Complex (MHC), a section of DNA remarkable for the number of encoded proteins with immunological functions. The strongest MHC association identified thus far is for the null allele of C4B in the class III region. The complex allelic composition of C4 as determined by immunoelectrophoresis is discussed. Low levels of C4 resulting from the null allele may be important in disease pathogenesis especially since C4 has been identified in developing brain neurons. The DNA region just telomeric to C4 has several genes including tumor necrosis factor which encode proteins with immunological functions. These proteins may act in concert with C4 in disease contribution and the genes should be more closely examined.

2. INTRODUCTION

Autism is a chronic developmental disorder characterized by abnormalities in the ability of the brain to collect and integrate information. Individuals with autism show grossly abnormal communicative and social skills and the disease occurs four-five times more frequently in boys than in girls. The cause or causes of autism remains unknown and current research is aimed at determining susceptibility genes since there is a very high concordance rate in monozygotic twins and siblings (1-3). One line of active autism research stems from immune system abnormalities commonly found in autistic subjects.

The human immune system can be divided into two arms: the innate (non-specific) and the adaptive (specific), both of which have humoral and cellular components (4) (Table 1). The innate immune response is constitutively expressed and germ-line encoded and therefore does not adapt to antigen. It is a primitive system that offers the first line of defense and effectively destroys most microorganisms within hours through complement activation. The innate system is also very important in antigen processing and antigen presentation. This is accomplished by antigen binding to HLA-proteins on antigen presenting cells as macrophages and B-lymphocytes. HLA-proteins thus serve as transducers in activating the adaptive immune system by binding to specialized T-cell receptors and B-cell receptors on the surfaces of T and B-lymphocytes respectively.

Adaptive changes result in specific antigen recognition through somatic mutations in T and B lymphocyte genes. Surface receptor and immunoglobulin genes rearrange to encode binding diversity. The adaptive arm of the immune system is thus in a state of flux to maximize antigen recognition. Both antibody producing B-cells’ and effector T-cells’ responses are influenced by the adaptive immune system and although exquisite antigen recognition is achieved, the process is relatively slow. Several days are required for clonal expansion of primed lymphocytes into cytotoxic effector-cells or antibody secreting-cells. Weeks can be required for a primary antigen challenge. This manuscript briefly discusses basic immunological concepts and reviews research on MHC-autism associations, in particular the C4B null allele in the class III region.

3. HUMAN COMPLEMENT C4 IN AUTISM

Stubbs observed over 20 years ago that children with autism often fail to produce antibodies following rubella vaccination (5). Others have reported higher antibody titers to brain proteins in subjects with autism (6). Observations such as these were important in bringing attention to a possible connection of immune system dysfunction with autism. It has also been shown that autistic subjects have decreased responsiveness to T-cell mitogens (7, 8), reduced numbers of helper T-cells and suppressor-inducer T-cells (9) as well as decreased natural killer cell activity (10). Plioplys et al. (11) and Warren et al. (12) have both reported that a significant number of autistic subjects have increased numbers of DR+ T-cells
HLA genes in autism

Table 1. Components of Adaptive and Innate Immune Systems

<table>
<thead>
<tr>
<th>Component</th>
<th>Innate</th>
<th>Adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular</td>
<td>Monocytes/macrophages</td>
<td>Alpha/beta T-lymphocytes</td>
</tr>
<tr>
<td></td>
<td>Gamma/delta T-lymphocytes</td>
<td>B-lymphocytes</td>
</tr>
<tr>
<td></td>
<td>NK cells, granulocytes</td>
<td>Complement, acute phase-globulins</td>
</tr>
<tr>
<td>Humoral</td>
<td>Complement</td>
<td>Antibodies-all classes</td>
</tr>
<tr>
<td></td>
<td>Mannose binding lectin</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Genomic organization of the human MHC, drawn to scale. Only genes of particular interest for the immune system are indicated. (With permission).

suggesting T-cell activation. This laboratory has also shown that certain extended haplotypes (13) and the null allele of the fourth component of complement C4B are increased in autistic subjects (14). An observation that subjects with autism have significantly higher urinary levels of neopterin and biopoterin provides direct biochemical evidence that monocytes are activated and producing lymphokines (15). The above evidence plus the genetic inheritance data suggests that this disorder has a complex etiology resulting from a combination of genetic, microbiological, immunological, and possibly environmental factors.

There is also evidence that autism shares several features of established autoimmune diseases. The ability to discriminate self from non-self in a highly specific manner is critical for survival. This phenomenon of self-tolerance can break down resulting in the immune system’s response to self-antigens. Autoimmunity, originally called horror autotoxicus by Paul Ehrlich, can take many forms and there is extensive research being done into the immunological aspects of the class III region around C4 since many of these genes are expressed in immune cells (25-27). Examination of the entire MHC indicates that polymorphic frozen blocks of DNA were developed by imperfect sequential duplication (28). The mechanisms for these duplications have not been determined. However, it appears that human endogenous retroviruses (HERV) and HERV fragments are involved (28, 29). The MHC contains about 10 times the number of retroelements as are detected in other areas of the human genome and it has been postulated that these elements provide sites for recombination, translocation and perhaps even control of gene expression (28). The long version of the C4 gene, for example, results from the integration of an HERV-K into intron 9 (31). Associations of these retroelements may well prove to be of major importance in autoimmune diseases (30-32). HLA class II involvement is the hallmark of typical autoimmune disease (16) as demonstrated by Gross et al (33) when they identified LFA-1 as the autoantigen in autoimmune Lyme arthritis. Immune reactivity to outer surface protein A (OspA) of Borrelia burgdorferi, the causative agent of Lyme disease, can lead to autoimmunity in individuals with the HLA class II DRB1*0401 allele. They elegantly showed that the OspA protein and the cell surface adhesion protein LFA-1 share highly related peptide epitopes both of which bind to DRB1*0401. Such molecular mimicry occurs when immunity against a

The MHC contains about 4 million base pairs of DNA and is divided into three regions (class I, class II, and class III) (Figure 1). The HLA region is remarkable in at least 3 respects:

(a) Allelic polymorphisms;
(b) Disease associations;
(c) Gene density, sequence duplications, insertions and deletions.

It is the most polymorphic region of DNA yet discovered with hundreds of alleles that encode proteins with various immunological functions (22). The best characterized genes in the class I region code for cell surface proteins which bind and present peptides to CD8+ T-cells. The entire HLA has been sequenced (23) and a research group in Japan has estimated that there are as many as 118 genes in this two million base pair region plus a complicated array of retroelements and microsatellites (24). This means that the gene density is comparable to other HLA regions and it is likely that investigators have only begun to uncover the important information encoded in this region. Genes (DR, DQ and DP) in the class II region have been thoroughly characterized for autoimmune-associations. The proteins encoded by these genes process, bind and present peptides to CD4+ T-cells.

The gene-rich class III region contains four complement genes (C4A, C4B, BF, C2), genes for tumor necrosis factor (TNF-alpha and TNF-beta), and many others (25). There is extensive research being done into the immunological aspects of the class III region around C4 since many of these genes are expressed in immune cells (25-27).
microbial antigen spreads to a related autoantigen. It is thought that such mimicry exists in many autoimmune diseases, however, in most cases the specific antigens are not known. Statistical calculations showing a significant HLA class II disease association must suffice for many autoimmune diseases. This laboratory has published several manuscripts showing HLA associations with autism. The first manuscript showed an increase in certain extended or ancestral haplotypes (13). Extended haplotypes are conserved regions of DNA, which demonstrate linkage disequilibrium by segregating together more than would be expected by their genetic distances. An extended haplotype as defined in our laboratory includes class I HLA-B, HLA-DR and DQ from class II, and the class III complement genes BF, C2, C4A, and C4B. We have seen 11 different extended haplotypes which include the above class I, class II, and class III genes in our subject population. Two later manuscripts indicated that the third hypervariable region of certain HLA-DRB1 alleles had stronger autism-associations than the specific extended haplotypes (34, 35). These publications were important in showing that MHC genes/proteins are somehow involved or associated with autism and should be studied to perhaps better delineate autism pathogenesis and autoimmune characteristics. It is interesting to note that after years of research into HLA associations, the often-overlooked class III complement components have the strongest autoimmune-associations (36).

Increased incidences of C4 null alleles have been associated with systemic lupus erythematosus (SLE), insulin dependent diabetes mellitus, membranoproliferative glomerulonephritis, bacterial meningitis, scleroderma (37, 38), and autism by this laboratory (14). Non expressed, or deleted C4 genes, are commonly given the notation (Q0) for quantity zero. About 15% of Caucasians have C4AQ0 and 20% have C4BQ0 (37). A publication from the Warren laboratory in 1991 (14) noted that 58% of autistic subjects (n=19) and their mothers had C4B null alleles compared to 27% in controls. Fathers had normal frequencies of the C4B null allele and all family members had normal frequencies of the C4A null allele. There were no abnormalities in the adjoining BF and C2 alleles. Plasma values for C4B but not C4A were significantly decreased (P=0.01) for the autistic subjects when compared to family members and normal controls (Figure 2). Current data from an additional 39 subjects show very similar results for the C4B null allele, however, current subjects also show an increased frequency in the C4A null allele (unpublished data). At this time, the C4B null allele is the strongest autism-association thus far discovered in the MHC. The complement system is composed of a complex series of interacting blood proteins that serve as mediators of the innate immune response for protection of the host against pathogens and to clear immune complexes. These proteins circulate as inactive precursors and are activated in a very precise order of biochemical reactions. Four of the 20 complement proteins (C2, BF, C4A, & C4B) are encoded in the class III region of HLA and are inherited as compact units called complementotypes. C4A is more reactive for targets with free amino groups as antibody-antigen complexes and C4B is more hemolytic and reactive with hydroxyl groups. Four amino acids at positions 1101-1106 code for these functional differences. Site-directed mutagenesis experiments have determined that the amino acid at position 1106 (aspartic acid in C4A and histidine in C4B) is the critical residue of functional differences between these proteins (39). The fourth component of complement (C4A and C4B) is a major factor in the complement cascade and is an acute phase globulin produced in the liver, macrophages and various brain cells including neurons (40,41).

The tandemly duplicated C4 genes encode single chain protein precursors (pro-C4) which result in three-chain native proteins (185 kDa). The C4 proteins represent a very complex family with over 40 alleles (42). It is difficult to identify nonexpressed C4B alleles by genetic methods (42) and typing is routinely done by non-denaturing protein immunoelectrophoresis (Figure 3), verified by a C4B ELISA developed in our laboratory. However, the native composition is even more complex as plasma samples are typically treated with neuraminidase and carboxypeptidase B, two enzymes that remove

Figure 2. The C4A (top) and C4B (bottom) protein concentrations in the plasmas of autistic subjects, siblings, mothers, fathers, and normal controls. The C4B values in autistic subjects were significantly (P=0.01) decreased compared with those of normal controls. (With permission).
null alleles of C4 have been associated with autoimmune diseases for over 15 yr (37). Even today little is known about C4-autoimmunity relationships. However, many immunologists are now looking at the complete immune system with the realization that collaboration between the innate and adaptive systems is more important than previously appreciated (4). C4 is a very important protein in the innate system and it is possible that even partial deficiencies during critical periods could leave the host with less than optimal defenses. Johnson et al. (40) has made an interesting observation in showing a inverse relationship of C4 mRNA to cell proliferation in the developing rat brain. They postulated that C4 may have a function in the brain outside of the complement cascade. It should be mentioned that only a small number of subjects with C4B null alleles have autism. It would thus appear that other etiological agents are required. Perhaps subjects with autism have another uncharacterized deficiency in the innate system or a partial C4 deficiency in certain children or mothers which allows unknown infectious agents to persist during pregnancy resulting in an abnormal immune stimulation and autism. It should be remembered that C4B reacts with carbohydrate hydroxyl groups seen on the surfaces of many microorganisms and that an impaired clearance of certain microorganisms could result in an immunological challenge leading to autism. Our laboratory is currently examining mannose binding lectin (MBL) genes to determine if autistic subjects also show deficiencies in another protein which binds to pathogens through surface carbohydrate groups.

Another possible explanation for the association of C4BQ0 with autism is simply due to its location in the MHC. The null allele of C4B is in strong linkage disequilibrium with other MHC genes and perhaps one of these genes alone or in combination with C4BQ0 is related to the development of autism. Gruen and Weissman (22) recently proposed that a MHC class IV region be considered due to the high concentration of genes at the telomeric end of class III which have various roles in inflammation or infection. Tumor necrosis factor (TNF) is the most studied gene in this region as it encodes an important cytokine that is involved in inflammation, microbial infections, tumor cachexia, and normal immune responses (22). The gene for leukocyte-specific transcript (LST-1), also known as B144, is located just centromeric to the TNF gene cluster and is expressed exclusively in macrophages, monocytes, and some T-cells (22). The putative protein has not been identified but must be

null alleles of C4 have been associated with autoimmune diseases for over 15 yr (37). Even today little is known about C4-autoimmunity relationships. However, many immunologists are now looking at the complete immune system with the realization that collaboration between the innate and adaptive systems is more important than previously appreciated (4). C4 is a very important protein in the innate system and it is possible that even partial deficiencies during critical periods could leave the host with less than optimal defenses. Johnson et al. (40) has made an interesting observation in showing a inverse relationship of C4 mRNA to cell proliferation in the developing rat brain. They postulated that C4 may have a function in the brain outside of the complement cascade. It should be mentioned that only a small number of subjects with C4B null alleles have autism. It would thus appear that other etiological agents are required. Perhaps subjects with autism have another uncharacterized deficiency in the innate system or a partial C4 deficiency in certain children or mothers which allows unknown infectious agents to persist during pregnancy resulting in an abnormal immune stimulation and autism. It should be remembered that C4B reacts with carbohydrate hydroxyl groups seen on the surfaces of many microorganisms and that an impaired clearance of certain microorganisms could result in an immunological challenge leading to autism. Our laboratory is currently examining mannose binding lectin (MBL) genes to determine if autistic subjects also show deficiencies in another protein which binds to pathogens through surface carbohydrate groups.
HLA genes in autism

Figure 5. Relationship between C4A and C4B plasma levels and C4 function. C4B values were positively correlated with C4 function at P <0.0001.

Figure 6. A. Relationship between DR+T cells in the autistic patients and plasma levels of the C4B protein. The inverse correlation was significant at p = 0.003; R = 0.627. (With permission) B. Relationship between DR+ T cells in the autistic patients and plasma levels of the C4A protein. (With permission.)

expressed at high levels as there are rather large amounts of mRNA under conditions of cell activation. Another gene in this area, AIF-1, is also expressed in subsets of monocytes and lymphocytes, but its function remains unknown (22). I kB-like is a very interesting protein encoded by a gene just telomeric to the TNF cluster (22). This protein has motifs similar to those seen in I-kB which is a negative regulator of NFkB, an important universal transcriptional control protein. It has been postulated that the I kB-like protein functions as a regulator of inflammatory responses in immune cells. It is also interesting to note that the expression of many of these genes is controlled by g-interferon, an important immune cytokine.

TNF genes are at the telomeric end of the class III region and the TNF protein is produced mainly by cells of the macrophage lineage and to a lesser degree by certain T-cell subsets. The coding regions of the TNF genes do not reveal a high degree of polymorphism; RFLPs show a biallelic system (44,45). However, microsatellite DNA sequences flanking the TNF genes are highly polymorphic and, more importantly, specific microsatellite polymorphisms have been shown to be associated with autoimmune diseases (46-55). Certain associations occur as part of extended haplotypes and others seem to be independently associated with disease (49,52, 56). It is interesting that some extended haplotypes can be differentiated into two separate haplotypes according to the TNF microsatellites even though the haplotypes are serologically identical (44). A Swedish study involving SLE patients showed an increase of TNFα2b3c1 and a decrease in TNFα4. They concluded that this TNF microsatellite haplotype is increased due to linkage disequilibrium with extended haplotypes previously associated with SLE (46). An SLE study in Greeks also reported linkage disequilibrium between DRB1*1501 and TNFα111 and DR3 with TNFα2b3d2. More importantly, specific microsatellites were found to be significantly associated with autoantibody production in SLE (50). A multiple sclerosis study concluded that an increased frequency of TNFα11b3 was likely due to linkage disequilibrium with DR2 (51). On the other hand, a study of Crohn’s disease patients showed an increase of TNFα2b1c2d4e1 and concluded that this TNF haplotype incurred the greatest risk along with DR1/DQ5 (53). TNFα2 was reported to be independently associated with celiac disease in an Irish population (55) while a Finnish study showed an increase of TNFα6c1 and a decrease of TNFα11c2 in HLA-B27 RA subjects (49). It is well to note that ethnicity is extremely important when studying disease associations as different TNF microsatellite associations are reported in different ethnic groups (56, 57). Sex of the patients can also be important as it has been reported that female RA patients have an increase of TNFα6b5c1 while males show an increase of TNFα2b1c2 (56). Comings (59) has postulated that microsatellite DNA is important in gene transcription by allowing the formation of z-DNA.

Our laboratory has evaluated the TNF microsatellite composition in 55 autistic subjects and 39 controls as well as TNF-alpha plasma levels. Briefly, the results in this study showed that autistic subjects have
lower levels of plasma TNF-alpha and TNF microsatellites that correlate with lower TNF-alpha levels than controls (unpublished data). This is in contradiction to Singh (60) who reported normal plasma TNF-alpha in subjects with autism.

4. CONCLUSIONS

Various research groups have reported numerous immune-autism associations over the last 20 years. This laboratory has published disease-associations with MHC class I, class II and class III genes/proteins. At this time, the C4B null allele in the class III region has the strongest disease association. It is unknown if the C4B locus or another segment of DNA in linkage disequilibrium with this locus is involved in the pathogenesis of autism. There are several class III genes on the telomeric side of C4 including tumor necrosis factor loci that may have various roles in inflammation and infection and may therefore be involved in disease pathogenesis. This gene rich region between the TNF locus and the HLA-B locus should be more closely studied for autism association. An additional area, which should be studied, is the function of C4 in the developing brain.

5. ACKNOWLEDGMENTS

This work was supported by Center for Persons with Disabilities, Utah State University, National Institute of Mental Health, Bethesda, MD, Grant MH42119, and National Institute of Child and Development, Bethesda, MD, Grant HD35476. We are indebted to Peggy L. Dahle for manuscript preparation.

6. REFERENCES


HLA genes in autism


HLA genes in autism


In memory of Dr. Reed P. Warren, deceased 1998

Key Words: Autism, HLA, Complement, TNF, Review

Send correspondence to: Anthony R. Torres, M.D., Utah State University, CPD, 6895 Old Main Hill, Logan, Utah 84322-6895, Tel: 435-797-2750, Fax: 435-797-4054, E-mail: rtorres@cpd2.usu.edu