Analysis of genotypes that alter protein structure as a possible cause of Multiple Sclerosis
Draft: October 26, 2006
Accelerated Cure Project, Inc.

Summary
Genetic variants that change the shape of a protein and thus affect its function may be involved in the development of MS. Many mutations known or presumed to alter protein function compared with wildtype have been investigated as potential causes of MS; so far none of these have emerged as strong candidates. Stronger evidence has been produced for immune system genes characterized by extensive functional variation, most notably the MHC class II genes. It has been hypothesized that variants of these genes influence MS risk by producing conformations with increased affinity to antigens such as myelin or microbial peptides. Further research into these immunological structures may provide valuable insights into the development of MS.

Hypothesis
Genetic defects or variants that alter the structure and function of a protein increase susceptibility to MS.

Experimental tests of the hypothesis
The structure of a protein is determined largely by its amino acid sequence, which in turn derives from the sequence of base pairs in its gene. Alterations to a gene’s sequence may therefore result in structural alterations in the protein it encodes, with potentially significant effects on that protein’s function. For instance, structural changes may increase a protein’s effectiveness or activity, decrease its effectiveness or activity, or even cause the protein to interfere with other functions.

It is possible that the risk of MS is influenced by one or more structural alterations to proteins that result from gene-level changes. Two types of studies have been performed to investigate this possibility. In one type of study, known functionally relevant genetic polymorphisms are studied for association or linkage with MS (as in the many studies of apolipoprotein E alleles). In another type of study, a gene thought to be a candidate for MS involvement is sequenced to look for previously undiscovered mutations that may be related to MS (as in a recent analysis of myelin-associated glycoprotein).

This document summarizes the outcomes of the searches for functionally relevant alleles that influence MS susceptibility. The first section below addresses the polymorphisms in genes known or suspected to alter structure that have been analyzed.
for a potential connection with MS. The second section deals with a set of immune system genes characterized by extensive functional polymorphism. The degree of polymorphism seen in these genes is normal and indeed helpful or even necessary as the large number of variants help to ensure effective responses against a wide variety of pathogens. However, particular polymorphisms may be predispositional to diseases such as MS.

**Genes with structure-altering mutations:** Genes containing structural alterations that have been investigated for association with MS are listed in Table 1 below. In most cases, the effect of the mutation on structure and/or function is known or suspected. Also included are a few mutations that may cause structural changes but for which the structural or functional consequences are not yet understood.

Table 1. Structure-modifying alleles studied for linkage or association with MS

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Allele studied for linkage/association with MS</th>
<th>Results of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABO blood group</td>
<td>A and B alleles that express glycosyltransferases converting H antigen into A or B antigen; the O allele is identical to the A allele except for a frame shift deletion which results in a protein incapable of antigen modification</td>
<td>Six studies have been performed, of which two found allele distribution differences between MS subjects and controls and four did not</td>
</tr>
</tbody>
</table>

Alpha-1-antitrypsin

Variety of alleles such as mutations that form insoluble intracellular inclusions

Of the two studies performed, one found an increased frequency of the normal allele M3 in MS subjects while the other found no association

Apolipoprotein E (APOE)

Amino acid substitutions that create several variants; the major isoforms are E2, E3, and E4. E2 binds less well to receptors on liver and peripheral cells; E4 binds more rapidly to beta amyloid than E3; E4 forms a stable folding intermediate more readily; microtubule associated protein tau also binds differently with different variants.

23 linkage or association studies have been performed for this gene, the overwhelming majority of which found no relationship between APOE genotype and MS risk

Apolipoprotein E (APOE)

Amino acid alterations that create several variants; the major isoforms are E2, E3, and E4. E2 binds less well to receptors on liver and peripheral cells; E4 binds more rapidly to beta amyloid than E3; E4 forms a stable folding intermediate more readily; microtubule associated protein tau also binds differently with different variants.

23 linkage or association studies have been performed for this gene, the overwhelming majority of which found no relationship between APOE genotype and MS risk

Apoptosis genes bax, bcl-2, and p53

Amino acid-altering SNPs in bax and p53; a SNP at codon 43 of bcl-2 previously associated with autoimmunity

None of these alleles were associated with MS in one study

Brain-derived neurotrophic factor (BDNF)

Val66Met polymorphism that alters intracellular transport and secretion of BDNF

No association has been found with MS for this allele

BTNL2

SNP rs2076530 that truncates the protein and disrupts its membrane localization

One study did not find an effect of this SNP on MS risk independent of DRB1*15

Complement component

Two main variants, C3S and

Three studies provide mixed
<table>
<thead>
<tr>
<th>Gene</th>
<th>Polymorphism</th>
<th>Description</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3f</td>
<td>C3f, that differ by a single amino acid substitution (effect on structure or function unknown)</td>
<td></td>
<td>results&lt;sup&gt;37-39&lt;/sup&gt;</td>
</tr>
<tr>
<td>COMT</td>
<td>G158A mutation that results in decreased activity</td>
<td>No statistically significant association with MS has been found&lt;sup&gt;40&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>FcgammaR IIA and IIB</td>
<td>Arg131His mutation in IIA that affects ability to ligate IgG2; V158 allotype in IIB that results in more efficient binding; NA1 and NA2 alleles in IIB that differ in their ability to bind opsonized particles</td>
<td>Two studies have been performed, which were unable to find significant allele frequency differences between MS subjects and controls&lt;sup&gt;41, 42&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Glucocorticoid receptor</td>
<td>Three polymorphisms that may influence glucocorticoid sensitivity (N363S, ER22/23EK, and the Bcl)</td>
<td>Only one study has been conducted to date, which did not show an association for any of these alleles with MS&lt;sup&gt;43&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Glutathione S-transferase pi 1 (GSTP1)</td>
<td>Substitution of Val for Ile at position 105 of GSTP1 that affects catalytic efficiency</td>
<td>Only one study has been conducted to date, which showed no significant differences for this polymorphism in MS subjects compared with controls&lt;sup&gt;44-46&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>G-protein beta3 (GNB3)</td>
<td>C825T mutation associated with enhanced signaling</td>
<td>The only study performed did not find this mutation to be a risk factor for MS&lt;sup&gt;45&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Hemochromatosis (HFE)</td>
<td>Variants H63D and C282Y that are thought to affect interaction with the transferrin receptor</td>
<td>No independent statistically significant associations were found for either allele with MS in three studies&lt;sup&gt;46-48&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>ICAM-1</td>
<td>G241R and K469E polymorphisms creating amino acid changes which may lead to altered interaction with the ligand</td>
<td>Eight studies have been conducted, producing mixed results&lt;sup&gt;49-56&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Interleukin 4 receptor</td>
<td>R551 variant which may decrease receptor function; R576 which enhances signaling</td>
<td>One study found an association for R551 with the primary progressive form of MS&lt;sup&gt;57&lt;/sup&gt;; another study found an association with this polymorphism in Basque but not two other populations&lt;sup&gt;58&lt;/sup&gt;; Q576R was found to be associated with MS in one study&lt;sup&gt;59&lt;/sup&gt; but not another&lt;sup&gt;56&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Interferon gene cluster</td>
<td>Premature stop codons in IFNA10 and IFNA17 and an amino acid substitution in IFNA17</td>
<td>Only one study has performed; it showed an association between increased risk of MS and a nonfunctional IFNA17 gene&lt;sup&gt;60&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Leukemia inhibitory factor (LIF)</td>
<td>SNP at position 2680 that is thought to reduce biological activity of LIF</td>
<td>The only study that has been conducted for this SNP in MS did not show an association&lt;sup&gt;61&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Methylenetetrahydrofolate reductase</td>
<td>C677T mutation that reduces enzyme activity</td>
<td>No statistically significant association with MS has been found&lt;sup&gt;40&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Myelin-associated glycoprotein (MAG)</td>
<td>Nonconservative amino acid substitution (Arg537Cys) in</td>
<td>Genotyping revealed this SNP in one MS subject but subsequent</td>
<td></td>
</tr>
<tr>
<td>Gene/Phenotype</td>
<td>Substitution/Polymorphism</td>
<td>Description</td>
<td>Results/Findings</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Myelin oligodendrocyte glycoprotein (MOG)</td>
<td>Exon 4 amino acid substitutions Val142Leu and Val145Ile (effect on structure or function unknown)</td>
<td>One study found Val142Leu underrepresented in MS subjects, but two others found no association.</td>
<td></td>
</tr>
<tr>
<td>Platelet-activating factor acetylhydrolase</td>
<td>Missense mutation G994T that causes loss of catalytic activity</td>
<td>One study found no relation between this mutation and susceptibility to MS.</td>
<td></td>
</tr>
<tr>
<td>Platelet-activating factor receptor</td>
<td>Point mutation A224D which reduces PAF-induced intracellular signaling</td>
<td>Genotype and phenotype associations for the D allele were found in the only MS study that has examined this gene.</td>
<td></td>
</tr>
<tr>
<td>Protein-tyrosine phosphatase receptor-type C</td>
<td>Novel point mutation (C59A) that alters the structure of the extracellular domain, possibly interfering with binding</td>
<td>Only one study has been performed, which reported the presence of this mutation in an MS multiplex family but found it to be rare in the MS population overall.</td>
<td></td>
</tr>
<tr>
<td>PTPN22</td>
<td>Missense allele (R620W) that may affect binding to Csk</td>
<td>Four studies did not find an association between this allele and MS.</td>
<td></td>
</tr>
<tr>
<td>Selectin E</td>
<td>Point mutation (A561C) that affects ligand binding specificity</td>
<td>One study did not find a connection between this allele and MS risk, but its presence may protect against conversion to SPMS.</td>
<td></td>
</tr>
<tr>
<td>Selectin P ligand</td>
<td>VNTR polymorphism that may affect binding capability</td>
<td>One study found the C allele to be overrepresented in cases with primary progressive MS.</td>
<td></td>
</tr>
<tr>
<td>Spinocerebellar ataxia (SCA) type 1, 2, 3, 6, and 7 genes</td>
<td>Expanded CAG trinucleotide repeats</td>
<td>Four studies have been performed; none showed expanded repeats to be associated with MS risk.</td>
<td></td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>Asp299Gly and Thr399Ile mutations which affect the extracellular domain of the receptor</td>
<td>Two studies have been performed, which found no difference in the frequencies of these mutations between MS subjects and controls.</td>
<td></td>
</tr>
<tr>
<td>Vitamin D binding protein (or Group-specific component, Gc)</td>
<td>Substitutions at codons 416 and 420 in exon 11 that result in three main phenotypes (Gc1f, Gc1s and Gc2)</td>
<td>Four studies found no connection between VDBP genotype and MS, although one found suggestive evidence for a relationship.</td>
<td></td>
</tr>
<tr>
<td>Vitamin D receptor (VDR)</td>
<td>Polymorphisms FokI (exon 2) and TaqI (exon 9) that may have functional significance</td>
<td>Mixed results are given by the four studies that have examined these alleles.</td>
<td></td>
</tr>
</tbody>
</table>

**Immune system genes characterized by substantial diversity:** Immunoglobulins (Igs), T-cell receptors (TCRs) and class I and II major histocompatibility complex (MHC) molecules are all immunological structures whose shape helps determine the nature and effectiveness of an immune response against an antigen. The affinity with which each of these proteins binds to antigens is determined largely by protein structure. To cope with the wide and changing set of antigens that exists in our environment, each of these types of proteins is characterized by a great deal of genetic diversity which results in a
broad repertoire of protein structures. This diversity helps ensure a robust immune response to new antigens, both for humans as individuals and for the species as a whole.

Genetic diversity in immunoglobulin and T-cell receptor genes is largely provided by an on-going process of spontaneous somatic gene rearrangement and/or mutation that produces a broad and dynamic arsenal of molecules inside the body. The MHC genes do not undergo this particular process – for these genes, diversity comes from the extremely polymorphic nature of the MHC region.

Each of these molecules (Igs, TCRs, and MHC molecules) has been investigated as candidates for inducing susceptibility to MS. (See the file phase2-genetic-studies.xls for details of these studies as well as the other studies referred to in this document.)

- **Immunoglobulins (Igs or antibodies):** Immunoglobulins, the antigen-binding proteins produced by B cells, are composed of “heavy” chains (proteins) and “light” chains, each of which are encoded by different genes. Germline polymorphisms in the genes encoding the heavy chain (the Gm system on chromosome 14) and kappa light chain (the Km system on chromosome 2) have been investigated numerous times but with very mixed results. (For example, see Wood, et al.86 and Raknes, et al.87.) At this time there is no one allele or allotype that is widely recognized to increase susceptibility to MS. In addition to germline variation, the Ig gene repertoire in each individual is continually changing through a dynamic process of somatic genetic rearrangements and mutations. While this process is normal and nonpathogenic, and helps us respond to a wide and changing assortment of antigens, it may nevertheless produce Ig structures that somehow predispose a person to MS, such as antibodies with affinities to certain antigens that contribute to the development of MS. However, there is no conclusive evidence yet that this process is involved in the etiology of MS. (See our document “Analysis of spontaneous genetic mutations as a possible cause of Multiple Sclerosis” for a more extensive review of the investigations into how this process may be involved in MS.)

- **T-cell receptors (TCRs):** Like immunoglobulins, the receptors on T cells that bind to antigen/MHC compounds are also composed of multiple proteins (an alpha and beta chain or a gamma and delta chain), each of which is encoded by a separate gene located on chromosome 7 (beta and gamma) or 14 (alpha and delta). These genes also undergo a process of somatic rearrangement that ensures a robust response to pathogenic threats. While some evidence has been produced suggesting that the nature of a person’s TCR repertoire helps determine MS risk, more evidence is needed before somatic rearrangement can be conclusively associated with MS. (See “Analysis of spontaneous genetic mutations as a possible cause of Multiple Sclerosis” for more details.) As for germline polymorphisms in the TCR genes, the genes for each chain have been studied for an association with MS. So far the results are inconclusive. The gamma and delta chain genes have been studied only twice (Briant, et al. and Droogan, et al.88,89), and the nine studies conducted on the alpha chain gene to date have produced mostly negative results (see for example Droogan, et al.89). Stronger evidence has been produced for the beta chain, in particular the V8 and V11 segments for which multiple studies have found an association with MS,
although not all studies have replicated these results. (See Martinez-Naves, et al, and Vandevyver, et al.) Some studies found increased association when results were stratified for the presence of the MHC class II DR2 haplotype (e.g., Hockertz, et al), raising the possibility that particular combinations of TCR and MHC genes may increase MS risk.

- **MHC molecules:** MHC class I and II molecules, which are located on the surface of cells for the purpose of presenting antigens to T cells, are also composed of separate chains that assemble into a compound. The genes for these chains are found in the major histocompatibility complex, a set of genes found on chromosome 6. Whereas genetic diversity for Igs and TCRs is primarily created by gene rearrangements and somatic mutations, diversity for MHC molecules is provided by a high degree of germline genetic polymorphism. Indeed, a tally published in 2006 by the Anthony Nolan Research Institute revealed that a total of 1585 class I alleles (HLA-A, B, and C) and 808 class II alleles (HLA-DP, DQ, and DR) have been identified and officially named. The MHC region has been extensively analyzed for possible connections with MS, with over 190 studies conducted to date (for instance, see Barcellos, et al). While several alleles and haplotypes have been studied over the years for an association with MS, the strongest evidence produced so far concerns the class II genetic haplotype DR2. (See “Analysis of genetic mutations or alleles on the autosomal chromosomes as possible causes of Multiple Sclerosis” for more information on the association between MS and DR2.) The MHC region is marked by extensive linkage disequilibrium, so it is possible that the alleles and haplotypes associated with MS to date are not themselves causal but rather linked with causal alleles. However, the hypothesis that class II molecules influence MS risk is certainly credible given the involvement of T cells in MS pathology, and scientists are now working to better understand the structure of these molecules to determine how their affinity to various antigens (such as myelin and viral peptides) might play a role in triggering MS.

**Conclusions**

Two types of structure-modifying polymorphisms have been studied for a possible link with MS: “normal” polymorphisms in immune system genes that enable the immune system to respond to a wide variety of antigens, and mutations that increase or decrease the effectiveness or activity of a protein compared with wildtype. At this time there is no conclusive evidence that a mutation in the latter category causes or helps to cause MS. The known structural mutations that have been investigated for association with MS have either produced negative or mixed results, or have not been studied sufficiently to make a determination. A stronger case can be made for the involvement of immune system genes in the development of MS, particularly MHC class II genes as alleles and haplotypes in this region have repeatedly demonstrated association with MS. A better understanding of how these alleles affect protein structure and therefore function has the potential to add greatly to our knowledge of the etiology of this disease.
References

Note: Details for each of the MS candidate gene studies listed here can be found in the file phase2-genetic-studies.xls available for download at www.acceleratedcure.org.


73 D. Scalabrini, D. Galimberti, C. Fenoglio, C. Comi, M. De Riz, E. Venturelli, L. Castelli, L. Piccio, M. Ronzoni, C. Lovati, C. Mariani, F. Monaco, N. Bresolin and E. Scarpini, P-selectin glycoprotein ligand-1 variable number of tandem repeats


