From THE DIVISION OF NEUROLOGY
DEPARTMENT OF CLINICAL NEUROSCIENCE
Karolinska Institutet, Stockholm, Sweden

GENETIC VARIATIONS INFLUENCING SUSCEPTIBILITY TO MULTIPLE SCLEROSIS

Kristina Duvefelt

Stockholm 2005

Published and printed by Repro Print AB, Stockholm, Sweden

ISBN 91-7140-573-9

If we knew what it was we were doing,
it would not be called research, would it?

Albert Einstein

To my family

ABSTRACT

Multiple sclerosis (MS) is a chronic neurological disease affecting approximately 12 000 individuals in Sweden. Epidemiological studies have shown that most likely several genetic variations contribute and interact with each other as well as with environmental factors in the development of MS. The disease is characterized by focal infiltration of inflammatory cells in the brain and spinal cord, demyelination of axons as well as neurodegeneration.

In this thesis we examined the risk of MS in relation to genetic variations in selected candidate genes and gene expression induced by neurodegeneration in an experimental animal model disease.

Study I did not show any association in MS with the two candidate genes *NOTCH4* and tumour necrosis factor α (*TNF*), both located in the human leukocyte antigen (HLA) complex on chromosome 6p21.3. We found a difference in carriage count for one single nucleotide polymorphism (SNP) studied in the *NOTCH4* gene but this was shown to be secondary to HLA-DRB1*15, the well known risk factor for MS in the HLA area.

Study II was a two-stage analysis, in which we determined genotypes, in up to 672 MS patients and 672 controls, for 123 SNPs in 66 genes. Genes were chosen based on their chromosomal positions, suggestively linked to MS or other autoimmune disorders, or on the basis of biological function hypothesized to be important in MS. In stage one, 22 genes contained at least one SNP for which the carriage rate for one allele differed significantly between patients and controls. After additional genotyping in the second stage, two genes conferred susceptibility to MS: *LAG3* on chromosome 12p13, and *IL7R* on 5p13. LAG3 inhibits activated T cells, while IL7R is necessary for the maturation of T and B cells.

In *Study III* we examined nerve injury-induced neurodegeneration in the ventral root avulsion (VRA) model using Affymetrix oligonucleotide arrays (interrogating approximately 7000 full-length rat sequences and 1000 EST clusters) in two inbred rat strains differing in degree of nerve cell loss after injury. Gene expression was determined in the spinal cord of naïve animals and at two time points after injury. Differential regulation was found in 278 genes, whereof 245 were regulated by the injury and 68 differed significantly between strains. Principal component analysis revealed a common injury response pattern significantly modified by genetic background. A central finding was that inflammatory genes comprised the largest group of genes induced by injury and that these transcripts prevailed in the strain most susceptible to nerve cell loss. In addition, levels of the strain regulated genes *C1qb* and *Timp1* correlated with degree of nerve cell loss in a group of genetically heterogeneous animals. These results suggest a link between the inflammatory response elicited by nerve injury and subsequent neurodegeneration.

In *Study IV* we investigated, the *CD74* (also known as invariant chain) gene on chromosome 5q33 and Metallothionein (*MT*) genes on chromosome 16q, for a possible importance in determining susceptibility to MS based on their upregulated transcription as well as difference in regulation influenced by genetic background in study III. Seven SNPs in the *CD74* gene and 9 SNPs in the *MT* cluster were analyzed in 890 MS patients and 775 controls. Three SNPs belonging to the same linkage disequilibrium block in the *MT* region were found to be associated with the risk of MS as well as one haplotype. MTs are heavy metal binding proteins that are transcriptionally regulated both by heavy metals and glucocorticoids. They are found throughout the mammalian body, including the CNS and are rapidly induced following many types of CNS insults where being strongly neuroprotective. One SNP in *CD74* showed borderline signs of association not supported by other SNPs in the same LD block.

LIST OF PUBLICATIONS

I. **Duvefelt K,** Anderson M, Fogdell-Hahn A, Hillert J.
A NOTCH4 association with multiple sclerosis is secondary to HLA DR*1501.
Tissue Antigens 2004;63(1):13-20.

II. Zhang Z, **Duvefelt K**, Svensson F, Masterman T, Jonasdottir G, Salter H, Emhaizon T, Hellgren D, Falk G, Olsson T, Hillert J, Anvret M.
Two genes encoding immune-regulatory molecules (*LAG3* and *IL7R*) confer susceptibility to multiple sclerosis.
Genes and Immunity 2005;6(2):145-52.

III. Swanberg M, **Duvefelt K**, Diez M, Hillert J, Olsson T[1], Piehl F, Lidman O.
Genetically determined susceptibility to neurodegeneration is associated with increased expression of inflammatory genes

Manuscript

IV. **Duvefelt K,** Swanberg M, Lindgren CM, Lidman O, Olsson T, Hillert J.
An association study of *CD74* and Metallothioneins in multiple sclerosis
Manuscript

CONTENTS

LIST OF ABBREVIATIONS

APC	antigen presenting cell
APL	altered peptide ligand
BBB	blood brain barrier
CARD15	caspase recruitment domain family, member 15
CDCV	common disease common variant
CNS	central nervous system
CSF	cerebrospinal fluid
DA	dark agouti
DHPLC	denaturing high-performance liquid chromatography
DZ	Dizygotic
EAE	experimental autoimmune encephalomyelitis
EDSS	Expanded Disability Status Scale
Gd	Gadolinium
HHV	human herpesvirus
HLA	human leukocyte antigen
IFN	interferon
Ig	immunoglobulin
IL	interleukin
LAG-3	lymphocyte activation gene 3
LD	linkage disequilibrium
MALDI-TOF	matrix-assisted laser desorption/ ionization time-of-flight
MBP	myelin basic protein
MHC	major histocompatibility complex
MOG	myelin oligodendrocyte protein
MRI	magnetic resonance imaging
MS	multiple sclerosis
MSFC	MS functional composite
MSSS	Multiple Sclerosis Severity Score
MZ	monozygotic
NCBI	National Center of Biotechnology Information
PCA	principal component analysis
PCR	polymerase chain reaction
PLP	proteolipid protein
PP	primary progressive
PTPN22	protein tyrosine phosphatase, non-receptor type 22 (lymphoid)
PVG	Pievald Virol Glaxo
QTL	quantitative trait locus
RA	rheumatoid arthritis
RR	relapsing remitting
SNP	single nucleotide polymorphism
TCR	T cell receptor
TNF	tumour necrosis factor
VRA	ventral root avulsion

1 INTRODUCTION

1.1 MULTIPLE SCLEROSIS

Multiple Sclerosis (MS) was first described in the 19th century by several investigators, in 1868 JM Charcot recognized it as a distinct entity (1). Accumulation of perivascular inflammatory cells within the brain and spinal cord white matter was detected in patients with episodes of neurologic dysfunction. Today's estimation is that approximately 12 000 people in Sweden suffer from MS (2). The worldwide figure is less certain, but at least 2.5 million cases of MS are believed to exist. The mean age of onset is 28 years usually ranging from 20 to 50 years of age.
The exact cause of MS is unknown; however, several lines of evidence suggest that it is an autoimmune disease directed against central nervous system (CNS) myelin-producing oligodendrocytes leading to breakdown of the myelin sheaths and substantial neuronal loss (reviewed in (3)).

1.1.1 Clinical features

MS predominantly affects women; about two third of the patients are women (reviewed in (4, 5)). The clinical course could either be a relapsing remitting form (RRMS) or a primary progressive (PPMS) variant. These categories are defined based on the clinical course rather than on known differences in genetics or pathogenesis. About 85 % of the patients have the RRMS type, which is characterized by relapses and between them episodes of recovery. The majority of these patients will eventually convert to a secondary progressive disease course which is characterized by slowly worsening of symptoms without periods of recovery. PPMS patients (about 15 % of the patient population) are directly be affected by a progressive phase (4, 5).

Since the progression phase of the two MS forms is very similar, the notion is gaining that these forms should not, as still common practice, be treated as two separate entities with regard to treatment and clinical studies (6).

A person with MS may experience a wide variety of symptoms including sensory disturbance, loss of balance, loss of vision, bladder dysfunction, fatigue, pain, temperature intolerance, sexual dysfunction and cognitive impairment (5).
The early course of MS is quite variable from patient to patient and for an individual MS patient it is difficult to establish (7). On average patients require a walking stick at a median of 15 years following diagnosis(8). At that point about 10% will need a wheelchair while 20-25% have no trouble walking without aid. In a Swedish study 31 % of 166 MS patients were retired due to medical reasons and 20 % were on full time sick leave (9). These figures demonstrate the profound impact the disease has on persons affected.

1.1.2 How is MS diagnosed?

The diagnosis of MS is based on both clinical symptoms as well as laboratory investigation. The latest guidelines for diagnosing MS were published 2001 by an international committee headed by McDonald (10). The emphasis is on the objective

demonstration of dissemination of brain lesions in both time and space. Magnetic resonance imaging (MRI) is integrated with clinical symptoms and investigation of the immunoglobulins (Ig) in the cerebrospinal fluid (CSF).
At MRI investigation the most common finding is multifocal white matter lesions of different age and size. Active lesions are demonstrated with the use of Gadolinium (Gd) enhancement. In the CSF analysis the hallmark of MS is the presence of oligoclonal IgG bands (without corresponding findings in the serum) and an elevated IgG index (IgG in CSF/IgG in serum) indicating expansion of B cells and resulting IgG synthesis in the CNS compartment. A consensus on how to perform the CSF analysis has recently been published (11).

The most used instrument for assessing disability is the Expanded Disability Status Scale (EDSS) (12). This ordinal scale is based on neurological examination and patient history. The scale ranges from 0 (assessed normal by neurological examination) to 10 (death due to MS). The EDSS has some weaknesses, it has for instance been considered to insensitive in certain respects and it has been suggested to use other scales in controlled clinical trials as the MS Functional Composite (MSFC) (13). This test combines the result of three performance tests: the nine hole peg test, the timed 25-foot walk and the three-second paced auditory serial additional test. So far, however its superiority to EDSS scale has not been proven.

Recently an additional measurement, Multiple Sclerosis Severity Score (MSSS) has been introduced (14) with the aim to better describe the progression of disability in MS. The core principle of MSSS – correcting single assessment score by comparison with the distribution in patients of similar disease duration – could be applied to both EDSS and MSFC. The score builds on an international cohort of 9892 EDSS evaluated patients with known disease duration. The MSSS score determined for a patient reflects the rate of progression from onset in comparison to what has been recorded for other patients with the same duration.

1.1.3 Health economics of MS

The total cost of MS in Sweden 1998 has been estimated to 4868 MSEK, giving an annual cost of 442 500 SEK (56 550 USD, 1 USD = 7.84 SEK) per patient (15). Direct costs accounted for about 67% of total cost, and these were dominated by the cost of personal assistants and pharmaceuticals. Indirect costs (loss of production) accounted for about 33% of total costs. To these economic costs, intangible costs due to reduced quality of life, costs of 2702 MSEK should be added as well. Direct, indirect and informal care costs all rose significantly with increased disability and were higher during a relapse. It has been shown in this Swedish study as well as several others, that costs increase considerably with increased disability (reviewed in (16)). It can be concluded that a high burden is imposed on the society, in terms of production losses by an essentially very young disabled patient population, as well as on families, with a very high informal care (17).

1.1.4 Treatment strategies for MS

There has been a dramatic change in the treatment of MS over the past decade. This is due to progress in several different areas, immunubiological knowledge about MS, advances in biotechnology, improvement in MRI techniques as well as better design of clinical trails. In clinical practice, two molecules, β-interferon and glatiramer acetate, dominate disease modifying treatment of RRMS.
β-interferon is marketed by three different pharmaceutical companies and the molecules differ in how they are produced. Betaferon®(18) is a β–interferon 1b (IFN β–1b) is produced in bacteria and sold by Schering. Rebif®(19) (sold by Serono) and Avonex®(20) (sold by Biogen) are β–interferon 1a (IFN β–1a) are produced in mammalian cells. Betaferon is non glycosylated and differs from the native β-interferon in one amino acid position as well as in the first amino acid that is not incorporated in the drug. The amino acid sequence of Rebif and Avonex are identical to that of natural human β-interferon and they are glycosylated by the cell they are produced in.

The effect of IFN- β compounds is shown by the reduced the number of clinical relapses (about 30% reduction) as well as reducing the number of Gd enhancing lesions in RRMS (reviewed in (21)). Antiviral effects, pleiotropic effects on the immune system and on the blood brain barrier could all contribute to a benefit in patients with MS (22). It is not known which of these mechanisms underlies the reported efficacy of interferons in the disease.

Glatiramer acetate, the active ingredient of Copaxone® (23, 24) (marketed by Aventis), consists of the acetate salts of synthetic polypeptides, containing four naturally occurring amino acids: L-glutamic acid, L-alanine, L-tyrosine, and L-lysine. Glatiramer acetate did not show any beneficial effect on the main outcome measures in MS, i.e. disease progression, and it did not substantially affect the risk of clinical relapses when a large meta analysis was performed on a large number of Copaxone clinical trials (25). Even so the compound has been shown to reduce the proportion of new lesions evolving into permanent black holes (26). It has also been shown to downregulate the T_H1 cytokines – believed to be detrimental in MS. Glatiramer acetate was thought to bring about this shift from T_H1 by acting like an altered peptide ligand but more recent work has shown that the drug notably affects the properties of antigen-presenting cells, such as monocytes and dendritic cells (reviewed in (27)).

In November 2004 another pharmaceutical compound, natalizumab, marketed under the name of Tysabri by Biogen/IDEC Elan, was approved for treatment of MS in the USA (28). This is a humanized monoclonal antibody to α4β1 Integrin which exerts its action by blockade of lymphocytes binding to the endothelium and thereby reduces the inflammatory infiltrate into the brain (reviewed in (29)). This drug reduced the frequency of clinical exacerbation to a higher degree than any previous drug on the market but was withdrawn from the market in February 2005 due to three cases of progressive multifocal leukoencephalopathy (PML). Two of the cases were MS patients who received a combination therapy of Tysabri and Avonex while the third patient suffered from Crohn's disease and received Natalizumab together with other immunomodulatory drugs (29). The proposed mechanism for this adverse event is an

opportunistic infection with JC virus, which causes PML, due to diminished immune surveillance of the brain as an effect of the α4β1 Integrin blockade by Tysabri.

A number of unsuccessful phase II clinical studies have been undertaken with a variety of agents, aiming at wide spectrum of targets. To only mention a few, trials have been performed with agents directed towards: T cell receptors (TCR), tumour necrosis factor α (TNF-α), CD4 and CD3.
There have also been clinical studies with agents aiming at intervening with a supposedly pathogenic immune response, for example altered peptide ligands (APL) and oral administration of myelin to provoke tolerance. Wiendl and Hohlfeld have reviewed these and other therapeutic approaches in MS (30). They conclude that: "Failed strategies are of high importance for a critical revision of assumed immunopathological mechanisms, their neuroimaging correlates, and for future trial design."

1.1.5 Pathogenesis of MS

In the name of the disease, multiple sclerosis, the hallmarks of its pathology are given: multiple = a large number, sclerosis = pathological hardening of tissue (scars). The scars are large confluent areas of demyelination accompanied with astrocytosis. Histopathological studies have shown that demyelination and inflammation involving B cells, T cells, macrophages and activated microglia are present in the acute MS lesion. The destruction of myelin leads to reduced velocity of the axonal action potential or even conduction block leading to clinical manifestation.

A question that has occupied the MS researchers for more than a hundred years has been WHY these large numbers of sclerotic plaque appear in the brain?
A number of hypotheses have been proposed and attempts have been made to test them. Since, in the brain of MS patients, i) large numbers of inflammatory infiltrating cells ii) demyelination of the axons as well as iii) degeneration of the neurons has been detected (3) three main hypotheses have crystallized during the years. According to them MS could be a consequence of either:

1. Autoimmunity
2. Infection
3. Primary neurodegeneration

Autoimmunity
Very early in the study of immunity it was realized that the powerful adaptive immune response effective in defending the host could, if turned against the host, can cause severe tissue damage. Ehrlich termed this *horror autotoxicus,* today known as autoimmunity.
According to the autoimmune hypothesis, T-cells reactive to myelin components or other brain antigens are activated in the periphery encountering foreign antigens (a process known as molecular mimicry). The T-cells could also become activated through other mechanisms of tolerance loss. These T-cells become activated and will be able, by expression of cell-surface adhesion molecules, to migrate through the blood brain barrier (BBB). In the brain they will be reactivated by local antigen presenting

cells (APC) presenting the "self" antigen and this activation will lead to cytokine release which in turn lead to tissue damage by direct action as well as antibody and CD8+ T-cell mediated cell death (reviewed in (31)).

The support for MS being an autoimmune disorder is strong: Myelin protein antigens are able to provoke a MS like disease, experimental allergic (autoimmune) encephalomyelitis (EAE), when given to susceptible animals, EAE can also develop by passive transfer of encephalitogenic CD4+ T-cells into naïve animals (32), MS susceptibility is genetically associated with major histocompatibility complex (MHC) class II, (reviewed in (33)) and ,at least in Scandinavia, to MHC class I (34, 35), MS patients respond to immunomodulatory drugs (se above). High avidity myelin specific T-cells are found at a higher frequency within MS patients than controls (36). Of course some of this evidence can also be in agreement with other hypotheses – as will be described below.

There is a lack of some critical evidence to make this hypothesis conclusive, particulary the lack of an autoantigen in MS, as well as lack of transfer experiments in animal model of autoantibodies that are able to mediate the pathogenesis. A great deal of research has been conducted in these fields without success.

Infection

Several hypotheses have been put forward on how infectious agents could influence MS. They could either be directly causative or be a contributing factor in the autoimmune process. Viruses have been shown to induce autoimmunity by different mechanisms, molecular mimicry, induction of MHC molecules, exposure of hidden cellular components, activation of superantigen and apoptosis dysregulation.

Support for a role of an infectious agent as being important in the pathogenesis largely originates from epidemiological studies (37). (See epidemiological section below). One possibility to explain these epidemiological data is the "prevalence" theory of MS, suggesting that the disease is most common where the causative agent is most widespread.

Other evidence for an influence by infectious agents in MS are the persistent oligoclonal IgG found in CSF from the vast majority of MS patients (38). One interpretation would be a chronic persistent infection of the CNS since similar oligoclonal banding pattern are found in several infectious diseases. Additionally, a number of viruses can induce demyelination in animal models. The most studied is infection of mice with Theiler's murine encephalomyelitis virus (TMEV) (39). Another evidence for infectious triggers stems from EAE studies. Transgenic mice expressing a TCR that is specific for an encephalitogenic myelin basic protein (MBP) peptide develop EAE when they are housed in an environment with pathogens, whereas the same animals remained free of disease in a pathogen free environment (40).

A large number of pathogens have been proposed to be involved in MS pathogenesis (41). The latest being the human herpesvirus 6 (HHV-6) (42) and *Chlamydia pneumonia* (reviewed in (43)). But the evidence for association between these

organisms and MS is weak. Further studies need to clarify their contribution. So as yet no pathogen has been identified as contributing to MS pathology.

Primary degeneration

Barnett and Pineas (44) have reported that changes in oligodendrocytes may be the initial events in formation of an acute lesion, serving to recruit an initial innate (microglia) response and subsequently an adaptive (T cell) immune response. From a study set out to assess early axonal pathogenesis it was concluded that a significant irreversible axonal damage occurs in patients at the earliest stage of multiple sclerosis (45). In considering this hypothesis one can have thoughts on how to fit in the HLA association as well as the extent and chronicity of the inflammation in the brain of MS patients.

As seen above these three are not in total contrast to each other, as combinations are possible, such as an autoimmune response with reactivation of tissue resident virus triggered by a vulnerability of neuronal cells induced by external stressors. The sequence of these events could quite obviously be the opposite, and such hypothesis is hard to test.

Lesion pathology

Distinct patterns of tissue injury in multiple sclerosis (MS) lesions, have been described by Lucchinetti and Lassman (46). They pointed out that the compositions of acute demyelinating lesions vary greatly among patients. Four patterns of lesions were defiend: Pattern I with macrophage mediated demyelination, Pattern II with antibody mediated demyelination, Pattern III with distal oligodendrogliopathy and pattern IV primary oligodendroglia degeneration and secondary demyelination. Pattern IV was only found in a few patients. Patterns I and II showed close similarities to T-cell-mediated or T-cell plus antibody-mediated autoimmune encephalomyelitis, respectively. The other patterns, III and IV, were more reminiscent of virus- or toxin-induced demyelination rather than autoimmunity. Different lesions of any given patient were of the same type, indicating that the heterogeneity does not reflect different stages but rather distinct mechanisms. This pathogenetic heterogeneity of plaques from different MS patients may implicate different mechanisms for the disease in different patients.

There is evidence that pronounced axonal injury occurs in actively demyelinating lesions (47, 48) but it has also been shown that a continuing axonal injury occurs in old inactive lesions (49) partly explaining the progression of disability even when evidence of inflammation is scarce. The destruction of axons, when past the threshold of compensation, is likely to be accompanied by irreversible clinical deficits.

1.1.6 Epidemiology

The epidemiology of MS has been intensively studied. The common conclusion from the studies is that genetic factors as well as environmental factors act together to cause disease. The disease tends to be rare in tropical areas and more common in temperate areas, although there are some exceptions (50).

Kurtzke (37) has defined different regions in the world depending on prevalence. High prevalence, >30 per 100 000, include northern Europe, the northern USA and Canada, southern Australia and New Zealand. Medium prevalence areas, 5-30 per 100 000, include southern Europe, the southern USA and Northern Australia. Low prevalence areas, <5 per 100 000, areas include Asia and South America. Within areas where the disease is common, some ethnic minorities have a low prevalence, exemplified by the Samis, Turkmen, Amerindians, Canadian Hutterites and New Zealand Maoris (50). The reason for the variation in incidence and prevalence is not known. Both genetic and environmental contributions have been put forward as a cause of this pattern. For example the north-to-south gradient in the US appears to match patterns of migration of Europeans at higher or lower risk. Also, the rarity of MS in some ethnic minorities (se above) point to a different ethnic susceptibility as being of importance (50).

Migration studies have shown that people who move from an area where the disease is common to an area where the disease is rare show a decrease in rate of disease. By contrast, people who migrate in the opposite direction tend to retain the low risk of their country of origin (reviewed in (51)). Results from studies which have examined the effect of age at migration on risk of multiple sclerosis suggest that an individual's risk of the disease is largely established during the first two decades of life. This is shown by the change in risk of people migrate (from high- to low incidence area) before 15-16 years of age and the preserved risk if the migration takes place after 15-16 years (reviewed in (52)).
The risk of MS can change rapidly between generations: although migrants from low risk countries to high risk countries retain their low risk, their children have a risk of multiple sclerosis that approaches that of the host country (reviewed in (53)).

The interpretation of this data could be that an exposure to a pathogen early in life gives a higher risk – it could also be interpreted as shortage of something during childhood that would protect from MS. Vitamin D has been proposed to be such an agent (reviewed in (54)). The shortage of vitamin D could be due to the decrease in sunlight depending on latitude. UV radiation influences the biosynthesis of vitamin D.

Other interesting reports regard the findings of “epidemic” MS. The most famous study is the report by Kurtzke of an apparent MS epidemic in the Faraoe Islands (55), but is has been challenged (56, 57) since the study has weaknesses both on statistical grounds as well as case ascertainment problems. In a recent review Marrie concludes that it is doubtful that the Faraoe Island report truly represents an epidemic of MS (51).

1.1.7 Genetic epidemiology

Before embarking on large genetic studies of a disease it is critical to assess whether there truly is a genetic component influencing the disease. Elucidation of the genetic contribution is often performed through studies assessing whether persons sharing genetic material to different degrees have the same or different risk of being affected by the disease. It is known that 15-20 % of all MS patients have a close relative diagnosed with MS (58). This could in theory both be due to a genetically encoded risk or environmental factors or a combination of both.

The genetic component in MS is demonstrated by the 25-30% concordance rate in monozygotic (MZ) twins (who share almost 100 % of their genetic material) (59-63). The concordance rate drops steeply to a rate below 5% for dizygotic twins (who on average share 50% of their genetic material) strongly indicating that many genes in combination, each probably exerting a relatively small effect, influence susceptibility to MS (64).

If the cause of getting a disease is totally coded within the genes the concordance rate for monozygotic twins would be almost 100 %, so one further conclusion that can be drawn from the twin studies is that, as the concordance rate in MZ-twins is considerably lower than 100%, environmental factors also play a role. But it should be noted here that identical twins are not completely genetically identical, they differ in their repertoire of antibodies and TCR, somatic mutations, the pattern of X-chromosome inactivation (in females) and in the number of mitochondrial DNA molecules.

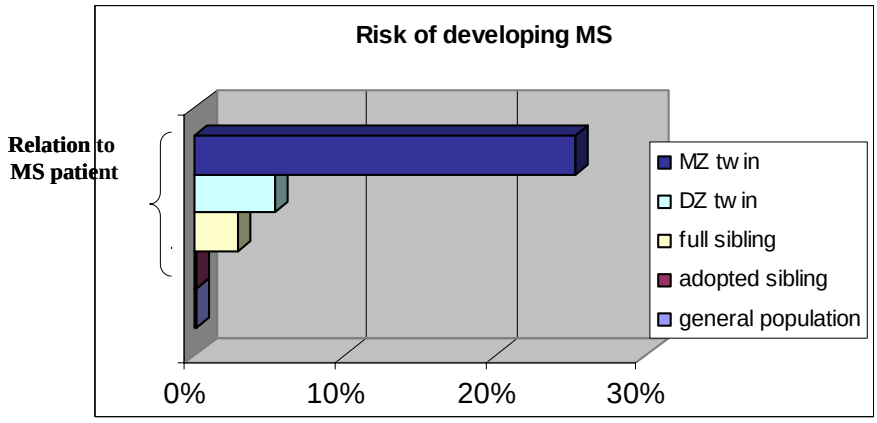
A first degree relative such as a sibling (siblings share on average 50 % of their genetic material) or child has a 3-5 % risk to develop MS which is 15-25 times the risk of the general population (58). For monozygotic twins the risk is increased 300 times (59, 61, 62), *Figure 1*. It has been shown that the risk for MS in offsprings from parents, where both suffer from MS, is very high (65, 66). The risk for these children was 30.5% compared with having one parent with MS – then the risk was 2-5%. These results also imply a genetic component of the disease.

An adoption study clearly showed that adoptive relatives although raised from infancy with the MS patient, were no more likely to develop MS than the general population (67). This gives an indication that the familial aggregation seen in MS is related to sharing of genetic material rather than shared environment.

Ebers and colleagues studied 13 000 spouses of multiple sclerosis patients and found that they had the same rate of MS as the general population (65). This will speak against a transmissible agent – if indeed there are such environmental factors, these needs to act equally in the whole population and not to be enriched in the family milieu.

The conclusion from the genetic epidemiology studies in MS is that there indeed is a genetic component contributing to the risk of MS but there are also other factors of considerable importance.

Figure 1. Concordance rates for relatives to MS patients



1.2 ANIMAL MODELS USED IN MS RESEARCH

A strategy to study complex diseases in humans is to use animal models that mimic the human disease in different aspects. As close mimicking as possible of the human disease characteristics is desirable. Even if not the same as in MS, genes identified in animals may unravel pathways with potential impact on MS. The understanding of molecular pathways regulating inflammation and neurodegeneration is of fundamental importance for unraveling disease mechanisms in MS and other human disorders. By the identification of genes predisposing to pathogenesis in the animals, the aim is to find clues to which genetic elements are also important in MS.

1.2.1 Experimental autoimmune encephalomyelitis

Experimental autoimmune (allergic) encephalomyelitis (EAE) is the prevailing experimental model for MS, recapitulating many features of the disease, and was established 70 years ago after an observation by Rivers and Remlinger that rabies vaccines contaminated with myelin components could provoke encephalomyelitis. EAE is an inflammatory disease model of the CNS with variable degrees of demyelination which is induced by immunization of susceptible animals with myelin antigens and Freund's adjuvant. It should be noted that EAE can be provoked in most rodents and some monkeys and the susceptibility as well as the course is highly strain dependent (68). EAE could also be induced by adoptive transfer of T cells. CD4+ T cells transmit the disease and among them the cells secreting T helper (T_H) 1 cytokines – interferon γ (IFN-γ), tumour necrosis factor-β (TNF- β) and interleukin-2 (IL-2) as well as the proinflammatory cytokine TNFα – are more likely to transfer the disease than other myelin specific T cells (reviewed in (69)). T cells of T_H2 type that secret IL-4, IL-5, IL-10 and IL-13 seem to protect animals from disease (70). This knowledge as well as the fact that administration of the T_H1 cytokine (IFN-γ) to MS patients gives rise to exacerbations has been important in the formation of the hypothesis that MS is a T_H1 mediated autoimmune disease.

In EAE it has also been shown that MBP specific $CD8^+$ T cell clones could be encephalitogenic and induce a disease phenotype that exhibits some similarities to MS

not seen in myelin specific $CD4^+$ T cell–mediated EAE (71). In this model lesions were confined to small blood vessels and the surrounding nervous tissue and the clinical signs differed from classical $CD4^+$ EAE in that they had a presence of ataxia and spasticity. This finding as well as the quite recent findings that oligoclonal CD8+ T cells are found in MS brain (72) and that $CD8^+$ T cells are more prevalent in MS brain than CD4+ T cells are (73) has led to the conclusion that both $CD4^+$- and $CD8^+$ T cell responses contribute to MS pathogenesis.

Another interesting observation is that spontaneous EAE can develop in transgenic mice expressing T cell receptors specific for the myelin components, myelin basic protein (MBP) (40, 74) or proteolipid protein (PLP) (75).

1.2.2 A neurodegeneration model

Ventral root avulsion (VRA) is a model of nerve injury-induced neurodegeneration and inflammation (reviewed in (76)). As described above MS lesions have a considerable amount of neurodegeneration. The degree of neuronal atrophy may be the best surrogate marker for permanent handicap (77). Dissection of potential nervous system factors of importance for neurodegeneration is not easily studied in the human. Neither are EAE models useful since the degree of CNS damage will depend on the systemic autoimmune response as well as target inherent factors. In the VRA model unilateral avulsion of ventral roots is performed and this leads to a retrograde reaction where microglia cells are activated, start to express MHC antigens and relocate to position around the cell bodies of axotomised motoneurons. With some delay astrocytes become activated as well and an influx, although limited, of T lymphocytes occurs. A large proportion of the lesioned nerve cells will subsequently die and the cell debris will be removed by phagocytic microglia. As in the EAE situation, different animal strains show different susceptibility and the degree of nerve cell loss differs greatly (78-80).

Interestingly it has been demonstrated that EAE induced by active immunization with an encephalitogenic myelin basic protein peptide dramatically reduces the loss of spinal motoneurons after ventral root avulsion in rats (81). Neurotrophic factors were detected in T- and NK-cells but also high levels of TNFα and IFNγ which are usually regarded as harmful to CNS cells. These results suggest that the potentially neurodamaging consequences of severe CNS inflammation are restricted by the production of several potent neurotrophic factors in leukocytes.

1.3 GENETIC STUDIES OF COMPLEX DISEASES

MS is in the literature described as a complex disease since the etiology is multifactorial, depending on both the carried genetic makeup as well as the environment. This leads to a complex inheritance pattern. The gene variations involved in a complex disease may neither be necessary nor sufficient for being affected by the disease. Gene products are thought to interact with each other as well as with environmental factors which could be pathogens, sunlight and exposure to toxins, just to mention a few possible candidates. Other diseases that are labeled as complex is for

example: rheumatoid arthritis (RA), diabetes mellitus, Alzheimer disease (AD), schizophrenia and hypertension.

On the other end of the spectrum we have the Mendelian disorders so named because they follow a mode of inheritance that was described by the Austrian monk, Gregor Mendel, in the mid 19th century who published his first work in 1865. Usually these diseases are monogenic: mutations within a single gene are the only cause of the disease. By that they will also follow one of the five the Mendelian inheritance pattern: autosomal dominant, autosomal recessive, X-linked dominant, X-linked recessive or Y-linked. Unlike for multifactorial diseases, environmental factors do not play such an important role in the development of the Mendelian diseases. Well-known examples of monogenic diseases of the central nervous system are Huntington's disease, familial Alzheimer's disease and spinocerebellar ataxias.
Positional cloning has been very successful in finding the mutated gene for many monogenic disorders. The success has relied on having access to large families, with several members affected by the disease, which can be studied by linkage analysis. Today more than 1400 genes for approximately 1200 Mendelian traits have been identified.

Performing genetic studies of complex disease is an ambitious endeavor. Genes influencing complex diseases generally *contribute* to the susceptibility rather than being *deterministic* as in the Mendelian diseases. Usually large pedigrees (useful for positional cloning), with many affected individuals, are not often found. If they are found they could be useful for unraveling pathogenic mechanisms – but they could also be problematic since there could well be a special mechanism in a given family that may not apply to the general disease population. In the case of a seemingly Mendelian inheritance pattern a very careful clinical investigation is vital to rule out other differential diagnoses.

How to find genetic variations in a complex disorder?

Naturally occurring variations in the DNA sequence are used to facilitate the study of the genetic component in a complex disease. These sequence variations can either be causative themselves or act as markers for linked variations.
Two main types of markers are usually utilized: microsatellites and single nucleotide polymorphisms.

Microsatellites
Microsatellites, which are tandemly repeated DNA sequences, were discovered in 1989 (82). The repeat units are generally di-, tri- tetra- or pentanucleotides. The repeat unit can occur up to 50 times in a row. Microsatellites are quite uniformly distributed throughout the genome and tend to occur in noncoding regions of DNA but there are examples of human disorders that are caused by microsatellites in the coding region, for example Huntington's disease (83).
Since they are highly polymorphic they are very useful as genetic markers, especially in studying co-segregation of a marker and disease status within families. The relatively high polymorphic nature of these markers is due to the slippage of DNA polymerase during replication. This mechanism will lead to a different amount of repeat units in the newly synthesized DNA strand in comparison with the template strand. The mutation

rate for microsatellites is 1/1000 per site per generation (84), which is more than 10 000 times higher in a non repeated sequence.

Single nucleotide polymorphisms
Single nucleotide polymorphisms (SNPs) are defined as: Single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in the normal population(s) wherein the least frequent allele has an abundance of 1% or greater. The SNPs can be directly causative of disease but are usually reflecting the natural variation in the species.
On average the human genome contains one SNP per 1200 nucleotides (www.hapmap.org). This will lead to several million base pair differences between any two unrelated individuals. SNPs have different impact depending on their location – in regulatory, coding or noncoding regions of the genome. If they are present in the coding region one differs between the synonomous SNPs which do not give rise to any amino acid change and nonsynonomous SNPs that will change the amino acid in the encoded protein. The mutation rate for SNPs is low, around 10^{-8} per site per generation, which means that most SNPs have arisen from a single historical mutation event.

During the last year the knowledge on these variations in the human has exploded due to the international Hapmap project (85) which is a multi-country collaborative effort to identify and catalogue genetic similarities and differences in human beings. The project is a collaboration between scientists and funding agencies from Japan, the United Kingdom, Canada, China, Nigeria, and the United States. All information generated by the project is released into the public domain. The latest updated version of the NCBI database, dbSNP build 125, contains 5 million validated SNPs.

These variations are also very popular since they are easy to genotype in a high throughput fashion – something that is becoming more and more important in the field of complex genetics.

Studies on genetic susceptibility in complex traits mainly rely on either of two principally different methods: Linkage analysis or association analysis

Linkage Analysis

Traditionally an unbiased selection of up to 3-400 markers covering the genome is utilized. In the more recent studies a larger number of markers are often utilized.
Linkage is based on co-segregation of a marker and disease status within families. For this type of analysis related affected individuals are needed, either multi-generational pedigrees or sib-pairs concordant for the disease.

Linkage is the tendency for two loci on the same chromosome to be inherited together more often than by chance alone. The most powerful method uses multigenerational pedigrees with multiple affected relatives where the mode of inheritance is known. Since such pedigrees not often exist in complex disorders the more commonly used method is the affected sib-pair analysis, reviewed in (86). The basis for this analysis is: if there is linkage, affected individuals will be more similar in those parts of the genome close to a disease susceptibility gene than would be expected by chance. The analysis performed in this case is nonparametric as there is no clear mode of inheritance.

This method has traditionally been very successfully in mapping genes that underlying monogenic diseases (87) but has not been as rewarding in the studies of complex traits (88). The lack of success could be due to several factors: low heritability in complex traits, the inability of the standard set of microsatellite markers -which are spaced at least 10 cM apart- to extract complete information, the imprecise definition of phenotypes and inadequately powered studies (89). Linkage analysis is also much less powerful for identifying common genetic variants that have modest effect on disease (90, 91).

Association analysis

Association analysis has so far mostly been hypothesis based as potential selected candidate genes are studied. The selection is based upon knowledge about pathogenic mechanisms and genome areas of interest. The distributions of alleles at polymorphic sites (usually SNPs) in the genome between persons with and without disease are studied. Without the sequencing of the human genome (92, 93) this kind of studies had been almost impossible to perform. It is now relatively easy to find your gene of interest as well as polymorphic markers within the gene due to the research performed by the Hapmap consortium (85) (se discussion under the heading of SNPs). Association analyses of candidate genes have identified genes that confer risk to common complex diseases (94-97).
New methods have recently been introduced that makes it possible to perform association analysis in a more hypothesis generating way by simultaneously analyzing 500 000 SNPs throughout the genome (http://www.affymetrix.com/products/arrays/specific/500k.affx).

Association studies are usually performed in a case/control material composed of cases and unrelated controls. An alternative strategy is to use family based controls. Family based controls are less sensitive to population stratification but on the downside they are less powerful than unrelated controls.

The most comprehensive approach is to re-sequence the entire gene of interest in patients and controls. This would give the most valid information but is still very laborious and costly. The most common approach is to use already discovered SNPs (microsatellites are considered to be too polymorphic, and its alleles possibly to young as well as not abundant enough for use in association studies) to perform the analysis. Since it is very unlikely that the investigated alleles are the causative ones, one has to rely on highly correlated alleles that are said to be in Linkage Disequilibrium (LD) with the causal allele. This approach is usually termed “LD mapping “or “indirect association”.

LD mapping

LD is the non-random association of alleles at adjacent loci. When a particular allele at one locus is found together with a specific allele at a second locus, more often than if the loci were segregating independently, the loci are in linkage disequilibrium. When a SNP variant is first introduced into a population by mutation it will be perfectly correlated with nearby variants, but over successive generations meiotic recombination will break up the correlations and LD will decay. Many factors can influence LD including: genetic drift, population structure, natural selection, variable recombination and mutation rates and gene conversion (98).

There are different measurements of LD – the two most common being D' and r^2. The underlying measure for both of them is the measure of disequilibrium, D. Consider two adjacent loci A and B with the alleles A,a and B,b. The observed haplotype frequency of the alleles A and B is P_{AB}. If the loci are independent the haplotype frequency would be P_A x P_B . (P_A = frequency of allele A, P_B = frequency of allele B) The measure of disequilibrium D is then calculated as $D= P_{AB}- P_A$ x P_B. This value is however not useful as measure of strength because it is dependant upon allele frequencies. D', on the other hand, is determined by dividing D by its maximum possible value given the allele frequencies at the two loci. So if D' is equal to 1 the two SNPs have not been separated by recombination during the history of the sample. But the magnitude of D' strongly depends on samples size for SNPs with rare alleles therefore values of D' close to one provide indication of minimal historical recombination but intermediate values should not be used for strength comparisons.

The measure of r^2 is in some way complementary to D'. It describes the correlation of alleles at two sites and is calculated as $r^2= D^2/P_A$ x P_a x P_B x P_b. r^2 will only get the value of 1 if the markers have not been separated by recombination and have exactly the same allele frequency. In LD mapping r^2 is frequently used because it correctly takes into account the differences in allele frequency of the two loci when estimating the sample size needed. The sample size must be increased by roughly $1/r^2$ when compared with the sample size for detecting association for the susceptibility locus itself. In trying to reduce the number of redundant SNPs typed r^2 is also used because it describes how much information one marker provides about another. The upper limit of r^2 for a given set of markers is the D' value for these markers. An excellent review of LD is written by Ardlie (98).

A quite new way of performing association analysis is to compare the haplotype frequencies between cases and controls. Haplotypes are the combination of alleles with high LD at one chromosome inherited as a unit. It has been shown that haplotype blocks exist in the human genome and that they capture most of the genetic variation in the region (99). Haplotypes can also code for functional units in the subsequent protein or in the regulation of transcription as shown in the case of leprosy (100).

Since many association studies use case/control material as their core, the parental origin of each allele is not known and haplotypes need to be estimated from the genotype data. The EM-algorithm (Expectation-Maximisation) (101) is widely used for this purpose.

Admixture mapping

This association method was proposed 50 years ago (102) but has not been possible to perform until very recently (103, 104) since reliable high density maps of the human genome as well as methods for analyzing the data have been lacking. Admixture mapping identifies genomic regions where individuals with the disease tend to have unusually high proportion of ancestry from a population where the disease is more prevalent.

What are the potential weaknesses of association analyses?

Any method is impaired by weaknesses – genetic association analysis is certainly not an exception, but rather an example where many difficulties need to be considered:

- *Ascertainment of cases* – extremely important since small interference of persons not carrying the disease could give spurious results when searching for variations leading to small effects.
- *Selection of control group* – The selection of controls has been extensively discussed (105, 106) since controls need to be drawn from the same population as the cases.
- *Selection of candidate genes* – The human genome contains about 30 000 genes. Selection based on knowledge from linkage studies and functional data can be very challenging to perform (107).
- *Statistics* - When is something statistically significant? Since we are searching for small effects we will need very large materials to get significance levels that are beyond doubt. Should we correct for multiple testing? This is a never ending discussion. Applying correction for every independent test is a very stringent way and will perhaps lead to false negative results- i.e. there is a true effect that we miss while not doing it perhaps will give false positives. Many studies have been underpowered and we can not say anything about the variations analyzed but to have power for detecting small contributions the studies need to be very large (108).
- *Genotyping method* - The accuracy of the genotyping is of course also extremely important since this is a cornerstone in the analysis. Genotyping errors are not uncommon (109).
- *Gene-gene interaction* - The hypothesis behind a complex disease is really that many factors are involved but still it is not usual to study interactions between different genes. Perhaps we will only find an association if more than one factor is present.
- *Replication* – Even if everything above is considered, small differences due to some unforeseen problem or a simple sample bias urge for replication studies in independent samples (110) before any genetic association could be considered as proven.

1.4 GENETIC KNOWLEDGE OF MULTIPLE SCLEROSIS

The history of MS genes begins already in the 1970-ies with the finding of an association within the HLA complex to the risk of being affected by MS (111, 112). The HLA complex spans approximately 3 Mb and harbors about 200 known genes many encoding proteins crucial in adaptive and innate immunity.
When molecular genetic analysis became available, the MS-associated HLA specificity was soon identified as the haplotype:
DRB1*1501, DRB5*0101,DQA1*0102,DQB1*0602 (113). This haplotype is a class II haplotype. The class II molecules are known to present externally derived antigens to $CD4^+$ T cells (114). The haplotype, commonly simplified DRB1*15, is present in 60% of the patients and 30 % of the controls, these figures stems from populations of Northern European descent. In Sardinia there also exists association with the DRB1 gene but in this population most clearly with DRB1*03 and DRB1*04 confers the risk for developing MS (115-117).

Two different studies have shown a dose effect of DRB1. Carrying two copies of the risk allele DRB1*15 gives a substantially higher risk than carrying one (118, 119). The risk augmentation if you are carrying one allele is about 3 times and if you are homozygous the risk is about 7-8 times higher than the risk for the general population. The HLA class II haplotype has been estimated to account for 14% of the total genetic effect in MS (120).

A question in this research area of HLA and MS has long been if DRB1*15 haplotype really is the causing determinant or are other genes in LD also of importance? The HLA class II area has an uncommonly high level of LD and it has not been possible to determine which part of the haplotype or if other genes in the vicinity also in LD could be of importance.
Recently Lincoln and colleagues published a report (121) where they genotyped 1185 Canadian and Finnish families with multiple sclerosis (n = 4203 individuals) with a high-density SNP panel spanning in the genes encoding the MHC and flanking genomic regions. Strong associations were observed with several blocks within the HLA class II genomic region but the strongest association was with HLA-DRB1. By conditioning on HLA-DRB1 they revealed that no additional block or SNP association was independent of the HLA class II genomic region. The authors claim that the study therefore indicates that MHC-associated susceptibility to multiple sclerosis is determined by HLA class II alleles, their interactions and closely neighboring variants. It should, however, be noted that they did not assess functional variants in this region (besides DRB1) and they applied a very stringent correction for multiple testing making it possible that functional variants as well as small effects could have been missed. They also had limited ability to exclude the effect of other genetic variant within approximately 100 kb from HLA-DRB1. This area includes HLA-DQB1. Admixture studies suggest that HLA-DRB1 (122) and HLA-DQB1 (123) may each have a role in distinct populations.

In the MHC area an independent effect of genes within the HLA class I region is also determining the risk of MS. This effect was most clearly shown by the HLA-A allele, HLA-A*0201 negatively associated to the risk of MS (34). Subsequently, Harbo and co-workers confirmed the importance of class I alleles in MS in the Nordic population (35). The biological function of HLA class I is primarily to present self antigens to $CD8^+$ T cells (124).

Before the completion of the human genome sequence project, the knowledge of both the physical position of genes as well as the information of their polymorphisms were quite scarce. In combination with the successful mapping of Mendelian disease genes using linkage analysis several groups started to use this strategy to find areas in the genome linked to MS.
In 1996 three linkage screens were reported (125-127) and a fourth followed the next year (128). These screens, however, all failed to identify a single locus with a significant statistic. A meta-analysis of raw data from the first three screens also failed to increase the power significantly (129), and even for the HLA locus the evidence for linkage was modest. When more screens became available, a new meta-analysis was performed based on the raw data from the screens of nine populations: American (127), Australian (130), British (125, 131), Canadian (126), Finnish (128), Italian (132), Sardinian (133), Scandinavian (134) and Turkish (135). This analysis proved the case of the HLA class II region with a significant linkage score (Sawcer, 2002) (136). For the rest of the genome, a few less significant peaks were found in particular the chromosomal areas 17q21 and 22q13.

In an additional attempt to exhaust the possibilities of linkage analysis in MS, Sawcer and colleagues (137) gathered, from Australian, Scandinavian, British and American collaborators, DNA samples from MS families (most of them had been included in the meta analysis described above). The resulting analysis contained data from 4506 markers and 2692 individuals and revealed highly significant linkage to the MHC on chromosome 6p21 and suggestive linkage on chromosomes 17q23 and 5q33. The authors conclude that the screen provides the most definitive linkage map currently available for MS. The result indicates that future studies attempting to identify genetic factors influencing the development of this disease will need to rely on association-based methods and must involve large patient cohorts.

What studies are performed in the field of association analysis apart from the MHC region? A large number of potential candidate genes have been investigated by association for their involvement in susceptibility for MS and also for their contribution to course and different phenotypes, reviewed in (5, 33, 138-140). The exact number of genes is impossible to state since many have been performed in concert with other diseases and presumably a large number of negative studies are not published but the number ought to be well over 200 genes.
Functional and positional candidates such as those coding for immunoglobulins, cytokines, chemokines, T-cell receptors, interleukins, interferons, HLA and myelin antigens have been investigated. A large number of studies have initially shown genes associated with MS but the result has not been possible to confirm in later studies.

These partly frustrating results may be due to genetic heterogeneity in which different genes predispose for the same phenotype, that is, MS, and that these genes may vary between families and populations. In addition, one problem is that each disease-predisposing gene probably contributes with only a small increase in the risk of developing MS. Therefore the studies as discussed above need to be powered for detecting very small contribution of each polymorphism. This has seldom been the case in the early association studies of MS. Other problems that can have occurred are discussed in the section of potential weaknesses of association studies. However, recently some interesting candidates, *NOS2A*, *TAC1*, *PRKCA*, *MHC2TA*, from reasonably well designed studies have emerged (141-144) but they have not yet passed the crucial step of confirmation.

Another finding worth mentioning is that the APOE4 allele of the *APOE* gene, also known for conferring risk for Alzheimers disease (145), has been shown to confer increased risk of severe MS and atrofi (146, 147) but conflicting result have also been obtained (148-150).

Hafler and coworkers recently published the first admixture scan in MS (151). Multiple sclerosis is described by the authors as an excellent candidate for admixture mapping since it is more prevalent in European Americans than in African Americans (152). They hypothesize that if there are genetic risk factors for MS explaining the epidemiology, they should be identifiable as regions with a high proportion of European ancestry in African Americans compared to the average. The analysis detected one area around the chromosome 1 centromere, containing 68 known genes, with increased risk due to European Ancestry. Only in one other disease, hypertension, an admixture mapping attempt has been published (104).

It is also worth mentioning that two genome wide screens for linkage have been performed in multigenerational pedigrees affected with MS (153, 154). Peaks with suggestive linkage were found but no further studies have revealed these areas as important for MS.

In conclusion MS genetics has a problematic past, but since the case for a genetic component in MS is solid it is likely that large-scale SNP genotyping efforts based on large clinical materials will in the end probably be successful in identifying risk or disease modifying genes. Reassuring positive examples from other complex traits includes the association of *CARD15* gene and Crohn's disease (95, 96) as well as the *PTPN22* gene association with rheumatoid arthritis (RA) (97).

1.5 GLOBAL GENE EXPRESSION PROFILING

Sequencing of the human genome (92, 93) enables a new scale in of research in biology. This progress made it possible to think about how to study biology in a more hypothesis generating way, for example how to study gene expression of all genes within an organism in parallel?

In 1996 Lockhart and colleagues published their first article about oligonucleotide arrays constructed to measure global gene expression (155). They stated that sequence information alone is insufficient for a full understanding of gene function, expression and regulation.

Because cellular processes are governed by the repertoire of expressed genes, and the levels and timing of expression, they deemed it important to have tools for the direct monitoring of large numbers of mRNAs in parallel. They had developed an approach based on hybridization to small, high-density arrays (usually called microarrays) containing tens of thousands of synthetic oligonucleotides synthesized *in situ* by photolithography and oligonucleotide chemistry.

Lockhart and colleagues were not the first to measure gene expression in parallel. Several techniques were already in use: Differential display (156), sequencing of cDNA libraries (157) , serial analysis of gene expression (SAGE) (157) , cDNA spotted on nylon membrane (158) or glass slides (159). The spotted cDNA methods also have the ability to measure expression from a large number of genes in parallel.

Since the introduction of these methods, a very fast development has been seen and global gene expression profiling (also referred to as transcriptomics) is today an important tool in medical science and basic biology research. Today we have platforms with the capability of measuring all known transcripts in a wide variety of species, the latest addition being oligonucleotide arrays interrogating different exons throughout the human genome enabling analysis of alternative splicing (www.affymetrix.com/products/arrays/exon_application.affx).
An ongoing discussion in this field has been if the results are reproducible across cDNA and oligonucleotide platforms. Two recent publications show that there are no major platform specific differences that mask underlying biological response (160, 161). Discrepancies in results in formerly obtained results can be due to a large number of things including problems with annotation. Can we be certain which genes are actually represented by the different probes? A recent publication draws attention to this problem; difficulties in annotation are most often due to cross-hybridization of several transcripts to one probe (162).
The golden standard in these experiments is still to confirm the regulations scored as significant by quantitative RT-PCR (reversed transcriptase polymerase chain reaction). This technique has a large dynamic range as well as high sensitivity and specificity.

Critical issues in gene expression profiling

The method is as described very powerful but there are some critical issues that have to be taken into consideration.

- *The design of the experiment* - With the large number of transcripts interrogated it is easy to get false positive results if experiments not are designed properly. A great deal of time needs to be spent on this part of the work. Using replicate samples is of utterly importance to reduce the number of false positives as well as providing enough statistical power. The method is comparable to its nature and the design needs to take this into consideration. It can be difficult to decide on what to compare with, for instance, one needs to define: what is normal?
- *The quality of RNA*- Since RNA is prone to degradation by RNases it is very important to take necessary precautions to keep the RNA as intact as possible. The different RNA populations that are compared should be prepared in the same way. The purity of the RNA after preparation is also of importance.
- *Performing the experiment* – The experiments being compared (for instance case compared to control) should preferentially be performed by the same trained person, with reagents from the same lot, as well as arrays from the same lot. This is to ensure that small differences are not introduced by differences in laboratory practice as well as differences of the used components.
- *The statistical analysis* – For ten years there has been a debate on this subject. Numerous approaches to the data analysis, both the raw data and analyses of global trends in the data, exist (163-168). It is crucial to be able to distinguish real changes in gene expression from normal variation and noise. Different strategies can be relevant depending on the experiment set up.
- *The cost* - The price is still quite high for this kind of experiments. If one can not afford a proper design and data analysis it is probably not worth pursuing the experiment – the resulting data will be too insecure. On the other hand, one should also consider that the cost per interrogated transcript is very low.

Some examples from the MS field

In the animal models set out to mimic MS or different features of the disease, gene expression profiling experiments have been performed in several studies. Preferentially brain or spinal cord tissue has been investigated in the animal models. Differences between animals afflicted with disease and unafflicted animals (169-174), as well as, response to different treatments (175-177) have been studied. The resulting expression profile after nerve injury has also been under investigation (178, 179). (This will be further discussed in the results section.)

In the human situation of MS a number of studies have been undertaken. Both brain tissue (173, 180-186) and peripheral blood lymphocytes have been the targeted (187-195). Also, in MS, disease versus control as well as responses to pharmaceutical treatments have been studied.

1.5.1 The Affymetrix® technology

The Affymetrix technology used in this thesis will briefly be described below.
The Affymetrix microarray consists of 25-mer oligonucleotides synthesized *in situ* on a silica surface by photolitographics at a defined location. For each transcript 11 to 20 oligonucleotides that are a perfect match of the interrogated transcript as well as mismatch control probes identical to the perfect match with the exception of single base difference in the middle of the oligonucleotide. In this way the background hybridization governed by the mismatch probe can be subtracted to reduce the background due to non specific hybridization. However, it is an ongoing discussion how appropriate the use of a mis- match probe is. Some data analysis software does not take the mismatch probes into account. Probes are selected with a bias towards the 3´-end of the gene to try to ensure that enough cRNA will be available for the hybridization. The RNA is prepared and used as template for reverse transcription using a poly-T primer with a T7 polymerase site incorporated. This reaction initiates from the 3'-end – thereof the importance of the probes being designed towards this end of the transcript. The resulting product is then in vitro transcribed in the presence of biotinylated oligonucleotides resulting in a labeled cRNA which is subsequently hybridized to the array. After washing of the array the hybridized RNA is stained with phycoerythrin-conjugated streptavidin which signal is detected by scanning with a laser scanner. One chip is used for each sample unless replicate samples are pooled.

2 AIMS OF PRESENT STUDIES

The overall aim of the present project was to advance our comprehension of the genetic contribution to multiple sclerosis and by that gaining a better understanding of the pathophysiology of the disease.

The specific aims for the different studies were the following:

Study I
To determine if genetic variations in two genes, located in the MHC area at 6p21.3, the *NOTCH4* gene, reported to be associated with schizophrenia, and the *TNF* gene, coding for the proinflammatory cytokine TNFα, influence the risk of being affected by MS.

Study II
To investigate the importance of genetic variations in 66 genes in the susceptibility to MS. Forty-five genes were selected on the basis of their location in regions linked to MS or other autoimmune diseases on chromosomes 5, 7, 12, 17 and 19, and the remaining 21 on the basis either of encoding proteins of potential functional importance in MS, or of having been the focus of previous association studies in MS.

Study III
To identify genes in the spinal cord, by expression profiling in two inbred rat strains, differentially expressed after nerve injury and to elucidate their differential regulation dependant on genetic background.

Study IV
To evaluate two gene loci, *CD74* on chromosome 5q33 and Metallothionein on chromosome 16q, for a possible importance in determining susceptibility to MS, based on their differential regulation over time and by genetic background, in the nerve injury model used in study III.

3 MATERIALS AND METHODS

For detailed delineation of materials and methods, please see the separate articles attached.

3.1 PATIENTS AND CONTROLS

All patients were diagnosed by neurologists according to the Poser or McDonald criteria at Karolinska University Hospital, Huddinge and Solna Sites. Study I consisted of 186 patients, study I of 672 patients and study IV of 890 patients.

The control group of Study I consisted of 88 employees of Karolinska University hospital and 104 healthy blood donors. In study II the first control group (used in the two screening stages) consisted of samples from a control group consisting of 288 blood donors and 384 controls randomly selected unrelated members of a set of mono- and dizygotic twin pairs enrolled in the Swedish Twin Registry. A second control set of 465 blood donors was used to evaluate four SNPs in the *HAVCR2* gene. In study IV the control group consisted of 775 blood donors.

All patients and controls were residing in the Stockholm area and originating from Sweden or other Nordic countries.

3.2 GENETIC ANALYSIS

3.2.1 DNA extraction

Genomic DNA was extracted from leukocytes using a modified salting out method (196), QiaAmp DNA Blood Maxi kit (Qiagen Gmbh Germany) or PureGene (Gentra Systems USA).

3.2.2 Genetic markers

Microsatellite
For study I a CTG repeat located 15 bp downstream of the transcription start in exon 1, coding for the signal peptide of the protein, of the *NOTCH4* gene was selected due to previously shown association of this genetic variant to schizophrenia (197).

Single nucleotide polymorphisms (SNPs)
In study I, two SNPs were selected in the promoter region of *NOTCH4* due to previously shown association of this genetic variant to schizophrenia. For *TNF* two SNPs were selected, both previously studied for association to MS (198-203). The SNPs are situated in the promoter region of the gene.

For study II, in total 123 SNPs were analyzed in 66 genes. Of the 66 candidate genes, 34 were genotyped for a single SNP (on account of the scarcity of known SNPs in or near the genes) and 32 for two or more SNPs. In the first stage, we found an association, at the 8% significance level, between SNPs in 22 genes and susceptibility to MS. In order to confirm these results, and to increase the power of detection, a

second stage of genotyping was performed on these 22 genes using a larger number of patient and control samples; further, several additional SNPs, located in or around the associated genes, were included in the analysis at this stage. Information regarding previously characterized SNPs genotyped in this study was accessed via public (The SNP Consortium [http://snp.cshl.org/],

dbSNP [http://www.ncbi.nlm.nih.gov/]) and proprietary (AstraZeneca) databases. In study IV ten SNPs in the gene area coding for CD74 and 12 SNPs from the Metallothionein (MT) cluster area were initially selected from dbSNP. Priority was given for markers that had been genotyped by the Hapmap consortium in the CEPH sample (Utah residents with ancestry from northern and western Europe) and showed a relatively high heterozygosity. Six SNPs were excluded after a first genotyping in 368 of the controls, based on that the genotyping assay did not work or that the markers were shown to be monomorphic or that they deviated from Hardy-Weinberg equilibrium in our material.

SNP discovery

In light of an initial finding in study II indicating association of one SNP located in *LAG3* with MS, all coding sequences of the gene, as well as its promoter and 5'- and 3'-untranslated regions, were amplified in 96 subjects. Denaturing high-performance liquid chromatography (DHPLC) was then performed using the Transgenomic WAVE System (Transgenomic, Omaha, Neb, USA). PCR products were separated on a reverse-phase column (DNASep; Transgenomic). Samples detected by DHPLC as being heterozygous were then sequenced using ABI PRISM Big Dye Terminator (Applied Biosystems, Foster City, CA, USA), and the sequencing products analyzed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Two novel SNPs, located in noncoding sequences of LAG3, were discovered in this manner; prior to the submission of study II, these SNPs were registered by others in dbSNP under the identification numbers rs2365095 and rs7488113.

3.2.3 Genotyping

Microsatellite

In study I the CTG repeat located + 15 bp from transcription start of NOTCH4 was amplified by PCR using fluorescently labeled primers. The resulting products were analyzed by electrophoresis on an ABI 377 DNA sequencer in the presence of a size standard (Genescan 100 TAMRA). The results were analyzed using GENESCAN/GENOTYPER software (version 1.1, Applied Biosystems)

Single nucleotide polymorphisms (SNPs)

In study I the two SNPs in the 5'- untranslated region of *NOTCH4* were amplified by allele specific PCR. For the two *TNF* promoter SNPs the method of 5' nuclease (TaqMan ®) technique was used for genotyping. Two probes, complementary to the different alleles, were used in the same PCR reaction (24). The primers and the probes were designed using Primer Express™ program (Applied Biosystems). The alleles for

the *TNF* SNPs were analyzed using the graphical view from the Sequencing Detection System (Applied Biosystems).The product sequences for the four different SNP products in study I, were verified by sequencing using ABI PRISM Big Dye terminator (Applied Biosystems). Controls for the different genotypes were included in all genotyping experiments.

In study II all genotyping was performed at the AstraZeneca laboratory by the Pyrosequencing (204) method according to the standard protocol provided by the manufacturer (Pyrosequencing AB, Uppsala, Sweden). Sequencing and PCR primers were designed using Oligo 5.0 software. Controls for the different genotypes were included in all genotyping experiments.

The SNP genotyping in Study IV was performed by the Mutation Analysis Faciltiy MAF core facility at Karolinska Institutet with matrix-assisted laser desorption/ ionization time-of-flight (MALDI-TOF) mass spectrometry (SEQUENOM Inc., San Diego, California) of allele-specific primer extension products. Multiplex SNP assays were designed using SpectroDESIGNER software (Sequenom Inc., San Diego, California). All reactions were run under the same conditions. Allele-specific primer extensions were conducted using an extension primer and the Mass EXTEND Reagents Kit. Primer extension products were analyzed using a MassARRAY mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). The resulting mass spectra were processed and analyzed for peak identification using the SpectroTYPER RT 2.0 software (Sequenom). A subset of DNA samples occurred more than once in the genotyping set-up to ensure the correctness of genotyping.

HLA–typing
In study I low resolution genotyping was performed for HLA-A and HLA-DRB1 genes. The resolution corresponded to standard serologic typing. PCR amplification with sequence-specific primers (SSP) (Olerup SSP^{TM} Saltsjöbaden, Sweden) was used. High-resolution typing was performed for individuals carrying DRB1*15, A*02, or A*03 alleles which allowed the identification of the allelic series DRB1*1501-DRB1*1511, A*0201-*0251, and A*0301-A*0308.

3.2.4 Statistical analysis

Single point analysis
In study I, II and IV Hardy-Weinberg calculations were performed to ensure that the allele frequencies in the control population were in Hardy –Weinberg equilibrium.
In study I and IV, P-values for two-by-two contingency tables comparing carriage counts in patients and controls were calculated with Fischer´s exact test for a two sided P-value (GraphPad Instat) and were not subjected to correction.
In study II P-values for two-by-two contingency tables comparing carriage counts in patients and controls were calculated by a χ^2 test and were not corrected.

In study II and IV, single-point power analysis was performed for detecting association in the SNPs, using a standard method whereby normality in the logarithm of the OR is assumed; power was calculated assuming a two-sided significance level of 0.05 and one-sided power for detecting an OR of at least 1.5 (or 0.67), in the absence of an adjustment for multiple testing.
In study I a Monte Carlo test for association between disease and microsatellite alleles was performed using the CLUMP program (205). To study the LD between loci and markers, and to analyze the interdependence of associations, the two-locus method proposed by Svejgaard and Ryder was applied (206).

Haplotype analysis
In study II pairwise measures of LD were obtained using EH (207). These measures were used to define haplotype blocks of closed gene-centric LD with |D' | values greater than 0.85. Markers within the blocks were then used to estimate haplotype frequencies using the SNPHAP program (http://www-gene.cimr.cam.ac.uk/clayton/software/snphap.txt). Haplotype frequencies were tested for significance using the T1 statistic of CLUMP with 10 000 simulations.

The haplotype analysis for study IV was performed in the following way: LD, haplotype block and haplotype frequencies were estimated using Haploview version 3.2 (http://www.broad.mit.edu/mpg/haploview/download.php) Haplotype association was tested by performing 1000 Monte-Carlo simulations using the Matlab Software (TheMathworks, Inc., Natick, USA).

OR calculation
Odds ratio for single point SNPs were calculated with the Instat (GraphPad Instat) software that uses the approximation of Woolf for the 95% confidence interval (CI). This was performed in study I, II and IV.

3.3 GENE EXPRESSION PROFILING

This whole section relates to methods used in study III.

3.3.1 Animals

Fourty-five DA($RT1^{av1}$) and 45 PVG(RT^{1c}) adult male rats (Scanbur BK AB, Sollentuna, Sweden) were used for the Affymetrix GeneChip expression analysis (n=36) and RT-PCR validations (n=9). The advanced intercross line (AIL) used was founded by reciprocal crossing of DA (RT1av1) and PVG.1AV1 (Harlan UK Ltd, Blackthorn, UK) rats. This AIL cohort has previously been used for genetic mapping of nerve lesion-induced MHC class II expression (208). Animals were kept in a barrier animal facility under specific pathogen-free and climate-controlled conditions with 12 h light/dark cycles, housed in polystyrene cages containing wood shavings, and fed standard rodent chow and water ad libitum. All experiments in this study were approved and performed in accordance with the guidelines from the Swedish National

Board for Laboratory Animals and the European Community Council Directive (86/609/EEC).

3.3.2 Nerve lesion model

Animals were subjected to unilateral avulsion of the left L3-L5 ventral roots under standardized conditions and in deep isoflurane anesthesia at an age of 8-10 weeks with a post-operative survival time of 5 or 14 days (+/- 2 hours). Rats of the two strains were handled alternately throughout the experiment. The meninges were removed and the ipsilateral L4 ventral quadrant dissected by dividing the cord vertically and horizontally at the level of the central canal. In DAxPVG intercross animals, the ipsilateral L3 ventral quadrant was dissected and the L4-L5 segments kept intact for histological analysis. The tissue was subsequently snap frozen and stored at -80°C until further use.

3.3.3 Nerve cell counts

Nerve cell counts were performed as described previously (209). Motoneurons were counted in cresyl-violet stained 14 μm section from the L4 segment (n=15 from each rat). Cell counts were corrected for atrophy according to the Abercormbie formula (210). The degree of neuronal survival is presented as a mean ratio between the number of motoneurons on the ipsilateral and contralateral sides.

3.3.4 cRNA preparation and gene chip hybridization

Total RNA from the ipsilateral L4 spinal cord segment of naïve and operated rats was extracted using TRIzol (GIBCO/BRL Life technologies, Rockville, MD) according to the supplier´s instructions with the inclusion of Pellet paint (Novagen, Madison, WI) as co-precipitant. The RNA was purified with RNeasy Mini kit (Qiagen, Hilden, Germany). The subsequent RNA processing procedures followed protocols in the Expression Analysis Technical Manual (Affymetrix, Santa Clara, CA). Fragmented cRNA, control cRNA and grid alignment oligonucleotide were added to the chip and hybridized for 16 hours to a test array and subsequently 16h to the RG-U34A array (Affymetrix). In total 18 arrays were used. The arrays were stained with streptavidin-phycoerythrin (SAPE, Molecular Probes), signals amplified with goat biotinylated anti-streptavidin (Vector Laboratories, Burlingame, CA), and finally stained with SAPE using the Affymetrix fluidics station 400. Fluorescent signals were measured with a Gene array scanner (Hewlett Packard Corporation, Palo Alto, CA).

3.3.5 Data processing and statistical analysis

Microarray data was analyzed by the Microarray Suite software version 5.0 (MAS 5.0, Affymetrix). The signal intensity of all probes on each chip was scaled to a single value, 100, to allow for comparisons between samples. The quality of array hybridization was assessed by the ratio between probe sets located in the 3´and 5´end for β-actin and glyceraldehyde-3-phosphate dehydrogenase. Pairwise comparisons were made between all chip pairs in the two groups, resulting in 9 comparisons (3 by 3 chips.) A probeset was defined as regulated when fulfilling the following criteria: Signal detected as present (“P”) in 2/3 test replicates for “increased”, or baseline replicates for “decreased”, change (“increase”/”decrease”) in (≥7/9) of the chip comparisons, and mean signal log ratio (MSLR) >0.5 or <-0.5.

Principal component analysis (PCA) (211) was performed on the entire set of data (mean signal values), using DecisionSite for Functional Genomics (Spotfire AB, Gothenburg, Sweden.

Classification of genes was based on information in the Ensembl database (http://www.ensembl.org), the NCBI Entrez Gene (http:// www .ncbi. nlm. nih. gov/ entrez) and the RGD (http://rgd.mcw.edu). Information on homologous genes in human and mouse (NCBI) was also used when available. Genes were classified according to their most informative function and/or localization under the conditions studied.

3.3.6 Quantitative RT-PCR

Reverse transcription of the RNA was performed with random hexamer primers (Gibco BRL) and Superscript Reverse Transcriptase (Gibco BRL). Amplification was performed on an ABI PRISM 7700 Sequence Detection System (Perkin Elmer, Norwalk, CT). Primers were designed with the Primer Express software. All RT-PCR runs were conducted using Qiagen QuantiTect™ SYBR® green (Hilden, Germany) according to the manufacturer's instructions (Perkin Elmer). Primer specificity was assessed by analyzing amplicon dissociation curves in each sample. Relative amounts of mRNA levels were calculated using the standard curve method. The mRNA level in each sample was calculated as the ratio between the relative amount of the specific target and the relative amount of the endogenous control, GAPDH or 18S mRNA.

3.3.7 Immunohistochemistry

The immunohistochemistry method that was used has been described previously (212, 213). Cryosections (14 µm) were cut at the level of the L4-L5 segment of the spinal cord. The specificity of the immunostainings was tested in control slides by omission of the primary antibody, incubation with pre-immune serum or incubation with unrelated isotype-matched antibody controls.

4 RESULTS AND INTERPRETATIONS

4.1 STUDY I

The association of MS with the MHC class II region has been known for a long time. But it has been impossible, due to extensive linkage disequilibrium, to determine if the causal variant is the suspected *HLA-DRB1* gene. Increased sharing of the HLA-DRB genes in MS affected sibling pairs even in absence of the known associated DRB1*15 haplotype had been reported (120) shortly before the embarking on this study. This finding suggests that the association to the *HLA-DRB1* gene could also be due to other alleles at the DRB1 locus and/or other genetic variations in LD with *HLA-DRB1*. An interesting report showed that polymorphisms within and in the vicinity of the *NOTCH4* gene were strongly associated to schizophrenia (197). This led us to select these polymorphisms for studying in MS. A recent meta analysis (214) of 9 linkage and association based studies however could only conclude that some of the polymorphism may influence risk for schizophrenia and only under certain circumstances. The Notch4 protein functions as a receptor for membrane-bound ligands Jagged1, Jagged2 and Delta1 to regulate cell-fate determination. The protein affects the implementation of differentiation, proliferation and apoptotic programs. Another interesting candidate gene located in the MHC region, *TNF*, that previously had been studied for association with the risk for being affected with MS, with conflicting results (198-203), was also chosen as candidate. The gene encodes the proinflammatory cytokine TNFα that has been implicated in a wide variety of diseases including MS. TNFα is present in MS-lesions, and elevated levels are found in cerebrospinal fluid and serum of MS patients and may indirectly injure oligodendrocytes, induce demyelination, and stimulate astrocyte proliferation as well as induce endothelial cell adhesion molecule expression to facilitate the entry of leukocytes into tissues (215).

In the study material, 186 MS patients and 192 controls, we did not find any association of the TNF polymorphisms and neither with the microsatellite studied for *NOTCH4*. We found a difference in carriage count for the SNP studied in the *NOTCH4* promoter but after conditioning on HLA-DRB1*15 , as described by Svejgaard (206), it was shown that the association was secondary to HLA-DRB1*15.

After submission for publication we extended the study to gain a greater power, in total 475 patients and 470 controls, for the SNP with the highest minor allele frequency used in Study I for the TNF locus. The total material did not show any association either.

After this attempt to study the question of the causative variant in the MHC region we realized that to perform this in an appropriate manner one would need a large material as well as a large number of markers. Typing of functional units would also have been needed to get closer to resolving the puzzle.

Recently (October 2005) Lincoln and coworkers published a paper in which they used a SNP panel of 1068 SNPs covering a 12MB region at chromosome 6 spanning the entire MHC region along with flanking sequences (121). In total they analyzed 4203

individuals. Even with this highly ambitious design it was impossible to conclude if HLA- DRB1 was the sole determinant.

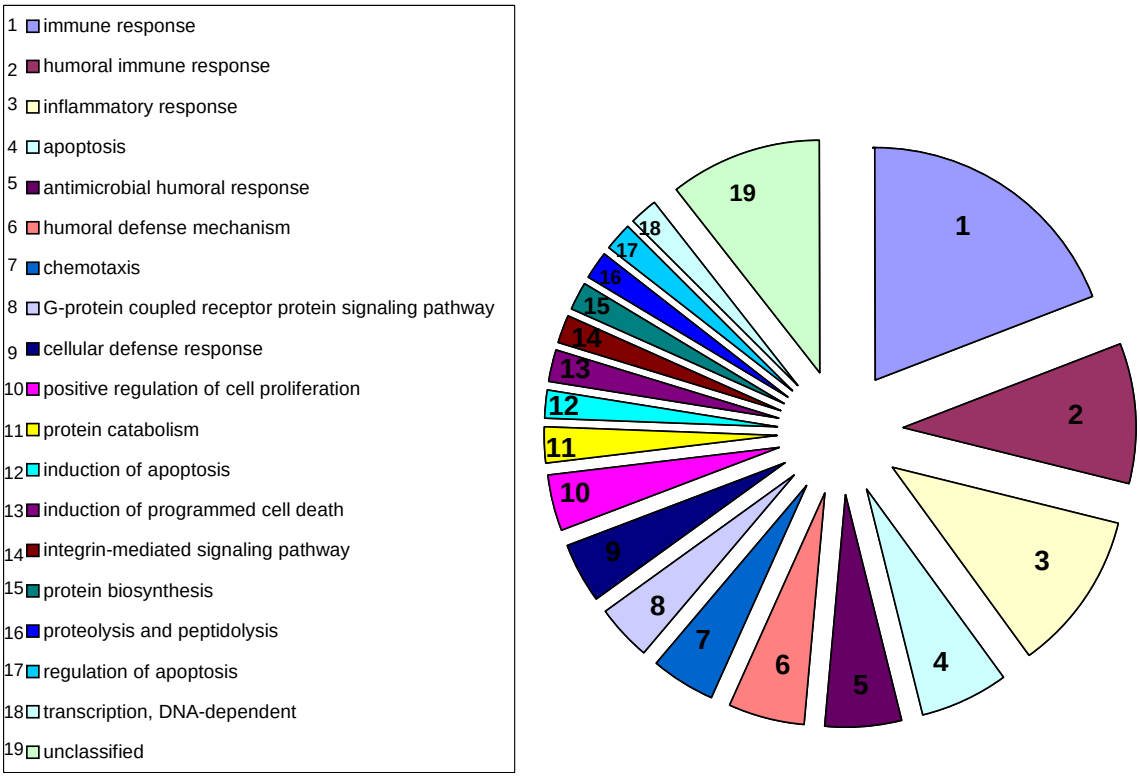
To further shed light on this question one would need to type either known functional units and/or re-sequence the region closest to DRB1 in a large cohort. Further functional studies of the different MHC determinants would be helpful. Functional studies have been performed regarding the DRB1*1501 effect, for example the study of functional and structural basis of TCR crossreactivity with DRB1*1501 restricted MBP and DRB5*0101 restricted Epstein Barr virus (EBV) peptide (216).

4.2 STUDY II

A large (at that time) collaborative study with a pharmaceutical company was started in 1999 and resulted in the publication of study II in year 2005.

With a two stage design of the analysis, we determined genotypes, in up to 672 MS patients and 672 controls, for 123 single-nucleotide polymorphisms (SNPs) in 66 genes. The genes were selected on the basis of suggestive susceptibility regions for MS and other autoimmune diseases at chromosome 5, 7, 12, 17 and 19, (136, 217, 218) and known or presumed biological functions considered to be important in MS. Twenty-one genes were selected solely on the basis of their biological function as well as previous association studies performed in MS cohorts (219-223). See Figure 2 for the biological processes these genes are involved in:

Figure 2. Gene ontology: biological processes for the 66 genes in Study II. Note that a gene can belong to more than one ontology.



In the first stage, we found an association, at the 8% significance level, between SNPs in 22 genes and susceptibility to MS. In order to confirm these results, and to increase the power of detection, a second stage of genotyping was performed for these 22 genes using a larger number of patient and control samples; further, several additional SNPs, located in or around the associated genes, were included in the analysis at this stage.

After genotyping in stage two, two genes- each containing at least three significantly ($P<0.05$) associated SNPs- conferred susceptibility to MS: *LAG3* on chromosome 12p13, and *IL7R* on 5p13. For *IL7R* the haplotype distribution also differed between MS cases and controls. In the haplotype analysis we also identified one predisposing and one protective haplotype. Interestingly, these two haplotypes differ at only one position—at rs987106, in intron 6 of IL7R— suggesting that a functional consequence of this particular polymorphism may underlie the haplotypes' positive and negative associations with the risk of MS.

LAG3 (CD223) is an MHC class II ligand evolutionarily related to CD4; it downregulates the activated T cells on which it is expressed through a high affinity interaction with its receptor that blocks the binding of CD4.

The interleukin 7 receptor (IL-7R; CD127), a member of the hematopoietin receptor family, is a type 1 membrane glycoprotein capable of binding alpha-helical cytokines (224).
Signaling via IL-7R induces somatic recombination of the T-cell-receptor and immunoglobulin genes, promoting the proliferation and survival of T and B lymphocytes (225). The IL-7R complex consists of the IL-7R alpha chain and the common cytokine-receptor gamma chain (CD132); the former molecule transduces trans-membrane signals through the recruitment of intracellular messengers to its cytoplasmic tail, while the latter activates this transduction (224, 226). IL-7Rα is expressed on immature B cells and T cells; in mice, administration of neutralizing antibodies or genetic ablation of IL7-Rα blocks lymphocyte development (225). In humans, germline mutations resulting in defective expression of IL-7R give rise to a subtype of the disorder severe combined immunodeficiency.

In analyzing the results from this study it is interesting to note the need for studies in large cohorts in a complex disease as MS. In the first stage of the analysis 22 genes showed association to the risk of developing MS while after the second stage, with more SNPs as well as further patients and controls added, only two genes were found to be significant. This is in agreement with the findings from Ioinnadis and collegues (227), who analyzed results from 55 meta-analyses comprising of 579 studies. They found that small initial studies rarely find the correct result.

Since quite few makers in each gene were studied and especially for those studied in only the first stage we are not able to conclude that genetic variations within these genes not are of importance.

To further corroborate these data other researchers in our group have studied a large Nordic cohort consisting of 1277 MS patients and 1396 controls. In this cohort they

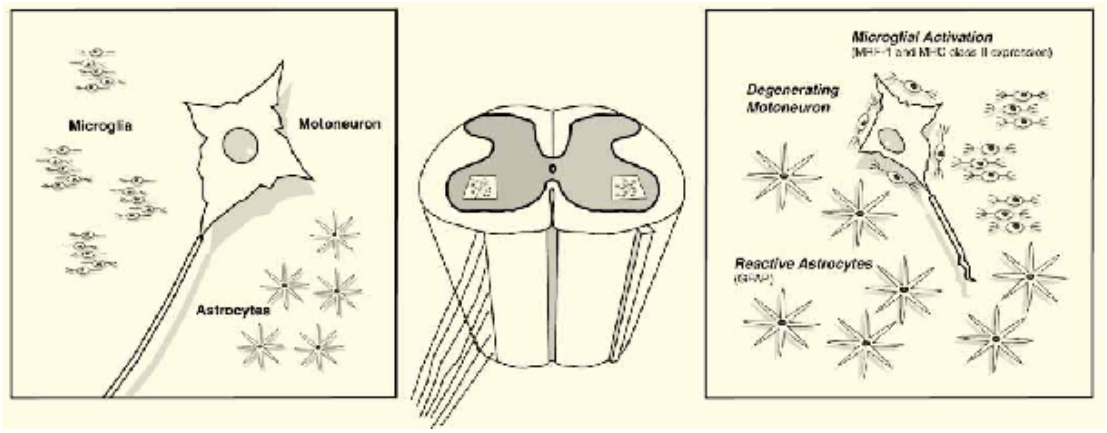
could confirm the association to *IL7R* but not to *LAG3*. This also shows the necessity to confirm findings in another dataset. The question here is if we can consider this a "candidate gene that has won the election" or if even further confirmation will be needed.

Interestingly other data have emerged that make IL7Rα of interest in MS research.
Gene expression profiling has shown significantly higher expression in peripheral-blood mononuclear cells in MS patients compared to controls (189). In another gene expression profiling experiment, IL7R was one of the transcripts, of a total of 53 genes, useful for deducing if the sample corresponded to an MS patient or a healthy donor (191). A recent gene expression study showed that IL7R was underexpressed in PPMS and upregulated in SPMS compared to healthy controls (228). The authors of this study also examined the genetic differences of a promoter polymorphism in the IL7R receptor and found an undertransmission of one allele in a trios material of PPMS. This was also confirmed in a material of PPMS cases and healthy controls.

4.3 STUDY III

Inflammation and neurodegeneration characterize MS. In this study we focused on the gene expression response after neurodegeneration. Theoretically, factors of relevance for nerve cell vulnerability may vary between different individuals depending on their genetic background. If so, identification of genetically encoded differences associated with nerve cell survival may be used to unravel critical pathogenic pathways. The primary signal mediating the reaction must originate from the injured cell but the exact mechanism remains unknown.

Figure 3. Schematic illustration of the VRA injury in the spinal cord. To the left a drawing of naïve tissue and to the right a substantial motor neuron loss is seen after lesion. Microglial and astrocytes are activated and some T-cells are also attracted to the lesion.



In this study we measured the gene expression profile after nerve trauma in the ventral root avulsion model in rat (Figure 3). We used two different strains known to have a different vulnerability to the injury.

The design of the experiment was to use two different time points after nerve lesion; the first before onset of nerve cell loss and the second at the midst of the neurodegenerative process. For each time point, naïve, 5 days, 14 days, three pools of RNA from four animals were used. This design was used in order to have enough biological replicate to be able to dissect the true difference in gene expression from biological variation and random noise. We used the Affymetrix genechip RG-U34A representing approximately 7000 full-length rat sequences and 1000 EST clusters.

Transcripts differentially regulated after VRA (the VRA response) were defined by comparing the groups from 5 and 14 days after VRA within each strain with the respective control group, whereas transcripts differentially expressed between the two strains (the strain response) were defined by comparing the strains at control, 5 and 14 days.
We found 278 annotated genes that were regulated after injury (5 and/or 14 days) and 68 genes that were regulated by strain. Of these were 45 regulated both by VRA and strain. To further dissect the different profiles the regulated genes were classified into 12 groups based on biological function and cellular localization. This approach revealed that a large number of genes upregulated after VRA were found in the immunology and inflammation group of genes, this group also showed the largest strain dependant response. The DA strain being more vulnerable to nerve injury had a more abundant expression of genes involved in the inflammatory response. The fact that many of these genes were upregulated early after injury before no significant loss of nerve cells had occurred argues against an increased inflammatory activation in the DA strain only as a consequence of increased neuronal degeneration. Instead, the extent and nature of the inflammatory reactions could determine nerve cell survival.
In contrast, enzymes, cell cycle proteins and some growth factors were expressed at higher levels in the more resistant PVG strain.

PCA was used to analyze mean signal values from the three replicate microarray chips for all probe sets in order to identify major gene expression patterns in the two strains after VRA. This statistical method reduces dimensionality in a multivariate dataset and the outcome defines theoretical independent axes ordered so that the first explains most of the variability. The first principal component (PC1) accounted for 56.2% of the variability, primarily segregating naïve versus injured animals. The second most prominent profile (PC2) in the data explained 28.1% of the variability and segregated groups primarily by strain, i.e. difference in genetic background, Figure 4.

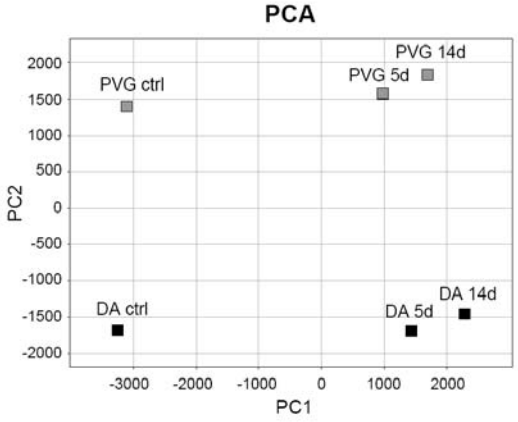


Figure. 4 Graph demonstrating the results from principal component analysis (PCA). Mean signal values, were used in the calculations. The first two principal components are displayed.

To validate the microarray data quantitative RT-PCR was performed on a subset of regulated genes. The mRNA levels of C3, C1q, Aif1, and Mt1a and Timp1 were analyzed in DA and PVG spinal cord tissue from controls, 5 and 14 days after VRA. The analyzed cohort of animals was not used in microarray experiment thus giving the validation even more strength. In all cases the regulation was very similar to that judged by the microarray experiment.

Studies on the protein level were performed by immunohistochemistry. C3, MHCII and Timp1 were analyzed on spinal cord sections 14 days after nerve lesion. The protein level for the first two were markedly upregulated in the DA strain in comparison to the PVG strain while for Timp1 no significant differences could be found.

In order to gain further knowledge on some of the interesting findings the complement gene C1qb and the Timp1, gene expression was analyzed in a set of heterogenous animals of an advanced intercross line (AIL) originally established by intercrossing DA and PVG. These animals show differences in neurodegeneration and it was shown that the expression of these genes was negatively correlated with neuronal survival. In other words, a high level of these transcripts gives a more pronounced neuronal celldeath.

During the work with study III, a publication from Velardo and coworkers (178) reported transcriptional profiling in a spinal cord contusion model performed in Sprague Dawley (SD) and Athymic Nude (AN) rats. Only minor strain differences were found (2% of the regulated genes), which contrasts to 24% of the regulated genes showing strain differences in the study III. A likely explanation is that the SD/AN strain combination is less genetically heterogeneous than DA/PVG, however, an exact comparison is not possible to make, since SD is an outbred strain.

The results from study III suggest a link between the inflammatory response elicited by nerve injury and subsequent neurodegeneration. The exact definition of the genetic variants underlying differences in gene expression pattern can shed light on disease mechanisms in neurological conditions where neurons are damaged and die.

4.4 STUDY IV

In this work we were guided by the findings from study III selecting candidate genes for a large case control study consisting of 890 MS patients and 775 controls. We aimed for designing a study of as high power as possible within our patient population. We focused on genes that were regulated both by time and strain after the nerve injury hypothesizing that expression influenced by the genetic background in the rat model also could show genetic variations between MS cases and healthy controls. A potential weakness of this strategy is that we cannot with certainty know that gene expression of any gene is regulated by variations in or close to the gene. Another strategy that could have been applied is to first define the QTL in the rat and study the genes homologous to the found rat locus in the human situation. This strategy will be applied for selected genes found in study III.
To further refine the selection of candidate genes other available biological information, especially linked to MS or EAE, was used.

This resulted in the selection of two gene loci, CD74 (also known as invariant chain) on chromosome 5q33 and Metallothionein (MT) on chromosome 16q, for their potential importance in determining susceptibility to MS.

The CD74 protein plays a critical role in the regulation of antigen presentation by MHC class II, (reviewed in (229)). It has been demonstrated that this molecule is an absolute requirement for a functioning class II-restricted Ag processing pathway in the CNS for the initiation of EAE (230). In the light of the knowledge on MHC class II and MS this was an appealing candidate gene.

Metallothioneins (MT) are heavy metal binding proteins that are transcriptionally regulated both by heavy metals and glucocorticoids. Human metallothioneins I-IV are located in a cluster at chromosome 16q. They are found throughout the mammalian body, including the CNS, MT-I and MT-II are rapidly induced following many types of CNS insults, and are strongly neuroprotective, reviewed by West et al (231).
MT-I and MT-II are induced in the brain of mice afflicted with EAE (232) and MT-I+II knockouts are more susceptible to and suffer from a more severe EAE than wildtype controls (233). A small study by Penkowa et al showed that MT-I and MT-II expression was significantly increased in astrocytes and activated machrophages /microglia in MS lesions (234).

In rats and mice, only one gene encoding for MT-I is present while in human at least 10 different *MT1* genes are found (Entrez gene, www.ncbi.nlm.nih.gov/entrez). Some of these genes are probably pseudogenes while some are expressed as functional transcripts (235). The *MT1* genes span an area of 58 kilobases at chromosome 16q13. The current study focused on *MT1E, K, J* and *A*.

Ten SNPs in the gene area coding for CD74 and 12 SNPs from the MT cluster area were initially selected from dbSNP. Six SNPs were excluded (three from each area) after a first genotyping in 368 controls. The reasons for exclusion were either that the genotyping assay did not work or that the markers were shown to be monomorphic or that they deviated from Hardy-Weinberg equilibrium in our control material.

The remaining 7 SNPs for the CD74 gene and 9 SNPs in the MT cluster were analyzed. Three SNPs belonging to the same linkage disequilibrium (LD) block in the MT region were found to be associated with the risk of MS ($P<0.05$, OR:1.24-1.47) as well as one haplotype ($P=0.018$) consisting of the three associated SNPs and two additional SNPs. One SNP in CD74 showed borderline signs of association, while the other SNPs in the same LD block did not show any association.

While writing this thesis summary genotyping of the associated SNPs from the *MT* area was performed in the large material from our Nordic collaborators (the same that was used in study I) were obtained but failed to confirm our initial finding of an association in our Swedish material. This seems to diminish the likelihood that the MT genes influence the susceptibility to MS.

5 GENERAL DISCUSSION

The development of the field of complex genetics, including MS genetics, has been quite rapid during the years, 2000-2005, which I have spent on my thesis work. In year 2000 there were several quite small linkage studies under way. Today, few would advise to embark on a linkage study without having a very large material, since it is increasingly recognized that linkage may be less effective at identifying common variants with modest effects.
The latest published linkage analysis in MS (137) analyzed over 2600 individuals and the authors conclude that future studies attempting to identify susceptibility to MS will need to rely on association based methods of large cohorts.

When I began my work one still believed that linkage would be the best solution of the problem. Association studies had been hampered by small sample sizes giving inconsistent results so they were "accused" of being unreliable. But also in the field of association analysis there has been a considerable change over the years. Our own approaches in the different studies reflect this quite well. In the candidate gene study of *NOTCH4* and *TNF* gene the investigation was only powered to detect a quite large genetic impact and also only a few selected SNPs were studied. In the second study of 66 different genes we applied a two stage design with quite few SNPs per gene in a first small cohort; if we detected any association we expanded both the number of polymorphisms as well as the cohort size. This study also contained haplotype block analysis. Afterwards our group also performed confirmation studies in an even larger independent material.
For the last association analysis we directly genotyped a quite large number of SNPs covering as well as possible, with the help of information from public resources, the areas we had selected. Also in this study a haplotype analysis was a part of the analysis. A confirmation study was performed before we even started to discuss performing more studies of the Metallothionein genes guided by the initial finding of an association.

So our studies have reflected the general trend towards increasingly larger studies, with more subjects as well as markers, and addition of an independent cohort for confirmation studies. Also the use of haplotype block analyses have been incorporated.

The usefulness of the haplotype approach is still an issue of debate – what extra do we gain from this approach? The hypothesis is that the haplotype blocks will capture most of the genetic variation in the region and therefore one would be able to localize haplotype tagging SNPs that will be informative for the majority of the found haplotypes, and reduce the SNPs needed to be genotyped for the analysis.

In our case a finding of a five-SNP LD block in the IL7R gene, with one predisposing and one protective haplotype differing at only one position led us to suggest that a functional consequence of this particular polymorphism may underlie the haplotypes' positive and negative associations with the risk of MS.

Our basic strategy for the association analyses has been study built on the 'common disease- common variant hypothesis' (CVCD). This hypothesis states that common genetic variations are involved as major players in contributing to common disease. Support for this hypothesis is reviewed in (236) but there are also major concerns that this not will be valid for a number of complex diseases. If the CDCV-hypothesis does not hold true and very rare alleles are causative it will remain difficult to test the gene sequence variations and their involvement in the causation of common disease. Very large studies that for the moment are virtually impossible to perform would be needed. Even if the CDCV-hypothesis is correct one has to consider that the scale of analyses needed could be quite large.

It has been calculated that, in a worst case scenario, 160 000 cases and controls would be needed to have power to detect a genetic effect with an odds ratio (OR) of 1.3, if you are studying markers that are only in LD and not causative by themselves, and if there is a large difference in allele frequencies between the causing variant and the marker, (this example was based on a p-value of 0.001, a D´ between the markers of 0.7, a disease allele frequency of 0.05 and a marker allele frequency 0.5) (108).
But even the best scenario, calculated in the same article, would need 2100 patients if one targets a P-value of 0.001.

Another issue worth considering is the correction for multiple testing. If you are performing a large number of statistical tests, a few will be significant purely by chance alone. The issue is how to get rid of false positives but avoid false negative results. Our approach to this problem has been to use uncorrected p-values but instead use a replication step – either in the same population or in another population. If the *IL7R* association is a true finding it is interesting to note that if we had applied even a very liberal correction in our first stage of Study II this finding would have been regarded as false positive. On the other hand in the case of *LAG3* and *MT-I* genes a correction of P-values had saved us the effort to repeat the analysis the Nordic material.

To elucidate if IL7R is important in the pathogenesis of MS we will embark on functional studies of this gene. Interestingly a number of publications already (as discussed in the result section) describe findings that make this gene and its product even more valid for further study. This is probably a reasonable approach when working with a candidate gene thought to contribute to the susceptibility of a disease.

As a last comment to the genetic discussion I would like to add that in the future we will probably need to analyze interactions between genetic variations in a more exhaustive way than today. It seems very likely that genetic variations interplay in giving their effect and this needs more attention and development of new statistical methods.

Also in the field of gene expression the development has been very fast – due both to the sequencing of the human genome as well the establishment of new powerful techniques. I did my first gene expression analyses during my degree project in 1989. At that time I was very satisfied if, by Northern analysis, the expression level of a few genes could be established during the course of a week. If anyone then would have told

me that in 15 years time you will be able to measure the expression level of *all* genes in a genome overnight – I would have thought that they were describing a utopia.

The global gene expression field has also developed rapidly since it was first introduced. In the beginning it was hampered by the lack of knowledge that even very small variations in the laboratory routines could introduce bias. It was not fully appreciated that the use of enough biological replicates are of utmost importance. Also the statistical analysis has been developed making it possible to have both a gene based view on expression but also a more global view on the regulation of pathways. More development will certainly be seen in this area.

A key factor of success for gene expression analysis is to be at the right place at the right time since what we study is an instant view on the expression. It is also worth noticing the well known fact, that a change in gene expression is not always accompanied with the same change in the proteins produced. Since the effect on the cell is largely exerted by the protein profile, the levels of the encoded proteins as well as the posttranslational modifications also need to be taken into consideration.

In this thesis we made use of the global gene expression in studying neurodegeneration in a rat model. Neurodegeneration is one feature of MS and the knowledge of this process in different rat strains with different vulnerability to nerve injury with the hypothesis that underlying genetic differences will be of importance. As always it is of course not easy to judge how much that can be concluded from an animal model and translated to human disease.

6 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

We have performed studies of the influence on multiple sclerosis by selected genetic variations.

Evidence was found for involvement of variations in the *IL7R*, *LAG3* and *MT1* genes in a material of Swedish patients with MS compared to a control material. The finding of *IL7R* has subsequently found further support in an independent material (to be published)

By the use of global gene expression profiling we have revealed that the inflammatory response, at the gene expression level, is highly induced after nerve injury and that these transcripts prevailed in the rat strain most susceptible to nerve cell loss.

What lies in the future you never know. One can speculate that to be able to dissect the intricate puzzle of which gene sequence variations as well as which other pathogenic mechanisms contribute to the susceptibility to complex diseases some critical issues need to be taken into consideration. There will be a need to gather large well characterized patient materials, of DNA but also of relevant tissues, to enable functional studies in the human situation. I believe that there will be a need for "omics" strategies, including genomics, transcriptomics, proteomics and metabolomics, all dealing with high throughput studies of different cell components. This will also necessitate highly collaborative studies – dependant on the amount of subjects needed as well as the cost involved. There will also be a requirement for researchers from many different disciplines taking part. Already today we can find out which variations a certain person carries of 500 000 SNPs as well as the transcriptional profile of 47 000 transcripts.

My dream is that we in the future will be able to record the whole genomic sequence as well as transcriptional and protein profiles from different tissues and timepoints from a large number of patients and controls. I have already seen one utopia being fulfilled, so maybe it is not totally unbelievable……

7 ACKNOWLEDGEMENTS

First of all I would like to thank my supervisor Professor Jan Hillert for giving me the opportunity to perform this project within your group. Always encouraging and inspiring and with an enormous knowledge in the MS field. You also have an outstanding skill in solving problems. Also many thanks for all the courses and conferences you have made possible for me to participate in.

I would also like to thank Professor Tomas Olsson for the invitation to work together with you and your group on the gene expression part of this project. You have created such a pleasant atmosphere for research in your laboratory and your addiction to research has inspired me.

Many thanks will go to Kjell Stenberg, my boss at Astra during the *Helicobacter pylori* time. You really taught me how wonderful research is and always thought that a dedicated person could make the difference. Your belief in me really made me, eventually, decide on performing this PhD thesis work.

I would like to thank all the MS patients and healthy volunteers who have participated in the studies included in my thesis.

I would also like to thank the healthcare staffs in Huddinge and Solna for participation in that sample collection. Also a special thank the staff at Kringeltappen in Södertälje.

I am very grateful to all past and present members of the Division of Neurology, especially from the wonderful Neurogenetics group.

- *Eva Åkesson* – my office mate and friend. For all interesting discussions, about research and life, and for invaluable help with the thesis writing. I would not have made it without you!
- *Helena Modin* – for company during travels and for collaboration on infrastructure
- *Thomas Masterman*- for your wisdom and true interest in research
- *Frida Lundmark* – for fruitful collaborations with many amusing and frustrating moments
- *Boel Bryndeal* – for HLA –A collaboration, deep discussions on all part of statistics and for understanding the need for chocolate during manuscript writing
- *Christina Sjöstrand* – for fruitful collaboration – I have learnt a lot about headache from you
- *Kerstin Imrell*- for giving things another perspective
- *Izaura Roos* - for interesting LD discussions
- *Viriginja Karrenbauer*- for your captivating energy
- *Cecilia Svarén Quiding* – for all nice chats about our Belgian friends as well as help with laboratory related issues

Also many thanks to, Sanjaya Adikari, Jenny Ahlqvist, Anna Cunningham, Mabel Cruz, Yamei Dai, Ulrika Einarsson, Gabriel von Euler, Vilmantas Giedraitis, Andreia Gomez, Kristina Gottberg, Anna Fogdell-Hahn, Eva Johansson, Kosta Kostulas, Mathilde Kouwenhoven, Arturs Ligers, Anneli Matisson, Susanna Mjörnheim,

Rayomand Press, Prabhakar Puthei, Uros Root, Ajith Sominanada, Leszek Stawiarz, Anna Steinberg, Dan Svensson, Mathula Thangarajh, Volkan Özenci. Our summer workers Pernilla Anderson, Karin Moriatis, Wiveka Olsson, Erik Nelsson and everybody else past and present for a very nice years together.
A special thank to Anna Ljungberg, Faezeh Vejdani and Merja Kanerva in the laboratory (Merja – I will soon leave now and everything will be so calm again….) and to Gunnel Larsson for much help with adminstration

Many thanks to my students Hilde Shah, Malin Anderson and Ulf Stille: there is no better way to learn things than to explain them for other people.

A large thank to my collaborators in Tomas Olsson's group in Solna:
Especially to Maria Swanberg for your enormous interest in research, fun but sweaty moment in the HS lab, discussion on microarray analysis, manuscript writing and friendship.
Also many thanks to Olle Lidman for your great knowledge and help with practical things for the thesis, about ethical permissions, latest version of manuscript, pictures and so on.
Many thanks also to Margerita Diez and Fredrik Piehl for great collaboration.
And special thanks to Mohsen Khademi and Erik Wallström for collaboration on a project that really will be something special!

My co authors on the genetic project at AstraZeneca for fruitful collaboration: Zhiping Zhang, Dennis Hellgren, Hugh Salter, Tesfai Emhaizon, Gunnar Falk and Maria Anvret.

A very warm and grateful thank to Gudrun Jonasdottir - a better guide to the mysterious world of statistics does not exist! Always prompt answers to strange questions from a poor biologist and for always performing the analyses in a moment. Also thank to Maria Grunewald and Juni Palmgren for statistical collaborations.

Lots of thanks to Johan Lidén, Marika Rönnholm and David Brodin at the Affymetrix core facilty. Never getting tired of my questions, need for prompt Agilent analyses and valuable discussion on microarray experiment design and analysis.

Many thanks to Cecilia Lindgren, Astrid Lindstedt and Linda Berglind at the mutation analysis facility for fruitful collaborations, nice discussions and excellent genotypings.

Many thanks to Andreas Kjerrman at Scania for help with designing the cover of this thesis – the cake is on its way!

Also many thanks to my "AstraZeneca" friends Anna-Karin Tidén, Christina Svensson, Margareta Berg, Åke Larsson, Bo Franzén and all the rest of the old gang.

A special thank to my friends Carina Jonsson and Inger Axberg for dinners, holidays and garden visits.

Also a warm thank to all friends from the dog world – especially all the members of "sökgänget" - letting me hide in the forest every Sunday and playing with my tricky Hrimm and HUS for being so engaged and nice people. A special thank to Sophia Trenkle Nyberg, participating in both groups, also for very nice lunch trainings, chats, dinners and friendship.

Thank to Sofie Lennmark for Friday swim-company and for the Volvo

Lots of thanks to my friends who I met during my most intensive horse years – Eva Netterberg and Karin Åsebring. Taking care of the party together with Anna is the best present anyone can get, thank you! Also thanks to Eva and Martin for taking so good care of Gnejs while my time was spent on research.

A thousand thanks to my dear friends Anna Gunnarsson and Ulrika Sullivan, you two are the best – what more can I say. Always supporting, caring and full of surprises.

Even if they not will read this (not even my animals are capable of that) thank to you all for keeping my mind on other things than research. Late Frosty, Hrimm, Gnejs and little Morris, not to forget Ettan, Tvåan, Fiona and Einar.

Then warm thanks to my family: My brother Sune, my sister in law Anne and there wonderful children: Andreas, Hampus, Joline and Ellinore. My brother Håkan. My wonderful mother and father, Gun and Alf. Thank you for all your support both mentally and practically during the years – thank you mum for proofreading the thesis. I will never learn if it should be was or were….. Tack pappa för att du hållit igång mina traktor- och hökörar-kunskaper under åren – man har alltid nytta av att ha fötterna på/i jorden!
Many thanks also to my "Enhörna family" – Ulla, Rolf, Anita, Maria, Juha, Hampus, Pontus and Malte. When will we go to Tenerife again?

Last and most of all I would like to thank *Tommy,* my husband, for all love and understanding – always. You have always let me be the one I am and you have always been the one you are. I love you so much for that! Thank you for everything.

8 REFERENCES

1. Charcot J. 1868. Histologic de la sclerose en plaque. *Gaz Hop* 41: 554-66
2. Akesson E, Fredrikson S. 2004. [Multiple sclerosis]. *Lakartidningen* 101: 2489-93
3. Hafler DA, Slavik JM, Anderson DE, et al. 2005. Multiple sclerosis. *Immunol Rev* 204: 208-31.
4. Keegan BM, Noseworthy JH. 2002. Multiple sclerosis. *Annu Rev Med* 53: 285-302.
5. Compston A, Coles A. 2002. Multiple sclerosis. *Lancet* 359: 1221-31.
6. Montalban X. 2005. Primary progressive multiple sclerosis. *Curr Opin Neurol* 18: 261-6.
7. Simone IL, Carrara D, Tortorella C, et al. 2002. Course and prognosis in early-onset MS: comparison with adult-onset forms. *Neurology* 59: 1922-8.
8. Weinshenker BG. 1995. The natural history of multiple sclerosis. *Neurol Clin* 13: 119-46
9. Einarsson U, Gottberg K, von Koch L, et al. In press. Cognitive and motor function in people with multiple sclerosis in Stockholm County. *Multiple Sclerosis*
10. McDonald WI, Compston A, Edan G, et al. 2001. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 50: 121-7
11. Freedman MS, Thompson EJ, Deisenhammer F, et al. 2005. Recommended standard of cerebrospinal fluid analysis in the diagnosis of multiple sclerosis: a consensus statement. *Arch Neurol* 62: 865-70.
12. Kurtzke JF. 1961. On the evalutation of disability in multiple sclerosis. *Neurology* 11: 686-94
13. Fischer JS, Rudick RA, Cutter GR, Reingold SC. 1999. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. *Mult Scler* 5: 244-50
14. Roxburgh RH, Seaman SR, Masterman T, et al. 2005. Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. *Neurology* 64: 1144-51
15. Henriksson F, Fredrikson S, Masterman T, Jonsson B. 2001. Costs, quality of life and disease severity in multiple sclerosis: a cross-sectional study in Sweden. *Eur J Neurol* 8: 27-35.
16. Patwardhan MB, Matchar DB, Samsa GP, et al. 2005. Cost of multiple sclerosis by level of disability: a review of literature. *Mult Scler* 11: 232-9
17. Kobelt G. 2004. Economic evidence in multiple sclerosis: a review. *Eur J Health Econ* 5 Suppl 1: S54-62
18. 1995. Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. The IFNB Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group. *Neurology* 45: 1277-85
19. 1998. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group. *Lancet* 352: 1498-504
20. Jacobs LD, Cookfair DL, Rudick RA, et al. 1996. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann Neurol* 39: 285-94
21. Filippini G, Munari L, Incorvaia B, et al. 2003. Interferons in relapsing remitting multiple sclerosis: a systematic review. *Lancet* 361: 545-52
22. Hohlfeld R. 1997. Biotechnological agents for the immunotherapy of multiple sclerosis. Principles, problems and perspectives. *Brain* 120 (Pt 5): 865-916

23. Johnson KP, Brooks BR, Cohen JA, et al. 1995. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. *Neurology* 45: 1268-76
24. Comi G, Filippi M, Wolinsky JS. 2001. European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging--measured disease activity and burden in patients with relapsing multiple sclerosis. European/Canadian Glatiramer Acetate Study Group. *Ann Neurol* 49: 290-7
25. Munari L, Lovati R, Boiko A. 2004. Therapy with glatiramer acetate for multiple sclerosis. *Cochrane Database Syst Rev*: CD004678
26. Filippi M. 2001. Non-conventional MR techniques to monitor the evolution of multiple sclerosis. *Neurol Sci* 22: 195-200
27. Farina C, Weber MS, Meinl E, et al. 2005. Glatiramer acetate in multiple sclerosis: update on potential mechanisms of action. *Lancet Neurol* 4: 567-75
28. Miller DH, Khan OA, Sheremata WA, et al. 2003. A controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med* 348: 15-23
29. Steinman L. 2005. Blocking adhesion molecules as therapy for multiple sclerosis: natalizumab. *Nat Rev Drug Discov* 4: 510-8
30. Wiendl H, Hohlfeld R. 2002. Therapeutic approaches in multiple sclerosis: lessons from failed and interrupted treatment trials. *BioDrugs* 16: 183-200
31. Markovic-Plese S, McFarland HF. 2001. Immunopathogenesis of the multiple sclerosis lesion. *Curr Neurol Neurosci Rep* 1: 257-62
32. Pettinelli CB, McFarlin DE. 1981. Adoptive transfer of experimental allergic encephalomyelitis in SJL/J mice after in vitro activation of lymph node cells by myelin basic protein: requirement for Lyt 1+ 2- T lymphocytes. *J Immunol* 127: 1420-3
33. Dyment DA, Ebers GC, Sadovnick AD. 2004. Genetics of multiple sclerosis. *Lancet Neurol* 3: 104-10
34. Fogdell-Hahn A, Ligers A, Gronning M, et al. 2000. Multiple sclerosis: a modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. *Tissue Antigens* 55: 140-8
35. Harbo HF, Lie BA, Sawcer S, et al. 2004. Genes in the HLA class I region may contribute to the HLA class II-associated genetic susceptibility to multiple sclerosis. *Tissue Antigens* 63: 237-47
36. Bielekova B, Sung MH, Kadom N, et al. 2004. Expansion and functional relevance of high-avidity myelin-specific CD4+ T cells in multiple sclerosis. *J Immunol* 172: 3893-904
37. Kurtzke JF. 1993. Epidemiologic evidence for multiple sclerosis as an infection. *Clin Microbiol Rev* 6: 382-427.
38. Kostulas VK, Link H, Lefvert AK. 1987. Oligoclonal IgG bands in cerebrospinal fluid. Principles for demonstration and interpretation based on findings in 1114 neurological patients. *Arch Neurol* 44: 1041-4
39. Lipton HL. 1975. Theiler's virus infection in mice: an unusual biphasic disease process leading to demyelination. *Infect Immun* 11: 1147-55
40. Goverman J, Woods A, Larson L, et al. 1993. Transgenic mice that express a myelin basic protein-specific T cell receptor develop spontaneous autoimmunity. *Cell* 72: 551-60
41. Granieri E, Casetta I, Tola MR, Ferrante P. 2001. Multiple sclerosis: infectious hypothesis. *Neurol Sci* 22: 179-85
42. Fotheringham J, Jacobson S. 2005. Human herpesvirus 6 and multiple sclerosis: potential mechanisms for virus-induced disease. *Herpes* 12: 4-9
43. Swanborg RH, Whittum-Hudson JA, Hudson AP. 2003. Infectious agents and multiple sclerosis--are Chlamydia pneumoniae and human herpes virus 6 involved? *J Neuroimmunol* 136: 1-8
44. Barnett MH, Prineas JW. 2004. Relapsing and remitting multiple sclerosis: pathology of the newly forming lesion. *Ann Neurol* 55: 458-68
45. Filippi M, Bozzali M, Rovaris M, et al. 2003. Evidence for widespread axonal damage at the earliest clinical stage of multiple sclerosis. *Brain* 126: 433-7

46. Lucchinetti C, Bruck W, Parisi J, et al. 2000. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Ann Neurol* 47: 707-17
47. Trapp BD. 2004. Pathogenesis of multiple sclerosis: the eyes only see what the mind is prepared to comprehend. *Ann Neurol* 55: 455-7
48. Ferguson B, Matyszak MK, Esiri MM, Perry VH. 1997. Axonal damage in acute multiple sclerosis lesions. *Brain* 120: 393-9
49. Kornek B, Storch MK, Weissert R, et al. 2000. Multiple sclerosis and chronic autoimmune encephalomyelitis: a comparative quantitative study of axonal injury in active, inactive, and remyelinated lesions. *Am J Pathol* 157: 267-76.
50. Pugliatti M, Sotgiu S, Rosati G. 2002. The worldwide prevalence of multiple sclerosis. *Clin Neurol Neurosurg* 104: 182-91
51. Marrie RA. 2004. Environmental risk factors in multiple sclerosis aetiology. *Lancet Neurol* 3: 709-18.
52. Compston A. 1997. Genetic epidemiology of multiple sclerosis. *J Neurol Neurosurg Psychiatry* 62: 553-61
53. Gale CR, Martyn CN. 1995. Migrant studies in multiple sclerosis. *Prog Neurobiol* 47: 425-48
54. VanAmerongen BM, Dijkstra CD, Lips P, Polman CH. 2004. Multiple sclerosis and vitamin D: an update. *Eur J Clin Nutr* 58: 1095-109
55. Kurtzke JF, Hyllested K. 1975. Multiple sclerosis: an epidemic disease in the Faeroes. *Trans Am Neurol Assoc* 100: 213-5
56. Poser CM, Hibberd PL, Benedikz J, Gudmundsson G. 1988. Analysis of the 'epidemic' of multiple sclerosis in the Faroe Islands. I. Clinical and epidemiological aspects. *Neuroepidemiology* 7: 168-80
57. Poser CM, Hibberd PL. 1988. Analysis of the 'epidemic' of multiple sclerosis in the Faroe Islands. II. Biostatistical aspects. *Neuroepidemiology* 7: 181-9
58. Sadovnick AD, Baird PA, Ward RH. 1988. Multiple sclerosis: updated risks for relatives. *Am J Med Genet* 29: 533-41
59. Ebers GC, Bulman DE, Sadovnick AD, et al. 1986. A population-based study of multiple sclerosis in twins. *N Engl J Med* 315: 1638-42
60. Mumford CJ, Wood NW, Kellar-Wood H, et al. 1994. The British Isles survey of multiple sclerosis in twins. *Neurology* 44: 11-5
61. Hansen T, Skytthe A, Stenager E, et al. 2005. Concordance for multiple sclerosis in Danish twins: an update of a nationwide study. *Mult Scler* 11: 504-10
62. Sadovnick AD, Armstrong H, Rice GP, et al. 1993. A population-based study of multiple sclerosis in twins: update. *Ann Neurol* 33: 281-5
63. Willer CJ, Dyment DA, Risch NJ, et al. 2003. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci U S A* 100: 12877-82
64. Ebers GC. 1996. Genetic epidemiology of multiple sclerosis. *Curr Opin Neurol* 9: 155-8
65. Ebers GC, Yee IM, Sadovnick AD, Duquette P. 2000. Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. Canadian Collaborative Study Group. *Ann Neurol* 48: 927-31
66. Robertson NP, O'Riordan JI, Chataway J, et al. 1997. Offspring recurrence rates and clinical characteristics of conjugal multiple sclerosis. *Lancet* 349: 1587-90
67. Ebers GC, Sadovnick AD, Risch NJ. 1995. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature* 377: 150-1
68. Griffiths MM, Encinas JA, Remmers EF, et al. 1999. Mapping autoimmunity genes. *Curr Opin Immunol* 11: 689-700
69. Zamvil SS, Steinman L. 1990. The T lymphocyte in experimental allergic encephalomyelitis. *Annu Rev Immunol* 8: 579-621
70. Rocken M, Racke M, Shevach EM. 1996. IL-4-induced immune deviation as antigen-specific therapy for inflammatory autoimmune disease. *Immunol Today* 17: 225-31
71. Huseby ES, Liggitt D, Brabb T, et al. 2001. A pathogenic role for myelin-specific CD8(+) T cells in a model for multiple sclerosis. *J Exp Med* 194: 669-76

72. Skulina C, Schmidt S, Dornmair K, et al. 2004. Multiple sclerosis: brain-infiltrating CD8+ T cells persist as clonal expansions in the cerebrospinal fluid and blood. *Proc Natl Acad Sci U S A* 101: 2428-33
73. Cabarrocas J, Bauer J, Piaggio E, et al. 2003. Effective and selective immune surveillance of the brain by MHC class I-restricted cytotoxic T lymphocytes. *Eur J Immunol* 33: 1174-82
74. Lafaille JJ, Nagashima K, Katsuki M, Tonegawa S. 1994. High incidence of spontaneous autoimmune encephalomyelitis in immunodeficient anti-myelin basic protein T cell receptor transgenic mice. *Cell* 78: 399-408
75. Waldner H, Whitters MJ, Sobel RA, et al. 2000. Fulminant spontaneous autoimmunity of the central nervous system in mice transgenic for the myelin proteolipid protein-specific T cell receptor. *Proc Natl Acad Sci U S A* 97: 3412-7
76. Piehl F, Lidman O. 2001. Neuroinflammation in the rat--CNS cells and their role in the regulation of immune reactions. *Immunol Rev* 184: 212-25
77. De Stefano N, Matthews PM, Fu L, et al. 1998. Axonal damage correlates with disability in patients with relapsing-remitting multiple sclerosis. Results of a longitudinal magnetic resonance spectroscopy study. *Brain* 121: 1469-77
78. Koliatsos VE, Cayouette MH, Berkemeier LR, et al. 1994. Neurotrophin 4/5 is a trophic factor for mammalian facial motor neurons. *Proceedings of the National Academy of Sciences of the United States of America* 91: 3304-8
79. Lidman O, Swanberg M, Horvath L, et al. 2003. Discrete gene loci regulate neurodegeneration, lymphocyte infiltration, and major histocompatibility complex class II expression in the CNS. *J Neurosci* 23: 9817-23.
80. Piehl F, Lundberg C, Khademi M, et al. 1999. Non-MHC gene regulation of nerve root injury-induced spinal cord inflammation and neuron death. *J. Neuroimmunol.* 101: 87-97
81. Hammarberg H, Lidman O, Lundberg C, et al. 2000. Neuroprotection by encephalomyelitis: rescue of mechanically injured neurons and neurotrophin production by CNS-infiltrating T and natural killer cells. *J Neurosci* 20: 5283-91.
82. Litt M, Luty JA. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* 44: 397-401
83. 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 72: 971-83
84. Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5: 435-45
85. 2003. The International HapMap Project. *Nature* 426: 789-96
86. Holmans P. 1998. Affected sib-pair methods for detecting linkage to dichotomous traits: review of the methodology. *Hum Biol* 70: 1025-40
87. Jimenez-Sanchez G, Childs B, Valle D. 2001. Human disease genes. *Nature* 409: 853-5
88. Altmuller J, Palmer LJ, Fischer G, et al. 2001. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 69: 936-50
89. Hirschhorn JN, Daly MJ. 2005. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6: 95-108
90. Risch N, Merikangas K. 1996. The future of genetic studies of complex human diseases. *Science* 273: 1516-7
91. Cardon LR, Bell JI. 2001. Association study designs for complex diseases. *Nat Rev Genet* 2: 91-9
92. Venter JC, Adams MD, Myers EW, et al. 2001. The sequence of the human genome. *Science* 291: 1304-51
93. Lander ES, Linton LM, Birren B, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409: 860-921
94. Altshuler D, Hirschhorn JN, Klannemark M, et al. 2000. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26: 76-80

95. Hugot JP, Chamaillard M, Zouali H, et al. 2001. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411: 599-603
96. Ogura Y, Bonen DK, Inohara N, et al. 2001. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 411: 603-6.
97. Gregersen PK. 2005. Pathways to gene identification in rheumatoid arthritis: PTPN22 and beyond. *Immunol Rev* 204: 74-86
98. Ardlie KG, Kruglyak L, Seielstad M. 2002. Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3: 299-309
99. Gabriel SB, Schaffner SF, Nguyen H, et al. 2002. The structure of haplotype blocks in the human genome. *Science* 296: 2225-9
100. Mira MT, Alcais A, Nguyen VT, et al. 2004. Susceptibility to leprosy is associated with PARK2 and PACRG. *Nature* 427: 636-40
101. Excoffier L, Slatkin M. 1995. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 12: 921-7
102. Rife DC. 1954. Populations of hybrid origin as source material for the detection of linkage. *Am J Hum Genet* 6: 26-33
103. Smith MW, Patterson N, Lautenberger JA, et al. 2004. A high-density admixture map for disease gene discovery in african americans. *Am J Hum Genet* 74: 1001-13
104. Zhu X, Luke A, Cooper RS, et al. 2005. Admixture mapping for hypertension loci with genome-scan markers. *Nat Genet* 37: 177-81
105. Cardon LR, Palmer LJ. 2003. Population stratification and spurious allelic association. *Lancet* 361: 598-604
106. Colhoun HM, McKeigue PM, Davey Smith G. 2003. Problems of reporting genetic associations with complex outcomes. *Lancet* 361: 865-72
107. McCarthy MI, Smedley D, Hide W. 2003. New methods for finding disease-susceptibility genes: impact and potential. *Genome Biol* 4: 119
108. Zondervan KT, Cardon LR. 2004. The complex interplay among factors that influence allelic association. *Nat Rev Genet* 5: 89-100
109. Abecasis GR, Cherny SS, Cardon LR. 2001. The impact of genotyping error on family-based analysis of quantitative traits. *Eur J Hum Genet* 9: 130-4
110. Risch NJ. 2000. Searching for genetic determinants in the new millennium. *Nature* 405: 847-56
111. Jersild C, Fog T. 1972. Histocompatibility (HL-A) antigens associated with multiple sclerosis. *Acta Neurol Scand Suppl* 51: 377
112. Jersild C, Dupont B, Fog T, et al. 1973. Histocompatibility-linked immune-response determinants in multiple sclerosis. *Transplant Proc* 5: 1791-6
113. Hillert J, Olerup O. 1993. Multiple sclerosis is associated with genes within or close to the HLA-DR-DQ subregion on a normal DR15,DQ6,Dw2 haplotype. *Neurology* 43: 163-8
114. Doyle C, Strominger JL. 1987. Interaction between CD4 and class II MHC molecules mediates cell adhesion. *Nature* 330: 256-9
115. Marrosu MG, Murru R, Murru MR, et al. 2001. Dissection of the HLA association with multiple sclerosis in the founder isolated population of Sardinia. *Hum Mol Genet* 10: 2907-16
116. Marrosu MG, Murru MR, Costa G, et al. 1998. DRB1-DQA1-DQB1 loci and multiple sclerosis predisposition in the Sardinian population. *Hum Mol Genet* 7: 1235-7
117. Marrosu MG, Murru MR, Costa G, et al. 1997. Multiple sclerosis in Sardinia is associated and in linkage disequilibrium with HLA-DR3 and -DR4 alleles. *Am J Hum Genet* 61: 454-7
118. Modin H, Olsson W, Hillert J, Masterman T. 2004. Modes of action of HLA-DR susceptibility specificities in multiple sclerosis. *Am J Hum Genet* 74: 1321-2
119. Barcellos LF, Oksenberg JR, Begovich AB, et al. 2003. HLA-DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. *Am J Hum Genet* 72: 710-6
120. Ligers A, Dyment DA, Willer CJ, et al. 2001. Evidence of linkage with HLA-DR in DRB1*15-negative families with multiple sclerosis. *Am J Hum Genet* 69: 900-3

121. Lincoln MR, Montpetit A, Cader MZ, et al. 2005. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat Genet* 37: 1108-12
122. Oksenberg JR, Barcellos LF, Cree BA, et al. 2004. Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. *Am J Hum Genet* 74: 160-7
123. Caballero A, Alves-Leon S, Papais-Alvarenga R, et al. 1999. DQB1*0602 confers genetic susceptibility to multiple sclerosis in Afro-Brazilians. *Tissue Antigens* 54: 524-6
124. Norment AM, Salter RD, Parham P, et al. 1988. Cell-cell adhesion mediated by CD8 and MHC class I molecules. *Nature* 336: 79-81
125. Sawcer S, Jones HB, Feakes R, et al. 1996. A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nat Genet* 13: 464-8
126. Ebers GC, Kukay K, Bulman DE, et al. 1996. A full genome search in multiple sclerosis. *Nat Genet* 13: 472-6
127. Haines JL, Ter-Minassian M, Bazyk A, et al. 1996. A complete genomic screen for multiple sclerosis underscores a role for the major histocompatability complex. The Multiple Sclerosis Genetics Group. *Nat Genet* 13: 469-71
128. Kuokkanen S, Gschwend M, Rioux JD, et al. 1997. Genomewide scan of multiple sclerosis in Finnish multiplex families. *Am J Hum Genet* 61: 1379-87
129. 2001. A meta-analysis of genomic screens in multiple sclerosis. The Transatlantic Multiple Sclerosis Genetics Cooperative. *Mult Scler* 7: 3-11
130. Ban M, Stewart GJ, Bennetts BH, et al. 2002. A genome screen for linkage in Australian sibling-pairs with multiple sclerosis. *Genes Immun* 3: 464-9
131. Hensiek AE, Roxburgh R, Smilie B, et al. 2003. Updated results of the United Kingdom linkage-based genome screen in multiple sclerosis. *J Neuroimmunol* 143: 25-30
132. Broadley S, Sawcer S, D'Alfonso S, et al. 2001. A genome screen for multiple sclerosis in Italian families. *Genes Immun* 2: 205-10
133. Coraddu F, Sawcer S, D'Alfonso S, et al. 2001. A genome screen for multiple sclerosis in Sardinian multiplex families. *Eur J Hum Genet* 9: 621-6
134. Akesson E, Oturai A, Berg J, et al. 2002. A genome-wide screen for linkage in Nordic sib-pairs with multiple sclerosis. *Genes Immun* 3: 279-85
135. Eraksoy M, Kurtuncu M, Akman-Demir G, et al. 2003. A whole genome screen for linkage in Turkish multiple sclerosis. *J Neuroimmunol* 143: 17-24
136. 2003. A meta-analysis of whole genome linkage screens in multiple sclerosis. *J Neuroimmunol* 143: 39-46
137. Sawcer S, Ban M, Maranian M, et al. 2005. A high-density screen for linkage in multiple sclerosis. *Am J Hum Genet* 77: 454-67
138. Oksenberg JR, Barcellos LF. 2005. Multiple sclerosis genetics: leaving no stone unturned. *Genes Immun* 6: 375-87
139. Giordano M, D'Alfonso S, Momigliano-Richiardi P. 2002. Genetics of multiple sclerosis: linkage and association studies. *Am J Pharmacogenomics* 2: 37-58.
140. Kenealy SJ, Pericak-Vance MA, Haines JL. 2003. The genetic epidemiology of multiple sclerosis. *J Neuroimmunol* 143: 7-12.
141. Barcellos LF, Begovich AB, Reynolds RL, et al. 2004. Linkage and association with the NOS2A locus on chromosome 17q11 in multiple sclerosis. *Ann Neurol* 55: 793-800
142. Cunningham S, Patterson CC, McDonnell G, et al. 2005. Haplotype analysis of the preprotachykinin-1 (TAC1) gene in multiple sclerosis. *Genes Immun* 6: 265-70
143. Barton A, Woolmore JA, Ward D, et al. 2004. Association of protein kinase C alpha (PRKCA) gene with multiple sclerosis in a UK population. *Brain* 127: 1717-22
144. Swanberg M, Lidman O, Padyukov L, et al. 2005. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nat Genet* 37: 486-94
145. Sandbrink R, Hartmann T, Masters CL, Beyreuther K. 1996. Genes contributing to Alzheimer's disease. *Mol Psychiatry* 1: 27-40

146. Enzinger C, Ropele S, Smith S, et al. 2004. Accelerated evolution of brain atrophy and "black holes" in MS patients with APOE-epsilon 4. *Ann Neurol* 55: 563-9
147. De Stefano N, Bartolozzi ML, Nacmias B, et al. 2004. Influence of apolipoprotein E epsilon4 genotype on brain tissue integrity in relapsing-remitting multiple sclerosis. *Arch Neurol* 61: 536-40
148. Zakrzewska-Pniewska B, Styczynska M, Podlecka A, et al. 2004. Association of apolipoprotein E and myeloperoxidase genotypes to clinical course of familial and sporadic multiple sclerosis. *Mult Scler* 10: 266-71
149. Zwemmer JN, van Veen T, van Winsen L, et al. 2004. No major association of ApoE genotype with disease characteristics and MRI findings in multiple sclerosis. *Mult Scler* 10: 272-7
150. Savettieri G, Andreoli V, Bonavita S, et al. 2003. Apolipoprotein E genotype does not influence the progression of multiple sclerosis. *J Neurol* 250: 1094-8
151. Reich D, Patterson N, Jager PL, et al. 2005. A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility. *Nat Genet* 37: 1113-8
152. Wallin MT, Page WF, Kurtzke JF. 2004. Multiple sclerosis in US veterans of the Vietnam era and later military service: race, sex, and geography. *Ann Neurol* 55: 65-71
153. Vitale E, Cook S, Sun R, et al. 2002. Linkage analysis conditional on HLA status in a large North American pedigree supports the presence of a multiple sclerosis susceptibility locus on chromosome 12p12. *Hum Mol Genet* 11: 295-300
154. Modin H, Masterman T, Thorlacius T, et al. 2003. Genome-wide linkage screen of a consanguineous multiple sclerosis kinship. *Mult Scler* 9: 128-34
155. Lockhart DJ, Dong H, Byrne MC, et al. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14: 1675-80.
156. Liang P, Pardee AB. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257: 967-71
157. Adams MD, Kelley JM, Gocayne JD, et al. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252: 1651-6
158. Zhao N, Hashida H, Takahashi N, et al. 1995. High-density cDNA filter analysis: a novel approach for large-scale, quantitative analysis of gene expression. *Gene* 156: 207-13
159. DeRisi J, Penland L, Brown PO, et al. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 14: 457-60
160. Larkin JE, Frank BC, Gavras H, et al. 2005. Independence and reproducibility across microarray platforms. *Nat Methods* 2: 337-44
161. Irizarry RA, Warren D, Spencer F, et al. 2005. Multiple-laboratory comparison of microarray platforms. *Nat Methods* 2: 345-50
162. Harbig J, Sprinkle R, Enkemann SA. 2005. A sequence-based identification of the genes detected by probesets on the Affymetrix U133 plus 2.0 array. *Nucleic Acids Res* 33: e31
163. Ringner M, Peterson C. 2003. Microarray-based cancer diagnosis with artificial neural networks. *Biotechniques* Suppl: 30-5
164. Peterson LE. 2002. CLUSFAVOR 5.0: hierarchical cluster and principal-component analysis of microarray-based transcriptional profiles. *Genome Biol* 3: SOFTWARE0002
165. Ringner M, Peterson C, Khan J. 2002. Analyzing array data using supervised methods. *Pharmacogenomics* 3: 403-15
166. Quackenbush J. 2001. Computational analysis of microarray data. *Nat Rev Genet* 2: 418-27
167. Irizarry RA, Bolstad BM, Collin F, et al. 2003. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31: e15
168. Li C, Wong WH. 2001. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 98: 31-6

169. Carmody RJ, Hilliard B, Maguschak K, et al. 2002. Genomic scale profiling of autoimmune inflammation in the central nervous system: the nervous response to inflammation. *J Neuroimmunol* 133: 95-107
170. Matejuk A, Hopke C, Dwyer J, et al. 2003. CNS gene expression pattern associated with spontaneous experimental autoimmune encephalomyelitis. *J Neurosci Res* 73: 667-78
171. Baranzini SE, Bernard CC, Oksenberg JR. 2005. Modular transcriptional activity characterizes the initiation and progression of autoimmune encephalomyelitis. *J Immunol* 174: 7412-22
172. Mix E, Ibrahim S, Pahnke J, et al. 2004. Gene-expression profiling of the early stages of MOG-induced EAE proves EAE-resistance as an active process. *J Neuroimmunol* 151: 158-70
173. Whitney LW, Ludwin SK, McFarland HF, Biddison WE. 2001. Microarray analysis of gene expression in multiple sclerosis and EAE identifies 5-lipoxygenase as a component of inflammatory lesions. *J Neuroimmunol* 121: 40-8
174. Alt C, Duvefelt K, Franzen B, et al. 2005. Gene and protein expression profiling of the microvascular compartment in experimental autoimmune encephalomyelitis in C57Bl/6 and SJL mice. *Brain Pathol* 15: 1-16
175. Spach KM, Pedersen LB, Nashold FE, et al. 2004. Gene expression analysis suggests that 1,25-dihydroxyvitamin D3 reverses experimental autoimmune encephalomyelitis by stimulating inflammatory cell apoptosis. *Physiol Genomics* 18: 141-51
176. Paintlia AS, Paintlia MK, Singh AK, et al. 2004. Regulation of gene expression associated with acute experimental autoimmune encephalomyelitis by Lovastatin. *J Neurosci Res* 77: 63-81
177. Matejuk A, Dwyer J, Zamora A, et al. 2002. Evaluation of the effects of 17beta-estradiol (17beta-e2) on gene expression in experimental autoimmune encephalomyelitis using DNA microarray. *Endocrinology* 143: 313-9
178. Velardo MJ, Burger C, Williams PR, et al. 2004. Patterns of gene expression reveal a temporally orchestrated wound healing response in the injured spinal cord. *J Neurosci* 24: 8562-76
179. Hu J, Fink D, Mata M. 2002. Microarray analysis suggests the involvement of proteasomes, lysosomes, and matrix metalloproteinases in the response of motor neurons to root avulsion. *Eur J Neurosci* 16: 1409-16.
180. Chabas D, Baranzini SE, Mitchell D, et al. 2001. The influence of the proinflammatory cytokine, osteopontin, on autoimmune demyelinating disease. *Science* 294: 1731-5
181. Graumann U, Reynolds R, Steck AJ, Schaeren-Wiemers N. 2003. Molecular changes in normal appearing white matter in multiple sclerosis are characteristic of neuroprotective mechanisms against hypoxic insult. *Brain Pathol* 13: 554-73
182. Tajouri L, Mellick AS, Ashton KJ, et al. 2003. Quantitative and qualitative changes in gene expression patterns characterize the activity of plaques in multiple sclerosis. *Brain Res Mol Brain Res* 119: 170-83
183. Lock C, Hermans G, Pedotti R, et al. 2002. Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat Med* 8: 500-8
184. Whitney LW, Becker KG, Tresser NJ, et al. 1999. Analysis of gene expression in mutiple sclerosis lesions using cDNA microarrays. *Ann Neurol* 46: 425-8
185. Mycko MP, Papoian R, Boschert U, et al. 2003. cDNA microarray analysis in multiple sclerosis lesions: detection of genes associated with disease activity. *Brain* 126: 1048-57
186. Lindberg RL, De Groot CJ, Certa U, et al. 2004. Multiple sclerosis as a generalized CNS disease--comparative microarray analysis of normal appearing white matter and lesions in secondary progressive MS. *J Neuroimmunol* 152: 154-67
187. Iglesias AH, Camelo S, Hwang D, et al. 2004. Microarray detection of E2F pathway activation and other targets in multiple sclerosis peripheral blood mononuclear cells. *J Neuroimmunol* 150: 163-77

188. Mayne M, Moffatt T, Kong H, et al. 2004. CYFIP2 is highly abundant in CD4+ cells from multiple sclerosis patients and is involved in T cell adhesion. *Eur J Immunol* 34: 1217-27
189. Ramanathan M, Weinstock-Guttman B, Nguyen LT, et al. 2001. In vivo gene expression revealed by cDNA arrays: the pattern in relapsing-remitting multiple sclerosis patients compared with normal subjects. *J Neuroimmunol* 116: 213-9.
190. Achiron A, Gurevich M, Magalashvili D, et al. 2004. Understanding autoimmune mechanisms in multiple sclerosis using gene expression microarrays: treatment effect and cytokine-related pathways. *Clin Dev Immunol* 11: 299-305
191. Bomprezzi R, Ringner M, Kim S, et al. 2003. Gene expression profile in multiple sclerosis patients and healthy controls: identifying pathways relevant to disease. *Hum Mol Genet* 12: 2191-9
192. Mandel M, Gurevich M, Pauzner R, et al. 2004. Autoimmunity gene expression portrait: specific signature that intersects or differentiates between multiple sclerosis and systemic lupus erythematosus. *Clin Exp Immunol* 138: 164-70
193. Weinstock-Guttman B, Badgett D, Patrick K, et al. 2003. Genomic effects of IFN-beta in multiple sclerosis patients. *J Immunol* 171: 2694-702
194. Koike F, Satoh J, Miyake S, et al. 2003. Microarray analysis identifies interferon beta-regulated genes in multiple sclerosis. *J Neuroimmunol* 139: 109-18
195. Sturzebecher S, Wandinger KP, Rosenwald A, et al. 2003. Expression profiling identifies responder and non-responder phenotypes to interferon-beta in multiple sclerosis. *Brain* 126: 1419-29
196. Miller SA, Dykes DD, Polesky HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16: 1215
197. Wei J, Hemmings GP. 2000. The NOTCH4 locus is associated with susceptibility to schizophrenia. *Nat Genet* 25: 376-7.
198. Lucotte G, Bathelier C, Mercier G. 2000. TNF-alpha polymorphisms in multiple sclerosis: no association with -238 and -308 promoter alleles, but the microsatellite allele a11 is associated with the disease in French patients. *Mult Scler* 6: 78-80
199. Mycko M, Kowalski W, Kwinkowski M, et al. 1998. Multiple sclerosis: the frequency of allelic forms of tumor necrosis factor and lymphotoxin-alpha. *J Neuroimmunol* 84: 198-206
200. Braun N, Michel U, Ernst BP, et al. 1996. Gene polymorphism at position -308 of the tumor-necrosis-factor-alpha (TNF-alpha) in multiple sclerosis and it's influence on the regulation of TNF-alpha production. *Neurosci Lett* 215: 75-8
201. He B, Navikas V, Lundahl J, et al. 1995. Tumor necrosis factor alpha-308 alleles in multiple sclerosis and optic neuritis. *J Neuroimmunol* 63: 143-7
202. Huizinga TW, Westendorp RG, Bollen EL, et al. 1997. TNF-alpha promoter polymorphisms, production and susceptibility to multiple sclerosis in different groups of patients. *J Neuroimmunol* 72: 149-53
203. Fernandez-Arquero M, Arroyo R, Rubio A, et al. 1999. Primary association of a TNF gene polymorphism with susceptibility to multiple sclerosis. *Neurology* 53: 1361-3
204. Ronaghi M, Uhlen M, Nyren P. 1998. A sequencing method based on real-time pyrophosphate. *Science* 281: 363, 5.
205. Sham PC, Curtis D. 1995. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet* 59: 97-105.
206. Svejgaard A, Ryder LP. 1994. HLA and disease associations: detecting the strongest association. *Tissue Antigens* 43: 18-27.
207. Zhao JH, Curtis D, Sham PC. 2000. Model-free analysis and permutation tests for allelic associations. *Hum Hered* 50: 133-9.
208. Swanberg M, Lidman O, Padyukov L, et al. 2005. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. In *Nat Genet*, pp. 486-94
209. Lundberg C, Lidman O, Holmdahl R, et al. 2001. Neurodegeneration and glial activation patterns after mechanical nerve injury are differentially regulated by non-MHC genes in congenic inbred rat strains. In *J Comp Neurol*, pp. 75-87.

210. Guillery RW, Herrup K. 1997. Quantification without pontification: choosing a method for counting objects in sectioned tissues. In *J Comp Neurol*, pp. 2-7.
211. Jolliffe IT. 2002. Principal component analysis. In *Springer series in statistics*. New York: Springer
212. Hammarberg H, Wallquist W, Piehl F, et al. 2000. Regulation of laminin-associated integrin subunit mRNAs in rat spinal motoneurons during postnatal development and after axonal injury. In *J Comp Neurol*, pp. 294-304.
213. Diez M, Danner S, Frey P, et al. 2003. Neuropeptide alterations in the hippocampal formation and cortex of transgenic mice overexpressing beta-amyloid precursor protein (APP) with the Swedish double mutation (APP23). In *Neurobiol Dis*, pp. 579-94
214. Glatt SJ, Wang RS, Yeh YC, et al. 2005. Five NOTCH4 polymorphisms show weak evidence for association with schizophrenia: evidence from meta-analyses. *Schizophr Res* 73: 281-90
215. Hauser SL. 1995. Tumor necrosis factor: immunogenetics and disease. *Ann Neurol* 38: 702-4
216. Lang HL, Jacobsen H, Ikemizu S, et al. 2002. A functional and structural basis for TCR cross-reactivity in multiple sclerosis. *Nat Immunol* 3: 940-3
217. Becker KG, Simon RM, Bailey-Wilson JE, et al. 1998. Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc Natl Acad Sci U S A* 95: 9979-84.
218. Oksenberg JR, Baranzini SE, Barcellos LF, Hauser SL. 2001. Multiple sclerosis: genomic rewards. *J Neuroimmunol* 113: 171-84.
219. Feakes R, Sawcer S, Broadley S, et al. 2000. Interleukin 1 receptor antagonist (IL-1ra) in multiple sclerosis. *J Neuroimmunol* 105: 96-101.
220. Schrijver HM, Crusius JB, Uitdehaag BM, et al. 1999. Association of interleukin-1beta and interleukin-1 receptor antagonist genes with disease severity in MS. *Neurology* 52: 595-9.
221. Hensiek AE, Roxburgh R, Meranian M, et al. 2003. Osteopontin gene and clinical severity of multiple sclerosis. *J Neurol* 250: 943-7.
222. Kotze MJ, de Villiers JN, Rooney RN, et al. 2001. Analysis of the NRAMP1 gene implicated in iron transport: association with multiple sclerosis and age effects. *Blood Cells Mol Dis* 27: 44-53.
223. Giedraitis V, He B, Huang WX, Hillert J. 2001. Cloning and mutation analysis of the human IL-18 promoter: a possible role of polymorphisms in expression regulation. *J Neuroimmunol* 112: 146-52.
224. Fry TJ, Mackall CL. 2002. Interleukin-7: from bench to clinic. *Blood* 99: 3892-904.
225. Ye SK, Maki K, Kitamura T, et al. 1999. Induction of germline transcription in the TCRgamma locus by Stat5: implications for accessibility control by the IL-7 receptor. *Immunity* 11: 213-23.
226. Lai SY, Molden J, Goldsmith MA, et al. 1997. Shared gamma(c) subunit within the human interleukin-7 receptor complex. A molecular basis for the pathogenesis of X-linked severe combined immunodeficiency. *J Clin Invest* 99: 169-77.
227. Ioannidis JP, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG. 2003. Genetic associations in large versus small studies: an empirical assessment. *Lancet* 361: 567-71
228. Booth DR, Arthur AT, Teutsch SM, et al. 2005. Gene expression and genotyping studies implicate the interleukin 7 receptor in the pathogenesis of primary progressive multiple sclerosis. *J Mol Med* 83: 822-30
229. Matza D, Kerem A, Shachar I. 2003. Invariant chain, a chain of command. *Trends Immunol* 24: 264-8
230. Tompkins SM, Padilla J, Dal Canto MC, et al. 2002. De novo central nervous system processing of myelin antigen is required for the initiation of experimental autoimmune encephalomyelitis. *J Immunol* 168: 4173-83
231. West AK, Chuah MI, Vickers JC, Chung RS. 2004. Protective role of metallothioneins in the injured mammalian brain. *Rev Neurosci* 15: 157-66

232. Espejo C, Carrasco J, Hidalgo J, et al. 2001. Differential expression of metallothioneins in the CNS of mice with experimental autoimmune encephalomyelitis. *Neuroscience* 105: 1055-65
233. Penkowa M, Espejo C, Martinez-Caceres EM, et al. 2001. Altered inflammatory response and increased neurodegeneration in metallothionein I+II deficient mice during experimental autoimmune encephalomyelitis. *J Neuroimmunol* 119: 248-60
234. Penkowa M, Espejo C, Ortega-Aznar A, et al. 2003. Metallothionein expression in the central nervous system of multiple sclerosis patients. *Cell Mol Life Sci* 60: 1258-66
235. Stennard FA, Holloway AF, Hamilton J, West AK. 1994. Characterisation of six additional human metallothionein genes. *Biochim Biophys Acta* 1218: 357-65
236. Lohmueller KE, Pearce CL, Pike M, et al. 2003. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 33: 177-82