RESEARCH

Anu Kemppinen

Studies on Causes of Multiple Sclerosis:
From Genes to Transcriptome

RESEARCH 65

Anu Kemppinen

Studies on Causes of Multiple Sclerosis:

From Genes to Transcriptome

ACADEMIC DISSERTATION

NATIONAL INSTITUTE
FOR HEALTH AND WELFARE

National Institute for Health and Welfare.

Cover graphic: "*One in 999*", Anu Kemppinen

Supervised by

Academician of Science, Professor Leena Peltonen-Palotie †
Institute for Molecular Medicine Finland (FIMM), University of Helsinki,
Helsinki, Finland
National Institute for Health and Welfare, Helsinki, Finland
Department of Medical Genetics, Haartman Institute, University of Helsinki,
Helsinki, Finland
Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

and

Adjunct Professor Janna Saarela
Institute for Molecular Medicine Finland (FIMM), University of Helsinki,
Helsinki, Finland

Reviewed by

Adjunct Professor Minna Männikkö
Department of Medical Biochemistry and Molecular Biology,
University of Oulu,
Oulu, Finland

and

Adjunct Professor Sakari Jokiranta
Department of Bacteriology and Immunology,
Haartman Institute,
University of Helsinki,
Helsinki, Finland

Opponent

Professor Rikard Holmdahl
Department of Medical Biochemistry and Biophysics,
Karolinska Institutet,
Stockholm, Sweden

†Academician, Professor Leena Peltonen-Palotie deceased in March 2010

In Memoriam of Leena

Abstract

Multiple sclerosis (MS) is the most common cause of neurological disability in young adults, affecting more than two million people worldwide. It manifests as a chronic inflammation in the central nervous system (CNS) and causes demyelination and neurodegeneration. Depending on the location of the demyelinated plaques and axonal loss, a variety of symptoms can be observed including deficits in vision, coordination, balance and movement. With a typical age of onset at 20-40 years, the social and economic impacts of MS on lives of the patients and their families are considerable. Unfortunately the current treatments are relatively inefficient and the development of more effective treatments has been impeded by our limited understanding of the causes and pathogenesis of MS.

Risk of MS is higher in biological relatives of MS patients than in the general population. Twin and adoption studies have shown that familial clustering of MS is explained by shared genetic factors rather than by shared familial environment. While the involvement of the human leukocyte antigen (HLA) genes was first discovered four decades ago, additional genetic risk factors have only recently been identified through genome-wide association studies (GWAS). Current evidence suggests that MS is a highly polygenic disease with perhaps hundreds of common variants with relatively modest effects contributing to susceptibility. Despite extensive research, the majority of these risk factors still remain to be identified.

In this thesis the aim was to identify novel genes and pathways involved in MS. Using genome-wide microarray technology, gene expression levels in peripheral blood mononuclear cells (PBMC) from 12 MS patients and 15 controls were profiled and more than 600 genes with altered expression in MS were identified. Three of five selected findings, *DEFA1A3*, *LILRA4* and *TNFRSF25*, were successfully replicated in an independent sample. Increased expression of *DEFA1A3* in MS is a particularly interesting observation, because its elevated levels have previously been reported also in several other autoimmune diseases.

A systematic review of seven microarray studies was then performed leading to identification of 229 genes, in which either decreased or increased expression in MS had been reported in at least two studies. In general there was relatively little overlap across the experiments: 11 of the 229 genes had been reported in three studies and only *HSPA1A* in four studies. Nevertheless, these 229 genes were associated with several immunological pathways including interleukin pathways related to type 2 and type 17 helper T cells and regulatory T cells. However, whether these pathways are involved in causing MS or related to secondary processes activated after disease onset remains to be investigated. The 229 genes were also compared with loci

identified in published MS GWASs. Single nucleotide polymorphisms (SNP) in 17 of the 229 loci had been reported to be associated with MS with P-value <0.0001 including variants in *CXCR4* and *SAPS2*, which were the only loci where evidence for correlation between the associated variant and gene expression was found. The *CXCR4* variant was further tested for association with MS in a large case-control sample and the previously reported suggestive association was replicated (P-value =0.0004).

Finally, common genetic variants in candidate genes, which had been selected on the basis of showing association with other autoimmune diseases (*MYO9B*) or showing differential expression in MS in our study (*DEFA1A3*, *LILRA4* and *TNFRSF25*), were tested for association with MS, but no evidence of association was found.

In conclusion, through a systematic review of genome-wide expression studies in MS we have identified several promising candidate genes and pathways for future studies. In addition, we have replicated a previously suggested association of a SNP variant upstream of *CXCR4* with MS.

Keywords: autoimmune disease, common variant, *CXCR4*, *DEFA1A3*, *HSPA1A*, gene expression, genetic association, GWAS, MS, multiple sclerosis, systematic review

Tiivistelmä

Pesäkekovettumatauti eli multippeli skleroosi (MS) on nuorten aikuisten yleisin neurologisiin häiriöihin johtava syy. Maailmanlaajuisesti sitä sairastaa yli kaksi miljoonaa ihmistä. Tauti ilmenee keskushermoston kroonisena tulehduksena ja johtaa myeliinikatoon ja keskushermoston rappeutumiseen. Oireet riippuvat leesioiden sijainnista keskushermostossa ja saattavat kohdistua esimerkiksi näköaistiin, tasapainoon ja koordinaatio- ja liikuntakykyyn. MS-tauti puhkeaa yleensa 20-40 vuoden iässä, joten sen sosiaaliset ja ekonomiset vaikutukset potilaisiin ja näiden lähimmäisiin ovat merkittävät. Tehokkaiden hoitojen kehittämistä on vaikeuttanut se, että taudin syyt ja mekanismit tunnetaan toistaiseksi huonosti.

Todennäköisyys sairastua MS-tautiin on suurempi MS-potilaiden biologisilla sukulaisilla kuin populaatiossa keskimäärin. Kaksos- ja adoptiotutkimuksissa on osoitettu, että taudin keskittyminen perheisiin selittyy jaetuilla perinnöllisillä alttiustekijöillä. Vaikka immunologisiin soluihin liittyvien HLA-varianttien yhteys MS-tautiin osoitettiin jo lähes neljä vuosikymmentä sitten, uusia geneettisiä alttiustekijöitä on onnistuttu paikantamaan vasta viime vuosina koko genomin laajuisten assosiaatiotutkimusten avulla. Nykyisten tutkimusten valossa MS-alttiuteen vaikuttavia geenivariantteja näyttäisi kuitenkin olevan jopa satoja ja suurin osa näistä on edelleen tunnistamatta.

Tämän väitöskirjan tavoitteena oli tunnistaa uusia MS-tautiin liittyviä geenejä ja reaktioreittejä. Geenien ilmentymistasoja verrattiin ensin periferaalisissa mononukleaarisissa immuunisoluissa MS-potilaiden ja verrokkien välillä käyttäen mikrosirutekniikkaa, joka mahdollistaa kaikkien tunnettujen geenien samanaikaisen tarkastelun. Yli 600 geenin ilmentymisen havaittiin poikkeavan merkitsevästi MS-potilaiden ja verrokkien välillä. Tunnistetuista geeneistä viisi valittiin toistotutkimukseen, jossa kolmen geenin (*DEFA1A3*, *LILRA4* ja *TNFRSF25*) tapauksessa ilmentymisero potilaiden ja verrokkien välillä toistui. Näistä geeneistä *DEFA1A3* on erityisen kiinnostava, sillä sen ilmentymistason on havaittu olevan verrokkeja korkeampi myös muissa autoimmuunisairauksissa.

Koska vastaavia mikosirututkimuksia on tehty MS-taudissa myös aiemmin, yhtenä väitöskirjan tavoitteena oli verrata tuloksia näistä tutkimuksista. Vertailuun otettiin kaikkiaan seitsemän tutkimusta, mukaan lukien tämän väitöskirjan osana tehty tutkimus. Tavoitteena oli tunnistaa geenit, joissa ilmentymisen oli raportoitu poikkeavan MS-potilaiden ja verrokkien välillä ainakin kahdessa tutkimuksessa. Vaikka tällaisia geenejä tunnistettiin 229, yleisesti ottaen tutkimustulosten välillä oli

kuitenkin vähän päällekkäisyyttä: vain 11 geeniä oli raportoitu kolmessa tutkimuksessa ja ainoastaan yksi, *HSPA1A*, neljässä tutkimuksessa. Reaktioreittianalyysin avulla osoitettiin, että tunnistetut 229 geeniä assosioituvat merkittävästi useisiin mielenkiintoisiin immunologisiin reaktioreitteihin. Monet näistä liittyvät tiettyihin T-solutyyppeihin, joiden merkitystä sekä MS-taudin alttiuden että taudin etenemisen ja patogeneesin kannalta on tulostemme perusteella syytä tutkia lisää.

229 geenistä ainoastaan 17 on aiemmin raportoitu assosioituvan MS-tautiin, mikä viittaisi siihen että nämä 229 geeniä eivät välttämättä ole merkittäviä taudin alttiuden lisäämisen kannalta, mutta voivat liittyä taudin mekanismeihin taudin puhkeamisen jälkeen. 17 geenistä ainoastaan *SAPS2-* ja *CXCR4*-geenien tapauksessa havaitsimme yhteyden geenin ilmentymisen ja geneettisen riskivariantin välillä siten, että molemmissa tapauksissa geenin ilmentyminen oli koholla riskialleelin kantajilla.

Lopuksi väitöskirjassa testattiin tämän *CXCR4*-lokuksessa sijaitsevan variantin assosiaatiota MS-tautiin tuhansien henkilöiden potilas-kontrolli aineistossa ja aiemmissa tutkimuksissa havaittu assosiaatio toistettiin (P-arvo =0.0004). Lisäksi tutkittiin valituissa ehdokasgeeneissä sijaitsevien yleisten geneettisten varianttien yhteyttä MS-tautiin. Edellä mikrosirututkimuksessa tunnistettujen *DEFA1A3-*, *LILRA4-* ja *TNFRSF25*-geenien lisäksi tarkasteltiin *MYO9B*-geeniä, jonka oli aiemmin havaittu assosioituvan muihin autoimmuunisairauksiin. Merkitsevää yhteyttä MS-taudin ja tarkasteltujen geneettisten varianttien välillä ei kuitenkaan havaittu.

Yhteenvetona, tässä väitöskirjassa on geenien ilmentymistutkimusten ja niiden systemaattisen katsauksen avulla tunnistettu useita lupaavia ehdokasgeenejä (mm. *DEFA1A3*, *LILRA4*, *TNFRSF25* ja *HSPA1A*) ja erityisesti T-soluihin liittyviä reaktioreittejä. Lisäksi väitöskirjassa on toistettu aiemmin raportoitu assosiaatio MS-taudin ja *CXCR4*-lokuksessa sijaitsevan geneettisen variantin välillä. Tunnistettujen geenien rooli MS-taudin patogeneesissä on kuitenkin toistaiseksi tuntematon, joten tarkempia jatkotutkimuksia tarvitaan.

Avainsanat: *CXCR4, DEFA1A3*, *HSPA1A*, geenien ilmentyminen, geneettinen assosiaatio, GWAS, MS-tauti, multippeli skleroosi, pesäkekovettumatauti, yleinen variantti

Table of contents

List of original papers

This thesis is based on the following original articles referred to in the text by their Roman numerals. In addition, some previously unpublished data are presented.

I   **Kemppinen A**, Suvela M, Tienari PJ, Elovaara I, Koivisto K, Pirttilä T, Reunanen M, Rautakorpi I, Hillert J, Lundmark F, Oturai A, Ryder L, Harbo HF, Celius EG, Palotie A, Peltonen L, Saarela J (2009). MYO9B polymorphisms in multiple sclerosis. *European Journal of Human Genetics* 17: 840-3.

II   **Kemppinen A**, Kaprio J, Palotie A, Saarela J (2011). Systematic review of genome-wide expression studies in multiple sclerosis. *BMJ Open* doi:10.1136/bmjopen-2011-000053.

III **Kemppinen A**, Leppä V, Harbo HF, Celius EG, Kockum I, Olsson T, Hillert J, Oturai A, Sawcer S, Saarela J. Replicating association of a variant in *CXCR4* locus with multiple sclerosis. Manuscript soon to be submitted.

Abbreviations

AA	Alopecia areata
APC	Antigen presenting cell
BBB	Blood-brain barrier
BBC	Biomedicum Biochip Center
bp	Base pair
CD	Celiac disease
cDNA	Complementary DNA
CEPH	Centre d'Etude du Polymorphisme Humain
ChD	Crohn's disease
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel test
CNS	Central nervous system
CNV	Copy number variant
CSF	Cerebrospinal fluid
Ct	Threshold cycle
DC	Dendritic cell
DEG	Differentially expressed gene
DNA	Deoxyribonucleic acid
DZ	Dizygotic
EAE	Experimental autoimmune encephalomyelitis
EBV	Epstein-Barr virus
EDSS	Expanded Disability Status Scale
eQTL	Expression quantitative trait locus
FITSA	Finnish Twin Study on Ageing
GC-RMA	GC-Robust Multi-Array normalization
GEO	Gene Expression Omnibus
GWAS	Genome-wide association study
GWS	Genome-wide significance (P-value $\leq 5 \times 10^{-08}$)
HGP	Human Genome Project
HLA	Human leukocyte antigen
IBD	Inflammatory bowel disease
IFN	Interferon
IMSGC	International Multiple Sclerosis Genetics Consortium
INDEL	Insertion-deletion polymorphism
IPA	Ingenuity Pathway Analysis
LD	Linkage disequilibrium
LOD	Logarithm of odds
kb	Kilobase
MAF	Minor allele frequency

Mb	Megabase
MBP	Myelin basic protein
mDC	Myeloid dendritic cell
MHC	Major Histocompatibility Complex
MHCI	Major Histocompatibility Complex Class I
MHCII	Major Histocompatibility Complex Class II
MM	Mismatch
MOG	Myelin oligodendrocyte protein
MRI	Magnetic resonance imaging
MSSS	Multiple Sclerosis Severity Score
MZ	Monozygotic
NCBI	National Center for Biotechnology Information
NHGRI	National Human Genome Research Institute
NK cell	Natural killer cell
OR	Odds ratio
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
pDC	Plasmacytoid dendritic cell
PLP	Proteolipid protein
PM	Perfect match
PML	Progressive multifocal leukoencephalopathy
PPMS	Primary-progressive MS
PS	Psoriasis
PSA	Psoriatic arthritis
RA	Rheumatoid arthritis
RAF	Risk allele frequency
RMA	Robust Multi-Array
RNA	Ribonucleic acid
RNA-Seq	RNA-sequencing
RRMS	Relapsing-remitting MS
SAGE	Serial Analysis of Gene Expression
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
SPMS	Secondary-progressive MS
SSR	Single sequence repeat
T1D	Type I diabetes
TCR	T cell receptor
TDT	Transmission disequilibrium test
Th	T helper
Treg	Regulatory T cell
UCSC	University of California, Santa Cruz
UTR	Untranslated region

VNTR   Variable number of tandem repeats
WTCCC   Wellcome Trust Case Control Consortium

1 Introduction

Multiple sclerosis (MS) is the most common cause of neurological disability in young people, affecting approximately 0.1% of individuals in populations of Northern European ancestry (Sadovnick and Ebers 1993, Dean 1994). Although the range in age of onset is broad, there is a peak at 20-40 years and MS therefore often has major economic and social impacts on lives of the patients and their families. MS manifests in the central nervous system (CNS) as demyelinated lesions in the myelin sheath and is also characterized by axonal loss. Myelin, or white matter, surrounds and protects the axons and is required for efficient transmission of nerve impulses. As a result of demyelination, scar tissue ("sclerosis") is formed at various places in CNS ("multiple"). Depending on the locations of the lesions and axonal loss, a variety of symptoms can be observed (Figure 1).

Currently available MS treatments can only slow the progression of the disease and are not effective in all patients. More effective treatments are therefore needed, but their development has been hindered by poor understanding of MS pathogenesis. Twin and adoption studies have provided evidence for a significant genetic component in MS predisposition and identifying genetic factors involved could therefore increase the understanding of underlying disease processes. Furthermore, as the environmental factors contributing to MS remain unknown, genetic research is perhaps more critical in MS than in complex disorders where controllable life style factors such as diet are known to play a significant role. However, as in other complex disorders, identifying MS predisposing genetic variants has proven to be a challenging task and for nearly four decades the Major Histocompatibility Complex (MHC) locus remained the only well-established genetic finding.

At the time of starting this thesis in 2005, researchers were hopeful that association studies on selected candidate genes would detect common variants predisposing to complex diseases. This was also one of the approaches taken in this thesis. However, six years later it is easy to conclude that most of these early candidate gene studies had a low chance of succeeding for two main reasons. First, picking candidate genes is challenging when the disease mechanisms are poorly understood. This issue has been resolved thanks to advancements in genotyping technologies, which now allow high resolution genome-wide association studies (GWAS) of common single nucleotide polymorphism (SNP) variants at a manageable cost. Second, common variants have proven to have lower effects on disease risk than perhaps initially was thought. Consequently extensive sample sizes are required to achieve sufficient power for detecting associations. International collaborations have therefore become a necessity, with possibly one of the most impressive examples in complex diseases

being the International Multiple Sclerosis Genetics Consortium (IMSGC), which now has samples from ~20,000 MS patients from fourteen countries. As a result of the aforementioned technical advancements and collaborative efforts close to 50 confirmed MS risk loci have been identified over the last four years.

But despite the rapid recent progress, the genetic architecture of MS is far from unraveled. Even with the now long list of confirmed risk variants, much of the heritability of MS still remains unexplained. As the costs of next generation sequencing technologies are decreasing, the focus in the next years is likely to be on rare variants with potentially larger individual effects. In addition, most of the associated loci extend over several genes and further genetic and functional studies are needed in order to identify the real culprits in these regions, and to explore how they actually cause increased risk for MS. Since many risk variants are likely to affect gene expression studies, correlating genetic variants with expression of their nearby genes is one way of identifying the most likely disease-related genes. Genome-wide expression profiling and comparison with GWAS findings was also one of the approaches taken in this thesis.
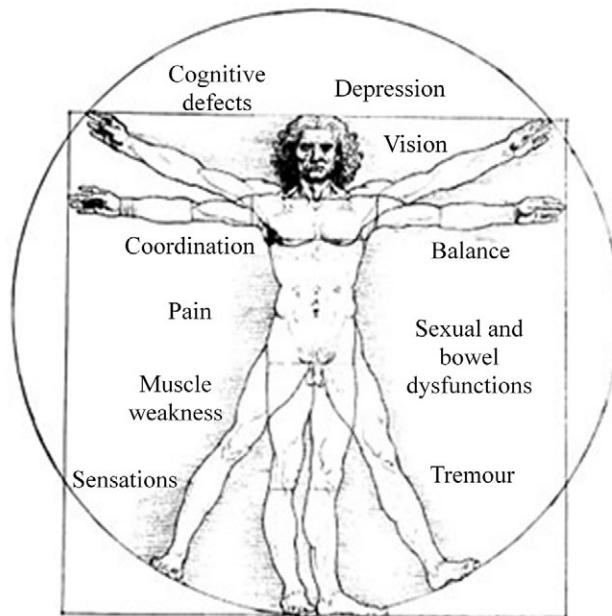
Figure 1. Common symptoms and functions affected in MS. Leonardo Da Vinci's "Vitruvian man" modified from a figure from Wikimedia commons.

2 Review of the literature

2.1 Multiple sclerosis

2.1.1 History, diagnosis and treatment

The first documentation of MS may date back to a medieval Norse saga from the end of 11th century, which describes a woman, Halladora, with symptoms reminiscent of MS (Poser 1995, Holmøy 2006). A Scottish physician and artist Robert Carswell was knowingly the first to describe and illustrate spinal cord lesions in a post-mortem subject who was described as having been paralysed (Carswell 1838), likely as a consequence of MS. The clinical features and pathology of MS (French sclérose en plaques) were, however, first systematically and comprehensively described by a French neurologist, Jean-Martin Charcot in the 19th century (Charcot 1865, 1868a, 1868b, 1868c). Today, the diagnosis of MS is based on Poser's criteria (Poser *et al.* 1983) or McDonald's criteria (McDonald *et al.* 2001, Polman *et al.* 2005), the latter of which incorporated direct detection of lesions through magnetic resonance imaging (MRI) (Table 1).

Majority of patients initially present with a relapsing-remitting MS (RRMS), in which relapses are followed by partial or complete recovery. Around 65% of RRMS patients later proceed to secondary progressive MS (SPMS), in which the condition is chronic and deteriorating without intervening remission periods (Compston and Coles 2002). Around 20% of all MS cases show continuous progression from the onset, displaying primary-progressive MS (PPMS) (Compston and Coles 2002). The severity of MS is commonly measured on the Expanded Disability Status Scale (EDSS) from 0.0-10.0 (Kurtzke 1983). However, the EDSS score alone is not informative of the rate of progression. Multiple sclerosis severity score (MSSS), which is based on the EDSS score but also takes into account the time from onset, was therefore developed and can be used to predict the patient's approximate disability level after 30 years with MS (Roxburgh *et al.* 2005).

Current MS treatments, such as interferon (IFN)-beta, natalizumab, glatiramer acetate and mitoxantrone, are immunomodulatory; they can decrease the inflammatory reaction, but do not directly protect against demyelination or neurodegeneration. They are also inefficient in treating progressive MS. Encouraging results in clinical trials have been obtained with rituximab, which targets B-cells (Chaudhuri and Behan 2008, De Palma and Sementa 2008, Schrijver 2008), and alemtuzumab (CAMMS223 Trial Investigators *et al.* 2008), which is an

antibody against CD52, a surface protein present on lymphocytes and monocytes. However, both can have serious side effects. Some patients treated with rituximab have become affected with a usually fatal demyelinating disease, progressive multifocal leukoencephalopathy (PML) (Freim Wahl *et al.* 2007). After alemtuzumab treatment infection-associated events were reported in 66% of the patients, while autoimmune-related thyroid disorders and immune thrombocytopenic purpura developed in 23% and 3% of patients, respectively (CAMMS223 Trial Investigators *et al.* 2008).

Table 1. McDonald's criteria for MS diagnosis. CSF=cerebrospinal fluid. VPE=visual evoked potential. Positive CSF refers to presence of oligoclonal immunoglobulin bands in CSF. While oligoclonal bands are characteristic for MS, their relevance for MS pathogenesis is not understood. Delayed VPE, i.e. delayed neurological response to a visual signal, is typical for MS.

Clinical Presentation	Additional Data Needed
Two or more attacks and objective clinical evidence of two or more lesions separated in time and necessarily separated in space	None
Two or more attacks and objective clinical evidence of one lesion	Dissemination in space demonstrated by MRI *or* two or more MRI-detected lesions consistent with MS plus positive CSF *or* await further clinical attack implicating a different site
One attack and objective clinical evidence of two or more lesions	Dissemination in time demonstrated by MRI *or* second clinical attack
One attack and objective clinical evidence of one lesion (clinically isolated syndrome)	Dissemination in space demonstrated by MRI *or* two or more MRI-detected lesions consistent with MS plus positive CSF *and* dissemination in time demonstrated by MRI *or* second clinical attack

2.1.2 Pathogenesis

Role of autoimmunity

Although direct evidence is lacking, MS is commonly considered to be an autoimmune disease. Myelin proteins (myelin basic protein (MBP), myelin oligodendrocyte protein (MOG), proteolipid protein (PLP)) are obvious candidates for auto-antigens in MS, but their role as such has not been unequivocally demonstrated. However, MS displays characteristics typical for autoimmune disorders such as having a higher prevalence in females and showing strong linkage and association with the MHC locus. Although the precise role of MHC in these diseases is not understood, MHC encodes several molecules involved in antigen presentation in the immune system and this process is therefore likely to be of importance in development of autoimmunity in general. There is also evidence for a higher risk of other autoimmune diseases in first degree relatives of MS patients (Broadley et al. 2000). Finally, many of the non-MHC genetic loci associated with MS are also associated with other autoimmune diseases (see 2.3.3).

Studies in animal model

According to the prevailing theory MS is mediated by T helper (Th) cells, which are CD4+ T cells that primarily activate other immune cells by producing various signaling molecules such as cytokines and chemokines. This hypothesis is largely based on studies in the animal model of MS, experimental autoimmune encephalomyelitis (EAE), which can be induced in rodents and primates by injection of myelin components or by adoptive transfer of activated myelin-recognizing T cells to an unaffected animal (Gold *et al.* 2006). However, while there are clear similarities between EAE and MS, with both demonstrating CNS inflammation and demyelination, it is not clear how well the pathogenesis of induced EAE corresponds to that of spontaneously developing and pathologically more heterogeneous MS (Gold *et al.* 2006). In addition, there is no EAE model that shows a similar disease course to PPMS.

Activation of immune response

Evidence from rats suggests that only activated T cells can cross the normally impermissible blood-brain barrier (BBB) and gain access to CNS (Hickey *et al.* 1991). However, it is not clear how and where the activation of auto-reactive T cells occurs in MS patients. It has been shown in mice that T cells can be activated in

Studies on causes of
multiple sclerosis

peripheral lymphoid organs by myelin-presenting antigen presenting cells (APC), specifically dendritic cells (DC), which are able to migrate out of the CNS (Karman *et al.* 2004). Another theory is that some viral or bacterial antigens with similarity to myelin antigens activate myelin-specific T cells (Wucherpfennig and Strominger 1995). However, due to lack of convincing evidence of association between MS and any specific pathogen, this molecular mimicry theory remains controversial (Libbey *et al.* 2007). Finally, although it seems to be the general view that T cells are activated prior to entry into CNS, it has been shown in EAE that also naive lymphocytes are able to gain access to CNS if it is already inflamed (McMahon *et al.* 2005).

Once lymphocytes have accessed CNS, auto-reactive T cells are thought to become re-activated upon encountering endogenous myelin antigens presented to them by APCs. Re-activated T cells begin to produce chemokines and cytokines which activate CNS cells such as microglia and astrocytes, and promote recruitment of other immune cells from peripheral blood, including CD8+ T-cells, B-cells, mast cells, monocytes and macrophages (Sospedra and Martin 2005) (Figure 2.). Their potential roles in MS pathogenesis are covered in more detail in 2.1.3. These cells further produce various pro-inflammatory molecules, complement proteins and nitrogen and oxygen radicals, although the relevance of these molecules in MS pathogenesis is not understood in detail (Sospedra and Martin 2005).

Demyelination

The direct cause of demyelination in MS is still unclear. Activation of the complement system, which is a central mechanism in clearing pathogens, may be responsible for myelin destruction as suggested by presence of complement proteins in demyelinated MS lesions (Compston *et al.* 1989, Storch *et al.* 1998, Ingram *et al.* 2009). Also nitric oxide has been shown to be toxic to myelin producing cells, oligodendrocytes, *in vitro* (Mitrovic *et al.* 1995).

Nevertheless, it has been proposed that as myelin becomes destructed, additional myelin peptides become exposed and auto-reactive T cells recognizing these particular epitopes (i.e. parts of antigens recognized by immune cells) are activated. This epitope spreading phenomena, where an initial response to one epitope results in tissue destruction and thereby in release of new epitopes and activation of cells recognizing these epitopes, could be responsible for chronicity of the inflammatory reaction in MS (Tuohy and Kinkel 2000).
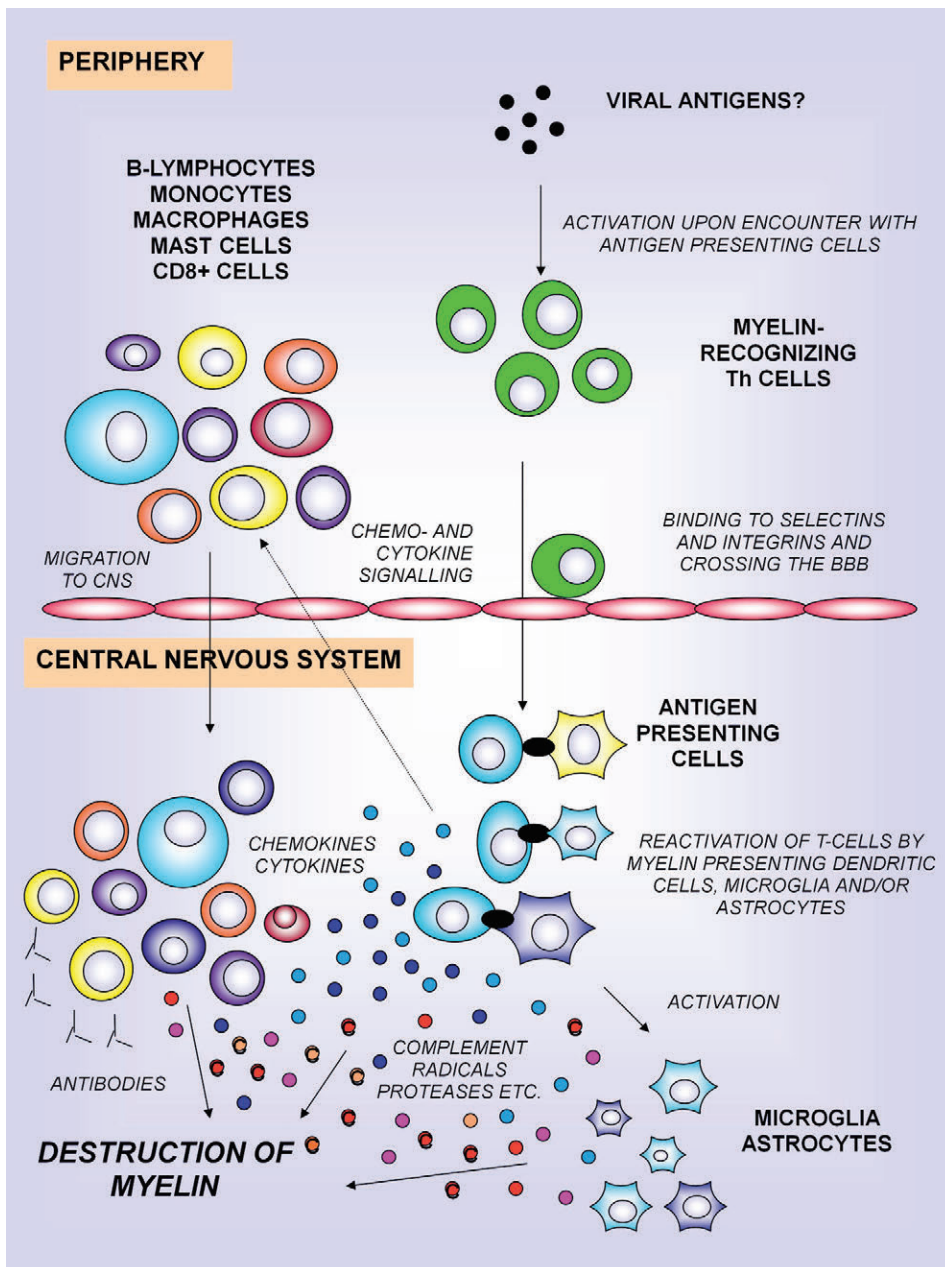
Figure 2. Hypothetic scenario of events leading to demyelination in MS. Based on text and figures in a review by Sospedra and Martin (2005).

2.1.3 Role of different cells in MS pathogenesis

Although the critical first event in MS pathogenesis is likely to be the activation of auto-reactive Th cells by DCs or other APCs expressing MHC, also other cells are likely to play important roles in MS pathogenesis. Some of these and their functions are described below. The immunological cell lineages are depicted in Figure 3.
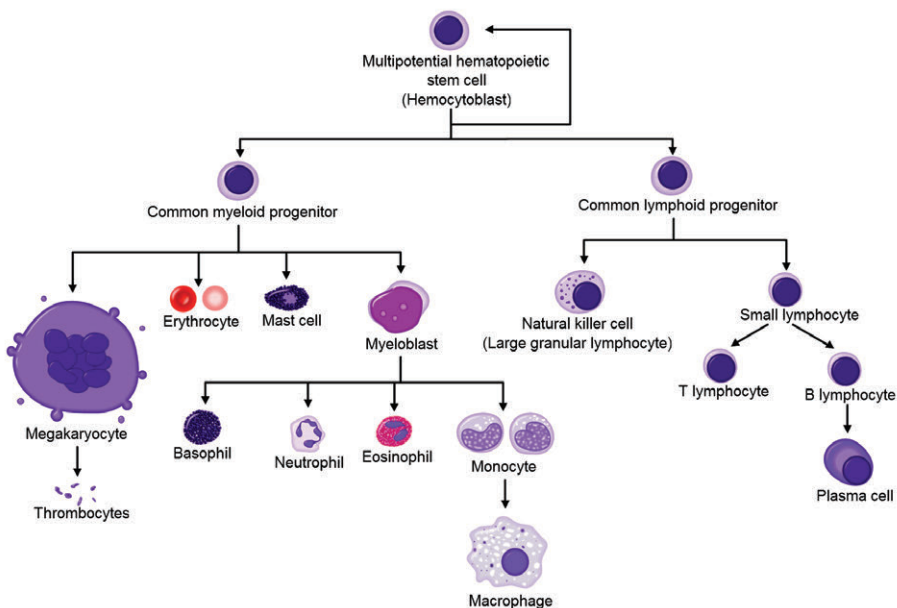
Figure 3. Lineages of immunological cells. Dendritic cells are not shown, but can develop from both myeloid and lymphoid lineage progenitors. Figure from Wikimedia Commons by Mikael Häggström from original by A. Rad.

Astrocytes

Astrocytes are abundant CNS glial cells which provide support for neurons and maintain and regulate BBB. Based on studies in both MS patients and EAE, there is evidence for the role of astrocytes in various processes: they can attract and activate

T cells by producing cytokines and chemokines, aid the access of peripheral cells through BBB and impede remyelination, but on the other hand have also been suggested to limit effects of proinflammatory factors and to provide support for oligodendrocytes (Nair *et al.* 2008). They also uptake glutamate, a neurotransmitter which is highly toxic to oligodendrocytes *in vitro* (Oka *et al.* 1993) and can thereby have a protective role for oligodendrocyte survival (Corley *et al.* 2001). Finally, astrocytes can also act as APCs and present MBP peptides to T cells in rats (Fontana *et al.* 1984). There are thereby several protective and promoting mechanisms by which astrocytes could contribute to MS pathogenesis.

B lymphocytes

Plasma cells, which are differentiated B cells, are specialised in antibody production. In MS, B cells and plasma cells are recruited from the peripheral blood to CNS where plasma cells produce antibodies presumably against some components of myelin. The production of antibodies is evidenced by oligoclonal immunoglobulin bands, which are characteristically detected in CSF of MS patients by protein electrophoresis. However, the antigen specificity of these antibodies and their relevance for MS pathogenesis remain unknown. In addition to being precursors for plasma cells, B cells may also have other important functions in MS such as acting as APCs and producing proinflammatory cytokines (McLaughlin and Wucherpfennig 2008). Successful clinical trials with rituximab which specifically targets CD20, a surface molecule expressed by naive, developing and memory B cells but not plasma cells, also suggests a role for B cells in MS pathogenesis (Bar-Or *et al.* 2008, Hauser *et al.* 2008).

CD4+ T cells

CD4+ T cells are a heterogeneous cell population including Th cells and regulatory T cells (Tregs). Three types of Th cells have now been characterised: Th1, Th2 and Th17 cells. The antigen specificity of Th cells is determined by the T cell receptor (TCR), which is expressed on cell surface. TCR recognizes its antigen when the antigen is presented in the context of Major Histocompatibility Complex Class II (MHCII) molecules on the surface of APCs such as DCs. Given that MHCII variants are a major risk factor for MS, antigen presentation and interactions between APCs and CD4+ T cells are likely to play an important role in MS. After antigen recognition, the naive T cells develop into mature Th1, Th2 or Th17 cells depending on cytokines present in the environment (Harrington *et al*. 2006). The different types of Th cells also have distinct cytokine expression profiles: among other cytokines Th1 cells are characterized by secretion of IFN-γ, Th2 cells by production of IL-4 and Th17 cells by production of cytokine IL-17A (Cherwinski *et al*. 1987,

Harrington *et al*. 2005, Park *et al*. 2005). Thereby the different Th cell types have different effects on other cells of the immune system.

Evidence from early studies in EAE suggested that Th1 cells are responsible for initiating the inflammatory response, while Th2 cytokines were regarded to act as counter-inflammatory by antagonising Th1 cells (Olsson 1992, Gold et al. 2006). The importance of Th1 cells was also supported by studies in human, as Th cells from CSF of MS patients were shown to produce IFN-γ (Benvenuto *et al*. 1992, Olsson 1992). In addition, increased Th1/Th2 balance has been associated with active inflammation in MS patients (Misu *et al.* 2001, Nakajima *et al.* 2004). On the other hand it has been shown that at least the major cytokine produced by Th1 cells, IFN-γ, is not required for development of EAE (Ferber *et al*. 1996). Finally, Th2 cells are also able to induce EAE under certain conditions (Lafaille *et al.* 1997).

Recently the role of Th17 cells has, however, been gaining increasing attention in studies of MS and other autoimmune diseases. Higher numbers of IL-17 producing cells have been reported in blood and CSF of MS patients (Matusevicius *et al.* 1999) and in EAE it was shown that a Th17-related cytokine IL-23, rather than IL-12 which drives Th1 differentiation, was required for development of EAE (Cua *et al*. 2003, Harrington *et al.* 2006). Various processes have been suggested through which these cells could contribute to MS, including affecting BBB breakdown, activation and recruitment of neutrophils and stimulation of production of proinflammatory molecules (Segal 2010).

Finally, unlike Th cells which have primarily pro-inflammatory effects, Tregs act in suppressing immune responses and are thereby important in controlling inflammation and in maintaining peripheral tolerance to auto-antigens. This has been demonstrated in mice, in which depletion of Tregs results in systemic autoimmunity (Sakaguchi 2000). Although there is no evidence for their decreased levels, Tregs isolated from MS patients show impaired suppressive functionality (Viglietta et al 2004).

CD8+ T cells

CD8+ T cells are a subset of T cells, also called cytotoxic T cells because of their capacity to induce cell death. Unlike CD4+ T cells which recognize MHCII molecules on the surface of APCs, CD8+ cells interact with Major Histocompatibility Complex Class I (MHCI) molecules which present primarily viral antigens. CD8+ T cells are found in MS lesions and are associated with axonal damage (Babbe *et al.* 2000). Myelin-specific CD8+ T cells are also able to induce CNS autoimmunity in mice and this condition was found to display MS-like features not observed in CD4+ cell -induced disease (Huseby *et al.* 2001).

Dendritic cells

DCs are professional APCs: their main function is to present antigens bound to MHC molecules on their cell surface to T cells. DCs can be divided into two main types: myeloid dendritic cells (mDC) and plasmacytoid dendritic cells (pDC), which differ with respect to their antigen uptake and cytokine secretion profiles (Shortman and Liu 2002). The potential role of DCs in MS is supported by findings that peripheral DCs, specifically mDCs, accumulate in CNS in EAE, acquire myelin debris and present it to CD4+ T-cells *in vivo* (Greter *et al.* 2005, Bailey *et al.* 2007). Furthermore, functional abnormalities and elevated CSF levels of pDCs have been reported in MS patients (Pashenkov *et al.* 2001, Stasiolek *et al.* 2006). In addition, IFN-beta and glatiramer acetate have been found to affect DC functions from MS patients *in vitro* (Hussien *et al.* 2001, Stasiolek *et al.* 2006).

Granulocytes

The role of granulocytes (eosinophils, basophils, neutrophils) in MS has so far been scarcely studied. However, there is evidence for elevated levels of neutrophils in peripheral blood of PPMS patients (McKay *et al.* 2008) and IFN-beta has been shown to prevent neutrophil infiltration to CNS in rat (Veldhuis *et al.* 2003), suggesting yet another potential mechanism by which IFN-beta treatment could ameliorate the clinical condition in MS. In addition, depletion of peripheral blood granulocytes delayed and in some cases completely prevented the development of EAE in mice (McColl *et al.* 1998).

Microglia and macrophages

Microglia are CNS glial cells, i.e. non-neuronal cells of the nervous system, whereas macrophages are found in peripheral blood and in MS are found to migrate to CNS. Both are critical players in innate immunity and have the ability to identify foreign agents and debris and phagocytose them. The functions of macrophages and microglia in MS and EAE have been reviewed by Raivich and Banati (2004). Both cells can act as APCs and interact with T cells in CNS. They also produce various substances such as cytokines, chemokines, complement components, cell adhesion glycoproteins, reactive oxygen species and neurotrophins, which may have both protective and damaging effects on neurons and oligodendrocytes.

Natural killer cells

Natural killer (NK) cells play an important role in causing death of virus-infected cells and cancer cells by producing and releasing toxic substances. While there is evidence for their involvement in MS from studies in both humans and EAE, the

results are somewhat controversial and suggest both protective and promoting mechanisms (Morandi *et al.* 2008). Proposed MS promoting mechanisms by NK cells include polarization of Th1 cells through cytokine release, direct damage to CNS tissue, and activation of DCs. Potential protective mechanisms include release of cytokines which induce Th2 cells or Tregs, killing of immature DCs or killing of myelin-specific T cells.

Oligodendrocytes

Oligodendrocytes, or oligodendroglia, are myelin-forming cells of CNS. They are normally capable of forming new myelin sheath around demyelinated axons in a process called remyelination. However, defects in remyelination in MS patients have been shown, possibly resulting from disabled proliferation and differentiation of oligodendrocyte precursor cells into mature oligodendrocytes (Wolswijk 1998).

2.2 Unravelling the genetics of complex diseases

2.2.1 Genetics of complex traits

Complex traits are phenotypes determined by multiple genetic and/or environmental factors and possibly also their interactions. While monogenic traits, which are controlled by single genes, show a Mendelian inheritance pattern within pedigrees (hence also called Mendelian traits), no obvious transmission pattern is seen in complex phenotypes. Although familial aggregation and correlation is commonly observed also in case of complex traits, this does not always imply that a significant genetic component is involved. The proportion of variation in a phenotype explained by genetic variation can be estimated with a parameter called heritability, h^2. Broad-sense heritability (h_b^2) is the ratio of total genetic variance to total phenotypic variance and can be estimated for example from concordance rates (r) in monozygotic (MZ) and dizygotic (DZ) twins according to: $h_b^2=2(r_{MZ}-r_{DZ})$ (Falconer and MacKay 1996). Narrow-sense heritability (h_n^2) is the proportion of phenotypic variance explained by additive genetic variance and hence does not take into account dominance and epistatic genetic effects, i.e. relationships between genetic variants within and between loci, respectively. Narrow-sense heritability can be estimated from concordances in siblings (r_{sib}) and adoptees (r_{ad}): $h_n^2 = r_{sib}-r_{ad}$.

There are many types of genetic variants, which can potentially explain the genetic component of a variable trait. Microsatellites, or single sequence repeats (SSR), are variants consisting of a short repeat sequence of 2-4 bp, while minisatellites

(variable number of tandem repeats, VNTR) are repeats of about 10-100 bp sequence (Ramel 1997). Allelic variants differ in the number of times this sequence is consecutively repeated in the locus. Micro- and minisatellites are generally highly mutable and thereby highly polymorphic. Microsatellites in particular have been found to cause rare disorders including Fragile X syndrome and Huntington's disease (Ramel 1997) and have also been useful in mapping of complex disease loci. It remains to be shown whether some of them also play a direct role in altering susceptibility to complex traits.

The first drafts of the human genome sequence were published in 2001 independently by the private company Celera Genomics (Venter *et al.* 2001) and the public-sector Human Genome Project (HGP) (Lander *et al.* 2001). Shortly after, HGP data together with data from the SNP Consortium project provided the first dense map of over 1.4 million SNPs (Sachidanandam *et al.* 2001). SNPs are allelic variants at a specific nucleotide position and are usually bi-allelic, showing two of the four bases (adenine (A), cytosine (C), guanine (G), thymine (T)). In comparison with micro- and minisatellites SNPs are relatively evenly distributed in the genome and are much more common within genes. The majority of genetically determined phenotypic variation is therefore likely to be explained by SNPs.

With the aim of cataloging common SNP variants in different ethnic populations, the International HapMap Project was initiated in 2002 (International HapMap Consortium 2003). The data produced by this effort has also allowed identification of "tagging SNPs", i.e. SNPs which are informative in predicting alleles at nearby SNPs due to correlation, or linkage disequilibrium (LD), between SNPs within so called haploblocks. This data was soon widely applied for selecting SNPs for association studies. However, the HapMap project was primarily targeting common SNP variants with minor allele frequency (MAF) >1%. In order to identify and catalogue rarer SNPs, the 1000 Genomes Project was launched in early 2008 by the international 1000 Genomes Consortium (http://www.1000genomes.org/page.php). Using next-generation re-sequencing techniques, the pilot phase of this project identified approximately 15 million SNPs, 1 million short (1-50 bp) insertion-deletion polymorphisms (INDEL) and 20,000 larger structural variants of >1 kb (including tandem repeats and tandem duplications, transposable elements, and large insertions and deletions) (1000 Genomes Project Consortium *et al.* 2010). The majority of all identified variants were previously undescribed; expectedly most of the common SNP variants were previously known, but most of the SNPs with lower frequencies (MAF <5%) as well as most of all INDELs and structural variants were novel. While the importance of rare SNPs, INDELs and structural variants in complex diseases remains to be studied, it seems that at least the more common and relatively large (>500 bp) structural variants, also called copy number variants (CNVs), are unlikely to play a major role in complex diseases: a recent genome-

28 Studies on causes of multiple sclerosis

wide CNV association study in 16,000 cases of eight common diseases and 3,000 controls identified only three associated CNV loci (WTCCC *et al.* 2010). Furthermore, the majority of common CNVs are well tagged by common SNPs, suggesting that they are unlikely to explain much of the void left by genome-wide SNP association studies (Conrad *et al.* 2010, WTCCC *et al.* 2010).

2.2.2 Linkage studies

High-resolution microsatellite maps which became available in the early 1990's paved the way for genome-wide linkage studies, in which microsatellite variants are screened for co-segregation with the phenotype in families with at least two affected individuals. A marker is linked to the phenotype if it is transmitted together with the phenotype significantly more often than would be expected if the marker and phenotype were inherited independently. Linkage can therefore only occur if the locus affecting the phenotype is located in the same chromosome with the tested marker, and close enough so that in most meiosis within a pedigree there is no recombination between the marker and the locus encoding the trait. The drawback of linkage studies is that the identified regions are large, possibly carrying hundreds or thousands of genes.

While linkage studies have been extremely successful in leading to identification of genes underlying monogenic disorders, this has not been the case with complex diseases (Altmüller *et al.* 2001). With realistic sample sizes the power to detect linkage is low if the effect of the contributing locus on the phenotype is modest (Risch and Merikangas 1996). The fact that widely replicated linkage signals are virtually lacking in complex diseases therefore implies that loci with major impact are unlikely to significantly contribute to these conditions.

2.2.3 Association studies

The first comprehensive SNP maps together with developments in SNP genotyping techniques allowed a shift to another approach, association studies. In testing for association, the allele frequencies are compared between affected individuals and healthy controls. Alternatively, family based association tests can be used to test whether an allele is transmitted to an affected offspring more often than would be expected under random Mendelian segregation (Figure 4).

Early association studies focused on regions with some evidence of linkage or on candidate genes selected on functional basis and were generally conducted in rather

small samples. However, only few of the associations in these early studies were replicated by others. This irreproducibility could be explained by a variety of factors including initial false positives due to chance or population stratification, variable LD structures between populations, population specific gene-gene and gene-environment interactions and false negative replication efforts (Hirschhorn *et al.* 2002), the first of which is probably the most likely explanation in most cases in the light of our current knowledge.

For the last few years it has, however, been possible to conduct genome-wide association studies (GWASs), where 300,000-1,000,000 SNPs are simultaneously genotyped and tested for association. While GWASs have been successful in revealing novel risk variants in a variety of complex disorders, the risk effects of these common variants have been found to be low (odds ratio (OR) <1.5) and additional genetic risk factors therefore remain to be discovered (McCarthy and Hirschhorn 2008, Altshuler *et al.* 2008). To what extent rarer SNP variants contribute to susceptibility to complex traits will be the focus of research in the next years. Data from the 1000 Genomes Project can be used to impute rare variants in existing GWAS datasets and as the costs of sequencing are rapidly decreasing, whole genome sequencing of large samples will eventually become feasible.

CASE-CONTROL STUDIES

CASES   CONTROLS

FAMILY STUDIES (TDT)

GENOTYPE

○ a/a
⬤ A/a
● A/A

NUMBER OF A ALLELES
TRANSMITTED TO THE
AFFECTED CHILD FROM
HETEROZYGOUS PARENTS

------   0
——   1
━━   2
✕   UNINFORMATIVE
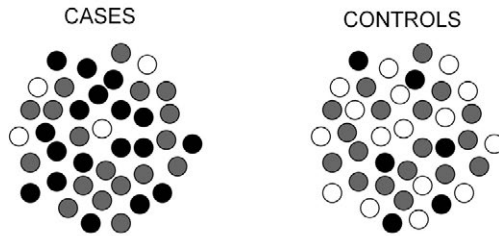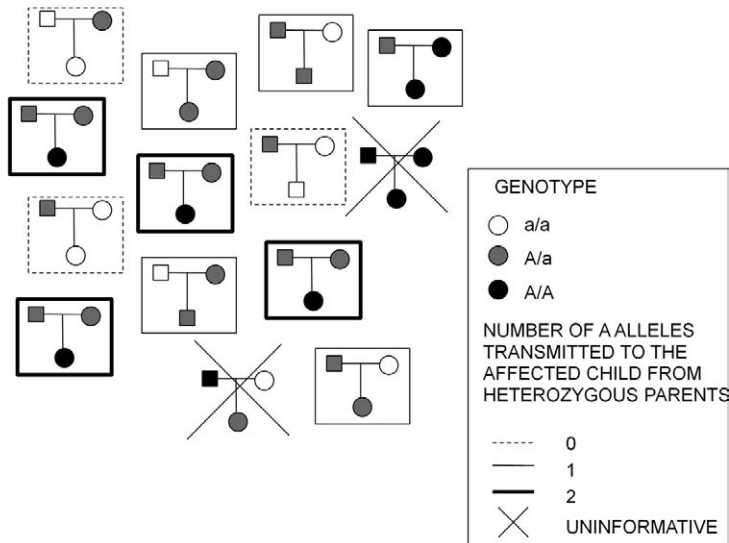
Figure 4. Principle of case-control and family based association studies. In a case-control setting, a two-tailed chi-square test is conducted in order to test the hypotheses that allele counts are different in the two groups, against the null hypotheses that there is no difference. One commonly used family based test is the transmission disequilibrium test (TDT), which counts the number of times alleles are transmitted to the affected offspring from heterozygous parents. Under the null hypotheses of no association, alleles are expected to be transmitted randomly according to Mendelian segregation. In this example 13 of 16 heterozygous parents transmit allele "A" to their offspring and 3 transmit allele "a". A chi-square test is conducted to assess, whether these observed counts differ significantly from expected.

2.2.4 Expression studies

Until recently, there were two main approaches for genome-wide analysis of gene expression: microarrays and Serial Analysis of Gene Expression (SAGE) (Velculescu *et al.* 1995). Shortly, SAGE is based on sequencing of short tags of complementary DNA (cDNA), which can then be identified by sequence searches against public databases and quantified directly based on sequence counts. The benefit of the technique is that basically any transcript present in the sample can be detected. Microarrays on the other hand are limited to detecting only transcripts for which there are probes present on the microarray, but have become the method of choice for genome-wide expression profiling. Latest microarray platforms cover all known human genes and also contain probes for detecting alternative isoforms, i.e. transcripts which are encoded from the same gene but contain a different set of exons or differ at their 5' or 3' untranslated regions (UTR). There are several different types of microarrays. Affymetrix microarrays, also called GeneChip arrays (Affymetrix, CA, USA) differ from other types of microarrays, because the short oligonucleotide probes are extended directly on the surface of silicon chips by photolithography. In other types of microarrays, the probes (either cDNA from a library or synthesized oligonucleotides) are attached on the surface of a solid support rather than being synthesized *in situ*. cDNA microarrays are usually hybridised to cDNA from two different samples (one of the samples can be a reference sample used in all arrays to be compared) labeled with different fluorescent dyes (usually Cy3 (green) and Cy5 (red)). The two labeled samples are mixed together and hybridised on the same array. Spot intensity is determined for the two wavelengths (detecting e.g. Cy3 and Cy5) and their ratio represents the difference in the amount of transcript in the two samples. In contrast, Affymetrix chips are hybridised with only one sample (see Figure 5 for Affymetrix workflow). After hybridisation with a labeled sample arrays are washed to remove unhybridised sample and scanned to measure the fluorescence signals. Because one-color experiments do not allow incorporation of a reference sample on the same array, normalisation and data analyses differ from that of two-color experiments. Raw data from most recently published microarray experiments is available in public data repositories such as the Gene Expression Omnibus (GEO) by National Center for Biotechnology Information (NCBI).

Microarrays have been widely used for comparing expression profiles between patients and healthy controls in order to identify genes, which show changes in transcript levels in affected individuals. However, genes with altered expression are not necessarily involved in causing the disease, since many genes may also show altered expression after disease onset due to activation of protective mechanisms or secondary pathogenic processes, for example. In addition to the potential effects of

the disease itself, measured gene expression levels can be influenced by various non-biological and biological factors including genetic variants (some of which may be disease risk variants) (Figure 6). Genetic loci which correlate with gene expression are called expression quantitative trait loci (eQTL). Studies correlating genome-wide SNP data with genome-wide expression data have now identified thousands of eQTLs and their target genes (Dimas *et al.* 2009, Dixon *et al.* 2007, Moffatt *et al.* 2007, Myers *et al.* 2007, Schadt *et al.* 2008, Stranger *et al.* 2007, Veyrieras *et al.* 2008, Zeller *et al.* 2010). eQTLs can be located close to the gene(s) they regulate in which case they are said to be *cis*-acting, or further away or in a different chromosome, in which case they are *trans*-acting. Although in most cases the mechanisms by which eQTLs exert their effects on gene expression are still unknown, *cis*-acting SNPs may for example have an impact on efficiency of transcription factor binding, while *trans*-acting SNPs may alter the expression or structure of nearby transcription factors or microRNAs (miRNA), i.e. short RNA molecules which bind to their target mRNAs and regulate their degradation and translation efficiency. This in turn has consequences on the expression of their target genes in other loci (in *trans*). Many eQTLs may be tissue- or cell type-specific (Dimas *et al.* 2009) and studies correlating disease risk variants with gene expression should therefore ideally focus on a tissue or cell population, which is known to be relevant for the disease.

Decreased costs of next-generation sequencing techniques have recently made it possible to sequence the transcriptome, i.e. the entire set of ribonucleic acid (RNA) molecules in a given sample. This technique called RNA-sequencing (RNA-Seq) enables not only quantitative analysis but also precise identification of transcriptional isoforms (Nagalakshmi *et al.* 2008). RNA-Seq is, however, still costly in comparison with microarrays.
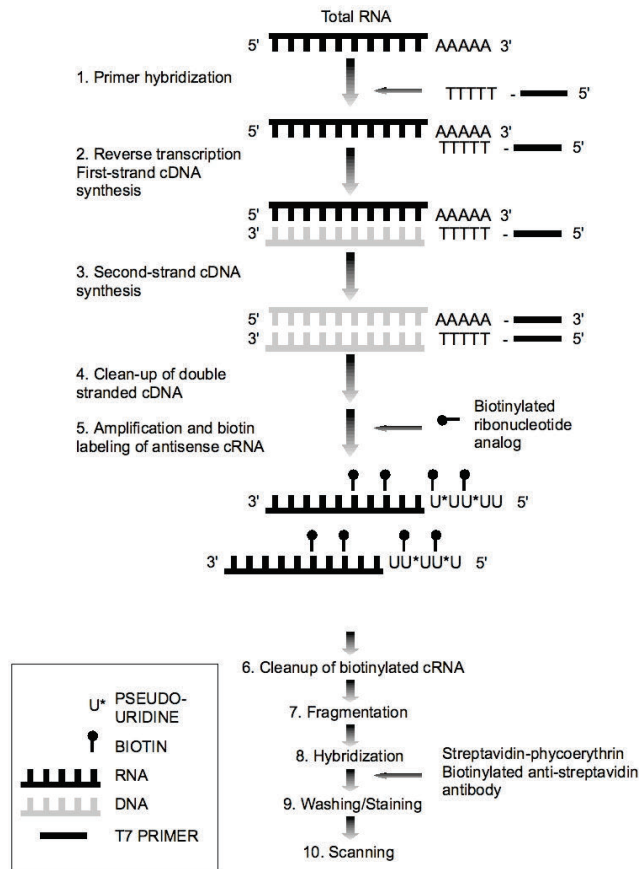
Figure 5. Affymetrix® microarray experiment workflow. Figure is based on text and figures in GeneChip® Expression Analysis Technical Manual (Affymetrix®) (www.affymetrix.com). In the first step the first strand of cDNA is created by using T7-Oligo(dT) promoter primer, which recognizes mRNA poly(A) and has a binding sequence for T7 RNA polymerase. Second cDNA strand is then synthesised to serve as template for complementary RNA (cRNA) strand amplification. In addition to normal ribonucleotides, biotin-labeled pseudouridine nucleotides are also added into the reaction and become incorporated in the cRNA at random uridine positions. After clean-up and fragmentation of cRNA molecules into 25-200 bp fragments, the fragments are hybridized on the chip containing complementary oligonucleotide probes. Streptavidin-phycoerythrin is then added into the reaction: streptavidin binds strongly with biotin present on the cRNA fragments while phycoerythrin acts as a fluorescence marker. Fluorescence signal is further amplified by adding biotinylated anti-streptavidin. Strength of the fluorescence signal reflects the amount of cRNA bound to each of the spots containing millions of copies of a specific probe sequence.
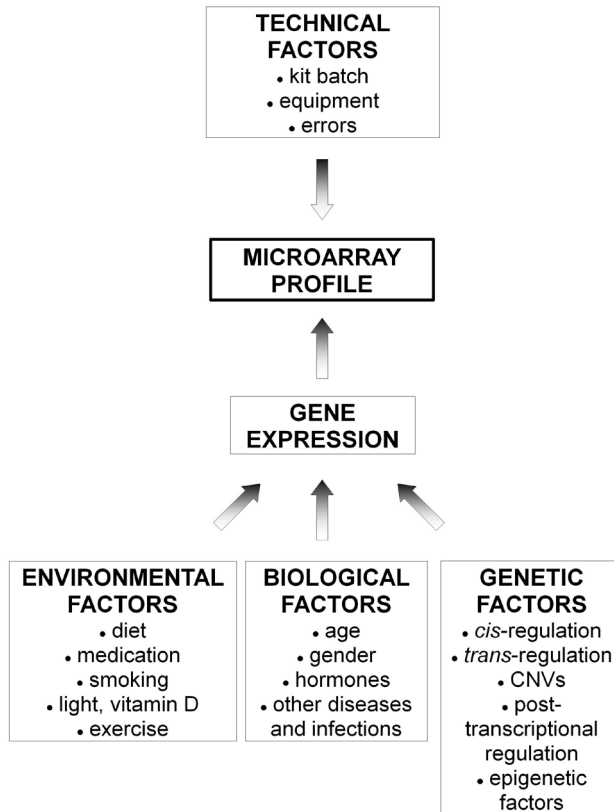
Figure 6. Factors affecting gene expression profiles in microarray experiments.

2.3 Epidemiology and genetics of MS

2.3.1 Epidemiological studies and role of environment in MS

The prevalence of MS is typically high $(100\text{-}200/10^5)$ in Northern Caucasian populations (Northern Europe and North America), but also for example in Sardinians, Sicilians, Indian Parsis and Palestinians (Figure 7, Pugliatti *et al.* 2002). Low prevalences are found in Southern Europe, Southern America and Australia and New Zealand and the disease is rarely encountered in black African, Asian and many indigenous populations such as Samis and Maoris (Pugliatti *et al.* 2002). These differences between populations are likely to be explained by both genetic and environmental factors as well as their interactions. Increased prevalence and incidence of MS during the past decades, particularly in females, seem to be rather universal phenomena, and suggest an increasing exposure to some environmental or lifestyle-related risk factors (Koch-Henriksen and Sørensen 2010).

Biological relatives of MS patients have a significantly increased risk for developing MS (Table 2). The role of genetic factors in explaining familial aggregation has been shown in twin and adoption studies: risk of developing MS is higher in MZ twins of affected individuals than in their DZ twins or siblings (Willer *et al.* 2003), while the risk is not increased in non-biological family members of adopted MS patients (Ebers *et al.* 1995). A shared family environment therefore does not seem to predispose to MS, suggesting that the environmental influences act on a population level.

The role of various environmental factors in MS has been investigated. Pathogens have been studied quite extensively, but no undisputable evidence has been obtained for their involvement. One of the most studied viruses is the Epstein-Barr virus (EBV), which is a ubiquitous herpes group virus with cross-reactivity with MBP (Lang *et al.* 2002). Serafini *et al.* (2007) found EBV infected B cells in 21 of 22 post-mortem MS brain samples, but not in 11 non-MS control samples, most of which were obtained from sufferers of other inflammatory neurological diseases. However, two more recent studies failed to replicate these findings (Willis *et al.* 2009, Fatima *et al*. 2011) and the role of EBV in MS remains elusive.

Based on the observed higher prevalence of MS in Northern countries, association between MS and sunlight exposure and D vitamin levels has also been hypothesized (Ascherio *et al.* 2010). While there is some evidence for a protective effect of D vitamin, the precise mechanisms and role of gene-environment interactions remain to be further investigated.

Given that MS is more common in females than in males, sex hormones have also been suggested to play a role in MS (Shuster 2008). Relating to this hypothesis, earlier age at menarche has been shown to increase the risk for MS (Ramagopalan *et al.* 2008).

Lifestyle factors commonly suggested to increase the risk for MS include smoking (Hawkes 2007, Sundström *et al.* 2008), possibly in interaction with MS risk variants at Human Leukocyte Antigen (HLA) genes (Hedström *et al.* 2011), and diet (Lauer 1994).

Table 2. Prevalence of MS in relatives of MS probands (Ebers 2008).

Relationship to MS patient	Prevalence (%)	Identity by descent (%)
Adoptive sibling	0.1	0
First cousin	0.7	12.5 (on average)
Paternal half sibling	1.3	25 (on average)
Half sibling reared apart	2.1	25 (on average)
Maternal half sibling	2.4	25 (on average)
Full sibling	3.5	50 (on average)
Sibling in consanguineous mating	8	50 (on average)
Offspring of conjugal mating	20	50
Monozygotic twin	27	100
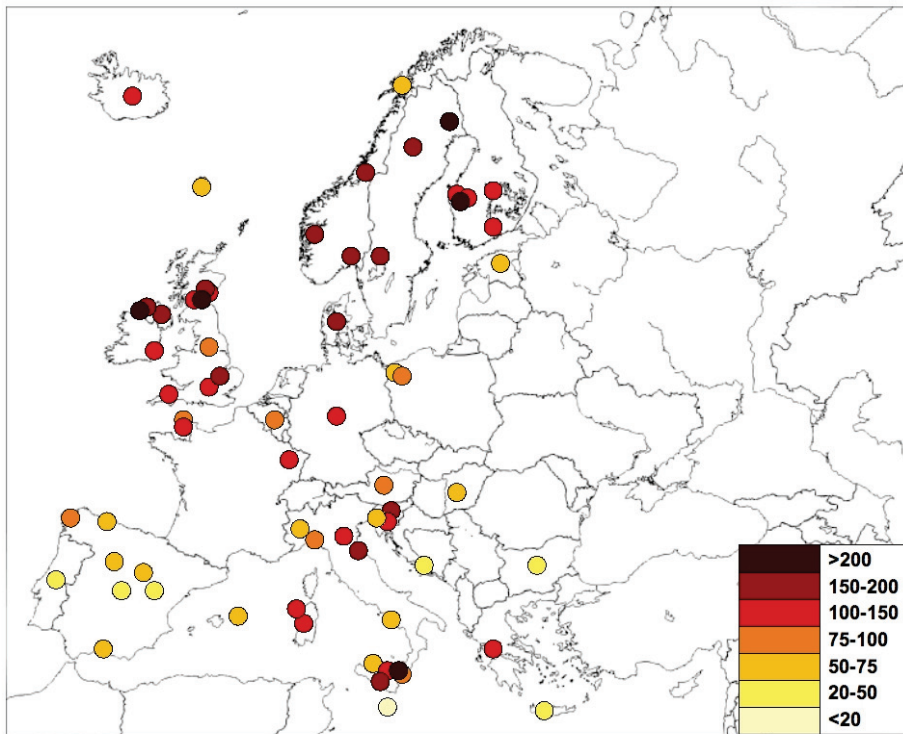
Figure 7. Prevalence of MS in European populations. Based on data from 1990-
2010 (Binzer *et al.* 1994, Cavalletti *et al.* 1994, Fernández *et al.* 1994, van
Ooteghem *et al.* 1994, Rice-Oxley *et al.* 1995, Sharpe *et al.* 1995, Robertson *et al.*
1996, Dean *et al.* 1997, Uria *et al.* 1997, Benito-León *et al.* 1998, Bencsik *et al.*
1998, Casetta *et al.* 1998, Ford *et al.* 1998, Grant *et al.* 1998, McDonnell and
Hawkins 1998, Pina *et al.* 1998, Rothwell and Charlton 1998, Forbes *et al.* 1999,
Gross-Paju *et al.* 1999, Milanov *et al.* 1999, Potemkowski 1999, Tola *et al.* 1999,
Grønlie *et al.* 2000, Hein and Hopfenmüller 2000, Moreau *et al.* 2000, Bencsik *et al.*
2001, Casquero *et al.* 2001, Grimaldi *et al.* 2001, Kurtzke and Heltberg 2001,
Sumelahti *et al.* 2001, Baumhackl *et al.* 2002, Benedikz *et al.* 2002, Dean *et al.* 2002,
Sundström *et al.* 2003, Dahl *et al.* 2004, Fox *et al.* 2004, McGuigan *et al.* 2004,
Murray *et al.* 2004, Ragonese *et al.* 2004, Sarasoja *et al.* 2004, Tienari *et al.* 2004,
Nicoletti *et al.* 2005a, Nicoletti *et al.* 2005b, Pugliatti *et al.* 2005, Solaro *et al.* 2005,
De Sá *et al.* 2006, Grytten *et al.* 2006, Peterlin *et al.* 2006, Ares *et al.* 2007,
Debouverie *et al.* 2007, Granieri *et al.* 2007, Grimaldi *et al.* 2007, Granieri *et al.*
2008, Gray *et al.* 2008, Iuliano and Napoletano 2008, Papathanasopoulos *et al.* 2008,
Smestad *et al.* 2008, Boström *et al.* 2009, Bentzen *et al.* 2010).

2.3.2 Linkage studies

The first three genome-wide MS linkage scans, conducted in UK, US and Canada, were published back to back in 1996 (Ebers *et al.* 1996, Haines *et al.* 1996, Sawcer *et al.* 1996). These were followed by genome-wide scans in Finland (Kuokkanen *et al.* 1997), Italy (Broadley *et al.* 2001), Sardinia (Coraddu *et al.* 2001), Scandinavia (Akesson *et al.* 2002), Australia (Ban *et al.* 2002) and Turkey (Eraksoy *et al.* 2003). None of these studies identified any loci with a genome-wide significant logarithm of odds (LOD) score >3, although several loci with suggestive evidence for linkage were identified. A meta-analysis of all nine linkage scans containing 719 families and on average 359 microsatellite markers per family did not identify any loci with LOD >3 apart from the MHC, although loci on 17q21 and 22q13 were suggestively linked (GAMES 2003).

Later, extended studies were published by the groups in UK, US-France and Canada (Hensiek *et al.* 2003, Dyment *et al.* 2004b, Kenealy *et al.* 2004). Significant evidence for linkage at the MHC locus was found in Canadian and US-French studies. Finally, a large linkage study in 730 pedigrees with over 4500 SNPs conducted by the IMSGC failed to identify any non-MHC regions with significant linkage (Sawcer *et al.* 2005). Based on these studies it was concluded that any common non-MHC alleles are likely to increase MS risk by a factor of <2.0 (Sawcer 2008). However, the poor overlap across linkage studies and the failure of larger scans to detect significant non-MHC loci may also to some extent reflect locus heterogeneity in MS between and within populations.

In Finnish MS pedigrees linkage studies identified four candidate regions with evidence of linkage: the MHC locus on 6p (LOD_{max}=6.4, Tienari *et al.* 1993), 5p12-p14 (LOD_{max}=3.4, Kuokkanen *et al.* 1996), 17q22-q24 (LOD_{max} score=2.8, Kuokkanen *et al.* 1997) and the myelin basic protein locus on 18q23 (LOD_{max}=3.3, Tienari *et al.* 1992). In addition, 19q13.1 showed weak evidence for linkage in HLA-DR15 negative families (LOD_{max}=1.8, Reunanen *et al.* 2002). The linked region on chromosome 17 was later narrowed down by fine mapping to a 2.5 Mb segment (LOD_{max}=2.9, Saarela *et al.* 2002) and the region on chromosome 5 to 5.3 Mb (LOD_{max}=3.5, Riise Stensland *et al.* 2005). Loci on 5p and 17q have emerged also in other studies (Dyment *et al.* 2004b; Eraksoy *et al.* 2003, Hensiek *et al.* 2003, Sawcer *et al.* 2005) with LOD scores 1.1-2.45 and are syntenic with susceptibility loci in EAE mouse and rat (Jagodic *et al.* 2001; Sundvall *et al.* 1995). However, all linkage regions are still wide and contain numerous candidate genes, some of which have been further investigated in association studies (see 2.3.3).

2.3.3 Association studies

Prior to platforms for genome-wide SNP genotyping, association studies needed to be targeted on specific loci or genes, usually selected on the basis of evidence of linkage or a functional hypothesis. Given that any gene with a neurological or immunological function would make a valid candidate for an MS susceptibility gene, it is not surprising that most association studies have failed to find association, while the initial positive findings have rarely been replicated, indicating high false positive rates in these early candidate gene studies. On the other hand, since the first studies were conducted in small samples, some true signals may have been missed as well. The only solid finding for a long time was the MHC region, which is indisputably the strongest genetic risk factor for MS. However, as a result of recent GWASs there are currently close to 50 non-MHC loci, which are associated with MS with genome-wide significance (GWS) (P-value $\leq 5 \times 10^{-08}$). Most of these are located at or near genes with immunological functions.

Major Histocompatibility Complex (MHC)

First associations of HLA protein variants, which are encoded by HLA-genes within the MHC region, with MS were described nearly four decades ago (Naito *et al.* 1972; Jersild *et al.* 1972; Bertrams *et al.* 1972). The since identified DNA risk haplotype HLA-DRB1*1501-DQA1*0102-DQB1*0602 has been widely replicated (Schmidt *et al.* 2007) and is estimated to explain 17%-62% of MS heritability (Haines *et al.* 1996). However, it still remains unclear which of the genes is the main determinant of increased risk for MS, and whether there are several susceptibility genes within the MHCII region. The HLA-genes within MHCII encode molecules expressed on the surface of APCs and present antigens to CD4+ cells, but the mechanisms by which the MS associated variants result in increased susceptibility are currently unknown. In addition to MHCII, there is also evidence for independent risk factors in the MHCI region (Harbo *et al.* 2004, Yeo *et al.* 2007, De Jager *et al.* 2009, IMSGC and WTCCC2 2011).

Non-MHC loci

As of April 2011 seventeen non-MHC loci had been reported to be associated with MS with GWS (Table 3). In addition, a study in nearly 10,000 MS patients and more than 17,000 controls conducted by the IMSGC and the Wellcome Trust Case Control Consortium 2 (WTCCC2) had been completed at the time of writing this thesis and had identified 29 novel risk loci (IMSGC and WTCCC2 2011) (Table 4). However, the reported ORs for all confirmed risk variants are low (1.1-1.4) and even together with the MHC region, much if not most of the genetic variants contributing

to MS heritability still remain to be identified. Further finemapping and functional studies are also required in order to pinpoint the functionally relevant variants and genes in the currently known loci, many of which extend over several hundred kb and contain several genes. Many candidate genes in these loci have immunological functions (*CD58*, *CD6*, *CD86, CLEC16A*, *CLECL1, GPR65, EOMES, IL12A*, *IL12B, IL2RA, IL22RA2, IL7R, IRF8, MALT1, MERTK, SP140, STAT3, TAGAP, TNFRSF1A*, *TNFSF14*, *TYK2*), although interestingly some genes relevant for neurological functions have also emerged (*AHI1*, *GALC*, *KIF21B*). Strikingly many of the established loci overlap with those identified in other autoimmune diseases (Table 5). These findings point towards partially shared causative mechanisms even in diseases with very different manifestations such as MS and celiac disease (CD).

The mechanisms through which MS variants exert their effects on disease susceptibility are currently poorly understood, but it is likely that many common risk variants modulate gene expression. At least 30% of the GWS MS risk variants have an effect on gene expression of a nearby gene (Table 6). However, it is likely that more such effects would be identified if larger studies and studies in different immune cell populations were conducted. Furthermore, only one of the published eQTL studies had sufficient power to test for *trans*-regulatory effects (Zeller *et al.* 2010), which may be important in regulating immunological pathways relevant in autoimmune diseases (Heinig *et al.* 2010).

Table 3. Non-MHC risk loci associated with MS with GWS as of April 2011.

SNP	Nearest Gene	Description	N(Cases);N(controls) in original study	Reference
rs9657904	*CBLB*	Cas-Br-M (murine) ecotropic retroviral transforming sequence b	1775; 2005	Sanna *et al.* 2010
rs2300747	*CD58*	CD58 molecule	4838; 9336	De Jager *et al.* 2009
rs17824933	*CD6*	CD6 molecule	4838; 9336	De Jager *et al.* 2009
rs12708716	*CLEC16A*	C-type lectin domain family 16, member A	5737; 10296	IMSGC 2009
rs17445836	*IRF8*	interferon regulatory factor 8	4838; 9336	De Jager *et al.* 2009
rs4680534	*IL12A*	interleukin 12A	10709; 14997	IMSGC 2010a
rs12722489	*IL2RA*	interleukin 2 receptor, alpha	2322; 5418 (+ 1540 trios)	IMSGC *et al.* 2007
rs6897932	*IL7R*	interleukin 7 receptor	11019; 13616	IMSGC 2008
rs10492972*	*KIF1B*	kinesin family member 1B	2679; 3125	Aulchenko *et al.* 2008
rs12122721	*KIF21B*	kinesin family member 21B	3507; 3395	IMSGC 2010b
rs703842	*METTL1*	methyltransferase like 1; cytochrome P450, family 27, subfamily B, polypeptide 1	3874; 5723	ANZgene 2009
rs1790100	*MPHOSPH9*	M-phase phosphoprotein 9/cyclin-dependent kinase 2 associated protein 1	10709; 14997	IMSGC 2010a
rs2760524	*RGS1*	regulator of G-protein signaling 1	10709; 14997	IMSGC 2010a
rs744166	*STAT3*	signal transducer and activator of transcription 3	3859; 9110	Jakkula *et al.* 2010
rs1132200	*TMEM39A*	transmembrane protein 39A	3507; 3395	IMSGC 2010b
rs1800693	*TNFRSF1A*	tumor necrosis factor receptor superfamily, member 1A	4838; 9336	De Jager *et al.* 2009
rs34536443	*TYK2*	tyrosine kinase 2	10642; 10620 (+2110 trios)	Mero *et al.* 2010

*This association was not replicated in a larger sample of 8391 cases, 8052 controls and 2137 trio families and is likely to have been false positive (IMSGC *et al.* 2010)

Table 4. Novel MS risk loci identified with GWS in the IMSGC and WTCCC2 GWAS.

SNP	Nearest Gene	Description
rs11154801	*AHI1*	Abelson helper integration site 1
rs12212193	*BACH2*	BTB and CNC homology 1, basic leucine zipper transcription factor 2
rs2300603	*BATF*	basic leucine zipper transcription factor, ATF-like
rs9282641	*CD86*	CD86 molecule
rs10466829	*CLECL1*	C-type lectin-like 1
rs669607	*CMC1*	COX assembly mitochondrial protein homolog (*S. cerevisiae*)
rs2248359	*CYP24A1*	cytochrome P450, family 24, subfamily A, polypeptide 1
rs2303759	*DKKL1*	dickkopf-like 1
rs11129295	*EOMES*	eomesodermin
rs2119704	*GPR65*	G protein-coupled receptor 65
rs7923837	*HHEX*	hematopoietically expressed homeobox
rs2546890	*IL12B*	interleukin 12B (natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40)
rs17066096	*IL22RA2*	interleukin 22 receptor, alpha 2
rs7238078	*MALT1*	mucosa associated lymphoid tissue lymphoma translocation gene 1
rs2283792	*MAPK1*	mitogen-activated protein kinase 1
rs17174870	*MERTK*	c-mer proto-oncogene tyrosine kinase
rs874628	*MPV17L2*	MPV17 mitochondrial membrane protein-like 2
rs4410871	*MYC*	v-myc myelocytomatosis viral oncogene homolog (avian)
rs2019960	*MYC*	v-myc myelocytomatosis viral oncogene homolog (avian)
rs140522	*ODF3B*	outer dense fiber of sperm tails 3B
rs11581062	*SLC30A7*	solute carrier family 30 (zinc transporter), member 7
rs10201872	*SP140*	SP140 nuclear body protein
rs12466022	*ZFP36L2*	zinc finger protein 36, C3H type-like 2
rs7595037	*PLEK*	pleckstrin
rs802734	*PTPRK*	protein tyrosine phosphatase, receptor type, K
rs1738074	*TAGAP*	T-cell activation RhoGTPase activating protein
rs1077667	*TNFSF14*	tumor necrosis factor (ligand) superfamily, member 14
rs4902647	*ZFP36L1*	zinc finger protein 36, C3H type-like 1
rs354033	*ZNF746*	zinc finger protein 746

Table 5. GWS MS risk loci identified in other autoimmune diseases with GWS. MS risk loci were listed in Tables 3 and 4. NHGRI Catalog of Published Genome-Wide Association Studies track in the UCSC Genome Browser (April 2011) was applied to identify SNPs associated with other autoimmune diseases and the loci were considered to overlap with MS risk loci if the SNPs were located within 500kb. AA= Alopecia areata, CD=celiac disease, ChD=Crohn's disease, PS=psoriasis, PSA=psoriatic arthritis, RA=rheumatoid arthritis, SLE=systemic lupus erythematosus, T1D=type 1 diabetes, UC=ulcerative colitis.

Gene nearest to the MS associated SNP	Associated SNP in other disease(s) (r^2 with MS SNP based on 1000 Genomes data)	Disease(s)
BACH2	rs10806425 (0.48)	CD (Dubois *et al.* 2010)
	rs11755527 (0.26)	T1D (Barrett *et al.* 2009, Cooper *et al.* 2008)
CLEC16A	rs12708716 (0.02)	T1D (Barrett *et al.* 2009, Cooper *et al.* 2008, Todd *et al.* 2007)
	rs12928822 (0.02)	CD (Dubois *et al.* 2010)
	rs2903692 (0.96)	T1D (Hakonarson *et al.* 2007)
CLECL1	rs3764021 (0.87)	T1D (WTCCC 2007)
	rs4763879 (0.47)	T1D (Barrett *et al.* 2009)
IL12A	rs17810546 (0.05)	CD (Dubois *et al.* 2010, Hunt *et al.* 2008)
IL12B	rs10045431 (0.06)	ChD (Barrett *et al.* 2008)
	rs12188300 (0.01)	PSA (Hüffmeier *et al.* 2010)
	rs2082412 (0.31)	PS (Nair *et al.* 2009)
	rs2546890 (0.46)	PS (Ellinghaus *et al.* 2010)
	rs3213094 (0.31)	PS (Zhang *et al.* 2009)
IL2RA	rs12251307 (0.02)	T1D (Barrett *et al.* 2009)
	rs3118470 (0.1)	AA (Petukhova *et al.* 2010)
	rs706778 (0.12)	RA (Stahl *et al.* 2010)
	rs706779 (0.16)	Vitiligo (Jin *et al.* 2010)
	rs947474 (0.02)	T1D (Cooper *et al.* 2008)
KIF21B	rs11584383 (0.96)	ChD (Barrett *et al.* 2008)
	rs296547 (0.77)	CD (Dubois *et al.* 2010)
MAPK1	rs131654 (0.05)	SLE (Han *et al.* 2009)
MYC	rs9792269 (0.02 with rs4410871, 0.05 with rs2019960)	CD (Dubois *et al.* 2010)
PLEK	rs17035378 (0.37)	CD (Dubois *et al.* 2010)
PTPRK	rs802734 (0.36)	CD (Dubois *et al.* 2010)
RGS1	rs2816316 (0.94)	CD (Dubois *et al.* 2010, Hunt *et al.* 2008)
STAT3	rs744166 (same SNP)	ChD (Barrett *et al.* 2008)
TAGAP	rs1738074 (0.54)	CD (Dubois *et al.* 2010)
TYK2	rs2304256 (information not available)	T1D (Wallace *et al.* 2009)
ZFP36L1	rs1465788 (0.31)	T1D (Barrett *et al.* 2009)

Table 6. GWS MS risk variants correlating with gene expression. The MS SNPs (or their best Affy 6.0 proxy) were searched in the monocyte eQTL database using their threshold of GWS. In addition the mRNA by SNP Browser (Dixon *et al.* 2007, Moffatt *et al.* 2007, database available at http://www.sph.umich.edu/csg/liang/asthma/) was searched for effects with LOD >3, again using the best proxy SNP (r^2 >0.8) if the associated SNP was not available. All variants are *cis*-acting unless specified as *trans*.

SNP	Target gene	Tissue/cell type (reference)
rs10201872	*SP110*	LCL (mRNA by SNP Browser)
	SP140	LCL (mRNA by SNP Browser)
rs10466829	*CLECL1*	LCL, monocytes (mRNA by SNP Browser, Zeller *et al.* 2010)
	IGFL3 (*trans*)	Monocytes (Zeller *et al.* 2010)
	KLRB1	LCL (mRNA by SNP Browser)
rs1077667	*TNFSF14*	LCL (mRNA by SNP Browser)
rs11154801	*AHI1*	LCL, monocytes (mRNA by SNP Browser, Zeller *et al.* 2010)
rs12708716	*DEXI*	LCL, monocytes (mRNA by SNP Browser, Zeller *et al.* 2010)
rs140522	*CPT1B*	Monocytes (Zeller *et al.* 2010)
	ECGF1	Monocytes (Zeller *et al.* 2010)
	KREMEN2 (*trans*)	Monocytes (Zeller *et al.* 2010)
rs17174870	*MERTK*	Monocytes (Zeller *et al.* 2010)
rs17824933	*CD6*	Monocytes (Zeller *et al.* 2010)
rs1790100	*MPOSHPH9*	Monocytes (Zeller *et al.* 2010)
	OGFOD2	LCL (mRNA by SNP Browser)
	SBNO1	LCL (mRNA by SNP Browser)
rs2300747	*CD58*	LCL (mRNA by SNP Browser)
rs2760524	*RGS1*	Monocytes (Zeller *et al.* 2010)
rs4902647	*ZFP36L1*	LCL (mRNA by SNP Browser)
rs703842	*FAM119B*	LCL, monocytes (mRNA by SNP Browser, Zeller *et al.* 2010)
	TSFM	LCL, monocytes (mRNA by SNP Browser, Zeller *et al.* 2010)
	TSPAN31	Monocytes (Zeller *et al.* 2010)
	XRCC6BP1	Monocytes (Zeller *et al.* 2010)
rs874628	*CHRM4* (*trans*)	Monocytes (Zeller *et al.* 2010)

Studies on causes of
multiple sclerosis

Association studies in the Finnish population

As a result of detailed screening of the candidate region 17q22-24 suggested by linkage studies, Saarela *et al.* (2006) identified the first non-MHC MS associated gene in the Finnish population, *PRKCA*, in 151 families (OR=1.34, 95% CI=1.07-1.68). In addition, they measured *PRKCA* transcript levels in CD4+ T cells and in non-CD4+T cell PBMCs (containing all PBMCs except CD4+ T cells, i.e. CD8+ T cells, B cells, NK cells, monocytes, macrophages and DCs) previously collected from MS patients and their family members. They detected a 2-fold decrease in *PRKCA* transcript levels in non-CD4+ T cell PBMCs in carriers of two copies of the associated two-SNP haplotype (*n*=9) versus carriers of one copy (*n*=7), but did not see any significant difference in CD4+ PBMCs. Association was also demonstrated in 554 Canadian families in the same study, although the strongest association was found with a different haplotype (OR=1.64, 95% CI=1.39-1.94). In addition, association at *PRKCA* has been detected in 184 cases and 340 controls from UK (P-value=0.002) (Barton *et al.* 2004).

The first Finnish GWAS, conducted in 68 distantly related patients and 136 controls from a Southern Ostrobothnian subisolate with high MS prevalence, revealed associations in *STAT3* and in a region on chromosome 16 in addition to *HLA-DRB1* (Jakkula *et al.* 2010). Association in *STAT3* was replicated with GWS in 3859 cases and 9110 controls from more heterogeneous populations in the same study. Furthermore, the same variant is associated with Crohn's disease (Barrett *et al.* 2008). *STAT3* is involved in various immunological signaling pathways. Interestingly, EAE does not develop in mice with a targeted STAT3 deletion in CD4+ T cells (Liu *et al.* 2008). The associated region on chromosome 16 does not harbour any known genes, the nearest gene being located 1.4 Mb away, but shows a high level of evolutionary conservation.

The GWAS data was also applied to investigate haplotype sharing at one of the most consistent linkage regions, 5p. This resulted in identification of a rare risk haplotype extending over the complement component 7 gene *C7* and a gene with unknown function, *FLJ40243* (P-value=0.0001), 5.1 Mb centromeric of *IL7R* (Kallio *et al.* 2009). Association was replicated in samples from the same subisolate and suggestive evidence was detected also in more heterogeneous Northern European populations. The risk allele was also found to correlate with higher total complement activity. This finding gives support for the role of innate immunity and complement system in MS pathogenesis, although the role of *FLJ40243* as the causative gene cannot be excluded at this stage.

While associations at *PRKCA* and *C7* have not been confirmed, of the currently known GWS MS loci *IL7R* and *STAT3* map to the originally identified wider linkage

regions on chromosomes 5 and 17, respectively, and may explain some of the linkage signal. Of the candidate genes on the linked 19q13 region, Sulonen *et al.* (2009) tested the association of *TYROBP* with MS in Finnish samples based on the fact that loss-of-function mutations in this gene are found in a rare disease affecting bone and brain white matter called polycystic lipomembranous osteodysplasia with sclerosing leucoencephalopathy (PLOSL) (Paloneva *et al*. 2000). However, common genetic variants in this gene were found to be unlikely to play a role in MS.

2.3.4 Expression studies

Several microarray studies have been conducted on samples from MS patients, both using CNS biopsies and peripheral blood cells as the source of RNA. As with most microarray experiments direct comparison of these experiments is difficult. Studies conducted in samples from peripheral blood cells from human MS patients and healthy controls and including at least 10 samples in both groups are described below.

Using Affymetrix GeneChip U95Av2, Achiron *et al.* (2004) studied expression of genes in PBMCs from 26 MS patients and 18 healthy controls. They identified a statistically significant transcriptional MS-signature of 1109 genes, which was not dependent on disease activity or treatment. These contained genes involved in T-cell activation and expansion, inflammation, and apoptosis. In addition, a signature of 721 genes differed between MS patients in relapse versus those in remission. These genes included genes involved in cellular recruitment, epitope spreading and escape from regulatory immune surveillance.

Arthur *et al.* (2008) compared whole blood expression profiles of 10 RR-MS patients in relapse and remission with healthy controls. Of 9381 cDNAs present on the array, 989 genes were significantly upregulated in relapse, including *ALOX5*, and 536 were significantly downregulated. In remission, 655 genes including *ALOX5* were upregulated and 662 downregulated. *ALOX5*, encoding for arachidonate 5-lipoxygenase, was considered of particular interest, since it had been identified also in previous studies (Whitney *et al.* 2001, Booth *et al.* 2005). Over-represented pathways among the genes upregulated during relapse were mostly related to apoptosis and inflammation, and genes upregulated in remission were found to be related to protein localization, for example.

Using a cDNA microarray with 6500 (Set 1) or 7500 (Set 2) cDNA clones, Bomprezzi *et al.* (2003) conducted an experiment with fresh PBMCs from 14 MS patients and 7 controls and frozen blood samples from 3 patients and 2 controls (Set

1) in addition to another 10 frozen patient and 10 frozen control samples (Set 2). They reported 303 differentially expressed genes between MS patients and healthy controls, including *IL7R*. In general, several genes mapping to 6p21 (harboring the HLA genes, histone cluster and heatshock protein 70) were found to be differentially expressed.

Brynedal *et al.* (2010) studied expression profiles in PBMCs and CSF from 26 MS patients and 18 controls using Affymetrix Human Genome U133 Plus 2.0 array covering over 47,000 transcripts. Surprisingly, no differentially expressed genes were identified in PBMCs, while 939 probesets were differentially expressed in CSF. However, patients in relapse showed differential expression in 266 probesets in comparison with those in remission in PBMCs. Based on these observations the authors conclude that MS is a CNS disease, but that events outside the CNS are important in triggering the relapse.

The largest study so far was conducted by Gandhi *et al.* (2010) in whole blood samples from 99 untreated patients and 45 matched healthy controls. Study was conducted using human HT-12 expression beadchip (Illumina, Inc., CA, USA), which profiles 48,804 transcripts. Genome-wide SNP data was also available for 115 of the samples from GWAS conducted by ANZgene (2009). They found a strong upregulation of T cell-related genes in MS, in addition to genes involved in translational regulation, oxidative phosphorylation, immune synapse and antigen presentation pathways. They also identified a significant association between *CD40* expression and *CD40* MS risk haplotype, which had been identified in the ANZGene GWAS. There was also a strong correlation between the MS risk allele at 12q13-14, another locus identified in the same GWAS, and expression of *FAM119B*, which is one of 13 leukocyte expressed genes in the 12q13-14 candidate region.

Mandel *et al.* (2004) studied expression in 13 RR-MS patients, 8 SLE patients and 10 healthy controls using Affymetrix Genechip array U95Av2. 541 genes showed differential expression in both MS and SLE patients compared to healthy controls including genes involved in apoptosis, cell cycle, inflammation and matrix metalloproteinase pathways. In addition, 1031 genes were specifically differentially expressed in MS. Genes with increased expression in MS included several genes encoding adhesion molecules, whereas heat shock protein genes showed decreased expression in MS.

Ramanathan *et al.* (2001) compared gene expression in >4000 genes in monocyte-depleted PBMCs from 15 RR-MS patients and 15 age- and sex matched controls and identified 34 genes with significantly differential expression. Expression levels of 25 genes were increased and 9 decreased in MS. One of the upregulated genes was *IL7R*.

48 Studies on causes of
multiple sclerosis

Satoh *et al.* (2005, 2006) conducted microarray experiments in T-cells and non-T-cells comparing gene expression in 72 MS patients with 22 healthy unrelated controls. However, the used cDNA array contained only 1258 genes, thereby with a relatively modest coverage of the transcriptome. In T cells they identified 286 genes with differential expression between patients and controls, and further showed that MS patients could be divided in four subgroups based on expression of these genes (Satoh *et al.* 2006). In another study they identified 173 genes in T cells and 50 genes in non-T cells as differentially expressed between cases and controls. More than 80% of the top 30 most significant genes were classified as apoptosis signaling-related (Satoh *et al.* 2005).

Studies on causes of
multiple sclerosis

3 Aims of the study

The aim of this thesis was to elucidate genes and pathways involved in susceptibility and pathogenesis of multiple sclerosis. This goal was addressed in three projects, with the following specific aims:

I. Testing the association of SNP variants in a candidate gene, *MYO9B*, with multiple sclerosis (Study I).

II. Identifying novel candidate genes and pathways through analysis of genome-wide expression in MS patients and controls (Study II, unpublished data).

III. Replicating association of *CXCR4* with multiple sclerosis (Study III).

4 Materials and methods

4.1 Study samples

The study samples used in studies I-III are presented in Tables 7-9 and Figure 8.

DNA samples

In study I genotyping was first conducted in 730 unrelated Finnish MS cases and 1425 available family members. 19 families had more than one (2-4) affected individual and were therefore informative for linkage. Later, 1325 control samples were obtained from unrelated healthy Finnish subjects and genotyped together with additional 169 Finnish cases and case-control collections from Denmark, Norway and Sweden, which were obtained through collaborators at Copenhagen University Hospital, Ullevål University Hospital, Oslo, and Karolinska University Hospital, Stockholm.

In study II 803 unrelated patients and approximately 1080 unrelated healthy controls were genotyped.

In study III 4104 cases and 4128 controls from Denmark, Norway, Sweden and the United Kingdom were genotyped. The Scandinavian samples were obtained through the same collaborators as in study I and the samples from United Kingdom through collaborators at the University of Cambridge.

All Finnish patients had been previously recruited through the central hospitals in Helsinki, Kuopio, Tampere, Oulu, Rovaniemi and Seinäjoki under ethical approval from the Helsinki University Hospital Ethical Committee of Ophthalmology, Otorhinolaryngology, Neurology and Neurosurgery (permit 192/E9/02). All patients had been diagnosed according to Poser's criteria for MS and had provided their informed consent.

RNA samples

RNA samples were collected on two occasions in 2006 and 2007. In 2006, a sample collection was organised in Helsinki, and PBMCs were collected from 55 MS patients (38 females, 16 males) using BD Vacutainer® CPT™ Tubes (Becton, Dickinson and Company, NJ, USA). A portion of each sample was further separated into CD4+ T cells and non-CD4+ T cells using the magnetic AutoMACS cell

separator (Miltenyi Biotec, Germany). In 2007, PBMC samples were collected from 19 patients in Seinäjoki. Control RNA samples were either obtained from the Finnish Twin Study on Ageing (FITSA), an existing collection of samples from elderly Finnish female twins, or collected from volunteer female laboratory staff members. RNA was extracted from the FITSA PBMCs and Seinäjoki patient PBMCs with TRIzol@ (Invitrogen, CA, USA) followed by purification with Qiagen Mini Rneasy kit (Qiagen, Hilden, Germany). RNA from PBMCs from the Helsinki patients and staff members was isolated with either Qiagen Mini or Qiagen Plus Mini Rneasy kits (Qiagen, Hilden, Germany). Sample concentrations and 260/280 and 260/230 ratios were determined with Nano Drop spectrophotometer (Thermo Scientific, DE, USA). The integrity of RNA was confirmed with 2100 Bioanalyzer (Agilent, CA, USA). All samples were stored in -80˚C.

Since all control subjects were female, only samples from female patients were selected for the microarray experiment and replication stage.

Table 7. Study samples in study I.

A. Finnish MS families	
Total Cases	730
Cases with both parents (and healthy sibling(s))	340
Cases with one parent (and healthy sibling(s))	286
Cases with no parents, at least 1 healthy sibling	104
B. Finnish case-controls	
Total Cases	899
Cases from Finnish MS families	730
Additional cases	169
Controls (Helsinki, Kuopio, Seinäjoki)	1325
C. Combined Scandinavian case-controls	
Total Cases	1521
Denmark	488
Norway	348
Sweden	685
Total Controls	1476
Denmark	468
Norway	369
Sweden	639

Table 8. Finnish samples in study II. SNPs were genotyped in two multiplex assays. Control samples genotyped with the first and second multiplex differed by 96 samples and two different figures are therefore provided. n.a. = not available.

Region	Cases	Controls
Seinäjoki	132	365/385
Helsinki	115	361/413
Kuopio	98	359/279
Other (Oulu, Rovaniemi, Tampere)	458	n.a.
Total	803	1085/1077

Table 9. Samples in study III.

Population	Cases	Controls
Denmark	580	560
Norway	617	1018
Sweden	678	579
United Kingdom	2229	1971
Total	4104	4128
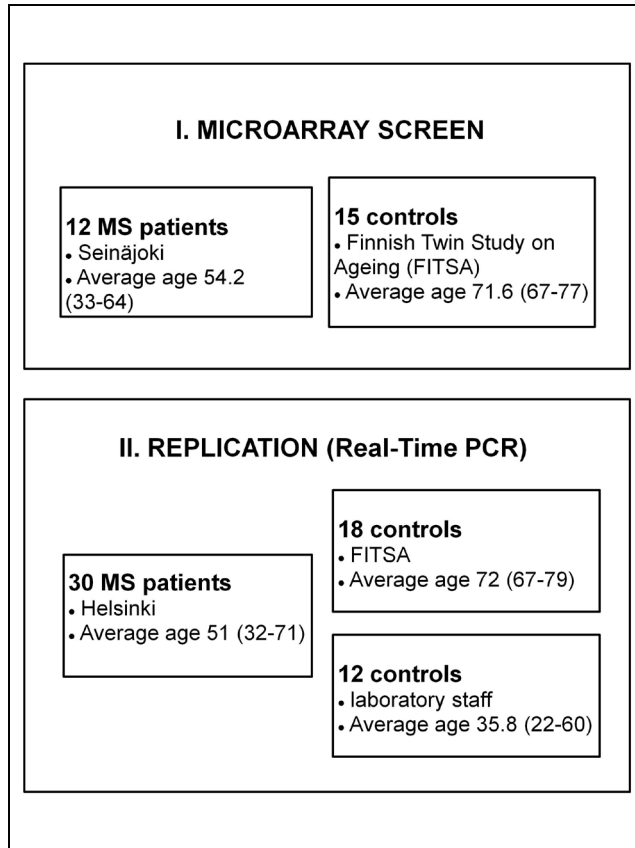
Figure 8. Samples in the Finnish microarray screen and replication in study II.

4.2 Genotyping

Genotyping in studies I and II was conducted using MassARRAY® iPLEX Gold genotyping platform, which is based on single nucleotide extension and separation of extension products by mass spectrometry (Sequenom, Inc., CA, USA) (Figure 9.). Shortly, after amplification of an approximately 100 bp segment of DNA flanking the SNP, an extension primer complementary to the sequence just upstream of the SNP binds to the amplified fragments and a complementary nucleotide is added at the SNP position (single nucleotide extension). The mass of the extension product depends on the added nucleotide, i.e. on the SNP allele on the template DNA. The extension products are separated by their mass using matrix assisted laser desorption ionization time-of-flight (MALDI-TOF) method and the allele(s) can then be identified based on the detected mass(es). Genotyping is conducted on 384 well plates using multiplex assays with currently up to 40 SNPs.

The Finnish genome-wide SNP data in study II was produced with the Illumina HumanHap300 genotyping chip using the Illumina Infinium assay (Illumina, Inc., CA, USA). This assay is based on hybridizing a whole-genome amplified sample to locus-specific 50-mer probes (covalently bound to beads on the chip), which are designed to stop one base before the SNP marker. Following single-base extension with differentially labeled nucleotides and fluorescent staining, alleles can be detected based on signal intensities.

TaqMan® allelic discrimination technology was used for genotyping in study III. Briefly, allele-specific probes recognizing the SNP region and labeled with different fluorophores (usually VIC and FAM, for example) at their 5' end are added into the PCR reaction with amplification primers flanking the probe binding site. The labeled probes also have a so called quencher attached to the 3' end, i.e. a molecule that eliminates the fluorophore signal when in close proximity. As the DNA strand extension reaches the probe during PCR amplification, Taq DNA polymerase's 5'-nuclease activity cleaves the probe and as a result, the fluorophore and quencher become dissociated and the fluorophore signals become detectable. Genotypes are then determined based on signals from the two fluorophores.
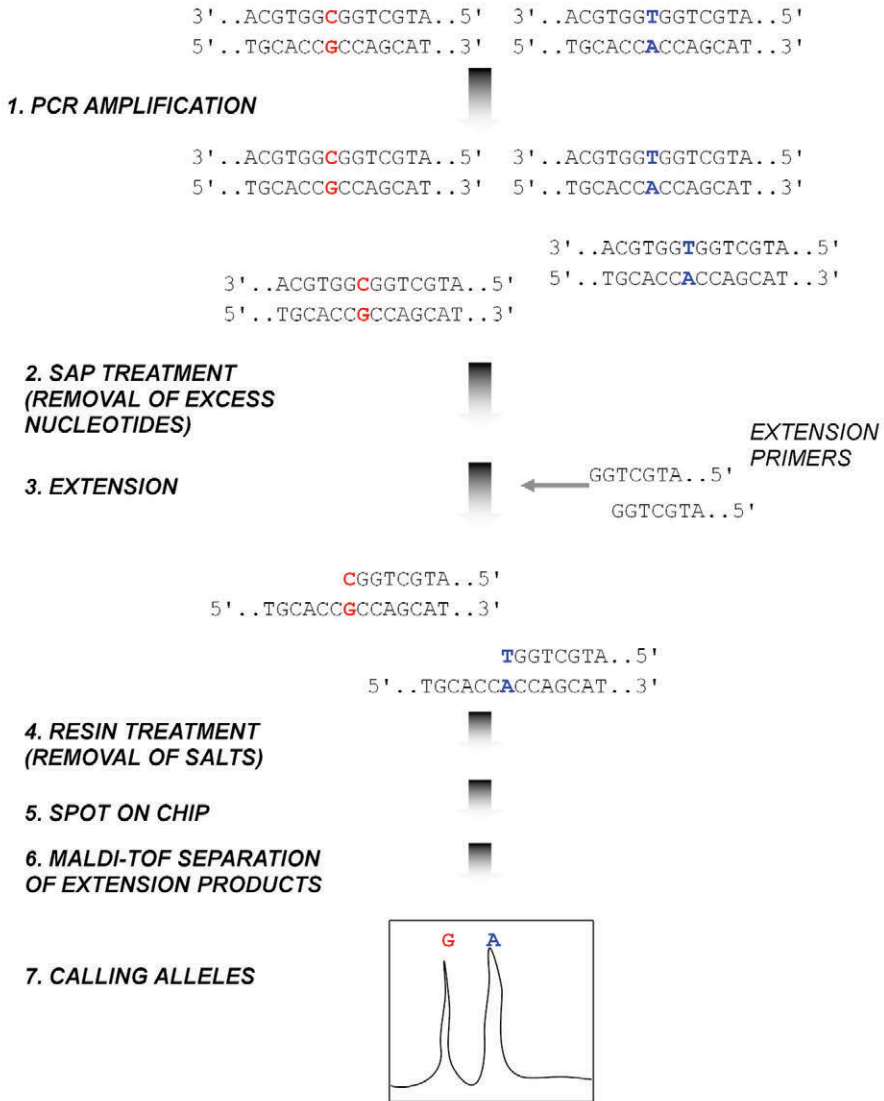
Figure 9. Sequenom MassARRAY® iPlex Gold genotyping workflow.

4.3 Microarray profiling

Genome-wide expression profiling was conducted using the Affymetrix Human Genome U133 Plus 2.0 microarray (Affymetrix, CA, USA), which measures the expression levels of over 47,000 transcripts, including 38,500 well-characterized human genes. Each so called probeset on this chip consists of 11 pairs of 25 bp long perfect match (PM) and mismatch (MM) probes, each present as millions of copies on each array. The MM probes differ from the PM probes by having the complement nucleotide at the 13th nucleotide position and are used for measuring unspecific binding. A probeset thereby consists of 22 probes, measuring 11 different 25-bp sequences within a given transcript. There can be several probesets for a single gene, in which case they might measure different splicing- or polyA-variants.

Total RNA samples were prepared for microarrays at two different laboratories by their respective staff. All 15 control samples were processed at the Biomedicum Biochip Center (BBC) core facility, while 12 patient samples were processed at the National Public Health Institute. In addition, one patient sample and technical replicates for two patient samples were processed at BBC. BBC equipment was used for hybridization, washing and scanning of all arrays. All samples were processed according to Affymetrix's recommendations.

4.4 Real-Time polymerase chain reaction

Real-time polymerase chain reaction (Real-Time PCR), also called quantitative PCR, is a PCR method based on using either TaqMan® probes labeled with a fluorescent dye (see 4.2) or SYBR® Green detector, which is a cyanine dye unselectively binding to all double-stranded DNA. The method differs from standard PCR in that the amplified product can be detected and quantified by measuring fluorescence at each amplification cycle. Because the level of fluorescence depends on the number of copies of the target sequence in the sample, the method can be used for estimating transcript and DNA copy numbers based on the amplification cycle at which a given fluorescence threshold is reached, i.e. the threshold cycle (Ct). In order to account for variation in the amount of total DNA or RNA in the reaction, the Ct values are corrected using an endogenous control gene such as *ACTB*, encoding actin protein, in which there is assumed to be no quantitative variation between samples. The fold-difference in the amount of target sequence copies between two samples can then be estimated using the comparative Ct method according to:

$2^{\wedge}(\Delta Ct(Sample1)-\Delta Ct(Sample2))$,

where ΔCt is the Ct(target gene)-Ct(endogenous control gene). This methods assumes, however, that the amplification efficiency is 100%, i.e. that in each cycle the amount of amplified product doubles. A more accurate alternative to this method would therefore be to use standard curves to quantify the amount of target sequence in each sample.

We used Real-time PCR for measuring transcript levels and relative genomic copy numbers in study II. In both cases, detection was based on SYBR® Green chemistry and amplification was conducted using the ABI7900 thermocycler (Applied Biosystems, CA, USA).

4.5 Data analysis

The most frequently used databases and analysis tools are listed in Tables 10 and 11.

Analysis of SNP variants (studies I-III)

Prior to analysis quality control checks were performed to identify any SNPs where genotyping problems may have occurred. First, genotyping success rate (percentage of samples with successfully called genotype) was assessed and any SNPs with success rate <95% were excluded. Second, a test for deviation from Hardy-Weinberg equilibrium was performed. Hardy-Weinberg equilibrium is the expected distribution of genotype counts given the observed allele frequencies in the population. Expected distribution is based on the assumption of random Mendelian segregation of chromosomes from parent to offspring, and significant deviations from this distribution may indicate problems in genotyping or genotype calling. Where family data was available, each SNP was also assessed for Mendelian errors, i.e. for situations where family members' genotypes are not compatible with inheritance. Again, multiple Mendelian errors at a single SNP would indicate problems in genotyping or genotype calling.

To test for association in families, two tests were applied. The principle of TDT, which requires that both parents of the affected individual have been genotyped, has been presented earlier in Figure 4. The TDT test was performed both on single SNPs as well as on haplotypes defined using Haploview's solid spine option (Barrett *et al*. 2005). Association at individual SNPs was also tested using the gamete-competition test (Sinsheimer *et al*. 2000), which is equivalent to TDT if all parents have been genotyped. However, in case of missing data the method estimates these genotypes according to their population frequencies. The gamete-competition test is also applicable to quantitative traits, and covariates can be incorporated into the analysis.

Finally, a non-parametric test for linkage was conducted in 19 families with more than one affected individual using the Merlin analysis package (Abecasis *et al*. 2002). The test is based on a LOD score calculation method developed by Kong and Cox (1997).

In case-control datasets SNPs were tested for association by a two-tailed X^2 test with one degree of freedom. Furthermore, meta-analyses of datasets were performed by a Cochran-Mantel-Haenszel (CMH) test (Mantel 1963) either using PLINK (Purcell *et al*. 2007) or the R 'Stats' package (www.r-project.org/). Because this test assumes that the effect of the variant is not significantly different across datasets, a Cochran's Q test for homogeneity of ORs (Cochran 1954) was performed before the data was combined. The Q test is implemented in R 'meta' package.

Analysis of microarray and Real-Time PCR expression data (study II)

In order to account for technical variation between microarrays (i.e. amount of starting RNA, and differences in efficiencies of reverse transcription, labeling and hybridization), raw signal intensities were first normalized. In this study, the data was normalized using the GC-RMA method available in GeneSpring 7.3 (Agilent, CA, USA). GC-RMA is a modification of the RMA (Robust Multi-Array) normalization method, but unlike RMA it takes into account the GC content of a probe (GC rich sequences are expected to show higher intensities due to a higher binding efficiency). Both RMA and GC-RMA perform a simultaneous within- and between-array normalization across all arrays in a given set and summarize signals of the 11 individual probes to a normalised probeset signal, which is used in the actual analysis.

After normalisation the data was filtered in order to exclude probesets with low expression and/or affected by differences between the laboratories (see Figure 13 under 5.2.1 for details). Differentially expressed genes (DEG) were identified by comparing average expression levels in cases and controls and by conducting a non-parametric Mann-Whitney test to obtain P-values for statistical significance. This test does not make any assumptions about distribution of data like parametric tests such as t-test, which should only be applied to normally distributed datasets. P-values were then corrected for multiple comparisons using Benjamini-Hochberg false discovery rate (FDR). All filtering steps and statistical analyses were conducted using GeneSpring 7.3.

Finally, pathway analysis was performed using the Core Analysis option in Ingenuity Pathway Analysis (IPA) tool (Ingenuity Systems, Inc., CA, USA). This option tests for association between a user-provided list of genes and genes in canonical pathways stored in IPA by performing Fisher's exact test, which provides

a probability that the given genes are associated with the genes in the pathway by chance alone.

Statistical significance for difference in gene expression between cases and controls in data obtained by Real-Time PCR was evaluated by Mann-Whitney test.

Table 10. Most frequently used databases. Website is provided if the database is freely accessible through this site.

Name (reference if available)	Website	Studies in which used
eQTL Browser	http://eqtl.uchicago.edu/Home.html	II
GeneCards	http://www.genecards.org	I-III
Gene Expression Omnibus	http://www.ncbi.nlm.nih.gov/geo	II
GHS_Express (Zeller et al. 2010)	http://genecanvas.ecgene.net/uploads/ForReview	II
HapMap (International HapMap Consortium 2005)	http://hapmap.ncbi.nlm.nih.gov	I-III
mRNA by SNP Browser (Dixon *et al*. 2007, Moffatt *et al*. 2007)	http://www.sph.umich.edu/csg/liang/asthma	II
NetAffx	http://www.affymetrix.com/analysis/index.affx	II
UCSC Genome Browser	http://genome.ucsc.edu/	I-III

Table 11. Most frequently used analysis tools. Website is provided if the tool is freely accessible through this site.

Name (reference if available)	Website if available	Studies in which used
AssayDesign (Sequenom, Inc., CA, USA)		I,II
GCOS (Affymetrix, CA, USA)		II
Genetic Power Calculator (Purcell *et al*. 2003)	http://pngu.mgh.harvard.edu/~purcell/gpc	I-III
GeneSpring (Agilent, CA, USA)		II
Haploview (Barrett *et al*. 2005)	http://www.broadinstitute.org	I-III
Ingenuity Pathway Analysis (Ingenuity Systems, Inc., CA, USA)		II
Mendel (Lange *et al*. 2001)	http://www.genetics.ucla.edu/software/mendel	I
PedCheck (O'Connell and Weeks 1998)	http://watson.hgen.pitt.edu/register/docs/pedcheck.html	I
PLINK (Purcell *et al*. 2007)	http://pngu.mgh.harvard.edu/~purcell/plink/gplink.shtml	I-III
Primer3 (Rozen and Skaletsky 2000)	http://frodo.wi.mit.edu/primer3	II
R software environment	http://www.r-project.org	II
SDS 2.2 (Applied Biosystems, CA, USA)		II,III
SNP Annotation and Proxy Search (SNAP) (Johnson *et al*. 2008)	http://www.broadinstitute.org/mpg/snap	II
SNPper (Riva and Kohane 2002)	http://snpper.chip.org	I,II
TypeAnalyzer (Sequenom, Inc., CA, USA)		I,II

5 Results and discussion

5.1 Testing association of *MYO9B* variants with MS (study I)

5.1.1 Association analysis in Finnish trios

At the time of initiating this project in 2005, replicated associations of SNP variants in *MYO9B* encoding myosin IXB protein had been reported in CD and UC (Monsuur *et al.* 2005, van Bodegraven *et al.* 2006). With the hypotheses that the same variants or loci may predispose to different autoimmune diseases, in the first part of this thesis the role of *MYO9B* in MS was investigated. While the previously reported associations mapped to the 3' end haploblock of *MYO9B*, variants from the entire gene were included in this study, taking into account the possibility of allelic heterogeneity.

First, 18 SNPs were genotyped in 730 Finnish MS families (Figure 10). Because these included only 340 full trios, i.e. trios where both parents had been genotyped, a Gamete Competition test which is able to use families with missing genotypes, was conducted in addition to TDT test which only uses full trios. No significant association was observed and only a single nominally significant association with a relatively rare variant (MAF=0.09) was identified (P-value =0.03).

The families were then stratified according to a previously identified two-SNP risk haplotype in the *PRKCA* gene (defined by SNPs rs887797 and rs1010544), which at that time was the only non-MHC risk variant identified in the Finnish population. There was also a function-based rationale behind this stratification: MYO9B protein is known to downregulate RhoA (Post *et al.* 1998), which is central in a pathway regulating brain endothelial permeability, and this pathway also involves PRKCA (Stamatovic *et al.* 2006) (Figure 11). After stratification, a significant association with the SNP rs12986130 was identified in families where the patient did not carry a copy of the *PRKCA* risk allele (P-value =0.0001) (A.K. *et al.* unpublished results).
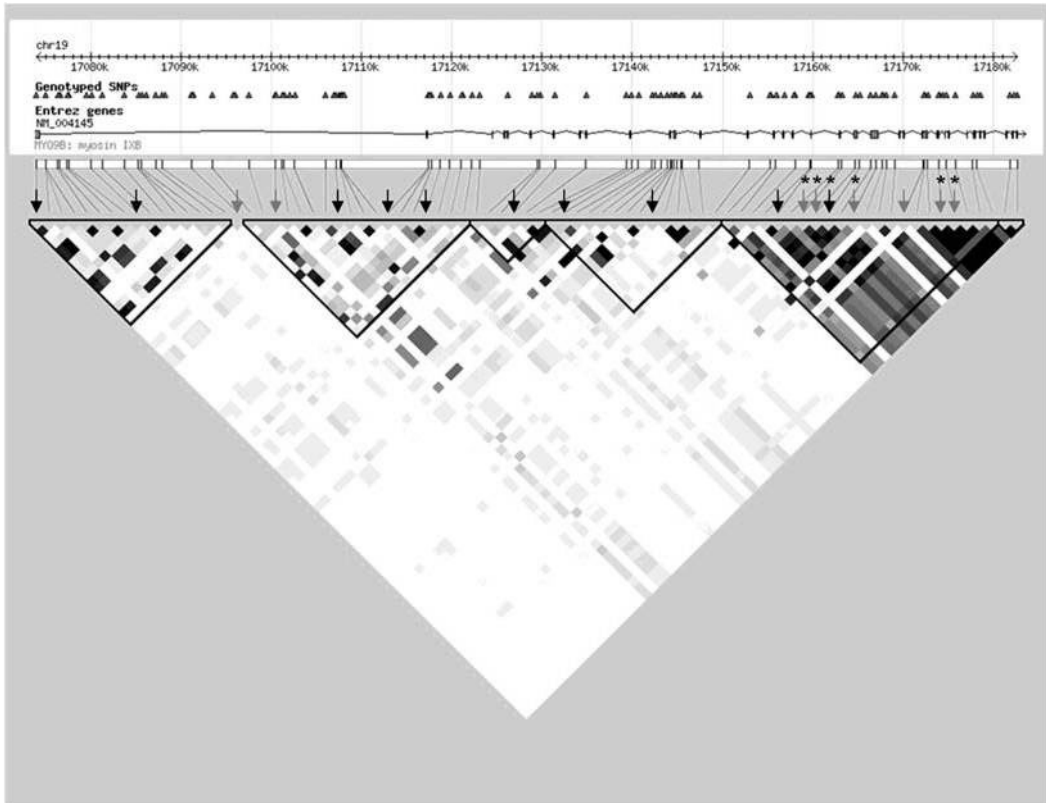
Figure 10. Haploblock structure of *MYO9B* in HapMap CEPH samples. Figure was produced with Haploview (Barrett *et al*. 2005). SNPs genotyped in all samples are indicated with black arrows and SNPs genotyped only in Finnish MS families are marked with grey arrows. SNPs associated with CD or UC are marked with asterisks. Colouring indicates the extent of linkage disequilibrium (LD) between SNP pairs, based on r^2 as the measure of LD. Darker color indicates higher LD. Haploblocks, i.e. regions of relatively high LD between SNPs, are defined with triangles. In this case haploblocks were defined using the "solid spine" option, which identifies blocks so that the first and last SNPs in a block are in strong LD with all intermediate markers, but these intervening markers are not necessarily in LD with each other.
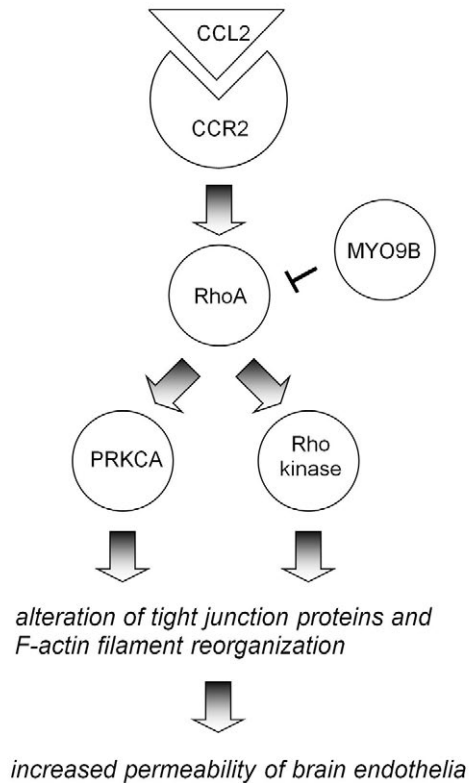
Figure 11. A potential scenario depicting how MYO9B may be involved in regulating BBB permeability via a CCL2 induced pathway. Modified from Stamatovic *et al.* (2006) by adding MYO9B, which is known to down-regulate RhoA (Post *et al.* 1998). Chemokine CCL2 binds to its receptor CCR2, which is present on surface of brain endothelial cells. This results in activation of RhoA which further activates Rho kinase and PRKCA. Rho kinase then induces the reorganization of cell actin cytoskeleton and tight junction proteins, which connect brain endothelial cells, while PRKCA causes direct phosphorylation of tight junction proteins and their redistribution. As a result, brain endothelial barrier permeability is increased, which facilitates leukocyte movement into CNS. CCL2=chemokine (C-C motif) ligand 2, CCR2=chemokine (C-C motif) receptor 2, PRKCA=protein kinase C alpha, RhoA=Ras homolog gene family, member A.

## 5.1.2 Association analysis in Northern European cases and controls

Acknowledging the low power of the family sample to detect association at variants with modest effect sizes, and in order to further investigate the potential PRKCA-MYO9B interaction, 10 of the variants (a single multiplex) were tested for association in a larger sample of 2420 cases and 2801 controls from Finland, Denmark, Norway and Sweden. No significant association was observed in any individual population or in a Cochran-Mantel-Haenszel-test treating populations as separate strata. The cases were then stratified according to the *PRKCA* risk allele carrier status by using two-SNP genotype combinations, and rs12986130 was tested for association with MS in the risk allele negative group. No significant association was observed in any individual population or in CMH-test.

5.1.3 Discussion

Although the results presented above do not provide support for a major role of *MYO9B* variants in MS, the possibility of an epistatic effect between *PRKCA* and *MYO9B* and its role in MS cannot be completely excluded. Neither gene has been reported in recent MS GWASs, but again this does not exclude the role of their potential interaction. However, as the initial finding in Finnish families suggesting association of *MYO9B* after conditioning on *PRKCA* MS-risk haplotype was not replicated, it is unlikely that this finding was a true positive.

Regarding the suggested role of *MYO9B* in other autoimmune diseases, after the first report of *MYO9B* association with CD (Monsuur *et al.* 2005), four studies have found association at *MYO9B* variants with inflammatory bowel disease (IBD, UC and/or ChD) (van Bodegraven *et al.* 2006, Núñez *et al.* 2007, Cooney *et al.* 2009, Latiano *et al.* 2008), one study with schizophrenia (Jungerius *et al.* 2008), one with refractory CD (Wolters *et al.* 2007), one with T1D (Santiago *et al.* 2008), and one study with CD, SLE and/or RA (Sánchez *et al.* 2007). On the other hand, six studies in CD (Latiano *et al.* 2007, Cirillo *et al.* 2007, Núñez *et al.* 2006, Giordano *et al.* 2006, Amundsen *et al.* 2006a, Hunt *et al.* 2006), one in UC (Amundsen *et al.* 2006b) and one in T1D (Persengiev *et al.* 2010) failed to find association. In addition, a Finnish-Hungarian study failed to find association with CD, although evidence for linkage was found (Koskinen *et al.* 2008). As of today, no functional studies have been conducted to further explore the potential role of *MYO9B* gene in these diseases, and according to the GWAS Catalogue (http://www.genome.gov/gwastudies/) the gene had not been identified in any GWASs (August 2011).

In conclusion, these results do not provide support for a major role of the tested *MYO9B* variants in MS, although the study would have been underpowered to detect association at these variants if their effect sizes were low, especially in case of SNPs with a relatively low MAF (Figure 12). For example, assuming homozygote OR =1.3 the power to detect association at a variant with RAF =20% was reasonable, 80%, but at a variant with RAF =5% only 35%. Therefore the power to detect association at two of the ten tested variants, in which MAF was <5%, was low unless their homozygote ORs were >1.5. With a lower homozygote OR =1.2 the power did not exceed ~60% even at common variants. Finally, as the SNPs genotyped in this study captured only approximately 50% of common variation (MAF $\geq$0.05) in the gene with r^2 $\geq$0.7, the role of remaining *MYO9B* variants in MS cannot be excluded.
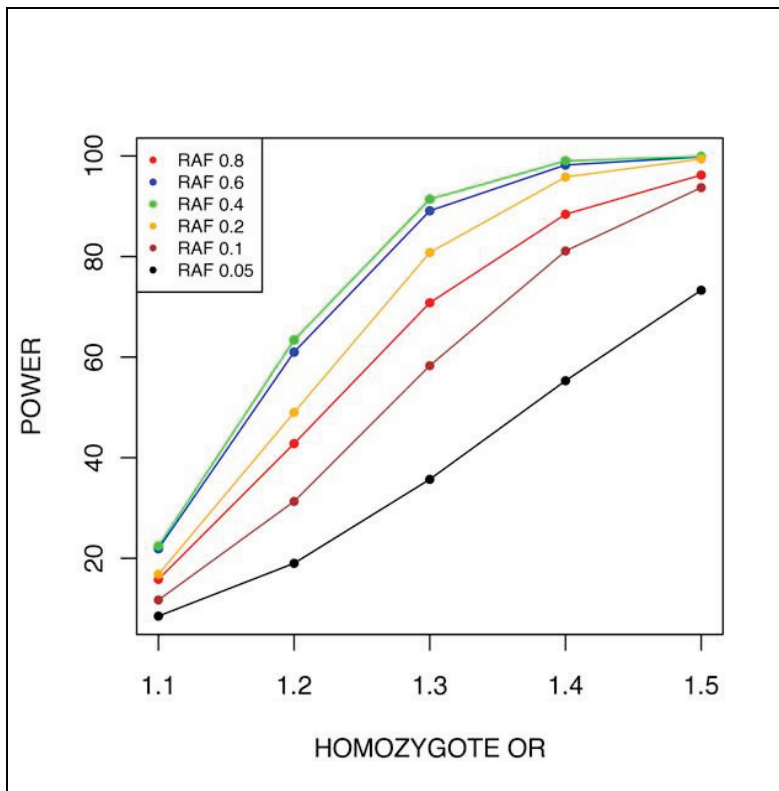
Figure 12. Power to detect association in the Northern European sample for different risk allele frequencies (RAF) as a function of risk allele homozygote OR.

5.2 Identifying MS-related genes and pathways through studies of genome-wide association (study II and unpublished results)

5.2.1 Finnish microarray study

In the second study the aim was to identify novel MS candidate genes and pathways through genome-wide expression profiling. First, a study in PBMC samples from 12 Finnish MS patients and 15 controls was performed. Outline of this study is presented in Figure 13. In this screen, 926 probesets showing a minimum fold change of 1.5 in average signal between cases and controls and statistically significant after correcting for multiple testing were identified. Only 692 of these had an unambiguous gene annotation based on comparing annotations obtained from GeneSpring and NetAffx. Of these, 301 probesets were upregulated and 392 downregulated in MS patients. Surprisingly, three of the four most strongly upregulated genes (*CLC*, *DEFA1A3* and *LTF*) are known to be highly expressed in granulocytes. Pathway analysis using IPA was also conducted on the DEGs and revealed a strong association with several pathways including PI3K signaling in B-lymphocytes and B cell development pathways (Table 12).

Studies on causes of
multiple sclerosis

15 CONTROLS
12 MS PATIENTS
2 TECHNICAL REPLICATES (MS)

DATA
NORMALIZATION

GC-RMA AND PER GENE
NORMALIZATION
54613 probesets

FILTERING

EXCLUDE PROBESETS WITH LOW
EXPRESSION
-> *19410 probesets*

EXCLUDE PROBESETS AFFECTED BY
TECHNICAL FACTORS
-> *15273 probesets*

IDENTIFYING
DIFFERENTIALLY
EXPRESSED
GENES

IDENTIFY PROBESETS WITH >1.5-FOLD
DIFFERENCE BETWEEN CASES AND
CONTROLS
->*1002 probesets*

IDENTIFY STATISTICALLY SIGNIFICANT
DIFFERENCES
-> *926 probesets*

INCLUDE ONLY PROBESETS WITH A
RELIABLE GENE ANNOTATION
-> *692 probesets*

SELECT CANDIDATES FOR REPLICATION
(*DEFA1A3, IGHM, IL8, LILRA4, TNFRSF25*)

FOLLOW-UP
(Real-Time PCR)

VALIDATE IN ORIGINAL SAMPLES

REPLICATE IN INDEPENDENT SAMPLES

SUCCESSFUL REPLICATION OF
DEFA1A3, LILRA4 AND *TNFRSF25*

Figure 13. Summary of the microarray study. To exclude probesets with low expression, all probesets in which normalised signal was <50 in at least 20 of 27 arrays were filtered out. To further exclude probesets potentially affected by technical factors, also probesets in which the expression showed >1.4-fold difference between the technical replicates in both replicate pairs were excluded. In addition, probesets where the signal in all three MS samples prepared at the BBC ranked among the four lowest or four highest among the MS samples were excluded. It is acknowledged that these criteria are arbitrary. However, whereas before filtering the technical replicates from the same sample clustered apart and the samples seemed to primarily cluster according to where the samples had been processed, after filtering the technical replicates clustered together, suggesting that the filtering steps were successful in removing a significant fraction of probesets in which the signal may have been affected by technical factors.

Table 12. Top 20 canonical pathways associated with genes showing differential expression in MS in the Finnish microarray study.

Pathway	P-value
PI3K Signaling in B Lymphocytes	1.3E-06
B Cell Development	2.6E-06
Altered T Cell and B Cell Signaling in Rheumatoid Arthritis	3.4E-06
Role of PKR in Interferon Induction and Antiviral Response	4.6E-05
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	7.8E-05
Death Receptor Signaling	1.2E-04
CD27 Signaling in Lymphocytes	3.4E-04
NF-κB Activation by Viruses	3.6E-04
Molecular Mechanisms of Cancer	3.9E-04
Glucocorticoid Receptor Signaling	4.4E-04
CD28 Signaling in T Helper Cells	4.6E-04
Dendritic Cell Maturation	5.0E-04
B Cell Receptor Signaling	5.0E-04
Type I Diabetes Mellitus Signaling	5.1E-04
TREM1 Signaling	5.2E-04
Sphingosine-1-phosphate Signaling	5.5E-04
Role of NFAT in Regulation of the Immune Response	5.8E-04
TNFR1 Signaling	6.8E-04
Apoptosis Signaling	9.8E-04
Induction of Apoptosis by HIV1	1.0E-03
Graft-versus-Host Disease Signaling	1.2E-03
Breast Cancer Regulation by Stathmin1	1.2E-03
CD40 Signaling	1.3E-03
IL-8 Signaling	1.4E-03
FcγRIIB Signaling in B Lymphocytes	1.6E-03
Docosahexaenoic Acid (DHA) Signaling	1.6E-03
fMLP Signaling in Neutrophils	1.7E-03

5.2.2 Integrating results from the Finnish microarray study and a Finnish GWAS (A.K. *et al.* unpublished results)

The list of DEGs was then compared with results from a Finnish GWAS, which had just been completed using the Illumina 300K platform (Jakkula *et al.* 2010). This study included 68 MS patients from a high risk subisolate in Southern Ostrobothnia and 136 genetically matched controls. Most patients could be traced back to one of two common ancestor couples in the 16th century, and it was therefore hypothesised that this sample would be enriched for shared MS susceptibility alleles. Of SNPs mapping within 100 kb of our DEGs, 20 with an association P-value <0.001 in the GWAS were identified. Eight of them mapped within 20 kb of their annotated gene and were included in replication in a more heterogeneous Finnish sample of 803 cases and ~1080 controls. However, none of the six successfully genotyped SNPs were replicated with statistical significance in the entire Finnish sample or in an expanded Southern Ostrobothnian sample apart from the SNP at *C6orf136*, which is located between MHCI and MHCII loci. This signal was therefore likely to reflect association coming from the MHCII risk variant, and indeed association at *C6orf136* was only seen in patients with at least one copy of the MS risk allele HLA-DRB1*1501.

5.2.3 Follow-up of findings from the Finnish microarray study (A.K. *et al.* unpublished results)

Replication

In order to experimentally confirm some of the findings from the Finnish microarray study, five genes were selected for replication by Real-Time PCR in 30 MS patients and 30 controls. *DEFA1A3* and *IGHM* were selected to represent granulocyte- and B cell-related genes, respectively, many of which were showing increased expression in MS in the microarray study. Of the genes showing increased expression in MS also *IL8*, which encodes interleukin 8, a cytokine acting as a chemotactic factor for various immunological cells and increasing migration of monocytes across BBB *in vitro*, was selected (Yang *et al.* 2009). Of the genes showing decreased expression in MS *LILRA4* and *TNFRSF25* were selected. *LILRA4* encodes a leukocyte-immunoglobulin-like receptor alpha 4, which is located in the leukocyte receptor gene cluster on 19q13. Many of the genes in this cluster bind with MHCI molecules and thereby likely play a role in antigen presentation (Brown *et al.* 2004). The ligand for *LILRA4* is unknown but interestingly MHCII encoded molecules have been suggested as potential candidates. *TNFRSF25* has been suggested to be involved in removal of auto-reactive T cells in the thymus (Wang *et al.* 2001).

Of the five genes, expression differences in *DEFA1A3*, *LILRA4* and *TNFRSF25* were statistically significant between cases and controls (Figure 14). In addition, *IL8* showed a 1.6-fold difference in expression between cases and controls, but the difference was not statistically significant. *IGHM* showed higher expression in MS patients in comparison with controls from the FITSA collection, but lower expression in comparison with control samples collected from staff members. This may suggest an age-related effect, since the average age in MS patients was lower than in FITSA controls, but higher than in staff controls. Unfortunately it was not possible to test the *IGHM* transcript levels in age-matched controls.
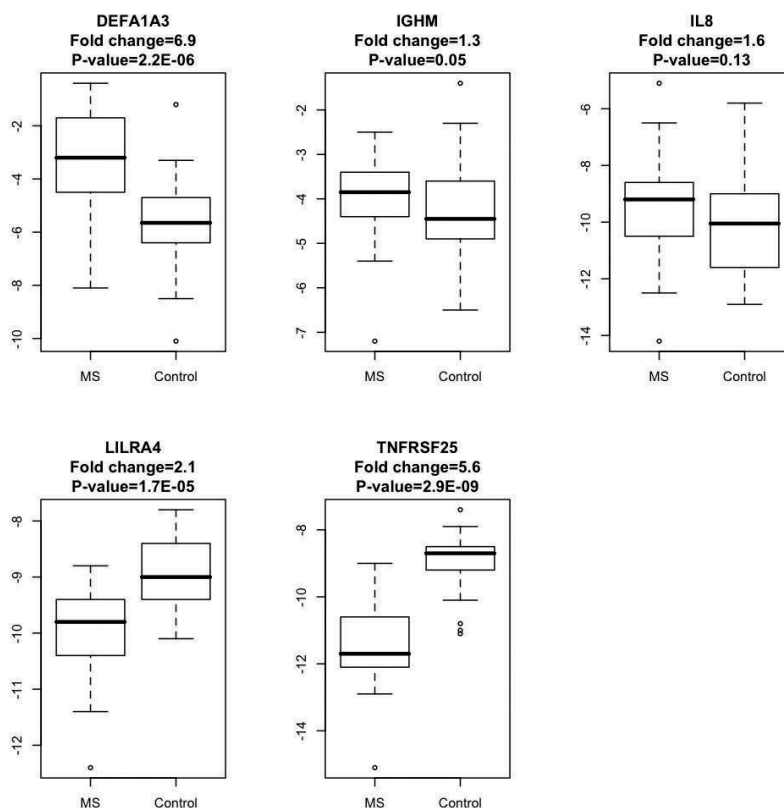
Figure 14. Results from Real-Time PCR replication. The replication set included 30 MS patient samples and 30 controls samples. Y-axis shows the negative of *ACTB* corrected Real-Time PCR Ct values such that expression increases along the axis. A difference of a single unit represents a two-fold difference in expression.

Analysis of genetic variants

The role of common CNV and SNP variants in *DEFA1A3*, *IL8*, *LILRA4* and *TNFRSF25* in MS was then tested. First, common SNP variants in *IL8*, *LILRA4* and *TNFRSF25* were investigated, while *DEFA1A3* was excluded due to previously reported extensive copy number variation in the general population (Aldred *et al*. 2005, Linzmeier *et al*. 2005). The selected SNPs tag 100%, 73% and 80% of variation within 5 kb of *IL8*, *LILRA4* and *TNFRSF25*, respectively, with $r^2 > 0.7$. No significant association was observed in 803 Finnish cases and ~1080 controls although two SNPs, one in *TNFRSF25* and one in *LILRA4*, were nominally significant (uncorrected P-values <0.05) (Table 13). Because some of the cases did not have geographically matched controls in the available control set, which may have distorted the result, these two SNPs were further tested in geographically matched subsets. A meta-analyses of samples from Southern Ostrobothnia, Southern Finland and Savonia revealed a significant association with *LILRA4* SNP rs11084330 (uncorrected CMH P-value =0.002). Interestingly, evidence was also found for an epistatic interaction between this variants and HLA-DRB1*1501 in cases (PLINK epistasis option P-value =0.02), but not in controls. It was then tested whether rs11084330 correlates with *LILRA4* transcript levels. Both genotype and Real-Time PCR expression data was available for 41 MS patients. No significant difference in *LILRA4* expression between individuals with no risk allele versus individuals with at least one risk allele was detected. Finally, rs11084330 was tested for association with MS in 3372 patients and 3720 controls from Denmark, Norway and Sweden. There was no significant association in any individual population or in meta-analysis of all four populations, when including Finland. There was also no evidence for an epistatic interaction with HLA-DRB1*1501 in cases from Denmark and Norway. HLA information was not available for the Swedish cases.

Table 13. Results from association analysis of *IL8*, *LILRA4* and *TNFRSF25*.

Gene	SNP	Alleles Major:Minor	Whole Finland			CMH P-value in three geographically matched subsets
			MAF Cases	MAF Controls	P-value	
IL8	rs2227306	C:T	0.39	0.38	0.65	-
TNFRSF25	rs3138158	C:T	0.09	0.11	0.17	-
TNFRSF25	rs3138156	T:C	0.10	0.08	0.02	0.25
TNFRSF25	rs2986754	C:G	0.04	0.03	0.33	-
LILRA4	rs7259731	G:T	0.15	0.14	0.62	-
LILRA4	rs6417165	C:T	0.21	0.20	0.24	-
LILRA4	rs1205322	A:G	0.33	0.33	0.99	-
LILRA4	rs11084330	A:G	0.20	0.17	0.01	0.002
LILRA4	rs7253755	C:T	0.14	0.16	0.35	-
LILRA4	rs7257994	T:C	0.17	0.18	0.33	-
LILRA4	rs3745419	T:A	0.17	0.18	0.45	-
LILRA4	rs10419832	T:C	0.12	0.12	0.75	-
LILRA4	rs2241385	T:C	0.30	0.32	0.16	-
LILRA4	rs8102662	A:G	0.49	0.48	0.51	-
LILRA4	rs8103715	T:C	0.36	0.37	0.28	-
LILRA4	rs10426462	T:C	0.37	0.37	0.86	-
LILRA4	rs4356595	G:C	0.37	0.37	0.72	-
LILRA4	rs2004431	T:G	0.12	0.10	0.07	-
LILRA4	rs1106659	C:A	0.34	0.37	0.08	-

CNVs had previously been reported in *DEFA1A3*, *LILRA4* and *TNFRSF25* (Iafrate *et al.* 2004, Database of Genomic Variants, http://projects.tcag.ca/variation/). It was therefore tested whether CNVs could contribute to their altered expression in MS. The relative genomic copy number was determined in 18 patients and 18 controls, for which transcript levels had been previously determined and showed a significant difference between the two groups (Figures 15-17). It is not clear whether there were CNVs in *LILRA4* and *TNFRSF25* or whether the observed variance in ΔCt represented technical variation. While *LILRA4* and *TNFRSF25* are expected to be present as two copies in the vast majority of individuals in the population, the

absolute genomic copy numbers of *DEFA1A3* (which in fact is a locus with a variable number of copies of two nearly identical genes, *DEFA1* and *DEFA3*) in the normal population have been shown to range from 5 to 14 in one study with 25% of people lacking *DEFA3* completely (Linzmeier and Ganz 2005), and from 4 to 11 in another study, in which approximately 10% lacked *DEFA3* (Aldred *et al*. 2005). Because reference samples with known *DEFA1A3* copy numbers were not available, determining the absolute copy numbers was not possible. Even with reference samples, the Real-Time PCR based method would not be sensitive enough to detect a difference between 8 and 10 copies, for example, as it's sensitivity is roughly limited to detecting 2-fold differences (for example, 2, 4, 8, 16 copies etc.). However, based on variance in ΔCt, the *DEFA1A3* copy number seemed to vary in the investigated samples, although there was no correlation between ΔCt(mRNA) and ΔCt(genomic DNA), or any significant difference in the average ΔCt(genomic DNA) between cases and controls in any of the three genes. This would indicate that CNVs are not responsible for the detected differences in gene expression between these cases and controls. The lack of correlation between copy number and expression in *DEFA1A3*, as shown in Figure 15, is consistent with previous findings by Aldred *et al*. (2005).
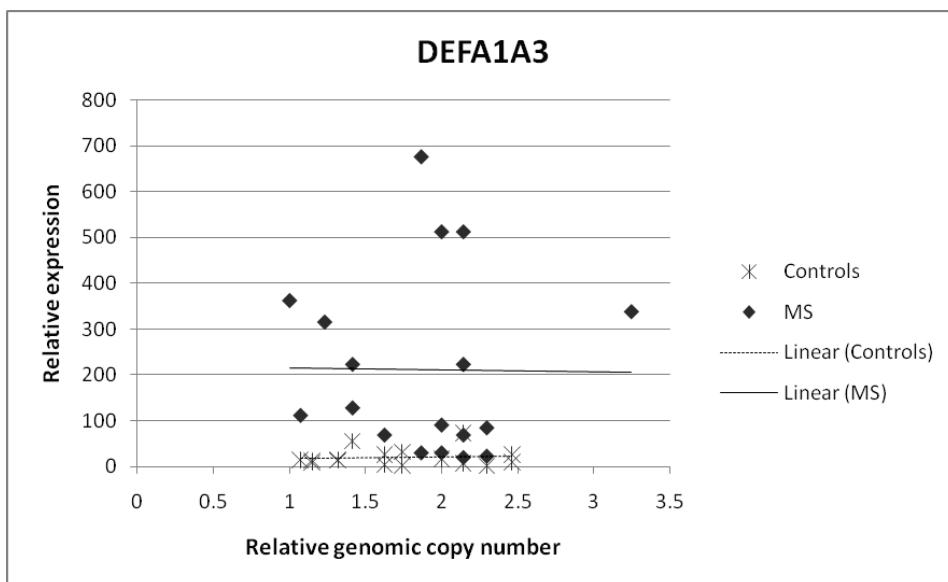
Figure 15. *DEFA1A3* expression levels as a function of genomic copy number. Values were scaled relative to the sample with lowest expression/copy number. For example, a sample with a relative copy number of 500 has a 500-fold expression compared to the sample with the lowest DEFA1A3 expression. Linear regression lines are shown separately for cases and controls.
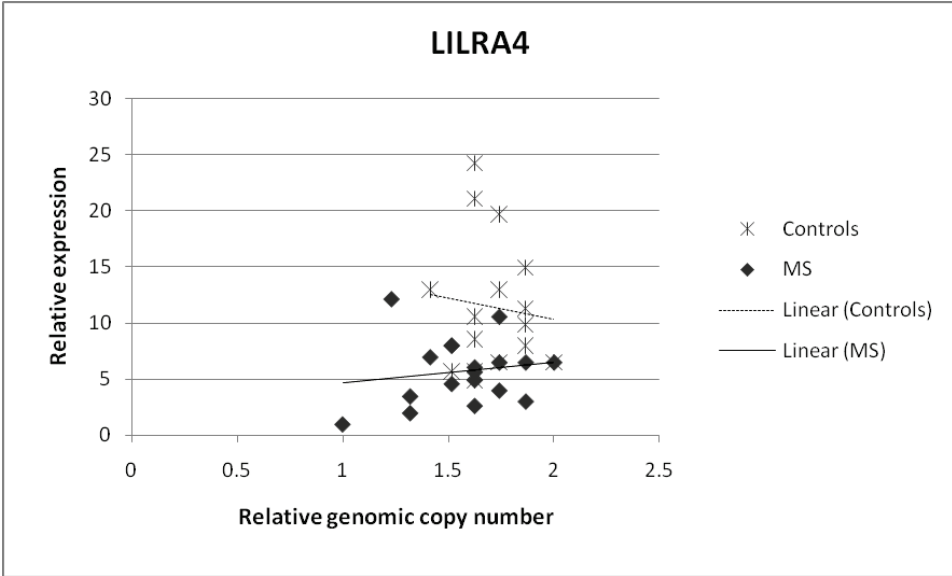
Figure 16. *LILRA4* **expression levels as a function of genomic copy number.** See Figure 15 legend for explanation.
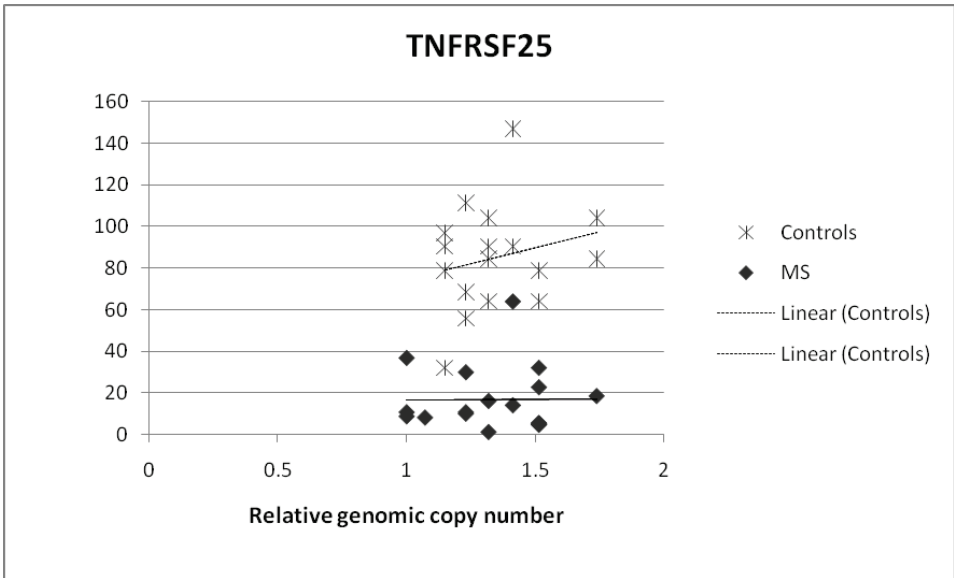
Figure 17. *TNFRSF25* **expression levels as a function of genomic copy number.** See Figure 15 legend for explanation.

 Studies on causes of multiple sclerosis

5.2.4 Systematic review of genome-wide expression studies in MS

In order to identify DEGs reported in several independent studies, genes reported as differentially expressed in MS in six previously published microarray experiments and the study described in 5.2.1 were listed. 229 genes were identified in at least two of the seven studies as either up- or downregulated. 11 of these *in silico* replicated genes were reported in three studies and one, *HSPA1A*, in four studies. Because the study described in 5.2.1 included some treated patients and may also have been confounded by the difference in average age between cases and controls as well as by technical differences in treating the samples, a systematic review was also conducted excluding this study. 135 genes remained as having been reported in at least two studies.

Pathway analysis on both the 135 and 229 DEGs revealed several highly significantly associated pathways, many of which are immune-related. The most highly significant pathway associated with both gene lists was the glucocorticoid receptor signaling pathway. Among the top findings were also several interleukin signaling pathways (IL-4, IL-6, IL-17), of which the IL-6 signaling pathway is shown in Figure 18 (the glucocorticoid receptor signaling pathway is too complex to be shown here).

The list of 229 DEGs was then compared to loci showing evidence of association (P-value ≤0.0001) in five previously published MS GWASs (ANZgene 2009, Baranzini *et al.* 2009, De Jager *et al.* 2009, IMSGC *et al.* 2007, Jakkula *et al.* 2010). Including all SNPs mapping within 100 kb of a DEG, 15 non-MHC loci with association P-value ≤0.0001 were found. These include three loci associated with MS with GWS: *CDK4*, *IL7R* and *TNFRSF1A*. None of the potent risk variants in these 15 loci correlated significantly with expression of the corresponding gene according to the mRNA by SNP Browser or the monocyte study by Zeller *et al*. (2010) apart from *CD40,* where the risk variant was, however, associated with lower expression of the gene, while higher expression in MS was reported in the expression studies. In addition, of the top 102 loci identified in the IMSGC and WTCCC2 GWAS (IMSGC and WTCCC2 2011), seven mapped within 100 kb of the 229 DEGs, of which *NFKB1* and *SAPS2* were not already listed among the 15 genes identified above. A correlation was seen between the risk variant near *SAPS2* and *SAPS2* expression in lymphoblastoid cell lines (mRNA by SNP Browser), and the effect was consistent with the increased expression of *SAPS2* in MS. Expression between risk allele carriers and non-carriers in the 17 genes was also compared using data from a lymphoblastoid cell line expression dataset from 60 HapMap CEPH parents (Spielman *et al*. 2007). Only *CXCR4* showed evidence for a difference in expression between risk allele carriers and non-carriers, with higher expression in the first

group (uncorrected one-sided Mann-Whitney test P-value =0.03). A concordant trend was also observed in the microarray dataset described in 5.2.1 (One-sided Mann-Whitney test P-value =0.06).
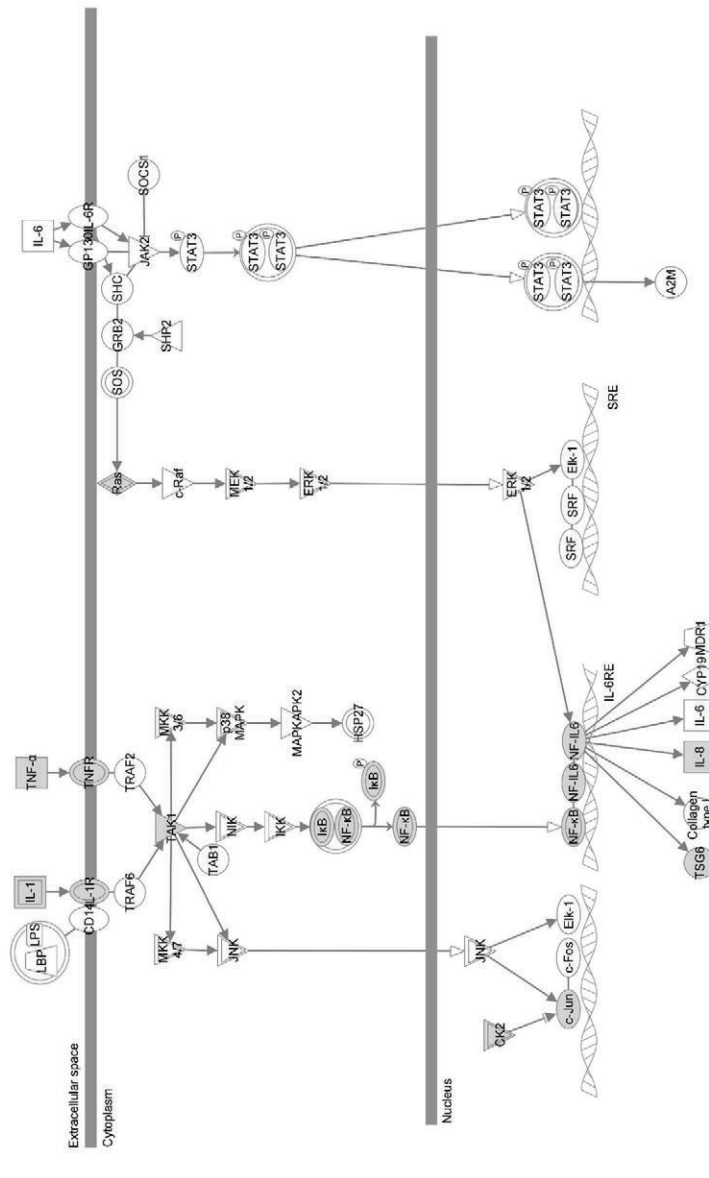
Figure 18. IL-6 signaling pathway. From Ingenuity Pathway Analysis (IPA) (Ingenuity Systems, Inc., CA, USA). Nodes shown in grey are among the *in silico* replicated 229 DEGs. Printed with permission.

5.2.5 Discussion

In this part of the thesis a genome-wide expression study was first conducted to identify genes with differential expression in MS, followed by replication and genetic studies in selected genes. In addition, a systematic review of seven studies on gene expression in MS was performed and the most consistent findings were analyzed in the context of published GWASs and using pathway analysis tools.

The conducted microarray study identified 609 DEGs, although it is likely that many of these could be false positives due to the small sample size and effects of confounding factors (difference in age of patients and controls, technical differences in sample processing). Three of five genes selected for replication in a larger independent sample were successfully replicated: *DEFA1A3*, *LILRA4*, and *TNFRSF25*. However, none of these showed any evidence of association with MS, nor have they emerged in published GWASs. This would suggest that their differential expression in MS is unlikely to be due to *cis*-acting common risk variants.

DEFA1A3, a copy number variable locus which contains two nearly identical genes, *DEFA1* and *DEFA3* (in this study it was not possible to distinguish between the expression of these two genes), showed highly increased expression in MS. These genes have previously been reported to be over-expressed on mRNA or protein level not only in MS, but also in SLE, T1D and IBD (Bennett *et al.* 2003, Ishii *et al.* 2005, Saraheimo *et al.* 2008, Sthoeger *et al.* 2009, Yamaguchi *et al.* 2009). The genes encode alpha defensins 1 and 3, which are secreted by neutrophils and can inactivate bacteria, fungi and viruses (Ganz *et al.* 1985). It is possible that a higher expression of *DEFA1A3* in MS could reflect a higher amount of granulocyte contamination in PBMC samples, potentially as a result of elevated levels of granulocytes in their peripheral blood. However, this is not supported by the observation that only one of 26 neutrophil markers identified by Whitney *et al.* (2003) showed significantly increased levels in MS in the microarray experiment.

The second replicated gene, *LILRA4*, is a marker for pDCs, a class of specialized DCs which produce high levels of type I interferons (Ju *et al.* 2004, Cho *et al.* 2008). *LILRA4* is downregulated when pDCs mature in response to viral or bacterial stimulation (Ju *et al.* 2004). It is therefore possible that the lower expression of *LILRA4* observed in MS patients could reflect lower levels of pDCs or higher pDC stimulation. pDCs are important in regulating peripheral T cell tolerance and in MS there is evidence for imbalance in two pDC subsets, which differentially induce Th2 and Th17 cells (Schwab *et al.* 2010). Whether *LILRA4* plays a role in causing this imbalance remains to be studied.

Finally, we replicated the decreased expression of full-length isoform of *TNFRSF25* in MS. This isoform encodes a transmembrane protein with a death-domain, which can transmit apoptotic signals (Screaton *et al.* 1997) and has been shown to be involved in regulating removal of auto-reactive thymocytes (Wang *et al.* 2001). The shorter isoforms encoding soluble proteins did not show differential expression in MS in our microarray screen and were not investigated in replication samples.

In order to identify additional DEGs supported by evidence from several experiments, six previously published expression studies were reviewed. 132 of the 609 DEGs (22%) found in the expression study described in this thesis had been identified in at least one other study. In the other six studies, the percentage of reported DEGs identified in other studies ranged from 11% to 47%. Most studies were relatively small (10-30 samples per group), which together with differences in the immune cells included in the profiling (whole blood, PBMCs, T cells, non-T cells) is likely to explain the poor overlap between studies. Only 11 genes had been reported to be differentially expressed in MS in three studies and only one gene, *HSPA1A*, in four studies. The *HSPA1A* gene is located in the Major Histocompatibility Complex Class III (MHCIII) region between MHCI and MHCII, and encodes a 70kDa heat shock protein HSP70-1A, which is part of the HSP70 chaperone. The HSP70 chaperone system is involved in various immunological processes: extracellular HSP70 can initiate innate immune responses, acting as a cytokine-like factor, while cellular HSP70 is involved in antigen presentation and thereby in the activation of the adaptive immune system (Chen and Cao 2010). How a lower expression of *HSPA1A* could relate to MS pathogenesis remains to be studied, but one possibility is that a reduced capacity to present auto-antigens to developing T cells in the thymus could result in impaired negative selection and increased release of auto-reactive T cells to the periphery.

Given that many disease risk variants are likely to affect gene expression, findings from expression studies were compared with those from GWASs. Suggestive evidence of association had been reported in 17 of the 229 DEGs, but only in case of *SAPS2* there was evidence for an effect of the risk variant on gene expression based on the mRNA by SNP Browser (listing eQTLs with P-value <0.001) or the GHS_Express monocyte eQTL database (Zeller *et al.* 2010) (listing eQTLs with P-value <0.00001). Not much is known about the function of *SAPS2*, but it has been shown to interact with *NFKBIE* (Stefansson and Brautigan 2006), which is an inhibitor of transcription factor NF-κB and is involved at least in regulating T cell proliferation and B cell differentiation (Doerre *et al*. 2005, Tergaonkar *et al*. 2005, Wang *et al*. 2008). The associated SNP is located 90 kb downstream of the gene next to a CpG island (according to the UCSC Genome Browser). It may therefore well be related to regulation of gene expression, although none of the eight genes

located between the SNP and *SAPS2* were among the 229 DEGs. Finally, in order to detect more modest SNP-expression correlations the 17 genes were examined in an expression dataset from 60 HapMap CEPH parents. Unfortunately *SAPS2* was not represented on this expression array. However, there was nominally significant evidence for higher *CXCR4* expression in carriers of the suggestively associated variant near *CXCR4*. Based on consistent data from the microarray study described in 5.2.1., it seems that the increased expression of *CXCR4* in MS may be contributed to alternative polyadenylation, because a difference in expression between patients and controls was only seen with a probeset extending to the 3' UTR and not with probesets measuring exonic sequences. However, the *CXCR4* risk variant is located relatively far from the gene (100 kb upstream) with no high LD to SNPs within the gene, and it is therefore not clear how it would affect its polyadenylation.

Although there was no evidence for correlation between the potent MS susceptibility variant and gene expression in the other 15 genes in available datasets, large studies in different cell populations can be expected to reveal additional correlations in the future. However, the overall limited overlap between the 229 DEGs and loci identified in GWASs would suggest that the altered expression in most of the DEGs identified is not explained by genetic MS risk variants. Instead, some of these findings may reflect immunological processes activated during MS. The DEGs identified in more than one study were significantly associated with several immunological pathways including the glucocorticoid receptor signaling pathway which is important in controlling inflammation, and several interleukin signaling pathways (IL-4, IL-6, IL-10, IL-17) which are primarily related to Th2, and Th17 cells and Tregs: IL-4 and IL-10 are secreted by Th2 cells (Cherwinski *et al*. 1987, Fiorentino *et al*. 1989), IL17-production is characteristic for Th17 cells (Harrington *et al*. 2005, Park *et al*. 2005) and IL-6 regulates differentiation of naive T cells into Th17 cells while inhibiting their differentiation into Tregs (Bettelli *et al*. 2006) However, at this stage not much more can be concluded from these findings and it should be further investigated whether these pathways are differentially regulated in MS patients already prior to disease onset, or whether their altered activity is a secondary phenomena.

Studies on causes of
multiple sclerosis

5.3 Replicating association of a variant in *CXCR4* locus with MS (study III)

Based on findings in the second part of this thesis, where a modest correlation between *CXCR4* expression and a suggested MS risk variant located 100 kb upstream was found, in the final part of the thesis the aim was to replicate the association of this variant, rs8823000, with MS. This variant showed strong evidence of association with MS in two previous studies, although it did not reach GWS (De Jager *et al*. 2009, IMSGC 2010a). In this thesis, the association of rs882300 with MS was replicated in a total of 4104 MS cases and 4128 controls (P-value =0.0004). However, the association was not confirmed with GWS even when combining with allele counts from 8085 cases and 7777 controls genotyped in a previous study (IMSGC 2010a) (P-value =1.8E-07).

Public eQTL databases (mRNA by SNP Browser, eQTL Browser and the monocyte eQTL database by Zeller *et al*. 2010) were then searched for evidence of correlation between rs882300 or its tagging SNPs (r^2 >0.8) and *CXCR4* expression (as of April 2011), but no reported correlations were found. Furthermore, there was no evidence of correlation in 585 whole blood samples (data from study by Inouye *et al*. 2010, personal communication with J. Kettunen, National Institute for Health and Walfare, Helsinki, Finland). However, none of the eQTL studies conducted in immune cells or lymphoblastoid cell lines were performed with a platform that would have measured *CXCR4* expression at its 3' UTR, where evidence for a difference in expression between risk allele carriers and non-carriers was seen in study II. Data from the published eQTL studies is therefore not contradictory to the observation in study II, although given the location of rs882300 100kb upstream of *CXCR4* with no known correlating variants in the 3' UTR, its effect on alternative polyadenylation doesn't seem very likely. It is also possible that rs882300 exerts its effects via some other gene than *CXCR4*, although none of the other genes within 1 Mb (*ZRANB3*, *R3HDM1*, *MIR128-1*, *UBXN4*, *LCT*, *MCM6*, *DARS*, *THSD7B*) have any known immunologically or neurologically relevant functions.

CXCR4 is an excellent MS candidate gene as it plays a role in various immunological and neurological processes, including thymic selection and T cell differentiation (Suzuki *et al*. 1998, Plotkin *et al*. 2003), B-lymphopoiesis and humoral immunity (Ma *et al*. 1998, Nie *et al*. 2004), and cerebellar development (Ma *et al*. 1998, Zou *et al*. 1998). CXCR4 and its ligand CXCL12 have also been shown to promote trans-endothelial migration of T cells *in vitro* (Liu *et al.* 2009), while antagonists for CXCR3 and CXCR4 inhibit EAE pathogenesis and reduce accumulation of CD4+ T cells in CNS (Kohler *et al.* 2008). Although the association of rs882300 with MS did not reach GWS, it is very likely to be a true MS

susceptibility variant given the consistency of association signal in several independent populations.

5.4 General discussion

In 2005 when this thesis was started, candidate gene studies were the only available approach for identifying SNPs associated with complex traits. Just six years later hundreds of GWASs have been conducted in a number of complex traits and have proven successful in identifying new disease risk variants, including close to fifty variants associated with MS. While the most important outcome of the hypotheses-free GWAS approach has been the identification of new candidate genes that otherwise would not have been considered of any special interest, these studies have failed to provide support for perhaps most of the associations reported in earlier candidate gene studies, including *MYO9B* investigated in this thesis. GWASs have therefore also helped to resolve many of these claimed, but debatable associations at candidate genes, and have set more widely accepted criteria for declaring association.

But what do all these recent genetic findings tell us about MS? Majority of the genes that have emerged in MS GWASs have immunological functions, with the presence of T cell related genes being particularly evident, but not much more can be concluded from genetic studies alone (IMSGC and WTCCC2 2011). Many of the identified candidate genes, including *CXCR4* in which a previously suggested association was replicated in this thesis, play a role in various processes and it will be a challenge to understand in detail how these individual genes contribute to MS susceptibility. Studies correlating the identified genetic variants with gene expression and cellular functions in different cell types will be needed to investigate the potential underlying mechanisms.

Meanwhile, data from existing expression studies comparing gene expression in MS cases and controls can provide further clues towards mechanisms of MS pathogenesis. In this thesis a systematic review of seven such studies showed that genes identified in at least two studies were associated with cytokine pathways related to Th2, Th17 and Treg cells. These findings are consistent with increasing supporting evidence for the role of Tregs and Th17 cells in MS (Matusevicius *et al*. 1999, Viglietta *et al*. 2004, Segal 2010) and with studies questioning the central role of Th1 cells (Ferber *et al*. 1996, Harrington *et al*. 2006). It was also shown in this thesis that a neutrophil-related gene *DEFA1A3* encoding alpha defensins 1 and 3 had highly elevated levels in MS patients. Interestingly, this gene has also been found to be highly expressed in SLE, T1D and IBD suggesting that the largely ignored

neutrophils and this cytotoxic protein in particular could be relevant for autoimmune disease pathogenesis.

However, whether these findings from expression studies reflect primary or secondary events is a more difficult issue to resolve; many genes and pathways become activated after disease onset as the primary event initiating an auto-reactive response results in activation of other components of the immune system. That is not to say, however, that these secondary events are not relevant in MS as some of these are likely to contribute to demyelination and chronic inflammation in MS. Therefore, characterizing not only the causes of MS but also these secondary processes is highly important from the perspective of developing new treatments. In this thesis very little overlap was found in genes identified in expression studies and genes identified in GWASs, which may suggest that most of the genes identified in these expression studies are not involved in processes that affect susceptibility to MS. However, it should be noted that most of the expression studies were conducted in small samples and given that there was little overlap between genes identified in these studies, majority of the genes reported to be differentially expressed in MS may have been false positives rather than MS-related. Much larger samples would be needed to decrease the false positive rate and to increase the power to detect true signals in expression studies, which are more prone to noise due to effects of technical and biological factors than studies on DNA variants.

Finally, although identifying novel candidate genes is always an important step forward, it should be noted that interpretation of these findings is limited and driven by available data on functions of the identified genes and by the current understanding of basic immunology. New immune cell subsets are constantly being characterized and previously unknown roles are associated with genes as they are being investigated in different cell populations and tissues. While our hypotheses and thinking are naturally influenced by existing knowledge, we should not forget that in complex diseases and MS anything is possible - at least until proven impossible.

6 Conclusions and future prospects

In this thesis the aim was to identify MS-related genes and pathways by investigating common genetic variants in selected candidate genes through genome-wide screening of gene expression in PBMCs from MS patients and controls and by systematically reviewing previous microarray studies. As a result of these studies it was shown that:

I.   Expression of *DEFA1A3* was highly elevated in MS patients in two independent study samples. Increased levels have also previously been found in MS and other autoimmune diseases, suggesting a role for neutrophils and innate defense mechanisms in the pathogenesis of autoimmunity.

II.   Findings from seven genome-wide expression studies in MS show relatively little overlap. However, altered expression of 229 genes had been reported in at least two studies, of which *HSPA1A* was reported in four studies.

III.   Genes showing altered expression in MS are associated with several immunological pathways including the glucocorticoid receptor signaling pathway and Th2-, Th17- and Treg-related interleukin signaling pathways.

IV.   Of 17 genes with altered expression in MS and harbouring confirmed or suggestively associated MS risk variants, only *CXCR4* and *SAPS2* showed evidence of correlation between the associated SNP and gene expression.

V.   Consistent with previous data, the variant upstream of *CXCR4* was shown to be strongly associated with MS and is likely to be a true disease risk variant.

Despite considerable progress in MS genetics during the past few years, MS remains perhaps one of the most enigmatic of complex diseases. While it is likely that most common risk variants with moderate effects have now been identified, a considerable fraction of the genetic component still remains unexplained by the current findings. But how much of the missing heritability is explained by additional common variants with more modest effects and how much can be contributed to rare variants with potentially larger effects on risk? The rarity of extended MS families in most populations and lack of consistent and strong findings in linkage studies would perhaps argue against a major role of rare variants in MS. However, families with high MS prevalence do exist and rare variants may play a larger role in such cases than in sporadic MS. Importantly, lack of consistent linkage signals across populations is not inconsistent with family- and population-specific variants. In fact,

it is likely that locus and allelic heterogeneity is an important factor in MS and other complex diseases, where possibly hundreds of genes and several different pathways as well as gene-gene and gene-environment interactions might have an effect on disease susceptibility. With lowering costs of re-sequencing, families with multiple MS patients are likely to be the first to be investigated. Even if rare variants would not play a major role in the majority of sporadic cases, their identification may give novel insights into MS pathogenesis.

In addition to rare variants, the role of CNVs in MS remains to be comprehensively studied. However, recent evidence suggests that common CNVs do not play a major role in complex traits (Conrad *et al.* 2010). Furthermore, many of the CNVs correlate with SNPs and would therefore have been indirectly detected in GWASs. It therefore now seems unlikely that common CNVs would end up explaining much of the missing heritability in MS.

Finally, in addition to inherited DNA variants, the role of epigenetics and somatic and *de novo* mutations needs to be investigated. A recent study sequencing the CD4+ T cell methylome in three MS discordant MZ twin pairs did not reveal any consistent and significant methylation differences, which would indicate that at least in these cells methylation variation is not a major factor influencing MS (Baranzini *et al.* 2010). The same study also failed to find any mutations and transcriptome differences that could have explained the discordance in these genetically "identical" individuals. Similar studies in larger samples and different immunological cell populations should however be conducted in order to confirm these initial observations.

Although the hunt for missing heritability continues, the primary goal should be to understand the disease and its pathogenesis so that more efficient treatments can be developed. Functional studies on the currently established MS candidate genes may already provide important clues towards achieving this goal, while identifying biomarkers or molecular profiles correlating with drug response could also improve the safety and efficiency of existing treatments. With the remarkable recent progress in MS research, the future is at last looking promising.

7 Acknowledgements

This study was carried out at the Public Health Genomics Unit, National Public Health Institute (later National Institute for Health and Welfare; THL) and Institute for Molecular Medicine Finland FIMM. I am grateful to Pekka Puska, director of THL, Anu Jalanko, the head of Public Health Genomics Unit, and Olli Kallioniemi, director of FIMM, for providing such excellent research facilities.

Almost exactly seven years ago I nearly had a heart attack: I received an e-mail from Professor Leena Peltonen-Palotie inviting me for an interview - and I had never been in touch with her! I still don't know exactly how that happened, but I suspect it had something to do with Janna Saarela, whom I had contacted for positions at the Biomedicum Biochip Center, not knowing that she was in any way connected to Leena. And suddenly, after just trying to find a job, I became a PhD student in a group of one of world's most acknowledged human geneticists. It was truly a privilege to have the opportunity to work for such a unique personality. Anyone who has been around Leena will never think that science is boring or for geeks only! However, as much as Leena has been, and always will be, a great inspiration, I would certainly not be here today without my other supervisor Janna. You have always been easy to approach with any stupid question and the tiniest problem, and always found the time to discuss and listen. You have been supportive of our own ideas, which has been incredibly valuable for my development as a researcher. I could not have asked for a better combination of supervisors. Both Leena and Janna are also to blame for another turn of life, Cambridge. I feel extremely lucky to have had the chance to come here. Wherever this road leads, in science or other aspects of life, I will never forget where and how it started.

I would also wish to thank Professor Aarno Palotie who has been closely involved in our MS research and always showed genuine interest in our projects. I greatly appreciate your help with my second article.

I am very grateful to Professor Rikard Holmdahl for accepting the role of opponent. I would also wish to thank my thesis reviewers Adjunct Professor Minna Männikkö and Adjunct Professor Sakari Jokiranta for their valuable comments, which led to

many considerable improvements. Professor Jaakko Kaprio is thanked for agreeing to act as the custos in my dissertation.

I would like to warmly thank my thesis committee members Professor Anu Wartiovaara, Adjunct Professor Päivi Saavalainen and Professor Joe Terwilliger for taking time from their busy schedules. Although we only had two meetings, these were very helpful - and good practice for the dissertation!

I would also like to thank the senior scientists at the department, Iiris Hovatta, Marjo Kestilä, Anu Loukola, Markus Perola, Samuli Ripatti, Kaisa Silander, Ismo Ulmanen and Teppo Varilo for their help over the years.

During this thesis I was fortunate enough to meet some MS patients. That was a humbling experience; there were patients unable to speak and walk and yet they had made their way to the hospital only to give blood in order to participate in our studies. I always try to keep in mind that without the thousands of patients and controls providing samples to research, none of what we know about MS genetics today would have been discovered, and that the ultimate goal of our research is not to publish in high impact journals, but to make discoveries that can eventually help the patients. The studies in this thesis have been based on samples collected from around Finland, Denmark, Norway, Sweden and United Kingdom. I wish to warmly thank the Finnish clinicians Pentti Tienari, Keijo Koivisto, Irina Elovaara, Mauri Reunanen, Ilkka Rautakorpi and the late Tuula Pirttilä. I'm also deeply grateful to our collaborators in the Nordic MS Genetics Network and to Dr. Stephen Sawcer for providing samples. The Wellcome Trust Case Control Consortium is acknowledged for providing control data. Professor Jaakko Kaprio is thanked for providing control samples from the FITSA sample collection.

I greatly appreciate the help of the always friendly and efficient THL secretaries, especially Sari Kivikko, Mika Kivimäki, Tuija Koski and Sanna Tossavainen. Sisko Lietola, thank you for your help with orderings. I also wish to thank Juri Ahokas, Juha Knuuttila, Jari Raikko, Juha Saharinen and Hannu Turunen for their help with IT and database issues.

As I started my thesis with practically no experience of lab work, these projects would never have happened without the help and guidance from the skilled THL lab technicians. I would especially wish to acknowledge Anne Nyberg, Eija Hämäläinen, Siv Knaappila and Lea Puhakka, who helped with RNA sample collections and microarrays, and Minna Suvela for her always efficient help with DNA samples and genotyping. Also Liisa Arala, Elli Kempas and Anne Vikman have kindly provided their help over the years. Lilja Jansson is thanked for collecting the patient blood samples.

I am deeply grateful to Professor Alastair Compston and Dr. Stephen Sawcer for providing the amazing opportunity to work as a "post doc" at the Department of Clinical Neurosciences at the University of Cambridge. You have been incredibly understanding with my PhD being delayed, delayed and once again, delayed... I also wish to thank the past and present members of our MS genetics group, Amie, Anna, Helen, Maria and Sheila, as well as all the visiting students and members of Dr. Alasdair Coles's group, for making it such a pleasant place to work!

A very special thanks to the KTL goes Cambridge unit, past and present: Eija, Heidi, Helena, Henna, Jenny, Johannes, Jonna, Jouni, Karola, Kati, Katja, Katta, Olli, Sanna (and Aada), Verneri and Virpi. Thanks to all of you moving here was easier than I ever expected and I don't think there's been a single day when I would have missed going back to Finland. A special thanks to Katta for listening to my moaning over my PhD at the gym and, perhaps more often, over a bottle of 5 pound cava... I have also met numerous great new people during my stay here. Thanks to all of you for all the memorable (and the less memorable) events, of which I'm sure there will be many more to come!

Missukat: Ansku, Eve, Suvi and Virpi. It was a pleasure to work with such nice, smart and fun people and I'm happy that we still stay in touch. Thanks to you girls, most of the time it didn't feel like being at work at all!

Anna, Annika, Annina, Annu, Antti, Emma N, Emma P, Emmi, Hanski, Ida, Jarkko, Jonas, Joni, Juho, Jussi, Kaisu, Kati, Kirsi, Liisa, Mari, Marika, Marine, Markus L., Mervi, Minttu, Nora, Outi, Pekka, Peter, Pia, PP, Sampo, Tea, Tero, Tiia, Tintti and others who I have already mentioned or may have forgotten: there are not that many jobs where you get to party all night in fancy dresses several times a year or to sip cocktails in Honolulu, Nice or Barcelona with your colleagues (between some science, of course). Thanks for making it such a unique, relaxed and fun working environment! Oh, and thank you also for your help with the various work-related things...

However, I would have gone insane if I didn't also have friends who have absolutely nothing to do with science. I especially want to thank the Peikko and the four M's: Maija, Maitis, Miia and my godson Ossi's mum Misa. I'm grateful to have you crazy people as my closest friends!

Finally, I want to thank my grandmother Raija, my parents Ritva and Jorma, my brother Ville and his wife Leena for all their support and encouragement. My grandmother, amongst other things a big Leena (and Formula 1) fan, has shared my stress and frustrations more than perhaps anyone else and I am so glad that at 93

years of age she is with us to see this process finally come to an end. As a daughter of two engineers I think I must have inherited some kind of need to solve problems. Thanks to my parents, our home has been full of sensibility, optimism, creativity, and, last but not least, humour, and with such a combination there was never an issue that would have been impossible to solve. This is certainly an example to live up to, both in science and in life in general. Finally, I'm afraid my dissertation will remain the number one topic in the family for only a couple of months before another more important family event. Ville and Leena, just make sure my nephew or niece will call me Dr. Aunt ;)

Without all of you I would not be finishing this chapter.

Cambridge, August 14th, 2011

Anu Kemppinen

8 References

1000 Genomes Project Consortium, Durbin, R. M., Abecasis, G. R., Altshuler, D. L., Auton, A., Brooks, L. D., Durbin, R. M., Gibbs, R. A., Hurles, M. E., and McVean, G. A. (2010). A map of human genome variation from population-scale sequencing. Nature 467, 1061-73.

Abecasis, G. R., Cherny, S. S., Cookson, W. O., and Cardon, L. R. (2002). Merlin— rapid analysis of dense maps using sparse gene flow trees. Nat Genet 30, 97-101.

Achiron, A., Gurevich, M., Friedman, N., Kaminski, N., and Mandel, M. (2004). Blood transcriptional signatures of multiple sclerosis: unique gene expression of disease activity. Ann Neurol 55, 410-7.

Airas, L., Nikula, T., Huang, Y. H., Lahesmaa, R., and Wiendl, H. (2007). Postpartum-activation of multiple sclerosis is associated with down-regulation of tolerogenic HLA-G. J Neuroimmunol 187, 205-11.

Akesson, E., Oturai, A., Berg, J., Fredrikson, S., Andersen, O., Harbo, H. F., Laaksonen, M., Myhr, K. M., Nyland, H. I., Ryder, L. P., Sandberg-Wollheim, M., Sorensen, P. S., Spurkland, A., Svejgaard, A., Holmans, P., Compston, A., Hillert, J., and Sawcer, S. (2002). A genome-wide screen for linkage in Nordic sib-pairs with multiple sclerosis. Genes Immun 3, 279-85.

Aldred, P. M., Hollox, E. J., and Armour, J. A. (2005). Copy number polymorphism and expression level variation of the human α-defensin genes DEFA1 and DEFA3. Hum Mol Genetics 14, 2045–52.

Altmüller, J., Palmer, L. J., Fischer, G., Scherb, H., and Wjst, M. (2001). Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet 69, 936-50.

Altshuler, D., Daly, M. J., and Lander, E. S. (2008). Genetic mapping in human disease. Science 322, 881-8.

Amundsen, S. S., Monsuur, A. J., Wapenaar, M. C., Lie, B. A., Ek, J., Gudjonsdottir, A. H., Ascher, H., Wijmenga, C., and Sollid, L. M. (2006). Association analysis of MYO9B gene polymorphisms with celiac disease in a Swedish/Norwegian cohort. Hum Immunol 67, 341-5.

Amundsen, S. S., Vatn, M., Wijmenga, C., Sollid, L. M., and Lie, B. A. (2006). Association analysis of MYO9B gene polymorphisms and inflammatory bowel disease in a Norwegian cohort. Tissue Antigens 68, 249-52.

Ares, B., Prieto, J. M., Lema, M., Dapena, D., Arias, M., and Noya, M. (2007). Prevalence of multiple sclerosis in Santiago de Compostela (Galicia, Spain). Mult Scler 13, 262-4.

Arthur, A. T., Armati, P. J., Bye, C., Heard, R. N., Stewart, G. J., Pollard, J. D., and Booth, D. R. (2008). Genes implicated in multiple sclerosis pathogenesis from

consilience of genotyping and expression profiles in relapse and remission. BMC Med Genet 9, 17.

Ascherio, A., Munger, K. L., and Simon, K. C. (2010). Vitamin D and multiple sclerosis. Lancet Neurol 9, 599-612.

Aulchenko, Y. S., Hoppenbrouwers, I. A., Ramagopalan, S. V., Broer, L., Jafari, N., Hillert, J., Link, J., Lundstrom, W., Greiner, E., Dessa Sadovnick, A., Goossens, D., Van Broeckhoven, C., Del-Favero, J., Ebers, G. C., Oostra, B. A., van Duijn, C. M., and Hintzen, R. Q. (2008). Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. Nat Genet 40, 1402-3.

Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). (2009). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet 41, 824-8.

Avasarala, J. R., Chittur, S. V., George, A. D., and Tine, J. A. (2008). Microarray analysis in B cells among siblings with/without MS - role for transcription factor TCF2. BMC Med Genomics 1, 2.

Babbe, H., Roers, A., Waisman, A., Lassmann, H., Goebels, N., Hohlfeld, R., Friese, M., Schroder, R., Deckert, M., Schmidt, S., Ravid, R., and Rajewsky, K. (2000). Clonal expansions of CD8(+) T cells dominate the T cell infiltrate in active multiple sclerosis lesions as shown by micromanipulation and single cell polymerase chain reaction. J Exp Med 192, 393-404.

Bailey, S. L., Schreiner, B., McMahon, E. J., and Miller, S. D. (2007). CNS myeloid DCs presenting endogenous myelin peptides 'preferentially' polarize CD4+ T(H)-17 cells in relapsing EAE. Nat Immunol 8, 172-80.

Ban, M., Stewart, G. J., Bennetts, B. H., Heard, R., Simmons, R., Maranian, M., Compston, A., and Sawcer, S. J. (2002). A genome screen for linkage in Australian sibling-pairs with multiple sclerosis. Genes Immun 3, 464-9.

Baranzini, S. E., Wang, J., Gibson, R. A., Galwey, N., Naegelin, Y., Barkhof, F., Radue, E. W., Lindberg, R. L., Uitdehaag, B. M., Johnson, M. R., Angelakopoulou, A., Hall, L., Richardson, J. C., Prinjha, R. K., Gass, A., Geurts, J. J., Kragt, J., Sombekke, M., Vrenken, H., Qualley, P., Lincoln, R. R., Gomez, R., Caillier, S. J., George, M. F., Mousavi, H., Guerrero, R., Okuda, D. T., Cree, B. A., Green, A. J., Waubant, E., Goodin, D. S., Pelletier, D., Matthews, P. M., Hauser, S. L., Kappos, L., Polman, C. H., and Oksenberg, J. R. (2009). Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet 18, 767-78.

Baranzini, S. E., Mudge, J., van Velkinburgh, J. C., Khankhanian, P., Khrebtukova, I., Miller, N. A., Zhang, L., Farmer, A. D., Bell, C. J., Kim, R. W., May, G. D., Woodward, J. E., Caillier, S. J., McElroy, J. P., Gomez, R., Pando, M. J., Clendenen, L. E., Ganusova, E. E., Schilkey, F. D., Ramaraj, T., Khan, O. A., Huntley, J. J., Luo, S., Kwok, P. Y., Wu, T. D., Schroth, G. P., Oksenberg, J. R., Hauser, S. L., and Kingsmore, S. F. (2010). Genome, epigenome and RNA

sequences of monozygotic twins discordant for multiple sclerosis. Nature 464, 1351-6.

Bar-Or, A., Calabresi, P. A., Arnold, D., Markowitz, C., Shafer, S., Kasper, L. H., Waubant, E., Gazda, S., Fox, R. J., Panzara, M., Sarkar, N., Agarwal, S., and Smith, C. H. (2008). Rituximab in relapsing-remitting multiple sclerosis: a 72-week, open-label, phase I trial. Ann Neurol 63, 395-400.

Barrett, J. C., Fry, B., Maller, J., and Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263-5.

Barrett, J. C., Hansoul, S., Nicolae, D. L., Cho, J. H., Duerr, R. H., Rioux, J. D., Brant, S. R., Silverberg, M. S., Taylor, K. D., Barmada, M. M., Bitton, A., Dassopoulos, T., Datta, L. W., Green, T., Griffiths, A. M., Kistner, E. O., Murtha, M. T., Regueiro, M. D., Rotter, J. I., Schumm, L. P., Steinhart, A. H., Targan, S. R., Xavier, R. J., Libioulle, C., Sandor, C., Lathrop, M., Belaiche, J., Dewit, O., Gut, I., Heath, S., Laukens, D., Mni, M., Rutgeerts, P., Van Gossum, A., Zelenika, D., Franchimont, D., Hugot, J. P., de Vos, M., Vermeire, S., Louis, E., Cardon, L. R., Anderson, C. A., Drummond, H., Nimmo, E., Ahmad, T., Prescott, N. J., Onnie, C. M., Fisher, S. A., Marchini, J., Ghori, J., Bumpstead, S., Gwilliam, R., Tremelling, M., Deloukas, P., Mansfield, J., Jewell, D., Satsangi, J., Mathew, C. G., Parkes, M., Georges, M., and Daly, M. J. (2008). Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 40, 955-62.

Barrett, J. C., Clayton, D. G., Concannon, P., Akolkar, B., Cooper, J. D., Erlich, H. A., Julier, C., Morahan, G., Nerup, J., Nierras, C., Plagnol, V., Pociot, F., Schuilenburg, H., Smyth, D. J., Stevens, H., Todd, J. A., Walker, N. M., Rich, S. S., and Type 1 Diabetes Genetics Consortium. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. (2009). Nat Genet 41, 703-7.

Barton, A., Woolmore, J. A., Ward, D., Eyre, S., Hinks, A., Ollier, W. E., Strange, R. C., Fryer, A. A., John, S., Hawkins, C. P., and Worthington, J. (2004). Association of protein kinase C alpha (PRKCA) gene with multiple sclerosis in a UK population. Brain 127, 1717-22.

Baumhackl, U., Eibl, G., Ganzinger, U., Hartung, H. P., Mamoli, B., Pfeiffer, K. P., Fazekas, F., and Vass, K. (2002). Prevalence of multiple sclerosis in Austria. Results of a nationwide survey. Neuroepidemiology 21, 226-34.

Bencsik, K., Rajda, C., Klivenyi, P., Jardanhazy, T., and Vecsei, L. (1998). The prevalence of multiple sclerosis in the Hungarian city of Szeged. Acta Neurol Scand 97, 315-9.

Bencsik, K., Rajda, C., Fuvesi, J., Klivenyi, P., Jardanhazy, T., Torok, M., and Vecsei, L. (2001). The prevalence of multiple sclerosis, distribution of clinical forms of the disease and functional status of patients in Csongrad County, Hungary. Eur Neurol 46, 206-9.

Benedikz, J., Stefansson, M., Guomundsson, J., Jonasdottir, A., Fossdal, R., Gulcher, J., and Stefansson, K. (2002). The natural history of untreated multiple sclerosis in Iceland. A total population-based 50 year prospective study. Clin Neurol Neurosurg 104, 208-10.

Benito-Leon, J., Martin, E., Vela, L., Villar, M. E., Felgueroso, B., Marrero, C., Guerrero, A., and Ruiz-Galiana, J. (1998). Multiple sclerosis in Mostoles, central Spain. Acta Neurol Scand 98, 238-42.

Benjamini, B., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc B 57, 289-300.

Bennett, L., Palucka, A. K., Arce, E., Cantrell, V., Borvak, J., Banchereau, J., and Pascual, V. (2003). Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. J Exp Med 197, 711-23.

Bentzen, J., Flachs, E. M., Stenager, E., Bronnum-Hansen, H., and Koch-Henriksen, N. (2010). Prevalence of multiple sclerosis in Denmark 1950--2005. Mult Scler 16, 520-5.

Benvenuto, R., Paroli, M., Buttinelli, C., Franco, A., Barnaba, V., Fieschi, C., and Balsano, F. (1992). Tumor necrosis factor-alpha and interferon-gamma synthesis by cerebrospinal fluid-derived T cell clones in multiple sclerosis. Ann N Y Acad Sci 650, 341-6.

Bertrams, J., Kuwert, E., and Liedtke, U. (1972). HL-A antigens and multiple sclerosis. Tissue Antigens 2, 405-8.

Bettelli, E., Carrier, Y., Gao, W., Korn, T., Strom, T. B., Oukka, M., Weiner, H. L., Kuchroo, and V. K. (2006). Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. Nature 441, 235-8.

Binzer, M., Forsgren, L., Holmgren, G., Drugge, U., and Fredrikson, S. (1994). Familial clustering of multiple sclerosis in a northern Swedish rural district. J Neurol Neurosurg Psychiatry 57, 497-9.

Bomprezzi, R., Ringner, M., Kim, S., Bittner, M. L., Khan, J., Chen, Y., Elkahloun, A., Yu, A., Bielekova, B., Meltzer, P. S., Martin, R., McFarland, H. F., and Trent, J. M. (2003). Gene expression profile in multiple sclerosis patients and healthy controls: identifying pathways relevant to disease. Hum Mol Genet 12, 2191-9.

Booth, D. R., Arthur, A. T., Teutsch, S. M., Bye, C., Rubio, J., Armati, P. J., Pollard, J. D., Heard, R. N., and Stewart, G. J. (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.

Boström, I., Callander, M., Kurtzke, J. F., and Landtblom, A. M. (2009). High prevalence of multiple sclerosis in the Swedish county of Varmland. Mult Scler 15, 1253-62.

Broadley, S. A., Deans, J., Sawcer, S. J., Clayton, D., and Compston, D. A. (2000). Autoimmune disease in first-degree relatives of patients with multiple sclerosis. A UK survey. Brain 123 (Pt 6), 1102-11.

Broadley, S., Sawcer, S., D'Alfonso, S., Hensiek, A., Coraddu, F., Gray, J., Roxburgh, R., Clayton, D., Buttinelli, C., Quattrone, A., Trojano, M., Massacesi, L., and Compston, A. (2001). A genome screen for multiple sclerosis in Italian families. Genes Immun 2, 205-10.

Brown, D., Trowsdale, J., and Allen, R. (2004). The LILR family: modulators of innate and adaptive immune pathways in health and disease. Tissue Antigens 64, 215-25.

Brynedal, B., Khademi, M., Wallstrom, E., Hillert, J., Olsson, T., and Duvefelt, K. (2010). Gene expression profiling in multiple sclerosis: a disease of the central nervous system, but with relapses triggered in the periphery? Neurobiol Dis 37, 613-21.

Carswell, R. (1838). Pathological anatomy: illustrations of the elementary forms of disease. London: Longman, Orme, Brown, Green and Longman.

Casetta, I., Granieri, E., Marchi, D., Murgia, S. B., Tola, M. R., Ticca, A., Lauria, G., Govoni, V., Murgia, B., and Pugliatti, M. (1998). An epidemiological study of multiple sclerosis in central Sardinia, Italy. Acta Neurol Scand 98, 391-4.

Casquero, P., Villoslada, P., Montalban, X., and Torrent, M. (2001). Frequency of multiple sclerosis in Menorca, Balearic islands, Spain. Neuroepidemiology 20, 129-33.

Cavalletti, S., Merelli, E., Cavazzuti, M., and Guidetti, D. (1994). Frequency of MS in the province of Modena, 1970-1990. Acta Neurol Scand 90, 377-81.

Charcot, J.M. (1865). Sclérose du cordons lateraux del la moelle épiniere chez une femme hysterique atteinte de contracture permaneste des quatres membres. L'Union Med 25, 451-72.

Charcot, J.M. (1868a). Comptes rendus des seances et memoires lus a la societe de Biologie 20, 13-4.

Charcot, J.M. (1868b). Histologie de la sclérose en plaques. Gazette Hôpitaux 41, 554-66.

Charcot, J.M. (1868c). Leçons sur les maladies chroniques du système nerveux. I: Des scléroses de la moelle épinière. Gazette Hôpitaux 41, 405-9.

Chaudhuri, A., and Behan, P. O. (2008). Rituximab in relapsing-remitting multiple sclerosis. N Engl J Med 358, 2646; author reply 2646-7.

Chen, T., and Cao, X. (2010). Stress for maintaining memory: HSP70 as a mobile messenger for innate and adaptive immunity. Eur J Immunol. 2010 Jun;40(6):1541-4.

Cherwinski, H. M., Schumacher, J. H., Brown, K. D., and Mosmann, T. R. (1987). Two types of mouse helper T cell clone. III. Further differences in lymphokine synthesis between Th1 and Th2 clones revealed by RNA hybridization,

functionally monospecific bioassays, and monoclonal antibodies. J Exp Med 166, 1229-44.

Cho, M., Ishida, K., Chen, J., Ohkawa, J., Chen, W., Namiki, S., Kotaki, A., Arai, N., Arai, K., and Kamogawa-Schifter, Y. (2008). SAGE library screening reveals ILT7 as a specific plasmacytoid dendritic cell marker that regulates type I IFN production. Int Immunol 20, 155-64.

Cirillo, G., Di Domenico, M. R., Corsi, I., Gagliardo, T., Del Giudice, E. M., Perrone, L., and Tolone, C. (2007). Do MYO9B genetic variants predispose to coeliac disease? An association study in a cohort of South Italian children. Dig Liver Dis 39, 228-31.

CAMMS223 Trial Investigators, Coles, A. J., Compston, D. A., Selmaj, K. W., Lake, S. L., Moran, S., Margolin, D. H., Norris, K., and Tandon, P. K. (2008). Alemtuzumab vs. interferon beta-1a in early multiple sclerosis. N Engl J Med 359, 1786-801.

Cochran, W. (1954). The combination of estimates from different experiments. Biometrics 10, 101-29.

Compston, D. A., Morgan, B. P., Campbell, A. K., Wilkins, P., Cole, G., Thomas, N. D., and Jasani, B. (1989). Immunocytochemical localization of the terminal complement complex in multiple sclerosis. Neuropathol Appl Neurobiol 15, 307-16.

Compston, A., and Coles, A. (2002). Multiple sclerosis. Lancet 359, 1221-31.

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T. D., Barnes, C., Campbell, P., Fitzgerald, T., Hu, M., Ihm, C. H., Kristiansson, K., Macarthur, D. G., Macdonald, J. R., Onyiah, I., Pang, A. W., Robson, S., Stirrups, K., Valsesia, A., Walter, K., Wei, J., Tyler-Smith, C., Carter, N. P., Lee, C., Scherer, S. W., and Hurles, M. E. (2010). Origins and functional impact of copy number variation in the human genome. Nature 464, 704-12.

Cooney, R., Cummings, J. R., Pathan, S., Beckly, J., Geremia, A., Hancock, L., Guo, C., Morris, A., and Jewell, D. P. (2009). Association between genetic variants in myosin IXB and Crohn's disease. Inflamm Bowel Dis 15, 1014-21.

Cooper, J. D., Smyth, D. J., Smiles, A. M., Plagnol, V., Walker, N. M., Allen, J. E., Downes, K., Barrett, J. C., Healy, B. C., Mychaleckyj, J. C., Warram, J. H., and Todd, J. A. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. (2008). Nat Genet 40, 1399-401.

Coraddu, F., Sawcer, S., D'Alfonso, S., Lai, M., Hensiek, A., Solla, E., Broadley, S., Mancosu, C., Pugliatti, M., Marrosu, M. G., and Compston, A. (2001). A genome screen for multiple sclerosis in Sardinian multiplex families. Eur J Hum Genet 9, 621-6.

Corley, S. M., Ladiwala, U., Besson, A., and Yong, V. W. (2001). Astrocytes attenuate oligodendrocyte death in vitro through an alpha(6) integrin-laminin-dependent mechanism. Glia 36, 281-94.

Cua, D. J., Sherlock, J., Chen, Y., Murphy, C. A., Joyce, B., Seymour, B., Lucian, L., To, W., Kwan, S., Churakova, T., Zurawski, S., Wiekowski, M., Lira, S. A., Gorman, D., Kastelein, R. A., and Sedgwick, J. D. (2003). Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. Nature 6924, 744-8.

Cummings, C. J., Sun, Y., Opal, P., Antalffy, B., Mestril, R., Orr, H. T., Dillmann, W. H., and Zoghbi, H. Y. (2001). Over-expression of inducible HSP70 chaperone suppresses neuropathology and improves motor function in SCA1 mice. Hum Mol Genet 10, 1511-8.

Dahl, O. P., Aarseth, J. H., Myhr, K. M., Nyland, H., and Midgard, R. (2004). Multiple sclerosis in Nord-Trondelag County, Norway: a prevalence and incidence study. Acta Neurol Scand 109, 378-84.

De Jager, P. L., Jia, X., Wang, J., de Bakker, P. I., Ottoboni, L., Aggarwal, N. T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C., Briskin, R., Romano, S., Baranzini, S. E., McCauley, J. L., Pericak-Vance, M. A., Haines, J. L., Gibson, R. A., Naeglin, Y., Uitdehaag, B., Matthews, P. M., Kappos, L., Polman, C., McArdle, W. L., Strachan, D. P., Evans, D., Cross, A. H., Daly, M. J., Compston, A., Sawcer, S. J., Weiner, H. L., Hauser, S. L., Hafler, D. A., and Oksenberg, J. R. (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet 41, 776-82.

De Palma, R., and Sementa, A. (2008). Rituximab in relapsing-remitting multiple sclerosis. N Engl J Med 358, 2645-6; author reply 2646-7.

De Sá, J., Paulos, A., Mendes, H., Becho, J., Marques, J., and Roxo, J. (2006). The prevalence of multiple sclerosis in the District of Santarem, Portugal. J Neurol 253, 914-8.

Dean, G. (1994). How many people in the world have multiple sclerosis? Neuroepidemiology 13, 1-7.

Dean, G., Aksoy, H., Akalin, T., Middleton, L., and Kyriallis, K. (1997). Multiple sclerosis in the Turkish- and Greek-speaking communities of Cyprus. A United Nations (UNHCR) Bicommunal Project. J Neurol Sci 145, 163-8.

Dean, G., Elian, M., de Bono, A. G., Asciak, R. P., Vella, N., Mifsud, V., and Aquilina, J. (2002). Multiple sclerosis in Malta in 1999: an update. J Neurol Neurosurg Psychiatry 73, 256-60.

Debouverie, M., Pittion-Vouyovitch, S., Louis, S., Roederer, T., and Guillemin, F. (2007). Increasing incidence of multiple sclerosis among women in Lorraine, Eastern France. Mult Scler 13, 962-7.

Dimas, A. S., Deutsch, S., Stranger, B. E., Montgomery, S. B., Borel, C., Attar-Cohen, H., Ingle, C., Beazley, C., Gutierrez Arcelus, M., Sekowska, M., Gagnebin, M., Nisbett, J., Deloukas, P., Dermitzakis, E. T., and Antonarakis, S. E. (2009). Common regulatory variation impacts gene expression in a cell type-dependent manner. Science 325, 1246-50.

Studies on causes of
multiple sclerosis

Dixon, A. L., Liang, L., Moffatt, M. F., Chen, W., Heath, S., Wong, K. C., Taylor, J., Burnett, E., Gut, I., Farrall, M., Lathrop, G. M., Abecasis, G. R., and Cookson, W. O. (2007). A genome-wide association study of global gene expression. Nat Genet 39, 1202-7.

Doerre, S., Mesires, K. P., Daley, K. M., McCarty, T., Knoetig, S., and Corley, R. B. (2005). Reductions in I kappa B epsilon and changes in NF-kappa B activity during B lymphocyte differentiation. J Immunol 174, 983-91.

Dubois, P. C., Trynka, G., Franke, L., Hunt, K. A., Romanos, J., Curtotti, A., Zhernakova, A., Heap, G. A., Adany, R., Aromaa, A., Bardella, M. T., van den Berg, L. H., Bockett, N. A., de la Concha, E. G., Dema, B., Fehrmann, R. S., Fernandez-Arquero, M., Fiatal, S., Grandone, E., Green, P. M., Groen, H. J., Gwilliam, R., Houwen, R. H., Hunt, S. E., Kaukinen, K., Kelleher, D., Korponay-Szabo, I., Kurppa, K., MacMathuna, P., Maki, M., Mazzilli, M. C., McCann, O. T., Mearin, M. L., Mein, C. A., Mirza, M. M., Mistry, V., Mora, B., Morley, K. I., Mulder, C. J., Murray, J. A., Nunez, C., Oosterom, E., Ophoff, R. A., Polanco, I., Peltonen, L., Platteel, M., Rybak, A., Salomaa, V., Schweizer, J. J., Sperandeo, M. P., Tack, G. J., Turner, G., Veldink, J. H., Verbeek, W. H., Weersma, R. K., Wolters, V. M., Urcelay, E., Cukrowska, B., Greco, L., Neuhausen, S. L., McManus, R., Barisani, D., Deloukas, P., Barrett, J. C., Saavalainen, P., Wijmenga, C., and van Heel, D. A. (2010). Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 42, 295-302.

Dyment, D. A., Ebers, G. C., and Sadovnick, A. D. (2004). Genetics of multiple sclerosis. Lancet Neurol 3, 104-10.

Dyment, D. A., Sadovnick, A. D., Willer, C. J., Armstrong, H., Cader, Z. M., Wiltshire, S., Kalman, B., Risch, N., and Ebers, G. C. (2004). An extended genome scan in 442 Canadian multiple sclerosis-affected sibships: a report from the Canadian Collaborative Study Group. Hum Mol Genet 13, 1005-15.

Ebers, G. C., Sadovnick, A. D., and Risch, N. J. (1995). A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. Nature 377, 150-1.

Ebers, G. C., Kukay, K., Bulman, D. E., Sadovnick, A. D., Rice, G., Anderson, C., Armstrong, H., Cousin, K., Bell, R. B., Hader, W., Paty, D. W., Hashimoto, S., Oger, J., Duquette, P., Warren, S., Gray, T., O'Connor, P., Nath, A., Auty, A., Metz, L., Francis, G., Paulseth, J. E., Murray, T. J., Pryse-Phillips, W., Nelson, R., Freedman, M., Brunet, D., Bouchard, J. P., Hinds, D., and Risch, N. (1996). A full genome search in multiple sclerosis. Nat Genet 13, 472-6.

Ebers, G. C. (2008). Environmental factors and multiple sclerosis. Lancet Neurol 7, 268-77.

Ellinghaus, E., Ellinghaus, D., Stuart, P. E., Nair, R. P., Debrus, S., Raelson, J. V., Belouchi, M., Fournier, H., Reinhard, C., Ding, J., Li, Y., Tejasvi, T., Gudjonsson, J., Stoll, S. W., Voorhees, J. J., Lambert, S., Weidinger, S., Eberlein,

B., Kunz, M., Rahman, P., Gladman, D. D., Gieger, C., Wichmann, H. E., Karlsen, T. H., Mayr, G., Albrecht, M., Kabelitz, D., Mrowietz, U., Abecasis, G. R., Elder, J. T., Schreiber, S., Weichenthal, M., and Franke A. (2010). Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2. Nat Genet 42, 991-5.

Eraksoy, M., Kurtuncu, M., Akman-Demir, G., Kilinc, M., Gedizlioglu, M., Mirza, M., Anlar, O., Kutlu, C., Demirkiran, M., Idrisoglu, H. A., Compston, A., and Sawcer, S. (2003). A whole genome screen for linkage in Turkish multiple sclerosis. J Neuroimmunol 143, 17-24.

Falconer, D. S., and MacKay, T. F. C. (1996). Introduction to Quantitative Genetics, 4th Ed. Longmans Green, Harlow, Essex, UK.

Fatima, N., Toscano, M. P., Hunter, S. B., and Cohen, C. (2011). Controversial role of Epstein-Barr virus in multiple sclerosis. Appl Immunohistochem Mol Morphol 19, 246-52.

Ferber, I. A., Brocke, S., Taylor-Edwards, C., Ridgway, W., Dinisco, C., Steinman, L., Dalton, D., and Fathman, C. G. (1996). Mice with a disrupted IFN-gamma gene are susceptible to the induction of experimental autoimmune encephalomyelitis (EAE). J Immunol 156, 5-7.

Fernández, O., Luque, G., San Román, C., Bravo, M., and Dean, G. (1994). The prevalence of multiple sclerosis in the Sanitary District of Velez-Malaga, southern Spain. Neurology 44, 425-9.

Fiorentino, D. F., Bond, M. W., and Mosmann, T. R. (1989). Two types of mouse T helper cell. IV. Th2 clones secrete a factor that inhibits cytokine production by Th1 clones. J Exp Med 170, 2081-95.

Fontana, A., Fierz, W., and Wekerle, H. (1984). Astrocytes present myelin basic protein to encephalitogenic T-cell lines. Nature 307, 273-6.

Forbes, R. B., Wilson, S. V., and Swingler, R. J. (1999). The prevalence of multiple sclerosis in Tayside, Scotland: do latitudinal gradients really exist? J Neurol 246, 1033-40.

Ford, H. L., Gerry, E., Airey, C. M., Vail, A., Johnson, M. H., and Williams, D. R. (1998). The prevalence of multiple sclerosis in the Leeds Health Authority. J Neurol Neurosurg Psychiatry 64, 605-10.

Fox, C. M., Bensa, S., Bray, I., and Zajicek, J. P. (2004). The epidemiology of multiple sclerosis in Devon: a comparison of the new and old classification criteria. J Neurol Neurosurg Psychiatry 75, 56-60.

Freim Wahl, S. G., Folvik, M. R., and Torp, S. H. (2007). Progressive multifocal leukoencephalopathy in a lymphoma patient with complete remission after treatment with cytostatics and rituximab: case report and review of the literature. Clin Neuropathol 26, 68-73.

GAMES; Transatlantic Multiple Sclerosis Genetics Cooperative. A meta-analysis of whole genome linkage screens in multiple sclerosis. (2003). J Neuroimmunol 143, 39-46.

Gandhi, K. S., McKay, F. C., Cox, M., Riveros, C., Armstrong, N., Heard, R. N., Vucic, S., Williams, D. W., Stankovich, J., Brown, M., Danoy, P., Stewart, G. J., Broadley, S., Moscato, P., Lechner-Scott, J., Scott, R. J., and Booth, D. R. (2010). The multiple sclerosis whole blood mRNA transcriptome and genetic associations indicate dysregulation of specific T cell pathways in pathogenesis. Hum Mol Genet 19, 2134-43.

Ganz, T., Selsted, M. E., Szklarek, D., Harwig, S. S., Daher, K., Bainton, D. F., and Lehrer, R. I. (1985). Defensins. Natural peptide antibiotics of human neutrophils. J Clin Invest 76, 1427-35.

Giordano, M., Marano, C., Mellai, M., Limongelli, M. G., Bolognesi, E., Clerget-Darpoux, F., Momigliano-Richiardi, P., and Greco, L. (2006). A family-based study does not confirm the association of MYO9B with celiac disease in the Italian population. Genes Immun 7, 606-8.

Gold, R., Linington, C., and Lassmann, H. (2006). Understanding pathogenesis and therapy of multiple sclerosis via animal models: 70 years of merits and culprits in experimental autoimmune encephalomyelitis research. Brain 129, 1953-71.

Granieri, E., Economou, N. T., De Gennaro, R., Tola, M. R., Caniatti, L., Govoni, V., Fainardi, E., and Casetta, I. (2007). Multiple sclerosis in the province of Ferrara : evidence for an increasing trend. J Neurol 254, 1642-8.

Granieri, E., Monaldini, C., De Gennaro, R., Guttmann, S., Volpini, M., Stumpo, M., Fazio, P., and Casetta, I. (2008). Multiple sclerosis in the Republic of San Marino: a prevalence and incidence study. Mult Scler 14, 325-9.

Grant, R. M., Carver, A. D., and Sloan, R. L. (1998). Multiple sclerosis in Fife. Scott Med J 43, 44-7.

Gray, O. M., McDonnell, G. V., and Hawkins, S. A. (2008). Factors in the rising prevalence of multiple sclerosis in the north-east of Ireland. Mult Scler 14, 880-6.

Greter, M., Heppner, F. L., Lemos, M. P., Odermatt, B. M., Goebels, N., Laufer, T., Noelle, R. J., and Becher, B. (2005). Dendritic cells permit immune invasion of the CNS in an animal model of multiple sclerosis. Nat Med 11, 328-34.

Grimaldi, L. M., Salemi, G., Grimaldi, G., Rizzo, A., Marziolo, R., Lo Presti, C., Maimone, D., and Savettieri, G. (2001). High incidence and increasing prevalence of MS in Enna (Sicily), southern Italy. Neurology 57, 1891-3.

Grimaldi, L. M., Palmeri, B., Salemi, G., Giglia, G., D'Amelio, M., Grimaldi, R., Vitello, G., Ragonese, P., and Savettieri, G. (2007). High prevalence and fast rising incidence of multiple sclerosis in Caltanissetta, Sicily, southern Italy. Neuroepidemiology 28, 28-32.

Grønlie, S. A., Myrvoll, E., Hansen, G., Grønning, M., and Mellgren, S. I. (2000). Multiple sclerosis in North Norway, and first appearance in an indigenous population. J Neurol 247, 129-33.

Gross-Paju, K., Oopik, M., Luus, S., Kalbe, I., and Kaasik, A. E. (1999). The risk of motor neurone disease and multiple sclerosis is different in Estonians and Russians. Data from South Estonia. Eur J Neurol 6, 187-93.

Grytten, N., Glad, S. B., Aarseth, J. H., Nyland, H., Midgard, R., and Myhr, K. M. (2006). A 50-year follow-up of the incidence of multiple sclerosis in Hordaland County, Norway. Neurology 66, 182-6.

Haines, J. L., Ter-Minassian, M., Bazyk, A., Gusella, J. F., Kim, D. J., Terwedow, H., Pericak-Vance, M. A., Rimmler, J. B., Haynes, C. S., Roses, A. D., Lee, A., Shaner, B., Menold, M., Seboun, E., Fitoussi, R. P., Gartioux, C., Reyes, C., Ribierre, F., Gyapay, G., Weissenbach, J., Hauser, S. L., Goodkin, D. E., Lincoln, R., Usuku, K., Oksenberg, J. R., and *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.

Hakonarson, H., Grant, S. F., Bradfield, J. P., Marchand, L., Kim, C. E., Glessner, J. T., Grabs, R., Casalunovo, T., Taback, S. P., Frackelton, E. C., Lawson, M. L., Robinson, L. J., Skraban, R., Lu, Y., Chiavacci, R. M., Stanley, C. A., Kirsch, S. E., Rappaport, E. F., Orange, J. S., Monos, D. S., Devoto, M., Qu, H. Q., and Polychronakos, C. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. (2007). Nature 448, 591-4.

Han, J. W., Zheng, H. F., Cui, Y., Sun, L. D., Ye, D. Q., Hu, Z., Xu, J. H., Cai, Z. M., Huang, W., Zhao, G. P., Xie, H. F., Fang, H., Lu, Q. J., Xu, J. H., Li, X. P., Pan, Y. F., Deng, D. Q., Zeng, F. Q., Ye, Z. Z., Zhang, X. Y., Wang, Q. W., Hao, F., Ma, L., Zuo, X. B., Zhou, F. S., Du, W. H., Cheng, Y. L., Yang, J. Q., Shen, S. K., Li, J., Sheng, Y. J., Zuo, X. X., Zhu, W. F., Gao, F., Zhang, P. L., Guo, Q., Li, B., Gao, M., Xiao, F. L., Quan, C., Zhang, C., Zhang, Z., Zhu, K. J., Li, Y., Hu, D. Y., Lu, W. S., Huang, J. L., Liu, S. X., Li, H., Ren, Y. Q., Wang, Z. X., Yang, C. J., Wang, P. G., Zhou, W. M., Lv, Y. M., Zhang, A. P., Zhang, S. Q., Lin, D., Li, Y., Low, H. Q., Shen, M., Zhai, Z. F., Wang, Y., Zhang, F. Y., Yang, S., Liu, J. J., and Zhang, X. J. (2009). Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet 41, 1234-7.

Harbo, H. F., Lie, B. A., Sawcer, S., Celius, E. G., Dai, K. Z., Oturai, A., Hillert, J., Lorentzen, A. R., Laaksonen, M., Myhr, K. M., Ryder, L. P., Fredrikson, S., Nyland, H., Sørensen, P. S., Sandberg-Wollheim, M., Andersen, O., Svejgaard, A., Edland, A., Mellgren, S. I., Compston, A., Vartdal, F., and Spurkland A. (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.

Harrington, L. E., Hatton, R. D., Mangan, P. R., Turner, H., Murphy, T. L., Murphy, K. M., and Weaver, C. T. (2005). Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. Nat Immunol 6, 1123-32.

Harrington, L. E., Mangan, P. R., and Weaver, C. T. (2006). Expanding the effector CD4 T-cell repertoire: the Th17 lineage. Curr Opin Immunol 18, 349-56.

Hauser, S. L., Waubant, E., Arnold, D. L., Vollmer, T., Antel, J., Fox, R. J., Bar-Or, A., Panzara, M., Sarkar, N., Agarwal, S., Langer-Gould, A., and Smith, C. H. (2008). B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. N Engl J Med 358, 676-88.

Hawkes, C. H. (2007). Smoking is a risk factor for multiple sclerosis: a metanalysis. Mult Scler 13, 610-5.

Hedström, A. K., Sundqvist, E., Bäärnhielm, M., Nordin, N., Hillert, J., Kockum, I., Olsson, T., and Alfredsson, L. (2011). Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. Brain 134, 653-64.

Hein, T., and Hopfenmuller, W. (2000). [Projection of the number of multiple sclerosis patients in Germany]. Nervenarzt 71, 288-94.

Heinig, M., Petretto, E., Wallace, C., Bottolo, L., Rotival, M., Lu, H., Li, Y., Sarwar, R., Langley, S. R., Bauerfeind, A., Hummel, O., Lee, Y. A., Paskas, S., Rintisch, C., Saar, K., Cooper, J., Buchan, R., Gray, E. E., Cyster, J. G., Cardiogenics Consortium, Erdmann, J., Hengstenberg, C., Maouche, S., Ouwehand, W. H., Rice, C. M., Samani, N. J., Schunkert, H., Goodall, A. H., Schulz, H., Roider, H. G., Vingron, M., Blankenberg, S., Münzel, T., Zeller, T., Szymczak, S., Ziegler, A., Tiret, L., Smyth, D. J., Pravenec, M., Aitman, T. J., Cambien, F., Clayton, D., Todd, J. A., Hubner, N., and Cook, S. A. (2010). A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. Nature 467, 460-4.

Hensiek, A. E., Roxburgh, R., Smilie, B., Coraddu, F., Akesson, E., Holmans, P., Sawcer, S. J., and Compston, D. A. (2003). Updated results of the United Kingdom linkage-based genome screen in multiple sclerosis. J Neuroimmunol 143, 25-30.

Hickey, W. F., Hsu, B. L., and Kimura, H. (1991). T-lymphocyte entry into the central nervous system. J Neurosci Res 28, 254-60.

Hirschhorn, J. N., Lohmueller, K., Byrne, E., and Hirschhorn, K. (2002). A comprehensive review of genetic association studies. Genet Med 4, 45-61.

Holmøy, T. (2006). A Norse contribution to the history of neurological diseases. Eur Neurol 55, 57-8.

Hüffmeier, U., Uebe, S., Ekici, A. B., Bowes, J., Giardina, E., Korendowych, E., Juneblad, K., Apel, M., McManus, R., Ho, P., Bruce, I. N., Ryan, A. W., Behrens, F., Lascorz, J., Böhm, B., Traupe, H., Lohmann, J., Gieger, C., Wichmann, H. E., Herold, C., Steffens, M., Klareskog, L., Wienker, T. F., Fitzgerald, O., Alenius, G. M., McHugh, N. J., Novelli, G., Burkhardt, H., Barton, A., and Reis, A. (2010). Common variants at TRAF3IP2 are associated with susceptibility to psoriatic arthritis and psoriasis. Nat Genet 42, 996-9.

Hunt, K. A., Monsuur, A. J., McArdle, W. L., Kumar, P. J., Travis, S. P., Walters, J. R., Jewell, D. P., Strachan, D. P., Playford, R. J., Wijmenga, C., and van Heel, D. A. (2006). Lack of association of MYO9B genetic variants with coeliac disease in a British cohort. Gut 55, 969-72.

Hunt, K. A., Zhernakova, A., Turner, G., Heap, G. A., Franke, L., Bruinenberg, M., Romanos, J., Dinesen, L. C., Ryan, A. W., Panesar, D., Gwilliam, R., Takeuchi, F., McLaren, W. M., Holmes, G. K., Howdle, P. D., Walters, J. R., Sanders, D. S., Playford, R. J., Trynka, G., Mulder, C. J., Mearin, M. L., Verbeek, W. H., Trimble, V., Stevens, F. M., O'Morain, C., Kennedy, N. P., Kelleher, D., Pennington, D. J., Strachan, D. P., McArdle, W. L., Mein, C. A., Wapenaar, M. C., Deloukas, P., McGinnis, R., McManus, R., Wijmenga, C., and van Heel, D. A. (2008). Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40, 395-402.

Huseby, E. S., Liggitt, D., Brabb, T., Schnabel, B., Ohlen, C., and Goverman, J. (2001). A pathogenic role for myelin-specific CD8(+) T cells in a model for multiple sclerosis. J Exp Med 194, 669-76.

Hussien, Y., Sanna, A., Soderstrom, M., Link, H., and Huang, Y. M. (2001). Glatiramer acetate and IFN-beta act on dendritic cells in multiple sclerosis. J Neuroimmunol 121, 102-10.

Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., Scherer, S. W., and Lee, C. (2004). Detection of large-scale variation in the human genome. Nat Genet 36, 949-51.

Iglesias, A. H., Camelo, S., Hwang, D., Villanueva, R., Stephanopoulos, G., and Dangond, F. (2004). Microarray detection of E2F pathway activation and other targets in multiple sclerosis peripheral blood mononuclear cells. J Neuroimmunol 150, 163-77.

Ingram, G., Hakobyan, S., Robertson, N. P., and Morgan, B. P. (2009). Complement in multiple sclerosis: its role in disease and potential as a biomarker. Clin Exp Immunol 155, 128-39.

Inouye, M., Silander, K., Hämäläinen, E., Salomaa, V., Harald, K., Jousilahti, P., Männistö, S., Eriksson, J. G., Saarela, J., Ripatti, S., Perola, M., van Ommen, G. J., Taskinen, M. R., Palotie, A., Dermitzakis, E. T., and Peltonen, L. (2010). An immune response network associated with blood lipid levels. PLoS Genet 6, e1001113.

International HapMap Consortium. (2003). The International HapMap Project. Nature 426, 789-96.

International HapMap Consortium. (2005). A haplotype map of the human genome. Nature 437, 1299-320.

International Multiple Sclerosis Genetics Consortium (IMSGC), Hafler, D. A., Compston, A., Sawcer, S., Lander, E. S., Daly, M. J., De Jager, P. L., de Bakker, P. I., Gabriel, S. B., Mirel, D. B., Ivinson, A. J., Pericak-Vance, M. A., Gregory, S. G., Rioux, J. D., McCauley, J. L., Haines, J. L., Barcellos, L. F., Cree, B., Oksenberg, J. R., and Hauser, S. L. (2007). Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 357, 851-62.

International Multiple Sclerosis Genetics Consortium (IMSGC). (2008). Refining genetic associations in multiple sclerosis. Lancet Neurol 7, 567-9.

International Multiple Sclerosis Genetics Consortium (IMSGC). (2009). The expanding genetic overlap between multiple sclerosis and type I diabetes. Genes Immun 10, 11-4.

International Multiple Sclerosis Genetics Consortium (IMSGC). (2010a). IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci. Genes Immun 11, 397-405.

International Multiple Sclerosis Genetics Consortium (IMSGC). (2010b). Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. Hum Mol Genet 19, 953-62.

International Multiple Sclerosis Genetics Consortium (IMSGC), Booth, D. R., Heard, R. N., Stewart, G. J., Cox, M., Scott, R. J., Lechner-Scott, J., Goris, A., Dobosi, R., Dubois, B., Saarela, J., Leppä, V., Peltonen, L., Pirttila, T., Cournu-Rebeix, I., Fontaine, B., Bergamaschi, L., D'Alfonso, S., Leone, M., Lorentzen, A. R., Harbo, H. F., Celius, E. G., Spurkland, A., Link, J., Kockum, I., Olsson, T., Hillert, J., Ban, M., Baker, A., Kemppinen, A., Sawcer, S., Compston, A., Robertson, N. P., De Jager, P. L., Hafler, D. A., Barcellos, L. F., Ivinson, A. J., McCauley, J. L., Pericak-Vance, M. A., Oksenberg, J. R., Hauser, S. L., Sexton, D., and Haines, J. (2010). Lack of support for association between the KIF1B rs10492972[C] variant and multiple sclerosis. Nat Genet 42, 469-70.

The International Multiple Sclerosis Genetics Consortium (IMSGC) and The Wellcome Trust Case Control Consortium 2 (WTCCC2). (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476, 214-9.

Ishii, T., Onda, H., Tanigawa, A., Ohshima, S., Fujiwara, H., Mima, T., Katada, Y., Deguchi, H., Suemura, M., Miyake, T., Miyatake, K., Kawase, I., Zhao, H., Tomiyama, Y., Saeki, Y., and Nojima, H. (2005). Isolation and expression profiling of genes upregulated in the peripheral blood cells of systemic lupus erythematosus patients. DNA Res 12, 429-39.

Iuliano, G., and Napoletano, R. (2008). Prevalence and incidence of multiple sclerosis in Salerno (southern Italy) and its province. Eur J Neurol 15, 73-6.

Jagodic, M., Kornek, B., Weissert, R., Lassmann, H., Olsson, T., and Dahlman, I. (2001). Congenic mapping confirms a locus on rat chromosome 10 conferring strong protection against myelin oligodendrocyte glycoprotein-induced experimental autoimmune encephalomyelitis. Immunogenetics 53, 410-5.

Jakkula, E., Leppa, V., Sulonen, A. M., Varilo, T., Kallio, S., Kemppinen, A., Purcell, S., Koivisto, K., Tienari, P., Sumelahti, M. L., Elovaara, I., Pirttila, T., Reunanen, M., Aromaa, A., Oturai, A. B., Sondergaard, H. B., Harbo, H. F., Mero, I. L., Gabriel, S. B., Mirel, D. B., Hauser, S. L., Kappos, L., Polman, C., De Jager, P. L., Hafler, D. A., Daly, M. J., Palotie, A., Saarela, J., and Peltonen, L. (2010). Genome-wide association study in a high-risk isolate for multiple

sclerosis reveals associated variants in STAT3 gene. Am J Hum Genet 86, 285-91.

Jersild, C., Svejgaard, A., and Fog, T. (1972). HL-A antigens and multiple sclerosis. Lancet 1, 1240-1.

Jin, Y , Birlea, S. A., Fain, P. R., Gowan, K., Riccardi, S. L., Holland, P. J., Mailloux, C. M., Sufit, A. J., Hutton, S. M., Amadi-Myers, A., Bennett, D. C., Wallace, M. R., McCormack, W. T., Kemp, E. H., Gawkrodger, D. J., Weetman, A. P., Picardo, M., Leone, G., Taïeb, A., Jouary, T., Ezzedine, K., van Geel, N., Lambert, J., Overbeck, A., and Spritz, R. A. (2010). Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. N Engl J Med 362, 1686-97.

Johnson, A. D., Handsaker, R. E., Pulit, S., Nizzari, M. M., O'Donnell, C. J., and de Bakker, P. I. W. (2008). SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 24, 2938-9.

Ju, X. S., Hacker, C., Scherer, B., Redecke, V., Berger, T., Schuler, G., Wagner, H., Lipford, G. B., and Zenke, M. (2004). Immunoglobulin-like transcripts ILT2, ILT3 and ILT7 are expressed by human dendritic cells and down-regulated following activation. Gene 331, 159-64.

Jungerius, B. J., Bakker, S. C., Monsuur, A. J., Sinke, R. J., Kahn, R. S., and Wijmenga, C. (2008). Is MYO9B the missing link between schizophrenia and celiac disease? Am J Med Genet B Neuropsychiatr Genet 147, 351-5.

Kallio, S. P., Jakkula, E., Purcell, S., Suvela, M., Koivisto, K., Tienari, P. J., Elovaara, I., Pirttila, T., Reunanen, M., Bronnikov, D., Viander, M., Meri, S., Hillert, J., Lundmark, F., Harbo, H. F., Lorentzen, A. R., De Jager, P. L., Daly, M. J., Hafler, D. A., Palotie, A., Peltonen, L., and Saarela, J. (2009). Use of a genetic isolate to identify rare disease variants: C7 on 5p associated with MS. Hum Mol Genet 18, 1670-83.

Karman, J., Ling, C., Sandor, M., and Fabry, Z. (2004). Initiation of immune responses in brain is promoted by local dendritic cells. J Immunol 173, 2353-61.

Kenealy, S. J., Babron, M. C., Bradford, Y., Schnetz-Boutaud, N., Haines, J. L., Rimmler, J. B., Schmidt, S., Pericak-Vance, M. A., Barcellos, L. F., Lincoln, R. R., Oksenberg, J. R., Hauser, S. L., Clanet, M., Brassat, D., Edan, G., Yaouanq, J., Semana, G., Cournu-Rebeix, I., Lyon-Caen, O., and Fontaine, B. (2004). A second-generation genomic screen for multiple sclerosis. Am J Hum Genet 75, 1070-8.

Koch-Henriksen, N., and Sørensen, P. S. (2010). The changing demographic pattern of multiple sclerosis epidemiology. Lancet Neurol 9, 520-32.

Kohler, R. E., Comerford, I., Townley, S., Haylock-Jacobs, S., Clark-Lewis, I., and McColl, S. R. (2008). Antagonism of the chemokine receptors CXCR3 and CXCR4 reduces the pathology of experimental autoimmune encephalomyelitis. Brain Pathol 18, 504-16.

Kong, A., and Cox, N. J. (1997). Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61, 1179-88.

Koskinen, L. L., Korponay-Szabo, I. R., Viiri, K., Juuti-Uusitalo, K., Kaukinen, K., Lindfors, K., Mustalahti, K., Kurppa, K., Adany, R., Pocsai, Z., Szeles, G., Einarsdottir, E., Wijmenga, C., Maki, M., Partanen, J., Kere, J., and Saavalainen, P. (2008). Myosin IXB gene region and gluten intolerance: linkage to coeliac disease and a putative dermatitis herpetiformis association. J Med Genet 45, 222-7.

Kuokkanen, S., Gschwend, M., Rioux, J. D., Daly, M. J., Terwilliger, J. D., Tienari, P. J., Wikstrom, J., Palo, J., Stein, L. D., Hudson, T. J., Lander, E. S., and Peltonen, L. (1997). Genomewide scan of multiple sclerosis in Finnish multiplex families. Am J Hum Genet 61, 1379-87.

Kuokkanen, S., Sundvall, M., Terwilliger, J. D., Tienari, P. J., Wikstrom, J., Holmdahl, R., Pettersson, U., and Peltonen, L. (1996). A putative vulnerability locus to multiple sclerosis maps to 5p14-p12 in a region syntenic to the murine locus Eae2. Nat Genet 13, 477-80.

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 33, 1444-52.

Kurtzke, J. F., and Heltberg, A. (2001). Multiple sclerosis in the Faroe Islands: an epitome. J Clin Epidemiol 54, 1-22.

Lafaille, J. J., Keere, F. V., Hsu, A. L., Baron, J. L., Haas, W., Raine, C. S., and Tonegawa, S. (1997). Myelin basic protein-specific T helper 2 (Th2) cells cause experimental autoimmune encephalomyelitis in immunodeficient hosts rather than protect them from the disease. J Exp Med 186, 307-12.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., *et al.*

(2001). Initial sequencing and analysis of the human genome. Nature 409, 860-921.

Lang, H. L., Jacobsen, H., Ikemizu, S., Andersson, C., Harlos, K., Madsen, L., Hjorth, P., Sondergaard, L., Svejgaard, A., Wucherpfennig, K., Stuart, D. I., Bell, J. I., Jones, E. Y., and Fugger, L. (2002). A functional and structural basis for TCR cross-reactivity in multiple sclerosis. Nat Immunol 3, 940-3.

Lange, K., Cantor, R., Horvath, S., Perola, M., Sabatti, C., Sinsheimer, J., and Sobel, E. (2001). Mendel version 4.0: A complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. Amer J Hum Genetics 69(supplement), 504.

Latiano, A., Mora, B., Bonamico, M., Megiorni, F., Mazzilli, M. C., Cucchiara, S., Palmieri, O., Valvano, M. R., and Annese, V. (2007). Analysis of candidate genes on chromosomes 5q and 19p in celiac disease. J Pediatr Gastroenterol Nutr 45, 180-6.

Latiano, A., Palmieri, O., Valvano, M. R., D'Inca, R., Caprilli, R., Cucchiara, S., Sturniolo, G. C., Bossa, F., Andriulli, A., and Annese, V. (2008). The association of MYO9B gene in Italian patients with inflammatory bowel diseases. Aliment Pharmacol Ther 27, 241-8.

Lauer, K. (1994). The risk of multiple sclerosis in the U.S.A. in relation to sociogeographic features: a factor-analytic study. J Clin Epidemiol 47, 43-8.

Libbey, J. E., McCoy, L. L., and Fujinami, R. S. (2007). Molecular mimicry in multiple sclerosis. Int Rev Neurobiol 79, 127-47.

Linzmeier, R. M., and Ganz, T. (2005). Human defensin gene copy number polymorphisms: Comprehensive analysis of independent variation in α- and β-defensin regions at 8p22–p23. Genomics 86, 423–30.

Liu, K. K., and Dorovini-Zis, K. (2009). Regulation of CXCL12 and CXCR4 expression by human brain endothelial cells and their role in CD4+ and CD8+ T cell adhesion and transendothelial migration. J Neuroimmunol 215, 49-64.

Liu, X., Lee, Y. S., Yu, C. R., and Egwuagu, C. E. (2008). Loss of STAT3 in CD4+ T cells prevents development of experimental autoimmune diseases. J Immunol 180, 6070-6.

Ma, Q., Jones, D., Borghesani, P. R., Segal, R. A., Nagasawa, T., Kishimoto, T., Bronson, R. T., and Springer, T. A. (1998). Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice. Proc Natl Acad Sci U S A 95, 9448-53.

Mandel, M., Gurevich, M., Pauzner, R., Kaminski, N., and Achiron, A. (2004). Autoimmunity gene expression portrait: specific signature that intersects or differentiates between multiple sclerosis and systemic lupus erythematosus. Clin Exp Immunol 138, 164-70.

Mantel, N. (1963). Chi-square tests with 1 degree of freedom – extensions of the Mantel-Haenszel procedure. J Am Stat Assoc 58, 690–700.

Studies on causes of
multiple sclerosis

Matusevicius, D., Kivisakk, P., He, B., Kostulas, N., Ozenci, V., Fredrikson, S., and Link, H. (1999). Interleukin-17 mRNA expression in blood and CSF mononuclear cells is augmented in multiple sclerosis. Mult Scler 5, 101-4.

Mayne, M., Moffatt, T., Kong, H., McLaren, P. J., Fowke, K. R., Becker, K. G., Namaka, M., Schenck, A., Bardoni, B., Bernstein, C. N., and Melanson, M. (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.

McCarthy, M. I., and Hirschhorn, J. N. (2008). Genome-wide association studies: potential next steps on a genetic journey. Hum Mol Genet 17, R156-65.

McDonald, W. I., Compston, A., Edan, G., Goodkin, D., Hartung, H. P., Lublin, F. D., McFarland, H. F., Paty, D. W., Polman, C. H., Reingold, S. C., Sandberg-Wollheim, M., Sibley, W., Thompson, A., van den Noort, S., Weinshenker, B. Y., and Wolinsky, J. S. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol 50, 121-7.

McDonnell, G. V., and Hawkins, S. A. (1998). An epidemiologic study of multiple sclerosis in Northern Ireland. Neurology 50, 423-8.

McGuigan, C., McCarthy, A., Quigley, C., Bannan, L., Hawkins, S. A., and Hutchinson, M. (2004). Latitudinal variation in the prevalence of multiple sclerosis in Ireland, an effect of genetic diversity. J Neurol Neurosurg Psychiatry 75, 572-6.

McKay, F. C., Swain, L. I., Schibeci, S. D., Rubio, J. P., Kilpatrick, T. J., Heard, R. N., Stewart, G. J., and Booth, D. R. (2008). Haplotypes of the interleukin 7 receptor alpha gene are correlated with altered expression in whole blood cells in multiple sclerosis. Genes Immun 9, 1-6.

McLaughlin, K. A., and Wucherpfennig, K. W. (2008). B cells and autoantibodies in the pathogenesis of multiple sclerosis and related inflammatory demyelinating diseases. Adv Immunol 98, 121-49.

McMahon, E. J., Bailey, S. L., Castenada, C. V., Waldner, H., and Miller, S. D. (2005). Epitope spreading initiates in the CNS in two mouse models of multiple sclerosis. Nat Med 11, 335-9.

Mero, I. L., Lorentzen, A. R., Ban, M., Smestad, C., Celius, E. G., Aarseth, J. H., Myhr, K. M., Link, J., Hillert, J., Olsson, T., Kockum, I., Masterman, T., Oturai, A. B., Sondergaard, H. B., Sellebjerg, F., Saarela, J., Kemppinen, A., Elovaara, I., Spurkland, A., Dudbridge, F., Lie, B. A., and Harbo, H. F. (2010). A rare variant of the TYK2 gene is confirmed to be associated with multiple sclerosis. Eur J Hum Genet 18, 502-4.

Milanov, I., Topalov, N., and Kmetski, T. (1999). Prevalence of multiple sclerosis in Gypsies and Bulgarians. Neuroepidemiology 18, 218-22.

Misu, T., Onodera, H., Fujihara, K., Matsushima, K., Yoshie, O., Okita, N., Takase, S., and Itoyama, Y. (2001). Chemokine receptor expression on T cells in blood

and cerebrospinal fluid at relapse and remission of multiple sclerosis: imbalance of Th1/Th2-associated chemokine signaling. J Neuroimmunol 114, 207-12.

Mitrovic, B., Ignarro, L. J., Vinters, H. V., Akers, M. A., Schmid, I., Uittenbogaart, C., Merrill, J. E. (1995). Nitric oxide induces necrotic but not apoptotic cell death in oligodendrocytes. Neuroscience 65, 531-9.

Moffatt, M. F., Kabesch, M., Liang, L., Dixon, A. L., Strachan, D., Heath, S., Depner, M., von Berg, A., Bufe, A., Rietschel, E., Heinzmann, A., Simma, B., Frischer, T., Willis-Owen, S. A., Wong, K. C., Illig, T., Vogelberg, C., Weiland, S. K., von Mutius, E., Abecasis, G. R., Farrall, M., Gut, I. G., Lathrop, G. M., and Cookson, W. O. (2007). Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 448, 470-3.

Monsuur, A. J., de Bakker, P. I., Alizadeh, B. Z., Zhernakova, A., Bevova, M. R., Strengman, E., Franke, L., van't Slot, R., van Belzen, M. J., Lavrijsen, I. C., Diosdado, B., Daly, M. J., Mulder, C. J., Mearin, M. L., Meijer, J. W., Meijer, G. A., van Oort, E., Wapenaar, M. C., Koeleman, B. P., and Wijmenga, C. (2005). Myosin IXB variant increases the risk of celiac disease and points toward a primary intestinal barrier defect. Nat Genet 37, 1341-4.

Morandi, B., Bramanti, P., Bonaccorsi, I., Montalto, E., Oliveri, D., Pezzino, G., Navarra, M., and Ferlazzo, G. (2008). Role of natural killer cells in the pathogenesis and progression of multiple sclerosis. Pharmacol Res 57, 1-5.

Moreau, T., Manceau, E., Lucas, B., Lemesle, M., Urbinelli, R., and Giroud, M. (2000). Incidence of multiple sclerosis in Dijon, France: a population-based ascertainment. Neurol Res 22, 156-9.

Murray, S., Bashir, K., Penrice, G., and Womersley, S. J. (2004). Epidemiology of multiple sclerosis in Glasgow. Scott Med J 49, 100-4.

Myers, A. J., Gibbs, J. R., Webster, J. A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P., Zismann, V. L., Joshipura, K., Huentelman, M. J., Hu-Lince, D., Coon, K. D., Craig, D. W., Pearson, J. V., Holmans, P., Heward, C. B., Reiman, E. M., Stephan, D., and Hardy J. (2007). A survey of genetic human cortical gene expression. Nat Genet 39, 1494-9.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320, 1344-9.

Nair, A., Frederick, T. J., and Miller, S. D. (2008). Astrocytes in multiple sclerosis: a product of their environment. Cell Mol Life Sci 65, 2702-20.

Nair, R. P., Duffin, K. C., Helms, C., Ding, J., Stuart, P. E., Goldgar, D., Gudjonsson, J. E., Li, Y., Tejasvi, T., Feng, B. J., Ruether, A., Schreiber, S., Weichenthal, M., Gladman, D., Rahman, P., Schrodi, S. J., Prahalad, S., Guthery, S. L., Fischer, J., Liao, W., Kwok, P. Y., Menter, A., Lathrop, G. M., Wise, C. A., Begovich, A. B., Voorhees, J. J., Elder, J. T., Krueger, G. G., Bowcock, A. M., Abecasis, G. R., and Collaborative Association Study of Psoriasis. (2009).

Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. Nat Genet 41, 199-204.

Naito, S., Namerow, N., Mickey, M. R., and Terasaki, P. I. (1972). Multiple sclerosis: association with HL-A3. Tissue Antigens 2, 1-4.

Nakajima, H., Fukuda, K., Doi, Y., Sugino, M., Kimura, F., Hanafusa, T., Ikemoto, T., and Shimizu, A. (2004). Expression of TH1/TH2-related chemokine receptors on peripheral T cells and correlation with clinical disease activity in patients with multiple sclerosis. Eur Neurol 52, 162-8.

Nicoletti, A., Lo Fermo, S., Reggio, E., Tarantello, R., Liberto, A., Le Pira, F., Patti, F., and Reggio, A. (2005a). A possible spatial and temporal cluster of multiple sclerosis in the town of Linguaglossa, Sicily. J Neurol 252, 921-5.

Nicoletti, A., Patti, F., Lo Fermo, S., Sorbello, V., Reggio, E., Maimone, D., Zappia, M., and Reggio, A. (2005b). Possible increasing risk of multiple sclerosis in Catania, Sicily. Neurology 65, 1259-63.

Nie, Y., Waite, J., Brewer, F., Sunshine, M. J., Littman, D. R., and Zou, Y. R. (2004). The role of CXCR4 in maintaining peripheral B cell compartments and humoral immunity. J Exp Med 200, 1145-56.

Núñez, C., Marquez, A., Varade, J., Martinez, A., Polanco, I., Maluenda, C., Fernandez-Arquero, M., de la Concha, E. G., and Urcelay, E. (2006). No evidence of association of the MYO9B polymorphisms with celiac disease in the Spanish population. Tissue Antigens 68, 489-92.

Núñez, C., Oliver, J., Mendoza, J. L., Gomez-Garcia, M., Pinero, A., Taxonera, C., Diaz-Rubio, M., Lopez-Nevot, M. A., de la Concha, E. G., Nieto, A., Urcelay, E., Martinez, A., and Martin, J. (2007). MYO9B polymorphisms in patients with inflammatory bowel disease. Gut 56, 1321-2.

O'Connell, J. R., and Weeks, D. E. (1998). PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63, 259-66.

Oka, A., Belliveau, M. J., Rosenberg, P. A., and Volpe, J. J. (1993). Vulnerability of oligodendroglia to glutamate: pharmacology, mechanisms, and prevention. J Neurosci 13, 1441-53.

Olsson, T. (1992). Cytokines in neuroinflammatory disease: role of myelin autoreactive T cell production of interferon-gamma. J Neuroimmunol 40, 211-8.

Paloneva, J., Kestilä, M., Wu, J., Salminen, A., Böhling, T., Ruotsalainen, V., Hakola, P., Bakker, A. B., Phillips, J. H., Pekkarinen, P., Lanier, L. L., Timonen, T., and Peltonen, L. (2000). Loss-of-function mutations in TYROBP (DAP12) result in a presenile dementia with bone cysts. Nat Genet 25, 357-61.

Papathanasopoulos, P., Gourzoulidou, E., Messinis, L., Georgiou, V., and Leotsinidis, M. (2008). Prevalence and incidence of multiple sclerosis in western Greece: a 23-year survey. Neuroepidemiology 30, 167-73.

Park, H., Li, Z., Yang, X. O., Chang, S. H., Nurieva, R., Wang, Y. H., Wang, Y., Hood, L., Zhu, Z., Tian, Q., and Dong, C. (2005). A distinct lineage of CD4 T

cells regulates tissue inflammation by producing interleukin 17. Nat Immunol 6, 1133-41.

Persengiev, S., Koeleman, B. P., Downes, K., Valdigem, G., van der Slik, A. R., Eerligh, P., Monsuur, A., Bruining, G. J., Wijmenga, C., Todd, J. A., Roep, B. O., and Alizadeh, B. Z. (2010). Association analysis of myosin IXB and type 1 diabetes. Hum Immunol 71, 598-601.

Peterlin, B., Ristic, S., Sepcic, J., Vracko, B. K., Rako, A., Lovrecic, L., Brajenovic-Milic, B., Rudez, J., Materljan, E., and Kapovic, M. (2006). Region with persistent high frequency of multiple sclerosis in Croatia and Slovenia. J Neurol Sci 247, 169-72.

Petukhova, L., Duvic, M., Hordinsky, M., Norris, D., Price, V., Shimomura, Y., Kim, H., Singh, P., Lee, A., Chen, W. V., Meyer, K. C., Paus, R., Jahoda, C. A., Amos, C. I., Gregersen, P. K., and Christiano, A. M. (2010). Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. Nature 466, 113-7.

Pina, M. A., Ara, J. R., Modrego, P. J., Morales, F., and Capablo, J. L. (1998). Prevalence of multiple sclerosis in the sanitary district of Calatayud, Northern Spain: is Spain a zone of high risk for this disease? Neuroepidemiology 17, 258-64.

Plotkin, J., Prockop, S. E., Lepique, A., and Petrie, H. T. (2003). Critical role for CXCR4 signaling in progenitor localization and T cell differentiation in the postnatal thymus. J Immunol 171, 4521-7.

Polman, C. H., Reingold, S. C., Edan, G., Filippi, M., Hartung, H. P., Kappos, L., Lublin, F. D., Metz, L. M., McFarland, H. F., O'Connor, P. W., Sandberg-Wollheim, M., Thompson, A. J., Weinshenker, B. G., and Wolinsky, J. S. (2005). Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol 258, 840-6.

Poser, C. M., Paty, D. W., Scheinberg, L., McDonald, W. I., Davis, F. A., Ebers, G. C., Johnson, K. P., Sibley, W. A., Silberberg, D. H., and Tourtellotte, W. W. (1983). New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 13, 227-31.

Poser, C. M. (1995). Viking voyages: the origin of multiple sclerosis? An essay in medical history. Acta Neurol Scand Suppl 161, 11-22.

Post, P. L., Bokoch, G. M., and Mooseker, M. S. (1998). Human myosin-IXb is a mechanochemically active motor and a GAP for rho. J Cell Sci 111 (Pt 7), 941-50.

Potemkowski, A. (1999). [Epidemiology of multiple sclerosis in the region of Szczecin: prevalence and incidence 1993-1995]. Neurol Neurochir Pol 33, 575-85.

Pugliatti, M., Sotgiu, S., and Rosati, G. (2002). The worldwide prevalence of multiple sclerosis. Clin Neurol Neurosurg 104, 182-91.

Pugliatti, M., Riise, T., Sotgiu, M. A., Sotgiu, S., Satta, W. M., Mannu, L., Sanna, G., and Rosati, G. (2005). Increasing incidence of multiple sclerosis in the province of Sassari, northern Sardinia. Neuroepidemiology 25, 129-34.

Purcell, S., Cherny, S. S., and Sham, P. C. (2003). Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19, 149-50.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., and Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81, 559-75.

Ragonese, P., Salemi, G., D'Amelio, M., Gammino, M., Aridon, P., and Savettieri, G. (2004). Multiple sclerosis in southern Europe: Monreale City, Italy. A twenty-year follow-up incidence and prevalence study. Neuroepidemiology 23, 306-9.

Raivich, G., and Banati, R. (2004). Brain microglia and blood-derived macrophages: molecular profiles and functional roles in multiple sclerosis and animal models of autoimmune demyelinating disease. Brain Res Brain Res Rev 46, 261-81.

Ramagopalan, S. V., Valdar, W., Criscuoli, M., DeLuca, G. C., Dyment, D. A., Orton, S. M., Yee, I. M., Ebers, G. C., and Sadovnick, A. D. (2009). Age of puberty and the risk of multiple sclerosis: a population based study. Eur J Neurol 16, 342-7.

Ramanathan, M., Weinstock-Guttman, B., Nguyen, L. T., Badgett, D., Miller, C., Patrick, K., Brownscheidle, C., and Jacobs, L. (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.

Ramel, C. (1997). Mini- and microsatellites. Environ Health Perspect 105, 781-789.

Rice-Oxley, M., Williams, E. S., and Rees, J. E. (1995). A prevalence survey of multiple sclerosis in Sussex. J Neurol Neurosurg Psychiatry 58, 27-30.

Riise Stensland, H. M., Saarela, J., Bronnikov, D. O., Parkkonen, M., Jokiaho, A. J., Palotie, A., Tienari, P. J., Sumelahti, M. L., Elovaara, I., Koivisto, K., Pirttila, T., Reunanen, M., Sobel, E., and Peltonen, L. (2005). Fine mapping of the multiple sclerosis susceptibility locus on 5p14-p12. J Neuroimmunol 170, 122-33.

Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science 273, 1516-7.

Riva, A., and Kohane, I. S. (2002). SNPper: retrieval and analysis of human SNPs. Bioinformatics 18, 1681-5.

Robertson, N., Deans, J., Fraser, M., and Compston, D. A. (1996). Multiple sclerosis in south Cambridgeshire: incidence and prevalence based on a district register. J Epidemiol Community Health 50, 274-9.

Rothwell, P. M., and Charlton, D. (1998). High incidence and prevalence of multiple sclerosis in south east Scotland: evidence of a genetic predisposition. J Neurol Neurosurg Psychiatry 64, 730-5.

Roxburgh, R. H., Seaman, S. R., Masterman, T., Hensiek, A. E., Sawcer, S. J., Vukusic, S., Achiti, I., Confavreux, C., Coustans, M., le Page, E., Edan, G., McDonnell, G. V., Hawkins, S., Trojano, M., Liguori, M., Cocco, E., Marrosu, M. G., Tesser, F., Leone, M. A., Weber, A., Zipp, F., Miterski, B., Epplen, J. T., Oturai, A., Sorensen, P. S., Celius, E. G., Lara, N. T., Montalban, X., Villoslada, P., Silva, A. M., Marta, M., Leite, I., Dubois, B., Rubio, J., Butzkueven, H., Kilpatrick, T., Mycko, M. P., Selmaj, K. W., Rio, M. E., Sa, M., Salemi, G., Savettieri, G., Hillert, J., and Compston, D. A. (2005). Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. Neurology 64, 1144-51.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132, 365-86.

Saarela, J., Schoenberg Fejzo, M., Chen, D., Finnila, S., Parkkonen, M., Kuokkanen, S., Sobel, E., Tienari, P. J., Sumelahti, M. L., Wikstrom, J., Elovaara, I., Koivisto, K., Pirttila, T., Reunanen, M., Palotie, A., and Peltonen, L. (2002). Fine mapping of a multiple sclerosis locus to 2.5 Mb on chromosome 17q22-q24. Hum Mol Genet 11, 2257-67.

Saarela, J., Kallio, S. P., Chen, D., Montpetit, A., Jokiaho, A., Choi, E., Asselta, R., Bronnikov, D., Lincoln, M. R., Sadovnick, A. D., Tienari, P. J., Koivisto, K., Palotie, A., Ebers, G. C., Hudson, T. J., and Peltonen, L. (2006). PRKCA and multiple sclerosis: association in two independent populations. PLoS Genet 2, e42.

Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G., Sherry, S., Mullikin, J. C., Mortimore, B. J., Willey, D. L., Hunt, S. E., Cole, C. G., Coggill, P. C., Rice, C. M., Ning, Z., Rogers, J., Bentley, D. R., Kwok, P. Y., Mardis, E. R., Yeh, R. T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R. H., McPherson, J. D., Gilman, B., Schaffner, S., Van Etten, W. J., Reich, D., Higgins, J., Daly, M. J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M. C., Linton, L., Lander, E. S., and Altshuler, D. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409, 928-33.

Sadovnick, A. D., and Ebers, G. C. (1993). Epidemiology of multiple sclerosis: a critical overview. Can J Neurol Sci 20, 17-29.

Sakaguchi, S. (2000). Regulatory T cells: key controllers of immunologic self-tolerance. Cell 101, 455-8.

Sánchez, E., Alizadeh, B. Z., Valdigem, G., Ortego-Centeno, N., Jiménez-Alonso, J., de Ramón, E., Garcia, A., López-Nevot, M. A., Wijmenga, C., Martin, J., and Koeleman, B. P. (2007). MYO9B gene polymorphisms are associated with autoimmune diseases in Spanish population. Hum Immunol 68, 610-5.

Sanna, S., Pitzalis, M., Zoledziewska, M., Zara, I., Sidore, C., Murru, R., Whalen, M. B., Busonero, F., Maschio, A., Costa, G., Melis, M. C., Deidda, F., Poddie, F., Morelli, L., Farina, G., Li, Y., Dei, M., Lai, S., Mulas, A., Cuccuru, G., Porcu, E.,

Liang, L., Zavattari, P., Moi, L., Deriu, E., Urru, M. F., Bajorek, M., Satta, M. A., Cocco, E., Ferrigno, P., Sotgiu, S., Pugliatti, M., Traccis, S., Angius, A., Melis, M., Rosati, G., Abecasis, G. R., Uda, M., Marrosu, M. G., Schlessinger, D., and Cucca, F. Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. (2010). Nat Genet 42, 495-7.

Santiago, J. L., Martínez, A., Núñez, C., de la Calle, H., Fernández-Arquero, M., de la Concha, E. G., and Urcelay, E. (2008). Association of MYO9B haplotype with type 1 diabetes. Hum Immunol 69, 112-5.

Saraheimo, M., Forsblom, C., Pettersson-Fernholm, K., Flyvbjerg, A., Groop, P. H., and Frystyk, J. (2008). Increased levels of alpha-defensin (-1, -2 and -3) in type 1 diabetic patients with nephropathy. Nephrol Dial Transplant 23, 914-8.

Sarasoja, T., Wikstrom, J., Paltamaa, J., Hakama, M., and Sumelahti, M. L. (2004). Occurrence of multiple sclerosis in central Finland: a regional and temporal comparison during 30 years. Acta Neurol Scand 110, 331-6.

Särkijärvi, S., Kuusisto, H., Paalavuo, R., Levula, M., Airla, N., Lehtimäki, T., Kaprio, J., Koskenvuo, M., and Elovaara, I. (2006). Gene expression profiles in Finnish twins with multiple sclerosis. BMC Med Genet 7, 11.

Satoh, J., Nakanishi, M., Koike, F., Miyake, S., Yamamoto, T., Kawai, M., Kikuchi, S., Nomura, K., Yokoyama, K., Ota, K., Kanda, T., Fukazawa, T., and Yamamura, T. (2005). Microarray analysis identifies an aberrant expression of apoptosis and DNA damage-regulatory genes in multiple sclerosis. Neurobiol Dis 18, 537-50.

Satoh, J., Nakanishi, M., Koike, F., Onoue, H., Aranami, T., Yamamoto, T., Kawai, M., Kikuchi, S., Nomura, K., Yokoyama, K., Ota, K., Saito, T., Ohta, M., Miyake, S., Kanda, T., Fukazawa, T., and Yamamura, T. (2006). T cell gene expression profiling identifies distinct subgroups of Japanese multiple sclerosis patients. J Neuroimmunol 174, 108-18.

Sawcer, S., Ban, M., Maranian, M., Yeo, T. W., Compston, A., Kirby, A., Daly, M. J., De Jager, P. L., Walsh, E., Lander, E. S., Rioux, J. D., Hafler, D. A., Ivinson, A., Rimmler, J., Gregory, S. G., Schmidt, S., Pericak-Vance, M. A., Akesson, E., Hillert, J., Datta, P., Oturai, A., Ryder, L. P., Harbo, H. F., Spurkland, A., Myhr, K. M., Laaksonen, M., Booth, D., Heard, R., Stewart, G., Lincoln, R., Barcellos, L. F., Hauser, S. L., Oksenberg, J. R., Kenealy, S. J., and Haines, J. L. (2005). A high-density screen for linkage in multiple sclerosis. Am J Hum Genet 77, 454-67.

Sawcer, S., Jones, H. B., Feakes, R., Gray, J., Smaldon, N., Chataway, J., Robertson, N., Clayton, D., Goodfellow, P. N., and Compston, A. (1996). A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. Nat Genet 13, 464-8.

Sawcer, S. (2008). The complex genetics of multiple sclerosis: pitfalls and prospects. Brain 131, 3118-31.

Schadt, E. E., Molony, C., Chudin, E., Hao, K., Yang, X., Lum, P. Y., Kasarskis, A., Zhang, B., Wang, S., Suver, C., Zhu, J., Millstein, J., Sieberts, S., Lamb, J., GuhaThakurta, D., Derry, J., Storey, J. D., Avila-Campillo, I., Kruger, M. J., Johnson, J. M., Rohl, C. A., van Nas, A., Mehrabian, M., Drake, T. A., Lusis, A. J., Smith, R. C., Guengerich, F. P., Strom, S. C., Schuetz, E., Rushmore, T. H., and Ulrich, R. (2008). Mapping the genetic architecture of gene expression in human liver. PLoS Biol 6, e107.

Schmidt, H., Williamson, D., and Ashley-Koch, A. (2007). HLA-DR15 haplotype and multiple sclerosis: a HuGE review. Am J Epidemiol 165, 1097-109.

Schrijver, H. M. (2008). Rituximab in relapsing-remitting multiple sclerosis. N Engl J Med 358, 2645; author reply 2646-7.

Schwab, N., Zozulya, A. L., Kieseier, B. C., Toyka, K. V., and Wiendl, H. (2010). An imbalance of two functionally and phenotypically different subsets of plasmacytoid dendritic cells characterizes the dysfunctional immune regulation in multiple sclerosis. J Immunol 184, 5368-74.

Screaton, G. R., Xu, X. N., Olsen, A. L., Cowper, A. E., Tan, R., McMichael, A. J., and Bell, J. I. (1997). LARD: a new lymphoid-specific death domain containing receptor regulated by alternative pre-mRNA splicing. Proc Natl Acad Sci U S A 94, 4615-9.

Segal, B. M. (2010). Th17 cells in autoimmune demyelinating disease. Semin Immunopathol 32, 71-7.

Serafini, B., Rosicarelli, B., Franciotta, D., Magliozzi, R., Reynolds, R., Cinque, P., Andreoni, L., Trivedi, P., Salvetti, M., Faggioni, A., and Aloisi, F. (2007). Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain. J Exp Med 204, 2899-912.

Sharpe, G., Price, S. E., Last, A., and Thompson, R. J. (1995). Multiple sclerosis in island populations: prevalence in the Bailiwicks of Guernsey and Jersey. J Neurol Neurosurg Psychiatry 58, 22-6.

Sinsheimer, J. S., Blangero, J., and Lange, K. (2000). Gamete-competition models. Am J Hum Genet 66, 1168-72.

Shortman, K., and Liu, Y. J. (2002). Mouse and human dendritic cell subtypes. Nat Rev Immunol 2, 151-61.

Shuster, E. A. (2008). Hormonal influences in multiple sclerosis. Curr Top Microbiol Immunol 318, 267-311.

Smestad, C., Sandvik, L., Holmoy, T., Harbo, H. F., and Celius, E. G. (2008). Marked differences in prevalence of multiple sclerosis between ethnic groups in Oslo, Norway. J Neurol 255, 49-55.

Solaro, C., Allemani, C., Messmer Uccelli, M., Canevari, E., Dagnino, N., Pizio, R., Regesta, G., Tanganelli, P., Battaglia, M. A., and Mancardi, G. L. (2005). The prevalence of multiple sclerosis in the north-west Italian province of Genoa. J Neurol 252, 436-40.

Sospedra, M., and Martin, R. (2005). Immunology of multiple sclerosis. Annu Rev Immunol 23, 683-747.

Spielman, R. S., Bastone, L. A., Burdick, J. T., Morley, M., Ewens, W. J., and Cheung, V. G. (2007). Common genetic variants account for differences in gene expression among ethnic groups. Nat Genet 39, 226-31.

Stahl, E. A., Raychaudhuri, S., Remmers, E. F., Xie, G., Eyre, S., Thomson, B. P., Li, Y., Kurreeman, F. A., Zhernakova, A., Hinks, A., Guiducci, C., Chen, R., Alfredsson, L., Amos, C. I., Ardlie, K. G., BIRAC Consortium, Barton, A., Bowes, J., Brouwer, E., Burtt, N. P., Catanese, J. J., Coblyn, J., Coenen, M. J., Costenbader, K. H., Criswell, L. A., Crusius, J. B., Cui, J., de Bakker, P. I., De Jager, P. L., Ding, B., Emery, P., Flynn, E., Harrison, P., Hocking, L. J., Huizinga, T. W., Kastner, D. L., Ke, X., Lee, A. T., Liu, X., Martin, P., Morgan, A. W., Padyukov, L., Posthumus, M. D., Radstake, T. R., Reid, D. M., Seielstad, M., Seldin, M. F., Shadick, N. A., Steer, S., Tak, P. P., Thomson, W., van der Helm-van Mil, A. H., van der Horst-Bruinsma, I. E., van der Schoot, C. E., van Riel, P. L., Weinblatt, M. E., Wilson, A. G., Wolbink, G. J., Wordsworth, B. P., YEAR Consortium, Wijmenga, C., Karlson, E. W., Toes, R. E., de Vries, N., Begovich, A. B., Worthington, J., Siminovitch, K. A., Gregersen, P. K., Klareskog, L., and Plenge, R. M. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42, 508-14.

Stamatovic, S. M., Dimitrijevic, O. B., Keep, R. F., and Andjelkovic, A. V. (2006). Protein kinase Calpha-RhoA cross-talk in CCL2-induced alterations in brain endothelial permeability. J Biol Chem 281, 8379-88.

Stasiolek, M., Bayas, A., Kruse, N., Wieczarkowiecz, A., Toyka, K. V., Gold, R., and Selmaj, K. (2006). Impaired maturation and altered regulatory function of plasmacytoid dendritic cells in multiple sclerosis. Brain 129, 1293-305.

Stefansson, B., and Brautigan, D. L. (2006). Protein phosphatase 6 subunit with conserved Sit4-associated protein domain targets IkappaBepsilon. J Biol Chem 281, 22624-34.

Sthoeger, Z. M., Bezalel, S., Chapnik, N., Asher, I., and Froy, O. (2009). High alpha-defensin levels in patients with systemic lupus erythematosus. Immunology 127, 116-22.

Stranger, B. E., Nica, A. C., Forrest, M. S., Dimas, A., Bird, C. P., Beazley, C., Ingle, C. E., Dunning, M., Flicek, P., Koller, D., Montgomery, S., Tavaré, S., Deloukas, P., and Dermitzakis, E. T. (2007). Population genomics of human gene expression. Nat Genet 39, 1217-24.

Storch, M. K., Piddlesden, S., Haltia, M., Iivanainen, M., Morgan, P., and Lassmann, H. (1998). Multiple sclerosis: in situ evidence for antibody- and complement-mediated demyelination. Ann Neurol 43, 465-71.

Sulonen, A. M., Kallio, S. P., Ellonen, P., Suvela, M., Elovaara, I., Koivisto, K., Pirttilä, T., Reunanen, M., Tienari, P. J., Palotie, A., Peltonen, L., and Saarela J.

Studies on causes of
multiple sclerosis

(2009). No evidence for shared etiology in two demyelinative disorders, MS and PLOSL. J Neuroimmunol 206, 86-90.

Sumelahti, M. L., Tienari, P. J., Wikstrom, J., Palo, J., and Hakama, M. (2001). Increasing prevalence of multiple sclerosis in Finland. Acta Neurol Scand 103, 153-8.

Sundstrom, P., Nystrom, L., and Forsgren, L. (2003). Incidence (1988-97) and prevalence (1997) of multiple sclerosis in Vasterbotten County in northern Sweden. J Neurol Neurosurg Psychiatry 74, 29-32.

Sundstrom, P., Nystrom, L., and Hallmans, G. (2008). Smoke exposure increases the risk for multiple sclerosis. Eur J Neurol 15, 579-83.

Sundvall, M., Jirholt, J., Yang, H. T., Jansson, L., Engstrom, A., Pettersson, U., and Holmdahl, R. (1995). Identification of murine loci associated with susceptibility to chronic experimental autoimmune encephalomyelitis. Nat Genet 10, 313-7.

Suzuki, G., Nakata, Y., Dan, Y., Uzawa, A., Nakagawa, K., Saito, T., Mita, K., and Shirasawa, T. (1998). Loss of SDF-1 receptor expression during positive selection in the thymus. Int Immunol 10, 1049-56.

Tergaonkar, V., Correa, R. G., Ikawa, M., and Verma, I. M. (2005). Distinct roles of IkappaB proteins in regulating constitutive NF-kappaB activity. Nat Cell Biol 7, 921-3.

Tienari, P. J., Wikstrom, J., Sajantila, A., Palo, J., and Peltonen, L. (1992). Genetic susceptibility to multiple sclerosis linked to myelin basic protein gene. Lancet 340, 987-91.

Tienari, P. J., Wikstrom, J., Koskimies, S., Partanen, J., Palo, J., and Peltonen, L. (1993). Reappraisal of HLA in multiple sclerosis: close linkage in multiplex families. Eur J Hum Genet 1, 257-68.

Tienari, P. J., Sumelahti, M. L., Rantamaki, T., and Wikstrom, J. (2004). Multiple sclerosis in western Finland: evidence for a founder effect. Clin Neurol Neurosurg 106, 175-9.

Todd, J. A., Walker, N. M., Cooper, J. D., Smyth, D. J., Downes, K., Plagnol, V., Bailey, R., Nejentsev, S., Field, S. F., Payne, F., Lowe, C. E., Szeszko, J. S., Hafler, J. P., Zeitels, L., Yang, J. H., Vella, A., Nutland, S., Stevens, H. E., Schuilenburg, H., Coleman, G., Maisuria, M., Meadows, W., Smink, L. J., Healy, B., Burren, O. S., Lam, A. A., Ovington, N. R., Allen, J., Adlem, E., Leung, H. T., Wallace, C., Howson, J. M., Guja, C., Ionescu-Tîrgovişte, C., Genetics of Type 1 Diabetes in Finland, Simmonds, M. J., Heward, J. M., Gough, S. C., Wellcome Trust Case Control Consortium, Dunger, D. B., Wicker, L. S., and Clayton, D. G. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. (2007). Nat Genet 39, 857-64.

Tola, M. A., Yugueros, M. I., Fernandez-Buey, N., and Fernandez-Herranz, R. (1999). Prevalence of multiple sclerosis in Valladolid, northern Spain. J Neurol 246, 170-4.

Tuohy, V. K., and Kinkel, R. P. (2000). Epitope spreading: a mechanism for progression of autoimmune disease. Arch Immunol Ther Exp (Warsz) 48, 347-51.

Uria, D. F., Abad, P., Calatayud, M. T., Virgala, P., Diaz, A., Chamizo, C., and Dean, G. (1997). Multiple sclerosis in Gijon health district, Asturias, northern Spain. Acta Neurol Scand 96, 375-9.

van Bodegraven, A. A., Curley, C. R., Hunt, K. A., Monsuur, A. J., Linskens, R. K., Onnie, C. M., Crusius, J. B., Annese, V., Latiano, A., Silverberg, M. S., Bitton, A., Fisher, S. A., Steinhart, A. H., Forbes, A., Sanderson, J., Prescott, N. J., Strachan, D. P., Playford, R. J., Mathew, C. G., Wijmenga, C., Daly, M. J., Rioux, J. D., and van Heel, D. A. (2006). Genetic variation in myosin IXB is associated with ulcerative colitis. Gastroenterology 131, 1768-74.

van Ooteghem, P., D'Hooghe, M. B., Vlietinck, R., and Carton, H. (1994). Prevalence of multiple sclerosis in Flanders, Belgium. Neuroepidemiology 13, 220-5.

Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995). Serial analysis of gene expression. Science 270, 484-7.

Veldhuis, W. B., Floris, S., van der Meide, P. H., Vos, I. M., de Vries, H. E., Dijkstra, C. D., Bar, P. R., and Nicolay, K. (2003). Interferon-beta prevents cytokine-induced neutrophil infiltration and attenuates blood-brain barrier disruption. J Cereb Blood Flow Metab 23, 1060-9.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., *et al.* (2001). The sequence of the human genome. Science 291, 1304-51.

Veyrieras, J. B., Kudaravalli, S., Kim, S. Y., Dermitzakis, E. T., Gilad, Y., Stephens, M., and Pritchard, J. K. (2008). High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet 4, e1000214.

Viglietta, V., Baecher-Allan, C., Weiner, H. L., and Hafler, D. A. (2004). Loss of functional suppression by CD4+CD25+ regulatory T cells in patients with multiple sclerosis. J Exp Med 199, 971-9.

Wallace, C., Smyth, D. J., Maisuria-Armer, M., Walker, N. M., Todd, J. A., and Clayton, D. G. (2010). The imprinted DLK1-MEG3 gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. Nat Genet 42, 68-71.

Wang, E. C., Thern, A., Denzel, A., Kitson, J., Farrow, S. N., and Owen, M. J. (2001). DR3 regulates negative selection during thymocyte development. Mol Cell Biol 21, 3451-61.

Wang, M., Windgassen, D., and Papoutsakis, E. T. (2008). A global transcriptional view of apoptosis in human T-cell activation. BMC Med Genomics 1, 53.

Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661-78.

Wellcome Trust Case Control Consortium, Craddock, N., Hurles, M. E., Cardin, N., Pearson, R. D., Plagnol, V., Robson, S., Vukcevic, D., Barnes, C., Conrad, D. F., Giannoulatou, E., Holmes, C., Marchini, J. L., Stirrups, K., Tobin, M. D., Wain, L. V., Yau, C., Aerts, J., Ahmad, T., Andrews, T. D., Arbury, H., Attwood, A., Auton, A., Ball, S. G., Balmforth, A. J., Barrett, J. C., Barroso, I., Barton, A., Bennett, A. J., Bhaskar, S., Blaszczyk, K., Bowes, J., Brand, O. J., Braund, P. S., Bredin, F., Breen, G., Brown, M. J., Bruce, I. N., Bull, J., Burren, O. S., Burton, J., Byrnes, J., Caesar, S., Clee, C. M., Coffey, A. J., Connell, J. M., Cooper, J. D., Dominiczak, A. F., Downes, K., Drummond, H. E., Dudakia, D., Dunham, A., Ebbs, B., Eccles, D., Edkins, S., Edwards, C., Elliot, A., Emery, P., Evans, D. M., Evans, G., Eyre, S., Farmer, A., Ferrier, I. N., Feuk, L., Fitzgerald, T., Flynn, E., Forbes, A., Forty, L., Franklyn, J. A., Freathy, R. M., Gibbs, P., Gilbert, P., Gokumen, O., Gordon-Smith, K., Gray, E., Green, E., Groves, C. J., Grozeva, D., Gwilliam, R., Hall, A., Hammond, N., Hardy, M., Harrison, P., Hassanali, N., Hebaishi, H., Hines, S., Hinks, A., Hitman, G. A., Hocking, L., Howard, E., Howard, P., Howson, J. M., Hughes, D., Hunt, S., Isaacs, J. D., Jain, M., Jewell, D. P., Johnson, T., Jolley, J. D., Jones, I. R., Jones, L. A., Kirov, G., Langford, C. F., Lango-Allen, H., Lathrop, G. M., Lee, J., Lee, K. L., Lees, C., Lewis, K., Lindgren, C. M., Maisuria-Armer, M., Maller, J., Mansfield, J., Martin, P., Massey, D. C., McArdle, W. L., McGuffin, P., McLay, K. E., Mentzer, A., Mimmack, M. L., Morgan, A. E., Morris, A. P., Mowat, C., Myers, S., Newman, W., Nimmo, E. R., O'Donovan, M. C., Onipinla, A., Onyiah, I., Ovington, N. R., Owen, M. J., Palin, K., Parnell, K., Pernet, D., Perry, J. R., Phillips, A., Pinto, D., Prescott, N. J., Prokopenko, I., Quail, M. A., Rafelt, S., Rayner, N. W., Redon, R., Reid, D. M., Renwick, Ring, S. M., Robertson, N., Russell, E., St Clair, D., Sambrook, J. G., Sanderson, J. D., Schuilenburg, H., Scott, C. E., Scott, R., Seal, S., Shaw-Hawkins, S., Shields, B. M., Simmonds, M. J., Smyth, D. J., Somaskantharajah, E., Spanova, K., Steer, S., Stephens, J., Stevens, H. E., Stone,

M. A., Su, Z., Symmons, D. P., Thompson, J. R., Thomson, W., Travers, M. E., Turnbull, C., Valsesia, A., Walker, M., Walker, N. M., Wallace, C., Warren-Perry, M., Watkins, N. A., Webster, J., Weedon, M. N., Wilson, A. G., Woodburn, M., Wordsworth, B. P., Young, A. H., Zeggini, E., Carter, N. P., Frayling, T. M., Lee, C., McVean, G., Munroe, P. B., Palotie, A., Sawcer, S. J., Scherer, S. W., Strachan, D. P., Tyler-Smith, C., Brown, M. A., Burton, P. R., Caulfield, M. J., Compston, A., Farrall, M., Gough, S. C., Hall, A. S., Hattersley, A. T., Hill, A. V., Mathew, C. G., Pembrey, M., Satsangi, J., Stratton, M. R., Worthington, J., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W., Parkes, M., Rahman, N., Todd, J. A., Samani, N. J., and Donnelly. P. (2010). Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464, 713-20.

Whitney, A. R., Diehn, M., Popper, S. J., Alizadeh, A. A., Boldrick, J. C., Relman, D. A., and Brown, P. O. (2003). Individuality and variation in gene expression patterns in human blood. Proc Natl Acad Sci U S A 100, 1896-901.

Whitney, L. W., Ludwin, S. K., McFarland, H. F., and Biddison, W. E. (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.

Willer, C. J., Dyment, D. A., Risch, N. J., Sadovnick, A. D., and Ebers, G. C. (2003). Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci U S A 100, 12877-82.

Willis, S. N., Stadelmann, C., Rodig, S.J., Caron, T., Gattenloehner, S., Mallozzi, S. S., Roughan, J. E., Almendinger, S. E., Blewett, M. M., Brück, W., Hafler, D. A., and O'Connor, K. C.(2009). Epstein-Barr virus infections is not a characteristic feature of multiple sclerosis brain. Brain 132, 3318-28.

Wolswijk, G. (1998). Chronic stage multiple sclerosis lesions contain a relatively quiescent population of oligodendrocyte precursor cells. J Neurosci 18, 601-9.

Wolters, V. M., Verbeek, W. H., Zhernakova, A., Onland-Moret, C., Schreurs, M. W., Monsuur, A. J., Verduijn, W., Wijmenga, C., and Mulder, C. J. (2007). The MYO9B gene is a strong risk factor for developing refractory celiac disease. Clin Gastroenterol Hepatol 5, 1399-405, 1405 e1-2.

Wucherpfennig, K. W., and Strominger, J. L. (1995). Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell 80, 695-705.

Yamaguchi, N., Isomoto, H., Mukae, H., Ishimoto, H., Ohnita, K., Shikuwa, S., Mizuta, Y., Nakazato, M., and Kohno, S. (2009). Concentrations of alpha- and beta-defensins in plasma of patients with inflammatory bowel disease. Inflamm Res 58, 192-7.

Yang, B., Akhter, S., Chaudhuri, A., and Kanmogne, G. D. (2009). HIV-1 gp120 induces cytokine expression, leukocyte adhesion, and transmigration across the

blood-brain barrier: modulatory effects of STAT1 signaling. Microvasc Res 77, 212-9.

Yeo, T. W., De Jager, P. L., Gregory, S. G., Barcellos, L. F., Walton, A., Goris, A., Fenoglio, C., Ban, M., Taylor, C. J., Goodman, R. S., Walsh, E., Wolfish, C. S., Horton, R., Traherne, J., Beck, S., Trowsdale, J., Caillier, S. J., Ivinson, A. J., Green, T., Pobywajlo, S., Lander, E. S., Pericak-Vance, M. A., Haines, J. L., Daly, M. J., Oksenberg, J. R., Hauser, S. L., Compston, A., Hafler, D. A., Rioux, J. D., and Sawcer, S. (2007). A second major histocompatibility complex susceptibility locus for multiple sclerosis. Ann Neurol 61, 228-36.

Zeller, T., Wild, P., Szymczak, S., Rotival, M., Schillert, A., Castagne, R., Maouche, S., Germain, M., Lackner, K., Rossmann, H., Eleftheriadis, M., Sinning, C. R., Schnabel, R. B., Lubos, E., Mennerich, D., Rust, W., Perret, C., Proust, C., Nicaud, V., Loscalzo, J., Hübner, N., Tregouet, D., Münzel, T., Ziegler, A., Tiret, L., Blankenberg, S., and Cambien, F. (2010). Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. PLoS One 5, e10693.

Zhang, X. J., Huang, W., Yang, S., Sun, L. D., Zhang, F. Y., Zhu, Q. X., Zhang, F. R., Zhang, C., Du, W. H., Pu, X. M., Li, H., Xiao, F. L., Wang, Z. X., Cui, Y., Hao, F., Zheng, J., Yang, X. Q., Cheng, H., He, C. D., Liu, X. M., Xu, L. M., Zheng, H. F., Zhang, S. M., Zhang, J. Z., Wang, H. Y., Cheng, Y. L., Ji, B. H., Fang, Q. Y., Li, Y. Z., Zhou, F. S., Han, J. W., Quan, C., Chen, B., Liu, J. L., Lin, D., Fan, L., Zhang, A. P., Liu, S. X., Yang, C. J., Wang, P. G., Zhou, W. M., Lin, G. S., Wu, W. D., Fan, X., Gao, M., Yang, B. Q., Lu, W. S., Zhang, Z., Zhu, K. J., Shen, S. K., Li, M., Zhang, X. Y., Cao, T. T., Ren, W., Zhang, X., He, J., Tang, X. F., Lu, S., Yang, J. Q., Zhang, L., Wang, D. N., Yuan, F., Yin, X. Y., Huang, H. J., Wang, H. F., Lin, X. Y., and Liu, J. J. (2009). Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21. Nat Genet 41, 205-10.

Zou, Y. R., Kottmann, A. H., Kuroda, M., Taniuchi, I., and Littman, D. R. (1998). Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. Nature 393, 595-9.