GENETICS OF MULTIPLE SCLEROSIS

Virpi Leppä

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine of the University of Helsinki, for public examination in lecture hall 2, Biomedicum 1, on 4 January 2013, at 12 noon.

Institute for Molecular Medicine Finland FIMM,
University of Helsinki, Helsinki, Finland
and
Department of Medical Genetics,
Haartman Institute, University of Helsinki, Helsinki, Finland
and
Public Health Genomics Unit
National Institute for Health and Welfare, Helsinki, Finland

Finland 2012

Helsinki University Biomedical Dissertations Series No. 181

ISSN 1457-8433
ISBN 978-952-10-8560-4 (pbk.)
ISBN 978-952-10-8561-1 (PDF)
http://ethesis.helsinki.fi/

Unigrafia Oy
Helsinki, Finland 2012

Supervised By

Docent Janna Saarela
Institute for Molecular Medicine Finland FIMM, University of Helsinki,
Helsinki, Finland

Professor Aarno Palotie
The Wellcome Trust Sanger Institute, Cambridge, United Kingdom
Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland

Academician of Science, Professor Leena Palotie
The Wellcome Trust Sanger Institute, Cambridge, United Kingdom
Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland
National Institute for Health and Welfare (THL), Helsinki, Finland

Reviewed by

Professor Hannes Lohi
Department of Veterinary Biosciences, Faculty of Veterinary Medicine,
Research Programs Unit, Molecular Medicine, Faculty of Medicine,
University of Helsinki and Folkhälsan Research Center, Helsinki, Finland

Professor Marju Orho-Melander
Department of Clinical Sciences Malmö,
Diabetes and Cardiovascular Disease Genetic Epidemiology Unit,
University Hospital in Skåne,
Lund University, Malmö, Sweden

Opponent

Professor Paolo Gasparini
S.C. Medical Genetics, Institute for Maternal and Child Health, IRCCS ''Burlo Garofolo''
Trieste, Italy

Läheisilleni
To my dearest

ABSTRACT

Multiple sclerosis (MS) is a complex autoimmune disease of the central nervous system. Current available treatments can slow down the disease progression in the majority of patients, but can neither stop the progression nor cure the patients. MS is a debilitating neurological disease of young adults and it affects 0.1% of the populations of Northern European descent. The prevalence of MS varies globally and locally, and there are several isolated high-risk populations. One such high-risk population is in the Southern Ostrobothnia region of Finland, where the incidence and prevalence are approximately twice those of the surrounding regions. Both environmental and genetic factors are thought to contribute to disease pathogenesis. Twin, half-sib and adoption studies all point towards underlying genetic factors that predispose to MS. In recent years, genome-wide association studies have revealed over fifty genes that are associated with MS. At the start of this thesis project, only one MS locus, *HLA-DRB1*, was known, whilst shortly after the start two more loci, *IL2RA* and *IL7R*, were reported.

The aim of this thesis was to find new associated loci, confirm newly found loci and assess the relevance of these known loci within the Southern Ostrobothnian high-risk isolate, and in the Finnish population overall. The association analyses were performed using genome-wide association analysis methods, haplotype analysis and meta-analysis methods. In the first study, 68 Southern Ostrobothnian cases and 136 identity-by-state matched controls were analyzed in a genome-wide association analysis, where an association signal from the SNP rs744166 in the *STAT3* gene was discovered. The association signal was replicated in an international sample set of 4487 MS cases and 9778 controls (OR 1.18, 95% CI 1.12-1.24, $p=2.75\times10^{-10}$). In order to further understand the association signal, a haplotype analysis was performed, revealing that the association signal was on a common haplotype. Further, in a genome-wide homozygosity analysis, excess homozygosity was found in the 68 cases compared to the 136 controls in three genomic loci, 1q42.12, 2q24.3 and 12q24.33. The importance of these loci requires further study.

The second study aimed to replicate loci that had previously been observed to associate with MS in an international genome-wide meta-analysis. Three SNPs in three loci, *IRF8, TNFRSF1A* and *CD6,* were genotyped in 608 trios, 8439 MS cases, and 9280 controls from 11 populations of European origin, replicating the associations in all three loci. The SNP in *IRF8* was strongly associated with MS, while the SNPs in *TNTRSF1A* and *CD6* indicated slightly weaker association. The odds ratios for all the three loci were similarly modest: *IRF8* OR 1.11, *TNFRSF1A* OR 1.12, and *CD6* OR 1.11.

Since the prevalence and incidence of MS is two-fold in the Southern Ostrobothnian region and there is evidence for founder effect, the aim of the last study was to assess the accumulation of recently identified genome-wide significant MS-associated common risk alleles in the Southern Ostrobothnian isolate MS cases

and in familial samples. However, we could not detect accumulation of common variants in the isolate region or in the families compared to general Finnish samples or non-familial samples.

Keywords: genome-wide association, multiple sclerosis, autoimmunity, meta-analysis, *STAT3*, replication, *IRF8*, *TNFRSF1A*, *CD6* genetic

ACKNOWLEDGEMENTS

This study was carried out at the Institute for Molecular Medicine FIMM, University of Helsinki, and at the Public Health Genomics Unit at the National Institute for Health and Welfare (THL, former KTL). I would like to acknowledge professor Olli-Pekka Kallioniemi, the director of FIMM, Pekka Puska, the director of the National Institute of Health and Welfare (THL) and Anu Jalanko, the Head of Public Health Genomics unit in THL, for providing an excellent and inspiring working environment.

I would like to acknowledge the Helsinki Biomedical Graduate Program, The Sigrid Jusélius Foundation, the Emil Aaltonen Foundation, the Biomedicum Helsinki Foundation, The Research Foundation of Helsinki University Central Hospital (EVO), the Neuropromise EU project and the University of Helsinki Foundation of Medicine for providing financial support, that made this thesis possible.

Thank you for all the MS patients, who have participated in this study. I wish to thank the Finnish clinicians, who have made the study possible together with the patients: Pentti Tienari, Keijo Koivisto, Irina Elovaara, Mauri Reunanen, Ilkka Rautakorpi and late Tuula Pirttilä. In addition, I would like to acknowledge the Finnish cohorts Young Finns, Health2000, Jaakko Kaprio and the TWIN cohort, Helsinki Birth Cohort Study and all the people involved for collaboration. And thank you for all the countless clinicians, technicians, PIs and students in the collaboration efforts. The study samples for this effort have been collected from patients and controls from Finland, Sweden, Norway, Denmark, United Kingdom, United States, Belgium, France, Italy, Spain, Germany, Netherlands and Switzerland.

I sincerely wish to thank Professor Paolo Gasparini for accepting the invitation to be the opponent at my thesis defense. Professor Hannes Lohi and Professor Marju Orho-Melander are acknowledged for their review of this thesis. I sincerely value your constructive comments that have improved this work. In addition, I would like to acknowledge the honored custos, Professor Seppo Meri. I warmly thank my wonderful thesis committee that has both supported and instructed me: Iiris Hovatta, Samuli Ripatti and Maija Wolf.

The greatest thanks to my supervisors, Janna Saarela, Aarno Palotie and Leena Peltonen. When I started as a genotyping specialist at the Finnish Genome Center (FGC) in 2004, I had no idea where the road would take me. I was fortunate to be recruited to do a pro gradu with Leena and Janna, who later became my PhD thesis supervisors. Leena was a spectacular lady, whose piercing intellect, catching enthusiasm, energy and talent for collaboration was inspiring. Fashion, good looks, intellect and cutting edge science were all combined in Leena. She pushed me and others to go beyond our perceived limits and she opened the world to me. She truly was inspirational.

Janna, has been the supervisor, who has supported and encouraged me throughout the thesis. You have been the senior, PI and "science big sister", who has

gently guided my way. Your office door has always been open, and you have given support when I needed it.

I wish to thank Aarno for being my second supervisor after Leena's passing. You have been involved in the MS project and provided your time, help and effort in a very generous way. I thank you for the discussions about the weird ideas, the opportunities and doors you have opened for me and the care with which you have supervised.

I wish to thank the senior scientists and the PIs at FIMM and at the THL department: Samuli Ripatti for numerous consultations on statistics, William Hennah for the endless discussions about the future, science and genetics, Elisabet Widén, Kaisa Silader for invaluable help with laboratory and genotyping, Teppo Varilo for the work on the pedigrees, Markus Perola, Ismo Ulmanen, Marjo Kestilä and Anu Loukola. Life at the university or the institute is full of practicalities, also impossible ones. My sincerest thanks to the secretaries Sari Kivikko, Tuija Koski, Heidi Arling Tripepi, Susanna Rosas, Riitta Alatalo, Sanni Hyppönen, Riitta Koskinen and Marja Medina for all your help.

All I have learned about the lab work has been from the wonderful laboratory technicians at the department. I sincerely thank Minna Suvela, Elli Kempas, Anne Nyberg, Liisa Arala, Anne Vikman, Anu Yli-Perttula, Sirkku Ekström, Siv Knaappila and other laboratory personnel for your invaluable help. Whenever the Tecan has crashed, or the spotter has been too intimidating, you have been there to help and guide. Thank you Päivi Lahermo, Mari Kaunisto and others form the FIMM Technology center, Minttu Sauramo, Päivi Laiho and colleagues from the DNA unit and all others.

Missukat: Ansku, Eve, Anu, Suvi, and our OA member Annu. Thank you for the sauna evenings, the discussions and memorable moments. You have made this journey worthwhile.

My special thanks to the coffee break group and the Cambridge division: Anu, Anni, Annu, Ansku, Antti, Diane, Emmi, Hanski, Heidi, Helena, Himanshu, Ida, Jaakko, Jarkko, Johannes & co., Jonna, Juho, Jussi, Jonas, Kakku, Kati & Jouni, Luca, Mari, Marika, Marine, Marjis, Liisa, Olli, Peter, P-P, Pietari, Saana, Tea, Tero, Tiia, Verneri, Will and all others: thank you for the weird discussions, conference trips and wild parties. You have created an exceptional(ly hilarious) working and after-work environment. I wish all the best for all of you and wish to meet all of you again. I have had great times with you guys.

Thank you for my non-science friends: thank you tykkisiskot, extracurricular friends, Conitea friends, and all others. You keep reminding me of how much there is in life, besides work. Thank you, my parents, for your support, skills that you have given me and for the foundation. Thank you Anna, for being there, although you could not care less about science. You have given me another perspective. And finally, Johannes, thank you for being there for me and standing by me for better or worse. Way to go, our team!

CONTENTS

Abstract....6
Acknowledgements....8
Contents....10
List of original publications....12
Abbreviations....13
1 Introduction....15
2 Review of the literature....16
2.1 Complex disease genetics....16
2.1.1 Characteristics of the human genome....16
2.1.2 Gene mapping approaches....20
2.1.3 Population structure and population isolates....25
2.2 Autoimmunity and multiple sclerosis....27
2.2.1 Autoimmunity....28
2.2.2 The immune cells in autoimmunity....29
2.2.3 Multiple sclerosis – clinical characteristics....32
2.3 Genetics of complex autoimmune disorders....37
3 Aims of the study....41
4 Materials and methods....42
4.1 Study cohorts....42
4.1.1 Finnish study cohorts....42
4.1.2 International study cohorts....44
4.2 Genotyping methods....47
4.3 Statistical methods....49
4.3.1 Study I: Genome-wide association study in a Southern Ostrobothian isolate....49

4.3.2 Study II: Meta-analysis of international replication cohorts 52

4.3.3 Study III: Genetic burden analysis in families and in the Southern Ostrobothnian isolate 52

5 Results and discussion 54

5.1 Association between variants in *STAT3* and MS, Study I 54

5.1.1 Association between common haplotypes in *STAT3* and MS 58

5.1.2 Three regions of homozygosity in MS 60

5.1.3 Copy-number variants in MS patients 61

5.2 Meta-analysis of association between rs1800693 (*TNFRSF1A*), rs17445836 (*IRF8*) and rs17824933 (*CD6*) and MS, Study II 61

5.3 Accumulation of common MS associated alleles in Southern Ostrobothnia, Study III 63

5.3.1 Genetic burden score of MS associated alleles was not increased in the Southern Ostrobothian isolate 63

5.3.2 No evidence for accumulation of common MS associated alleles in Finnish multiplex families 65

6 Conclusions and future prospects 67

Web-based resources 69

References 70

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Jakkula E[1], Leppä V[1], Sulonen A-M, Varilo T, Kallio S, Kemppinen A, Purcell S, Koivisto K, Tienari P, Sumelahti M-L, Elovaara I, Pirttilä T, Reunanen M, Aromaa A, Oturai AB, Søndergaard HB, Harbo HF, Mero I-L, Gabriel SB, Mirel DB, Hauser SL, Kappos L, Polman C, De Jagger PL, Hafler DA, Daly MJ, Palotie A, Saarela J, Peltonen L (2010) Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in *STAT3* gene. American Journal of Human Genetics 86(2):285-91

II International Multiple Sclerosis Genetics Consortium[2] (2011) The genetic association of variants in *CD6*, *TNFRSF1A* and *IRF8* to multiple sclerosis: a multicenter case-control study. *PLoS ONE* 6(4):e18813

III Leppä V, Jakkula E, Ripatti S, Gourraud PA, Kaprio J, Eriksson J, Lehtimäki T, Koskinen S, Palotie A, Saarela J (2012) Genetic Burden of Common Variants in a Population Isolate and in MS families *submitted*

The publications are referred to in the text by their roman numerals.
These articles are reproduced with the kind permission of their copyright holders.

[1] The authors contributed equally to this work
[2] First consortium author and corresponding author. Consortium authors: Leppä V, Surakka I, Tienari P, Elovaara I, Compston A, Sawcer S, Robertson N, De Jager PL, Aubin C, Hafler DA, Oturai AB, Søndergaard HB, Sellebjerg F, Sørensen PS, Hemmer B, Cepok S, Winkelmann J, Wichmann H-E, Comabella M, Bustamante MF, Montalban X, Olsson T, Kockum I, Hilert J, Alfredsson L, Goris A, Dubois B, Mero I-L, Smestad C, Celius EG, Harbo HF, D'Alfonso S, Bergamaschi L, Leone M, Ristori G, Kappos L, Hauser SL, Cornu-Rebeix I, Fontaine B, Boonen S, Polman C, Palotie A, Peltonen L, Saarela J

ABBREVIATIONS

ANZgene Australia and New Zealand Multiple Sclerosis Genetics Consortium
BWH Boston Women's Hospital
EAE experimental autoimmune encephalomyelitis
CD Crohn's disease
CD6 CD6 molecule
CEU People of Northern and Western European ancestry, Utah, USA (HapMap population)
CHB Han Chinese, Beijing, China (HapMap population)
CHR chromosome
CI confidence interval
CNS central nervous system
DNA deoxyribonucleic acid
EDSS Expanded Disability Status Scale
e.g. exempli gratia
GWAS genome-wide association study
HLA human leukocyte adhesion
IBD identical by descent
IBS identical by state
IL2RA interleukin 2 receptor, alpha
IL7R interleukin 7 receptor
IMSGC International Multiple Sclerosis Genetics consortium
IRF8 interferon regulatory factor 8
JPT Japanese, Tokyo, Japan (HapMap population)
LD linkage disequilibrium
mRNA messenger ribonucleic acid
MAF minor allele frequency
MRI magnetic resonance imaging
MS multiple sclerosis
OMIM Online Mendelian Inheritance in Man
OR odds ratio
PCR polymerase chain reaction
PPMS primary progressive multiple sclerosis
RA rheumatoid arthritis
RNA ribonucleic acid
ROH region of homozygosity
RRMS relapsing - remitting multiple sclerosis
SCID severe combined immunodeficiency
SD standard deviation
SLE systemic lupus erythematosus
SNP single nucleotide polymorphism
SO Southern Ostrobothnia

SPMS secondary progressive multiple sclerosis
STAT3 Signal Transducer and Activator of Transcription 3
STAT5a signal transducer and activator of transcription 5A
STAT5b signal transducer and activator of transcription 5B
TNFRSF1A tumor necrosis factor receptor precursor
TRAPS tumor necrosis factor receptor-associated periodic syndrome
T1D type 1 diabetes
UC ulcerative colitis
YRI The Yoruba people of Ibadan, Nigeria (HapMap population)

1 INTRODUCTION

Multiple sclerosis (MS) is a chronic complex autoimmune disorder of the central nervous system. MS patients develop multiple plaques in their central nervous system, where the axons of neurons have been demyelinated (Lucchinetti et al., 2000). Symptoms can vary according to the extent and location of the demyelination. There is no cure for MS and the disease can lead to disability and sometimes death (Kurtzke, 1983, Bronnum-Hansen et al., 2004, Koch-Henriksen et al., 1998, Sadovnick et al., 1992, Sumelahti et al., 2002, Roxburgh et al., 2005). Although the age of onset for MS varies greatly, most patients receive their diagnosis of definite MS between 20 and 45 years of age (Hammond et al., 1988).

MS is considered to be an important debilitating neurological disease of the central nervous system in young adults. The prevalence of MS is approximately one in a thousand individuals in the Finnish population (Sumelahti et al., 2001), which is similar to the prevalence in other populations of Northern European descent (Bentzen et al., 2010, Benedikz et al., 2002, Dahl et al., 2004, Forbes et al., 1999, Ford et al., 1998, Grytten et al., 2006, McGuigan et al., 2004, Sarasoja et al., 2004, Sumelahti et al., 2001). In the Southern Ostrobothnia region of Finland, the prevalence and incidence of MS is approximately two-fold that of the surrounding regions (Sumelahti et al., 2000). There is evidence for a founder effect in this region (Tienari et al., 2004), which suggests the accumulation of unknown heritable factors.

MS is a complex disease: both environmental and genetic factors are thought to affect the predisposition (Ebers et al., 1986, Hansen et al., 2005, Sadovnick et al., 1996). So far, the etiology of MS is largely unknown. Only a few environmental factors have been identified to contribute to MS predisposition despite intensive research (Riise et al., 2003, Haahr et al., 1995, Kim et al., 2000, Munger et al., 2004).

The genetic studies have had more success. Linkage and association at the *HLA-DRB1* locus on chromosome 6p was first published in 1972 (Jersild et al., 1972) and the risk allele has a reasonably high odds ratio (OR 2-3) (Sawcer et al., 2011). Recent genome-wide association studies (GWAS) and the following meta-analyses have identified at least 54 loci (ANZgene, 2009, IMSGC, 2007, Jakkula et al., 2010, Jersild et al., 1972, Mero et al., 2009, Patsopoulos et al., 2011, Sanna et al., 2010).

Before this thesis project, only three MS loci were known: *HLA-DRB1*, *IL2RA* and *IL7R* (IMSGC, 2007, Jersild et al., 1972). The aim of this thesis was to identify new loci for MS, replicate recently identified MS loci, and evaluate the accumulation of the recently identified MS loci in the Finnish population, especially in the high-risk isolate and patients with a familial background for MS.

2 REVIEW OF THE LITERATURE

2.1 Complex disease genetics

The phenomenon of inheritance of traits has been known for thousands of years and has been used to breed domestic animals and cultivate plants. The studies of Gregor Mendel clarified the laws of inheritance and defined modes of inheritance for monogenic traits. However, most traits are neither monogenic nor strictly ruled by the genes. In complex traits many genetic loci act together and with environmental factors to create complex phenotypes.

The recognition of deoxyribonucleic acid (DNA) polymer as the molecular base of inheritance, and the discovery of its three dimensional structure in 1953, facilitated the birth of the field of genetics (Watson and Crick, 1953). The human genome consists of approximately 3.2 billion base pairs, packed into 22 pairs of autosomal chromosomes and one pair of sex determining chromosomes (XX or XY). One of the sister chromosomes from each pair is from the mother, and the other from the father.

At each generation, or gamete formation, one of the two sister chromosomes from each pair is selected to the gametes, and to offspring, randomly. During gamete formation, the sister chromosomes align in the first cell division and can recombine or exchange their genetic material in a crossing-over event (Figure 1). Recombination creates new combinations of variants, breaks old connections and gives rise to unique chromosomes, which creates new possibilities for genetic variation to make new phenotypes or new phenotype combinations.

2.1.1 Characteristics of the human genome

The first draft of the sequence of the human genome was launched in 2001 and the first complete version was launched in 2004 (Lander et al., 2001, Venter et al., 2001, International Human Genome Sequencing Consortium, 2004). The human genome project discovered that only 1.1-3% of the genome is protein coding with the number of protein coding genes estimated to be only around 21,000 (Clamp et al., 2007, Venter et al., 2001, Harrow et al., 2012). In addition to the protein coding sequence, the human genome contains functional elements, such as regulatory elements, structural elements and ribonucleic acid (RNA) genes, such as micro-RNA (miRNA) coding genes, and sequence of unknown functions (Lindblad-Toh et al., 2005, Siepel et al., 2005).

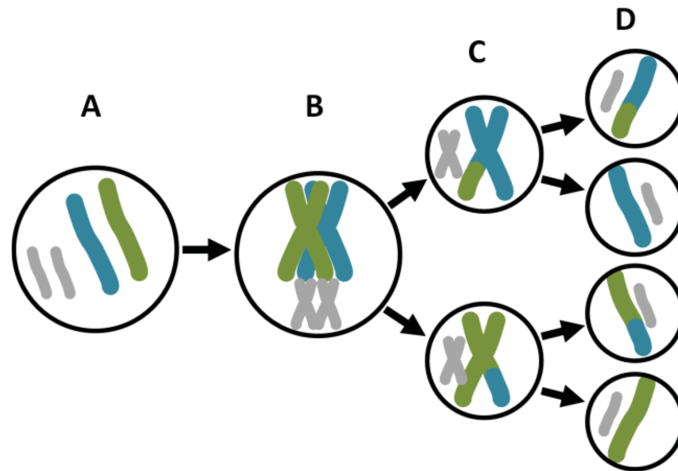

Figure 1 Recombination. The diploid nucleus of one cell (A) is divided into four haploid gametes during gametogenesis (D). B) The chromosomes first replicate and create a copy of themselves that remains attached at the centromere. Then the chromosomes align with their homologous pair during metaphase. At this time, the chromosomes can form chiasmata, or sites where the chromosomes cross over (the green and the blue chromosome), and can exchange material. C) Next, the chromosomes are divided into daughter cells. D) These cells then divide once more so that the sister chromatides are pulled into separate cells. The recombinant chromosomes (with both green and blue) carry new allele combinations not present in either of the original chromosomes.

Genetic Variation

Mutations are the source of genetic variation, and variation is the key to inherited phenotypic flexibility and alteration. The probability of a mutation per base pair per generation is $1.2\text{-}2.5\times10^{-8}$ (Nachman and Crowell, 2000, 1000 Genomes Project Consortium, 2010). This means that each individual is expected to carry approximately 30 - 60 *de novo* germline mutations, on average, compared to the parents. The number of germline mutations is affected by the age of parents (Sanders et al., 2012), and majority of *de novo* mutations originate from the father (O'Roak et al., 2012).

The genetic variation can be roughly divided into nucleotide sequence and structural variation in DNA. Single base pair polymorphisms (SNPs) represent the sequence variation, where the number of bases stays the same, but the sequence is different. SNPs are usually biallelic and they are very common in the human genome (Figure 2). On average 1 in every 1000 bases is expected to be variable in a human genome due to SNPs (1000 Genomes Project Consortium, 2010). Another source of variation is structural variation. It includes small changes in base pair compositions, such as insertion-deletion polymorphisms of one to a few bases (indels) or copy number variation (CNV) of larger segments of DNA (Figure 2). In October 2012, there were more than 38 million SNPs and indels in the dbSNP database. The structural variations have a large size scale from 1bp (indels) to millions of base pairs (large CNVs).

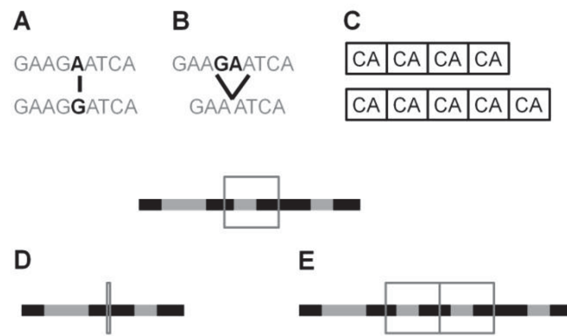

Figure 2 Genetic variation. Panel A presents a single nucleotide polymorphism (SNP), where one base pair (A) in the sequence has been mutated into another (G) in another sequence. Panel B presents an indel, where one or a few base pairs have been inserted or deleted compared to the other sequence. Panel C presents a VNTR, specifically a short microsatellite, where different sequences can have different number of short sequence repeats, here either 4 or 5 'CA' repeats. Finally, panels D and E present CNVs. Panel D presents a deletion, where a segment of DNA is missing compared to the reference sequence. Panel E presents a duplication of the same segment.

CNVs are much rarer but contribute to an additional 0.4% difference in DNA sequence between any two individuals (Redon et al., 2006, Sebat et al., 2004) (Figure 2).

The current massive parallel next generation sequencing techniques (NGS) allow the investigation of entire individual genomes at the base pair level. The 1000 genomes project uses NGS to find rare variants and it aims to discover most of the human variation with at least 1% minor allele frequency (MAF) (1000 Genomes Project Consortium, 2010). Therefore, the NGS projects will increase the number of known genetic variants substantially.

Structural repeat sequences such as the variable number of tandem repeats or VNTRs include the abundant micro- and minisatellites that are sequential repeats of 1-4 and 5 to ~ a dozen nucleotides, respectively. They can have phenotypic consequences, such as Huntington disease (Huntington's Disease Collaborative Research Group, 1993) or fragile X syndrome (Kremer et al., 1991), but are thought to be mostly neutral and have been used as genetic markers in gene mapping.

Transposons and retrotransposons are mobile elements in the genome that can create variation in the human genome by moving or copying genomic material from one location to another. In addition, they can act as sites for non-allelic homologous recombination that can cause misalignment of the genome and create, for example, CNVs. The majority of the human transposable elements are thought to be stable and inactive (Djebali et al., 2012), but some are still active and are thought to be able to move or copy DNA from one location to another. The biallelic SNPs are the most commonly used genomic markers because of their abundance, easy genotyping and stability, but the microsatellites are still used, for example, in forensic and paternity testing because of they are multiallelic and highly informative.

It is expected that most of the variation in the genome is rare (<1% MAF) and has not had the time to rise in frequency or to disappear from the population (1000 Genomes Project Consortium, 2010, Keinan and Clark, 2012). However, some

ancient variants, that are mostly expected to be neutral, and therefore are not subjected to selection, have become common over the course of human population history. An international effort, The HapMap consortium, has catalogued common SNP variation and ancestral haplotypes that are present in the human populations. Most common SNPs with >5% MAF are thought to be present in all populations, but in some occasions a variant can rise to high frequency in one population but not in others (Pickrell et al., 2009, Xing et al., 2009, Rosenberg et al., 2002). In some instances there has been selection either for or against a variant, but more often the population-specific rise in frequency is thought to be caused by drift. Two well known examples of positive selection are the variants that encode the lactase persistence (Tishkoff et al., 2007) and malaria resistance phenotypes (Ayodo et al., 2007).

Genome regulation and dynamics

The genome is a dynamic structure that is differentially expressed and regulated in different cells and tissues: approximately 22% of the genome is transcribed in each cell line, and no cell line expresses more than 57% of all the observed transcripts (Djebali et al., 2012). The expression of genes or transcription of DNA has many layers of regulation. The activity of the proteins that are involved in the transcription is regulated by external factors and environmental signals. The access of transcription factors to the regulatory regions can be altered by methylation of CpG islands or histone modifications that regulate the packaging of DNA (Birney et al., 2007). In addition, the expression of genes can be regulated in many ways during translation, for example, microRNAs (miRNA) can alter the half-life of the protein coding messenger-RNAs (Pasquinelli et al., 2000).

Variants in transcription factor binding sites, miRNA coding sites, miRNA binding sites, methylation sites, or histone modifications can affect gene and genome regulation. These variations and modifications can affect, for example, gene expression, gene-environment interactions, and genome accessibility, and they can be tissue specific (Dimas et al., 2009, Xiong et al., 2011, Du et al., 2009, Fire et al., 1998, Reinhart et al., 2000, Gervin et al., 2012, He et al., 2012). These types of variations could have more subtle effects than direct protein sequence affecting DNA-alterations. The regulatory variation in the human genome is under intensive research and many sites of tissue-specific regulatory sites have been discovered using the correlations and combinations of SNP genotypes and gene expression data (Dimas et al., 2009). For example, alleles of the SNP rs2867316 on chromosome 17 are correlated with the expression levels of *EFTUD2* in T cells, but not in fibroblasts (Dimas et al., 2009). The SNPs that correlate with gene expression are called expression quantitative trait locus (eQTL) SNPs (Dimas et al., 2009).

The Encyclopedia of DNA Elements (ENCODE) project was founded to identify all functional regions of the human genome, including all regions of transcription, transcription factor binding sites, chromatin structure, chromatin accessibility, gene expression and RNA processing, non-coding RNA species, DNA methylation and histone modification, among others (Encode Project Consortium et al., 2012). The

efforts of the ENCODE Project Consortium have revealed that at least 62% of the genome is transcribed into RNA, including non-polyadenylated RNAs (Djebali et al., 2012). The ENCODE project has also discovered, novel non-coding RNAs (Howald et al., 2012, Djebali et al., 2012), novel exons in known genes (Howald et al., 2012), potential novel protein coding genes and transcripts, novel gene isoforms, decrease in the intergenic regions, and an increasing overlap between protein coding regions, which might create a need to reconsider the concept of a gene (Djebali et al., 2012). In addition, the ENCODE project has revealed that at least 80% of the genome is involved in at least one biochemical function in at least one cell type, such as transcription, methylation, DNA regulation, histone modification or transcription factor binding (Encode Project Consortium et al., 2012). This is assumed to be the lower limit, since the study did not assess all cell types or transcription factors, and cell lines have specific active sites, active transcription factor binding sites and cell specific expression patterns (Djebali et al., 2012, Encode Project Consortium et al., 2012).

2.1.2 Gene mapping approaches

Heritability

Traits or phenotypes that have a heritable component can be subjected to gene mapping. The proportion of the genetic effect on a phenotype is usually described using heritability. Heritability is a mathematical estimate of the proportion of the total phenotypic variance of a trait in a population that can be explained by genetic factors. In humans, heritability is usually calculated from the phenotypic concordance of monozygotic (MZ) and dizygotic (DZ) twins, who share 100% or 50% of their genomes, respectively. MZ twins essentially share their entire genome, and therefore phenotypic differences between MZ twins can be expected to be largely caused by environmental effects (Merriman, 1924, Liew et al., 2005, Siemens, 1924, Galton, 1876).

Linkage analysis

Linkage analysis was the first genome-wide method for positional cloning, or gene mapping (Morton, 1955). It is a pedigree based method to map trait affecting loci. Linkage analysis follows the co-segregation of a trait and a marker in a pedigree in large chromosomal segments (Figure 3). Markers that are close to the trait affecting locus are linked with it, and appear to be co-segregated. Recombination events break the connections between the markers and the trait over generations, and this is used to narrow down the locus of interest on the chromosome. Linkage studies have been successful in mapping monogenic traits but the results in complex disease mapping have been discouraging. Most linkage results in complex diseases have not been replicated in subsequent studies. The challenges in linkage studies in complex diseases are thought to be due to multiple reasons, for example, locus heterogeneity,

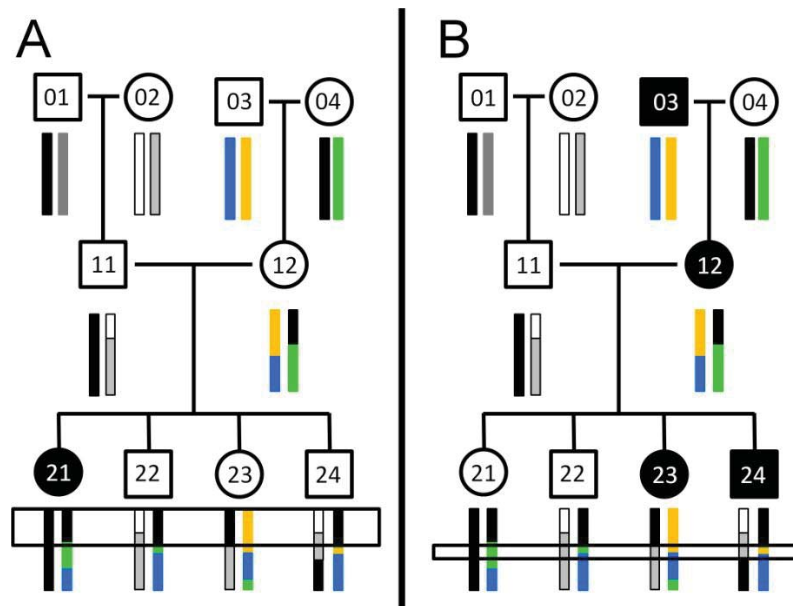

Figure 3 Linkage. Panel A represents monogenic autosomal recessive inheritance with full penetrance, where two causal alleles are required for a trait (individual 21). In panel B the inheritance pattern is monogenic autosomal dominant, and individuals carrying a segment of the yellow chromosome have the trait. Chromosomes are presented as colorful lines under each individual. Grand-parents chromosomes (01-04) are passed on to the grand children (21-24) through parents (11-12). Males are represented as squares and females as circles. Individuals with the trait are indicated with solid black shapes.

incomplete penetrance, small sample size, and no clear pattern of inheritance (Altmuller et al., 2001, Risch, 2000). It has been pointed out, that association studies and linkage studies have the power to detect very different type of connections between a phenotype and genetic variation.

Linkage disequilibrium and association studies can be used to map genes

Let's consider a mutation that has occurred on an ancestral chromosome in the proximity of another variant. If the subsequent recombination events and mutations in the following generations have not had time to separate the two markers in most genomes, the two markers appear together more often than would be expected by chance, which is to say they are non-randomly segregated in a population (Thompson and Neel, 1997). These two markers are said to be in linkage disequilibrium (LD) (Thompson and Neel, 1997, Maniatis et al., 2002). Markers that are in LD reside on the same chromosome and are usually close to one another physically (Peterson et al., 1995). LD can be used to identify ancestral haplotypes, i.e. blocks of DNA that are shared in the population and have been inherited from a common ancestor. From an evolutionary perspective, recombination gradually breaks the ancestral chromosomes into shorter segments in the subsequent generations eroding the connection between markers (International HapMap Consortium, 2005). Linkage disequilibrium can be measured using the correlation between genetic markers in a population with each haplotype recognized by a

distinctive set of marker alleles (Johnson et al., 2001, International HapMap Consortium, 2005). In a given population at a given time, there usually exist a limited number of common ancestral haplotypes (with MAF > 5%) that can be identified by genotyping common SNPs (Johnson et al., 2001, International HapMap Consortium, 2005).

The International HapMap consortium launched a project in 2003 to create a catalogue of common variation and haplotype structure within human populations. The first version of the HapMap data was released in 2005 (International HapMap Consortium, 2005). The availability of dense marker maps and LD information has facilitated the association studies, especially genome-wide association studies (GWAS) (International HapMap Consortium, 2005). Common haplotype variation in the human genome can be tagged by using from 300,000 to 1,000,000 SNP markers, depending on the desired haplotype frequencies and the extent of LD in the population of interest (Gabriel et al., 2002, Kruglyak, 1999). The haplotype map can be used to predict LD between markers in a population and to estimate haplotypes without any knowledge of parental chromosomes (International HapMap Consortium, 2005, Chapman et al., 2003). The HapMap project revealed that the length of LD differs between populations and the ancestral haplotypes are shorter in the more diverse populations of African descent. The median haplotype length in populations of European or Asian origin is expected to be ~44 kb, but only ~22 kb in African populations (Gabriel et al., 2002). This reflects the bottle necks in the non-African populations, where the recombination events have not yet had time to break down the ancestral haplotypes.

Genetic association studies

Association analysis tests the statistical dependence between one or more phenotypes and genetic polymorphisms (Gabriel et al., 2002, International HapMap Consortium, 2005). In association studies, the tested markers (usually SNPs) are often selected based on LD, to tag as much regional genetic variation as possible with as few markers, or polymorphisms, as possible (Johnson et al., 2001). The polymorphisms are then genotyped and the statistical dependence between the genotypes and the phenotype is measured statistically. Association is considered significant if the likelihood of falsely rejecting the null hypothesis is less than 5% (p-value < 0.05).

Association studies can be performed in population cohorts, by comparing unrelated cases and controls, using family-based association analyses or combined studies. Family based studies can use, for example, transmission-disequilibrium test (TDT) that can test for genetic linkage in the presence of association (Spielman et al., 1993). Combined studies can use both families and independent cases and controls (Hiekkalinna et al., 2011). In general, association studies can be divided into two categories: candidate gene studies and genome-wide association studies. Candidate gene association studies are generally based on information from previous linkage analyses or on biological function of gene products. The candidate gene region is tagged with genetic markers and analyzed for association. There have been

many candidate gene studies in complex disease genetics, but few loci have been successfully replicated in independent follow-up studies. This could be due to several reasons, for example, unsuccessful choice of candidate genes, focusing on genes and not including regulatory regions, inadequate sample size, or population stratification.

In genome-wide association analysis, genetic markers (usually SNPs) are selected to cover the entire genome in an as unbiased way as possible. In the case of GWAS from hundreds of thousands to millions of markers are tested. Current genotyping platforms use microarrays that can hold up to 4.5 million probes for SNP and copy-number variation detection, and a single array can be used to genotype up to four samples. Multiple testing is especially notable in GWAS, but it needs to be considered in all studies, including candidate gene studies.

Since multiple testing increases the risk of false positive findings, or by chance observations, of a p-value less than 0.05, the level of significance needs to be readjusted (Risch and Merikangas, 1996). Methods for correcting for multiple testing in genetic studies include, for example, Bonferroni correction that considers the number of tests, permutation methods for empirical p-values, and the Nyholt method that takes into account the LD between the markers (Nyholt, 2004). The limit for a significant uncorrected p-value has been calculated to be 5×10^{-8} in a genome-wide association analysis, assuming 1,000,000 independent association tests and 95% probability of no false positives (Risch and Merikangas, 1996).

General applications that arise from the use of LD are haplotype analysis, and genotype imputation. Haplotypes can be estimated in populations based on reference frequencies and the correlations of genetic polymorphisms (Gabriel et al., 2002, Wang et al., 2002). Analysis of haplotypes can be more informative than analyzing only single markers, especially for rare variants in LD with the tagging SNPs (Johnson et al., 2001, Browning and Browning, 2007, de Bakker et al., 2005). Imputation is a method, where the genotypes of ungenotyped markers are computationally estimated based on the LD information from variation in a reference population (Marchini et al., 2007, Servin and Stephens, 2007). Both of these approaches increase the number of genetic markers that can be assessed in the analysis.

Meta-analysis

Association study results from different populations and study sets can be combined using meta-analysis methods (Kazeem and Farrall, 2005, Mantel, 1963). The meta-analyses combine results across different studies and populations. Therefore, different meta-analysis methods exist, that differ from one another, for example, in the way they weigh the contribution of each population. These methods can generally account for population stratification, caused by differences in allele frequencies between populations (Nelis et al., 2009, WTCCC, 2007, Devlin and Roeder, 1999, Pritchard et al., 2000, Kazeem and Farrall, 2005, Mantel, 1963). Combining association studies increases the sample size, which in turn increases the statistical power to detect association signals.

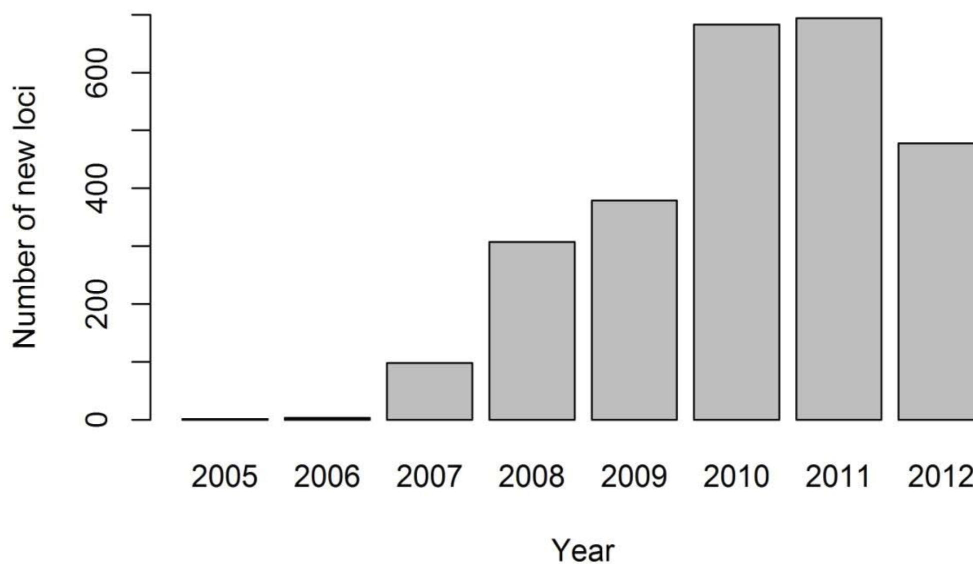

Figure 4 Number of new loci identified in GWAS studies has increased from 2005 to October 12th 2012. The data is from the Catalog of Published Genome-Wide Association Studies (https://www.genome.gov/gwastudies/).The year of publication is on the X-axis and the number of novel loci reported in that year are shown on the Y-axis.

Power

Power is the likelihood of finding a true association or difference of a defined magnitude in a data set with a certain number of samples, certain assumed allele frequencies, and desired false positive and negative rates. The power of a study depends on the number of samples, correct classification of cases and controls, the expected odds ratio (OR) or effect size, estimated penetrance, the expected frequency of the alleles, and the extent of LD in the population (WTCCC, 2007, Zeggini et al., 2008, Buyske et al., 2009). Roughly, the larger these values are, the greater the power (WTCCC, 2007). These parameters can compensate for one another, with, for example, extensive LD and higher odds ratio in theory making up for smaller sample sizes. A good example is genetic mapping in dogs (Wilbe et al., 2010). Dogs have traits and diseases that run in pedigrees or breeds, and have extensive LD due to inbreeding within breeds (Lindblad-Toh et al., 2005). Therefore, only 15,000 markers and fewer samples, compared to human studies, are needed to map genetic loci (Lindblad-Toh et al., 2005, Wilbe et al., 2010). However, it should also be noted that allelic heterogeneity can in some cases lower the power to detect associations (Slager et al., 2000).

Genome-wide association studies have been effective in finding loci with small effect sizes

The well-powered GWASs have proved to be very efficient in finding common variants with small effect sizes in complex diseases (Barrett et al., 2008, WTCCC, 2007, De Jager et al., 2009b, IMSGC, 2011, Barrett et al., 2009a, Cooper et al., 2008, Patsopoulos et al., 2011) (Figure 4). However, these individual variants explain only a fraction of the heritability of most complex traits. A large portion of the heritability remains unexplained despite dozens of identified loci (Visscher, 2008, Zeggini et al., 2008, Barrett et al., 2008, Manolio et al., 2009). Epistasis, structural variation, gene-environment interactions, rare(er) high-impact variants, epigenetics, incomplete LD with the detected markers and the causal variant, simply more unidentified common low-risk loci, and inaccurate heritability estimates have all been speculated to explain some of the missing heritability, amongst others.

So far there is some evidence for rare potentially high-impact alleles in some common diseases and new more well-powered GWAS studies have increased the number of known common low-risk loci (Pinto et al., Sanders et al., Cohen et al., 2004, Nejentsev et al., 2009, Mero et al., 2009). Epigenetics and genetic interactions are under intensive research, but they have been proven to be challenging fields of research.

Next generation sequencing techniques

The advent of NGS technologies has created new possibilities for the identification of rare and de novo variants (Sanders et al., 2012, Neale et al., Girard et al., 2011). The costs of sequencing individual genomes at reasonable accuracy and coverage have come down enough to allow large-scale sequencing projects. Sequencing techniques have already shed light on the autism spectrum disorders, schizophrenia and specific cancers (Neale et al., 2012, Sanders et al., 2012, Girard et al., 2011, Pasqualucci et al., 2011). The major challenges for NGS are the storage and analysis of the massive amounts of data that are generated during the sequencing process, alongside the analysis of the vast number of variants identified. These modern sequencing techniques have not only opened an avenue for scrutinizing individual genomes, but have also allowed the large-scale sequencing of the RNA content, for analysis of transcript level modification and regulation, through using RNA-seq method, and to identify transcription factor binding sites, by Chip-Seq (Visel et al., 2009, Ross-Innes et al., 2012, Wilhelm et al., 2008, Steidl et al., 2011).

2.1.3 Population structure and population isolates

Population history affects population genetics: it can cause differences in allele frequencies and lead to accumulation of disease causing variants and their nearby genetic markers (Bersaglieri et al., 2004, Tishkoff et al., 2007, Thompson and Neel, 1997, Varilo et al., 2003, Savukoski et al., 1994). Therefore, population history should be taken into account when selecting research methods. Drift, migration, selection, founder effect and other phenomena affect the genetic structure of all

populations. Humans have most likely dispersed from Africa to inhabit most of the planet and several subsequent migration waves and bottle neck effects have affected the genetic structure in human populations (Ramachandran et al., 2005, Tishkoff et al., 2009). Therefore, genetic variation is greatest within Africa, while populations outside Africa tend to have only a sub-group of the African variation (Conrad et al., 2006, Ramachandran et al., 2005, Tishkoff et al., 2009).

Whenever a sub-population has left from an existing habitat to colonize new regions, the new population has carried with them only an essentially random subset of the existing genetic variation (Conrad et al., 2006, Ramachandran et al., 2005, Tishkoff et al., 2009). Similarly, a population that has risen or re-risen from a small number of individuals, called founders, carries a subset of the original variation. If a small number of individuals establish a population, the population is said to have gone through a bottle neck. Human populations, in general, are estimated to have less variation than other primates, which is thought to be caused by multiple bottle necks throughout the course of human population history (Yu et al., 2004, Chimpanzee Sequencing and Analysis Consortium, 2005).

The frequencies of alleles and variants in a population that has gone through a bottle neck, or in populations that have been separated, can differ substantially from the population of origin (Price et al., 2009, Norio, 2003a). This is demonstrated in the Finnish population with the consequence of the Finnish disease heritage, which is the increase of certain diseases in Finland due to inherited factors. At least 35 globally rare known mono- or oligogenic inherited diseases are known to have significantly increased frequency in Finland (Norio, 2003b, Lahtela et al., 2010), (www.findis.org). The bottle neck effect can not only cause the increase of a certain allele, but it can also decrease the frequency of other alleles and inherited diseases. An example of a disease that has become rare in Finland is the autosomal recessive disorder phenylketonuria. The prevalence of phenylketonuria is approximately 1:26 000 of live births in Sweden (Holmgren et al., 1976), 1:17 000 in France (Abadie et al., 2001) and 1:10 000 in the UK (Hardelid et al., 2008), but only 1:100 000 in Finland (Pastinen et al., 2001). The overall unequal distribution of alleles, like the phenylketonuria allele, is called population structure. Population structure can be detected both between and within countries or populations, i.e. between people from Eastern and Western Finland or European populations (Jakkula et al., 2008, Nelis et al., 2009, Price et al., 2009).

Several populations can be said to be population isolates. Such isolated populations are often characterized by having different minor allele frequencies, reduced haplotype diversity, and increased LD, compared to the populations of origin. (Hovatta et al., 1997, Jakkula et al., 2008, Tienari et al., 2004, Service et al., 2006, Huyghe et al., 2010b). Isolated populations are often either geographically isolated, for example on islands, such as the Sardinian, Icelandic and Orkney populations, or linguistically isolated, such as the Saami (Price et al., 2009, Huyghe et al., 2010a, Huyghe et al., 2010b, Francalacci et al., 2003). In addition to geographically isolated populations, a population isolate can exist without physical or linguistic barriers, as is the case in Kuusamo and Southern Ostrobothnia (Jakkula et al., 2008, Tienari et al., 2004, Hovatta et al., 1997). The enrichment of certain rare

alleles and diseases in, for example Southern Ostrobothnia, is demonstrated by the enrichment of a rare monogenic disorder called late infantile neuronal ceroid lipofuschinosis in patients with Southern Ostrobothnian ancestry (Norio, 2003b, Santavuori et al., 1982). There are also marked differences in the frequency of certain complex phenotypes in different regions of Finland. For example, there is enrichment of autism in Central Finland compared to the overall Finnish population, whilst Southern Ostrobothnia has an increased prevalence in multiple sclerosis (MS) (Tienari et al., 2004, Kilpinen et al., 2009). It can be hypothesized, that predisposing alleles for these complex diseases could have been enriched in these regions, although the effect of environmental factors cannot be excluded.

In association studies the population structure can create spurious associations that are related to the allele frequency differences between populations (or drift) rather than the disease status. For example, if the phenotype of interest is more frequent in one sub-population, then more phenotype carriers can originate from the sub-isolate. If the population structure is not controlled, the markers that are more frequent in the high-risk sub-population can appear to be associated with the disease independent of the true effect (Pritchard et al., 2000, Clayton et al., 2005). Thoroughly matching cases and controls for population of origin, meta-analysis of different populations (or sub-populations), and other methods can be used to account for the population stratification in order to avoid false positive results (Pritchard et al., 2000, Clayton et al., 2005, Price et al., 2006).

2.2 Autoimmunity and multiple sclerosis

Humans and other multi-cellular organisms are a living environment and a resource for a number of commensal, symbiotic and pathogenic micro-organisms, such as bacteria, virus, fungi and parasites. A number of mechanisms have evolved to protect humans against harmful micro-organisms, pathogens, and to regulate the commensal and symbiotic species. Epithelia are the first line of defense that shields the human body from invasion, and a number of epithelial secretions have antimicrobial components. Physical barriers are not enough on their own, pathogens can breach them and cause infection. When a pathogen has entered the human body, the immune system is activated.

The immune system consists of a number of leucocytes that circulate the human body or guard the epi- and endothelia. Their role is to recognize and destroy potentially harmful pathogens as well as dangerously altered host self structures, such as proto-cancer cells. The immune system uses two major routes to overcome the infection: it either removes them directly from the extracellular spaces through phagocytosis, or induces controlled cell death, apoptosis, in the invaded or altered host cells (Medana et al., 2000, Yannelli et al., 1986). To avoid destruction of the healthy normal tissues (self), it is crucial for the immune system to be able to tolerate what is self and to recognize the difference between normal self and non-self or altered self structures (Nemazee and Buerki, 1989, Nemazee and Burki, 1989, Ramsdell and Fowlkes, 1990, Hodes et al., 1989, Fry et al., 1989).

2.2.1 Autoimmunity

Autoimmunity is a phenomenon where the immune system initiates an immune response against a normal, usually tolerated, self tissue. Autoimmunity can be targeted to a specific tissue or can be more general, depending on the self-structure that has been falsely recognized as foreign. In MS the immune system is thought to target the myelin around the neuronal axons in the central nervous system, leading to the destruction of this insulation, of oligodendrocytes, and damaging the axons (Lucchinetti et al., 2000, Reindl et al., 1999, Zamvil et al., 1985) Similarly, in type 1 diabetes (T1D) the tolerance to the insulin producing β-cells in pancreatic islets of Langerhans, is thought to be lost, and thus the β-cells are destroyed. Several islet-specific auto-antigens have been suggested, including glutamic acid decarboxylase (GAD) and insulin (Tisch et al., 1993, Atkinson et al., 1992, Palmer et al., 1983, Daniel et al., 1995, Sobel and Creswell, 2006). Systemic lupus erythematosus (SLE) is an example of a more general autoimmune disorder where several tissues are affected. The auto-antibodies in SLE are thought to be targeted against, for example, nuclear components such as DNA and RNA that are present in nearly all cell types, which could partly explain the systemic effect (Kamat et al., 2003, Kowal et al., 2006, Pollak, 1964). In addition to MS, T1D and SLE, several other common complex chronic autoimmune disorders exist. For example, rheumatoid arthritis (RA) is an autoimmune disorder of the joints and the skin is affected in psoriasis. The autoimmune inflammatory bowel diseases can be divided to ulcerative colitis (UC), a continuous inflammation of the colonic and rectal mucosal layers, and Crohn's disease (CD), a discontinuous transmural inflammation anywhere in the gastrointestinal tract. Each of these individual autoimmune diseases have prevalences of 0.2-3% in populations of European origin (Kappelman et al., 2007, Kurd and Gelfand, 2009, Gelfand et al., 2005, Johnson et al., 1995, Symmons et al., 2002, Bentzen et al., 2010).

Common autoimmune susceptibility -hypothesis

It has been suggested that common complex autoimmune disorders could share common factors that predispose to the loss of tolerance, or to an over-reactive immune system (Gregsen et al., 2006). According to the hygiene hypothesis, the lack of parasites and infections in modern Western societies could contribute to autoimmunity (Okada et al., 2010). It has been suggested, that persistent helminth infections, in particular, might prevent autoimmunity (Rook et al., 2008). MS, T1D, SLE, inflammatory bowel diseases, psoriasis and RA have common elements in addition to distinctive features. Some autoimmune disorders accumulate in certain families. For example, Grave’s disease, RA, MS and psoriasis have all been reported to have a tendency to accumulate in the same families (Broadley et al., 2000, McCombe et al., 1990, Midgard et al., 1996). Further, the co-occurrence of MS and inflammatory bowel diseases has been suggested, although results are currently inconsistent (Broadley et al., 2000). In addition to familial accumulation, autoimmune diseases share common features, which have lead to the common

autoimmune susceptibility hypothesis (Broadley et al., 2000, McCombe et al., 1990, Midgard et al., 1996).

Several common complex autoimmune diseases, like MS, psoriasis, SLE and inflammatory bowel diseases have a sub-type with a relapsing-remitting type of disease course (Lublin and Reingold, 1996, Petri et al., 1991, Nevitt and Hutchinson, 1996, Lapidus et al., 1998). The female to male ratio is often skewed with many autoimmune disorders being more common in females (Confavreux and Vukusic, 2006, Johnson et al., 1995, Symmons et al., 2002). However, this is not always the case as, for example, UC seems to be more common in males (Kappelman et al., 2007). Hormonal regulation has been suggested to be behind the uneven gender distribution, since the disease activity seems to be at least partially hormonally regulated. Some diseases can ease in women during pregnancy, although this could be due to other pregnancy related tolerance mechanisms (Confavreux et al., 1998, de Man et al., 2008, Aluvihare et al., 2004, Morgan et al., 2005, Raychaudhuri et al., 2003). Autoimmune diseases are chronic and there is no cure for any of them at the moment although immunosuppressants can be used in some diseases to alleviate some of the symptoms (Prisms Study Group and the University of British Columbia, 2001, Hartung et al., 2002).

2.2.2 The immune cells in autoimmunity

The innate immune system

The immune system can be divided into two major sections: innate and adaptive immunity. Both are thought to play a role in autoimmunity. The adaptive immune system is capable of creating an effective response against an extensive range of specific epitopes or antigens, whereas innate immunity is essential in activating the adaptive immunity. The innate immune system is considered to be the first active cellular response to pathogens. The innate immune cells carry germline-coded cell-surface receptors that can recognize a fixed set of non-self structures that are common in many pathogens. These receptors include, for example, the Toll-like receptors, which have a limited capability for recognition (Lemaitre et al., 1996, Rock et al., 1998). Therefore, the innate immune system has a more limited ability to detect epitopes. The cell-surface recognition receptors receive the signal that leads to activation, proliferation and differentiation of the cells, from their resting state precursor cells to their active state (Rock et al., 1998, Lemaitre et al., 1996). The major effect of innate immunity on autoimmunity is, in many cases, thought to be mediated through antigen presentation. Antigen presenting cells (APCs) are essential for the activation of the adaptive immune system (Wilde et al., 1983, Hickey and Kimura, 1988, Huitinga et al., 1995, Austyn et al., 1988). The innate immunity APCs includes cells with phagocytic activity: dendritic cells, macrophages and microglia, amongst others. Microglia are present in the central nervous system (CNS), where as the macrophages are mainly present in the periphery. Dendritic

cells, in turn, are the major APC type that presents antigens to the adaptive immune system.

Both microglia and macrophages have been observed in MS lesions, and are thus considered to be important in MS pathogenesis, having been suggested to clear the plaques from debris (Huitinga et al., 1995, Kuhlmann et al., 2002, De Simone et al., 1995, Takahashi et al., 2007, Weinger et al., 2011). In SLE, inefficient clearance of apoptotic cells by macrophages and the subsequent uptake of nuclear material by dendritic cells have been suggested to contribute to the disease pathogenesis (Baumann et al., 2002, Herrmann et al., 1998, Kenyon et al., 2011). Additionally, the depletion of macrophages has been suggested to ameliorate RA and experimental autoimmune encephalomyelitis (EAE) in rodents (Li et al., 2012a, Gerritse et al., 1996). Dendiritic cells are the major APC type. Studies of EAE, a rodent model for MS, have shown that myelin specific dendritic cells play an important role in EAE, and potentially also in the mouse model of T1D (Subramanian et al., 2001, Hoglund et al., 1999) For example, dendritic cells have been found to present neuronal antigens in the cervical lymph nodes in both EAE and MS (van Zwam et al., 2009). On the other hand, regulatory dendritic cells have been reported to be involved in suppression of autoimmune responses, especially in CD (Kwon et al., Rimoldi et al., 2005). In some autoimmune disorders local microglial production of the complement also seems to be active (Dietzschold et al., 1995, Kenyon et al., 2011, Compston et al., 1989).

The adaptive immune system

The adaptive immune system consists of lymphocytes, which can be divided into two major categories, B cells and T cells. The B cells are responsible for the humoral immune response, and they produce antibodies against specific targets. The T cells can be divided into different groups depending on their function: CD8 carrying cytotoxic T cells (CD8+ T cells), CD4 carrying T helper (Th) cells (CD4+ T), and CD4+ regulatory T cells (Tregs). The cytokine environment after antigen presentation guides the development of naïve CD4+ helper T cells into Th1, Th2 or Th17 cells (Harrington et al., 2005). All these cell types secrete a distinctive set of cytokines. The adaptive immune cells depend on their recognition receptors called B cell receptors (BCR) and T cell receptors (TCR). Each T and B cell carries a unique receptor with a unique recognition pattern that defines its activating target. One cell is thought to only respond to a specific type of molecular pattern recognized by its receptor. However, the vast amount of different cell clones, with their unique receptors, creates an immense potential for adaptation. The T and B cells have the potential to recognize self antigens through their receptors and launch a strong immune attack against the host itself. Therefore, their proliferation, maturation and activation are strictly controlled (Nemazee and Buerki, 1989, Ramsdell et al., 1989, Hodes et al., 1989). Mechanisms that induce tolerance to the self are an essential part of the development of the adaptive immune system (Hodes et al., 1989, Ramsdell et al., 1989, Nemazee and Buerki, 1989, Nemazee and Burki, 1989) However, quiescent self-reactive T and B cell clones are thought to be present in all

individuals, but on rare occasions something can trigger them and this could lead to autoimmunity.

B cells are specialized to produce antibodies, with each B cell clone producing one type of antibody that binds to one epitope. B cells produce one class of antibodies at a time (IgA, IgG, IgM or IgE) and switch antigen classes in certain sequential order. Oligoclonal antibodies against self-antigens have been found in many autoimmune diseases (Kowal et al., 2006, Pollak, 1964, Daniel et al., 1995, Palmer et al., 1983, Tisch et al., 1993, Reindl et al., 1999, Johnson et al., 2005, Rantapaa-Dahlqvist et al., 2003, Tagami et al., 1983, Aoki et al., 1989). In MS, oligoclonal immunoglobulin bands are present in the cerebrospinal fluid of patients (Delmotte and Gonsette, 1977, Johnson et al., 1977, McDonald et al., 2001). The exact nature of these antibodies is unknown, although auto-antibodies against myelin basic protein (MBP) and myelin oligodendrocyte glycoprotein (MOG) have been found in patients (Reindl et al., 1999, Zhou et al., 2006). B cell activation and antibody production usually requires an activation signal from CD4+ Th1 or Th2 cells, which, in turn, have been activated by APCs or antigen presenting B cells (Doyle and Strominger, 1987, McLaughlin and Wucherpfennig, 2008).

CD4+ T cells use their TCR and accessory molecules to recognize antigens that are presented by APC on major histocompatibility complex II (HLA-II) molecules, also called human leukocyte antigen complex II. Certain haplotypes of the *MHC-II* locus are known risk factors in MS and in many other autoimmune disorders, which makes CD4+ T cells a good candidate cell population for contributing to the autoimmune process (Asano et al., 2009, Barrett et al., 2008, WTCCC, 2007, Franke et al., 2008b, Franke et al., 2010, Han et al., 2009, Harley et al., 2008, Helms et al., 2005, Jersild et al., 1972, Nair et al., 2009, Nejentsev et al., 2009, Raychaudhuri et al., 2012, IMSGC, 2011, Silverberg et al., 2009, Todd et al., 1987, Wordsworth et al., 1989). In MS, the CD4+ T cells have been found in post mortem brain lesions and biopsies, and can be used to induce EAE in rodent models (Jones et al., 1999, Owhashi and Heber-Katz, 1988). Similar results have been seen in psoriasis (Nickoloff and Wrone-Smith, 1999).

Recent observation of different T cell sub-types in autoimmune disorders has directed attention towards Treg and interleukin-17 (IL17) producing Th17 T cell sub-populations (Langrish et al., 2005). Myelin specific Th17 cells were shown to induce EAE in mice, rather than the initially thought Th1 population (Langrish et al., 2005). Additionally, Th17 deficient or Th17 inducing IL23 deficient mice have been found to be resistant to both EAE and collagen-induced arthritis (CIA), which is a rodent model of RA (Nakae et al., 2003, Langrish et al., 2005, Cua et al., 2003, Murphy et al., 2003). MS patients have been shown to have a higher number of Th17 cells in their blood and cerebrospinal fluid, especially during exacerbation (Matusevicius et al., 1999). A recent study showed that Th17 regulating miR-326 was correlated with MS severity (Du et al., 2009).

Recent studies have emphasized the role of Treg cells in autoimmune diseases (Lepault and Gagnerault, 2000, Martinez et al., 2012, Chaudhry et al., 2009). It is thought that the Treg cells are responsible for maintaining or inducing the inactive state of the autoreactive B and T cells (Jang et al., 2011, Martinez et al., 2012),

especially the Th17 cells (Chaudhry et al., 2009). It has been suggested that Treg cells would induce anergy in the arthritogenic self-reactive T cells, and that pro-inflammatory signals in the inflamed joint would induce reduced suppressive capability in Tregs (Martinez et al., 2012, van Amelsfort et al., 2007). Tregs have been suggested to play a role as anti-autoimmunity regulators, at least in RA, MS, UC, and T1D (Lepault and Gagnerault, 2000, Martinez et al., 2012, Chaudhry et al., 2009).

Cytotoxic T cells recognize antigens presented on the major histocompatibility complex 1 (HLA-I) molecules, also called human leukocyte antigen complex I, that present intracellular pathogen antigens on all cells. Upon recognition of the MHC-I – antigen complex, they become activated and induce apoptosis in the presenting cell. CD8+ T cells have been observed to play an important role in T1D and they are also present in MS lesions (Lucchinetti et al., 2011, Babbe et al., 2000, Bulek et al., 2012). The number of CD8+ T cells and macrophages has been reported to correlate with the extent of axon damage (Kuhlmann et al., 2002). Interestingly, myelin specific CD8+ T cells have been reported to be able to induce CNS inflammation in rodents, although the type of inflammation is different from CD4+ induced EAE (Huseby et al., 2001).

Regulation of the immune system varies in different compartments of the body, which could explain some of the differences between autoimmune diseases. The gastrointestinal tract immune cells have to tolerate a multitude of non-self antigens, especially the commensal microbiota and antigens present through nutrition. Therefore, the gut immune system has evolved to induce tolerance against ingested antigens. Some studies have even suggested that probiotic supplements could be useful in managing autoimmune diseases (Kwon et al., 2010). The other extreme is the brain, which is isolated from systemic immunity by the blood brain barrier. Immune cells are not known to pass the barrier unless they are activated or during inflammation. It has been suggested that one cause leading to MS could be leaking of the blood brain barrier. A recent study found support for this hypothesis, reporting that Th17 cells could mediate the blood—brain barrier disruption and lead to CNS inflammation in MS (Kebir et al., 2007). Despite intensive research, the etiology and triggers for autoimmune reactions remain elusive and hard to disentangle.

2.2.3 Multiple sclerosis – clinical characteristics

MS is a complex autoimmune disease affecting the central nervous system (CNS). Most patients receive a definitive diagnosis of MS between the ages of 20 and 45 years of age (Hammond et al., 1988). The disease is gradually disabling and is potentially debilitating. Disease progression is estimated using the Expanded Disability Status Scale (EDSS), ranging from no disability (EDSS 0.0) to death due to MS (EDSS 10.0) (Kurtzke, 1983). The multiple sclerosis severity score (MSSS) has been developed to assess the severity of the disease, and to predict the level of disability later on (Roxburgh et al., 2005). Patients with MS have a decreased life expectancy and it has been estimated that 55-70% of patients die of MS related causes (Bronnum-Hansen et al., 2004, Koch-Henriksen et al., 1998, Sadovnick et al.,

1992, Sumelahti et al., 2002, Roxburgh et al., 2005). The severity of MS varies greatly from patient to patient.

Relapsing remitting MS (RRMS) is the most common form of MS, with approximately 67-90% of patients initially diagnosed with it (Sumelahti et al., 2002, Confavreux and Vukusic, 2006, Dahl et al., 2004, Grytten et al., 2006). The disease course in RRMS is episodic with bouts of MS, and worsening of the symptoms, followed by partial or complete recovery (Figure 5). In primary progressive MS (PPMS) the disease course is a continuous progression from the beginning, without bouts or any recovery (Figure 5). Only 9-20% of patients are diagnosed with PPMS (Sumelahti et al., 2003, Confavreux and Vukusic, 2006, Dahl et al., 2004, Grytten et al., 2006). Approximately two thirds of MS patients with RRMS will later develop secondary progressive MS (SPMS), where the initial RRMS disease type turns to steady progression (Figure 5) (Lublin and Reingold, 1996). As in other autoimmune diseases, the gender distribution in MS patients is uneven. The distribution is particularly skewed in patients with RRMS where the female to male ratio is 2-3:1, whilst the ratio in PPMS patients is closer to 1:1 (Sumelahti et al., 2003, Hammond et al., 1988, Confavreux and Vukusic, 2006, Ford et al., 1998).

Table 1. McDonald's criteria for MS diagnosis for relapsing-remitting MS (RR) and primary progressive MS (PP) based on the presence of MS attacks (relapses), lesions in the central nervous system, and on the possible presence of oligoclonal bands in the cerebrospinal fluid (CSF).

Disease type	**Clinical presentation**	**Additional Evidence Needed**
RR	≥2 attacks and objective evidence for ≥2 clinical lesions separated in time	No additional evidence needed
RR	1 attack and objective evidence for ≥2 clinical lesions separated in time	A second clinical attack or, for example, MRI[1] evidence for dissemination in time
RR	≥2 attacks and objective evidence for 1 clinical lesion	Positive evidence for CSF+[2] and evidence for ≥2 lesions in the MRI or evidence for dissemination in space in MRI
RR	1 attack and evidence for 1 clinical lesion	MRI evidence for dissemination in time or space or MRI evidence for ≥2 lesions and CSF+[2]
PP	≥1 year of insidious neurological progression suggestive of MS	Two of the following: evidence for MS in MRI in the brain, evidence for MS in MRI in the spinal cord, and CSF+[2]

[1]MRI= magnetic resonance imaging, [2]CSF+= oligoclonal immunoglobulin bands present in the cerebrospinal fluid.

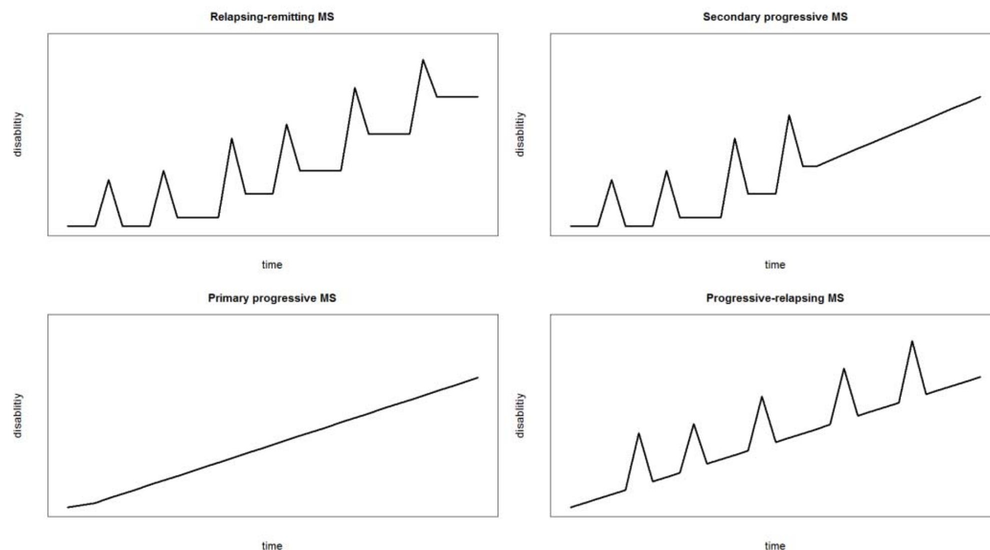

Figure 5 Types of MS progression. The four main types of disease progression in MS are divided into relapsing-remitting MS (RRMS), secondary progressive MS (SPMS), primary progressive MS (PPMS) and relapsing-progressive MS (RPMS) described according to Lublin and Reingold, 1996.

The first systematic description of MS was by Jean-Martin Charcot, who published his observations in a series of publications in 1868. Today, most patients are diagnosed using either the Poser or the McDonald criteria (McDonald et al., 2001, Poser et al., 1983) (Table 1). McDonald criteria is a revised version of the Poser criteria. Symptoms in MS can be transient and variable and include, for example, dizziness, cognitive dysfunction, spasticity and loss of vision. Therefore, receiving a correct diagnosis can take many years, especially for RRMS patients (Sumelahti et al., 2002). Current treatments for MS cannot cure or stop the progression, but can slow down the disease.

Pathogenesis

Multifocal demyelinating inflammation in the CNS is a hallmark of MS and it can be observed as plaques using magnetic resonance imaging (MRI). The plaques are visible due to loss of myelin, which forms the white matter of the brain. Oligodendroctyes are the primary cell type affected by the demyelination, since they produce the myelin sheath that is thought to be the target (Lucchinetti et al., 2000). The damage is not restricted to the oligodendrocytes, but extends to the axons (Kuhlmann et al., 2002). The damage to axons correlates with duration and course of the disease (Kuhlmann et al., 2002). It is thought that this compromises the neurons ability to conduct signals, causing ephaptic signaling or cross-talk between neighboring axons and the various symptoms of MS.

Prevalence

The prevalence of MS in populations of Northern European descent is typically around 1-2/1000) (Bentzen et al., 2010, Benedikz et al., 2002, Dahl et al., 2004, Forbes et al., 1999, Ford et al., 1998, Grytten et al., 2006, McGuigan et al., 2004, Sarasoja et al., 2004, Sumelahti et al., 2001). The prevalence varies around the world and correlates with latitude: MS is rare near the Equator and increases towards the poles (Hammond et al., 1988, McGuigan et al., 2004, Kurtzke et al., 1979). It has been suggested that the gradient is due to environmental factors, such as sun exposure or pathogens. Interestingly, however, the Sami populations that inhabit northernmost Europe, have been reported to have exceptionally low prevalence of MS (Gray et al., 2008). There have been reports of increasing prevalence and incidence of MS in recent years, especially in women (Debouverie et al., 2007, Granieri et al., 2007, Gray et al., 2008, Grimaldi et al., 2001, Sarasoja et al., 2004, Bentzen et al., 2010). This has been speculated to be caused by increased survival, more accurate diagnosis and changes in the hormonal environment in women, for example, oral contraceptives and older age at first child birth (Granieri et al., 2007, Sarasoja et al., 2004, Bentzen et al., 2010).

The prevalence varies not only according to latitude, but also within regions and countries (Tienari et al., 2004, Rothwell and Charlton, 1998, Bostrom et al., 2009, Visser et al., 2012). Certain isolated populations are known to have a notably increased prevalence of MS compared to the surrounding populations. The Sardinian island in Italy, the Orkney Islands off the coast of Scotland (Visser et al., 2012), the Lothian Region of Scotland (Rothwell and Charlton, 1998), Värmland in Sweden (Bostrom et al., 2009) and Southern Ostrobothnia in Finland (Tienari et al., 2004) are all such populations. It has been speculated, that specific high-risk MS predisposing genetic variants could have been enriched in these regions.

Risk factors for MS

The etiology of MS remains elusive despite decades of intensive research. At present, only a few factors are known to influence MS predisposition. Migration studies support the importance of environment in the development of MS (Hammond et al., 2000, Ahlgren et al., Dean and Elian, 1997). It has been suggested that child migrants acquire the average risk of the population in the new region, whereas, the risk of adult migrants would be closer to the population of origin (Dean and Elian, 1997). However, there have been conflicting results (Hammond et al., 2000).

Pathogens have been suggested to play a role in MS (Kurtzke and Heltberg, 2001), but only Epstein-Barr virus and human herpes virus 6 (HHV-6) have shown promising results, so far. Epstein-Barr virus, or mononucleosis, has been associated with MS in several studies (Haahr et al., 1995, Ascherio et al., 2001, Marrie et al., 2000). The presence of antibodies against the Epstein-Barr virus has been suggested to correlate with disease activity (Farrell et al., 2009). Additionally, the HHV-6 is suspected to either affect or exacerbate MS (Kim et al., 2000, Alvarez-Lafuente et al., 2006). The precise role of these infectious agents has not yet been resolved.

Exposure to sun, especially in childhood, has been shown to decrease the risk of MS (van der Mei et al., 2001, Islam et al., 2007, Freedman et al., 2000). This beneficial effect has been hypothesized to be linked to both increased vitamin D levels and the Treg promoting effects of ultraviolet radiation. While a number of studies report the protective effect for vitamin D (Munger et al., 2004, Munger et al., 2006), others have suggested that the effects of sun exposure and vitamin D could in fact be independent (Lucas et al., 2011). Recent genetic studies of MS have found associations to genes that are involved in vitamin D metabolism (Ramagopalan et al., 2011, IMSGC, 2011), but the link between vitamin D, sun exposure and MS still needs further studies before definite conclusions can be made.

Some studies report that smoking increases the risk of MS, but the increase in risk is small (relative risk 1.7-1.8) (Riise et al., 2003). However, the subtype of MS and the extent of disability correlate with smoking. Smokers are more likely to have PPMS than people who have never smoked (Healy et al., 2009).

Genetic factors contribute to MS predisposition

Several twin studies in MS have suggested that genetic factors contribute to the disease predisposition. The concordance between monozygotic twins has ranged from 13 to 34%, whereas the concordance of dizygoic twins has been closer to 2-5% (Ebers et al., 1986, Hansen et al., 2005, Islam et al., 2006, Mumford et al., 1994, Thorpe et al., 1994). The concordance between siblings is similar to that of the dizygotic twins, ~3% (Sadovnick et al., 1996). The prevalence of both twinning and MS are low, which means that the number of individuals in these studies has been low, and therefore the confidence intervals are large. The estimated heritability (H^2) from twin studies ranges from 0.25 to 0.76 (Hawkes and Macgregor, 2009). In general, the risk of MS is increased in close relatives of MS patients, varying proportionally to the fraction of shared genes (Sadovnick et al., 1996, Ebers et al., 1995).

Half-sib, family and adoption studies in populations of European descent also suggest that inherited factors may contribute to MS predisposition (Ebers et al., 2000, Ebers et al., 1995, Sadovnick et al., 1996). The risk of MS in adopted individuals with affected adoptive family members is similar to the population level (Ebers et al., 1995). Multigenerational families with multiple affected family members are rare in MS. In Finland and Southern Ostrobothnia there is an exceptional level of clustering of MS in families (Wikström et al., 1984), which could indicate the enrichment of rare MS predisposing genes in the region and in these families.

2.3 Genetics of complex autoimmune disorders

Major common complex autoimmune disorders such as RA, psoriasis, T1D, MS, inflammatory bowel diseases and SLE have been intensively studies for decades. These diseases have moderate heritability and there have been many linkage and association studies that have aimed to identify predisposing genetic factors. Until the advent of GWAS, the only established locus in many autoimmune diseases was the *HLA*-locus (Table 2). This has strengthened the view that these diseases are immune-mediated, because the HLA molecules are essential in antigen presentation and in priming adaptive immune responses. In many cases, the *HLA* risk haplotypes are responsible for the strongest known increase in risk. Due to the structure and tight LD across the *HLA* region, the linkage and association signals can cover long haplotypes with alleles from multiple *HLA* genes (de Bakker et al., 2006, Horton et al., 2008).

Many candidate gene studies in autoimmune disorders have been unsuccessful. However, at least one exception exists, the insulin gene in T1D (Bell et al., 1984, Hakonarson et al., 2007). One major issue in many candidate gene studies has been a limited sample size and therefore an inadequate power to detect association. In contrast to candidate studies, GWASs with large sample sizes (usually > 1000 samples) and the following meta-analyses have been very successful in identifying associated loci with modest odds ratios (Barrett et al., 2009a, Barrett et al., 2009b, Barrett et al., 2008, Okada et al., 2012, Stahl et al., 2010, Patsopoulos et al., 2011, IMSGC, 2011).

Table 2. *The established* HLA *loci that have been linked or associated with RA, MS, UC, CD, psoriasis and T1D.*

Disease	HLA locus	Reference
Rheumatoid arthritis	*HLA-DRB1, HLA-B, HLA-DPB1*	Wordsworth et al., 1989, Raychaudhuri et al., 2012
Multiple sclerosis	*HLA-DRB1, HLA-A*	Jersild et al., 1972, IMSGC, 2011
Ulcerative colitis	*HLA-DRB1- HLA-DQB1- HLA-DRA*	Franke et al., 2008b, Silverberg et al., 2009
Crohn's disease	*HLA-DQA, HLA-DRA2*	Franke et al., 2010, Barrett et al., 2008
Psoriasis	*HLA-C*	Helms et al., 2005, Nair et al., 2009
Type 1 Diabetes	*HLA-DRB1- HLA-DQB1- HLA-DQA1 HLA-A, HLA-B*	Todd et al., 1987, WTCCC, 2007, Nejentsev et al., 2007, Asano et al., 2009
SLE	*HLA-DRB1-HLA-DQB1-HLA-DRA*	Graham et al., 2009, Han et al., 2009, Harley et al., 2008

With the emergence of associated loci in GWAS, it has become clear that many of the associated loci in different autoimmune diseases are located near to one another or are even at the same loci (Table 3). The effects of the associated alleles of SNPs can be opposite, but the loci remain the same (Jakkula et al., 2010, Barrett et al., 2008, Cotsapas et al., 2011, Ramos et al., 2011). This has added to the discussion about the common autoimmune hypothesis (Cotsapas et al., 2011, Forabosco et al., 2009, Ramos et al., 2011). There have been genetic studies about autoimmune disease predisposing loci, but more comprehensive studies with larger sample sizes are needed to identify polymorphisms and pathways that can shed light on the biology underlying the common autoimmune phenotype (Cotsapas et al., 2011, Ramos et al., 2011).

So far, 192 genes have been associated at a genome-wide significant level across seven autoimmune diseases: RA, psoriasis, SLE, MS, T1D, and the two inflammatory bowel diseases (CD and UC). Forty of these loci are associated with multiple autoimmune disorders: 25 genes are associated with two autoimmune diseases, 10 with three, 5 with four, and one with five (Table 3). Furthermore, there is suggestive evidence for some of the involved loci in the remaining autoimmune disorders. Most of the genes in these loci have functions in the immune system. However, since neither the exact causative variants nor the affected gene(s) are definitely known, one should be cautious when interpreting the association results. Once analyzed in greater detail it may be found that some of the signals in these regions originate from different loci for the different diseases, whilst others are from the same loci (Cotsapas et al., 2011). It would seem reasonable to speculate that the list of associated genes is likely to grow in the near future and more overlap will undoubtedly emerge.

The SNPs in the five loci that are associated with at least four autoimmune disorders are located in or near the following genes: the protein tyrosine phosphatase, non-receptor type 22 gene (*PTPN22*) in the first locus, the tumor necrosis factor, the alpha-induced protein 3 (*TNFAIP3*) and oligodendrocyte transcription factor 3 (*OLIG3*) genes in the second locus, the T-cell activation RhoGTPase activating protein gene (*TAGAP*) in the third locus, the v-rel reticuloendtheliosis viral oncogene homolog (*REL*) and the pseudouridylate synthase 10 genes (*PUS10*) in the fourth locus, and the ORM1-like 3 (S. cerevisiae) (*ORMDL3*) and the IKAROS family zinc finger 3 (Aiolos) genes (*IKZF3*) in the fifth locus. All five loci contain genes that regulate immune functions and are plausible candidates for autoimmune disorders. SNPs in the *PTPN22* locus are associated with RA, CD, SLE and T1D. The protein encoded by *PTPN22* has been suggested to be involved in T cell, B cell and dendritic cell hyper-responsiveness and is a negative regulator of T cell activation (Zhang et al., 2011, Begovich et al., 2004). The *TNFAIP3-OLIG3* locus is associated with psoriasis, RA, MS, SLE and possibly with CD. *TNFAIP3* regulates TNF-mediated apoptosis and immune signaling, and its expression changes in response to the inflammatory cytokine TNFA. Little is known about human *OLIG3*, but the mouse ortholog is expressed in the developing neural tube (Takebayashi et al., 2002). The other members of this protein family, Olig1 and Olig 2, both promote oligodendrocyte differentiation in mice. The last locus contains

two genes *ORMDL3,* which negatively regulates sphingolipid synthesis, and *IKZF3*, a B lymphocyte development regulating transcription factor. This locus is associated with RA, CD, UC, T1D, and potentially with MS. *TAGAP* is associated with RA, MS, T1D, and UC, and it is thought to regulate T cell activation (Mao et al., 2004). *REL* is associated with psoriasis, RA, UC, and CD, and is thought to be involved in the development of regulatory T cells (Ruan et al., 2009). The *TYK2* gene is associated with psoriasis, MS, SLE, and T1D (Table 3). Mutations in *TYK2* can cause hyper immunoglobulin-E syndrome (HIES) and can potentially accelerate Th2 development (Minegishi et al., 2006). Many of the associated genes have been suggested to be involved in common functions, such as Th17 cell development or function (e.g. *STAT3, IL23R, IL7R, ETS1, CCR6, IL21*) (Chaudhry et al., 2009, Langrish et al., 2005, McGeachy et al., 2009, Liu et al., 2010, Du et al., 2009, Yamazaki et al., 2008, Nurieva et al., 2007), macrophage (e.g. *IL-23A*, *CD40*) (Piskin et al., 2006, Becker et al., 2002) or B cell functions (eg. *CD40*) (Gerritse et al., 1996), dendritic cell functions (eg. *MHC-II*, *IL23A*) (Piskin et al., 2006), and in Treg cell functions (e.g. *REL*, *STAT3, CCR6*, *IL2RA*, *CTLA4*, *IL10*) (Ruan et al., 2009, Chaudhry et al., 2009, Yamazaki et al., 2008, Asano et al., 1996, Fallarino et al., 2003, Asseman et al., 1999).

Of the 192 autoimmune disorder associated genes, more than 130 genes associate only with one specific autoimmune disorder. These include, for example, the insulin gene in T1D, the gene encoding for matrix metalloprotease 27 in psoriasis, and the gene encoding for extracellular matrix protein 1 in UC. However, many of the distinctive loci have common roles in immunology, such as the T cell activating *CD86* receptor in MS, and the T cell chemotactic *CCL21* in RA. Due to the fact that all autoimmune diseases share currently 26-55% of their loci with at least one other autoimmune disease, and that the diseases have common features, it is likely that some of the genetic predisposition could be caused by common autoimmune susceptibility. On the other hand, at least half of the associated loci in each disease are unique to one disorder. Some of these loci may be found in other autoimmune disorders in the future, but a substantial part of the association signals seem to be distinct. Therefore it is unlikely, that the susceptibility to autoimmune disorders would be entirely a result of a common autoimmune susceptibility that is directed towards different epitopes. In MS, 36 of the 54 non-*HLA* loci are unique to the disease, whilst only 18 loci are associated with at least one other autoimmune disease (Mero et al., 2009, Patsopoulos et al., 2011, IMSGC, 2011).

Many of the associated genes have known mutations that cause severe monogenic immunological disorders. For example, mutations in *IL2RA* can cause immunodeficiency (Sharfe et al., 1997) (OMIM #606367), mutations in *IL7R* cause severe combined immunodeficiency (SCID) (Puel et al., 1998), mutations in either *STAT3* or *TYK2* lead to hyper immunoglobulin-E syndrome (HIES) (OMIM #147060) (Minegishi et al., 2006, Minegishi et al., 2007, Holland et al., 2007), and mutations in *TNFRSF1A* cause familial Hibernian fever (OMIM# 142680) (Aksentijevich et al., 2001). One can hypothesize, that these genes could prove to be a good starting point for further functional studies also in complex autoimmune diseases.

Table 3. *List of chromosomal positions and genes in or near loci that have been associated with at least three autoimmune diseases.*

Position	Genes	Disease	Refrences
6q23.3	*TNFAIP3, OLIG3*	PS, RA, MS, SLE, CD	Nair et al., 2009, Thomson et al., 2007, IMSGC, 2011, Graham et al., 2008, Franke et al., 2010
1p13.2	*PTPN22*	RA, CD, SLE, T1D	Begovich et al., 2004, Kyogoku et al., 2004, Gateva et al., 2009, Barrett et al., 2008, Todd et al., 2007
2p16.1	*REL, PUS10*	PS, RA, UC, CD	Strange et al., 2010, Gregersen et al., 2009, McGovern et al., 2010, Festen et al., 2011
6q25.3	*TAGAP*	RA, MS, T1D, UC	Raychaudhuri et al., 2009, IMSGC, 2011, Smyth et al., 2008, Festen et al., 2011
17q12	*ORMDL3, IKZF3*	RA, CD, UC, T1D	Kurreeman et al., 2012, Barrett et al., 2008, McGovern et al., 2010, Barrett et al., 2009b
19p13.2	*TYK2*	PS, MS, SLE, T1D	Strange et al., 2010, Mero et al., 2009, Sigurdsson et al., 2005, Wallace et al. 2010
1p31.3	*IL23R*	PS, CD, UC	Cargill et al., 2007, Duerr et al., 2006, Franke et al., 2008b
1p36.32	*MMEL1, TNFRSF14*	RA, MS, SLE	Raychaudhuri et al., 2008, Ban et al., 2010, Kurreeman et al., 2012
1q32.1	*IL10*	UC, SLE, T1D	Franke et al., 2008b, Gateva et al., 2009, Barrett et al., 2009a
1q32.1	*KIF21B*	CD, UC, MS	Barrett et al., 2008, McGovern et al., 2010, Barrett et al., 2009b, IMSGC, 2011
5q33.3	*IL12B*	PS, CD, MS	Cargill et al., 2007, Parkes et al., 2007, IMSGC, 2011
6q21	*PRDM1, ATG5*	CD, RA, SLE	Barrett et al., 2008, Raychaudhuri et al., 2009, Gateva et al., 2009
10p15.1	*IL2RA*	RA, MS, T1D	Stahl et al., 2010, IMSGC, 2007, Vella et al., 2005
17q21.2	*STAT3*	CD, UC, MS	Barrett et al., 2008, Franke et al., 2008a, Jakkula et al., 2010
18p11.2	*PTPN2*	RA, CD, T1D	Okada et al., 2012, Parkes et al., 2007, Todd et al., 2007
20q13.12	*CD40*	RA, MS, SLE	Raychaudhuri et al., 2008, ANZgene, 2009, Vazgiourakis et al., 2011

3 AIMS OF THE STUDY

The availability of comprehensive and affordable genome-wide methods, together with large international collaborative efforts, have increased the potential to identify new variants that are associated with complex phenotypes by increasing the sample size and power to detect loci with modest odds ratios. The overall objective of this thesis was to use SNP-based methods to identify, assess and compare genetic factors contributing towards MS predisposition in the Finnish and international populations. Our specific aims were:

I To utilize the specific characteristics of the Southern Ostrobothnian case control cohort, such as an increased LD and a higher disease incidence, to identify new MS associated loci by using genome-wide SNP based, CNV and homozygozity analyses (Study I).

II To perform a replication meta-analysis study in 11 populations of European descent for three risk markers that were previously identified in an international genome-wide meta-analysis study (Study II).

III To investigate the role of common MS predisposing variants in the Finnish high-risk Southern Ostrobothnian isolate population, and in Finnish family samples (Study III).

4 MATERIALS AND METHODS

4.1 Study cohorts

4.1.1 Finnish study cohorts

Finnish multiple sclerosis samples and anonymous population controls

The Finnish multiple sclerosis study was approved by the Helsinki University Hospital ethics committee of ophthalmology, otorhinolaryngology, neurology and neurosurgery (approval 192/E9/02). Altogether 792 patient samples and 12 families with 139 additional family members were included in this study. The MS patients and their families were recruited from the University Hospitals of Helsinki, Kuopio, Oulu, and Tampere, and the Seinäjoki Central Hospital. The MS cases were diagnosed according to Poser criteria for clinically definite multiple sclerosis. All recruited individuals provided written informed consent. The 1,077 anonymous Finnish population controls were random visitors to either, the Seinäjoki Central Hospital, the University Hospital of Kuopio, or the University Hospital of Helsinki. 71% of all available Finnish MS patients were female. Information on additional phenotypic parameters was available for 434 patients (Table 4). 9.3% of the patients had a diagnosis of PPMS, of which 46% were female. In the anonymous population sample 52% were female.

Helsinki Birth Cohort Study cohort

The Helsinki Birth Cohort Study (HBCS) sample is a part of the IDEFIX-study (Identifying Early Factors In Syndrome X); a cohort study of growth, cardiovascular diseases and living conditions, and psychological outcomes. The subjects were recruited based on their year and place of birth (Helsinki, Finland, 1934-1944). All participants gave written informed consent, and the study was approved by the Ethics Committee of Epidemiology and Public Health at Helsinki University Hospital and by the Helsinki and Uusimaa Hospital District (approvals 140/E3/2004, 344/E3/2000 and 334/12/03/00/08). The sample set used in this study contained 1,574 subjects after quality control.

Table 4. *Characteristics of the Finnish MS patients.*

	Mean	Median	Skew	SD
Age at onset (yrs)	30.36	30	0.422	8.59
Duration (yrs)	14.16	13	0.773	9.08
EDSS[1]	4.22	4	0.167	2.55

[1]*Expanded Disability Status Scale (Kurtzke, 1983)*

The Cardiovascular Risk in Young Finns Study cohort

The Cardiovascular Risk in Young Finns study is a cross-sectional follow-up cohort of Finns born between 1962 and 1977 (Raitakari et al., 2008). The subjects were recruited from Helsinki, Kuopio, Oulu, Tampere, and Turku (Raitakari et al., 2008). The cohort included in this study consisted of 2,333 individuals. The institutional ethics committees of the Universities of Helsinki, Kuopio, Oulu, Tampere, and Turku all approved the Young Finns study (Raitakari et al., 2008). Informed consent was obtained from all participants or from their guardians in the case of subjects under the age of 18 years (Raitakari et al., 2008). The Young Finns samples were included as controls in Study III.

Health 2000 GenMets Study cohort

The Health 2000 GenMets is a study of metabolic syndrome in the Finnish population (Perttila et al., 2009). The 2,194 subjects for the general Health 2000 study were recruited from 80 regions across Finland. All participants gave their written consent and the Health 2000 study was approved by the Ethics Committees of the National Institute for Health and Welfare and the Ethics Committee for Research in Epidemiology and Public Health at the Hospital District of Helsinki and Uusimaa (Perttila et al., 2009). The GenMets study sample is a subset of Health 2000, which were selected either as metabolic syndrome cases or matched controls (Perttila et al., 2009). We included a subset of 104 individuals from the GenMets study based on their place of birth and included as controls in Studies I and III.

The Finnish Twin Cohort Study cohort

The Finnish Twin Cohort Study is a longitudinal population-based sample set of twins. It has been collected for the study of genetic and environmental risk factors in chronic disorders. The twins were from across Finland and had been recruited using the Central Population Registry (Kaprio et al., 2002). A subset of 1,492 samples from the cohort, one twin per pair, was included as controls in Study III. The samples had been selected for a genome-wide study of either smoking or alcohol use. The study has been approved by the Ethics Committee of Epidemiology and Public Health at the University Helsinki Hospital and the Helsinki and Uusimaa Hospital District and the Institute Review Board of Indiana University. All participants provided informed consent.

4.1.2 International study cohorts

Belgian cohort

The Belgian study was approved by the Ethics committee of the University of Leuven (approval number ML4733). The Belgian cohort consisted of 776 patients that had been recruited through two centers: the National Multiple Sclerosis Center Melsbroek and the University Hospital Leuven (UZLeuven). All patients had MS either according to the Poser or the McDonald criteria (McDonald et al., 2001, Poser et al., 1983). In addition, the sample set included 1,021 unrelated controls from Belgium. All participants gave informed consent. The cohort was included in Study II.

Danish cohort

The Danish Research Ethics Committee approved the Danish study (approval KF 01314 009). All participants gave informed consent. The 650 patient samples were recruited from The Danish MS Center in Copenhagen, Denmark and had multiple sclerosis according to the revised McDonald's criteria (Polman et al., 2005, McDonald et al., 2001). The Danish sample included 1,099 healthy Danish control samples. The cohort was included in Studies I and II.

French trio cohort

The French Ministry of Research approved the French MS study (approvals DC-2008-539 and AC-2008-548). The cohort consisted of 608 trios, an affected child and their parents, recruited through a national media campaign. All patients fulfilled the Poser and revised McDonald criteria for MS (Poser et al., 1983, McDonald et al., 2001, Polman et al., 2005, Goodkin et al., 1991). All participants gave informed consent. The cohort was included in Study II.

German cohort

The German study was approved by the Ethics committees of the Universities of Marburg, Düsseldorf, and Munich (approval 1856/07). The German cohort included 930 patients with MS according to either Poser or McDonald criteria, and 911 controls (Poser et al., 1983, McDonald et al., 2001). All study subjects gave their written informed consent. The cohort was included in Study II.

Italian cohort

The Comitato Etico (CE) Interaziendale in Novara, Italy, approved the Italian study (approval 570/CE, N. CE 38/0). The cohort included 629 patients with MS, according to the Poser criteria, and 828 regionally matched population controls. Informed consent was obtained from all patients (Poser et al., 1983). The cohort was included in Study II.

Norwegian cohort

The Norwegian cohort consisted of 662 samples with MS, according to either the Poser or the McDonald criteria, and 1,027 healthy controls. The patients were recruited through neurological departments in Norway and the healthy individuals were randomly recruited through the Norwegian Bone Marrow Donor Registry. The study was accepted by the Ethics Committee South Eastern Norway (approval S-08234a) and Local Data Inspectorate at Oslo University Hospital (2476). The cohort was included in Studies I and II.

Cohorts from United States of America and United Kingdom

The Partners Healthcare Institutional Review Board approved the USA and United Kingdom study (approval 2002-p-000434). The cohorts consisted of 656 MS cases and 714 controls from the UK, and 644 MS cases and 587 controls from the USA. The cases were diagnosed using McDonald criteria for multiple sclerosis (McDonald et al., 2001). All participants provided informed consent. The cohorts were included in Study II.

Spanish cohort

The cohort included 501 cases and 501 controls from Spain. All cases were diagnosed with clinically definite MS according to Poser's criteria. Control samples were anonymous samples from the Hospital Universitari Vall d´Hebron blood bank. All subjects gave written informed consent. The Spanish study was approved by the Ethics Committee of the Institut Català, Hospital Vall d'Hebron (approval PR(AG)30/2007). The cohort was included in Study II.

Swedish cohort

The Swedish samples were recruited from clinics all over Sweden. The cohort consisted of 2,016 patients with multiple sclerosis, according to either McDonald's or Poser's criteria (McDonald et al., 2001, Poser et al., 1983), and 1,723 controls. All patients gave informed consent. The ethics committees of Regionala Etikprövningsnämnde, Stockholm, Sweden and Karolinska Institutets Regionala Forskningskomitté approved the study (approvals 04-252/1-4; 2006/845-31/1 and 00-052, 04-375; 02-548). The cohort was included in Study II.

Gene MSA cohorts

The GeneMSA cohorts were from the United States of America (USA), Switzerland and the Netherlands (Baranzini et al., 2009b). The Dutch cohort consisted of 230 MS cases and 232 controls, the Swiss cohort of 253 cases and 208 controls and the USA cohort had 486 cases and 431 controls. The MS cases were recruited and collected at Vrije Universiteit Medical Center (Netherlands), University Hospital Basel (Switzerland), and University of California San Francisco (USA). All cases were

diagnosed using the new International Panel criteria for RRMS and the criteria for PPMS (McDonald et al., 2001, Polman et al., 2005, Baranzini et al., 2009). All participants gave informed consent prior to participation to the study and the studies were accepted by local Committees for Human Research. The cohorts were included in Studies I and II.

Boston Women's Hospital cohort, USA

The Boston Women's Hospital (BWH) cohort included 860 MS cases and 1,720 controls (De Jager et al., 2009b). All patients, and part of the healthy controls (spouses and friends of MS patients), were recruited at the Partners MS Center in Boston, Massachusetts, USA. The remaining controls were from the MIGen study, and were either healthy individuals or had a history of early myocardial infarction (De Jager et al., 2009b). The cases were matched to two controls from a pool of 2,951 subjects using EIGENSTRAT first principal component distance (De Jager et al., 2009b). All cases were diagnosed using the McDonald's criteria for MS (De Jager et al., 2009b). All individuals provided informed consent. This cohort was included in Studies I and II.

International Multiple Sclerosis Genetics Consortium cohort

The IMSGC study was reviewed by the Multi-Centre Research Ethics Committee (MREC) in the United Kingdom, and at the University of California Santa Fé and BWH Institutional Review Boards in the United States (IMSGC, 2007). All recruited individuals provided written informed consent. The MS cases from the United Kingdom were recruited from across the country, with the controls being the 2,950 shared healthy controls from the Wellcome Trust Case Control Consortium (WTCCC) study (IMSGC, 2007, WTCCC, 2007). The cases from the United States were collected at multiple specialized clinical sites across the country, while the 1,679 controls were healthy subjects from the National Institute of Mental Health cohort (IMSGC, 2007, De Jager et al., 2009b). The cases were diagnosed according to McDonald's criteria (Polman et al., 2005, McDonald et al., 2001, IMSGC, 2007). The cohort was included in Studies I and II.

4.2 Genotyping methods

Genome-wide genotyping

The Finnish samples were genotyped in two phases. First, 72 MS cases from Southern Ostrobothnia were genotyped using Illumina HumanHap310 gene chip (Illumina Inc., 5200 Research Place, San Diego, CA, 92122 USA) at the Broad Institute of MIT and Harvard's Center for Genotyping (Cambridge, MA, USA). The Health 2000 GenMets sample was genotyped at DeCode Genetics (Reykjavik, Iceland) using the Illumina HumanHap610-quad chip (Illumina Inc., 5200 Research Place, San Diego, CA, 92122 USA) (Perttila et al., 2009). Only 297,343 SNPs that were common to both the Illumina HumanHap 310 and HumanHap610-quad platforms were kept for the analysis. Quality control was performed using PLINK. All SNPs with less than 95% of genotypes were excluded from the analysis. Individuals with <95% of genotypes, gender discrepancies or cryptic relatedness were excluded. Further, all SNPs with Hardy-Weinberg equilibrium test p<0.001 were flagged as low quality markers for future analyses, but not removed.

In the second genome-wide genotyping set, 651 cases and 139 additional family members from 12 families were genotyped using the Illumina HumaHap670 custom chip (Illumina Inc., 5200 Research Place, San Diego, CA, 92122 USA) at the Wellcome Trust Sanger Institute (Cambridge, UK). The Finnish population sample cohorts Young Finns and HBCS were genotyped at the same institute using the same genotyping chip. Eight samples from the families were re-genotyped using the Illumina HumanHap660W-quad chip (Illumina Inc., 5200 Research Place, San Diego, CA, 92122 USA) at the Technology Center of the Institute for Molecular Medicine FIMM (Helsinki, Finland). The genotypes were filtered using the following thresholds: MAF >1%, <5% missing genotypes, and Hardy-Weinberg $p > 10^{-6}$ in the Young Finns cohort. Individuals were excluded if they had >5% missing genotypes, excess or low heterozygosity compared to the rest of the strata, shared >20% by IBD with another sample, or showed cryptic relatedness. Samples with discrepancies between the reported and the observed gender were removed.

The international sample sets BWH, IMSGC, and GeneMSA were genotyped at local institutes. The BWH was genotyped at the Broad Institute's Center for Genotyping (Cambridge, MA, USA) using the Affymetrix Genechip 6.0 platform (Affymetrix Inc., 3420 Central Expressway, Santa Clara, CA 95951, USA)(De Jager et al., 2009b). The IMSGC sample sets were genotyped using the Affymetrix GeneChip Human Mapping 500K Array set (Affymetrix Inc., 3420 Central Expressway, Santa Clara, CA 95951, USA) at the Broad Institute of MIT and Harvard (Cambridge, CA, USA) and the controls were genotyped at the Affymetrix Services Lab (West Sacramento, CA, USA) (IMSGC, 2007, WTCCC, 2007). The GeneMSA sample sets were genotyped at the Illumina facilities using Sentix® HumanHap550 BeadChip (Illumina Inc., 5200 Research Place, San Diego, CA, 92122 USA) (Baranzini et al., 2009). These international sample sets were imputed by De Jager et al. 2009 using MACH version 1.0.5 and phased using the Utah residents with ancestry from northern and western Europe (CEU) population in

HapMap release 21 as a reference (De Jager et al., 2009b). The imputed genotypes were filtered using standard quality metrics: SNPs with <5% missing genotypes, >1% MAF, and Hardy-Weinberg p> 10^{-6} were included (De Jager et al., 2009b).

Sequenom® genotyping in Studies I and II

Sequenom® iPLEX and iPLEX Gold MassARRAY® systems were used to genotype the samples for Studies I and II. The primers were designed using the MassARRAY® AssayDesigner (Sequenom®, 3595 John Hopkins Court, San Diego, CA 92121-1331). The default settings were used, except for the following modification: mass range from 4,300 to 8,500, and a maximum iPLEX size of 40 SNPs. The primers were ordered from Metabion (Metabion GmbH, Lena-Christ-strasse 44, 82152 Martinsried, Germany). The iPLEX and iPLEX Gold protocol provided by the manufacturer (Sequenom®, 3595 John Hopkins Court, San Diego, CA 92121-1331) was used for genotyping. The reagents were purchased from Sequenom (Sequenom®, 3595 John Hopkins Court, San Diego, CA 92121-1331). Genotypes were inspected visually using MassARRAY® Analyzer. All samples with <50% of genotypes per iPLEX were excluded. All markers with water contamination, <90% success rate, or discrepant duplicate results were excluded. The genotyping was performed in the Institute for Molecular Medicine Finland (FIMM) Technology Centre. In Study II the genotyping was performed in the participating laboratories, except for the Danish and Norwegian cohorts that were genotyped in the Institute for Molecular Medicine (FIMM) together with the Finnish cohort. All genotypes were combined to a single file, together with the TaqMan genotypes for the quality control and analysis steps.

Taqman genotyping in Study II

The SNP rs1800693 was genotyped using the c__2645714_10 TaqMan® assay (Life Technologies Corporation, 5791 Van Allen Way, Carlsbad, CA 92008, USA) and the TaqMan® end point genotyping protocol by Applied Biosystems (Life Technologies Corporation, 5791 Van Allen Way, Carlsbad, CA 92008, USA). The PCR was optimized to 35 cycles and the annealing temperature was set to 63ºC. An ABI 7900HT instrument was used for genotyping, and to visually analyze the genotype clustering. The genotyping was performed in the individual participating laboratories using either TaqMan or Sequenom genotyping, except for the Danish and Norwegian cohorts, that were genotyped together with the Finnish cohort in the Institute for Molecular Medicine Finland (FIMM). All genotypes were combined to a single file for the quality control and analysis steps.

Copy-number variation genotyping in Study I

The QuantiSNP program was used to genotype CNVs in the Study I sample sets using the defaults for the following: GC correction, Bayesian factor > 10 and length ≥ 3 SNPs. All genotype calls were verified visually using Bead Studio 3.3 (Illumina

Inc., 5200 Research Place, San Diego, CA, 92122 USA) with CNVs in the centromeric regions being excluded. CNVs that had undergone replications were genotyped using either a multiplex fragment PCR (CNVs near the *DLG2* and *NRG3* genes), or a single fragment PCR (CNV near the *ERBB4* gene). PCR reactions were performed in a 10µl reaction volume with final concentrations of 1x GeneAmp® PCR Gold buffer (Life Technologies, Carlsbad, California, USA), 1.50 mM $MgCl_2$ (Qiagen N.V., Spoorstraat 50, 5911 KJ Venlo, Netherlands), 200µM dNTPs, 0.55U AmpliTaq Gold® DNA ploymerase (Life Technologies, Carlsbad, California, USA), 300 nM of each forward primer, and 200nM of the labeled reverse primer. Primer sequences were ordered from Sigma-Aldrich (Sigma-Aldrich, 3050 Spruce Street, St. Louis, Missouri 63103, USA):

DLG2, labeled forward primer FAM-CAACTGCAATTTCCTTCTGGA,
DLG2, reverse primer 1 AGAGTAGAGGCAAGGCAGCA,
DLG2, reverse primer 2 TTTGAAGGGCAGTTTGCAC,
ERBB4, labeled forward primer VIC-GTAAGTCTTGCCCGAAGCTG,
ERBB4, reverse primer 1 GGAGGTGGGTGTATTTGTTCC,
ERBB4, reverse primer 2 (deletion) TGTGAGAACAGGCCTTGGA,
NRG3, labeled forward primer NED-GGGGAAATGATTGTGGTTCA,
NRG3, reverse primer 1 (deletion) AAATGCCTGGATCAAACCAA,
NRG3, reverse primer 2 AGGGGTGTGGAGGATATAGGA.

The PCR program included an initial denaturation step (12 minutes, 94°C), 30 cycles of amplification (30 s in 94°C, 15s in 61°C, 45 s in 72°C), and a final extension step (10minutes in 72°C). The PCR products were diluted in Hi-Di™ Formamide (Life Technologies, Carlsbad, California, USA) and run on the Applied Biosystems ABI 3730xl DNA Analyzer (Life Technologies, Carlsbad, California, USA) with the GeneScan™ -500 LIZ® standard (Life Technologies, Carlsbad, California, USA). The alleles were called using the GeneMapper 4.0 software (Life Technologies, Carlsbad, California, USA). The CNV genotyping was performed at the Institute for Molecular Medicine Finland FIMM Technology center.

4.3 Statistical methods

4.3.1 Study I: Genome-wide association study in a Southern Ostrobothian isolate

In order to control for cryptic relatedness in the 72 Southern Ostrobothnian cases and 2,194 GenMets samples, identity-by-descent (IBD) analysis was performed using PLINK (Purcell et al., 2007). First, a pruned genotype set was obtained using the PLINK nearest neighbor clustering method (pair wise population concordance = 0.05). Next, a multidimensional scaling (MDS) analysis was used to assess the clustering, or potential discrepancies in reported ancestry (Figure 6). The MDS analysis was used to collapse the IBD matrix data (Purcell et al., 2007). MDS results for the first and second dimension were plotted using Microsoft MS Office Excel.

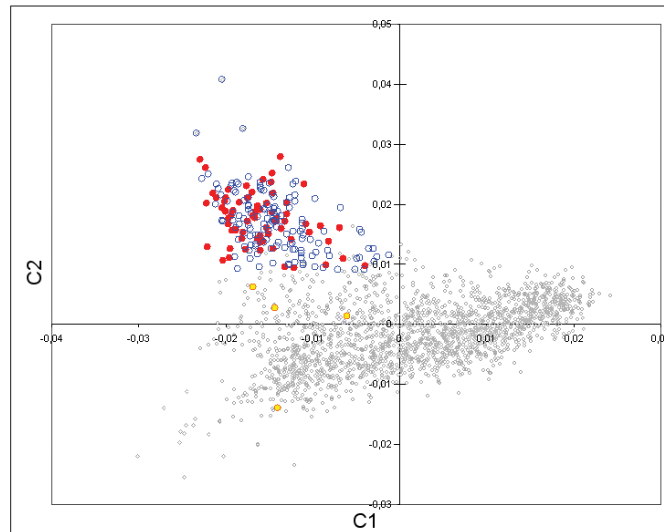

Figure 6 Multi-dimensional scaling MDS plot. The first two dimensions from the MDS analysis of the 72 cases and 2,194 population samples from the GWA study are shown. The MS samples that were included in the analysis are marked with red dots and the IBD selected controls are marked with blue dots. MS cases and population samples that were excluded based on this analysis are marked with yellow and grey dots, respectively.

The distances between cases and controls in the first two dimensions are presented in Figure 3. Outlying cases were removed based on a visual inspection of the graph. The IBD matrix was then used to select the two closest controls for each case, by using the clustering option in PLINK. After quality control (QC) and clustering, 68 of the 72 cases and 136 controls from the GenMets sample set remained in the analysis. The genomic inflation factor for the IBS matched sample set was 1.08 compared to the 1.57 for the non-matched sample set.

First, the sample set was analyzed for long homozygous stretches in the genome, in order to find potential large shared recessive risk haplotypes by using PLINK (Purcell et al., 2007). The minimum length was set to at least 50 consecutive SNPs and 500 kb per individual, per locus. The samples were then inspected for overlapping homozygous segments. The most promising overlapping homozygous segments were analyzed for statistical significance using permutation. The case-control labels were permuted 10,000 times to obtain an empirical p-value, to correct for the number of independent tests in the ROH analysis. As the PLINK software does not analyze the haplotypes within the homozygous regions. Therefore, the overlapping homozygous regions with an empirical $p< 10^{-3}$ were assessed for haplotype consistency.

Next, standard association test and association tests with dominant, recessive, and additive models were used to identify the most interesting loci with a $p < 10^{-4}$. These loci were then validated in a Finnish replication set of 83 cases and 365 controls from Southern Ostrobothnia, and 628 cases and 668 controls from elsewhere in Finland. These samples were analyzed as separate clusters using a Cochran-Mantel-Haensel (CMH) meta-analysis, in order to control for population

stratification. The two loci that passed the threshold of $p<0.05$ were analyzed in the international sample sets described in Table 5. The analysis was performed using CMH analysis and treating each sample set and nationality as a separate cluster. The *STAT3* locus haplotype structure was analyzed, to assess the haplotype background of the associated variant. The haploblock length and its constituent haplotypes were estimated using Haploview 4.0 in HapMap populations (Barrett et al., 2005). The information of the CEU population in HapMap release 23a was used to estimate the haplotype block structure using the Gabriel method (Barrett et al., 2005, Gabriel et al., 2002). The same haplotype was estimated in the YRI and CHB populations, as well. Three tagging SNPs from the rs744166 containing haplotype block were selected for haplotype analysis using the default parameters for pair-wise tagging in Haploview 4.0 and they were verified visually from the CEU haplotypes (Barrett et al., 2005). Haplotypes were estimated for all individuals from data sets with available SNP genotypes using PLINK. The CMH meta-analysis of the haplotypes was conducted, and each population was treated as a separate cluster so as to control for population stratification.

Since no systematic studies of CNVs in MS have been reported, our aim was to identify CNVs that could affect MS predisposition. The CNVs were genotyped using the intensity data from the microarray and the QuantiSNP program. Those CNVs that were observed in the 68 MS cases were analyzed using the Ingenuity® Pathway Analysis. The UCSF Genome Browser March 2006 release (NCBI36/hg18) was used to search for genes within the CNVs. All genes were analyzed for interactions and common functions with the Ingenuity Pathway Analysis software (Ingenuity® Systems, Redwood City, CA, USA) using the default settings.

Table 5. *List of international replication sample sets included in the genome wide association study in Study I.*

Country of origin	**Sample set**	**Number of MS cases**	**Number of controls**
Norway	NO	607	816
Denmark	DK	628	1074
Netherlands	GeneMSA NL	230	232
Switzerland	GeneMSA CH	253	208
United States	GeneMSA US	486	431
United Kingdom	IMSGC UK	453	2950
United States	IMSGC US	342	1679
United States	BWH	860	1720
Total		*3859*	*9110*

4.3.2 Study II: Meta-analysis of international replication cohorts

The meta-analysis method by Kazeem and Farrall, for a combined analysis of trios and case-control data, was used in study II (Kazeem and Farrall, 2005). We analyzed the cohorts presented in Table 6 as separate clusters, to control for differences in population structures. Additionally, the Finnish sample sets were split into two clusters to separate out the Southern Ostrobothnia isolate from other Finnish samples, resulting in a total of 12 clusters. Analysis was performed in R 2.9.0 by using the formulas provided in the meta-analysis article by Kazeem and Farrall 2005. The method weights each cluster taking into account both the sample size and the effect size.

4.3.3 Study III: Genetic burden analysis in families and in the Southern Ostrobothnian isolate

We used a genetic burden analysis based on a previously published weighted log-additive score (De Jager et al., 2009a, Gourraud et al.). We adapted the score to include 50 non-*HLA* SNPs that have been previously reported to associate with MS, at a genome-wide significant level, and the *HLA-DRB1**1501 allele tagging the SNP rs9271366 (de Bakker et al., 2006, IMSGC, IMSGC, 2007). In the CEU HapMap population (release 27), the SNP rs9271366 is in full LD (r^2 1, D’ 1) with rs3135388, which in turn is in high LD (r^2 0.966, D’ 0.993) with the *HLA-DRB1**1501 allele (de Bakker et al., 2006, IMSGC, 2007). We used the SNPs, risk alleles, and ORs reported in the IMSGC 2011 paper to calculate the genetic burden score, since it covered all of the genome-wide significant loci up to that date and were genotyped using the same platform as our cohort (IMSGC). For the rs9271366 SNP we used the OR reported for rs3135388 (OR 1.99) in the first multiple sclerosis GWAS that used trio samples (IMSGC, 2007).

Table 6. *List of cohorts in the meta-analysis in Study II.*

Sample set	Number of trios	Number of MS cases	Number of controls
Belgium (BE)	-	776	1021
Denmark (DK)	-	634	1090
Finland (FI)	-	792	1077
France (FR)	608	0	0
Germany (DE)	-	930	911
Italy (IT)	-	828	629
Norway (NO)	-	662	1027
Spain (ES)	-	501	501
Sweden (SE)	-	2016	1723
United Kingdom (UK)	-	656	714
United States (US)	-	644	587
Total	*608*	*8439*	*9280*

Genetic burden scores were calculated to each sample individually. R 2.15.1 was used to analyze the genetic burden score data (R Core Team, 2012). The distribution of the genetic burden scores were both drawn as histograms, and evaluated using the Shapiro-Wilkins test for normality. Although the distributions appeared normal, or did not differ significantly from normal distribution in the Shapiro–Wilkins test, statistical differences between sample groups were assessed using the non-parametric Kolmogorov-Smirnov test. Interactions between the affection status and region of origin against the genetic burden score were calculated using both an additive and a multiplicative regression models.

DSS Researcher's toolkit power and sample size analyses were used to estimate power together with simulations in R 1.15.1. The observed number of samples, genetic burden score averages, and standard deviations in each sample group were used when possible. The familial cases and Southern Ostrobothnian samples were kept at the observed constant and the values for the other group were altered according to simulation or calculation. The simulations for the power in the familial samples versus sporadic samples were calculated assuming normal distribution. For group one, the values were as follows: number of samples 63, average 6.8 and standard deviation 0.64. For group 2 the input values were as follows: number of samples 522, standard deviation 0.62 and average was either 6.6 or 6.53, depending on the assumed difference between the populations in the simulation. There were 1,000 to 10,000 simulation rounds during which random values were drawn from two normal distributions. Each rounds averages, KS-test, Welch t-test, and the average1-average2 difference were calculated and those values were stored in a table, and the stored values were observed to estimate the power and chance of a false negative finding.

5 RESULTS AND DISCUSSION

5.1 Association between variants in *STAT3* and MS, Study I

Genetic variants can be enriched in isolate populations and can cause unusually high prevalence of a phenotype. Such a phenomenon is profound for monogenic phenotypes, such as the diseases of the Finnish disease heritage. It has been hypothesized that a similar phenomenon could be involved in complex diseases, although the effect of genetics on disease predisposition is complex. The Southern Ostrobothnian population history supports the prospect of a founder effect in the region, and there is also a two-fold increase in prevalence and incidence of MS in the area (Sumelahti et al., 2000, Sumelahti et al., 2001, Tienari et al., 2004). Therefore, it is reasonable to hypothesize that MS predisposing variants could exist in this isolated region.

The 72 MS cases from Southern Ostrobothnia, and the 2,194 population samples from the GenMets cohort were initially included in our GWAS analysis. All samples were analyzed for IBD sharing, with no excess sharing being observed. Multidimensional scaling (MDS) analysis revealed that the Southern Ostrobothnian samples formed a distinct cluster from the other Finnish samples (Figure 3), which supports the isolate nature of the region. The first two clustering dimensions were then used to estimate the origin of individual samples, and 4 of the 72 cases were excluded from the initial isolate sample (Figure 3) (Jakkula et al., 2008). In total, 68 cases were included in the final analyses. Next, PLINK genome-wide clustering options were used to select the two closest controls for the remaining cases, 136 controls in total, from the available pool of 2,194 population selected samples (Purcell et al., 2007). The places of birth were then verified for the IBD matched controls with all samples originating from Southern Ostrobothnia.

Association between *STAT3* SNP rs744166 and MS in Finnish samples

Association analyses were performed using PLINK. The genomic inflation factor was found to be reasonable with $\lambda = 1.078$. The *HLA*-locus was the only locus that reached a genome-wide significant association ($p=1.35\times10^{-10}$ for rs3135338) (Figure 7). In total, 37 SNPs in 27 loci displayed at least modest association with MS (uncorrected $p < 10^{-4}$). A minimum of one SNP per locus was included in the first follow up step. A total of 28 SNPs were genotyped in an independent Finnish sample set of 83 cases and 365 controls from Southern Ostrobothnia, and 628 cases and 664 controls from elsewhere in Finland. Of the 27 loci, three displayed modest association with MS: rs3135338 in the *HLA*-region ($p=1.6\times10^{-25}$, OR=3.43), rs744166 in the first intron of the signal transducer and activator of transcription 3 (*STAT3*) gene on chromosome 17q21.2 (p=0.0012, OR=1.27), and rs1364194 on

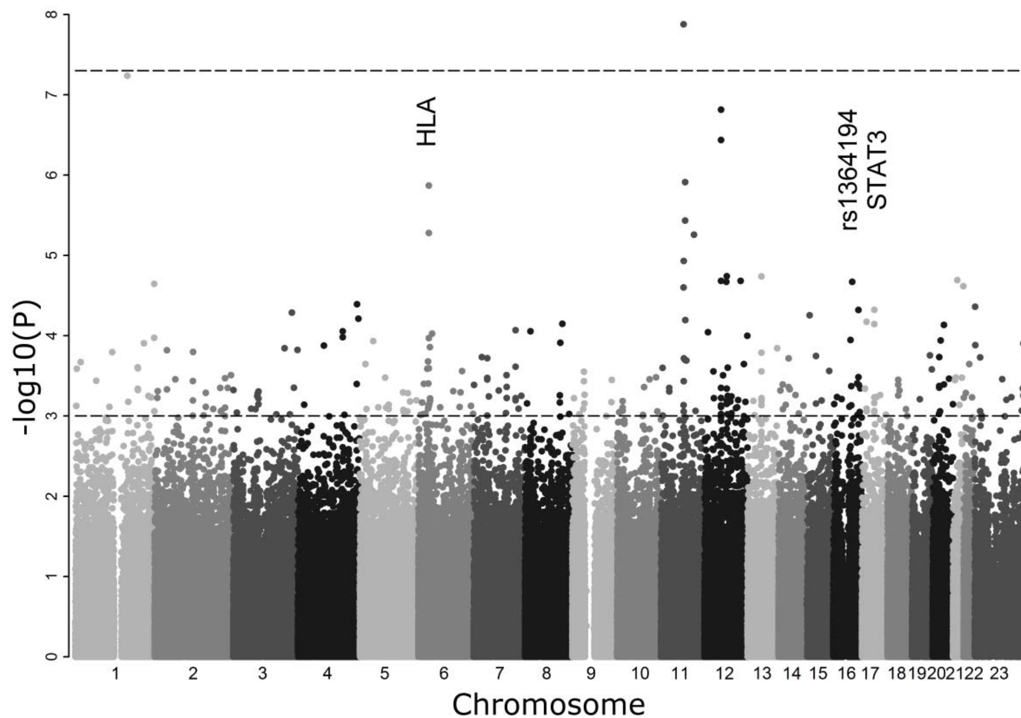

Figure 7 Manhattan plot of the GWA results. The loci that passed the first replication step are indicated in the figure: the *HLA* locus on chr 6p (including rs3135338), the rs1364194 on chr 16q and the rs744166 in the *STAT3* gene on chr 17q. The other regions were excluded, since the results could not be replicated in the Finnish sample set.

chromosome 16q21 (p=0.0047, OR=1.48). The closest genes to rs1364194 were 1 Mb away.

Next, we aimed to assess the significance of the two promising initial results, the associations between rs744166 and rs1364194 and MS, in a large international set of independent samples. The well known HLA locus was excluded from the analysis. Altogether, 3,859 MS cases and 9,110 controls from six populations and eight sample sets were analyzed using CMH meta-analysis (see Material and methods, section 4.3.1, Table 5). The genotypes for the two SNPs had been imputed in the IMSGC data sets, but directly genotyped in the other sets. No significant heterogeneity of odds ratios was observed, using the Breslow-Day test for heterogeneity, and the CMH analysis of all independent replication and validation sets confirmed the association with rs744166 in the *STAT3* gene ($p=2.75\times10^{-10}$, OR=1.18), but not with rs1364194, on chr16 (p=0.0793) (Figure 8).

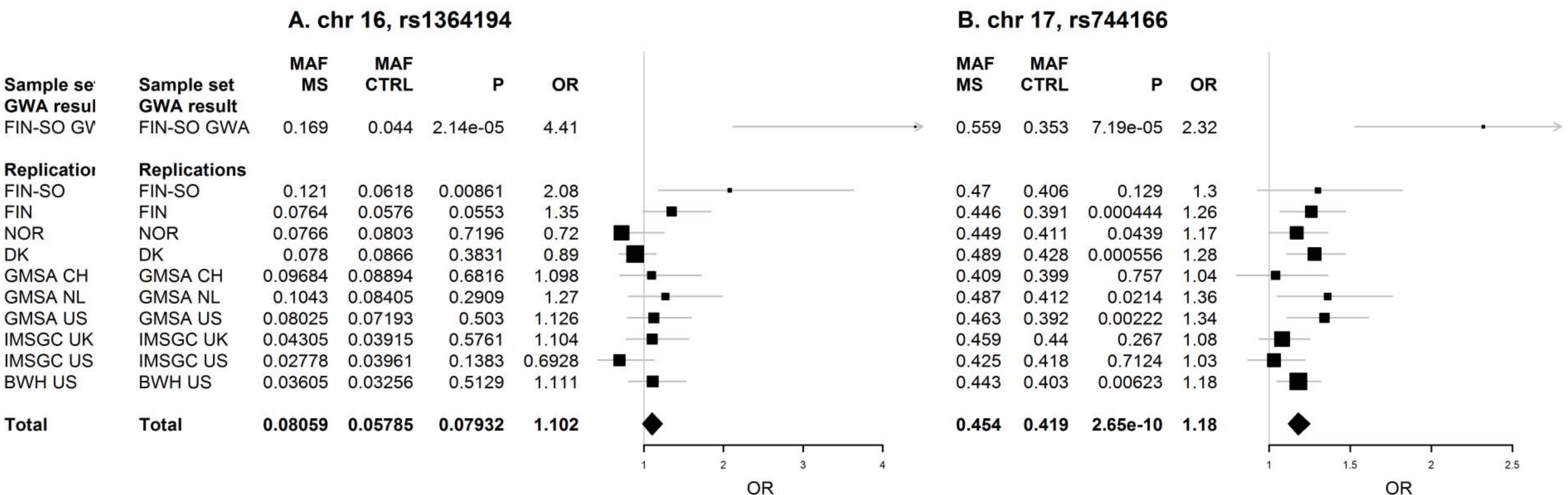

Figure 8 Results for the two follow-up SNPs from the GWAS. The figure shows the replication results for the two SNPs that were followed up after replication step I. The results are shown for the original GWAS, that was not included in the replication analysis (GWA result, FIN-SO GWA), for all individual replication cohorts (under the section Replications), and for the combined replication sets (Total). Each study sample is shown on its own line and the order of the sample sets is shown in the column Sample set. Part A shows the result for the chromosome 16 SNP rs1364149, while part B shows the results for the SNP rs744166 (in STAT3) on chromosome 17. Minor allele frequencies are shown for both cases (MAF MS) and controls (MAF CTRL). Individual p-values are listed in column P, and the odds ratios are indicated in column OR for each sample set. In addition, the odds ratios with the 95% confidence intervals (CI) are shown in the forest plot that follows the OR column. The size of the box in the Forest plot is relative to the number of samples in each study set: sample sets with larger boxes have more samples.

The associated SNP rs3135338 is located in the *HLA-II DRB1*-locus and is in modest LD with SNP rs3135388 (r^2=0.51, D'=1.0) in the CEU population (HapMap release 27) (International HapMap Consortium, 2003). The rs3135388 SNP is, additionally, in tight LD with the *HLA-DRB1*1501* allele (r^2=0.966, D'=0.993) (de Bakker et al., 2006), which has been consistently linked and associated with MS (Olerup and Hillert, 1991, Sawcer et al., 1996, Tienari et al., 1993, IMSGC, 2007, Jersild et al., 1972, Patsopoulos et al., 2011, Jakkula et al., 2010, IMSGC, 2011, ANZgene, 2009). The *HLA* region has strong LD and a complex structure (de Bakker et al., 2006, Horton et al., 2008) and only recent well-powered studies have been able to refine association signals in the region (IMSGC, 2011). Today, at least four different *HLA*-alleles have been associated with MS: *HLA-DRB1*1501*, *HLA-DRB1*0301*, *HLA-A*0201* and *HLA-DRB1*1303* (IMSGC, 2011, Fogdell-Hahn et al., 2000, Harbo et al., 2004). Some of the alleles are in strong LD with other loci in the *HLA* region, and therefore, their effects cannot be separated at this stage (IMSGC, 2011).

The SNP rs2293152 in the *STAT3* gene, was suggestively associated with MS in a previously published meta-analysis, but the association was not replicated (De Jager et al., 2009b). The same group recently extended their previous meta-analysis, and reported that rs2293152 is indeed associated with MS (Patsopoulos et al., 2011). However, the cohorts in the meta-analyses significantly overlapped with each other and with those of our study (De Jager et al., 2009b, Jakkula et al., 2010, Patsopoulos et al., 2011). The Gene MSA, IMSGC and parts of the BWH cohort are included in all three studies, and further, all the samples from the first meta-analysis were included in the follow up step of our study (Patsopoulos et al., 2011, De Jager et al., 2009b, Jakkula et al., 2010). The associated SNP in the meta-analyses is only in moderate LD (r^2=0.352, D'=0.861) with rs744166 in the CEU population (HapMap release 27) (International HapMap Consortium, 2005).

A Spanish study with 1,540 cases and 1,720 controls did not find evidence for association between rs744166 in *STAT3* and MS (Cenit et al., 2010). However, two more recent studies detected an association with MS at the *STAT3* locus (IMSGC, 2011, Lill et al., 2012). An independent German study reported a weak association with rs744166 (p=0.012, OR=1.09) in a reasonably large German cohort with 2,932 MS cases and 2,972 controls (Lill et al., 2012). Additionally, an international GWAS reported a strong association with rs9891119 ($p=1.80\times10^{-10}$, OR=1.11), which is in LD with rs744166 (r^2=0.65, D'=1.0) in the CEU population (HapMap release 27) (IMSGC, 2011). It should be noted, that the international GWAS is not completely independent of the current study (IMSGC, 2011). At least some of the samples are overlapping with the Finnish, Norwegian and Danish sample sets, however, the overlapping cohorts represent a minority of the total samples analyzed (maximum of 1,520 of 9,772 cases and 121 of 16,849 controls overlap) (IMSGC, 2011). The associated SNP rs9891119 resides in the same LD block as rs744166 and is in LD with it (r^2=0.69, D'=1.0) in the CEU population (HapMap release 27). It is therefore possible, that both association signals may point towards the same effective variant or variants.

The association between rs1364194 on chromosome 16 and MS was not replicated in the international replication meta-analysis. The chromosome 16 SNP

could represent a false positive initial finding, and is unlikely to be relevant in MS predisposition in general. However, we cannot exclude the option, that it might be relevant in the Southern Ostrobothnian isolate population. Due to lack of plausible candidate genes or functional elements, we cannot speculate its potential function or functional relevance.

5.1.1 Association between common haplotypes in *STAT3* and MS

We analyzed the haplotype structure around rs744166 in the HapMap 2 populations (release 23a) using Haploview (Barrett et al., 2005, International HapMap Consortium, 2003). Using the Gabriel et al. block definition method, the rs744166 was found to reside in a LD-block that covers the beginning of *STAT3* from the immediate promoter region to exon 4 (Barrett et al., 2005). In the CEU population the haplotype block is 40 kb long and, using the HapMap 2 release SNP resolution, five different haplotype backgrounds for the block could be defined. The rs744166 A-allele was present on only one haplotype in the Utah residents with ancestry from Northern and Western Europe, USA, (CEU), the Japanese from Tokyo, Japan, (JPT), and the Han Chinese from Beijing, China, (CHB) HapMap populations, but on four different haplotypes in the Yoruba from Ibadan, Nigeria, (YRI) population.

We used the Haploview haplotype tagger to select three additional SNPs to cover all of the variation in the CEU population. One SNP failed in both genotyping and the QC phase, leaving a total of three SNPs, including rs744166, for the haplotype analysis. In addition to the Finnish samples, the genotypes for the three SNPs were available in six other cohorts, (from the US, the UK, Netherlands, and Switzerland). In total, 3,377 MS cased and 8,295 controls were included in the haplotype analysis using PLINK. Association analysis was performed for each sample set, and the combined data set using CMH meta-analysis (Purcell et al., 2007). The two most common haplotypes were associated with MS (Table 7). The analysis could not discern if there was an associated and a reciprocal haplotype, or if both haplotypes were truly associated with MS. Interestingly, the protective A-allele that tags the protective haplotype in populations outside of Africa was common in the CEU population (frequency 0.56), the CHB population (frequency 0.65), and the JPT population (frequency 0.57), whilst it was rarer in the YRI population (frequency 0.25) (HapMap release 27). The haplotype distributions were similar in all studied populations, including Southern Ostrobothnia and elsewhere in Finland.

Pickrell *et al.* searched for regions under selection in the human genome. The SNP rs744166 was among the analyzed SNPs and its allele distribution varied throughout the worlds populations (Pickrell et al., 2009). However, rs744166 did not show significant evidence for positive selection in the study by Pickrell *et al*., 2009, although its frequency has increased in the populations outside of Africa (Pickrell et al., 2009).

The *STAT3*, *STAT5A* and *STAT5B* genes each encode transcription factors of the STAT-family, whose members are involved in the immune system functions, and

Table 7. *The two most common haplotypes and their frequencies in the* STAT3 *locus were associated with MS in the haplotype analysis of 3,377 cases and 8,295 controls from five populations.*

Haplotype by SNP alleles of rs6503695, rs744166, rs957970	Frequency in MS cases	Frequency in controls	OR	P-value
CGG	0.308	0.276	1.17	7.51×10^{-7}
TAA	0.549	0.582	0.87	4.26×10^{-6}
TGG	0.086	0.081	1.07	0.181
CGA	0.057	0.061	0.93	0.169

regulation of cell growth and apoptosis (Yao et al., 2006, Imataki et al., 2012, McLoughlin et al., 2005, Li et al., 2012b, Friedrichsen et al., 2001, Tronche et al., 2004). The STAT3 protein is phosphorylated and thus activated in response to cytokines or growth factors, for example IL-6 (Li et al., 2012b, McLoughlin et al., 2005). The activated STATs form homo- or heterodimers with other STAT-family members and are transported to the nucleus where they activate transcription (McLoughlin et al., 2005, Yuan et al., 2005, Wang et al., 1996).

STAT3 is thought to be involved in the regulation of apoptosis (Li et al., 2012b), supporting CD8+ dependent T cell activation (Imataki et al., 2012), fetal astrocyte development (Nakashima et al., 1999), Treg cell regulatory functions (Chaudhry et al., 2009), and as a key factor in Th17 cell differentiation (Liu et al., 2010). There is increasing evidence for the importance of Th17 cells in MS (Du et al., 2009, Kebir et al., 2007), in rodent EAE (Du et al., 2009, Cua et al., 2003), and in other autoimmune diseases (Nakae et al., 2003, Murphy et al., 2003). Mutations in *STAT3* have been reported to cause hyperimmunoglobulin E recurrent infection syndrome (HIES), also known as Job's disease (Holland et al., 2007, Minegishi et al., 2007). HIES is a monogenic autosomal dominant immune disorder characterized by inflammation and elevated immunoglobulin E levels (Holland et al., 2007, Minegishi et al., 2007).

The rs744166 in *STAT3* has also been associated with CD (Barrett et al., 2008). Interestingly, the MS risk increasing rs744166 C-allele (C) from our study was indicated to be the protective allele in CD (Barrett et al., 2008). In addition to the genetic associations, some functional evidence suggests that *STAT3* is involved in other autoimmune diseases, such as psoriasis (Sano et al., 2005) and CD (Lovato et al., 2003). The STAT3 pathway has been reported to be involved in the regulation of neuroinflammation and autoimmune demyelination in mice, possibly through Th17 cells (Mycko et al., 2012, Qin et al., 2012).

Further, rs744166 has been reported to be a tissue specific eQTL that regulates the expression of the adjacent Signal transducer and activator if transcription 5A and 5B (*STAT5A* and *STAT5B)* genes in a tissue specific manner (Dimas et al., 2009). Therefore, it is possible that the effect of the *STAT3* MS locus could be mediated through the *STAT5A* and *STAT5B* genes. Knock-out studies in mice have shown that Stat5A and Stat5B are required for the development of hematopoietic cell lines during fetal development (Snow et al., 2003, Socolovsky et al., 1999, Yao et al.,

2006). Knock-down and knock-out studies of the Stat5A and B genes in mice have been reported to cause fetal anemia, autoimmunity, decreased lifespan, perinatal lethality and a potential severe combined immunodeficiency (SCID) -type phenotype (Snow et al., 2003, Socolovsky et al., 1999, Yao et al., 2006). Interestingly, STAT5 has been suggested to be required for the survival and expansion of Th17 cells in response to activation on the IL7-IL7R signaling pathway (Liu et al., 2010).

5.1.2 Three regions of homozygosity in MS

Extended regions of homozygosity (ROHs) were analyzed in the Finnish GWAS sample set of 68 cases and 136 controls to find potential recessive loci. The minimum length for an individual ROH in an individual sample was set to 500kb to find large ROHs that could have been inherited from a common ancestor, and each ROH was required to cover at least 50 consecutive SNPs. All Finnish GWAS samples were compared to one another to identify potential overlap between the individual ROHs. On average, each sample had 149 ROHs (standard deviation 12 in cases, 10 in controls) that were 1030kb long (range 0.5-31.3Mb) in cases, and 1018kb (range 0.5-49.6Mb) in controls. All overlapping regions were permuted 10,000 times to find an empirical p-value for the enrichment of ROH in cases compared to controls.

Three regions of interest emerged (empirical $p < 10^{-3}$): 1q42.12 (242kb, p=0.0003), 2q24.3 (512kb, p=0.00008) and 12q24.33 (573kb, p=0.0003). In 1q42.12 and 2q24.3, a single haplotype was more frequent in the MS cases compared to controls (13% vs. 9% and 37% vs. 20%, respectively). The Health 2000 GenMets population samples were similar to the isolate controls: haplotype frequency in the 1q42.12 was 6% and haplotype frequency in the 2q24.3 was 20%. This indicates that these haplotypes could be enriched in the MS cases, at least in the isolate region. The third locus, in 12q24.33, had multiple associated haplotypes, and no single haplotype stood out. The 12q24.33 locus has been suggestively linked to MS in a previous study (Haines et al., 2002), but no strong evidence for association or linkage has been found for any of the loci (Patsopoulos et al., 2011, IMSGC, 2011). The 1q42.12 locus includes an open reading frame, and the 5' end of *DISP1* gene that is involved in embryonic development. The 2q24.3 locus includes the 3' end of the *XIRP2* gene that is involved in actin filaments. There are several genes in the 12q24.33 locus: *SFRS8*, *MMP17*, *ULK1*, *PUS1*, *EP400*, *EP400NL*, *DDX51*, *NOC4L*, and *GALNT9*. These genes are involved in RNA processing, extracellular matrix, autophagy, and nuclear membrane. None of these genes are obvious candidate genes for MS. Further studies on these regions and genes are needed to assess their relevance to MS predisposition.

5.1.3 Copy-number variants in MS patients

The Illumina HumanHap 300 platform is sparse, and most common CNV regions have been excluded from the chip. Therefore, our aim was to identify potential rare CNVs enriched in the Southern Ostrobothnian isolate region. CNV analysis of the 68 patients revealed 106 CNV regions. Nine of the 106 CNVs were previously unidentified, and each one was only observed once.

We hypothesized that the genes within, or next to, the CNVs could be a part of a common pathway that might play a role in MS etiology. Ingenuity Pathway Analysis revealed a single pathway, potentially regulating oligodendrocyte differentiation and myelin sheet formation, that involved the genes neuregulin 3 (*NRG3*), v-erb-a erythroblastic leukemia viral oncogene homolog 4 (*ERBB4*), discs, large homolog 2 (*DLG2*), utrophin (*UTRN*), and like-glycosyltransferase (*LARGE*). The CNVs in question were genotyped in a set of 703 MS cases and 1051 controls using a fragment analysis method. However, the CNV frequencies were not significantly different between the cases and controls. Eleven per cent of the MS cases carried a deletion in the *ERBB4* gene compared to 12% of controls (p=0.388), and the deletion in *NRG3* was present in 4% both in of cases and controls (p=0.90). The deletion in *DLG2* was present in 1% of cases, but was not detected at all in controls, and thus, needs to be further studied in order to define its significance. We were unable to define the borders for one deletion and one duplication, which lead to their exclusion. Based on this analysis, we could not identify any large rare CNVs that were enriched in the MS patients.

5.2 Meta-analysis of association between rs1800693 (*TNFRSF1A*), rs17445836 (*IRF8*) and rs17824933 (*CD6*) and MS, Study II

A recent meta-analysis of five independent data sets identified three genome-wide significant SNPs: rs1800693 in the tumor necrosis factor receptor superfamily, member 1A gene (*TNFRSF1A*), rs17445836 that is located 61.5 kb from the interferon regulatory factor 8 gene (*IRF8*) also know as interferon consensus sequence-binding protein gene (*ICSBP*), and rs17824933 in *CD6* (De Jager et al., 2009b). Our aim was to confirm the association of these loci with MS in another large international sample set. The study set included 8,047 cases, 9,174 controls and 608 trios from 11 populations across Europe and the USA (Materials and methods 4.3.2, Table 6). If a sample set didn't meet the quality control criteria for one of the SNPs, the sample set was excluded from the analysis of that particular SNP. We excluded the Danish and French samples from the analysis of *TNFRSF1A*, and the Spanish and German samples from the analysis of *IRF8*.

The data was analyzed using a meta-analysis method by Kazeem and Farrall 2005 that allowed us to include both case-control data and trios into the same analysis (Kazeem and Farrall, 2005). The strongest association with MS was observed at the *IRF8* locus ($p=5.34\times10^{-10}$, OR=0.84, 95% CI 0.80-0.89). The associations at the *TNFRSF1A* and *CD6* loci were more moderate ($p=4.19\times10^{-7}$,

OR=1.12, 95% CI 1.07-1.18 and p=2.19×10^{-5}, OR=1.11, 95% CI 1.06-1.11, respectively). Combining these results with the results from the original publication strengthened the association results: *TNFRSF1A* p=8.12×10^{-12} (OR 1.15), *IRF8* p=3.35×10^{-15} (OR=0.83), and *CD6* p=8.88×10^{-12} (OR=1.16). The results for *CD6* were the least significant, and the OR varied between populations. These loci have been further confirmed in recent studies (IMSGC, 2011).

The *IRF8* gene encodes a transcription factor that binds to the interferon-stimulated response element (ISRE) in response to the type I interferons (e.g interferon-α), (Weisz et al., 1992, Nelson et al., 1993). This protein has been reported to be important in the regulation of macrophage development, B cell differentiation (Tamura et al., 2000, Wang et al., 2008). The CD6 molecule is a cell-cell signaling and adhesion molecule (Singer et al., 2002, Whitney et al., 1995). The CD6 molecule has been suggested to play a role in apoptosis and positive selection of thymocytes in the thymus (Singer et al., 2002).

TNFRSF1A encodes a major TNFα receptor that can activate NF-κβ, mediated apoptosis (Micheau and Tschopp, 2003), and regulate inflammation (Gimenez et al., 2006). Mutations in *TNFRSF1A* have been reported to cause autosomal dominant familial Hibernian fever, or tumor necrosis factor receptor-associated periodic syndrome (TRAPS) (Ryan and Aksentijevich, 2009, Aksentijevich et al., 2001). A recent study found that 24% of MS patients (6 patients out of 24) with familial Hibernian fever-like symptoms carried the reported mutation (R92Q, rs4149584 A-allele), and they could have been misdiagnosed with MS (Kumpfel et al., 2007). In general, the mutation was present in 4.7% of the MS cases and 3.0% of the healthy controls (Kumpfel et al., 2007). The rs4149584 A-allele frequency is reported to be 4% in the CEU population, which could suggest that the familial Hibernian fever mutation might not be as penetrant or important to the disease as previously thought. Interestingly, familial Hibernian fever and MS both can have a relapsing-remitting disease course. Studies have shown that mice with no functional p55 (TNFR1/Tnfrsf1a/CD120a) receptor were resistant to EAE and lenercept (a recombinant TNF receptor p55 immunoglobulin fusion protein) was able to block EAE in preclinical studies (Suvannavejh et al., 2000). In humans the effect was opposite, and patients treated with lenercept experienced symptom exacerbation, compared to patients treated with placebo (van Oosten et al., 1996, Group, 1999).

5.3 Accumulation of common MS associated alleles in Southern Ostrobothnia, Study III

Current GWAS technologies have enabled the identification of over 50 common MS associated loci. Recent studies have observed an enrichment of a subset of these alleles both in MS patients, and especially in familial MS cases (De Jager et al., 2009a, Gourraud et al., 2011, D'Netto et al., 2009). We used the method described by De Jager et al. to calculate an updated genetic burden score to assess the accumulation of common genetic variants in the Southern Ostrobothnian MS patients, compared to population samples and non-isolate cases. In addition to the regional analysis, we assessed the accumulation of genetic burden in multiplex families compared to sporadic cases and population samples. MS patients in general had a significantly higher genetic burden score compared to population samples. The phenomenon was statistically significant in all comparisons between MS cases and population samples, when all loci were included in the analysis (Figures 9 and 11). These findings were consistent with the previous studies (De Jager et al.1 2009a, Gourraud et al., 2011, D'Netto et al., 2009).

5.3.1 Genetic burden score of MS associated alleles was not increased in the Southern Ostrobothian isolate

The MS cases in the isolate and the non-isolate groups had, on average, a significantly higher genetic burden score compared to the population samples (Figure 6). Despite the two-fold increase in the prevalence and incidence of MS, the genetic burden score did not differ significantly between the 111 Southern Ostrobothnian MS cases (average 6.64±0.64) and the general 497 Finnish MS patients (average 6.77±0.62) (p=0.16)(Figure 6). Similarly, the 135 Southern Ostrobothnian population samples (average 6.31±0.61) were not significantly different from the 5399 general population controls (average 6.23±0.62) (p=0.18) (Figure 9). The number of samples was relatively small for the isolate groups. However, according to power calculations, we had 80% power (alpha 0.05) to detect a difference of 0.19 in the genetic burden score between the groups, and 90% power to detect a difference of 0.22 with the analyzed sample set.

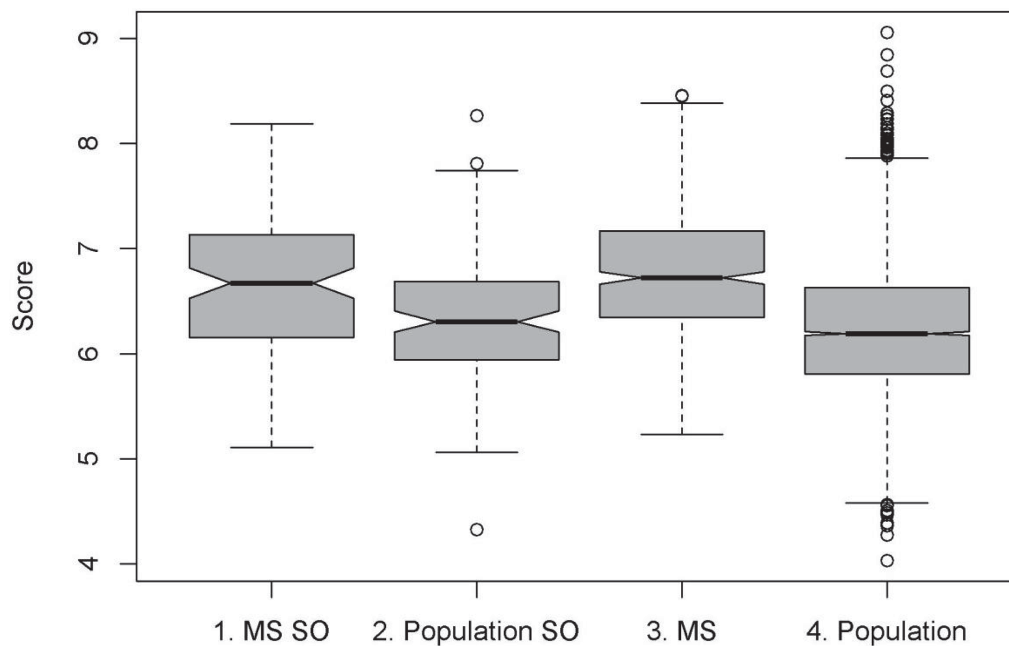

Figure 9 Distribution of the genetic burden score by region. The y-axis indicates the genetic burden score, the median of each sample group is represented by a thick horizontal line, boxes indicate the middle 50% of the data points, and the whiskers indicate the 1.5×inter-quartile range. The notches indicate the significance of the differences between the samples (±1.58 (interquartile range / √n): non-overlapping notches suggest a significant difference.

The differences between isolate and non-isolate samples were not significant. Therefore, we analyzed the interaction between affection status and the isolate region against genetic burden score. Affection status was significantly associated with the genetic burden score ($p < 2\times10^{-16}$) in linear regression analysis. In an additional linear model analysis of the Southern Ostrobothnian origin and affection status against the genetic burden score, the Southern Ostrobothnian origin was not significant (p=0.994). The multiplicative interaction analysis revealed that there was a borderline significant interaction between affection status and Southern Ostrobothnian origin (β=-0.211, se=0.841, p=0.0121), and that the genetic burden might in fact be reduced in isolate cases (Figure 9). It should be noted, however, that 25% of the isolate cases in this analysis had a familial background compared to only 7% of the non-isolate cases.

The *HLA* locus is the best known MS locus and it was first discovered in families. When the *HLA* locus was excluded from the analysis, the genetic burden score in the Southern Ostrobothnian isolate cases decreased to the population level (Figure 10). Despite the limited power, it could be speculated that in the isolate region, most of the accumulated risk beyond the population level could be caused by the *HLA* region and potentially by as yet unknown, and possibly rare, MS loci.

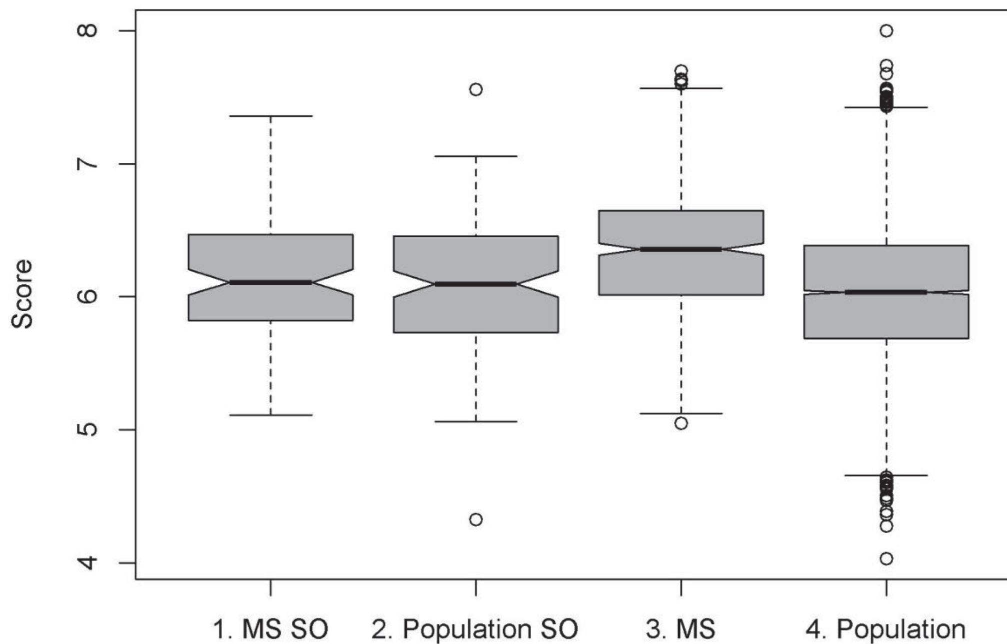

Figure 10 Distribution of the genetic burden score without the HLA proxy SNP by region. The genetic burden score in MS samples from the isolate (1. MS SO) is close to both the isolate population level (2. Population SO) and general population level (4. Population), whereas the non-isolate MS cases (3. MS) have a significantly higher genetic burden score than all other samples. The y-axis indicates the genetic burden score, and the median of each sample group is represented by a thick horizontal line in the middle of the box.

5.3.2 No evidence for accumulation of common MS associated alleles in Finnish multiplex families

Analysis of the genetic burden score in sporadic MS cases, familial cases, and population samples was performed using the statistical program R 1.15.1. Sixty-three individuals with at least one relative with MS and 522 sporadic cases were included. Twenty-three cases reported a relative with MS, but were excluded due to an unconfirmed affection status of the reported relative. Both the sporadic MS cases (n=522, average=6.731, sd=0.621) and the familial cases (n=63, average=6.795, sd=0.636) were compared to each other and the general population sample (n=5534, average=6.230, sd=0.616), (Figure 11). Both non-familial and familial MS cases were significantly different from the population sample ($p<2.2\times10^{-16}$ and $p=2.98\times10^{-7}$ respectively), (Figure 8). Unlike previous studies, we did not observe a significant difference between the familial and sporadic cases (p=0.216).

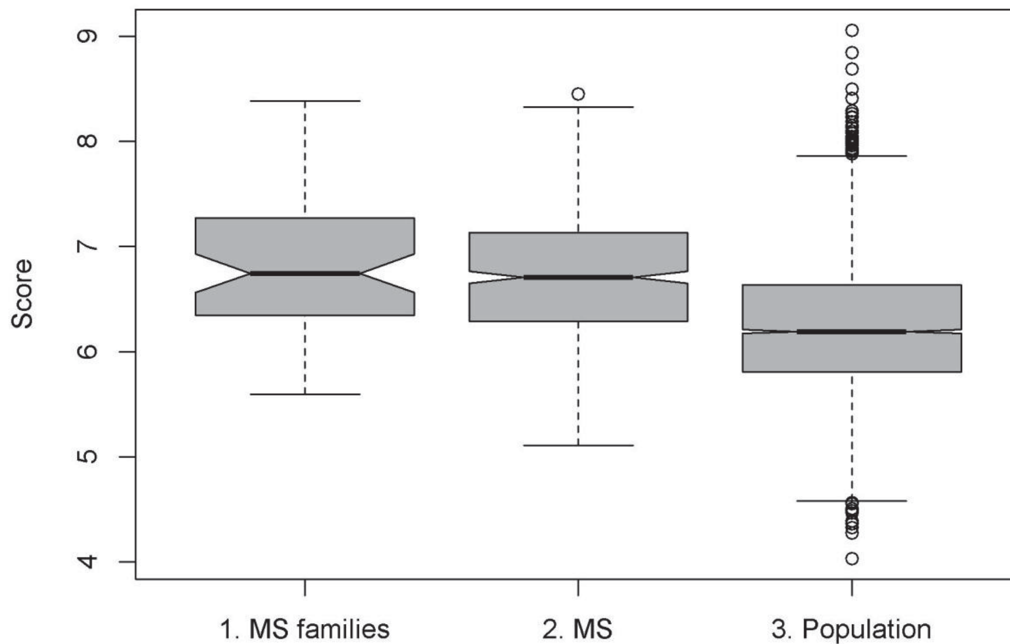

Figure 11 Distribution of the genetic burden score by familial background. The y-axis represents the genetic burden score.

Since 28 (44%) of the 63 familial samples originated from the isolate, we plotted the familial cases according to their isolate status. There was a slight trend towards higher risk in the non-isolate familial cases, but the number of samples was too low to draw conclusions, and the difference was not statistically significant. No correlation was found with either affection status or family status in the regression analyses, and the number of samples was too small to conduct a combined analysis of isolate and familial status.

A previous publication reported a significant difference of 0.27 in the genetic burden score between familial and sporadic cases (Gourraud et al., 2011). The difference observed in our study was much smaller, 0.064. Our sample size is small, but we had an estimated power of over 80% to detect a difference of 0.27. When we simulated a similar set of data using a difference of 0.27, and used the observed number of samples and standard deviations, only 0.8% of the 10,000 simulations produced a difference that was less than 0.065. This increases the likelihood that the lack of difference observed in our study may not just be due to small sample size and lack of power, but instead may indicate that the genetic burden might not be increased in the Southern Ostrobothnia.

6 CONCLUSIONS AND FUTURE PROSPECTS

We have utilized the Finnish population structure, genome-wide methods, haplotype analysis and meta-analysis methods, to identify an association between *STAT3* and MS in the Southern Ostrobothnian high-risk isolate population in Study I. Together with the replicated variants in *TNFRSF1A*, *IRF8,* and *CD6* from Study II, the variants in *STAT3,* alongside the other currently known MS associated variants (n>50), are beginning to form a picture of immunological pathways that are relevant in certain cell types and might be affected by genetic variation that is relevant for MS (IMSGC, 2011, Maurano et al., 2012). The odds ratios for individual variants are modest and many associated SNPs lie outside of the coding regions of the genes, which is common for all complex diseases. Therefore, more studies are needed to identify causative functional variants and their role in MS disease pathogenesis.

MS has an increased prevalence and incidence in the Southern Ostrobothnian isolate of Western Finland (Tienari et al., 2004). However, we could not find evidence for overall accumulation of common MS associated variants in this high-risk isolate. Therefore, it is reasonable to hypothesize, that there could be population specific high-risk variants in the isolate region. The Finnish population structure, population history and isolate populations can be useful in gene mapping, especially of potential rare high-risk variants, and have been successfully used in mapping of monogenic Mendelian disorders (Tallila et al., 2009). The usefulness is not self evident in common complex diseases, although a recent study predicted a significant role for rare variants in complex diseases (Tallila et al., 2009). For example, although the increased prevalence of MS in Southern Ostrobothnia could be due to enrichment of rare genetic high-risk variants, it is not possible to exclude the effect of other factors, such as the environment. Another consideration is that if the predisposing genetic variants would have been enriched in the isolate, it seems reasonable to expect them to be enriched both in cases and in controls. Therefore, the odds ratio needs to be relatively large to provide enough power to detect such variants. One could hypothesize that this is possible in late-onset diseases, such as MS, if the variant does not significantly affect the number of offspring. Then, rare high-risk variants should be seen as familial aggregation. Indeed, there are several extended multiplex families with MS in Southern Ostrobothnia (Wikström et al., 1984) and these families should be carefully studied in the future.

Many studies use genome-wide genotyping chips to genotype common SNPs that are unlikely to have high odds ratios for common diseases and traits. On the other hand, searching for rare variants of importance by using common SNP genotypes is challenging. Imputation approaches have been developed, but they depend on previously identified variants, the observed population, the reference population, available genotype coverage, and they create genotypes with some degree of uncertainty. Another option to increase resolution could be a genome-wide haplotype

analysis, but also this approach would be limited by the observed genotypes and genotype combinations.

The price for next generation sequencing is becoming affordable, and this technique can provide a way to directly measure most variation in individual genomes. There are already some examples, where studies have found rare likely causative *de novo* mutations using the NGS exome sequencing technology (Neale et al., 2012, O'Roak et al., 2012, Sanders et al., 2012). However, since most of the genome-wide hits in MS and other complex disorders lie outside protein coding regions, one could speculate that whole-genome sequencing might prove to be more informative than exome sequencing. So far, lack of functional annotation has hindered the interpretation of most variants outside coding regions, but annotations from the ENCODE project will likely improve this in the near future (Encode Project Consortium et al., 2012).

Many MS associated SNPs are located in regulatory regions (Maurano et al., 2012). For example, RNA sequencing in individuals who carry a potential risk allele or haplotype could give insights not only about the transcript levels, but also the transcribed isoforms. Since regulatory regions can be tissue specific (Dimas et al., 2009, Djebali et al., 2012, Encode Project Consortium et al., 2012), the tissues of interest should be selected carefully, perhaps based on previous pathway analyses (Maurano et al., 2012, IMSGC, 2007).

We have initiated targeted re-sequencing of the *STAT3* locus in MS patients, aiming to identify variants that could explain the association signal. We have focused our attention to the individuals who carry the *STAT3* risk haplotype. The verification or replication of identified candidate variants may, however, prove to be tricky. Especially variants that are private to a family, or a region, will require sophisticated analysis methods, large cohorts from potential isolate populations, and different functional studies. A study in large granular lymphocytic leukemia, which can have autoimmune disorder-like symptoms, has identified mutations in the STAT3 functional domain that were associated with phosphorylation status and activation (Koskela et al., 2012). Therefore, we are also planning to assess the phosphorylation status of STAT3 in MS patients and healthy individuals.

Insight into the genetic etiology of complex diseases has increased significantly over the past few years. The work presented in this thesis has contributed to the collective knowledge and understanding of multiple sclerosis. Ultimately, these genetic studies will lead to the better comprehension of MS disease pathogenesis, biological disease mechanisms and to better treatment options.

WEB-BASED RESOURCES

HapMap http://www.hapmap.org
UCSC Genome Browser http://genome.ucsc.edu
Roadmap Epigenomics Project http://www.roadmapepigenomics.org/
Finnish Disease Database http://www.findis.org
eQTL resources @ the prithcard lab http://eqtl.uchicago.edu/

Catalog of Published Genome-Wide Association Studies
https://www.genome.gov/gwastudies/.

REFERENCES

1000 GENOMES PROJECT CONSORTIUM 2010. A map of human genome variation from population-scale sequencing. *Nature,* 467**,** 1061-73.

ABADIE, V., BERTHELOT, J., FEILLET, F., MAURIN, N., MERCIER, A., DE BAULNY, H. O. & DE PARSCAU, L. 2001. Neonatal screening and long-term follow-up of phenylketonuria: the French database. *Early Hum Dev,* 65**,** 149-58.

AHLGREN, C., ODEN, A. & LYCKE, J. 2012. A nationwide survey of the prevalence of multiple sclerosis in immigrant populations of Sweden. *Mult Scler*.

AKSENTIJEVICH, I., GALON, J., SOARES, M., MANSFIELD, E., HULL, K., OH, H. H., GOLDBACH-MANSKY, R., DEAN, J., ATHREYA, B., REGINATO, A. J., HENRICKSON, M., PONS-ESTEL, B., O'SHEA, J. J. & KASTNER, D. L. 2001. The tumor-necrosis-factor receptor-associated periodic syndrome: new mutations in TNFRSF1A, ancestral origins, genotype-phenotype studies, and evidence for further genetic heterogeneity of periodic fevers. *Am J Hum Genet,* 69**,** 301-14.

ALTMULLER, J., PALMER, L. J., FISCHER, G., SCHERB, H. & WJST, M. 2001. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet,* 69**,** 936-50.

ALUVIHARE, V. R., KALLIKOURDIS, M. & BETZ, A. G. 2004. Regulatory T cells mediate maternal tolerance to the fetus. *Nat Immunol,* 5**,** 266-71.

ALVAREZ-LAFUENTE, R., GARCIA-MONTOJO, M., DE LAS HERAS, V., BARTOLOME, M. & ARROYO, R. 2006. Clinical parameters and HHV-6 active replication in relapsing-remitting multiple sclerosis patients. *J Clin Virol,* 37 Suppl 1**,** S24-6.

ANZGENE, A. A. N. Z. M. S. G. C. 2009. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet,* 41**,** 824-8.

AOKI, S., YAOITA, H. & KITAJIMA, Y. 1989. An elevated level of autoantibodies against 48- to 50-kd keratins in the serum of patients with psoriasis. *J Invest Dermatol,* 92**,** 179-83.

ASANO, K., MATSUSHITA, T., UMENO, J., HOSONO, N., TAKAHASHI, A., KAWAGUCHI, T., MATSUMOTO, T., MATSUI, T., KAKUTA, Y., KINOUCHI, Y., SHIMOSEGAWA, T., HOSOKAWA, M., ARIMURA, Y., SHINOMURA, Y., KIYOHARA, Y., TSUNODA, T., KAMATANI, N., IIDA, M., NAKAMURA, Y. & KUBO, M. 2009. A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. *Nat Genet,* 41**,** 1325-9.

ASANO, M., TODA, M., SAKAGUCHI, N. & SAKAGUCHI, S. 1996. Autoimmune disease as a consequence of developmental abnormality of a T cell subpopulation. *J Exp Med,* 184**,** 387-96.

ASCHERIO, A., MUNGER, K. L., LENNETTE, E. T., SPIEGELMAN, D., HERNAN, M. A., OLEK, M. J., HANKINSON, S. E. & HUNTER, D. J. 2001. Epstein-Barr virus antibodies and risk of multiple sclerosis: a prospective study. *JAMA,* 286**,** 3083-8.

ASSEMAN, C., MAUZE, S., LEACH, M. W., COFFMAN, R. L. & POWRIE, F. 1999. An essential role for interleukin 10 in the function of regulatory T cells that inhibit intestinal inflammation. *J Exp Med,* 190**,** 995-1004.

ATKINSON, M. A., KAUFMAN, D. L., CAMPBELL, L., GIBBS, K. A., SHAH, S. C., BU, D. F., ERLANDER, M. G., TOBIN, A. J. & MACLAREN, N. K. 1992. Response of peripheral-blood mononuclear cells to glutamate decarboxylase in insulin-dependent diabetes. *Lancet,* 339**,** 458-9.

AUSTYN, J. M., WEINSTEIN, D. E. & STEINMAN, R. M. 1988. Clustering with dendritic cells precedes and is essential for T-cell proliferation in a mitogenesis model. *Immunology,* 63**,** 691-6.

AYODO, G., PRICE, A. L., KEINAN, A., AJWANG, A., OTIENO, M. F., ORAGO, A. S., PATTERSON, N. & REICH, D. 2007. Combining evidence of natural selection with

association analysis increases power to detect malaria-resistance variants. *Am J Hum Genet,* 81**,** 234-42.

BABBE, H., ROERS, A., WAISMAN, A., LASSMANN, H., GOEBELS, N., HOHLFELD, R., FRIESE, M., SCHRODER, R., DECKERT, M., SCHMIDT, S., RAVID, R. & 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.

BAN, M., MCCAULEY, J. L., ZUVICH, R., BAKER, A., BERGAMASCHI, L., COX, M., KEMPPINEN, A., D'ALFONSO, S., GUERINI, F. R., LECHNER-SCOTT, J., DUDBRIDGE, F., WASON, J., ROBERTSON, N. P., DE JAGER, P. L., HAFLER, D. A., BARCELLOS, L. F., IVINSON, A. J., SEXTON, D., OKSENBERG, J. R., HAUSER, S. L., PERICAK-VANCE, M. A., HAINES, J., COMPSTON, A. & SAWCER, S. 2010. A non-synonymous SNP within membrane metalloendopeptidase-like 1 (MMEL1) is associated with multiple sclerosis. *Genes Immun,* 11**,** 660-4.

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. & OKSENBERG, J. R. 2009. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet,* 18**,** 767-78.

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. 2009a. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet,* 41**,** 703-7.

BARRETT, J. C., FRY, B., MALLER, J. & 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. & 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., LEE, J. C., LEES, C. W., PRESCOTT, N. J., ANDERSON, C. A., PHILLIPS, A., WESLEY, E., PARNELL, K., ZHANG, H., DRUMMOND, H., NIMMO, E. R., MASSEY, D., BLASZCZYK, K., ELLIOTT, T., COTTERILL, L., DALLAL, H., LOBO, A. J., MOWAT, C., SANDERSON, J. D., JEWELL, D. P., NEWMAN, W. G., EDWARDS, C., AHMAD, T., MANSFIELD, J. C., SATSANGI, J., PARKES, M., MATHEW, C. G., DONNELLY, P., PELTONEN, L., BLACKWELL, J. M., BRAMON, E., BROWN, M. A., CASAS, J. P., CORVIN, A., CRADDOCK, N., DELOUKAS, P., DUNCANSON, A., JANKOWSKI, J., MARKUS, H. S., MCCARTHY, M. I., PALMER, C. N., PLOMIN, R., RAUTANEN, A., SAWCER, S. J., SAMANI, N., TREMBATH, R. C., VISWANATHAN, A. C., WOOD, N., SPENCER, C. C., BELLENGUEZ, C., DAVISON, D., FREEMAN, C., STRANGE, A., LANGFORD, C., HUNT, S. E., EDKINS, S., GWILLIAM, R., BLACKBURN, H., BUMPSTEAD, S. J., DRONOV, S., GILLMAN, M., GRAY, E.,

HAMMOND, N., JAYAKUMAR, A., MCCANN, O. T., LIDDLE, J., PEREZ, M. L., POTTER, S. C., RAVINDRARAJAH, R., RICKETTS, M., WALLER, M., WESTON, P., WIDAA, S., WHITTAKER, P., ATTWOOD, A. P., STEPHENS, J., SAMBROOK, J., OUWEHAND, W. H., MCARDLE, W. L., RING, S. M. & STRACHAN, D. P. 2009b. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet,* 41**,** 1330-4.

BAUMANN, I., KOLOWOS, W., VOLL, R. E., MANGER, B., GAIPL, U., NEUHUBER, W. L., KIRCHNER, T., KALDEN, J. R. & HERRMANN, M. 2002. Impaired uptake of apoptotic cells into tingible body macrophages in germinal centers of patients with systemic lupus erythematosus. *Arthritis Rheum,* 46**,** 191-201.

BECKER, T., HARTL, F. U. & WIELAND, F. 2002. CD40, an extracellular receptor for binding and uptake of Hsp70-peptide complexes. *J Cell Biol,* 158**,** 1277-85.

BEGOVICH, A. B., CARLTON, V. E., HONIGBERG, L. A., SCHRODI, S. J., CHOKKALINGAM, A. P., ALEXANDER, H. C., ARDLIE, K. G., HUANG, Q., SMITH, A. M., SPOERKE, J. M., CONN, M. T., CHANG, M., CHANG, S. Y., SAIKI, R. K., CATANESE, J. J., LEONG, D. U., GARCIA, V. E., MCALLISTER, L. B., JEFFERY, D. A., LEE, A. T., BATLIWALLA, F., REMMERS, E., CRISWELL, L. A., SELDIN, M. F., KASTNER, D. L., AMOS, C. I., SNINSKY, J. J. & GREGERSEN, P. K. 2004. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am J Hum Genet,* 75**,** 330-7.

BELL, G. I., HORITA, S. & KARAM, J. H. 1984. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes,* 33**,** 176-83.

BENEDIKZ, J., STEFANSSON, M., GUOMUNDSSON, J., JONASDOTTIR, A., FOSSDAL, R., GULCHER, J. & 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.

BENTZEN, J., FLACHS, E. M., STENAGER, E., BRONNUM-HANSEN, H. & KOCH-HENRIKSEN, N. 2010. Prevalence of multiple sclerosis in Denmark 1950--2005. *Mult Scler,* 16**,** 520-5.

BERSAGLIERI, T., SABETI, P. C., PATTERSON, N., VANDERPLOEG, T., SCHAFFNER, S. F., DRAKE, J. A., RHODES, M., REICH, D. E. & HIRSCHHORN, J. N. 2004. Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet,* 74**,** 1111-20.

BIRNEY, E., STAMATOYANNOPOULOS, J. A., DUTTA, A., GUIGO, R., GINGERAS, T. R., MARGULIES, E. H., WENG, Z., SNYDER, M., DERMITZAKIS, E. T., THURMAN, R. E., KUEHN, M. S., TAYLOR, C. M., NEPH, S., KOCH, C. M., ASTHANA, S., MALHOTRA, A., ADZHUBEI, I., GREENBAUM, J. A., ANDREWS, R. M., FLICEK, P., BOYLE, P. J., CAO, H., CARTER, N. P., CLELLAND, G. K., DAVIS, S., DAY, N., DHAMI, P., DILLON, S. C., DORSCHNER, M. O., FIEGLER, H., GIRESI, P. G., GOLDY, J., HAWRYLYCZ, M., HAYDOCK, A., HUMBERT, R., JAMES, K. D., JOHNSON, B. E., JOHNSON, E. M., FRUM, T. T., ROSENZWEIG, E. R., KARNANI, N., LEE, K., LEFEBVRE, G. C., NAVAS, P. A., NERI, F., PARKER, S. C., SABO, P. J., SANDSTROM, R., SHAFER, A., VETRIE, D., WEAVER, M., WILCOX, S., YU, M., COLLINS, F. S., DEKKER, J., LIEB, J. D., TULLIUS, T. D., CRAWFORD, G. E., SUNYAEV, S., NOBLE, W. S., DUNHAM, I., DENOEUD, F., REYMOND, A., KAPRANOV, P., ROZOWSKY, J., ZHENG, D., CASTELO, R., FRANKISH, A., HARROW, J., GHOSH, S., SANDELIN, A., HOFACKER, I. L., BAERTSCH, R., KEEFE, D., DIKE, S., CHENG, J., HIRSCH, H. A., SEKINGER, E. A., LAGARDE, J., ABRIL, J. F., SHAHAB, A., FLAMM, C., FRIED, C., HACKERMULLER, J., HERTEL, J., LINDEMEYER, M., MISSAL, K., TANZER, A., WASHIETL, S., KORBEL, J., EMANUELSSON, O., PEDERSEN, J. S., HOLROYD, N., TAYLOR, R., SWARBRECK, D., MATTHEWS, N., DICKSON, M. C., THOMAS, D. J., WEIRAUCH, M. T., GILBERT, J., et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature,* 447**,** 799-816.

BOSTROM, I., CALLANDER, M., KURTZKE, J. F. & 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. & COMPSTON, D. A. 2000. Autoimmune disease in first-degree relatives of patients with multiple sclerosis. A UK survey. *Brain,* 123 (Pt 6)**,** 1102-11.

BRONNUM-HANSEN, H., KOCH-HENRIKSEN, N. & STENAGER, E. 2004. Trends in survival and cause of death in Danish patients with multiple sclerosis. *Brain,* 127**,** 844-50.

BROWNING, B. L. & BROWNING, S. R. 2007. Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. *Genet Epidemiol,* 31**,** 365-75.

BULEK, A. M., COLE, D. K., SKOWERA, A., DOLTON, G., GRAS, S., MADURA, F., FULLER, A., MILES, J. J., GOSTICK, E., PRICE, D. A., DRIJFHOUT, J. W., KNIGHT, R. R., HUANG, G. C., LISSIN, N., MOLLOY, P. E., WOOLDRIDGE, L., JAKOBSEN, B. K., ROSSJOHN, J., PEAKMAN, M., RIZKALLAH, P. J. & SEWELL, A. K. 2012. Structural basis for the killing of human beta cells by CD8(+) T cells in type 1 diabetes. *Nat Immunol,* 13**,** 283-9.

BUYSKE, S., YANG, G., MATISE, T. C. & GORDON, D. 2009. When a case is not a case: effects of phenotype misclassification on power and sample size requirements for the transmission disequilibrium test with affected child trios. *Hum Hered,* 67**,** 287-92.

CARGILL, M., SCHRODI, S. J., CHANG, M., GARCIA, V. E., BRANDON, R., CALLIS, K. P., MATSUNAMI, N., ARDLIE, K. G., CIVELLO, D., CATANESE, J. J., LEONG, D. U., PANKO, J. M., MCALLISTER, L. B., HANSEN, C. B., PAPENFUSS, J., PRESCOTT, S. M., WHITE, T. J., LEPPERT, M. F., KRUEGER, G. G. & BEGOVICH, A. B. 2007. A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am J Hum Genet,* 80**,** 273-90.

CENIT, M. C., ALCINA, A., MARQUEZ, A., MENDOZA, J. L., DIAZ-RUBIO, M., DE LAS HERAS, V., IZQUIERDO, G., ARROYO, R., FERNANDEZ, O., DE LA CONCHA, E. G., MATESANZ, F. & URCELAY, E. 2010. STAT3 locus in inflammatory bowel disease and multiple sclerosis susceptibility. *Genes Immun,* 11**,** 264-8.

CHAPMAN, J. M., COOPER, J. D., TODD, J. A. & CLAYTON, D. G. 2003. Detecting disease associations due to linkage disequilibrium using haplotype tags: a class of tests and the determinants of statistical power. *Hum Hered.,* 56**,** 18-31.

CHAUDHRY, A., RUDRA, D., TREUTING, P., SAMSTEIN, R. M., LIANG, Y., KAS, A. & RUDENSKY, A. Y. 2009. CD4+ regulatory T cells control TH17 responses in a Stat3-dependent manner. *Science,* 326**,** 986-91.

CHIMPANZEE SEQUENCING AND ANALYSIS CONSORTIUM 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature,* 437**,** 69-87.

CLAMP, M., FRY, B., KAMAL, M., XIE, X., CUFF, J., LIN, M. F., KELLIS, M., LINDBLAD-TOH, K. & LANDER, E. S. 2007. Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci U S A,* 104**,** 19428-33.

CLAYTON, D. G., WALKER, N. M., SMYTH, D. J., PASK, R., COOPER, J. D., MAIER, L. M., SMINK, L. J., LAM, A. C., OVINGTON, N. R., STEVENS, H. E., NUTLAND, S., HOWSON, J. M., FAHAM, M., MOORHEAD, M., JONES, H. B., FALKOWSKI, M., HARDENBOL, P., WILLIS, T. D. & TODD, J. A. 2005. Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat Genet,* 37**,** 1243-6.

COHEN, J. C., KISS, R. S., PERTSEMLIDIS, A., MARCEL, Y. L., MCPHERSON, R. & HOBBS, H. H. 2004. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science,* 305**,** 869-72.

COMPSTON, D. A., MORGAN, B. P., CAMPBELL, A. K., WILKINS, P., COLE, G., THOMAS, N. D. & JASANI, B. 1989. Immunocytochemical localization of the terminal complement complex in multiple sclerosis. *Neuropathol Appl Neurobiol,* 15**,** 307-16.

CONFAVREUX, C., HUTCHINSON, M., HOURS, M. M., CORTINOVIS-TOURNIAIRE, P. & MOREAU, T. 1998. Rate of pregnancy-related relapse in multiple sclerosis. Pregnancy in Multiple Sclerosis Group. *N Engl J Med,* 339**,** 285-91.

CONFAVREUX, C. & VUKUSIC, S. 2006. Natural history of multiple sclerosis: a unifying concept. *Brain,* 129**,** 606-16.

CONRAD, D. F., JAKOBSSON, M., COOP, G., WEN, X., WALL, J. D., ROSENBERG, N. A. & PRITCHARD, J. K. 2006. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet,* 38**,** 1251-60.

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. & TODD, J. A. 2008. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat Genet,* 40**,** 1399-401.

COTSAPAS, C., VOIGHT, B. F., ROSSIN, E., LAGE, K., NEALE, B. M., WALLACE, C., ABECASIS, G. R., BARRETT, J. C., BEHRENS, T., CHO, J., DE JAGER, P. L., ELDER, J. T., GRAHAM, R. R., GREGERSEN, P., KLARESKOG, L., SIMINOVITCH, K. A., VAN HEEL, D. A., WIJMENGA, C., WORTHINGTON, J., TODD, J. A., HAFLER, D. A., RICH, S. S. & DALY, M. J. 2011. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet,* 7**,** e1002254.

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. & SEDGWICK, J. D. 2003. Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature,* 421**,** 744-8.

D'NETTO, M. J., WARD, H., MORRISON, K. M., RAMAGOPALAN, S. V., DYMENT, D. A., DELUCA, G. C., HANDUNNETTHI, L., SADOVNICK, A. D. & EBERS, G. C. 2009. Risk alleles for multiple sclerosis in multiplex families. *Neurology,* 72**,** 1984-8.

DAHL, O. P., AARSETH, J. H., MYHR, K. M., NYLAND, H. & MIDGARD, R. 2004. Multiple sclerosis in Nord-Trondelag County, Norway: a prevalence and incidence study. *Acta Neurol Scand,* 109**,** 378-84.

DANIEL, D., GILL, R. G., SCHLOOT, N. & WEGMANN, D. 1995. Epitope specificity, cytokine production profile and diabetogenic activity of insulin-specific T cell clones isolated from NOD mice. *Eur J Immunol,* 25**,** 1056-62.

DE BAKKER, P. I., MCVEAN, G., SABETI, P. C., MIRETTI, M. M., GREEN, T., MARCHINI, J., KE, X., MONSUUR, A. J., WHITTAKER, P., DELGADO, M., MORRISON, J., RICHARDSON, A., WALSH, E. C., GAO, X., GALVER, L., HART, J., HAFLER, D. A., PERICAK-VANCE, M., TODD, J. A., DALY, M. J., TROWSDALE, J., WIJMENGA, C., VYSE, T. J., BECK, S., MURRAY, S. S., CARRINGTON, M., GREGORY, S., DELOUKAS, P. & RIOUX, J. D. 2006. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet,* 38**,** 1166-72.

DE BAKKER, P. I., YELENSKY, R., PE'ER, I., GABRIEL, S. B., DALY, M. J. & ALTSHULER, D. 2005. Efficiency and power in genetic association studies. *Nat Genet,* 37**,** 1217-23.

DE JAGER, P. L., CHIBNIK, L. B., CUI, J., REISCHL, J., LEHR, S., SIMON, K. C., AUBIN, C., BAUER, D., HEUBACH, J. F., SANDBRINK, R., TYBLOVA, M., LELKOVA, P., STUDY, S. C. O. T. B., STUDY, S. C. O. T. B., STUDY, S. C. O. T. L., STUDY, S. C. O. T. C., HAVRDOVA, E., POHL, C., HORAKOVA, D., ASCHERIO, A., HAFLER, D. A. & KARLSON, E. W. 2009a. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. *Lancet Neurol,* 8**,** 1111-9.

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. & OKSENBERG, J. R. 2009b. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet,* 41**,** 776-82.

DE MAN, Y. A., DOLHAIN, R. J., VAN DE GEIJN, F. E., WILLEMSEN, S. P. & HAZES, J. M. 2008. Disease activity of rheumatoid arthritis during pregnancy: results from a nationwide prospective study. *Arthritis Rheum,* 59**,** 1241-8.

DE SIMONE, R., GIAMPAOLO, A., GIOMETTO, B., GALLO, P., LEVI, G., PESCHLE, C. & ALOISI, F. 1995. The costimulatory molecule B7 is expressed on human microglia in culture and in multiple sclerosis acute lesions. *J Neuropathol Exp Neurol,* 54**,** 175-87.

DEAN, G. & ELIAN, M. 1997. Age at immigration to England of Asian and Caribbean immigrants and the risk of developing multiple sclerosis. *J Neurol Neurosurg Psychiatry,* 63**,** 565-8.

DEBOUVERIE, M., PITTION-VOUYOVITCH, S., LOUIS, S., ROEDERER, T. & GUILLEMIN, F. 2007. Increasing incidence of multiple sclerosis among women in Lorraine, Eastern France. *Mult Scler,* 13**,** 962-7.

DELMOTTE, P. & GONSETTE, R. 1977. Biochemical findings in multiple sclerosis IV. Isoelectric focusing of the CSF gamma globulins in multiple sclerosis (262 cases) and other neurological diseases (272 cases). *J Neurol,* 215**,** 27-37.

DEVLIN, B. & ROEDER, K. 1999. Genomic control for association studies. *Biometrics,* 55**,** 997-1004.

DIETZSCHOLD, B., SCHWAEBLE, W., SCHAFER, M. K., HOOPER, D. C., ZEHNG, Y. M., PETRY, F., SHENG, H., FINK, T., LOOS, M., KOPROWSKI, H. & WEIHE, E. 1995. Expression of C1q, a subcomponent of the rat complement system, is dramatically enhanced in brains of rats with either Borna disease or experimental allergic encephalomyelitis. *J Neurol Sci,* 130**,** 11-6.

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. & ANTONARAKIS, S. E. 2009. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science,* 325**,** 1246-50.

DJEBALI, S., DAVIS, C. A., MERKEL, A., DOBIN, A., LASSMANN, T., MORTAZAVI, A., TANZER, A., LAGARDE, J., LIN, W., SCHLESINGER, F., XUE, C., MARINOV, G. K., KHATUN, J., WILLIAMS, B. A., ZALESKI, C., ROZOWSKY, J., RODER, M., KOKOCINSKI, F., ABDELHAMID, R. F., ALIOTO, T., ANTOSHECHKIN, I., BAER, M. T., BAR, N. S., BATUT, P., BELL, K., BELL, I., CHAKRABORTTY, S., CHEN, X., CHRAST, J., CURADO, J., DERRIEN, T., DRENKOW, J., DUMAIS, E., DUMAIS, J., DUTTAGUPTA, R., FALCONNET, E., FASTUCA, M., FEJES-TOTH, K., FERREIRA, P., FOISSAC, S., FULLWOOD, M. J., GAO, H., GONZALEZ, D., GORDON, A., GUNAWARDENA, H., HOWALD, C., JHA, S., JOHNSON, R., KAPRANOV, P., KING, B., KINGSWOOD, C., LUO, O. J., PARK, E., PERSAUD, K., PREALL, J. B., RIBECA, P., RISK, B., ROBYR, D., SAMMETH, M., SCHAFFER, L., SEE, L. H., SHAHAB, A., SKANCKE, J., SUZUKI, A. M., TAKAHASHI, H., TILGNER, H., TROUT, D., WALTERS, N., WANG, H., WROBEL, J., YU, Y., RUAN, X., HAYASHIZAKI, Y., HARROW, J., GERSTEIN, M., HUBBARD, T., REYMOND, A., ANTONARAKIS, S. E., HANNON, G., GIDDINGS, M. C., RUAN, Y., WOLD, B., CARNINCI, P., GUIGO, R. & GINGERAS, T. R. 2012. Landscape of transcription in human cells. *Nature,* 489**,** 101-8.

DOYLE, C. & STROMINGER, J. L. 1987. Interaction between CD4 and class II MHC molecules mediates cell adhesion. *Nature,* 330**,** 256-9.

DU, C., LIU, C., KANG, J., ZHAO, G., YE, Z., HUANG, S., LI, Z., WU, Z. & PEI, G. 2009. MicroRNA miR-326 regulates TH-17 differentiation and is associated with the pathogenesis of multiple sclerosis. *Nat Immunol,* 10**,** 1252-9.

DUERR, R. H., TAYLOR, K. D., BRANT, S. R., RIOUX, J. D., SILVERBERG, M. S., DALY, M. J., STEINHART, A. H., ABRAHAM, C., REGUEIRO, M., GRIFFITHS, A., DASSOPOULOS, T., BITTON, A., YANG, H., TARGAN, S., DATTA, L. W., KISTNER,

E. O., SCHUMM, L. P., LEE, A. T., GREGERSEN, P. K., BARMADA, M. M., ROTTER, J. I., NICOLAE, D. L. & CHO, J. H. 2006. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science,* 314**,** 1461-3.

EBERS, G. C., BULMAN, D. E., SADOVNICK, A. D., PATY, D. W., WARREN, S., HADER, W., MURRAY, T. J., SELAND, T. P., DUQUETTE, P., GREY, T. & ET AL. 1986. A population-based study of multiple sclerosis in twins. *N Engl J Med,* 315**,** 1638-42.

EBERS, G. C., SADOVNICK, A. D. & RISCH, N. J. 1995. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature,* 377**,** 150-1.

EBERS, G. C., YEE, I. M., SADOVNICK, A. D. & DUQUETTE, P. 2000. Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. Canadian Collaborative Study Group. *Ann Neurol,* 48**,** 927-31.

ENCODE PROJECT CONSORTIUM, BERNSTEIN, B. E., BIRNEY, E., DUNHAM, I., GREEN, E. D., GUNTER, C. & SNYDER, M. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.

FALLARINO, F., GROHMANN, U., HWANG, K. W., ORABONA, C., VACCA, C., BIANCHI, R., BELLADONNA, M. L., FIORETTI, M. C., ALEGRE, M. L. & PUCCETTI, P. 2003. Modulation of tryptophan catabolism by regulatory T cells. *Nat Immunol,* 4**,** 1206-12.

FARRELL, R. A., ANTONY, D., WALL, G. R., CLARK, D. A., FISNIKU, L., SWANTON, J., KHALEELI, Z., SCHMIERER, K., MILLER, D. H. & GIOVANNONI, G. 2009. Humoral immune response to EBV in multiple sclerosis is associated with disease activity on MRI. *Neurology,* 73**,** 32-8.

FESTEN, E. A., GOYETTE, P., GREEN, T., BOUCHER, G., BEAUCHAMP, C., TRYNKA, G., DUBOIS, P. C., LAGACE, C., STOKKERS, P. C., HOMMES, D. W., BARISANI, D., PALMIERI, O., ANNESE, V., VAN HEEL, D. A., WEERSMA, R. K., DALY, M. J., WIJMENGA, C. & RIOUX, J. D. 2011. A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. *PLoS Genet,* 7**,** e1001283.

FIRE, A., XU, S., MONTGOMERY, M. K., KOSTAS, S. A., DRIVER, S. E. & MELLO, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature,* 391**,** 806-11.

FOGDELL-HAHN, A., LIGERS, A., GRONNING, M., HILLERT, J. & OLERUP, O. 2000. Multiple sclerosis: a modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. *Tissue Antigens,* 55**,** 140-8.

FORABOSCO, P., BOUZIGON, E., NG, M. Y., HERMANOWSKI, J., FISHER, S. A., CRISWELL, L. A. & LEWIS, C. M. 2009. Meta-analysis of genome-wide linkage studies across autoimmune diseases. *Eur J Hum Genet,* 17**,** 236-43.

FORBES, R. B., WILSON, S. V. & 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. & WILLIAMS, D. R. 1998. The prevalence of multiple sclerosis in the Leeds Health Authority. *J Neurol Neurosurg Psychiatry,* 64**,** 605-10.

FRANCALACCI, P., MORELLI, L., UNDERHILL, P. A., LILLIE, A. S., PASSARINO, G., USELI, A., MADEDDU, R., PAOLI, G., TOFANELLI, S., CALO, C. M., GHIANI, M. E., VARESI, L., MEMMI, M., VONA, G., LIN, A. A., OEFNER, P. & CAVALLI-SFORZA, L. L. 2003. Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. *Am J Phys Anthropol,* 121**,** 270-9.

FRANKE, A., BALSCHUN, T., KARLSEN, T. H., HEDDERICH, J., MAY, S., LU, T., SCHULDT, D., NIKOLAUS, S., ROSENSTIEL, P., KRAWCZAK, M. & SCHREIBER, S. 2008a. Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nat Genet,* 40**,** 713-5.

FRANKE, A., BALSCHUN, T., KARLSEN, T. H., SVENTORAITYTE, J., NIKOLAUS, S., MAYR, G., DOMINGUES, F. S., ALBRECHT, M., NOTHNAGEL, M., ELLINGHAUS, D., SINA, C., ONNIE, C. M., WEERSMA, R. K., STOKKERS, P. C., WIJMENGA, C., GAZOULI,

M., STRACHAN, D., MCARDLE, W. L., VERMEIRE, S., RUTGEERTS, P., ROSENSTIEL, P., KRAWCZAK, M., VATN, M. H., MATHEW, C. G. & SCHREIBER, S. 2008b. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet,* 40**,** 1319-23.

FRANKE, A., MCGOVERN, D. P., BARRETT, J. C., WANG, K., RADFORD-SMITH, G. L., AHMAD, T., LEES, C. W., BALSCHUN, T., LEE, J., ROBERTS, R., ANDERSON, C. A., BIS, J. C., BUMPSTEAD, S., ELLINGHAUS, D., FESTEN, E. M., GEORGES, M., GREEN, T., HARITUNIANS, T., JOSTINS, L., LATIANO, A., MATHEW, C. G., MONTGOMERY, G. W., PRESCOTT, N. J., RAYCHAUDHURI, S., ROTTER, J. I., SCHUMM, P., SHARMA, Y., SIMMS, L. A., TAYLOR, K. D., WHITEMAN, D., WIJMENGA, C., BALDASSANO, R. N., BARCLAY, M., BAYLESS, T. M., BRAND, S., BUNING, C., COHEN, A., COLOMBEL, J. F., COTTONE, M., STRONATI, L., DENSON, T., DE VOS, M., D'INCA, R., DUBINSKY, M., EDWARDS, C., FLORIN, T., FRANCHIMONT, D., GEARRY, R., GLAS, J., VAN GOSSUM, A., GUTHERY, S. L., HALFVARSON, J., VERSPAGET, H. W., HUGOT, J. P., KARBAN, A., LAUKENS, D., LAWRANCE, I., LEMANN, M., LEVINE, A., LIBIOULLE, C., LOUIS, E., MOWAT, C., NEWMAN, W., PANES, J., PHILLIPS, A., PROCTOR, D. D., REGUEIRO, M., RUSSELL, R., RUTGEERTS, P., SANDERSON, J., SANS, M., SEIBOLD, F., STEINHART, A. H., STOKKERS, P. C., TORKVIST, L., KULLAK-UBLICK, G., WILSON, D., WALTERS, T., TARGAN, S. R., BRANT, S. R., RIOUX, J. D., D'AMATO, M., WEERSMA, R. K., KUGATHASAN, S., GRIFFITHS, A. M., MANSFIELD, J. C., VERMEIRE, S., DUERR, R. H., SILVERBERG, M. S., SATSANGI, J., SCHREIBER, S., CHO, J. H., ANNESE, V., HAKONARSON, H., DALY, M. J. & PARKES, M. 2010. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet,* 42**,** 1118-25.

FREEDMAN, D. M., DOSEMECI, M. & ALAVANJA, M. C. 2000. Mortality from multiple sclerosis and exposure to residential and occupational solar radiation: a case-control study based on death certificates. *Occup Environ Med,* 57**,** 418-21.

FRIEDRICHSEN, B. N., GALSGAARD, E. D., NIELSEN, J. H. & MOLDRUP, A. 2001. Growth hormone- and prolactin-induced proliferation of insulinoma cells, INS-1, depends on activation of STAT5 (signal transducer and activator of transcription 5). *Mol Endocrinol,* 15**,** 136-48.

FRY, A. M., JONES, L. A., KRUISBEEK, A. M. & MATIS, L. A. 1989. Thymic requirement for clonal deletion during T cell development. *Science,* 246**,** 1044-6.

GABRIEL, S. B., SCHAFFNER, S. F., NGUYEN, H., MOORE, J. M., ROY, J., BLUMENSTIEL, B., HIGGINS, J., DEFELICE, M., LOCHNER, A., FAGGART, M., LIU-CORDERO, S. N., ROTIMI, C., ADEYEMO, A., COOPER, R., WARD, R., LANDER, E. S., DALY, M. J. & ALTSHULER, D. 2002. The structure of haplotype blocks in the human genome. *Science,* 296**,** 2225-9.

GALTON, F. 1876. The History of Twins, as a Criterion of the Relative Powers of Nature and Nurture. *The Journal of the Anthropological Institute of Great Britain and Ireland,* 5**,** 391-406.

GATEVA, V., SANDLING, J. K., HOM, G., TAYLOR, K. E., CHUNG, S. A., SUN, X., ORTMANN, W., KOSOY, R., FERREIRA, R. C., NORDMARK, G., GUNNARSSON, I., SVENUNGSSON, E., PADYUKOV, L., STURFELT, G., JONSEN, A., BENGTSSON, A. A., RANTAPAA-DAHLQVIST, S., BAECHLER, E. C., BROWN, E. E., ALARCON, G. S., EDBERG, J. C., RAMSEY-GOLDMAN, R., MCGWIN, G., JR., REVEILLE, J. D., VILA, L. M., KIMBERLY, R. P., MANZI, S., PETRI, M. A., LEE, A., GREGERSEN, P. K., SELDIN, M. F., RONNBLOM, L., CRISWELL, L. A., SYVANEN, A. C., BEHRENS, T. W. & GRAHAM, R. R. 2009. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet,* 41**,** 1228-33.

GELFAND, J. M., WEINSTEIN, R., PORTER, S. B., NEIMANN, A. L., BERLIN, J. A. & MARGOLIS, D. J. 2005. Prevalence and treatment of psoriasis in the United Kingdom: a population-based study. *Arch Dermatol,* 141**,** 1537-41.

GERRITSE, K., LAMAN, J. D., NOELLE, R. J., ARUFFO, A., LEDBETTER, J. A., BOERSMA, W. J. & CLAASSEN, E. 1996. CD40-CD40 ligand interactions in experimental allergic encephalomyelitis and multiple sclerosis. *Proc Natl Acad Sci U S A,* 93**,** 2499-504.

GERVIN, K., VIGELAND, M. D., MATTINGSDAL, M., HAMMERO, M., NYGARD, H., OLSEN, A. O., BRANDT, I., HARRIS, J. R., UNDLIEN, D. E. & LYLE, R. 2012. DNA methylation and gene expression changes in monozygotic twins discordant for psoriasis: identification of epigenetically dysregulated genes. *PLoS Genet,* 8**,** e1002454.

GIMENEZ, M. A., SIM, J., ARCHAMBAULT, A. S., KLEIN, R. S. & RUSSELL, J. H. 2006. A tumor necrosis factor receptor 1-dependent conversation between central nervous system-specific T cells and the central nervous system is required for inflammatory infiltration of the spinal cord. *Am J Pathol,* 168**,** 1200-9.

GIRARD, S. L., GAUTHIER, J., NOREAU, A., XIONG, L., ZHOU, S., JOUAN, L., DIONNE-LAPORTE, A., SPIEGELMAN, D., HENRION, E., DIALLO, O., THIBODEAU, P., BACHAND, I., BAO, J. Y., TONG, A. H., LIN, C. H., MILLET, B., JAAFARI, N., JOOBER, R., DION, P. A., LOK, S., KREBS, M. O. & ROULEAU, G. A. 2011. Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat Genet,* 43**,** 860-3.

GOODKIN, D. E., DOOLITTLE, T. H., HAUSER, S. S., RANSOHOFF, R. M., ROSES, A. D. & RUDICK, R. A. 1991. Diagnostic criteria for multiple sclerosis research involving multiply affected families. *Arch Neurol,* 48**,** 805-7.

GOURRAUD, P. A., MCELROY, J. P., CAILLIER, S. J., JOHNSON, B. A., SANTANIELLO, A., HAUSER, S. L. & OKSENBERG, J. R. 2011. Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. *Ann Neurol,* 69**,** 65-74.

GRAHAM, R. R., COTSAPAS, C., DAVIES, L., HACKETT, R., LESSARD, C. J., LEON, J. M., BURTT, N. P., GUIDUCCI, C., PARKIN, M., GATES, C., PLENGE, R. M., BEHRENS, T. W., WITHER, J. E., RIOUX, J. D., FORTIN, P. R., GRAHAM, D. C., WONG, A. K., VYSE, T. J., DALY, M. J., ALTSHULER, D., MOSER, K. L. & GAFFNEY, P. M. 2008. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat Genet,* 40**,** 1059-61.

GRAHAM, R. R., HOM, G., ORTMANN, W. & BEHRENS, T. W. 2009. Review of recent genome-wide association scans in lupus. *J Intern Med,* 265**,** 680-8.

GRANIERI, E., ECONOMOU, N. T., DE GENNARO, R., TOLA, M. R., CANIATTI, L., GOVONI, V., FAINARDI, E. & CASETTA, I. 2007. Multiple sclerosis in the province of Ferrara : evidence for an increasing trend. *J Neurol,* 254**,** 1642-8.

GRAY, O. M., MCDONNELL, G. V. & HAWKINS, S. A. 2008. Factors in the rising prevalence of multiple sclerosis in the north-east of Ireland. *Mult Scler,* 14**,** 880-6.

GREGERSEN, P. K., AMOS, C. I., LEE, A. T., LU, Y., REMMERS, E. F., KASTNER, D. L., SELDIN, M. F., CRISWELL, L. A., PLENGE, R. M., HOLERS, V. M., MIKULS, T. R., SOKKA, T., MORELAND, L. W., BRIDGES, S. L., JR., XIE, G., BEGOVICH, A. B. & SIMINOVITCH, K. A. 2009. REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nat Genet,* 41**,** 820-3.

GRIMALDI, L. M., SALEMI, G., GRIMALDI, G., RIZZO, A., MARZIOLO, R., LO PRESTI, C., MAIMONE, D. & SAVETTIERI, G. 2001. High incidence and increasing prevalence of MS in Enna (Sicily), southern Italy. *Neurology,* 57**,** 1891-3.

GROUP, T. L. M. S. S. G. A. T. U. O. B. C. M. M. A. 1999. TNF neutralization in MS: results of a randomized, placebo-controlled multicenter study. The Lenercept Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group. *Neurology,* 53**,** 457-65.

GRYTTEN, N., GLAD, S. B., AARSETH, J. H., NYLAND, H., MIDGARD, R. & MYHR, K. M. 2006. A 50-year follow-up of the incidence of multiple sclerosis in Hordaland County, Norway. *Neurology,* 66**,** 182-6.

HAAHR, S., KOCH-HENRIKSEN, N., MOLLER-LARSEN, A., ERIKSEN, L. S. & ANDERSEN, H. M. 1995. Increased risk of multiple sclerosis after late Epstein-Barr virus infection: a historical prospective study. *Mult Scler,* 1**,** 73-7.

HAINES, J. L., BRADFORD, Y., GARCIA, M. E., REED, A. D., NEUMEISTER, E., PERICAK-VANCE, M. A., RIMMLER, J. B., MENOLD, M. M., MARTIN, E. R., OKSENBERG, J. R., BARCELLOS, L. F., LINCOLN, R. & HAUSER, S. L. 2002. Multiple susceptibility loci for multiple sclerosis. *Hum Mol Genet,* 11**,** 2251-6.

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. & POLYCHRONAKOS, C. 2007. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. *Nature,* 448**,** 591-4.

HAMMOND, S. R., ENGLISH, D., DE WYTT, C., MAXWELL, I. C., MILLINGEN, K. S., STEWART-WYNNE, E. G., MCLEOD, J. G. & MCCALL, M. G. 1988. The clinical profile of MS in Australia: a comparison between medium- and high-frequency prevalence zones. *Neurology,* 38**,** 980-6.

HAMMOND, S. R., ENGLISH, D. R. & MCLEOD, J. G. 2000. The age-range of risk of developing multiple sclerosis: evidence from a migrant population in Australia. *Brain,* 123 (Pt 5)**,** 968-74.

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., 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., LOW, H. Q., SHEN, M., ZHAI, Z. F., WANG, Y., ZHANG, F. Y., YANG, S., LIU, J. J. & 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.

HANSEN, T., SKYTTHE, A., STENAGER, E., PETERSEN, H. C., BRONNUM-HANSEN, H. & KYVIK, K. O. 2005. Concordance for multiple sclerosis in Danish twins: an update of a nationwide study. *Mult Scler,* 11**,** 504-10.

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., SORENSEN, P. S., SANDBERG-WOLLHEIM, M., ANDERSEN, O., SVEJGAARD, A., EDLAND, A., MELLGREN, S. I., COMPSTON, A., VARTDAL, F. & 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.

HARDELID, P., CORTINA-BORJA, M., MUNRO, A., JONES, H., CLEARY, M., CHAMPION, M. P., FOO, Y., SCRIVER, C. R. & DEZATEUX, C. 2008. The birth prevalence of PKU in populations of European, South Asian and sub-Saharan African ancestry living in South East England. *Ann Hum Genet,* 72**,** 65-71.

HARLEY, J. B., ALARCON-RIQUELME, M. E., CRISWELL, L. A., JACOB, C. O., KIMBERLY, R. P., MOSER, K. L., TSAO, B. P., VYSE, T. J., LANGEFELD, C. D., NATH, S. K., GUTHRIDGE, J. M., COBB, B. L., MIREL, D. B., MARION, M. C., WILLIAMS, A. H., DIVERS, J., WANG, W., FRANK, S. G., NAMJOU, B., GABRIEL, S. B., LEE, A. T., GREGERSEN, P. K., BEHRENS, T. W., TAYLOR, K. E., FERNANDO, M., ZIDOVETZKI, R., GAFFNEY, P. M., EDBERG, J. C., RIOUX, J. D., OJWANG, J. O., JAMES, J. A., MERRILL, J. T., GILKESON, G. S., SELDIN, M. F., YIN, H., BAECHLER,

E. C., LI, Q. Z., WAKELAND, E. K., BRUNER, G. R., KAUFMAN, K. M. & KELLY, J. A. 2008. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. *Nat Genet,* 40**,** 204-10.

HARRINGTON, L. E., HATTON, R. D., MANGAN, P. R., TURNER, H., MURPHY, T. L., MURPHY, K. M. & 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.

HARROW, J., FRANKISH, A., GONZALEZ, J. M., TAPANARI, E., DIEKHANS, M., KOKOCINSKI, F., AKEN, B. L., BARRELL, D., ZADISSA, A., SEARLE, S., BARNES, I., BIGNELL, A., BOYCHENKO, V., HUNT, T., KAY, M., MUKHERJEE, G., RAJAN, J., DESPACIO-REYES, G., SAUNDERS, G., STEWARD, C., HARTE, R., LIN, M., HOWALD, C., TANZER, A., DERRIEN, T., CHRAST, J., WALTERS, N., BALASUBRAMANIAN, S., PEI, B., TRESS, M., RODRIGUEZ, J. M., EZKURDIA, I., VAN BAREN, J., BRENT, M., HAUSSLER, D., KELLIS, M., VALENCIA, A., REYMOND, A., GERSTEIN, M., GUIGO, R. & HUBBARD, T. J. 2012. GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res,* 22**,** 1760-74.

HARTUNG, H. P., GONSETTE, R., KONIG, N., KWIECINSKI, H., GUSEO, A., MORRISSEY, S. P., KRAPF, H., ZWINGERS, T. & MITOXANTRONE IN MULTIPLE SCLEROSIS STUDY, G. 2002. Mitoxantrone in progressive multiple sclerosis: a placebo-controlled, double-blind, randomised, multicentre trial. *Lancet,* 360**,** 2018-25.

HAWKES, C. H. & MACGREGOR, A. J. 2009. Twin studies and the heritability of MS: a conclusion. *Mult Scler,* 15**,** 661-7.

HE, Y., YU, Y., ZHANG, Y., SONG, J., MITRA, A., WANG, Y., SUN, D. & ZHANG, S. 2012. Genome-Wide Bovine H3K27me3 Modifications and the Regulatory Effects on Genes Expressions in Peripheral Blood Lymphocytes. *PLoS One,* 7**,** e39094.

HEALY, B. C., ALI, E. N., GUTTMANN, C. R., CHITNIS, T., GLANZ, B. I., BUCKLE, G., HOUTCHENS, M., STAZZONE, L., MOODIE, J., BERGER, A. M., DUAN, Y., BAKSHI, R., KHOURY, S., WEINER, H. & ASCHERIO, A. 2009. Smoking and disease progression in multiple sclerosis. *Arch Neurol,* 66**,** 858-64.

HELMS, C., SACCONE, N. L., CAO, L., DAW, J. A., CAO, K., HSU, T. M., TAILLON-MILLER, P., DUAN, S., GORDON, D., PIERCE, B., OTT, J., RICE, J., FERNANDEZ-VINA, M. A., KWOK, P. Y., MENTER, A. & BOWCOCK, A. M. 2005. Localization of PSORS1 to a haplotype block harboring HLA-C and distinct from corneodesmosin and HCR. *Hum Genet,* 118**,** 466-76.

HERRMANN, M., VOLL, R. E., ZOLLER, O. M., HAGENHOFER, M., PONNER, B. B. & KALDEN, J. R. 1998. Impaired phagocytosis of apoptotic cell material by monocyte-derived macrophages from patients with systemic lupus erythematosus. *Arthritis Rheum,* 41**,** 1241-50.

HICKEY, W. F. & KIMURA, H. 1988. Perivascular microglial cells of the CNS are bone marrow-derived and present antigen in vivo. *Science,* 239**,** 290-2.

HIEKKALINNA, T., SCHAFFER, A. A., LAMBERT, B., NORRGRANN, P., GORING, H. H. & TERWILLIGER, J. D. 2011. PSEUDOMARKER: a powerful program for joint linkage and/or linkage disequilibrium analysis on mixtures of singletons and related individuals. *Hum Hered,* 71**,** 256-66.

HODES, R. J., SHARROW, S. O. & SOLOMON, A. 1989. Failure of T cell receptor V beta negative selection in an athymic environment. *Science,* 246**,** 1041-4.

HOGLUND, P., MINTERN, J., WALTZINGER, C., HEATH, W., BENOIST, C. & MATHIS, D. 1999. Initiation of autoimmune diabetes by developmentally regulated presentation of islet cell antigens in the pancreatic lymph nodes. *J Exp Med,* 189**,** 331-9.

HOLLAND, S. M., DELEO, F. R., ELLOUMI, H. Z., HSU, A. P., UZEL, G., BRODSKY, N., FREEMAN, A. F., DEMIDOWICH, A., DAVIS, J., TURNER, M. L., ANDERSON, V. L., DARNELL, D. N., WELCH, P. A., KUHNS, D. B., FRUCHT, D. M., MALECH, H. L.,

GALLIN, J. I., KOBAYASHI, S. D., WHITNEY, A. R., VOYICH, J. M., MUSSER, J. M., WOELLNER, C., SCHAFFER, A. A., PUCK, J. M. & GRIMBACHER, B. 2007. STAT3 mutations in the hyper-IgE syndrome. *N Engl J Med,* 357**,** 1608-19.

HOLMGREN, G., LARSSON, A., PALMSTIERNA, H. & ALM, J. 1976. The frequency of PKU and hyperphenylalaninemia in Sweden - a study in institutions for the mentally retarded as well as in neonates. *Clin Genet,* 10**,** 313-8.

HORTON, R., GIBSON, R., COGGILL, P., MIRETTI, M., ALLCOCK, R. J., ALMEIDA, J., FORBES, S., GILBERT, J. G., HALLS, K., HARROW, J. L., HART, E., HOWE, K., JACKSON, D. K., PALMER, S., ROBERTS, A. N., SIMS, S., STEWART, C. A., TRAHERNE, J. A., TREVANION, S., WILMING, L., ROGERS, J., DE JONG, P. J., ELLIOTT, J. F., SAWCER, S., TODD, J. A., TROWSDALE, J. & BECK, S. 2008. Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. *Immunogenetics,* 60**,** 1-18.

HOWALD, C., TANZER, A., CHRAST, J., KOKOCINSKI, F., DERRIEN, T., WALTERS, N., GONZALEZ, J. M., FRANKISH, A., AKEN, B. L., HOURLIER, T., VOGEL, J. H., WHITE, S., SEARLE, S., HARROW, J., HUBBARD, T. J., GUIGO, R. & REYMOND, A. 2012. Combining RT-PCR-seq and RNA-seq to catalog all genic elements encoded in the human genome. *Genome Res,* 22**,** 1698-710.

HOVATTA, I., TERWILLIGER, J. D., LICHTERMANN, D., MAKIKYRO, T., SUVISAARI, J., PELTONEN, L. & LONNQVIST, J. 1997. Schizophrenia in the genetic isolate of Finland. *Am J Med Genet,* 74**,** 353-60.

HUITINGA, I., RUULS, S. R., JUNG, S., VAN ROOIJEN, N., HARTUNG, H. P. & DIJKSTRA, C. D. 1995. Macrophages in T cell line-mediated, demyelinating, and chronic relapsing experimental autoimmune encephalomyelitis in Lewis rats. *Clin Exp Immunol,* 100**,** 344-51.

HUNTINGTON'S DISEASE COLLABORATIVE RESEARCH GROUP 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell,* 72**,** 971-83.

HUSEBY, E. S., LIGGITT, D., BRABB, T., SCHNABEL, B., OHLEN, C. & GOVERMAN, J. 2001. A pathogenic role for myelin-specific CD8(+) T cells in a model for multiple sclerosis. *J Exp Med,* 194**,** 669-76.

HUYGHE, J. R., FRANSEN, E., HANNULA, S., VAN LAER, L., VAN EYKEN, E., MAKI-TORKKO, E., AIKIO, P., SORRI, M., HUENTELMAN, M. J. & VAN CAMP, G. 2010a. A genome-wide analysis of population structure in the Finnish Saami with implications for genetic association studies. *Eur J Hum Genet,* 19**,** 347-52.

HUYGHE, J. R., FRANSEN, E., HANNULA, S., VAN LAER, L., VAN EYKEN, E., MAKI-TORKKO, E., LYSHOLM-BERNACCHI, A., AIKIO, P., STEPHAN, D. A., SORRI, M., HUENTELMAN, M. J. & VAN CAMP, G. 2010b. Genome-wide SNP analysis reveals no gain in power for association studies of common variants in the Finnish Saami. *Eur J Hum Genet,* 18**,** 569-74.

IMATAKI, O., ANSÉN, S., TANAKA, M., BUTLER, M. O., BEREZOVSKAYA, A., MILSTEIN, M. I., KUZUSHIMA, K., NADLER, L. M. & N., H. 2012. IL-21 can supplement suboptimal Lck-independent MAPK activation in a STAT-3-dependent manner in human CD8(+) T cells. *J Immunol.,* 188**,** 1609-19.

IMSGC, I. M. S. G. C. 2007. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med,* 357**,** 851-62.

IMSGC, I. M. S. G. C. 2011. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature,* 476**,** 214-9.

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 HUMAN GENOME SEQUENCING CONSORTIUM 2004. Finishing the euchromatic sequence of the human genome. *Nature,* 431**,** 931-45.

ISLAM, T., GAUDERMAN, W. J., COZEN, W., HAMILTON, A. S., BURNETT, M. E. & MACK, T. M. 2006. Differential twin concordance for multiple sclerosis by latitude of birthplace. *Ann Neurol,* 60**,** 56-64.

ISLAM, T., GAUDERMAN, W. J., COZEN, W. & MACK, T. M. 2007. Childhood sun exposure influences risk of multiple sclerosis in monozygotic twins. *Neurology,* 69**,** 381-8.

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. & 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.

JAKKULA, E., REHNSTROM, K., VARILO, T., PIETILAINEN, O. P., PAUNIO, T., PEDERSEN, N. L., DEFAIRE, U., JARVELIN, M. R., SAHARINEN, J., FREIMER, N., RIPATTI, S., PURCELL, S., COLLINS, A., DALY, M. J., PALOTIE, A. & PELTONEN, L. 2008. The genome-wide patterns of variation expose significant substructure in a founder population. *Am J Hum Genet,* 83**,** 787-94.

JANG, E., CHO, W. S., CHO, M. L., PARK, H. J., OH, H. J., KANG, S. M., PAIK, D. J. & YOUN, J. 2011. Foxp3+ regulatory T cells control humoral autoimmunity by suppressing the development of long-lived plasma cells. *J Immunol,* 186**,** 1546-53.

JERSILD, C., SVEJGAARD, A. & FOG, T. 1972. HL-A antigens and multiple sclerosis. *Lancet,* 1**,** 1240-1.

JOHNSON, A. E., GORDON, C., PALMER, R. G. & BACON, P. A. 1995. The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. Relationship to ethnicity and country of birth. *Arthritis Rheum,* 38**,** 551-8.

JOHNSON, G. C., ESPOSITO, L., BARRATT, B. J., SMITH, A. N., HEWARD, J., DI GENOVA, G., UEDA, H., CORDELL, H. J., EAVES, I. A., DUDBRIDGE, F., TWELLS, R. C., PAYNE, F., HUGHES, W., NUTLAND, S., STEVENS, H., CARR, P., TUOMILEHTO-WOLF, E., TUOMILEHTO, J., GOUGH, S. C., CLAYTON, D. G. & TODD, J. A. 2001. Haplotype tagging for the identification of common disease genes. *Nat Genet,* 29**,** 233-7.

JOHNSON, K. P., ARRIGO, S. C., NELSON, B. J. & GINSBERG, A. 1977. Agarose electrophoresis of cerebrospinal fluid in multiple sclerosis. A simplified method for demonstrating cerebrospinal fluid oligoclonal immunoglobulin bands. *Neurology,* 27**,** 273-7.

JOHNSON, S. R., SCHENTAG, C. T. & GLADMAN, D. D. 2005. Autoantibodies in biological agent naive patients with psoriatic arthritis. *Ann Rheum Dis,* 64**,** 770-2.

JONES, R. E., MASS, M. & BOURDETTE, D. N. 1999. Myelin basic protein-specific T lymphocytes induce chronic relapsing experimental autoimmune encephalomyelitis in lymphocyte-deficient (SCID) mice. *J Neuroimmunol,* 93**,** 92-101.

KAMAT, S. S., PEPMUELLER, P. H. & MOORE, T. L. 2003. Triplets with systemic lupus erythematosus. *Arthritis Rheum,* 48**,** 3176-80.

KAPPELMAN, M. D., RIFAS-SHIMAN, S. L., KLEINMAN, K., OLLENDORF, D., BOUSVAROS, A., GRAND, R. J. & FINKELSTEIN, J. A. 2007. The prevalence and geographic distribution of Crohn's disease and ulcerative colitis in the United States. *Clin Gastroenterol Hepatol,* 5**,** 1424-9.

KAPRIO, J., PULKKINEN, L. & ROSE, R. J. 2002. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Res,* 5**,** 366-71.

KAZEEM, G. R. & FARRALL, M. 2005. Integrating case-control and TDT studies. *Ann Hum Genet,* 69**,** 329-35.

KEBIR, H., KREYMBORG, K., IFERGAN, I., DODELET-DEVILLERS, A., CAYROL, R., BERNARD, M., GIULIANI, F., ARBOUR, N., BECHER, B. & PRAT, A. 2007. Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. *Nat Med,* 13**,** 1173-5.

KEINAN, A. & CLARK, A. G. 2012. Recent Explosive Human Population Growth Has Resulted in an Excess of Rare Genetic Variants. *Science,* 336**,** 740-3.

KENYON, K. D., COLE, C., CRAWFORD, F., KAPPLER, J. W., THURMAN, J. M., BRATTON, D. L., BOACKLE, S. A. & HENSON, P. M. 2011. IgG autoantibodies against deposited C3 inhibit macrophage-mediated apoptotic cell engulfment in systemic autoimmunity. *J Immunol,* 187**,** 2101-11.

KILPINEN, H., YLISAUKKO-OJA, T., REHNSTROM, K., GAAL, E., TURUNEN, J. A., KEMPAS, E., VON WENDT, L., VARILO, T. & PELTONEN, L. 2009. Linkage and linkage disequilibrium scan for autism loci in an extended pedigree from Finland. *Hum Mol Genet,* 18**,** 2912-21.

KIM, J. S., LEE, K. S., PARK, J. H., KIM, M. Y. & SHIN, W. S. 2000. Detection of human herpesvirus 6 variant A in peripheral blood mononuclear cells from multiple sclerosis patients. *Eur Neurol,* 43**,** 170-3.

KOCH-HENRIKSEN, N., BRONNUM-HANSEN, H. & STENAGER, E. 1998. Underlying cause of death in Danish patients with multiple sclerosis: results from the Danish Multiple Sclerosis Registry. *J Neurol Neurosurg Psychiatry,* 65**,** 56-9.

KOSKELA, H. L., ELDFORS, S., ELLONEN, P., VAN ADRICHEM, A. J., KUUSANMAKI, H., ANDERSSON, E. I., LAGSTROM, S., CLEMENTE, M. J., OLSON, T., JALKANEN, S. E., MAJUMDER, M. M., ALMUSA, H., EDGREN, H., LEPISTO, M., MATTILA, P., GUINTA, K., KOISTINEN, P., KUITTINEN, T., PENTTINEN, K., PARSONS, A., KNOWLES, J., SAARELA, J., WENNERBERG, K., KALLIONIEMI, O., PORKKA, K., LOUGHRAN, T. P., JR., HECKMAN, C. A., MACIEJEWSKI, J. P. & MUSTJOKI, S. 2012. Somatic STAT3 mutations in large granular lymphocytic leukemia. *N Engl J Med,* 366**,** 1905-13.

KOWAL, C., DEGIORGIO, L. A., LEE, J. Y., EDGAR, M. A., HUERTA, P. T., VOLPE, B. T. & DIAMOND, B. 2006. Human lupus autoantibodies against NMDA receptors mediate cognitive impairment. *Proc Natl Acad Sci U S A,* 103**,** 19854-9.

KREMER, E. J., PRITCHARD, M., LYNCH, M., YU, S., HOLMAN, K., BAKER, E., WARREN, S. T., SCHLESSINGER, D., SUTHERLAND, G. R. & RICHARDS, R. I. 1991. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science,* 252**,** 1711-4.

KRUGLYAK, L. 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet,* 22**,** 139-44.

KUHLMANN, T., LINGFELD, G., BITSCH, A., SCHUCHARDT, J. & BRUCK, W. 2002. Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time. *Brain,* 125**,** 2202-12.

KUMPFEL, T., HOFFMANN, L. A., RUBSAMEN, H., POLLMANN, W., FENEBERG, W., HOHLFELD, R. & LOHSE, P. 2007. Late-onset tumor necrosis factor receptor-associated periodic syndrome in multiple sclerosis patients carrying the TNFRSF1A R92Q mutation. *Arthritis Rheum,* 56**,** 2774-83.

KURD, S. K. & GELFAND, J. M. 2009. The prevalence of previously diagnosed and undiagnosed psoriasis in US adults: results from NHANES 2003-2004. *J Am Acad Dermatol,* 60**,** 218-24.

KURREEMAN, F. A., STAHL, E. A., OKADA, Y., LIAO, K., DIOGO, D., RAYCHAUDHURI, S., FREUDENBERG, J., KOCHI, Y., PATSOPOULOS, N. A., GUPTA, N., SANDOR, C., BANG, S. Y., LEE, H. S., PADYUKOV, L., SUZUKI, A., SIMINOVITCH, K., WORTHINGTON, J., GREGERSEN, P. K., HUGHES, L. B., REYNOLDS, R. J., BRIDGES, S. L., JR., BAE, S. C., YAMAMOTO, K. & PLENGE, R. M. 2012. Use of a multiethnic approach to identify rheumatoid- arthritis-susceptibility loci, 1p36 and 17q12. *Am J Hum Genet,* 90**,** 524-32.

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

KURTZKE, J. F., BEEBE, G. W. & NORMAN, J. E., JR. 1979. Epidemiology of multiple sclerosis in U.S. veterans: 1. Race, sex, and geographic distribution. *Neurology,* 29**,** 1228-35.

KURTZKE, J. F. & HELTBERG, A. 2001. Multiple sclerosis in the Faroe Islands: an epitome. *J Clin Epidemiol,* 54**,** 1-22.

KWON, H. K., LEE, C. G., SO, J. S., CHAE, C. S., HWANG, J. S., SAHOO, A., NAM, J. H., RHEE, J. H., HWANG, K. C. & IM, S. H. 2010. Generation of regulatory dendritic cells and CD4+Foxp3+ T cells by probiotics administration suppresses immune disorders. *Proc Natl Acad Sci U S A,* 107**,** 2159-64.

KYOGOKU, C., LANGEFELD, C. D., ORTMANN, W. A., LEE, A., SELBY, S., CARLTON, V. E., CHANG, M., RAMOS, P., BAECHLER, E. C., BATLIWALLA, F. M., NOVITZKE, J., WILLIAMS, A. H., GILLETT, C., RODINE, P., GRAHAM, R. R., ARDLIE, K. G., GAFFNEY, P. M., MOSER, K. L., PETRI, M., BEGOVICH, A. B., GREGERSEN, P. K. & BEHRENS, T. W. 2004. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am J Hum Genet,* 75**,** 504-7.

LAHTELA, J., NOUSIAINEN, H. O., STEFANOVIC, V., TALLILA, J., VISKARI, H., KARIKOSKI, R., GENTILE, M., SALORANTA, C., VARILO, T., SALONEN, R. & KESTILA, M. 2010. Mutant CHUK and severe fetal encasement malformation. *N Engl J Med,* 363**,** 1631-7.

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.

LANGRISH, C. L., CHEN, Y., BLUMENSCHEIN, W. M., MATTSON, J., BASHAM, B., SEDGWICK, J. D., MCCLANAHAN, T., KASTELEIN, R. A. & CUA, D. J. 2005. IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J Exp Med,* 201**,** 233-40.

LAPIDUS, A., BERNELL, O., HELLERS, G. & LOFBERG, R. 1998. Clinical course of colorectal Crohn's disease: a 35-year follow-up study of 507 patients. *Gastroenterology,* 114**,** 1151-60.

LEMAITRE, B., NICOLAS, E., MICHAUT, L., REICHHART, J. M. & HOFFMANN, J. A. 1996. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell,* 86**,** 973-83.

LEPAULT, F. & GAGNERAULT, M. C. 2000. Characterization of peripheral regulatory CD4+ T cells that prevent diabetes onset in nonobese diabetic mice. *J Immunol,* 164**,** 240-7.

LI, J., HSU, H. C., YANG, P., WU, Q., LI, H., EDGINGTON, L. E., BOGYO, M., KIMBERLY, R. P. & MOUNTZ, J. D. 2012a. Treatment of arthritis by macrophage depletion and immunomodulation: testing an apoptosis-mediated therapy in a humanized death receptor mouse model. *Arthritis Rheum,* 64**,** 1098-109.

LI, S., WANG, N. & BRODT, P. 2012b. Metastatic cells can escape the proapoptotic effects of TNF-α through increased autocrine IL-6/STAT3 signaling. *Cancer Res,* 74**,** 865-75.

LIEW, S. H., ELSNER, H., SPECTOR, T. & HAMMOND, C. J. 2005. The first "classical" twin study? Analysis of refractive error using monozygotic and dizygotic twins published in 1922. *Twin Res Hum Genet,* 8**,** 198-200.

LILL, C. M., SCHJEIDE, B. M., AKKAD, D. A., BLASCHKE, P., WINKELMANN, A., GERDES, L. A., HOFFJAN, S., LUESSI, F., DORNER, T., LI, S. C., STEINHAGEN-THIESSEN, E., LINDENBERGER, U., CHAN, A., HARTUNG, H. P., AKTAS, O., LOHSE, P., KUMPFEL, T., KUBISCH, C., EPPLEN, J. T., ZETTL, U. K., BERTRAM, L. & ZIPP, F. 2012. Independent replication of STAT3 association with multiple sclerosis risk in a large German case-control sample. *Neurogenetics,* 13**,** 83-6.

LINDBLAD-TOH, K., WADE, C. M., MIKKELSEN, T. S., KARLSSON, E. K., JAFFE, D. B., KAMAL, M., CLAMP, M., CHANG, J. L., KULBOKAS, E. J., 3RD, ZODY, M. C., MAUCELI, E., XIE, X., BREEN, M., WAYNE, R. K., OSTRANDER, E. A., PONTING, C. P., GALIBERT, F., SMITH, D. R., DEJONG, P. J., KIRKNESS, E., ALVAREZ, P., BIAGI, T., BROCKMAN, W., BUTLER, J., CHIN, C. W., COOK, A., CUFF, J., DALY, M. J., DECAPRIO, D., GNERRE, S., GRABHERR, M., KELLIS, M., KLEBER, M., BARDELEBEN, C., GOODSTADT, L., HEGER, A., HITTE, C., KIM, L., KOEPFLI, K. P., PARKER, H. G., POLLINGER, J. P., SEARLE, S. M., SUTTER, N. B., THOMAS, R., WEBBER, C., BALDWIN, J., ABEBE, A., ABOUELLEIL, A., AFTUCK, L., AIT-ZAHRA, M., ALDREDGE, T., ALLEN, N., AN, P., ANDERSON, S., ANTOINE, C., ARACHCHI, H., ASLAM, A., AYOTTE, L., BACHANTSANG, P., BARRY, A., BAYUL, T., BENAMARA, M., BERLIN, A., BESSETTE, D., BLITSHTEYN, B., BLOOM, T., BLYE, J., BOGUSLAVSKIY, L., BONNET, C., BOUKHGALTER, B., BROWN, A., CAHILL, P., CALIXTE, N., CAMARATA, J., CHESHATSANG, Y., CHU, J., CITROEN, M., COLLYMORE, A., COOKE, P., DAWOE, T., DAZA, R., DECKTOR, K., DEGRAY, S., DHARGAY, N., DOOLEY, K., DORJE, P., DORJEE, K., DORRIS, L., DUFFEY, N., DUPES, A., EGBIREMOLEN, O., ELONG, R., FALK, J., FARINA, A., FARO, S., FERGUSON, D., FERREIRA, P., FISHER, S., FITZGERALD, M., FOLEY, K., et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature,* 438**,** 803-19.

LIU, X., LEUNG, S., WANG, C., TAN, Z., WANG, J., GUO, T. B., FANG, L., ZHAO, Y., WAN, B., QIN, X., LU, L., LI, R., PAN, H., SONG, M., LIU, A., HONG, J., LU, H. & ZHANG, J. Z. 2010. Crucial role of interleukin-7 in T helper type 17 survival and expansion in autoimmune disease. *Nat Med,* 16**,** 191-7.

LOVATO, P., BRENDER, C., AGNHOLT, J., KELSEN, J., KALTOFT, K., SVEJGAARD, A., ERIKSEN, K. W., WOETMANN, A. & ODUM, N. 2003. Constitutive STAT3 activation in intestinal T cells from patients with Crohn's disease. *J Biol Chem,* 278**,** 16777-81.

LUBLIN, F. D. & REINGOLD, S. C. 1996. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology,* 46**,** 907-11.

LUCAS, R. M., PONSONBY, A. L., DEAR, K., VALERY, P. C., PENDER, M. P., TAYLOR, B. V., KILPATRICK, T. J., DWYER, T., COULTHARD, A., CHAPMAN, C., VAN DER MEI, I., WILLIAMS, D. & MCMICHAEL, A. J. 2011. Sun exposure and vitamin D are independent risk factors for CNS demyelination. *Neurology,* 76**,** 540-8.

LUCCHINETTI, C., BRUCK, W., PARISI, J., SCHEITHAUER, B., RODRIGUEZ, M. & LASSMANN, H. 2000. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Ann Neurol,* 47**,** 707-17.

LUCCHINETTI, C. F., POPESCU, B. F., BUNYAN, R. F., MOLL, N. M., ROEMER, S. F., LASSMANN, H., BRUCK, W., PARISI, J. E., SCHEITHAUER, B. W., GIANNINI, C., WEIGAND, S. D., MANDREKAR, J. & RANSOHOFF, R. M. 2011. Inflammatory cortical demyelination in early multiple sclerosis. *N Engl J Med,* 365**,** 2188-97.

MANIATIS, N., COLLINS, A., XU, C. F., MCCARTHY, L. C., HEWETT, D. R., TAPPER, W., ENNIS, S., KE, X. & MORTON, N. E. 2002. The first linkage disequilibrium (LD) maps: delineation of hot and cold blocks by diplotype analysis. *Proc Natl Acad Sci U S A,* 99**,** 2228-33.

MANOLIO, T. A., COLLINS, F. S., COX, N. J., GOLDSTEIN, D. B., HINDORFF, L. A., HUNTER, D. J., MCCARTHY, M. I., RAMOS, E. M., CARDON, L. R., CHAKRAVARTI, A., CHO, J. H., GUTTMACHER, A. E., KONG, A., KRUGLYAK, L., MARDIS, E., ROTIMI, C. N., SLATKIN, M., VALLE, D., WHITTEMORE, A. S., BOEHNKE, M., CLARK, A. G., EICHLER, E. E., GIBSON, G., HAINES, J. L., MACKAY, T. F., MCCARROLL, S. A. & VISSCHER, P. M. 2009. Finding the missing heritability of complex diseases. *Nature,* 461**,** 747-53.

MANTEL, N. 1963. Chi-Square Tests with One Degree of Freedom; Extensions of the Mantel-Haenszel Procedure. *J Am Stat Assoc,* 58**,** 690-700.

MAO, M., BIERY, M. C., KOBAYASHI, S. V., WARD, T., SCHIMMACK, G., BURCHARD, J., SCHELTER, J. M., DAI, H., HE, Y. D. & LINSLEY, P. S. 2004. T lymphocyte activation gene identification by coregulated expression on DNA microarrays. *Genomics,* 83**,** 989-99.

MARCHINI, J., HOWIE, B., MYERS, S., MCVEAN, G. & DONNELLY, P. 2007. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet,* 39**,** 906-13.

MARRIE, R. A., WOLFSON, C., STURKENBOOM, M. C., GOUT, O., HEINZLEF, O., ROULLET, E. & ABENHAIM, L. 2000. Multiple sclerosis and antecedent infections: a case-control study. *Neurology,* 54**,** 2307-10.

MARTINEZ, R. J., ZHANG, N., THOMAS, S. R., NANDIWADA, S. L., JENKINS, M. K., BINSTADT, B. A. & MUELLER, D. L. 2012. Arthritogenic self-reactive CD4+ T cells acquire an FR4hiCD73hi anergic state in the presence of Foxp3+ regulatory T cells. *J Immunol,* 188**,** 170-81.

MATUSEVICIUS, D., KIVISAKK, P., HE, B., KOSTULAS, N., OZENCI, V., FREDRIKSON, S. & LINK, H. 1999. Interleukin-17 mRNA expression in blood and CSF mononuclear cells is augmented in multiple sclerosis. *Mult Scler,* 5**,** 101-4.

MAURANO, M. T., HUMBERT, R., RYNES, E., THURMAN, R. E., HAUGEN, E., WANG, H., REYNOLDS, A. P., SANDSTROM, R., QU, H., BRODY, J., SHAFER, A., NERI, F., LEE, K., KUTYAVIN, T., STEHLING-SUN, S., JOHNSON, A. K., CANFIELD, T. K., GISTE, E., DIEGEL, M., BATES, D., HANSEN, R. S., NEPH, S., SABO, P. J., HEIMFELD, S., RAUBITSCHEK, A., ZIEGLER, S., COTSAPAS, C., SOTOODEHNIA, N., GLASS, I., SUNYAEV, S. R., KAUL, R. & STAMATOYANNOPOULOS, J. A. 2012. Systematic localization of common disease-associated variation in regulatory DNA. *Science,* 337**,** 1190-5.

MCCOMBE, P. A., CHALK, J. B. & PENDER, M. P. 1990. Familial occurrence of multiple sclerosis with thyroid disease and systemic lupus erythematosus. *J Neurol Sci,* 97**,** 163-71.

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. & 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.

MCGEACHY, M. J., CHEN, Y., TATO, C. M., LAURENCE, A., JOYCE-SHAIKH, B., BLUMENSCHEIN, W. M., MCCLANAHAN, T. K., O'SHEA, J. J. & CUA, D. J. 2009. The interleukin 23 receptor is essential for the terminal differentiation of interleukin 17-producing effector T helper cells in vivo. *Nat Immunol,* 10**,** 314-24.

MCGOVERN, D. P., GARDET, A., TORKVIST, L., GOYETTE, P., ESSERS, J., TAYLOR, K. D., NEALE, B. M., ONG, R. T., LAGACE, C., LI, C., GREEN, T., STEVENS, C. R., BEAUCHAMP, C., FLESHNER, P. R., CARLSON, M., D'AMATO, M., HALFVARSON, J., HIBBERD, M. L., LORDAL, M., PADYUKOV, L., ANDRIULLI, A., COLOMBO, E., LATIANO, A., PALMIERI, O., BERNARD, E. J., DESLANDRES, C., HOMMES, D. W., DE JONG, D. J., STOKKERS, P. C., WEERSMA, R. K., SHARMA, Y., SILVERBERG, M. S., CHO, J. H., WU, J., ROEDER, K., BRANT, S. R., SCHUMM, L. P., DUERR, R. H., DUBINSKY, M. C., GLAZER, N. L., HARITUNIANS, T., IPPOLITI, A., MELMED, G. Y., SISCOVICK, D. S., VASILIAUSKAS, E. A., TARGAN, S. R., ANNESE, V.,

WIJMENGA, C., PETTERSSON, S., ROTTER, J. I., XAVIER, R. J., DALY, M. J., RIOUX, J. D. & SEIELSTAD, M. 2010. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet,* 42**,** 332-7.

MCGUIGAN, C., MCCARTHY, A., QUIGLEY, C., BANNAN, L., HAWKINS, S. A. & 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.

MCLAUGHLIN, K. A. & WUCHERPFENNIG, K. W. 2008. B cells and autoantibodies in the pathogenesis of multiple sclerosis and related inflammatory demyelinating diseases. *Adv Immunol,* 98**,** 121-49.

MCLOUGHLIN, R. M., JENKINS, B. J., GRAIL, D., WILLIAMS, A. S., FIELDING, C. A., PARKER, C. R., ERNST, M., TOPLEY, N. & JONES, S. A. 2005. IL-6 trans-signaling via STAT3 directs T cell infiltration in acute inflammation. *Proc Natl Acad Sci U S A,* 102**,** 9589-94.

MEDANA, I. M., GALLIMORE, A., OXENIUS, A., MARTINIC, M. M., WEKERLE, H. & NEUMANN, H. 2000. MHC class I-restricted killing of neurons by virus-specific CD8+ T lymphocytes is effected through the Fas/FasL, but not the perforin pathway. *Eur J Immunol,* 30**,** 3623-33.

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. & HARBO, H. F. 2009. A rare variant of the TYK2 gene is confirmed to be associated with multiple sclerosis. *Eur J Hum Genet,* 18**,** 502-4.

MERRIMAN, C. 1924. The Intellectual Resemblance of Twins. *Psychological Monographs: General and Applied,* 33**,** 1-57.

MICHEAU, O. & TSCHOPP, J. 2003. Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. *Cell,* 114**,** 181-90.

MIDGARD, R., GRONNING, M., RIISE, T., KVALE, G. & NYLAND, H. 1996. Multiple sclerosis and chronic inflammatory diseases. A case-control study. *Acta Neurol Scand,* 93**,** 322-8.

MINEGISHI, Y., SAITO, M., MORIO, T., WATANABE, K., AGEMATSU, K., TSUCHIYA, S., TAKADA, H., HARA, T., KAWAMURA, N., ARIGA, T., KANEKO, H., KONDO, N., TSUGE, I., YACHIE, A., SAKIYAMA, Y., IWATA, T., BESSHO, F., OHISHI, T., JOH, K., IMAI, K., KOGAWA, K., SHINOHARA, M., FUJIEDA, M., WAKIGUCHI, H., PASIC, S., ABINUN, M., OCHS, H. D., RENNER, E. D., JANSSON, A., BELOHRADSKY, B. H., METIN, A., SHIMIZU, N., MIZUTANI, S., MIYAWAKI, T., NONOYAMA, S. & KARASUYAMA, H. 2006. Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity. *Immunity,* 25**,** 745-55.

MINEGISHI, Y., SAITO, M., TSUCHIYA, S., TSUGE, I., TAKADA, H., HARA, T., KAWAMURA, N., ARIGA, T., PASIC, S., STOJKOVIC, O., METIN, A. & KARASUYAMA, H. 2007. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. *Nature,* 448**,** 1058-62.

MORGAN, M. E., FLIERMAN, R., VAN DUIVENVOORDE, L. M., WITTEVEEN, H. J., VAN EWIJK, W., VAN LAAR, J. M., DE VRIES, R. R. & TOES, R. E. 2005. Effective treatment of collagen-induced arthritis by adoptive transfer of CD25+ regulatory T cells. *Arthritis Rheum,* 52**,** 2212-21.

MORTON, N. E. 1955. Sequential tests for the detection of linkage. *Am J Hum Genet,* 7**,** 277-318.

MUMFORD, C. J., WOOD, N. W., KELLAR-WOOD, H., THORPE, J. W., MILLER, D. H. & COMPSTON, D. A. 1994. The British Isles survey of multiple sclerosis in twins. *Neurology,* 44**,** 11-5.

MUNGER, K. L., LEVIN, L. I., HOLLIS, B. W., HOWARD, N. S. & ASCHERIO, A. 2006. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *JAMA,* 296**,** 2832-8.

MUNGER, K. L., ZHANG, S. M., O'REILLY, E., HERNAN, M. A., OLEK, M. J., WILLETT, W. C. & ASCHERIO, A. 2004. Vitamin D intake and incidence of multiple sclerosis. *Neurology,* 62**,** 60-5.

MURPHY, C. A., LANGRISH, C. L., CHEN, Y., BLUMENSCHEIN, W., MCCLANAHAN, T., KASTELEIN, R. A., SEDGWICK, J. D. & CUA, D. J. 2003. Divergent pro- and antiinflammatory roles for IL-23 and IL-12 in joint autoimmune inflammation. *J Exp Med,* 198**,** 1951-7.

MYCKO, M. P., CICHALEWSKA, M., MACHLANSKA, A., CWIKLINSKA, H., MARIASIEWICZ, M. & SELMAJ, K. W. 2012. microRNA-301a regulation of a T-helper 17 immune response controls autoimmune demyelination. *Proc Natl Acad Sci U S A*.

NACHMAN, M. W. & CROWELL, S. L. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics,* 156**,** 297-304.

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. 2009. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat Genet,* 41**,** 199-204.

NAKAE, S., NAMBU, A., SUDO, K. & IWAKURA, Y. 2003. Suppression of immune induction of collagen-induced arthritis in IL-17-deficient mice. *J Immunol,* 171**,** 6173-7.

NAKASHIMA, K., YANAGISAWA, M., ARAKAWA, H., KIMURA, N., HISATSUNE, T., KAWABATA, M., MIYAZONO, K. & TAGA, T. 1999. Synergistic signaling in fetal brain by STAT3-Smad1 complex bridged by p300. *Science,* 284**,** 479-82.

NEALE, B. M., KOU, Y., LIU, L., MA'AYAN, A., SAMOCHA, K. E., SABO, A., LIN, C. F., STEVENS, C., WANG, L. S., MAKAROV, V., POLAK, P., YOON, S., MAGUIRE, J., CRAWFORD, E. L., CAMPBELL, N. G., GELLER, E. T., VALLADARES, O., SCHAFER, C., LIU, H., ZHAO, T., CAI, G., LIHM, J., DANNENFELSER, R., JABADO, O., PERALTA, Z., NAGASWAMY, U., MUZNY, D., REID, J. G., NEWSHAM, I., WU, Y., LEWIS, L., HAN, Y., VOIGHT, B. F., LIM, E., ROSSIN, E., KIRBY, A., FLANNICK, J., FROMER, M., SHAKIR, K., FENNELL, T., GARIMELLA, K., BANKS, E., POPLIN, R., GABRIEL, S., DEPRISTO, M., WIMBISH, J. R., BOONE, B. E., LEVY, S. E., BETANCUR, C., SUNYAEV, S., BOERWINKLE, E., BUXBAUM, J. D., COOK, E. H., JR., DEVLIN, B., GIBBS, R. A., ROEDER, K., SCHELLENBERG, G. D., SUTCLIFFE, J. S. & DALY, M. J. 2012. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature,* 485**,** 242-5.

NEJENTSEV, S., HOWSON, J. M., WALKER, N. M., SZESZKO, J., FIELD, S. F., STEVENS, H. E., REYNOLDS, P., HARDY, M., KING, E., MASTERS, J., HULME, J., MAIER, L. M., SMYTH, D., BAILEY, R., COOPER, J. D., RIBAS, G., CAMPBELL, R. D., CLAYTON, D. G. & TODD, J. A. 2007. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature,* 450**,** 887-92.

NEJENTSEV, S., WALKER, N., RICHES, D., EGHOLM, M. & TODD, J. A. 2009. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science,* 324**,** 387-9.

NELIS, M., ESKO, T., MAGI, R., ZIMPRICH, F., ZIMPRICH, A., TONCHEVA, D., KARACHANAK, S., PISKACKOVA, T., BALASCAK, I., PELTONEN, L., JAKKULA, E., REHNSTROM, K., LATHROP, M., HEATH, S., GALAN, P., SCHREIBER, S., MEITINGER, T., PFEUFER, A., WICHMANN, H. E., MELEGH, B., POLGAR, N., TONIOLO, D., GASPARINI, P., D'ADAMO, P., KLOVINS, J., NIKITINA-ZAKE, L., KUCINSKAS, V., KASNAUSKIENE, J., LUBINSKI, J., DEBNIAK, T., LIMBORSKA, S., KHRUNIN, A., ESTIVILL, X., RABIONET, R., MARSAL, S., JULIA, A., ANTONARAKIS, S. E., DEUTSCH, S., BOREL, C., ATTAR, H., GAGNEBIN, M., MACEK, M., KRAWCZAK, M., REMM, M. & METSPALU, A. 2009. Genetic structure of Europeans: a view from the North-East. *PLoS One,* 4**,** e5472.

NELSON, N., MARKS, M. S., DRIGGERS, P. H. & OZATO, K. 1993. Interferon consensus sequence-binding protein, a member of the interferon regulatory factor family, suppresses interferon-induced gene transcription. *Mol Cell Biol,* 13**,** 588-99.

NEMAZEE, D. & BUERKI, K. 1989. Clonal deletion of autoreactive B lymphocytes in bone marrow chimeras. *Proc Natl Acad Sci U S A,* 86**,** 8039-43.

NEMAZEE, D. A. & BURKI, K. 1989. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature,* 337**,** 562-6.

NEVITT, G. J. & HUTCHINSON, P. E. 1996. Psoriasis in the community: prevalence, severity and patients' beliefs and attitudes towards the disease. *Br J Dermatol,* 135**,** 533-7.

NICKOLOFF, B. J. & WRONE-SMITH, T. 1999. Injection of pre-psoriatic skin with CD4+ T cells induces psoriasis. *Am J Pathol,* 155**,** 145-58.

NORIO, R. 2003a. Finnish Disease Heritage I: characteristics, causes, background. *Hum Genet,* 112**,** 441-56.

NORIO, R. 2003b. The Finnish Disease Heritage III: the individual diseases. *Hum Genet,* 112**,** 470-526.

NURIEVA, R., YANG, X. O., MARTINEZ, G., ZHANG, Y., PANOPOULOS, A. D., MA, L., SCHLUNS, K., TIAN, Q., WATOWICH, S. S., JETTEN, A. M. & DONG, C. 2007. Essential autocrine regulation by IL-21 in the generation of inflammatory T cells. *Nature,* 448**,** 480-3.

NYHOLT, D. R. 2004. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet,* 74**,** 765-9.

O'ROAK, B. J., VIVES, L., GIRIRAJAN, S., KARAKOC, E., KRUMM, N., COE, B. P., LEVY, R., KO, A., LEE, C., SMITH, J. D., TURNER, E. H., STANAWAY, I. B., VERNOT, B., MALIG, M., BAKER, C., REILLY, B., AKEY, J. M., BORENSTEIN, E., RIEDER, M. J., NICKERSON, D. A., BERNIER, R., SHENDURE, J. & EICHLER, E. E. 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature,* 485**,** 246-50.

OKADA, Y., TERAO, C., IKARI, K., KOCHI, Y., OHMURA, K., SUZUKI, A., KAWAGUCHI, T., STAHL, E. A., KURREEMAN, F. A., NISHIDA, N., OHMIYA, H., MYOUZEN, K., TAKAHASHI, M., SAWADA, T., NISHIOKA, Y., YUKIOKA, M., MATSUBARA, T., WAKITANI, S., TESHIMA, R., TOHMA, S., TAKASUGI, K., SHIMADA, K., MURASAWA, A., HONJO, S., MATSUO, K., TANAKA, H., TAJIMA, K., SUZUKI, T., IWAMOTO, T., KAWAMURA, Y., TANII, H., OKAZAKI, Y., SASAKI, T., GREGERSEN, P. K., PADYUKOV, L., WORTHINGTON, J., SIMINOVITCH, K. A., LATHROP, M., TANIGUCHI, A., TAKAHASHI, A., TOKUNAGA, K., KUBO, M., NAKAMURA, Y., KAMATANI, N., MIMORI, T., PLENGE, R. M., YAMANAKA, H., MOMOHARA, S., YAMADA, R., MATSUDA, F. & YAMAMOTO, K. 2012. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat Genet,* 44**,** 511-6.

OLERUP, O. & HILLERT, J. 1991. HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. *Tissue Antigens,* 38**,** 1-15.

OWHASHI, M. & HEBER-KATZ, E. 1988. Protection from experimental allergic encephalomyelitis conferred by a monoclonal antibody directed against a shared idiotype on rat T cell receptors specific for myelin basic protein. *J Exp Med,* 168**,** 2153-64.

PALMER, J. P., ASPLIN, C. M., CLEMONS, P., LYEN, K., TATPATI, O., RAGHU, P. K. & PAQUETTE, T. L. 1983. Insulin antibodies in insulin-dependent diabetics before insulin treatment. *Science,* 222**,** 1337-9.

PARKES, M., BARRETT, J. C., PRESCOTT, N. J., TREMELLING, M., ANDERSON, C. A., FISHER, S. A., ROBERTS, R. G., NIMMO, E. R., CUMMINGS, F. R., SOARS, D., DRUMMOND, H., LEES, C. W., KHAWAJA, S. A., BAGNALL, R., BURKE, D. A., TODHUNTER, C. E., AHMAD, T., ONNIE, C. M., MCARDLE, W., STRACHAN, D., BETHEL, G., BRYAN, C., LEWIS, C. M., DELOUKAS, P., FORBES, A., SANDERSON, J., JEWELL, D. P., SATSANGI, J., MANSFIELD, J. C., CARDON, L. & MATHEW, C. G.

2007. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet,* 39**,** 830-2.

PASQUALUCCI, L., TRIFONOV, V., FABBRI, G., MA, J., ROSSI, D., CHIARENZA, A., WELLS, V. A., GRUNN, A., MESSINA, M., ELLIOT, O., CHAN, J., BHAGAT, G., CHADBURN, A., GAIDANO, G., MULLIGHAN, C. G., RABADAN, R. & DALLA-FAVERA, R. 2011. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet,* 43**,** 830-7.

PASQUINELLI, A. E., REINHART, B. J., SLACK, F., MARTINDALE, M. Q., KURODA, M. I., MALLER, B., HAYWARD, D. C., BALL, E. E., DEGNAN, B., MULLER, P., SPRING, J., SRINIVASAN, A., FISHMAN, M., FINNERTY, J., CORBO, J., LEVINE, M., LEAHY, P., DAVIDSON, E. & RUVKUN, G. 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature,* 408**,** 86-9.

PASTINEN, T., PEROLA, M., IGNATIUS, J., SABATTI, C., TAINOLA, P., LEVANDER, M., SYVANEN, A. C. & PELTONEN, L. 2001. Dissecting a population genome for targeted screening of disease mutations. *Hum Mol Genet,* 10**,** 2961-72.

PATSOPOULOS, N. A., ESPOSITO, F., REISCHL, J., LEHR, S., BAUER, D., HEUBACH, J., SANDBRINK, R., POHL, C., EDAN, G., KAPPOS, L., MILLER, D., MONTALBAN, J., POLMAN, C. H., FREEDMAN, M. S., HARTUNG, H. P., ARNASON, B. G., COMI, G., COOK, S., FILIPPI, M., GOODIN, D. S., JEFFERY, D., O'CONNOR, P., EBERS, G. C., LANGDON, D., REDER, A. T., TRABOULSEE, A., ZIPP, F., SCHIMRIGK, S., HILLERT, J., BAHLO, M., BOOTH, D. R., BROADLEY, S., BROWN, M. A., BROWNING, B. L., BROWNING, S. R., BUTZKUEVEN, H., CARROLL, W. M., CHAPMAN, C., FOOTE, S. J., GRIFFITHS, L., KERMODE, A. G., KILPATRICK, T. J., LECHNER-SCOTT, J., MARRIOTT, M., MASON, D., MOSCATO, P., HEARD, R. N., PENDER, M. P., PERREAU, V. M., PERERA, D., RUBIO, J. P., SCOTT, R. J., SLEE, M., STANKOVICH, J., STEWART, G. J., TAYLOR, B. V., TUBRIDY, N., WILLOUGHBY, E., WILEY, J., MATTHEWS, P., BONESCHI, F. M., COMPSTON, A., HAINES, J., HAUSER, S. L., MCCAULEY, J., IVINSON, A., OKSENBERG, J. R., PERICAK-VANCE, M., SAWCER, S. J., DE JAGER, P. L., HAFLER, D. A. & DE BAKKER, P. I. 2011. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol,* 70**,** 897-912.

PERTTILA, J., MERIKANTO, K., NAUKKARINEN, J., SURAKKA, I., MARTIN, N. W., TANHUANPAA, K., GRIMARD, V., TASKINEN, M. R., THIELE, C., SALOMAA, V., JULA, A., PEROLA, M., VIRTANEN, I., PELTONEN, L. & OLKKONEN, V. M. 2009. OSBPL10, a novel candidate gene for high triglyceride trait in dyslipidemic Finnish subjects, regulates cellular lipid metabolism. *J Mol Med (Berl),* 87**,** 825-35.

PETERSON, A. C., DI RIENZO, A., LEHESJOKI, A. E., DE LA CHAPELLE, A., SLATKIN, M. & FREIMER, N. B. 1995. The distribution of linkage disequilibrium over anonymous genome regions. *Hum Mol Genet,* 4**,** 887-94.

PETRI, M., GENOVESE, M., ENGLE, E. & HOCHBERG, M. 1991. Definition, incidence, and clinical description of flare in systemic lupus erythematosus. A prospective cohort study. *Arthritis Rheum,* 34**,** 937-44.

PICKRELL, J. K., COOP, G., NOVEMBRE, J., KUDARAVALLI, S., LI, J. Z., ABSHER, D., SRINIVASAN, B. S., BARSH, G. S., MYERS, R. M., FELDMAN, M. W. & PRITCHARD, J. K. 2009. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res,* 19**,** 826-37.

PINTO, D., PAGNAMENTA, A. T., KLEI, L., ANNEY, R., MERICO, D., REGAN, R., CONROY, J., MAGALHAES, T. R., CORREIA, C., ABRAHAMS, B. S., ALMEIDA, J., BACCHELLI, E., BADER, G. D., BAILEY, A. J., BAIRD, G., BATTAGLIA, A., BERNEY, T., BOLSHAKOVA, N., BOLTE, S., BOLTON, P. F., BOURGERON, T., BRENNAN, S., BRIAN, J., BRYSON, S. E., CARSON, A. R., CASALLO, G., CASEY, J., CHUNG, B. H., COCHRANE, L., CORSELLO, C., CRAWFORD, E. L., CROSSETT, A., CYTRYNBAUM, C., DAWSON, G., DE JONGE, M., DELORME, R., DRMIC, I., DUKETIS, E., DUQUE, F., ESTES, A., FARRAR, P., FERNANDEZ, B. A., FOLSTEIN, S. E., FOMBONNE, E., FREITAG, C. M., GILBERT, J., GILLBERG, C., GLESSNER, J. T.,

GOLDBERG, J., GREEN, A., GREEN, J., GUTER, S. J., HAKONARSON, H., HERON, E. A., HILL, M., HOLT, R., HOWE, J. L., HUGHES, G., HUS, V., IGLIOZZI, R., KIM, C., KLAUCK, S. M., KOLEVZON, A., KORVATSKA, O., KUSTANOVICH, V., LAJONCHERE, C. M., LAMB, J. A., LASKAWIEC, M., LEBOYER, M., LE COUTEUR, A., LEVENTHAL, B. L., LIONEL, A. C., LIU, X. Q., LORD, C., LOTSPEICH, L., LUND, S. C., MAESTRINI, E., MAHONEY, W., MANTOULAN, C., MARSHALL, C. R., MCCONACHIE, H., MCDOUGLE, C. J., MCGRATH, J., MCMAHON, W. M., MERIKANGAS, A., MIGITA, O., MINSHEW, N. J., MIRZA, G. K., MUNSON, J., NELSON, S. F., NOAKES, C., NOOR, A., NYGREN, G., OLIVEIRA, G., PAPANIKOLAOU, K., PARR, J. R., PARRINI, B., PATON, T., PICKLES, A., PILORGE, M., et al. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature,* 466**,** 368-72.

PISKIN, G., SYLVA-STEENLAND, R. M., BOS, J. D. & TEUNISSEN, M. B. 2006. In vitro and in situ expression of IL-23 by keratinocytes in healthy skin and psoriasis lesions: enhanced expression in psoriatic skin. *J Immunol,* 176**,** 1908-15.

POLLAK, V. E. 1964. Antinuclear Antibodies in Families of Patients with Systemic Lupus Erythematosus. *N Engl J Med,* 271**,** 165-71.

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. & WOLINSKY, J. S. 2005. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol,* 58**,** 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. & TOURTELLOTTE, W. W. 1983. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol,* 13**,** 227-31.

PRICE, A. L., HELGASON, A., PALSSON, S., STEFANSSON, H., ST CLAIR, D., ANDREASSEN, O. A., REICH, D., KONG, A. & STEFANSSON, K. 2009. The impact of divergence time on the nature of population structure: an example from Iceland. *PLoS Genet,* 5**,** e1000505.

PRICE, A. L., PATTERSON, N. J., PLENGE, R. M., WEINBLATT, M. E., SHADICK, N. A. & REICH, D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet,* 38**,** 904-9.

PRISMS STUDY GROUP & THE UNIVERSITY OF BRITISH COLUMBIA, M. S. M. R. I. A. G. 2001. PRISMS-4: Long-term efficacy of interferon-beta-1a in relapsing MS. *Neurology,* 56**,** 1628-36.

PRITCHARD, J. K., STEPHENS, M., ROSENBERG, N. A. & DONNELLY, P. 2000. Association mapping in structured populations. *Am J Hum Genet,* 67**,** 170-81.

PUEL, A., ZIEGLER, S. F., BUCKLEY, R. H. & LEONARD, W. J. 1998. Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency. *Nat Genet,* 20**,** 394-7.

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. & SHAM, P. C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet,* 81**,** 559-75.

QIN, H., YEH, W. I., DE SARNO, P., HOLDBROOKS, A. T., LIU, Y., MULDOWNEY, M. T., REYNOLDS, S. L., YANAGISAWA, L. L., FOX, T. H., 3RD, PARK, K., HARRINGTON, L. E., RAMAN, C. & BENVENISTE, E. N. 2012. Signal transducer and activator of transcription-3/suppressor of cytokine signaling-3 (STAT3/SOCS3) axis in myeloid cells regulates neuroinflammation. *Proc Natl Acad Sci U S A,* 109**,** 5004-9.

R CORE TEAM 2012. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

RAITAKARI, O. T., JUONALA, M., RONNEMAA, T., KELTIKANGAS-JARVINEN, L., RASANEN, L., PIETIKAINEN, M., HUTRI-KAHONEN, N., TAITTONEN, L., JOKINEN,

E., MARNIEMI, J., JULA, A., TELAMA, R., KAHONEN, M., LEHTIMAKI, T., AKERBLOM, H. K. & VIIKARI, J. S. 2008. Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol,* 37**,** 1220-6.

RAMACHANDRAN, S., DESHPANDE, O., ROSEMAN, C. C., ROSENBERG, N. A., FELDMAN, M. W. & CAVALLI-SFORZA, L. L. 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc Natl Acad Sci U S A,* 102**,** 15942-7.

RAMAGOPALAN, S. V., DYMENT, D. A., CADER, M. Z., MORRISON, K. M., DISANTO, G., MORAHAN, J. M., BERLANGA-TAYLOR, A. J., HANDEL, A., DE LUCA, G. C., SADOVNICK, A. D., LEPAGE, P., MONTPETIT, A. & EBERS, G. C. 2011. Rare variants in the CYP27B1 gene are associated with multiple sclerosis. *Ann Neurol,* 70**,** 881-6.

RAMOS, P. S., CRISWELL, L. A., MOSER, K. L., COMEAU, M. E., WILLIAMS, A. H., PAJEWSKI, N. M., CHUNG, S. A., GRAHAM, R. R., ZIDOVETZKI, R., KELLY, J. A., KAUFMAN, K. M., JACOB, C. O., VYSE, T. J., TSAO, B. P., KIMBERLY, R. P., GAFFNEY, P. M., ALARCON-RIQUELME, M. E., HARLEY, J. B. & LANGEFELD, C. D. 2011. A comprehensive analysis of shared loci between systemic lupus erythematosus (SLE) and sixteen autoimmune diseases reveals limited genetic overlap. *PLoS Genet,* 7**,** e1002406.

RAMSDELL, F. & FOWLKES, B. J. 1990. Clonal deletion versus clonal anergy: the role of the thymus in inducing self tolerance. *Science,* 248**,** 1342-8.

RAMSDELL, F., LANTZ, T. & FOWLKES, B. J. 1989. A nondeletional mechanism of thymic self tolerance. *Science,* 246**,** 1038-41.

RANTAPAA-DAHLQVIST, S., DE JONG, B. A., BERGLIN, E., HALLMANS, G., WADELL, G., STENLUND, H., SUNDIN, U. & VAN VENROOIJ, W. J. 2003. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum,* 48**,** 2741-9.

RAYCHAUDHURI, S., REMMERS, E. F., LEE, A. T., HACKETT, R., GUIDUCCI, C., BURTT, N. P., GIANNINY, L., KORMAN, B. D., PADYUKOV, L., KURREEMAN, F. A., CHANG, M., CATANESE, J. J., DING, B., WONG, S., VAN DER HELM-VAN MIL, A. H., NEALE, B. M., COBLYN, J., CUI, J., TAK, P. P., WOLBINK, G. J., CRUSIUS, J. B., VAN DER HORST-BRUINSMA, I. E., CRISWELL, L. A., AMOS, C. I., SELDIN, M. F., KASTNER, D. L., ARDLIE, K. G., ALFREDSSON, L., COSTENBADER, K. H., ALTSHULER, D., HUIZINGA, T. W., SHADICK, N. A., WEINBLATT, M. E., DE VRIES, N., WORTHINGTON, J., SEIELSTAD, M., TOES, R. E., KARLSON, E. W., BEGOVICH, A. B., KLARESKOG, L., GREGERSEN, P. K., DALY, M. J. & PLENGE, R. M. 2008. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. *Nat Genet,* 40**,** 1216-23.

RAYCHAUDHURI, S., SANDOR, C., STAHL, E. A., FREUDENBERG, J., LEE, H. S., JIA, X., ALFREDSSON, L., PADYUKOV, L., KLARESKOG, L., WORTHINGTON, J., SIMINOVITCH, K. A., BAE, S. C., PLENGE, R. M., GREGERSEN, P. K. & DE BAKKER, P. I. 2012. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet,* 44**,** 291-6.

RAYCHAUDHURI, S., THOMSON, B. P., REMMERS, E. F., EYRE, S., HINKS, A., GUIDUCCI, C., CATANESE, J. J., XIE, G., STAHL, E. A., CHEN, R., ALFREDSSON, L., AMOS, C. I., ARDLIE, K. G., BARTON, A., BOWES, J., BURTT, N. P., CHANG, M., COBLYN, J., COSTENBADER, K. H., CRISWELL, L. A., CRUSIUS, J. B., CUI, J., DE JAGER, P. L., DING, B., EMERY, P., FLYNN, E., HARRISON, P., HOCKING, L. J., HUIZINGA, T. W., KASTNER, D. L., KE, X., KURREEMAN, F. A., LEE, A. T., LIU, X., LI, Y., MARTIN, P., MORGAN, A. W., PADYUKOV, L., 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., WEINBLATT, M. E., WILSON, A. G., WOLBINK, G. J., WORDSWORTH, P., ALTSHULER, D., KARLSON, E. W., TOES, R. E., DE VRIES, N., BEGOVICH, A. B., SIMINOVITCH, K. A., WORTHINGTON, J., KLARESKOG, L., GREGERSEN, P. K., DALY, M. J. & PLENGE, R. M. 2009. Genetic

variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. *Nat Genet,* 41**,** 1313-8.

RAYCHAUDHURI, S. P., NAVARE, T., GROSS, J. & RAYCHAUDHURI, S. K. 2003. Clinical course of psoriasis during pregnancy. *Int J Dermatol,* 42**,** 518-20.

REDON, R., ISHIKAWA, S., FITCH, K. R., FEUK, L., PERRY, G. H., ANDREWS, T. D., FIEGLER, H., SHAPERO, M. H., CARSON, A. R., CHEN, W., CHO, E. K., DALLAIRE, S., FREEMAN, J. L., GONZALEZ, J. R., GRATACOS, M., HUANG, J., KALAITZOPOULOS, D., KOMURA, D., MACDONALD, J. R., MARSHALL, C. R., MEI, R., MONTGOMERY, L., NISHIMURA, K., OKAMURA, K., SHEN, F., SOMERVILLE, M. J., TCHINDA, J., VALSESIA, A., WOODWARK, C., YANG, F., ZHANG, J., ZERJAL, T., ARMENGOL, L., CONRAD, D. F., ESTIVILL, X., TYLER-SMITH, C., CARTER, N. P., ABURATANI, H., LEE, C., JONES, K. W., SCHERER, S. W. & HURLES, M. E. 2006. Global variation in copy number in the human genome. *Nature,* 444**,** 444-54.

REINDL, M., LININGTON, C., BREHM, U., EGG, R., DILITZ, E., DEISENHAMMER, F., POEWE, W. & BERGER, T. 1999. Antibodies against the myelin oligodendrocyte glycoprotein and the myelin basic protein in multiple sclerosis and other neurological diseases: a comparative study. *Brain,* 122 (Pt 11)**,** 2047-56.

REINHART, B. J., SLACK, F. J., BASSON, M., PASQUINELLI, A. E., BETTINGER, J. C., ROUGVIE, A. E., HORVITZ, H. R. & RUVKUN, G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature,* 403**,** 901-6.

RIISE, T., NORTVEDT, M. W. & ASCHERIO, A. 2003. Smoking is a risk factor for multiple sclerosis. *Neurology,* 61**,** 1122-4.

RIMOLDI, M., CHIEPPA, M., SALUCCI, V., AVOGADRI, F., SONZOGNI, A., SAMPIETRO, G. M., NESPOLI, A., VIALE, G., ALLAVENA, P. & RESCIGNO, M. 2005. Intestinal immune homeostasis is regulated by the crosstalk between epithelial cells and dendritic cells. *Nat Immunol,* 6**,** 507-14.

RISCH, N. & MERIKANGAS, K. 1996. The future of genetic studies of complex human diseases. *Science,* 273**,** 1516-7.

RISCH, N. J. 2000. Searching for genetic determinants in the new millennium. *Nature,* 405**,** 847-56.

ROCK, F. L., HARDIMAN, G., TIMANS, J. C., KASTELEIN, R. A. & BAZAN, J. F. 1998. A family of human receptors structurally related to Drosophila Toll. *Proc Natl Acad Sci U S A,* 95**,** 588-93.

ROSENBERG, N. A., PRITCHARD, J. K., WEBER, J. L., CANN, H. M., KIDD, K. K., ZHIVOTOVSKY, L. A. & FELDMAN, M. W. 2002. Genetic structure of human populations. *Science,* 298**,** 2381-5.

ROSS-INNES, C. S., STARK, R., TESCHENDORFF, A. E., HOLMES, K. A., ALI, H. R., DUNNING, M. J., BROWN, G. D., GOJIS, O., ELLIS, I. O., GREEN, A. R., ALI, S., CHIN, S. F., PALMIERI, C., CALDAS, C. & CARROLL, J. S. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature,* 481**,** 389-93.

ROTHWELL, P. M. & 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. & COMPSTON, D. A. 2005. Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. *Neurology,* 64**,** 1144-51.

RUAN, Q., KAMESWARAN, V., TONE, Y., LI, L., LIOU, H. C., GREENE, M. I., TONE, M. & CHEN, Y. H. 2009. Development of Foxp3(+) regulatory t cells is driven by the c-Rel enhanceosome. *Immunity,* 31**,** 932-40.

RYAN, J. G. & AKSENTIJEVICH, I. 2009. Tumor necrosis factor receptor-associated periodic syndrome: toward a molecular understanding of the systemic autoinflammatory diseases. *Arthritis Rheum,* 60**,** 8-11.

SADOVNICK, A. D., EBERS, G. C., DYMENT, D. A. & RISCH, N. J. 1996. Evidence for genetic basis of multiple sclerosis. The Canadian Collaborative Study Group. *Lancet,* 347**,** 1728-30.

SADOVNICK, A. D., EBERS, G. C., WILSON, R. W. & PATY, D. W. 1992. Life expectancy in patients attending multiple sclerosis clinics. *Neurology,* 42**,** 991-4.

SANDERS, S. J., MURTHA, M. T., GUPTA, A. R., MURDOCH, J. D., RAUBESON, M. J., WILLSEY, A. J., ERCAN-SENCICEK, A. G., DILULLO, N. M., PARIKSHAK, N. N., STEIN, J. L., WALKER, M. F., OBER, G. T., TERAN, N. A., SONG, Y., EL-FISHAWY, P., MURTHA, R. C., CHOI, M., OVERTON, J. D., BJORNSON, R. D., CARRIERO, N. J., MEYER, K. A., BILGUVAR, K., MANE, S. M., SESTAN, N., LIFTON, R. P., GUNEL, M., ROEDER, K., GESCHWIND, D. H., DEVLIN, B. & STATE, M. W. 2012. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature,* 485**,** 237-41.

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. & CUCCA, F. 2010. Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nat Genet,* 42**,** 495-7.

SANO, S., CHAN, K. S., CARBAJAL, S., CLIFFORD, J., PEAVEY, M., KIGUCHI, K., ITAMI, S., NICKOLOFF, B. J. & DIGIOVANNI, J. 2005. Stat3 links activated keratinocytes and immunocytes required for development of psoriasis in a novel transgenic mouse model. *Nat Med,* 11**,** 43-9.

SANTAVUORI, P., RAPOLA, J., SAINIO, K. & RAITTA, C. 1982. A variant of Jansky-Bielschowsky disease. *Neuropediatrics,* 13**,** 135-41.

SARASOJA, T., WIKSTROM, J., PALTAMAA, J., HAKAMA, M. & 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.

SAWCER, S., JONES, H. B., FEAKES, R., GRAY, J., SMALDON, N., CHATAWAY, J., ROBERTSON, N., CLAYTON, D., GOODFELLOW, P. N. & COMPSTON, A. 1996. A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nat Genet,* 13**,** 464-8.

SAVUKOSKI, M., KESTILA, M., WILLIAMS, R., JARVELA, I., SHARP, J., HARRIS, J., SANTAVUORI, P., GARDINER, M. & PELTONEN, L. 1994. Defined chromosomal assignment of CLN5 demonstrates that at least four genetic loci are involved in the pathogenesis of human ceroid lipofuscinoses. *Am J Hum Genet,* 55**,** 695-701.

SEBAT, J., LAKSHMI, B., TROGE, J., ALEXANDER, J., YOUNG, J., LUNDIN, P., MANER, S., MASSA, H., WALKER, M., CHI, M., NAVIN, N., LUCITO, R., HEALY, J., HICKS, J., YE, K., REINER, A., GILLIAM, T. C., TRASK, B., PATTERSON, N., ZETTERBERG, A. & WIGLER, M. 2004. Large-scale copy number polymorphism in the human genome. *Science,* 305**,** 525-8.

SERVICE, S., DEYOUNG, J., KARAYIORGOU, M., ROOS, J. L., PRETORIOUS, H., BEDOYA, G., OSPINA, J., RUIZ-LINARES, A., MACEDO, A., PALHA, J. A., HEUTINK, P., AULCHENKO, Y., OOSTRA, B., VAN DUIJN, C., JARVELIN, M. R., VARILO, T., PEDDLE, L., RAHMAN, P., PIRAS, G., MONNE, M., MURRAY, S., GALVER, L.,

PELTONEN, L., SABATTI, C., COLLINS, A. & FREIMER, N. 2006. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat Genet,* 38**,** 556-60.

SERVIN, B. & STEPHENS, M. 2007. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet,* 3**,** e114.

SHARFE, N., DADI, H. K., SHAHAR, M. & ROIFMAN, C. M. 1997. Human immune disorder arising from mutation of the alpha chain of the interleukin-2 receptor. *Proc Natl Acad Sci U S A,* 94**,** 3168-71.

SIEMENS, H. W. 1924. *Die Zwillingspathologie: Ihre Bedeutung, ihre Methodik, ihre bisherigen Ergebnisse (Twin Pathology: Its Importance, Its Methodology, Its Previous Results),* Berlin, Springer.

SIEPEL, A., BEJERANO, G., PEDERSEN, J. S., HINRICHS, A. S., HOU, M., ROSENBLOOM, K., CLAWSON, H., SPIETH, J., HILLIER, L. W., RICHARDS, S., WEINSTOCK, G. M., WILSON, R. K., GIBBS, R. A., KENT, W. J., MILLER, W. & HAUSSLER, D. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res,* 15**,** 1034-50.

SIGURDSSON, S., NORDMARK, G., GORING, H. H., LINDROOS, K., WIMAN, A. C., STURFELT, G., JONSEN, A., RANTAPAA-DAHLQVIST, S., MOLLER, B., KERE, J., KOSKENMIES, S., WIDEN, E., ELORANTA, M. L., JULKUNEN, H., KRISTJANSDOTTIR, H., STEINSSON, K., ALM, G., RONNBLOM, L. & SYVANEN, A. C. 2005. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am J Hum Genet,* 76**,** 528-37.

SILVERBERG, M. S., CHO, J. H., RIOUX, J. D., MCGOVERN, D. P., WU, J., ANNESE, V., ACHKAR, J. P., GOYETTE, P., SCOTT, R., XU, W., BARMADA, M. M., KLEI, L., DALY, M. J., ABRAHAM, C., BAYLESS, T. M., BOSSA, F., GRIFFITHS, A. M., IPPOLITI, A. F., LAHAIE, R. G., LATIANO, A., PARE, P., PROCTOR, D. D., REGUEIRO, M. D., STEINHART, A. H., TARGAN, S. R., SCHUMM, L. P., KISTNER, E. O., LEE, A. T., GREGERSEN, P. K., ROTTER, J. I., BRANT, S. R., TAYLOR, K. D., ROEDER, K. & DUERR, R. H. 2009. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet,* 41**,** 216-20.

SINGER, N. G., FOX, D. A., HAQQI, T. M., BERETTA, L., ENDRES, J. S., PROHASKA, S., PARNES, J. R., BROMBERG, J. & SRAMKOSKI, R. M. 2002. CD6: expression during development, apoptosis and selection of human and mouse thymocytes. *Int Immunol,* 14**,** 585-97.

SLAGER, S. L., HUANG, J. & VIELAND, V. J. 2000. Effect of allelic heterogeneity on the power of the transmission disequilibrium test. *Genet Epidemiol,* 18**,** 143-56.

SMYTH, D. J., PLAGNOL, V., WALKER, N. M., COOPER, J. D., DOWNES, K., YANG, J. H., HOWSON, J. M., STEVENS, H., MCMANUS, R., WIJMENGA, C., HEAP, G. A., DUBOIS, P. C., CLAYTON, D. G., HUNT, K. A., VAN HEEL, D. A. & TODD, J. A. 2008. Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N Engl J Med,* 359**,** 2767-77.

SNOW, J. W., ABRAHAM, N., MA, M. C., HERNDIER, B. G., PASTUSZAK, A. W. & GOLDSMITH, M. A. 2003. Loss of tolerance and autoimmunity affecting multiple organs in STAT5A/5B-deficient mice. *J Immunol,* 171**,** 5042-50.

SOBEL, D. O. & CRESWELL, K. 2006. Characterization of anti-islet cytotoxic human T-cell clones from patients with type 1 diabetes mellitus. *Autoimmunity,* 39**,** 323-32.

SOCOLOVSKY, M., FALLON, A. E., WANG, S., BRUGNARA, C. & LODISH, H. F. 1999. Fetal anemia and apoptosis of red cell progenitors in Stat5a-/-5b-/- mice: a direct role for Stat5 in Bcl-X(L) induction. *Cell,* 98**,** 181-91.

SPIELMAN, R. S., MCGINNIS, R. E. & EWENS, W. J. 1993. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet,* 52**,** 506-16.

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., 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., WIJMENGA, C., KARLSON, E. W., TOES, R. E., DE VRIES, N., BEGOVICH, A. B., WORTHINGTON, J., SIMINOVITCH, K. A., GREGERSEN, P. K., KLARESKOG, L. & PLENGE, R. M. 2010. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet,* 42**,** 508-14.

STEIDL, C., SHAH, S. P., WOOLCOCK, B. W., RUI, L., KAWAHARA, M., FARINHA, P., JOHNSON, N. A., ZHAO, Y., TELENIUS, A., NERIAH, S. B., MCPHERSON, A., MEISSNER, B., OKOYE, U. C., DIEPSTRA, A., VAN DEN BERG, A., SUN, M., LEUNG, G., JONES, S. J., CONNORS, J. M., HUNTSMAN, D. G., SAVAGE, K. J., RIMSZA, L. M., HORSMAN, D. E., STAUDT, L. M., STEIDL, U., MARRA, M. A. & GASCOYNE, R. D. 2011. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. *Nature,* 471**,** 377-81.

STRANGE, A., CAPON, F., SPENCER, C. C., KNIGHT, J., WEALE, M. E., ALLEN, M. H., BARTON, A., BAND, G., BELLENGUEZ, C., BERGBOER, J. G., BLACKWELL, J. M., BRAMON, E., BUMPSTEAD, S. J., CASAS, J. P., CORK, M. J., CORVIN, A., DELOUKAS, P., DILTHEY, A., DUNCANSON, A., EDKINS, S., ESTIVILL, X., FITZGERALD, O., FREEMAN, C., GIARDINA, E., GRAY, E., HOFER, A., HUFFMEIER, U., HUNT, S. E., IRVINE, A. D., JANKOWSKI, J., KIRBY, B., LANGFORD, C., LASCORZ, J., LEMAN, J., LESLIE, S., MALLBRIS, L., MARKUS, H. S., MATHEW, C. G., MCLEAN, W. H., MCMANUS, R., MOSSNER, R., MOUTSIANAS, L., NALUAI, A. T., NESTLE, F. O., NOVELLI, G., ONOUFRIADIS, A., PALMER, C. N., PERRICONE, C., PIRINEN, M., PLOMIN, R., POTTER, S. C., PUJOL, R. M., RAUTANEN, A., RIVEIRA-MUNOZ, E., RYAN, A. W., SALMHOFER, W., SAMUELSSON, L., SAWCER, S. J., SCHALKWIJK, J., SMITH, C. H., STAHLE, M., SU, Z., TAZI-AHNINI, R., TRAUPE, H., VISWANATHAN, A. C., WARREN, R. B., WEGER, W., WOLK, K., WOOD, N., WORTHINGTON, J., YOUNG, H. S., ZEEUWEN, P. L., HAYDAY, A., BURDEN, A. D., GRIFFITHS, C. E., KERE, J., REIS, A., MCVEAN, G., EVANS, D. M., BROWN, M. A., BARKER, J. N., PELTONEN, L., DONNELLY, P. & TREMBATH, R. C. 2010. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet,* 42**,** 985-90.

SUBRAMANIAN, S., BOURDETTE, D. N., CORLESS, C., VANDENBARK, A. A., OFFNER, H. & JONES, R. E. 2001. T lymphocytes promote the development of bone marrow-derived APC in the central nervous system. *J Immunol,* 166**,** 370-6.

SUMELAHTI, M. L., TIENARI, P. J., HAKAMA, M. & WIKSTROM, J. 2003. Multiple sclerosis in Finland: incidence trends and differences in relapsing remitting and primary progressive disease courses. *J Neurol Neurosurg Psychiatry,* 74**,** 25-8.

SUMELAHTI, M. L., TIENARI, P. J., WIKSTROM, J., PALO, J. & HAKAMA, M. 2000. Regional and temporal variation in the incidence of multiple sclerosis in Finland 1979-1993. *Neuroepidemiology,* 19**,** 67-75.

SUMELAHTI, M. L., TIENARI, P. J., WIKSTROM, J., PALO, J. & HAKAMA, M. 2001. Increasing prevalence of multiple sclerosis in Finland. *Acta Neurol Scand,* 103**,** 153-8.

SUMELAHTI, M. L., TIENARI, P. J., WIKSTROM, J., SALMINEN, T. M. & HAKAMA, M. 2002. Survival of multiple sclerosis in Finland between 1964 and 1993. *Mult Scler,* 8**,** 350-5.

SUVANNAVEJH, G. C., LEE, H. O., PADILLA, J., DAL CANTO, M. C., BARRETT, T. A. & MILLER, S. D. 2000. Divergent roles for p55 and p75 tumor necrosis factor receptors in the pathogenesis of MOG(35-55)-induced experimental autoimmune encephalomyelitis. *Cell Immunol,* 205**,** 24-33.

SYMMONS, D., TURNER, G., WEBB, R., ASTEN, P., BARRETT, E., LUNT, M., SCOTT, D. & SILMAN, A. 2002. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford),* 41**,** 793-800.

TAGAMI, H., IWATSUKI, K. & YAMADA, M. 1983. Profile of anti-stratum corneum autoantibodies in psoriatic patients. *Arch Dermatol Res,* 275**,** 71-5.

TAKAHASHI, K., PRINZ, M., STAGI, M., CHECHNEVA, O. & NEUMANN, H. 2007. TREM2-transduced myeloid precursors mediate nervous tissue debris clearance and facilitate recovery in an animal model of multiple sclerosis. *PLoS Med,* 4**,** e124.

TAKEBAYASHI, H., OHTSUKI, T., UCHIDA, T., KAWAMOTO, S., OKUBO, K., IKENAKA, K., TAKEICHI, M., CHISAKA, O. & NABESHIMA, Y. 2002. Non-overlapping expression of Olig3 and Olig2 in the embryonic neural tube. *Mech Dev,* 113**,** 169-74.

TALLILA, J., SALONEN, R., KOHLSCHMIDT, N., PELTONEN, L. & KESTILA, M. 2009. Mutation spectrum of Meckel syndrome genes: one group of syndromes or several distinct groups? *Hum Mutat,* 30**,** E813-30.

TAMURA, T., NAGAMURA-INOUE, T., SHMELTZER, Z., KUWATA, T. & OZATO, K. 2000. ICSBP directs bipotential myeloid progenitor cells to differentiate into mature macrophages. *Immunity,* 13**,** 155-65.

THOMPSON, E. A. & NEEL, J. V. 1997. Allelic disequilibrium and allele frequency distribution as a function of social and demographic history. *Am J Hum Genet,* 60**,** 197-204.

THOMSON, W., BARTON, A., KE, X., EYRE, S., HINKS, A., BOWES, J., DONN, R., SYMMONS, D., HIDER, S., BRUCE, I. N., WILSON, A. G., MARINOU, I., MORGAN, A., EMERY, P., CARTER, A., STEER, S., HOCKING, L., REID, D. M., WORDSWORTH, P., HARRISON, P., STRACHAN, D. & WORTHINGTON, J. 2007. Rheumatoid arthritis association at 6q23. *Nat Genet,* 39**,** 1431-3.

THORPE, J. W., MUMFORD, C. J., COMPSTON, D. A., KENDALL, B. E., MACMANUS, D. G., MCDONALD, W. I. & MILLER, D. H. 1994. British Isles survey of multiple sclerosis in twins: MRI. *J Neurol Neurosurg Psychiatry,* 57**,** 491-6.

TIENARI, P. J., SUMELAHTI, M. L., RANTAMAKI, T. & WIKSTROM, J. 2004. Multiple sclerosis in western Finland: evidence for a founder effect. *Clin Neurol Neurosurg,* 106**,** 175-9.

TIENARI, P. J., WIKSTROM, J., KOSKIMIES, S., PARTANEN, J., PALO, J. & PELTONEN, L. 1993. Reappraisal of HLA in multiple sclerosis: close linkage in multiplex families. *Eur J Hum Genet,* 1**,** 257-68.

TISCH, R., YANG, X. D., SINGER, S. M., LIBLAU, R. S., FUGGER, L. & MCDEVITT, H. O. 1993. Immune response to glutamic acid decarboxylase correlates with insulitis in non-obese diabetic mice. *Nature,* 366**,** 72-5.

TISHKOFF, S. A., REED, F. A., FRIEDLAENDER, F. R., EHRET, C., RANCIARO, A., FROMENT, A., HIRBO, J. B., AWOMOYI, A. A., BODO, J. M., DOUMBO, O., IBRAHIM, M., JUMA, A. T., KOTZE, M. J., LEMA, G., MOORE, J. H., MORTENSEN, H., NYAMBO, T. B., OMAR, S. A., POWELL, K., PRETORIUS, G. S., SMITH, M. W., THERA, M. A., WAMBEBE, C., WEBER, J. L. & WILLIAMS, S. M. 2009. The genetic structure and history of Africans and African Americans. *Science,* 324**,** 1035-44.

TISHKOFF, S. A., REED, F. A., RANCIARO, A., VOIGHT, B. F., BABBITT, C. C., SILVERMAN, J. S., POWELL, K., MORTENSEN, H. M., HIRBO, J. B., OSMAN, M., IBRAHIM, M., OMAR, S. A., LEMA, G., NYAMBO, T. B., GHORI, J., BUMPSTEAD, S., PRITCHARD, J. K., WRAY, G. A. & DELOUKAS, P. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet,* 39**,** 31-40.

TODD, J. A., BELL, J. I. & MCDEVITT, H. O. 1987. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature,* 329**,** 599-604.

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-TIRGOVISTE, C., SIMMONDS, M. J., HEWARD, J. M., GOUGH, S. C., DUNGER, D. B., WICKER, L. S. & CLAYTON, D. G. 2007. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet,* 39**,** 857-64.

TRONCHE, F., OPHERK, C., MORIGGL, R., KELLENDONK, C., REIMANN, A., SCHWAKE, L., REICHARDT, H. M., STANGL, K., GAU, D., HOEFLICH, A., BEUG, H., SCHMID, W. & SCHUTZ, G. 2004. Glucocorticoid receptor function in hepatocytes is essential to promote postnatal body growth. *Genes Dev,* 18**,** 492-7.

WALLACE, C., SMYTH, D. J., MAISURIA-ARMER, M., WALKER, N. M., TODD, J. A. & 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.

VAN AMELSFORT, J. M., VAN ROON, J. A., NOORDEGRAAF, M., JACOBS, K. M., BIJLSMA, J. W., LAFEBER, F. P. & TAAMS, L. S. 2007. Proinflammatory mediator-induced reversal of CD4+,CD25+ regulatory T cell-mediated suppression in rheumatoid arthritis. *Arthritis Rheum,* 56**,** 732-42.

VAN DER MEI, I. A., PONSONBY, A. L., BLIZZARD, L. & DWYER, T. 2001. Regional variation in multiple sclerosis prevalence in Australia and its association with ambient ultraviolet radiation. *Neuroepidemiology,* 20**,** 168-74.

VAN OOSTEN, B. W., BARKHOF, F., TRUYEN, L., BORINGA, J. B., BERTELSMANN, F. W., VON BLOMBERG, B. M., WOODY, J. N., HARTUNG, H. P. & POLMAN, C. H. 1996. Increased MRI activity and immune activation in two multiple sclerosis patients treated with the monoclonal anti-tumor necrosis factor antibody cA2. *Neurology,* 47**,** 1531-4.

VAN ZWAM, M., HUIZINGA, R., MELIEF, M. J., WIERENGA-WOLF, A. F., VAN MEURS, M., VOERMAN, J. S., BIBER, K. P., BODDEKE, H. W., HOPKEN, U. E., MEISEL, C., MEISEL, A., BECHMANN, I., HINTZEN, R. Q., T HART, B. A., AMOR, S., LAMAN, J. D. & BOVEN, L. A. 2009. Brain antigens in functionally distinct antigen-presenting cell populations in cervical lymph nodes in MS and EAE. *J Mol Med (Berl),* 87**,** 273-86.

WANG, D., STRAVOPODIS, D., TEGLUND, S., KITAZAWA, J. & IHLE, J. N. 1996. Naturally occurring dominant negative variants of Stat5. *Mol Cell Biol,* 16**,** 6141-8.

WANG, H., LEE, C. H., QI, C., TAILOR, P., FENG, J., ABBASI, S., ATSUMI, T. & MORSE, H. C., 3RD 2008. IRF8 regulates B-cell lineage specification, commitment, and differentiation. *Blood,* 112**,** 4028-38.

WANG, N., AKEY, J. M., ZHANG, K., CHAKRABORTY, R. & JIN, L. 2002. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am J Hum Genet,* 71**,** 1227-34.

VARILO, T., PAUNIO, T., PARKER, A., PEROLA, M., MEYER, J., TERWILLIGER, J. D. & PELTONEN, L. 2003. The interval of linkage disequilibrium (LD) detected with microsatellite and SNP markers in chromosomes of Finnish populations with different histories. *Hum Mol Genet,* 12**,** 51-9.

WATSON, J. D. & CRICK, F. H. 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature,* 171**,** 737-8.

VAZGIOURAKIS, V. M., ZERVOU, M. I., CHOULAKI, C., BERTSIAS, G., MELISSOURGAKI, M., YILMAZ, N., SIDIROPOULOS, P., PLANT, D., TROUW, L. A., TOES, R. E., KARDASSIS, D., YAVUZ, S., BOUMPAS, D. T. & GOULIELMOS, G. N. 2011. A common SNP in the CD40 region is associated with systemic lupus erythematosus and correlates with altered CD40 expression: implications for the pathogenesis. *Ann Rheum Dis,* 70**,** 2184-90.

WEINGER, J. G., BROSNAN, C. F., LOUDIG, O., GOLDBERG, M. F., MACIAN, F., ARNETT, H. A., PRIETO, A. L., TSIPERSON, V. & SHAFIT-ZAGARDO, B. 2011. Loss of the receptor tyrosine kinase Axl leads to enhanced inflammation in the CNS and delayed removal of myelin debris during experimental autoimmune encephalomyelitis. *J Neuroinflammation,* 8**,** 49.

WEISZ, A., MARX, P., SHARF, R., APPELLA, E., DRIGGERS, P. H., OZATO, K. & LEVI, B. Z. 1992. Human interferon consensus sequence binding protein is a negative regulator of enhancer elements common to interferon-inducible genes. *J Biol Chem,* 267**,** 25589-96.

VELLA, A., COOPER, J. D., LOWE, C. E., WALKER, N., NUTLAND, S., WIDMER, B., JONES, R., RING, S. M., MCARDLE, W., PEMBREY, M. E., STRACHAN, D. P., DUNGER, D. B., TWELLS, R. C., CLAYTON, D. G. & TODD, J. A. 2005. Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. *Am J Hum Genet,* 76**,** 773-9.

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.

WHITNEY, G. S., STARLING, G. C., BOWEN, M. A., MODRELL, B., SIADAK, A. W. & ARUFFO, A. 1995. The membrane-proximal scavenger receptor cysteine-rich domain of CD6 contains the activated leukocyte cell adhesion molecule binding site. *J Biol Chem,* 270**,** 18187-90.

WIKSTRÖM, J., KINNUNEN, E. & PORRAS, J. 1984. The Age-Specific Prevalence Ratio of Familial Multiple Sclerosis. *Neuroepidemiology,* 3**,** 74-81.

WILBE, M., JOKINEN, P., TRUVE, K., SEPPALA, E. H., KARLSSON, E. K., BIAGI, T., HUGHES, A., BANNASCH, D., ANDERSSON, G., HANSSON-HAMLIN, H., LOHI, H. & LINDBLAD-TOH, K. 2010. Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex. *Nat Genet,* 42**,** 250-4.

WILDE, D. B., MARRACK, P., KAPPLER, J., DIALYNAS, D. P. & FITCH, F. W. 1983. Evidence implicating L3T4 in class II MHC antigen reactivity; monoclonal antibody GK1.5 (anti-L3T4a) blocks class II MHC antigen-specific proliferation, release of lymphokines, and binding by cloned murine helper T lymphocyte lines. *J Immunol,* 131**,** 2178-83.

WILHELM, B. T., MARGUERAT, S., WATT, S., SCHUBERT, F., WOOD, V., GOODHEAD, I., PENKETT, C. J., ROGERS, J. & BÄHLER, J. 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature,* 453**,** 1239-43.

VISEL, A., BLOW, M. J., LI, Z., ZHANG, T., AKIYAMA, J. A., HOLT, A., PLAJZER-FRICK, I., SHOUKRY, M., WRIGHT, C., CHEN, F., AFZAL, V., REN, B., RUBIN, E. M. & PENNACCHIO, L. A. 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature,* 457**,** 854-8.

VISSCHER, P. M. 2008. Sizing up human height variation. *Nat Genet,* 40**,** 489-90.

VISSER, E. M., WILDE, K., WILSON, J. F., YONG, K. K. & COUNSELL, C. E. 2012. A new prevalence study of multiple sclerosis in Orkney, Shetland and Aberdeen city. *J Neurol Neurosurg Psychiatry,* 83**,** 719-24.

WORDSWORTH, B. P., LANCHBURY, J. S., SAKKAS, L. I., WELSH, K. I., PANAYI, G. S. & BELL, J. I. 1989. HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. *Proc Natl Acad Sci U S A,* 86**,** 10049-53.

WTCCC 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature,* 447**,** 661-78.

XING, J., WATKINS, W. S., WITHERSPOON, D. J., ZHANG, Y., GUTHERY, S. L., THARA, R., MOWRY, B. J., BULAYEVA, K., WEISS, R. B. & JORDE, L. B. 2009. Fine-scaled human genetic structure revealed by SNP microarrays. *Genome Res,* 19**,** 815-25.

XIONG, F., WU, C., CHANG, J., YU, D., XU, B., YUAN, P., ZHAI, K., XU, J., TAN, W. & LIN, D. 2011. Genetic variation in an miRNA-1827 binding site in MYCL1 alters susceptibility to small-cell lung cancer. *Cancer Res,* 71**,** 5175-81.

YAMAZAKI, T., YANG, X. O., CHUNG, Y., FUKUNAGA, A., NURIEVA, R., PAPPU, B., MARTIN-OROZCO, N., KANG, H. S., MA, L., PANOPOULOS, A. D., CRAIG, S., WATOWICH, S. S., JETTEN, A. M., TIAN, Q. & DONG, C. 2008. CCR6 regulates the migration of inflammatory and regulatory T cells. *J Immunol,* 181**,** 8391-401.

YANNELLI, J. R., SULLIVAN, J. A., MANDELL, G. L. & ENGELHARD, V. H. 1986. Reorientation and fusion of cytotoxic T lymphocyte granules after interaction with target cells as determined by high resolution cinemicrography. *J Immunol,* 136**,** 377-82.

YAO, Z., CUI, Y., WATFORD, W. T., BREAM, J. H., YAMAOKA, K., HISSONG, B. D., LI, D., DURUM, S. K., JIANG, Q., BHANDOOLA, A., HENNIGHAUSEN, L. & O'SHEA, J. J. 2006. Stat5a/b are essential for normal lymphoid development and differentiation. *Proc Natl Acad Sci U S A,* 103**,** 1000-5.

YU, N., JENSEN-SEAMAN, M. I., CHEMNICK, L., RYDER, O. & LI, W. H. 2004. Nucleotide diversity in gorillas. *Genetics,* 166**,** 1375-83.

YUAN, Z. L., GUAN, Y. J., CHATTERJEE, D. & CHIN, Y. E. 2005. Stat3 dimerization regulated by reversible acetylation of a single lysine residue. *Science,* 307**,** 269-73.

ZAMVIL, S., NELSON, P., TROTTER, J., MITCHELL, D., KNOBLER, R., FRITZ, R. & STEINMAN, L. 1985. T-cell clones specific for myelin basic protein induce chronic relapsing paralysis and demyelination. *Nature,* 317**,** 355-8.

ZEGGINI, E., SCOTT, L. J., SAXENA, R., VOIGHT, B. F., MARCHINI, J. L., HU, T., DE BAKKER, P. I., ABECASIS, G. R., ALMGREN, P., ANDERSEN, G., ARDLIE, K., BOSTROM, K. B., BERGMAN, R. N., BONNYCASTLE, L. L., BORCH-JOHNSEN, K., BURTT, N. P., CHEN, H., CHINES, P. S., DALY, M. J., DEODHAR, P., DING, C. J., DONEY, A. S., DUREN, W. L., ELLIOTT, K. S., ERDOS, M. R., FRAYLING, T. M., FREATHY, R. M., GIANNINY, L., GRALLERT, H., GRARUP, N., GROVES, C. J., GUIDUCCI, C., HANSEN, T., HERDER, C., HITMAN, G. A., HUGHES, T. E., ISOMAA, B., JACKSON, A. U., JORGENSEN, T., KONG, A., KUBALANZA, K., KURUVILLA, F. G., KUUSISTO, J., LANGENBERG, C., LANGO, H., LAURITZEN, T., LI, Y., LINDGREN, C. M., LYSSENKO, V., MARVELLE, A. F., MEISINGER, C., MIDTHJELL, K., MOHLKE, K. L., MORKEN, M. A., MORRIS, A. D., NARISU, N., NILSSON, P., OWEN, K. R., PALMER, C. N., PAYNE, F., PERRY, J. R., PETTERSEN, E., PLATOU, C., PROKOPENKO, I., QI, L., QIN, L., RAYNER, N. W., REES, M., ROIX, J. J., SANDBAEK, A., SHIELDS, B., SJOGREN, M., STEINTHORSDOTTIR, V., STRINGHAM, H. M., SWIFT, A. J., THORLEIFSSON, G., THORSTEINSDOTTIR, U., TIMPSON, N. J., TUOMI, T., TUOMILEHTO, J., WALKER, M., WATANABE, R. M., WEEDON, M. N., WILLER, C. J., ILLIG, T., HVEEM, K., HU, F. B., LAAKSO, M., STEFANSSON, K., PEDERSEN, O., WAREHAM, N. J., BARROSO, I., HATTERSLEY, A. T., COLLINS, F. S., GROOP, L., MCCARTHY, M. I., BOEHNKE, M. & ALTSHULER, D. 2008. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet,* 40**,** 638-45.

ZHANG, J., ZAHIR, N., JIANG, Q., MILIOTIS, H., HEYRAUD, S., MENG, X., DONG, B., XIE, G., QIU, F., HAO, Z., MCCULLOCH, C. A., KEYSTONE, E. C., PETERSON, A. C. & SIMINOVITCH, K. A. 2011. The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. *Nat Genet,* 43**,** 902-7.

ZHOU, D., SRIVASTAVA, R., NESSLER, S., GRUMMEL, V., SOMMER, N., BRUCK, W., HARTUNG, H. P., STADELMANN, C. & HEMMER, B. 2006. Identification of a pathogenic antibody response to native myelin oligodendrocyte glycoprotein in multiple sclerosis. *Proc Natl Acad Sci U S A,* 103**,** 19057-62.