From the Neuroimmunology Unit
Department of Clinical Neuroscience
Karolinska Institutet, Stockholm, Sweden

INHERITANCE OF AUTOIMMUNE NEUROINFLAMMATION

Pernilla Stridh

Karolinska Institutet

Stockholm 2010

Published by Karolinska Institutet. Printed by Larserics Digital Print AB, Sweden
Cover by Graham Bower

ISBN 978-91-7409-827-3

To my family

We used to think that if we knew one, we knew two, because one and one are two. We are finding that we must learn a great deal more about 'and'.

- Arthur Stanley Eddington

ABSTRACT

Multiple sclerosis (MS) is a chronic neuro-inflammatory disease with anticipated complex etiology. Susceptibility to MS is conferred by numerous genes, with very low odds ratios that explain minute fractions of disease. This indicates that unknown factors are responsible for the remaining genetic contribution, termed the 'missing heritability'.

Due to the similarities to MS pathogenesis, we studied myelin oligodendrocyte glycoprotein (MOG)-induced experimental autoimmune encephalomyelitis (EAE) in rats as a model for autoimmune neuroinflammation. Inbred rat strains show varying susceptibility to MOG-EAE, which can be explored in different experimental populations to identify influences on disease, including constituents of the 'missing heritability'. This thesis aims to identify components contributing to the heritability of autoimmune neuroinflammation.

We established congenic strains, a backcross (BC) and an advanced intercross line (AIL) to genetically map influences on EAE. The polygenic nature of neuroinflammation was demonstrated in these populations (Papers I, II and III). The BC identified 16 quantitative trait loci (QTL) that regulate EAE (Paper III), while one AIL region was resolved to four QTLs (Paper I: *Eae24-Eae27*) and another was resolved to two QTLs (Paper II: *Eae23a* and *Eae23b*). This enabled identification of a candidate gene for *Eae23b*, *ZEB1* (Paper II), which is involved in interleukin 2 (*IL2*) regulation. In Paper I, we demonstrated that epistatic interactions influence EAE, and that allele combinations are more important than individual QTL effects. Additionally, we identified parent-of-origin effects, a likely component of the 'missing heritability', to significantly contribute to the inheritance of EAE (Paper III).

These findings illustrate the genetic complexity involved in inheritance of autoimmune neuroinflammation, and prompted us to explore the use of a heterogeneous stock (HS) of rats to map EAE. In pilot studies (Paper IV), we determined that the HS can deliver high-resolution mapping, and influence from non-major histocompatibility complex (MHC) can be mapped, enabling the study of epistatic interactions involving the MHC.

A mixed genetic and epigenetic model of inheritance for autoimmune neuroinflammation is beginning to emerge. This indicates that genes, environment and their interactions, mediated by epigenetic mechanisms, contribute to neuroinflammation. Identifying constituents of this inheritance model will help us understand autoimmune neuroinflammation, and by extrapolation MS.

LIST OF PUBLICATIONS

I. Monica Marta*, Pernilla Stridh*, Kristina Becanovic, Alan Gillett, Johan Ockinger, Johnny C. Lorentzen, Maja Jagodic*, Tomas Olsson*. **Multiple loci comprising immune-related genes regulate experimental neuroinflammation.**
Genes & Immunity. 2010;11(1):21-36.

II. Pernilla M. Stridh, Melanie Thessen Hedreul, Amennai D. Beyeen, Milena Adzemovic, Alan Gillett, Johan Öckinger, Monica Marta, Hans Lassmann, Kristina Becanovic, Maja Jagodic, Tomas Olsson. **Fine-mapping resolves *Eae23* into two QTLs and implicates *ZEB1* as a candidate gene regulating experimental neuroinflammation.**
Manuscript.

III. Pernilla Stridh, Melanie Thessen Hedreul, Sevasti Flytzani, Petra Bergman, Amennai D. Beyeen, Alan Gillett, Johan Öckinger, Maja Jagodic. **Significant parent-of-origin effects implicate epigenetic regulation of experimental neuroinflammation.**
Manuscript.

IV. Martina Johannesson, Regina Lopez-Aumatell, Pernilla Stridh, Margarita Diez, Jonatan Tuncel, Gloria Blázquez, Esther Martinez-Membrives, Toni Cañete, Elia Vicens-Costa, Delyth Graham, Richard R. Copley, Polinka Hernandez-Pliego, Amennai D. Beyeen, Johan Öckinger, Cristina Fernández-Santamaría, Percio S. Gulko, Max Brenner, Adolf Tobeña, Marc Guitart-Masip, Lydia Giménez-Llort, Anna Dominiczak, Rikard Holmdahl, Dominique Gauguier, Tomas Olsson, Richard Mott, William Valdar, Eva E. Redei, Alberto Fernández-Teruel, Jonathan Flint. **A resource for the simultaneous high-resolution mapping of multiple quantitative trait loci in rats: the NIH heterogeneous stock.**
Genome Research. 2009;19(1):150-8.

*These authors contributed equally to the work

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

α-MOG	Anti-MOG
ACI	A X C 9935, Irish Curtiss and Dunning
AIL	Advanced Intercross Line
APC	Antigen-Presenting Cell
BBB	Blood-Brain Barrier
BC	Backcross
BN	Brown Norway
bp	Basepair
BUF	Buffalo
CFA	Complete Freund's Adjuvant
cM	Centimorgan
CNS	Central Nervous System
CNV	Copy Number Variant
DA	Dark Agouti
DNA	Deoxyribonucleic Acid
DZ	Dizygotic
EAE	Experimental Autoimmune Encephalomyelitis
EBV	Epstein-Barr Virus
ELISA	Enzyme-Linked Immunosorbent Assay
F2	Intercross Generation 2
F344	Fischer Curtiss and Dunning
G10	Generation 10
GWAS	Genome-Wide Association Study
HDAC	Histone Deacytelase
HLA	Human Leukocyte Antigen
HS	Heterogeneous Stock (of Rats)
IBD	Inflammatory Bowel Disease
ICR	Imprinting Control Regions
IFA	Incomplete Freund's Adjuvant
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
ISR	Interval-Specific Recombinant
kb	Kilobasepair
LD	Linkage Disequilibrium
LEW	Lewis
LOD	Logarithm of Odds
M520	Marshall 520
Mb	Megabasepair
MBP	Myelin Basic Protein

MHC	Major Histocompatibility Complex
mRNA	Messenger RNA
miRNA	Micro RNA
MOG	Myelin Oligodendrocyte Glycoprotein
MR	Maudsely Reactive
MS	Multiple Sclerosis
MZ	Monozygotic
ncRNA	Non-Coding RNA
PCR	Polymerase Chain Reaction
PLP	Proteolipid Protein
p.i.	Post-Immunization
PP	Primary Progressive
PVG	Piebald Virol Glaxo
QTL	Quantitative Trait Loci
RA	Rheumatoid Arthritis
rMOG	Recombinant MOG (aa 1-125)
RNA	Ribonucleic Acid
RR	Relapsing-Remitting
SCH	Spinal Cord Homogenate
SLE	Systemic Lupus Erythematosus
SNP	Single Nucleotide Polymorphism
SP	Secondary Progressive
Th	T helper
TNF	Tumor Necrosis Factor
Treg	Regulatory T-cell
WN	Inbred Wistar
WKY	Wistar-Kyoto
ZFN	Zinc-Finger Nucleases

1 GENETICS AND INHERITANCE

1.1 MENDELIAN INHERITANCE

Genetics is the science that studies variation of characteristics (traits) in organisms and the process of their inheritance from parents to offspring. Mendelian genetics, first published by Gregor Mendel in 1865-1866, presented a set of principle ideas that explain inheritance [1]. Although Mendel did not know the physical source of heredity, he proposed that traits are transmitted through discrete units, which are now called genes. Genes correspond to regions within the deoxyribonucleic acid (DNA) molecule that provide instructions for molecules such as proteins (their structure and function), while the gene products actually carry out the work in the cell or organism. A change in the genetic code (polymorphism) can directly impact how well a protein or other molecules function. An individual has two copies of each gene (alleles, one inherited from each parent). If the alleles are identical, the individual is homozygous and if the alleles are different, the individual is said to be heterozygous.

Mendel's ideas were summarized into the principle of segregation and the principle of independent assortment. The principle of segregation proposes that the two alleles of a gene separate from each other (segregate) so that each gamete receives a single copy. The physical proof of segregation was later found when the process of meiosis became known, a reductional division to produce the gametes for reproduction, and it involves the separation of homologous pairs of chromosomes [2] (Figure 1). This means that a combination of alleles is randomly inherited from each parent. The principle of independent assortment states that genes for different traits assort independently of one another (this actually deals with chromosome assortment rather than genes in the strict Mendelian sense). However, this turns out to be true also for genes on the same chromosome when they are not linked, which means that recombinations have occurred that separate the genes from each other (as opposed to linked genes that are inherited together). These ideas were later integrated with Thomas Hunt Morgan's' chromosome theory of inheritance (also known as the Boveri-Sutton chromosome theory [3]), which proposed that genes are carried on chromosomes that provide the mechanical basis of inheritance, to become the foundation of genetics. The same principles are used to discover genes that regulate a particular trait.

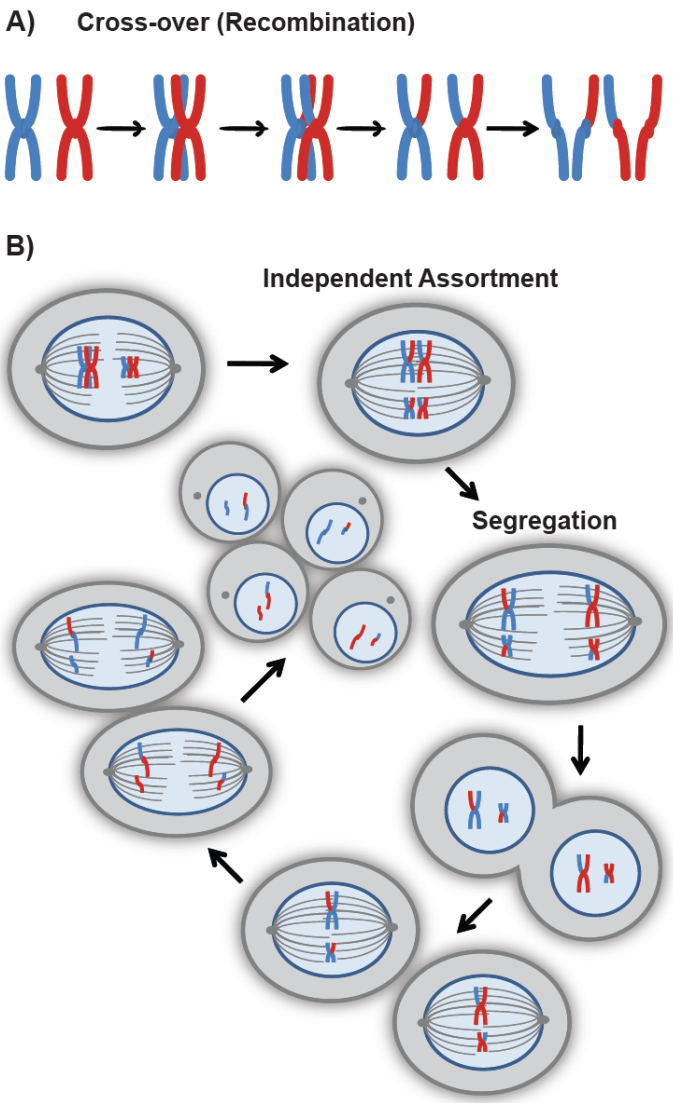

Figure 1. Transmission of genes (on chromosomes) from parent to offspring involves independent assortment and segregation. A) Genetic material is exchanged between homologous chromosomes by cross-overs: recombinations occur through introduction of breaks and exchange of DNA between chromatids at those breaks. Maternally (red, grand-maternal for the offspring) and paternally (blue, grand-paternal for the offspring) inherited chromosomes recombine to produce a chromosome with mixed parental DNA. B) The principle of independent assortment: homologous chromosomes (depicted by X and x) separate independently during meiosis (metaphase I). There is equal chance for the daughter cells to inherit either the mother's or father's homologue for each chromosome. The principle of segregation: each chromosome of a homologous pair (depicted by blue and red) is separated from the other during meiosis (anaphase I) so that each gamete receives a single copy.

William Bateson and Reginald Punnett's observation that some genes do not segregate independently during meiosis broke the principles of Mendelian inheritance and set forth the idea of genetic linkage. Genetic linkage describes the tendency of loci (genes or genetic regions) that are physically close on the same chromosome to be inherited together. Thomas Morgan observed that crossing-over events (recombinations) differed between linked genes, possibly reflecting the distance separating genes on the chromosome. His student, Alfred Sturtevant, proposed that the farther apart genes were the greater were the frequency of recombination between them. He developed genetic maps (linkage maps) with measures of distance, centimorgan (cM), determined by the number of recombinants between genetic markers. This provided a method to map traits to a location on the chromosome. Historically, phenotypic traits directly linked to genes (i.e. coat color) were used as markers, but today DNA sequences such as microsatellites and single nucleotide polymorphisms (SNP) are used instead.

1.2 COMPLEX TRAITS AND DISEASES

Mendelian genetics deals with monogenic traits and diseases, where a single gene is both necessary and sufficient to determine the trait. Single gene defects are responsible for many human disorders, however, the total frequency of susceptibility mutations for Mendelian diseases is usually below 1% [4]. Cystic fibrosis is the most common monogenic disorder and is caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene. Genetic approaches have been successfully used to identify genes responsible for a range of monogenic diseases [5]. However, most diseases depend on alleles at multiple genes that combine to confer susceptibility. Hallmark features of such complex diseases include etiological heterogeneity (where identical genes give different phenotypes), genetic heterogeneity (where different genes give the same phenotype), unknown mode of inheritance (dominance, incomplete dominance and codominance) and incomplete penetrance (where the risk allele has effect in only part of the individuals who carry it). Disease etiology is often largely unknown and many genes are involved, none of which is sufficient to cause disease on its own or that is essential for disease to develop [6].

Many polygenic diseases, such as Crohn's disease, type 1 diabetes rheumatoid arthritis (RA) and multiple sclerosis (MS), are complex in nature, which means that they arise from an interplay between genes, environment and unknown factors [4 7-9]. Family, adoption and twin studies concur in identifying a significant heritable component of inflammatory diseases [10-12]. Although genes have a critical role in this process, the combination with environmental and unknown factors determines the ultimate outcome.

Polymorphisms within risk genes are hypothesized to 'prime' an individual for disease and may cause the individual to pass a susceptibility threshold to develop disease. The threshold liability model [13] can be extended to illustrate this hypothesis (Figure 2). An assumption is that the genetic liability for a complex disease is normally distributed (Gaussian) among individuals within a population. Additionally, a threshold value must exists where individuals with genetic liabilities exceeding it converts from healthy to affected, equal to the population disease prevalence [14]. Accordingly, individuals who carry multiple susceptibility genes have a disease liability that is below, at or above threshold value. However, individuals may also carry disease resistance or modifier genes that alter the additive genetic liability [14]. Further, exposure to environmental and unknown factors could shift an individual's disease liability towards or away from the threshold. The advantage of this model is that it includes contributions of several factors to the onset of disease and can account for the genetic heterogeneity that exists within and between populations with complex diseases. Development of complex disorders can

thus be determined by genetic liability and environmental exposures together with yet unknown factors.

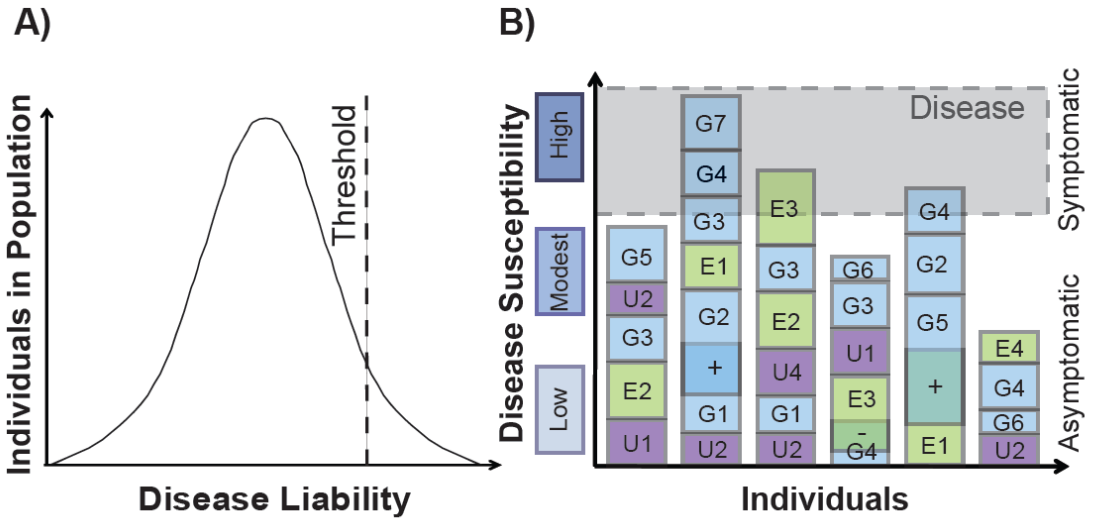

Figure 2. Threshold models for complex diseases. A) The threshold liability model shows the assumed Gaussian distribution of genetic liability for complex disease in the population. According to the model, individuals to the left of the liability threshold are healthy while individuals to the right have developed disease. The proportion of individuals that exceed the liability threshold is equal to the disease prevalence in the population. Adapted from Haegert 2004 [15]. B) An extended threshold model for susceptibility to complex disease. Susceptibility can be determined by an accumulation of genetic factors with weak to modest effects, environmental and unknown factors. Positive or negative interactions between the factors may alter individual effects. Factors: G = genetic, E = environmental and U = unknown.

1.2.1 Genetic Architecture and Inheritance

By 'genetic architecture', I refer to the number of genes that impact disease susceptibility and their spectrum of contribution (small to large effects), allelic effects at the given disease loci, penetrance of contributing genotype combinations and influence of gene-gene interactions. Hundreds of genetic variants have been identified by genome-wide association studies (GWAS) to be associated with complex diseases (http://www.genome.gov/26525384 [16]). These studies have identified risk loci in or near genes with no previous evidence for involvement in the etiology of the particular disease/trait, and have also identified loci that are shared among diseases without previously known common etiologic pathways. As expected, the identified genes demonstrated generally small effect sizes (odds ratios <1.5). Additionally, GWAS studies have demonstrated associations in many chromosomal regions that are currently annotated as gene poor [17] and the majority of associated variants fall outside of coding regions [16].

A popular model for the genetic architecture of complex disease is the Common Disease-Common Variant hypothesis, which proposes that the majority of genetic contribution to common diseases is attributable to disease loci with common variants (alleles present in more than 1-5% of the population)

[18 19]. However, this theory is criticized because most common variants confer relatively small increments of risk (1.1-1.5-fold) and explain only a small proportion of heritability. For example, 32 loci identified for Crohn's disease explain 20% of heritability [20] and six loci explain 15% of heritability in systemic lupus erythematosus (SLE) [21]. Critics of the Common Disease–Common Variant model have offered the alternative Common Disease-Rare Variant hypothesis, which approaches Mendelian heritability. This model favors the existence of multiple rare variants that contribute to disease with moderate to large effects [22]. Most of these variants will have low frequency (<5%) and low penetrance (<1.5-fold increased risk), although some may have moderate penetrance (1.5- to 5-fold risk) [23]. Proponents of this model generally concede that common variants may be the most important from the perspective of the population, but argue that rare variants are more important for individual risk.

There is evidence supporting both sides of the debate. Extreme examples of rare susceptibility variants with large effects manifest as rare Mendelian forms of complex disorders, such as maturity-onset diabetes in the young (MODY) [24]. More relevant may be the familial form of breast cancer, in which hundreds of rare disease-causing variations with large contributions to heritability have been identified in the BRCA1 and BRCA2 genes [25-27]. In inflammatory diseases, the common major risk factor is the human leukocyte antigen (HLA, MHC in rats), and rare variants have been indicated in inflammatory bowel disease (IBD) [28]. To date, hundreds of GWAS have been completed that have unequivocally identified common genetic variants associated with common diseases [17]. However, these may reflect rare variants with syntenic association to the identified SNP [29]. The Common Disease-Common or Rare Variant debate can be seen as a debate about the degree to which common and rare variations contribute to a particular disease phenotype [30]. Most complex diseases are likely to cover all parts of the spectrum between the extremes (Figure 3). Consistent with this, several genes that have been implicated in complex disease pathogenesis harbor multiple functionally significant rare and common variations [31]).

Ultimately, there is one concern at the heart of estimating heritability for complex disease. It may be the case that common diseases are not common at all. Because disease etiology is often largely not understood and biomarkers are often scarce, clinical symptoms are central to diagnosis. This means that a complex collection of symptoms can be grouped together to name a disease. However, if hundreds of genetic variants contribute to a single disease, and the genetic make-up that dictates susceptibility varies radically between individuals, the complex diseases we have today may actually be a collection of several different diseases. This could also explain the disparity in response to treatment among patients affected by complex diseases. The conundrum is

that without a more refined understanding of the genetic basis of disease and biomarkers thereof, more accurate categorization of the diseases is unrealistic.

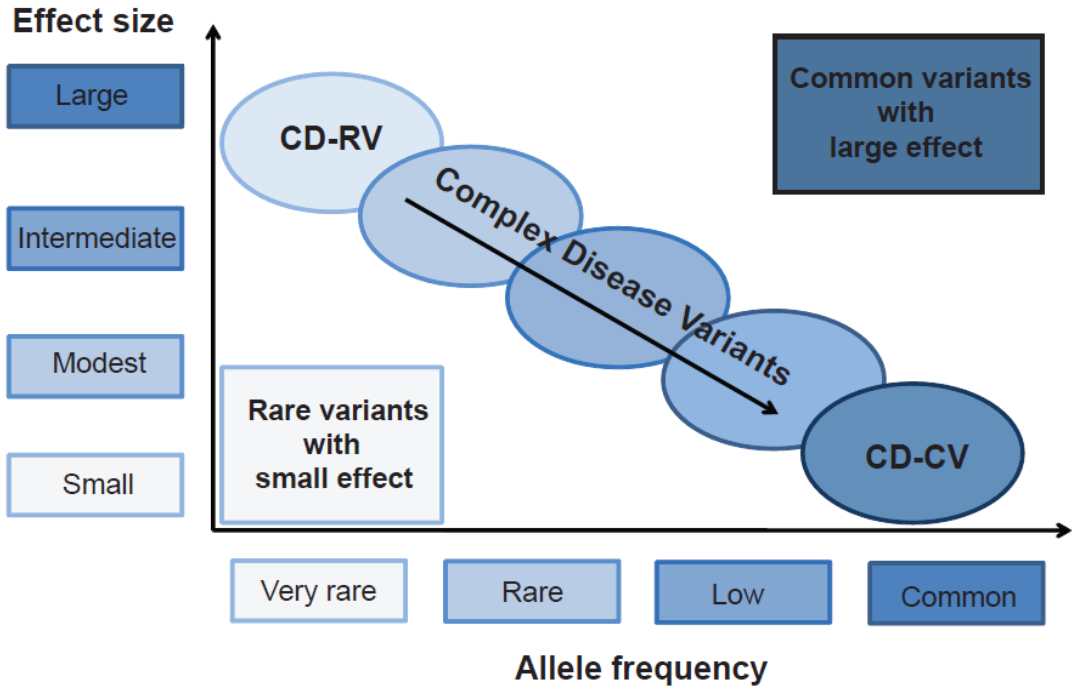

Figure 3. Complex Disease Variant Model. The Common Disease – Rare Variant (CD-RV) model fits the upper left side of the spectrum while the Common Disease – Common Variant hypothesis model fits the lower right. In contrast, complex diseases genes are hypothesized to span the spectrum between these models. Adapted from Manolino 2009 [32].

1.3 THE MISSING HERITABILITY

Despite hundreds of identified variants for several dozen traits, they fail to explain most of the heritability involved. Identification of additional risk variants will explain some of the deficit, but is not likely to explain the majority of complex disease etiology. This ultimately begs the question regarding the remaining genetic component that contribute to disease, termed the 'missing heritability' [33]. Many explanations for this missing heritability have been suggested, including larger numbers of variants than anticipated with smaller effects (yet to be found), rarer variants (possibly with larger effects) and structural variants, complicated gene-gene interactions and inadequately accounting for environment that is shared among relatives. Additionally, complex disease heritability that is not accounted for by genetic variants could be accounted for by subtle gene-environment interactions and epigenetic phenomena.

1.3.1 Numbers of Variants

One emerging hypothesis is that a significant proportion of this 'missing heritability' will be explained by low-frequency variants with intermediate penetrance effects [34]. There are relatively few examples of such variants

contributing to complex traits, possibly because they have largely escaped conventional gene-discovery approaches [32]. Hypothetically, a risk variant with a minor allele frequency of 1%, an odds ratio of 3 and ~45% penetrance would have a stronger effect on familial risk than most known common susceptibility variants. Yet, given a disease prevalence of 5% in the population, this variant would have too low penetrance to be detected by traditional linkage approaches and would have too low risk-allele frequency to be reliably detected by GWAS [8 35 36]. Familial risk is studied through linkage, which identifies genes inherited together with disease, while association studies compare allele distributions between affected and unaffected people. Possible contributors from these low frequency variants are so far included in the missing heritability component.

1.3.2 Structural Variants

Structural variations may also contribute to the 'missing heritability'. They can contribute large to modest effects on disease, but can remain undetected by standard arrays. Copy number variants (CNVs), which are ~1 kb fragments present in variable numbers across individuals (insertions and deletions) and copy neutral variation (CNPs, inversions and translocations), can also arise *de novo* in an individual without any family history of the disease. This would not reflect an inherited variant *per se* and would not contribute to heritability, but may account for some of the genetic variability between individuals and could explain some of the variation that is presently attributed to environment. CNV analysis may also aid the identification of additional and more prevalent risk variants in genes and pathways involved in disease. This rapidly expanding area of research holds a promise to further elucidate the missing component of heritability. The contributions of structural variants will not be explored in this thesis.

1.3.3 Technical Limitations

Some shortcomings in experimental design and statistical analyses of genetic studies may also contribute to the paucity of explained variance. For example, genetic markers used in studies are selected to represent a portion of the genome. However, the selected marker may not accurately represent the actual risk variant. This may also be the case with tag-SNPs, which are selected to represent a block of several variants. Additionally, many arrays have to date been focused on genes, and have therefore excluded sequence coverage of regulatory regions. Structural variants may also been missed by available genotyping arrays. To understand all of the 'missing heritability', better and more varied models of the entire network of genes and regulatory sequences are needed.

1.3.4 Gene-Environment Interactions

Complex diseases are influenced by environmental factors, which can either have a direct influence in the phenotype or act via gene-environment interactions. A role of environmental contribution is supported by geographic differences in disease incidence, variations in disease patterns in immigrant populations and variations in trends over time. Variable degrees of morbidity can be attributed to environmental exposures in some complex diseases [37-39]. The spectrum of involved environmental exposures includes, but is not limited to, risk behavior and/or lifestyle, community factors (the physical environment) and social factors (the psychological environment). An attempt to explain these variations resulted in the 'thrifty genotype hypothesis', which proposes that genes that evolved under particular selective pressures, which is a slow process, have instead become risk factors in a rapidly changing environment [40]. In other words, a relatively constant genetic architecture can respond to a changing environment to give an unfavorable genetic variant - environmental exposure combination.

The 'hygiene hypothesis' offers an alternative explanation to the development of 'new' diseases, as it proposes that the increased prevalence of inflammatory disease in the Western world is due to a decrease in exposure to microorganisms resulting from an increased emphasis on hygiene and the widespread use of antibiotics [41]. This theory is supported by a decreased occurrence of inflammatory disease with increased exposure of infants and young children to microorganisms [42], daycare and living with older siblings [43]. Whatever the mechanism, true gene-environment interactions involve a difference in the direction and magnitude of a gene variant's effect on disease depending on the environment.

Equally important to identifying environmental exposures that contribute to the etiology of complex disease may be the critical timing of such exposures [7]. For example, adverse intrauterine experiences have been shown to influence the occurrence of complex disease later in life [44]. Factors that contribute to the intrauterine effect have been postulated to include exposure to high amounts of endogenous estrogen [45] and growth factors [46], birth order and parental age [47]. Furthermore, cigarette smoking during pregnancy may exert its effect on future inflammatory disease through altered cytokine levels in offspring cord blood [48]. Collection of reliable data for the perinatal period is one of the major challenges in assessing the effect of intrauterine factors on complex disease risk later in life.

Measuring an individual's exposure to environmental factors is often complex, imprecise and subject to bias, even when assessment tools exist [49]. Further, teasing apart the contributions to heritability of environmental factors that are

shared among relatives is difficult. A positive family history captures a combination of genetic and environmental risk factors, both measured and unmeasured, that lead to disease expression. These influences may be attributed to inheritance, when they in fact originate in the environment. However, environmental factors that interact with genes to influence disease, and specifically with certain variants of genes, should accurately be included in heritability. This distinction will be difficult to tease apart. The availability of genome-wide markers can provide empirical estimates of identity-by-descent (IBD) allele sharing between pairs of relatives, and this may help to separate the components of heritability within families [50 51]. Identifying and characterizing environmental influences on complex diseases will probably elucidate part of the missing heritability that is due to environment.

1.3.5 Epistatic Interactions

Understanding of the genetic basis underlying complex diseases is further complicated by the presence of gene-gene interactions. These interactions may be either additive or epistatic. Additive interactions depend on the sum of gene effects that drive a phenotype in the same direction, while epistatic interactions give a phenotype that is not predicted by the sum of its known single-locus effects. This means that the outcome of one locus is masked or altered by the presence of another locus. As illustrated in Figure 3, when genes work in concert, they can amplify (+, synergistic) or diminish (-, antagonistic) their individual contributions. Epistasis can occur at the genomic level, where one gene could code for a protein that either promotes transcription of the other gene (synergistic) or prevents it (antagonistic). It may also involve other protein complexes and receptor-ligand interactions. If more than two loci are involved, the complex multi-way interactions between several loci make it even more difficult to distinguish the effect of each [52].

Although these effects have been considered to be of high importance in complex diseases, they have been largely unexplored until recently due to methodological and statistical limitations. One of the major challenges has been statistical power because the study of gene-gene interactions requires exponentially larger cohorts, something that is cost-prohibitive. Identifying effects originating from epistatic interactions has the potential to explain part of the missing heritability. Epistatic interactions may reduce the power to detect specific loci, if the effects of one variant cannot be found without the others. This would contribute to genuine disease loci remaining undetected; hence a proportion of heritability remains missing. Therefore, it is important to study epistatic interaction in order to accurately understand and recognize gene effects, so they can be properly accounted for. Additionally, these interactions may also explain some cases of incomplete penetrance. Variable penetrance of a gene can be caused by interactions with a modifier gene that masks or

diminish its effect, so that the gene's phenotype expression is not present in all individuals who carry the variant.

1.3.6 Parent-of-Origin Effects

Another phenomenon that may contribute to unidentified genetic variation underlying complex diseases is parent-of-origin effect, whereby the phenotype expression depends on whether the transmission originated from mother or father. Parent-of-origin effects can be caused by genetic imprinting, intrauterine effects, maternally inherited mitochondrial genes, or paternally inherited Y chromosome. These effects can mask genetic variation underlying diseases by reducing power for detection, since only the proportion of the population who has inherited the risk variant from the 'right' parent will be informative. This may explain the incomplete penetrance and variable expressivity of complex diseases.

Unequal transcription of parental alleles that results from epigenetic modification of the genome is called genomic imprinting. The resulting gene expression is dependent upon the sex of the parent from which the allele was inherited [53]. A preference for mutations in the germ line of either the father or mother can also contribute. Base substitutions tend to have paternal origin, due to the greater number of cell divisions in spermatogenesis contributing to errors during replication [54]. On the other hand, chromosomal abnormalities tend to be maternal in origin, likely due to the longer arrest in meiosis of oocytes contributing to nondisjunction events [55].

1.3.6.1 Genomic Imprinting

A broad definition for genomic imprinting is an epigenetic phenomenon that results in unequal expression of the maternally and paternally derived copies of a gene. This depends on epigenetic instructions (imprints) that mark the genes differently in egg and sperm, and inheritance of these epigenetic marks leads to parent-specific gene expression [56] [57]. Genomic imprinting is inherited via sequence elements called imprinting control regions (ICRs). ICRs generate differentially methylated regions that permit or silence expression of the imprinted genes [58]. Several molecular mechanisms can generate similar patterns of uniparental gene expression, including methylation of DNA, acetylation and methylation of histones, modifications of local chromosomal structure and noncoding RNA [59]. Aberrant imprinting disturbs development and is the cause of various disease syndromes.

Beckwith-Wiedemann syndrome (BWS) [60] is an overgrowth syndrome involving locus 11p15.5 , which encompasses one of the most well-known imprinted regions, the *H19- IGF2* (insulin-like growth factor 2) locus [61]. *H19* is maternally expressed (paternally imprinted). Further upstream is the *IGF2* gene, which in

contrast to *H19*, is expressed from the paternal allele [61]. The non-methylated maternal allele alters the local chromosomal structure to insulate *IGF2* from local enhancers that are required for expression [62 63]. Imprinting failure leads to bi-allelic expression of *IGF2*, which is one cause of BWS [64].

1.3.6.2 Intrauterine Effects

Epigenetic regulation of gene promoters that is responsible for transcriptional expression and silencing in the adult is established during development. Experimental studies have shown that disturbances to this process during the intra-uterine period can change gene expression patterns later in life. Systematic effects from an exposure during this period are called intrauterine effect. This represents a molecular mechanism for inducing enduring modification in phenotype by the environment. For example, feeding dietary methyl donor supplementation to yellow agouti mice that carry non-agouti alleles (a/a) changes the methylation pattern of the *Avy* locus in offspring, which changes the coat color from yellow to brown (agouti) [65-67]. This change is then inherited by the second generation offspring. Maternal behaviour after birth (during suckling) can also influence the epigenome. Poor nurturing by mothers induced hypermethylation of the GR gene, which lead to an increased stress response [68].

1.3.6.3 Mitochondrial Genome

Mitochondrial DNA (mtDNA) is a circular double-stranded molecule (16,569 bp long in humans), of which hundreds to thousands of copies are present in each cell. It codes for the oxidative phosphorylation system, rRNAs and tRNAs and is also involved in regulation of apoptosis [69]. The mitochondrial genome is different from the nuclear genome in that it has uniparental inheritance (maternal), a high copy number, a lack of recombination, and a generally higher mutation rate. The multiple copies of mtDNA are not necessarily identical within an individual (heteroplasmy).

The fact that mitochondrial genome is maternally inherited means that the female gamete contributes the functional mitochondrial genomes to the embryo. Mitochondrial genes that contribute to disease will therefore be detected as maternal parent-of-origin effect. Furthermore, risk variants in the genome that interacts with mitochondria to achieve its full effect will also be detected as dependant on maternally transmitted effect.

1.4 EPIGENETIC INFLUENCES

The incomplete penetrance and variable expressivity of complex disease may in part be explained by epigenetic differences [70]. Epigenetic marks comprise stable changes in gene expression that do not require changes in the DNA

sequence of the gene. Such states are mediated by DNA methylation, histone modifications, noncoding RNAs and large nuclear-transcription factor complexes [71]. Additionally, the molecular link between genes and environmental factors can be provided by the change in epigenetic marks induced by environmental exposures.

DNA methylation involves the covalent addition of a methyl group to cytosine bases in CpG dinucleotides. This is an essential mechanism for normal regulation of gene expression and disruptions are implicated in many cancers [72] [73]. Methylation of cytosines are, among other things, involved in the transcriptional silencing of transposons, imprinted genes, and genes on the inactive X chromosome [74]. Transcriptional inhibition may occur by methylated CpG sites blocking the binding of transcription factors to promoters or indirectly by methyl-CpG-binding-domain-containing proteins that recruit histone deacytelase (HDAC) activity to methylated DNA, resulting in a deacetylated repressive chromatin structure [75] [76].

The DNA is packaged around histones into highly compacted chromatin, which constrains the availability of genes for transcription and other chromatin-based processes. The covalent modifications of histone tails have an important role in defining the chromatin structure. Different histone residues can be modified by acetilation, methylation, phosphorylation, sumoylation, citrulination etc. The effects of modifications depend on the specific change and the specific residues that are modified, resulting in either gene activation or repression. In general, histone acetylation opens the chromatin structure to make the DNA accessible to the transcriptional machinery. Methylation of lysines 4 and and 36 in histone H3 correlates with transcriptional activation. Conversly, demethylation of lysine 4 and methylation of lysines 9 and 27 correlates with transcriptional repression [77]. There is cross-talk between DNA methylation and histone modifications to shape the epigenome [78].

Non-coding (nc)-RNAs can direct epigenetic change to loci that contain homology to their RNA sequence [79]. Long non-coding RNAs (ncRNA) epigenetically regulate gene expression by modifying histone structure to alter RNA polymerase accessibility, by means of defining domains of differential histone methylation along the target genes [80]. Long nc-RNA can also induce changes by associating with chromatin-modifying complexes, and are involved in genomic imprinting. Conversely, microRNAs (miRNAs) regulate gene expression post-transcriptionally [81]. Sequence-specific recruitment of RNA-induced silencing complexes (RISCs) to messages with complementary sequences leads to translational inhibition, accelerated exonucleolytic mRNA decay, or endonucleolytic cleavage of miRNA–mRNA pairs (slicing).

Alterations of epigenetic states result in changes of gene expression and are instrumental for regulating phenotypes, in health and disease. Thus, while genes confer the primary information for gene expression, epigenetic mechanisms can decide when and where the genetic information is going to be expressed. For example, epigenetic differences that accumulate during the life time of monozygotic (MZ) twins lead to differences in expression of their genes [82].

1.5 THE CHALLENGE OF MAPPING COMPLEX TRAITS

Genetic studies have been plagued by difficulties in replicating results. This is likely due to the previously described features of complex diseases in conjunction with the use of cohorts of insufficient size. The size was underastimated largely due to a historic over-estimation of effects size of contributing loci. Studies in several conditions have clearly demonstrated that the number of detected variants increases with increasing sample size [83]. Sample sizes used for the initial identification of sequence variants have generally been modest.

So, why do we study inheritance if it is so complicated?

The goal in genetic research of complex diseases is to better understand disease etiology, in order to achieve more effective means of diagnosis, treatment and prevention. We expect to accomplish this either through elucidating functional properties of known causative variants or identifying new variants in which true functionality lays. Identification of variants that contribute to disease risk, prognosis, or response to treatment should lead to the development of safe and effective diagnostic and interventional strategies. The ultimate objective, although very distant, is a complete inventory of the susceptibility architecture of disease and understanding of related functions and mechanisms with translation to the clinic.

2 PHENOTYPE IN FOCUS

2.1 MULTIPLE SCLEROSIS

MS, also known as encephalomyelitis disseminata, is a chronic neuroinflammatory disease that involves damage to myelin sheaths and nerve fibers [84]. There are large individual differences in clinical symptoms, severity and disease course, and the symptoms that a person experiences depend on the anatomical location of the inflammatory lesions in the central nervous system (CNS). Common symptoms include cognitive dysfunction, fatigue, speech problems, emotional liability, sensory loss, optic neuritis, difficulties with coordination/balance (ataxia), and muscle weakness and spasms [84]. Typically, disease onset occurs around 20-40 years of age and MS is more prevalent in women than men, with an approximate 1:2 ratio [85]. The prevalence of MS in Sweden is 125-140 affected individuals per 100 000 individuals [86 87] making MS the most common non-traumatic cause of neurological disability in young adults, often leading to a marked reduction in quality of life for those affected [88 89]. The etiology of MS remains elusive, and there is no single test or biomarker that is sufficient for diagnosis.

2.1.1 Clinical Features

MS manifests with a range of symptoms that can be grouped into three clinical disease courses [90]. Relapsing-remitting MS (RR-MS) involves bouts of inflammation interspersed with periods of recovery, and approximately 85% of cases have this diagnosis. Clinical exacerbation involves profound inflammation and blood-brain-barrier (BBB) breakdown, while remissions occur when inflammation is resolved [91]. Conversely, progressive MS is characterized by a steady worsening of symptoms, be it primary (PP-MS) or secondary (SP-MS) to the relapsing-remitting form. The absence of recovery in this phase may be due to repeated attacks of inflammatory demyelination of the same nerve, failed remyelination, and irreversible axonal damage [92 93]. Approximately 15% of cases have PP-MS, while half of RR-MS patients have entered into a progressive phase after 10 years [94].

2.1.2 Pathology and Immunology

The CNS in individuals with MS is characterized by inflammatory lesions, demyelination, glial scarring, axonal degeneration and varying degrees of remyelination, with substantial heterogeneity in pathology [95]. The inflammatory process involved in MS is hypothesized to be initiated by peripheral dysregulation of the immune response and activation of autoreactive myelin-specific T-cells. Activation then leads to an upregulation of adhesion molecules that facilitate migration of T-cells across the BBB into the CNS [96]. Once the autoreactive T-cells are on location, they recognize their antigen and become

reactivated through interaction with resident microglia [97 98]. Reactivation leads to secretion of cytokines and chemokines with subsequent recruitment of more inflammatory cells. Macrophages and microglia release inflammatory mediators to elicit profound inflammation, demyelination and axonal damage.

MS is sometimes referred to as autoimmune because of the immune cells and immune mediators present at the site of injury (described above). Autoimmune disorders develop when the physiological tolerance to ‘‘self’’ antigens is lost. Several features of MS are consistent with an autoimmune etiology. 1) The genetic association with HLA complex (described below) is coherent with other autoimmune diseases [99]. 2) The presence of antibodies directed against myelin proteins in the CNS, and the presence of myelin-reactive T-cell and B-cells in the serum and cerebrospinal fluid (CSF) of individuals with MS indicates an attack of “self” antigens [100 101]. 3) The infiltration of T-cells, B-cells, and macrophages into the CNS with associated CNS myelin destruction indicates a break in tolerance [102]. 4) Similarities in histopathology and clinical disease with autoimmune experimental models. The significance of antibodies detected in MS is still debated, but histopathological evidence of a antibody-mediated demyelination is evident in >50% of MS patients and is consistently associated with active demyelination [103]. Autoreactive T cells occur in the blood of both patients with MS and healthy individuals. This must imply that regulatory mechanisms exist to prevent autoreactive T cells from causing immune disorders, otherwise MS would be more widespread than it is. Active suppression by regulatory T (Treg) cells is one such mechanism for control of self-antigen-reactive T cells, and while the frequency of Tregs in MS patients is normal compared to controls, their suppressive function is diminished [104].

2.1.3 Genetic Contributions

MS is a genetically complex disease that depends on interactions between genetic and environmental factors. Observations in twin cohorts and familial aggregation studies have demonstrated the genetic component of MS etiology [10-12]. MZ twins, who theoretically share nearly 100% identical genomes, show a 25-30% concordance rate for MS compared to 2-5% concordance in dizygotic (DZ) twins, who share approximately 50% of the genome [105 106]. Conversely, adoptees or non-biological family members do not have increased risk for MS if other members are affected [10].

Many population based linkage and association studies have been conducted to identify MS risk genes. Until recently, the human leukocyte antigen (HLA) was the only region that unambiguously showed linkage and association to MS [107 108]. The HLA region contains over 200 genes, many of which are involved in immune development and function, and alleles at different loci are often inherited together in established haplotypes because of linkage disequilibrium

(LD) extending over large distances. The HLA class II haplotype DR15-DQB1*0602 is the strongest genetic risk factor for MS, conferring a 3-fold increase in risk [99]. A major advance in the search for MS genes came with the identification of the interleukin-7 receptor alpha (*IL7R*) and interleukin-2 receptor alpha (*IL2RA*) genes, the first non-HLA genes to unambiguously be associated with MS [109-112]. Since then, several other genes have been identified and confirmed, including *CD58, RPL5, CLEC16A, KIF21B and TREM39A* [112-116], with other candidates still to be confirmed (*TNFRSF1A, IRF8, CD6, TYK2, CD226, CYB27B1, PRKCA, KIF1B* [117-125]). Efforts to elucidate the potential role of these genes in MS susceptibility and to identify additional risk genes are ongoing.

Most of the genes implicated in MS code for proteins involved in adaptive immune functions, supporting a role for inflammatory pathogenesis. *IL2RA* is highly expressed on activated T helper (Th)1 cells and Treg. Expression of *IL2RA* is crucial for the delivery of IL-2 signals to Treg, which regulates the adaptive immune system and influence T-cell homeostasis [126]. *IL7R* is involved in T-cell survival and proliferation, and may influence MS by differential splice-variant expression of its membrane-bound and soluble forms [109 127 128].

2.1.4 Environmental Influences

Various environmental exposures may influence MS etiology, with an apparent geographic heterogeneity [129]. Individuals migrating from a low- to high-prevalence area before adolescence acquire the higher risk, but not if they move later in life [130-132]. There is also evidence for a seasonal effect, with most people affected by MS born in May (9.1% increase) and the fewest born in November (8.5% decrease) [133]. The annual variation in sunlight exposure has been proposed as a key environmental factor for MS, particularly because of its role in generating active vitamin D [134]. Vitamin D may influence HLA gene expression [135 136], and a lack of vitamin D in early life has been hypothesized to affect central deletion of self-reactive T-cells [137]. This is an example of how a gene-environment interaction can influence disease. Other environmental risk factors include smoking, which increases risk and worsen prognosis of MS [138 139], possibly due to chemical components other than nicotine [140 141] contributing to chronic cyanide intoxication, smoking-mediated infections and DNA methylation changes [141 142]. Common childhood infections, such as morbilli, rubella and Epstein-Barr virus (EBV), have also been associated with increased MS risk [143]. Although individuals who develop MS have an altered immune response to EBV antigen [144], extensive investigations have failed to confirm direct involvement of specific viral infections [145].

Despite sharing 50% of the genome, DZ co-twins have a higher concordance for MS (5.4%) compared to their non-twin siblings (2.9%) [146]. Thus, environmental factors seem to act at a young age [133] [147], either directly interacting with genes, as in the case of vitamin D and HLA, or indirectly by changing epigenetic marks and future gene expression.

2.1.5 Epigenetics

Recently, a number of studies have demonstrated that the inheritance of MS is much more complicated than initially anticipated. Epigenetic influences are consistent with the characteristics of MS and may provide a component in the missing heritability of chronic inflammatory disease. The maternal route is favored in disease transmission, with maternal half-siblings of MS-affected persons having a significantly higher risk for developing MS than paternal half-siblings [148] [149]. The *HLA* haplotype itself show similar parent-of-origin effects [150] [151] with higher conferred risk in families with affected second-degree relatives, implicating gene–environment interactions [151]. Sex-dependent multifactorial inheritance has been suggested in MS [152], while it is argued to reflect parental transmission by others [149] [153]. The parent-of-origin effects demonstrated in MS are likely to include inherited epigenetic changes induced by environmental triggers. The loss of genomic imprinting has influence in several disease phenotypes [154], but has not been directly studied in MS. The importance of epigenetic modifications in MS pathogenesis is also indicated by differential histone modifications and methylation status in the affected CNS [155]. Thus, these previous findings indicate that multifactorial genetic, environmental and epigenetic mechanisms are involved in the inheritance of MS.

2.1.6 Epistasis

MS susceptibility is influenced by epistatic interactions. For example, complex interactions between HLA haplotypes alter an individual's susceptibility to MS [156]. This means that the HLA association with MS is not straightforward, and understanding the epistatic interactions will be necessary to understand pathogenesis.

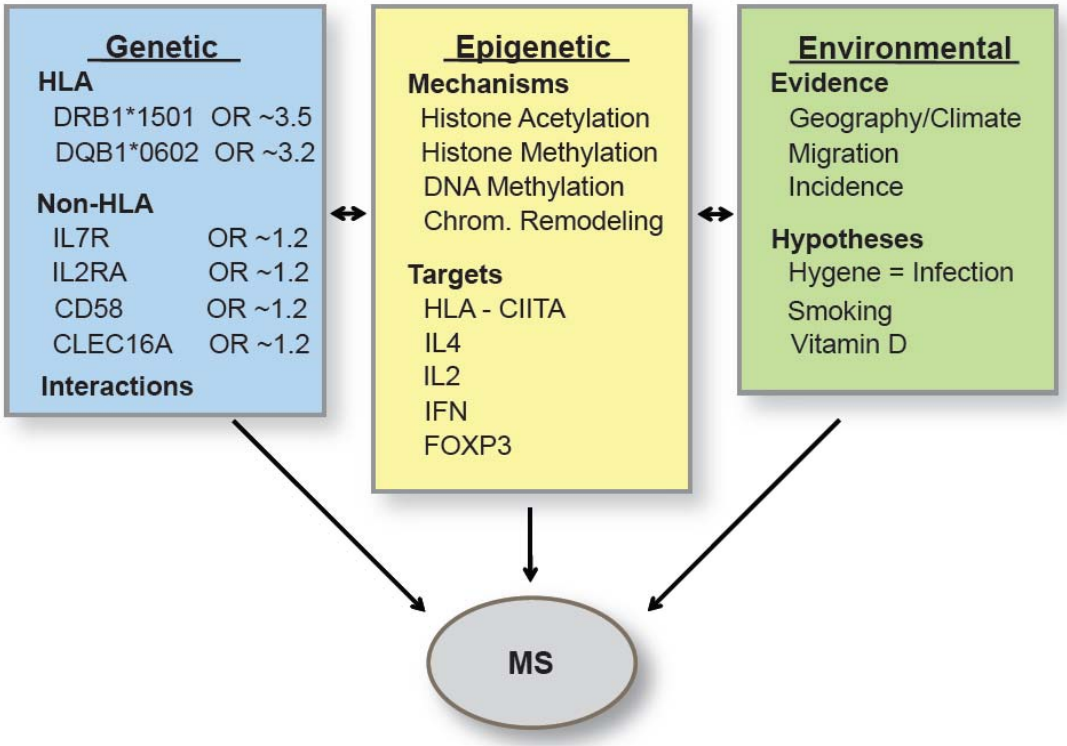

Figure 4. Risk factors for MS include genes (HLA [157] and non-HLA [114]), epigenetic regulation, environmental factors, and gene-gene and gene-environment intractions.

2.2 ANIMAL MODELS

Understanding the inheritance of MS and identifying the factors contributing to transmission of the disease remains a challenge. An approach to circumvent some of the obstacles of gene identification in humans is to use a simplified system. Animal models of MS represent such a system, where heterogeneity can be minimized and environmental conditions controlled. The primary advantage is that the genetic architecture of inbred animals is vastly simpler than the genetic architecture in human populations. Other advantages of this approach include theoretically unlimited sample sizes, control over disease kinetic and access to tissues and cells. This becomes particularly important when modeling a disease that affects an organ than can rarely be sampled in the human population, such as the CNS. The early events of MS cannot be studied in humans, since disease mechanisms have been operating for a while before diagnosis. Thus, animal models offer a unique opportunity to study the early events of disease, at least the artificially induced version of it.

The disease is genetically dissected in the model system, on the assumption that there are conserved mechanisms among species that lead to neuroinflammation. Several of the quantitative trait loci (QTL, regions that regulate quantitatively varying phenotypes) that have been identified in animal models are syntenic to human regions that have shown linkage or association in MS [158 159]. We see this as evidence that the QTL may constitute a gene or

pathway of relevance for both the animal model and MS. This further motivates the employment of animal models, because even if not identical an identified disease-regulating gene could resolve pathogenic mechanisms and pathways of importance in MS.

2.3 EXPERIMENTAL AUTOIMMUNE ENCEPHALOMYELITIS (EAE)

Experimental autoimmune encephalomyelitis (EAE) is an autoimmune neuroinflammatory disease with clinical and pathological similarities to MS [160]. EAE can be obtained in several species and strains including mice, rats, and guinea pigs, although there is no single experimental model that mimics all aspects of MS. The mode of induction, the genetic constitution, and the myelin autoantigen used if actively immunized, shape the clinical course and the histopathological and immunological features of MS captured. Other factors such as age, weight, and gender also influence the outcome. There is a great need for more efficient and safe MS treatments, which requires a better understanding of disease mechanisms. Therefore, disease-appropriate animal models are indispensable for further progress. Indeed, the most recent treatments approved for MS have been developed in EAE, demonstrating its predictive value when appropriately applied [161].

2.3.1 EAE Models and Symptoms

Several models can be used to induce EAE. EAE actively induced with CNS autoantigen and adjuvant (mineral oil, *Mycobacterium tuberculosis*, and sometimes *Bordetella pertussis*) recruits an autoimmune T- and B-cell response causing either acute monophasic or chronic relapsing course [160]. Active induction is usually achieved by subcutaneous immunization of the immunogen emulsified in an adjuvant. Commonly used immunogens include whole spinal cord homogenate (SCH), myelin basic protein (MBP), proteolipid protein (PLP) and myelin oligodendrocyte glycoprotein (MOG). *M. tuberculosis* can be added as a superantigen to boost the immune trigger and *pertussis* can be used to open the BBB, with the aim of achieving a more severe EAE induction. Rodents can display a monophasic bout of EAE, a relapsing-remitting form (mimicking RR-MS), or chronic EAE (mimicking PP-MS or SP-MS) (Figure 5). For example, a MS-like relapsing–remitting disease with plaques of demyelination can be induced in DA rats by MOG, while MBP in LEW rats gives an acute monophasic inflammatory disease with little demyelination [160].

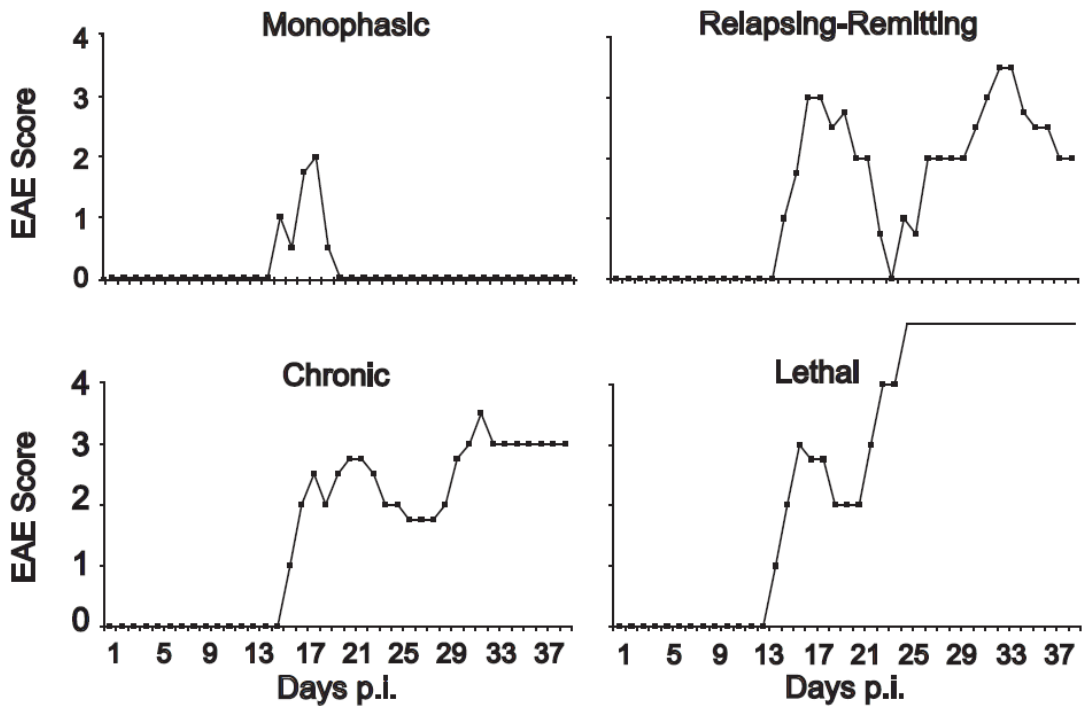

Figure 5. Clinical disease courses in EAE and corresponding MS disease course.

Adoptive transfer of myelin antigen reactive T-cells cause, in most cases, widespread CNS inflammation with little demyelination giving acute monophasic EAE [160]. This methodology is important for understanding central features of T-cell biology involved in neuroinflammation, but not for questions requiring a disease course and pathology closely mimicking MS. Cotransfer EAE, in which autoimmune T-cells are transferred together with myelin-specific antibodies, does not induce clinical disease, but can demonstrate the role played by autoantibodies [160]. Transgenic induction of EAE can be achieved in mice constructed to have high proportions of T-cells expressing human T-cell receptors for myelin autoantigens (or other manipulations), which lead to 'spontaneous' EAE [162]. Although they are somewhat artificial, these mice models are used for mechanistic studies [163-166].

The archetypical first clinical symptom is weakness of tail tonus that progresses to paralysis of the tail, followed by a progression up the body to affect the hind limbs and finally the forelimbs. However, similar to MS, the disease symptoms reflect the anatomical location of the inflammatory lesions, and may also include emotional liability, sensory loss, optic neuritis, difficulties with coordination/balance, and muscle weakness and spasms.

2.3.2 Pathology and Immunology

Just like MS, EAE is characterized by perivascular inflammation [160], primarily in the spinal cord. Most EAE models involve infiltration of the parenchyma by $CD4^+$ α-β cells and activated macrophages, with some $CD8^+$ α-β cells, natural killer (NK) cells, γ-δ T-cells, and B-cells present, and CNS resident cells

(especially microglia) are also activated [167]. Some models also involve demyelination and a high frequency of relapsing disease. However, manipulation of the induction protocol can produce a wide spectrum of neuropathological patterns including demyelination, remyelination, gliosis and loss of axons.

The environmental trigger (immunization) initiates EAE, which leads to activation of potentially autoaggressive T-cells that home to the CNS [168]. Secretion of proinflammatory cytokines, such as tumor necrosis factor alpha (TNF-α) and interferon-gamma (IFN-γ) cause endothelial upregulation of adhesion molecules, such as vascular cell adhesion molecules (VCAM), which help the T-cells migrate across the BBB [169]. Neuroantigen-specific $CD4^+$ T-cells are then reactivated by myelin antigens presented on MHC class II molecules by local antigen-presenting cells (APC) including macrophages, microglia, and less efficiently by astrocytes [170]. Subsequent release of proinflammatory cytokines cause an up-regulation of MHC molecules on a variety of resident APCs before the peak of EAE [171]. This also activates the endothelium, which leads to a second wave of recruitment of T-cells and macrophages that cause tissue damage.

In recent years it has become clear that besides IL-12, the pro-inflammatory cytokine IL-23 has an important role in neuroinflamation in EAE [172], and that this is linked to involvement of Th17 T cells that primarily produce the cytokine IL-17 [173]. However, it is important to note that both IFN-γ and IL-17 have a dual role in the full spectrum of disease, with the respective Th1 and Th17 cells being associated with homing to different locations within the CNS [174].

A simplistic description of the cytokine network involved in EAE is that proinflammatory cytokines, such as TNF, LT-α, IFN-γ, and IL-12, have a disease-promoting role, while anti-inflammatory cytokines, such as TGF-β and IL-10 may protect from disease [160]. However, a more accurate description is more complicated, and involves the nature of the APC, the local cytokine micro-environment, selective engagement with co-stimulatory molecules and interaction with altered forms of the immunizing antigen.

2.3.3 Genetic and Environmental Influence

Similar to MS, both MHC and non-MHC genes control the development and severity of EAE [175]. The model is also sensitive to environmental influences and gene-environment interactions. In fact, the induction of disease itself represents a potential point of interaction. *M. tuberculosis*, which is part of complete Freunds adjuvant frequently used to induce EAE, has been shown to significantly influence the genetic control of EAE [176], and can override the influence of genetic loci that regulate EAE. Additionally, interactions between

genes and age at immunization or season at induction have been shown in mice [177].

2.3.4 MOG-EAE

Due to the extreme similarities in pathogenesis, we study the genetic regulation of MOG-induced EAE in rats [160]. Induction of the disease is achieved by immunization with MOG, which is a minor glycoprotein exposed on the surface of the myelin sheath. In addition to the pathology described above, this model involves demyelinating plaques with depositions of IgG and complement component 9, and glial scar formation (also apparent in MS)[175 178-180]. Immunologically, there are signs of activation of both cellular and humoral anti-MOG specific response, which is also remeniscant of MS, where both T- and B- cell responses to MOG and other myelin antigens are present [181 182].

Inbred rat strains show varying susceptibility to MOG-EAE, demonstrating a difference dependent on genetic regulation. Consistent with MS, the MHC locus (HLA in human) is the strongest susceptibility locus in EAE [175]. Indeed, when rats face the same MOG challenge, the MHC haplotypes determine the severity of subsequent disease [175]. However, additional genes affect disease susceptibility and course [183] (Figure 6).

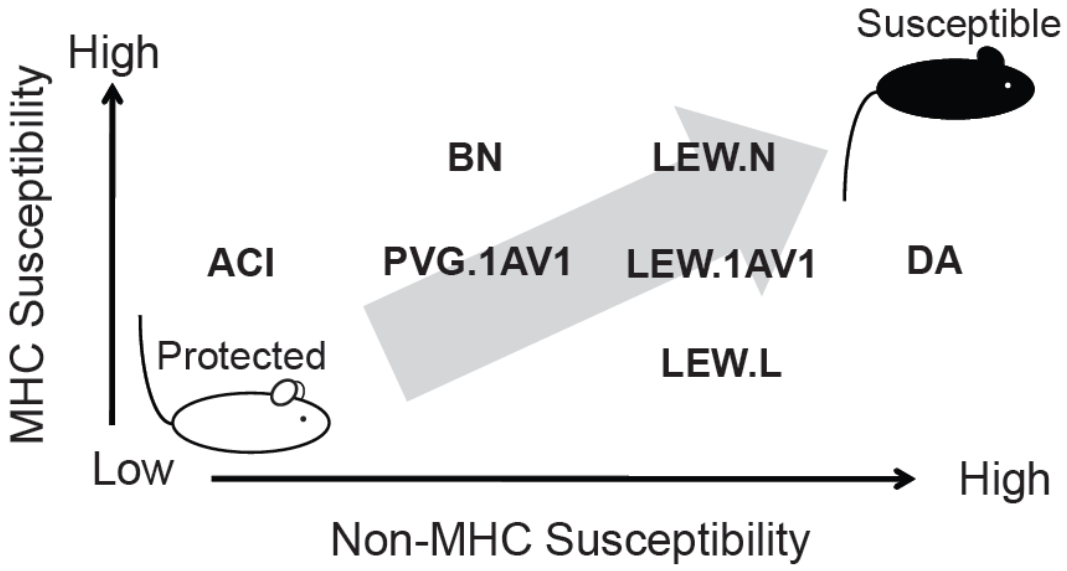

Figure 6. Schematic illustration of genetic susceptibility to MOG-EAE in inbred rat strains used in our Unit. The overall susceptibility is given by MHC and non-MHC influence.

With this model, several non-MHC genome regions have been identified that control either clinical susceptibility and severity, or that more specifically determine defined pathophysiological processes with regard to inflammation, demyelination or axonal loss. Loci that contribute to EAE with smaller effects have been present in several crosses [184-186], showing that the polygenic nature of MS is captured in the MOG-EAE model.

In this thesis, we use mainly EAE-susceptible DA and EAE-resistant PVG.1AV1 strains, or crosses thereof. To induce MOG-EAE in adult DA rats, inoculum containing incomplete adjuvant and MOG protein is immunized subcutaneously at the tail base [160]. MOG-EAE induction in PVG.1AV1 rats requires *M. tuberculosis* and a higher concentration of MOG protein to elicit a similar immune reaction. Typically, the DA rat will debut with clinical symptoms around two weeks after immunization and present with a relapsing-remitting EAE [160] (Figure 7). The severity of symptoms is scored on a scale ranging from 0 (healthy) to 5 (dead). Recovery from symptoms can be complete or partial and the relapse time varies with symptoms and disease severity. EAE phenotypes, such as incidence, disease onset, severity scores and duration of disease, can be evaluated for each strain or rat. Additionally, histopathological markers, risk gene expression, cytokine expression, cell populations and any feature of interest can be explored.

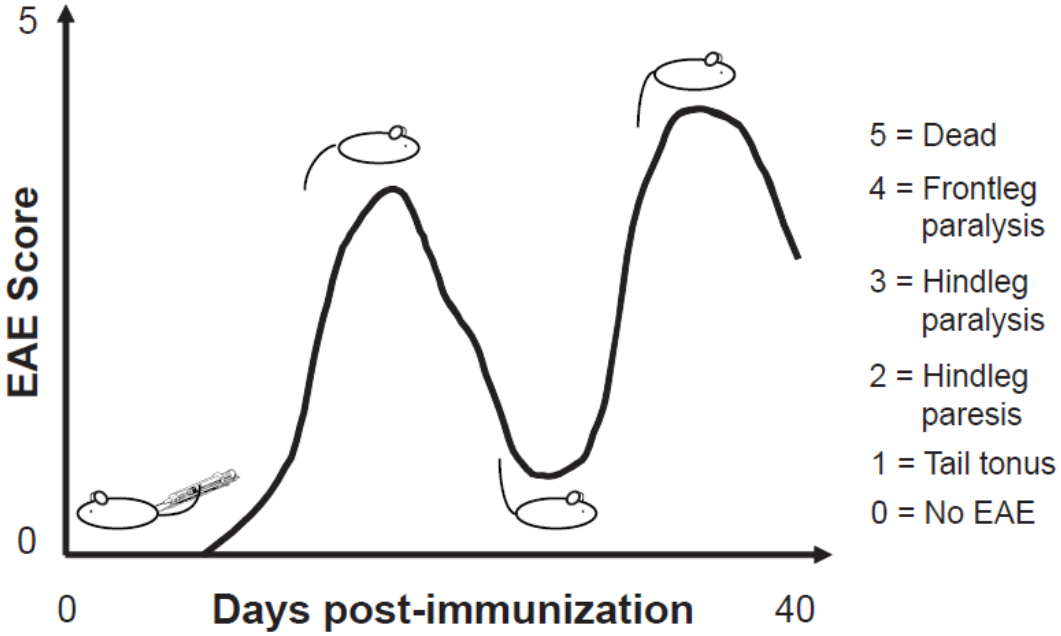

Figure 7. MOG-EAE model in DA rats. Tails are depicted with different degree of paralysis to represent the over-all status of the rat.

Identification of MOG-EAE-regulatory genes may improve the understanding of pathogenesis and resolve mechanisms of importance. Accordingly, this will clarify the inheritance of autoimmune neuroinflammation and, by extrapolation, of MS.

3 POPULATIONS UTILIZED

3.1 ANIMAL POPULATIONS IN GENETIC RESEARCH

Historically, experimental crosses provided increased heritability, flexibility and statistical power compared to human populations, and were used to identify genes to guide the search in the complex background of the human genome. The arrival of human GWAS and re-sequencing programs enabled screening of the human genome without *a priori* data, and therefore questioned the need for continued animal research. However, recent advances in human genetics research emphasize the necessity of multispecies platforms, as the plethora of susceptibility genes identified must be investigated for their roles and integrated into functional systems [187]. Additionally, many of the identified variants are located in gene deserts (~80%) and identifying the contributing mechanisms that underlie these associations remains a barrier to progress.

The use of animal models can complement human studies and help overcome these obstacles, since the genome can be manipulated to investigate the hypothesis in question. Animal strains that carry isolated genes (positionally cloned or disrupted) offer a unique opportunity to elucidate mechanisms underlying gene actions that contribute to disease. Reliable animal models that incorporate the genetic and environmental basis of disease can be used to predict drug responses in humans and potentially identify harmful or adverse effects. Additionally, environmental exposures can be investigated systematically under controlled conditions to chart their involvement in disease.

3.2 THE RAT AS MODEL ORGANISM

The aim of employing rat models is to better understand the complex systems involved in disease and translate that information back to humans. The laboratory rat (*Rattus norvegicus*) has for long been a favored organism for modeling physiology, pharmacology, toxicology, nutrition, behavior and immunology [188], where the size, ease of manipulation and breeding characteristics of the rat is advantageous. Conversely, the laboratory mouse (*Mus musculus*) became the species of choice for experimental genetics during the last century. However, the rat has become increasingly utilized in genetic research, as the characterization of the rat genome improved and resources developed [189].

Similar to many other laboratory animals, the breeding history of rat strains includes many rounds of inbreeding and interbreeding and several unknown relationships [190]. Due to several breeding stocks being initiated before inbreeding was completed, sub-strains are frequently not identical despite having the same strain designation, which can largely impact the phenotypic

outcome and reproducibility of experimental results [191-193]. To minimize this problem, we have established colonies at Karolinska Institutet that have been inbred for more than 20 generations (designation /Kini).

Selective breeding of rats who exhibit a phenotype of interest can be used to capture natural variation that lead to common traits and disease phenotypes. Alternatively, the genome can be manipulated to create a particular variant or disruption in a gene of interest. Gene targeting has been a challenge in the rat compared to the mouse, partly because of the absence of germline-competent rat embryonic stem (ES) cells. To overcome these obstacles, several alternative strategies were used, including N-ethyl-N-nitrosourea (ENU), transposon-mediated mutagenesis (reviewed in [194], [195-197]). However, the major advance in gene-targeting in the rat came with the establishment of germline-competent rat ES cells [198 199] and site-specific mutagenesis. By designing zinc-finger nucleases (ZFNs) engineered to target specific genes, it is now possible to create site-directed, heritable mutations in the rat genome [200 201]. Hence, the era of the knockout rats has arrived.

The current rat strain catalog contains more than 1000 strains, sub-strains and genetically modified rats (http://rgd.mcw.edu), including healthy inbred, transgenic, spontaneous mutant and complex trait model rats. It is often the case that a particular strain can be used to model several traits or diseases. Accordingly, the EURATools consortium has collected more than 100 phenotypes for behavior, metabolism, hemodynamics, hematology, immunology (Paper IV), morphology, and gene expression in the heterogeneous stock of rats (HS) [202]. Collecting multiple phenotypes in the same individual enables cross-comparisons and global characterization.

The genetic variation of rats contains approximately 3 million known SNPs [191], of which 10% are expected to have functional effects. Compared to the mouse, rat haplotype blocks cover less of the genome (12% compared to 35%), has a smaller proportion of informative markers (21% versus 56%) and smaller LD blocks (388 kb versus 648 kb) [191]. Pairwise correlations were observed among 0.2% of inter-chromosomal SNP pairs, likely reflecting epistatic interactions and ancestral relationships (since they disappeared when randomizing alleles). Estimates of mutation consequences revealed 56 SNPs likely to change protein function, 324 SNPs predicted to disrupt gene splicing, 1019 SNPs located in transcription factor binding sites and 132 SNPs predicted to affect microRNA targets. To summarize, the genetic architecture of the rat is complex with potential for mapping complex disease including epistatic interactions and epigenetic changes.

The rat genome was sequenced using the Brown Norway (BN) strain (BN/NHsdMcwi). BN is a founder strain for several important genetic

populations, including the HS [203]. However, BN turns out to be the most diverged strain, which may have consequences when comparing other strains to the reference sequence [191] (Figure 8). In any case, obtaining the rat sequence enabled detailed comparisons between mouse, rat and human. A subset of 20,000 SNPs have been genotyped in over 300 inbred strains and hybrid animals to construct high density maps, providing well characterized SNPs for quantitative trait locus (QTL) and disease gene mapping.

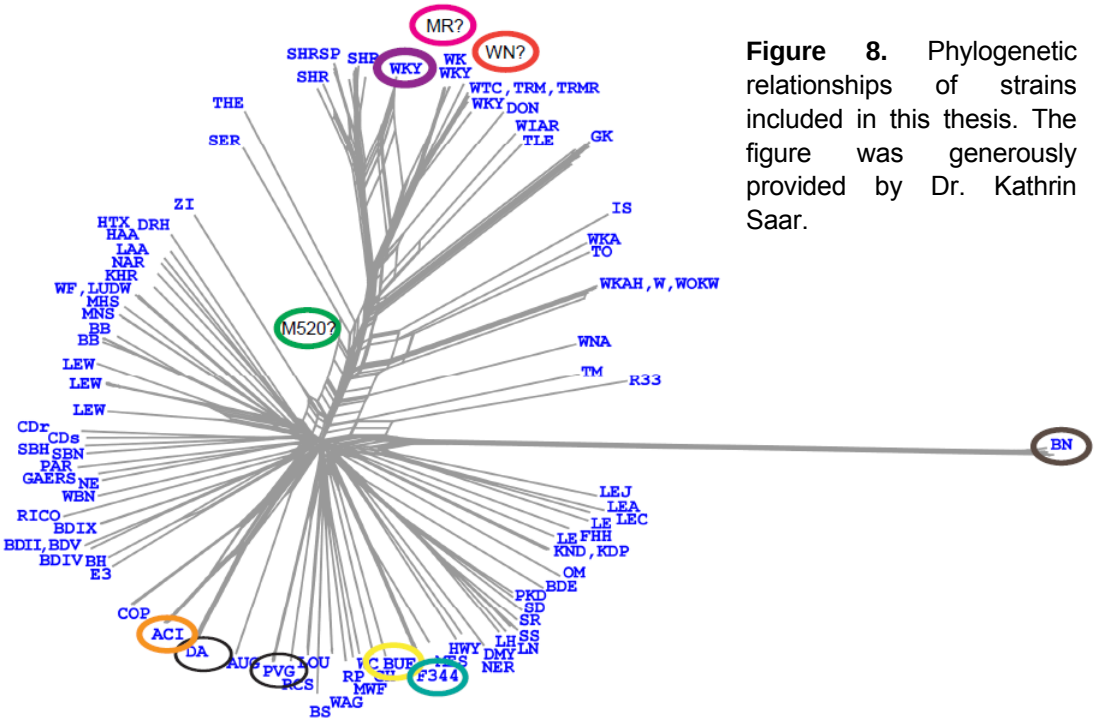

Figure 8. Phylogenetic relationships of strains included in this thesis. The figure was generously provided by Dr. Kathrin Saar.

In the past 20 years, the capacity of rat model research has evolved from positionally cloning monogenic traits to identifying genes that underlie complex diseases, including neuroinflammation [159] [204]. The rate of discovery will continue to accelerate as genome resources and mapping strategies improve.

3.3 INBRED RAT STRAINS

Inbred strains are families of rats where all members are genetically identical, or very close to identical. This is achieved by breeding brother and sister pairs for a minimum of 20 generations, which should achieve more than 99% identical genome [205].

The emphasis in studies of inbred strains is to identify a single, often extreme phenotype. The strain can then be exposed to different manipulations to study their effect. An example is traditional drug trials, where a group of individuals is exposed to a compound and compared to an unexposed group, to determine the efficacy of the treatment and potential hazards. It is also possible to extend this concept to two or more strains, in order to identify those who are sensitive

(susceptible) and those who are insensitive (resistant) to the same exposure, given that the strains are genetically different. This provides the building blocks for the more complex and diverse populations described below.

At the Neuroimmunology Unit, we use several inbred strains to evaluate neuroinflammatory phenotypes. DA/Kini and PVG.1AV1/Kini were used for Papers I-III. The 1AV1 designation means that PVG has been bred to carry the DA MHC II haplotype *(Rt1 av1)*, neutralizing the MHC effect to allow non-MHC genes to be investigated.

3.4 CONGENIC STRAINS

To study a locus or gene in isolation, congenic strains can be used. They are similar to inbred strains, with the exception that the genomic region of interest has been transferred from a donor strain (susceptible or resistant) onto a genetic background of different susceptibility (recipient strain).

A congenic strain can be produced by intercrossing two strains to create F1 hybrids, and then backcrossing the F1 to either parental strain (susceptible or resistant). The genetic recombinations will create unique animals and the aim is to select rats that carry the region of interest. They are backcrossed to the recipient strain for ten generations to ensure that the genomic background has minimum contamination with fragments of donor DNA (Figure 9). Alternatively, marker-assisted selection can be used (speed congenic), in which the background genome is screened to select the rat with least contaminating donor genome, to establish a homozygous congenic strain in 5-7 generations [206].

The purpose of studying congenic strains is to establish the function or effect of the gene/region that has been isolated. This is accomplished by comparing congenic rats to the parental strain, since the phenotypic differences reflect their genetic difference in the congenic region. This can thus be used to confirm the clinical relevance and biological impact of a QTL identified in a more complex population. Congenic rat strains can also be selectively bred to contain smaller overlapping portions of the original region, which is referred to as a panel of interval-specific recombinant (ISR) congenic strains [207]. Such panels can be used to narrow down the region and to in the final step positionally clone the gene of relevance [204 208].

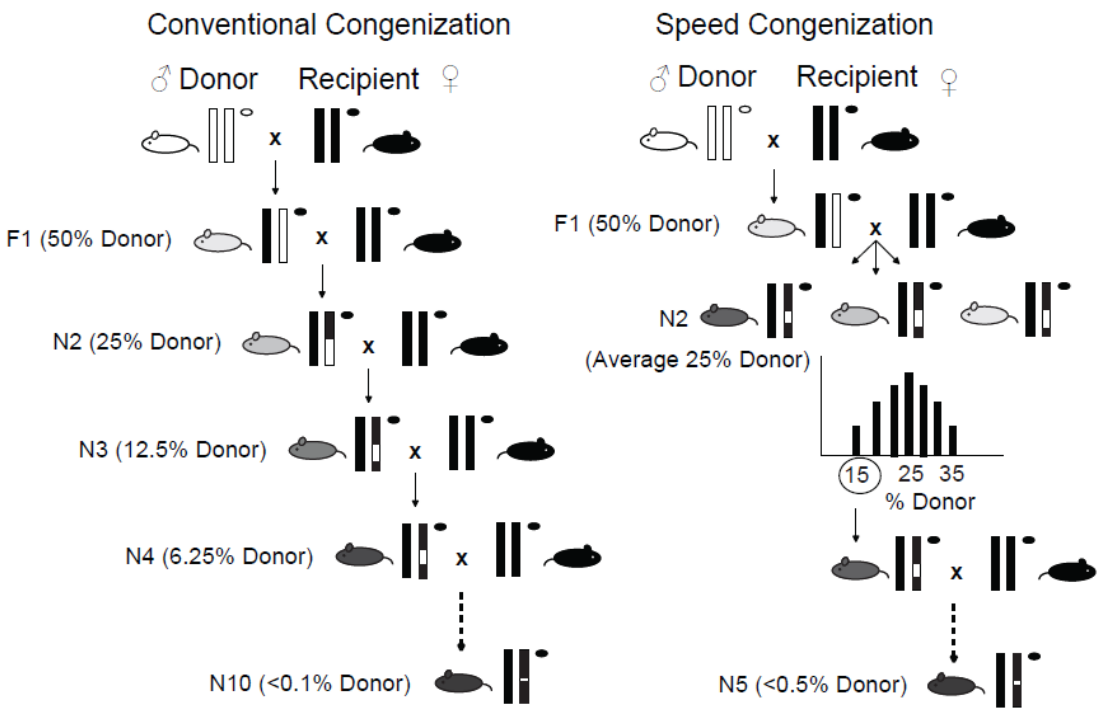

Figure 9. Schematic illustration of the conventional approach to construct congenic strains compared to the speed congenic approach [206]. One pair of autosomes is represented by vertical lines and mitochondria is represented by circles.

Congenic strains have been established for EAE-regulating QTLs on rat chromosome 4 (Paper I) and chromosome 17 (Paper II), with EAE-resistant PVG as donor strain and EAE-susceptible DA as recipient. The rats were exposed to MOG-EAE to study the effects on neuroinflammation exerted by the genes harbored in QTLs *Eae23-Eae27*. Additionally, a panel of ISR congenic strains containing overlapping portions of *Eae24-Eae27* was used to elucidate epistatic interactions that influence autoimmune neuroinflammation (Paper I).

3.5 INTERCROSS AND BACKCROSS

N2 backcross (BC) populations are used to identify genomic regions that are responsible for a phenotype, i.e. map a trait. The population is created by backcrossing F1 hybrids to one of the parental inbred strains [209]. The result is a population where each individual is genetically unique. However, the heterogeneity is reduced, because the uniqueness originates only with the F1 hybrid parent (Figure 10).

The aim of a cross population is to have phenotypic variation. Some individuals should exhibit a phenotype resembling the resistant parent and some should resemble the susceptible parent, while the majority of the population should continuously vary in phenotype. This kind of phenotype variation in a genetically heterogeneous population enables linkage studies [210], which identify genomic regions where a particular allele tends to be inherited together with the phenotype.

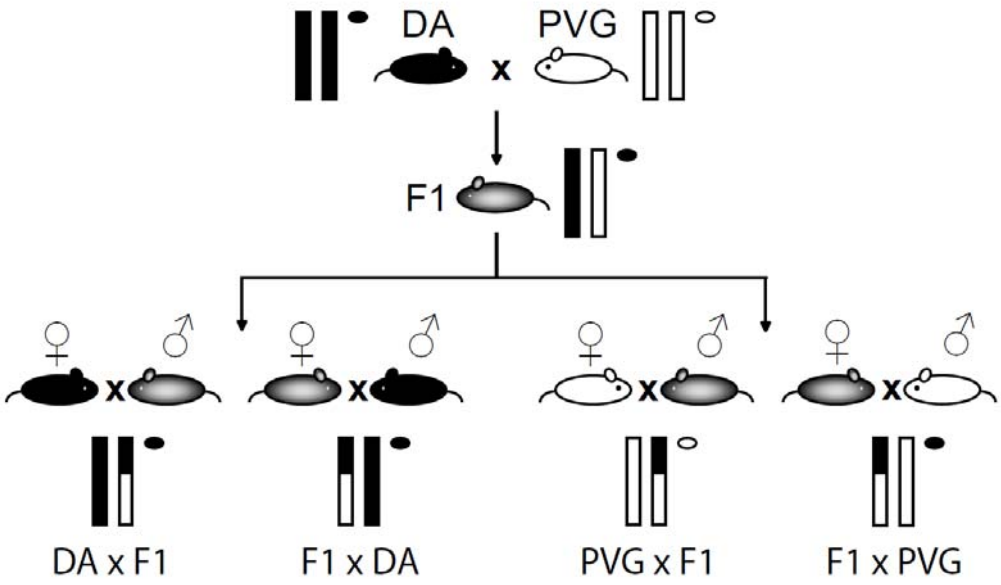

Figure 10. Schematic illustration of reciprocal BC design. This breeding set-up was used in Paper III. One pair of autosomes is represented by vertical lines and mitochondria is represented by circles.

In Paper III, we tweaked this approach to identify genomic regions that influence neuroinflammation in a parent-of-origin dependant manner. By using an F1 hybrid as one parent and an inbred strain as the other, we could determine which parent heterozygous alleles were inherited from [211]. Then, by using F1 hybrid mothers in half of the population and F1 hybrid fathers in the other half (a reciprocal cross), we could determine which parent the disease-predisposing alleles needed to be inherited from to have an effect (maternal, paternal or shared). Additionally, by creating one reciprocal cross with the susceptible strain and another with the resistant strain, we obtained all 3 possible genotypes to enable most allelic effects to occur. By manipulating which factors could be inherited from which parent in the separate crosses, we could create a comprehensive population where genetic and epigenetic influences on neuroinflammation could be ascertained.

Alternatively, all three genotypes can be directly achieved in a population by mating two F1 hybrid parents. This is called an intercross and the most commonly used generation is the second (F2) [210]. The mapping resolution can be slightly better compared to a BC, because both chromosomes in each pair are allowed to recombine. The drawback is that you have less information regarding which factors were inherited from whom. For example, it is no longer possible to identify which allele was inherited from a particular parent. An advantage for both types of crosses is that all genotypes can be obtained from the same set of parents, which means that the family structure of the population is known. This allows straight and accurate linkage mapping without adjustments to the data, and statistical methods can reliably calculate significance thresholds [212].

3.6 ADVANCED INTERCROSS LINE

The advanced intercross line (AIL) is a population that enables phenotype mapping at a higher resolution, compared to a conventional BC and F2 intercross [213]. QTLs mapped to a broad chromosomal region in F2 tend to carry several hundred genes, which makes gene identification challenging. An AIL can provide a more precise QTL location and can reduce the interval it spans, giving a better candidate list for gene identification [213].

The population is created much the same as F2 intercross populations, with the crucial difference that the breeding is continued in a pseudo-random fashion for several generations [213]. Accordingly, an increased number of recombinations accumulates over several generations, which provides the increased mapping resolution (Figure 11). For example, using the 10th generation (G10) provides an approximate 3.5-to-5 fold increase in resolution compared to an F2 intercross.

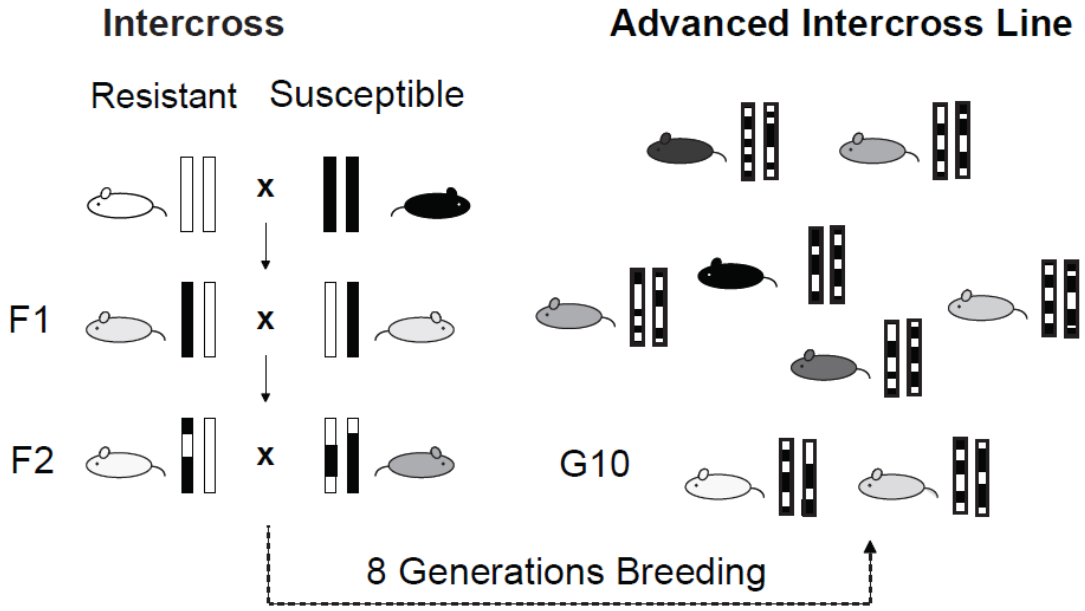

Figure 11. Schematic illustration of advanced intercross line breeding design. One pair of autosomes is represented by vertical lines.

The obvious advantage of the AIL is the increased resolution to accurately locate and define QTLs that influence neuroinflammation [162] [214]. Another advantage is that the scrambled (heterogeneous) genome can provide genotype combinations necessary for epistatic interactions. Since QTLs are allowed to interact with and be influenced by other loci, the AIL is likely to capture more realistic genetic effects in the complex setting.

A disadvantage in using a more complex population is that it does not meet the assumptions of established statistical methods. The statistical approach used to map linkage and establish significance thresholds and confidence intervals

must be adjusted for the underlying population stratification. Although this population is more complex than the BC/F2, it is important to emphasize that the heterogeneity is still vastly reduced by restricting the possible alleles for each gene to two.

An AIL between DA and PVG.1AV1 strains has been established at the Neuroimmunology Unit. This experimental population was utilized to fine-map previously identified genomic regions known to regulate MOG-EAE (Paper I and II). To define *Eae23-Eae27* and study the influence of these QTLs on neuroinflammation, the G10 rats were exposed to MOG-EAE and genotyped with densely spaced markers (approximately 1-5 Mb) in the defined regions. Interval-specific linkage analysis was used to refine the QTLs to several smaller regions, and the identified genetic effects were subsequently studied in congenic strains.

3.7 HETEROGENEOUS STOCK

The rat HS is a population that harbors recombinants derived from inbred strains that have accumulated over many generations of out-breeding to create a genetic mosaic. The HS can provide mapping resolution that is exponentially higher than an AIL, allowing fine-mapping of QTLs to intervals smaller than a cM [215]. Theoretically, it is possible to perform a GWAS of complex traits in the HS, to identify and fine-map QTLs in the same population. Thus, this approach combines the gene identification step usually performed in BC/F2 populations with the fine-mapping step usually done in interval-specific recombinant congenic strains and advanced populations [216].

The rat HS was established from eight inbred strains: ACI/N, BN/SsN, BUF/N, F344/N, M520/N, MR/N, WKY/N and WN/N (Figure 12) [203]. The MR, WN and WKY strains trace their ancestry to the original Wistar stock, the ACI strain is a hybrid between the August and Copenhagen strains, the BN strain trace its ancestry to the Wistar Institute stock of wild rats, and the M520, F344 and BUF strains are of unknown origin [203]. The European HS colony was established in 2004 at the Autonomous University of Barcelona obtained from the Northwestern University colony. The stock has been bred according to a standard pseudo-random out-breeding schedule through its 62th generation, using forty breeding pairs for each generation. The breeding scheme is designed to minimize inbreeding and maximize recombination density to reduce the size of inherited haplotypes [217].

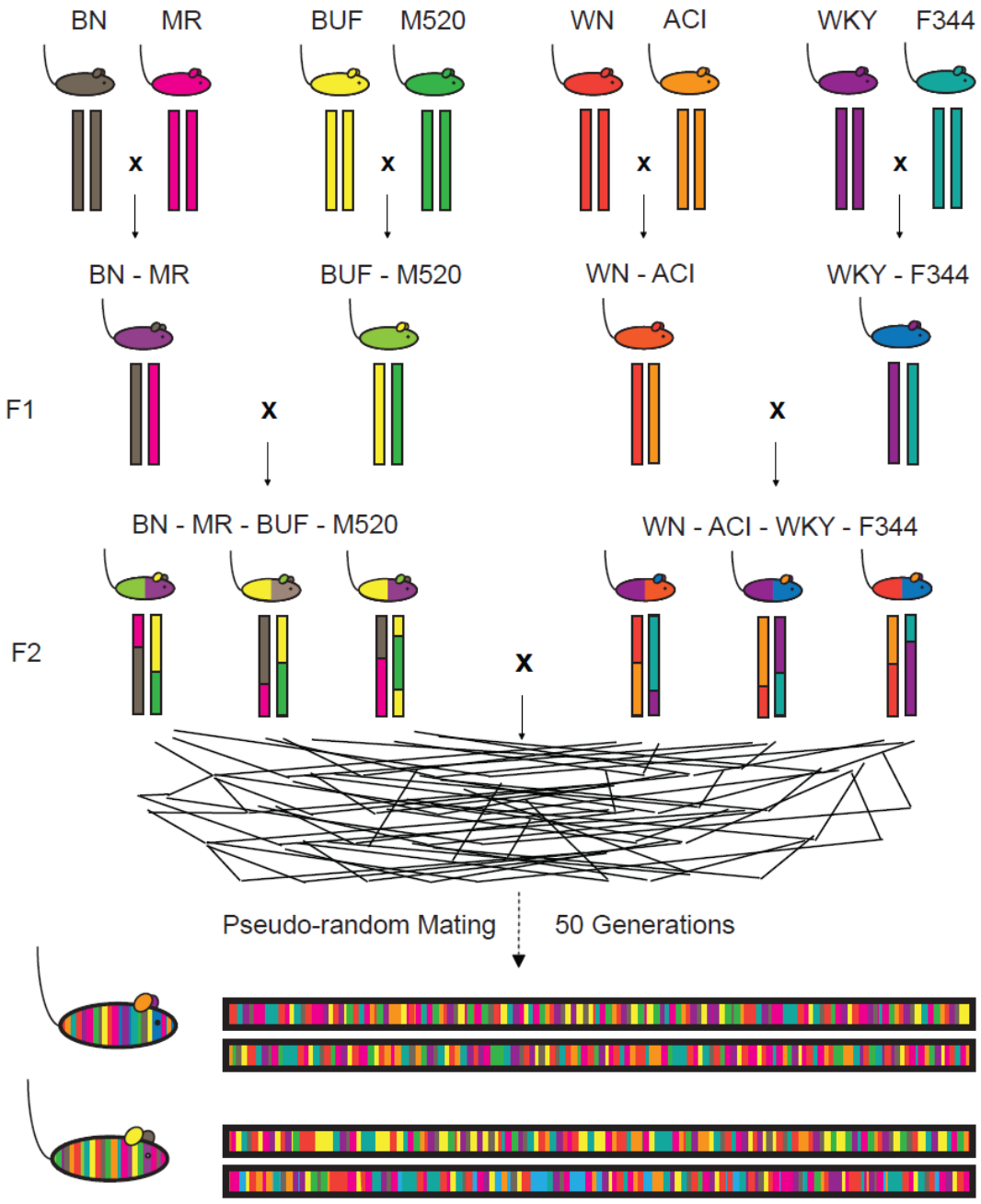

Figure 12. Schematic illustration of heterogeneous stock construction. One pair of autosomes is represented by vertical lines.

Complex trait analyses in a mouse HS has demonstrated successful fine mapping of approximately 100 phenotypes [215 218]. In total, 843 QTLs with an average 95% confidence interval of 2.8 Mb were identified. We therefore explored the rat HS to establish if this population could be used to investigate complex traits (Paper IV). The 'mappability' of the HS depends on its genetic constitution and on the quality of phenotypes it can deliver. Of 110 SNPs on two chromosomes, four were monomorphic, consistent with an expected rate of fixation of ~5% [215] and LD (measured as R^2) decayed to nonsignificant within 3 Mb. This is consistent with expectations that the rat HS can deliver high-resolution mapping and in agreement with the data from previous mouse

experiments [215 219]. Second, we needed to ascertain that the HS could deliver stable neuroinflammatory phenotypes equivalent to those we see in our inbred strains and crosses. It was also imperative to determine that there were no detrimental effects from serially collecting multiple phenotypes from each individual that could interact with the neuroinflammatory phenotypes of interest.

One particular concern was whether the HS could be used for dissecting EAE, because it contains several MHC types. This could potentially reduce the power for detecting non-MHC QTLs. These QTLs are the primary target for the HS studies since the MHC complex, and in particular the class II genes, are well characterized and studied by other means. Based on our findings and those reported in literature regarding EAE in the founder strains, we expected the HS to show variation in EAE susceptibility (Figure 13, [162 175 178 185 220-224]). However, it was important to perform pilot studies with the intended EAE model to establish if there was enough variation in disease outcome depending on non-MHC genes (Paper IV).

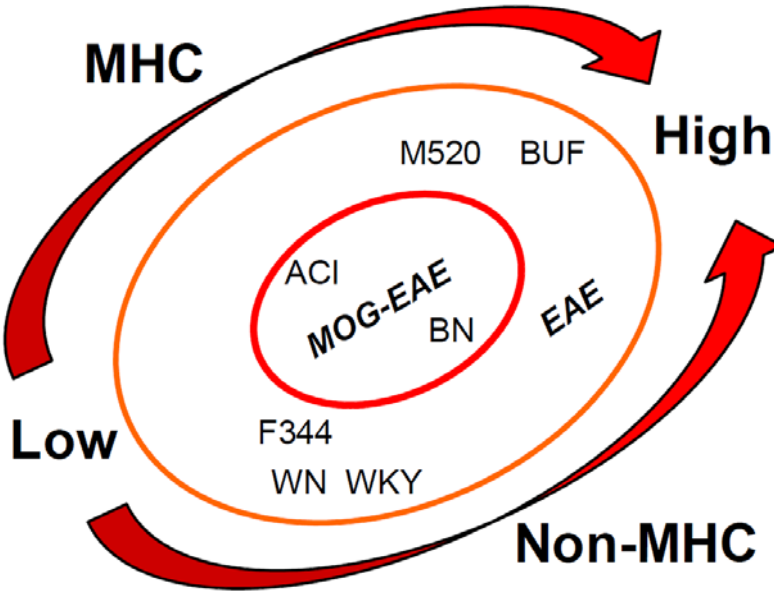

Figure 13. Predicted EAE susceptibility in HS founder strains, based on known MHC and non-MHC responses. The central circle show data for MOG-EAE while the peripherial circle shows data for other EAE models.

A considerable advantage of the HS compared to the AIL is that smaller regions are linked to disease in a system that more closely resembles a natural population (8 strains instead of 2). Identified QTLs of less than a cM may allow gene identification for some regions [225]. Another advantage is the capacity to elucidate small-effect QTLs, explaining even less than 2% of the variance [215 216 226 227]. A third advantage is that epistatic interactions and gene–environment interactions can be evaluated in a complex setting [215 228].

This is the first experimental population used in this thesis that even attempts to mimic features of a human population. The HS population is genetically much more diverse and complex than inbred crosses and has a wider range of phenotypes.

3.8 A COMBINED APPROACH

Using a combined approach to incorporate the strengths and weaknesses of each population described above into the study of autoimmune neuroinflammation allows us to identify QTLs that influence disease and functionally investigate their contribution to the pathophysiology. Genetic contribution to disease can be statistically investigated and determined in large intercross experiments, while the resulting biological effect can be elucidated in congenic strains harboring targeted loci. We can also model different aspects of disease in populations with differing properties and heterogeneity (BC, AIL and HS). The combination of studying biological differences and linkage analysis will define smaller and smaller QTLs with clinical relevance for neuroinflammation. The final goal is to identify and define a set of candidate genes or variants and understand how they operate in disease.

Overall, as genetic complexity increases to match complex phenotypes, experimental control and statistical simplicity decrease. The populations that offer the most control and statistical confidence are also the most genetically homogeneous populations (inbred and congenic strains), and therefore the most artificial models. On the other hand, populations that offer the highest probability of capturing the complexity involved in multifactorial traits are incredibly heterogeneous populations (HS) and are therefore more sensitive to the influence of family structure and confounders. To cope with this, rigorous statistical analyses using stringent criteria are necessary. This may partly explain why second generation crosses, which are in the middle of the complex genetics-statistical control continuum, are so widely used in genetic research. By incorporating all of these populations, we can identify genes and define their functions with high certainty (in the simpler populations), while also exploring their role in disease in combination with other genes, with less certainty but in a more relevant system (in complex populations).

4 AIMS OF THE THESIS

The initial aim of this thesis was to identify the autoimmune neuroinflammation-regulating genes responsible for the effects of previously identified QTLs and to define the mechanistic pathways involved. However, as the thesis work progressed, it became more evident that regulation of autoimmune neuroinflammation is much more complicated than initially anticipated and the genotype-phenotype relationships under study are influenced by a host of additional factors. It therefore became necessary to employ new strategies to discover such additional factors. The aim therefore progressed to include characterizing the inheritance of autoimmune neuroinflammation and identifying factors that contribute to the missing heritability.

More specifically, the aims were:

1) To fine-map EAE-susceptibility loci, primarily on rat chromosomes 4 and 17, by using an AIL and congenic strains.

2) To define mechanistic pathways and candidate genes by gene expression analysis, bioinformatics, and functional studies in relation to “disease genes”.

3) To dissect the extent of parent-of-origin effects to the etiology of autoimmune neuroinflammation.

4) To establish EAE in the heterogeneous stock of rats for GWAS study.

5 ANALYSIS OF GENOTYPE-PHENOTYPE RELATIONSHIPS

5.1 PRINCIPLES OF LINKAGE MAPPING

The search for genes that contribute to a disease often begins with a linkage study, with the aim of finding the locations of disease genes. Essentially, genetic linkage analysis is the statistical method of associating the phenotypic expression (or function) of genes to their location in the genome (Figure 14). The objective is to define markers that are linked to the phenotype in order to define QTLs where the disease genes are located.

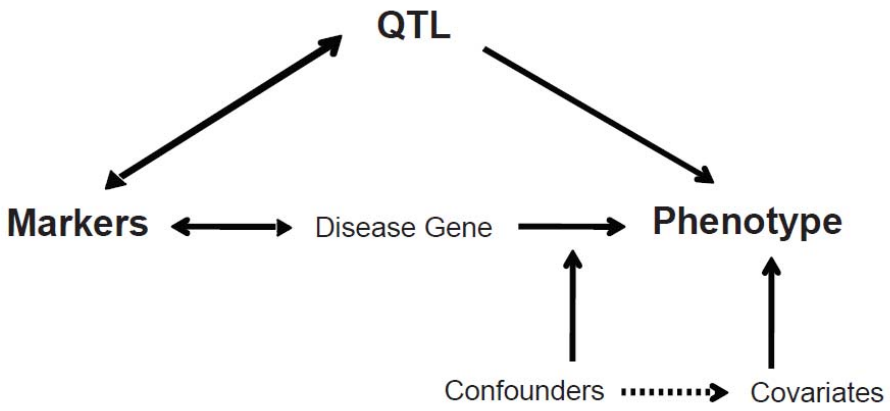

Figure 14. Schematic view of linkage analysis strategy.

When the disease gene is adjacent to a genetic marker, the disease phenotype will differ depending on the allele, represented by genotype at the specific marker. The closer the marker is to the gene, the larger the difference between alleles will be, reaching a maximum when the marker is at the exact location of the disease gene. By this approach, the whole genome can be systematically tested for QTLs, given a genetically segregating population that is variable for the phenotype, and that the phenotype is genetically regulated. The segregation of markers with the QTL and the association between QTL and phenotype are then modeled statistically to provide evidence for the QTL.

5.2 GENETIC MARKERS

Genome-wide screening became feasible with the advent of DNA markers (i.e. DNA sequences that vary in size or sequence), used with the polymerase chain reaction (PCR) amplification method. Microsatellites are highly heterozygous di-, tri-, or tetra-nucleotide repeats occurring relatively frequently in the genomic sequence (~10 kb), while SNPs are bi-allelic basepair substitutions that occur with high density (~800 bp). SNPs enable ultra high-throughput genotyping and development of dense genetic maps.

To be informative, a marker must be polymorphic between the parental strains of the population used for mapping (i.e. have different numbers of repeats or different nucleotide – A, T, C, G). The effect of the QTLs being mapped, together with the type and size of the population used, dictates the appropriate interval for marker spacing to achieve power to detect the QTLs. In general, marker intervals of 10-25 cM are appropriate for intercross and BC populations and approximately 1-2 cM marker intervals are appropriate for comparable power in the AIL. The HS requires 100 times more markers than a cross from inbred strains to allow haplotype reconstruction (see below) [227]. In this thesis, microsatellite markers were used for Papers I, II and III, while SNPs were used for Paper IV.

5.3 POPULATION CONSIDERATIONS

Phenotypic differences between inbred strains exclusively reflect their genetic differences, since they are exposed to the same environmental conditions. Finding the locations of QTLs responsible for these differences is possible in a cross because they carry mixed parental genomes with new combinations (described previously). Second generation crosses provide high power to detect linkage due to one round of recombination and all individuals being informative. The precision of QTL mapping is determined by the recombinant rate, and the average recombination fraction between two loci is r = 1/2 per meiosis. The population size required to detect QTLs depends on the phenotypic variance, the number of QTLs present, and the QTL effect size. These parameters can be difficult to predict, but generally, larger population have a better probability to detect several QTLs, including QTLs with small effects.

To estimate inheritance of autoimmune neuroinflammation, several genome wide linkage analyses in F2 intercrosses have been conducted in our laboratory [184-186 229 230]. Additionally, several identified QTLs have been fine mapped in the AIL [159 162 214]. In this thesis, the application of the G10 AIL to resolve two QTLs (Papers I and II) and the application of the BC to map parent-of-origin QTLs are reported (Paper III).

5.4 PHENOTYPE CONSIDERATIONS

Once the population is established, all rats are phenotyped. Phenotype selection is crucial in determining ‘mappability’ of the genes/QTLs. To enable a relevant, high-quality linkage analysis, phenotypes should reflect the investigated disease (or aspects thereof), screening should be objective and easy to perform, and statistical methods suitable for the phenotypic nature and distribution should be used. Phenotypes fulfilling these criteria should allow maximal extraction of information. The laboriousness of phenotype collection

becomes especially important for linkage analyses that demand large populations (such as our BC with ~900 rats and our HS with ~2000 rats).

Intermediate phenotypes, so called endo-phenotypes, usually have an improved signal-to-noise ratio due to fewer genetic factors influencing the trait, and are therefore likely to be detected with a higher significance and resolution [231]. They also have the potential to teach us about the mechanisms involved in pathogenesis. For example, rats usually lose weight 1-2 days before onset of neuroinflammation. Weight loss is a truly quantitative phenotype with normally distributed data, which provides additional information and accuracy, compared to other phenotypes. Incidence of neuroinflammation is a typical binary phenotype, designated as 0 (unaffected) or 1 (affected). Severity phenotypes are based on an ordinal scale that is not linear, since each increase in score does not reflect equivalent steps of deficits. In addition, these phenotypes display a discordant distribution, because all unaffected rats have phenotype values of 0 while affected rats follow a normal distribution. To appropriately map these phenotypes and avoid false positives, a model based on non-parametric statistics (that do not assume normal distribution) should be used.

5.5 MODELS FOR QTL MAPPING

Once the phenotype and genotype have been ascertained for all individuals in the population, the likelihood of existence, location and significance of QTLs are statistically determined by applying a model to the data. There is a difference in methods used to test single markers versus intervals, and in methods used to evaluate the existence of one QTL versus multiple QTLs. These different methods have advantages and disadvantages to consider when selecting the appropriate method to be used. It is important to remember that these models reduce a complex reality into a set of factors that can be tested. Therefore, model selection may never be perfect, but choosing an incorrect model can be detrimental.

5.5.1 Single Marker Test

The simplest method analyses single markers for differences in phenotype values between the marker genotype groups [232]. The comparison is made using a simple t-test, analysis of variance, or a suitable non-parametric test and a significant difference between genotypes indicate that the marker is linked to a QTL. Advantages of this method is the simplicity, that it does not require a genetic map or specialized software, and that it can be extended to include multiple QTLs and covariates. Disadvantages are that the QTL location is imprecise, because it depends on the marker location. In this thesis, we used single marker test to confirm our findings (Papers I-III).

5.5.2 Interval Mapping

The seminal article by Lander and Botstein introduced interval mapping, which tests a set of putative QTL locations along the genetic map [233]. Maximum likelihood is obtained where parameters are estimated to give the highest probability for the observed data (highest odds for linkage). The maximum likelihood solution is found in a reiterative process, by testing phenotype-genotype association based on a probability, and then re-evaluating the linkage based on the new parameters, until a QTL is detected. The strength of evidence for the presence of a QTL is given by a LOD score (logarithm (base 10) of odds favoring linkage). The LOD score is the odds between the likelihood of obtaining the given data if there is linkage compared to the likelihood of obtaining the same data purely by chance.

The advantage of interval mapping is that the QTL can be more precisely located, it gives a better estimation of the QTL effect, and accounts for missing genotypes and errors. Additionally, statistical stringency is used to correct for the multiple tests performed to screen an interval to decrease the risk of false positive QTLs. Haley and Knott developed a simple regression method for QTL mapping that approximates interval mapping very well, although it can be sensitive to strong epistasis and close linkage between QTLs [234]. This method can also be easily extended to mapping multiple QTLs and covariate analysis.

5.5.3 Multiple QTL Model

Both single marker and interval mapping methods model a single QTL. Multiple QTLs occurring simultaneously can be modeled instead, by using multiple regression methods. The simplest form is an extension of ANOVA and involves the same weaknesses as single marker tests. Multiple interval mapping [235] and composite interval mapping are extensions of interval mapping that account for the existence of multiple QTLs.

Composite interval mapping was developed to address the problem of mapping multiple linked QTLs [236], by conditioning the test on the effects from other selected markers. These are included in the model as covariates and increase the power to detect QTLs by reducing the variance attributed to markers. A disadvantage with composite interval mapping is that it is sensitive to uneven marker distributions, leading to a test statistic in a marker-dense region not being comparable to that of a sparse region. Additionally, this method suffers from difficulties in estimating the joint contribution of multiple linked QTLs and it cannot be directly extended to include epistasis.

To address the limitations of composite interval mapping, the multiple interval mapping model was developed, which facilitates multiple QTL identification, estimation of positions and effects, and discovery of epistatic interactions [235]

[237]. Multiple putative QTLs with associated epistatic effects are fitted directly into a model in four steps. First, an *evaluation* determines the likelihood of the data. This is followed by a *search* to select the best genetic model and an *estimation* of the best parameters for that model. Finally, a *prediction* is made regarding the genotypic values of individuals and their offspring.

An extension of the multiple interval mapping model uses stepwise regression with forward selection and backward elimination to construct the model in stages [238]. *Forward selection* identifies the most significant QTL, then adds additional QTLs sequentially that works best with the existing ones. *Reverse elimination* then sequentially removes the least significant QTL. The significance of each parameter is then tested with multiple interval mapping, and non-significant QTLs are dropped stepwise from the model. An advantage of this procedure is a dramatic reduction in numerical analysis burden, without losing accuracy of the likelihood evaluation. A disadvantage is the challenge of defining an appropriate stopping rule (how many QTLs to select for). The risk is that the model becomes 'overfit', catering too much to the given data to be generalizable to another data set. To prevent this, the complexity of the model is regulated by penalties applied for every QTL that is added to the model.

Multiple QTL models have several important advantages, with increased power, good positioning of QTLs, separation of linked QTLs, and the ability to map epistatic interactions, which are largely unpredictable and alters the genotype-phenotype relationship tested in single QTL analyses.

5.5.4 Confounding Factors

Confounders are factors that influence the phenotype, but that have not been considered and is therefore not accounted for in the model. They can be an important part of disease etiology with a true role in disease (unknown environmental factors, etc) or they can be extraneous factors that distort the result (experimenter differences, etc). To appropriately analyse a phenotype, it is important to identify confounding factors and either understand their role in disease or control them experimentally or statistically.

For example, sex, age and weight of the rat are known to influence experimental neuroinflammation [239] [240]. Therefore, we strive to reduce the variation in these factors. However, it is not always possible to control confounding factors experimentally. Therefore, we aim to record any extraneous factor with potential to introduce noise in the genotype-phenotype relationship (season, experimenter, time of day, etc) and perform a thorough cofounder analysis to identify such influences. To neutralize the cofounding effects, the phenotypic values can be adjusted to remove the extraneous influence. We used this approach in Paper III to neutralize differences between experimental sets and between the sexes. Another approach is to statistically

account for the cofounders' impact, by introducing these factors as covariates during the analyses. The genotype-phenotype relationship will then be conditioned on the covariate to analyze linkage using the selected model.

5.5.5 Significance of Detected QTL

The statistical analysis helps us evaluate the degree of confidence with which we can distinguish between a true genotype - phenotype association and the null hypothesis (there is no genotype-phenotype association). Significance, given by the p-value, is the probability of obtaining a different (greater) LOD score than the LOD given if there is no QTL in the population. A conventional p-value for significance is 0.05, indicating that there is 95% probability that the identified QTL is true. Corrections to this p-value must be made for testing a hypothesis on a number of locations. Another approach to setting significance thresholds is the permutation test [241]. The phenotypes and genotypes in the given data is disconnected and scrambled to randomly rematch one individuals phenotypes with one individuals genotype. The maximum LOD score acquired from this data and the frequency with which it occurs (95%) in the population is used to estimate the significance LOD score. This provides a significance threshold that is explicit for the given data.

To set significance thresholds in the AIL, we computed the mean for each family and the residual value for each individual from their family mean. This data transformation was applied to control for within family variance (inheritance of phenotype with the causing genotype, i.e what linkage is based on) and allow between family variance (representing random effects). The analysis was then repeated on this data to record the maximum LOD score acquired. To be stringent, we used the absolute max (100%) for each phenotype. We also knew the QTL was significant from the previous mapping studies.

5.6 GENOME-WIDE ASSOCIATION IN HS

The association studies in the HS involve a more specialized statistical approach, because of the population structure of individuals with different degrees of genetic relatedness [227]. The multitude of methods developed for genome-wide analyses in classical intercrosses are not applicable to this population [242-244]. A genome-wide analysis involves testing more loci than individuals, which makes fitting all markers simultaneously impossible.

To properly analyse the HS experiment, we use novel analytical methods and statistical packages developed specifically for the equivalent population in mouse [215 216]. To separate genotypes in the experimental population, parental haplotypes are reconstructed, using a hidden Markov-chain approach, to predict probabilities of inheritance from each of the 8 progenitor strains for each

SNP. A multiple QTL model is then fitted using a model averaging method to obtain a posterior probability that a QTL will be included in the model [215]. This is accomplished by repeatedly re-sampling the data and in each resample test which set of markers best explains the variation in the phenotype. Hence, the association between phenotype and genotype at any one locus is corrected by the pattern of associations over the rest of the genome.

The statistical packages that have been developed for analysis of the HS includes R/HAPPY and R/Bagphenotype. The haplotype reconstruction phase of analysis is carried out using the HAPPY software. The QTLs are then fitted using R/Bagphenotype.

5.7 FROM QTL TO CANDIDATE GENE

Our strategy has been to reduce QTL intervals to provide a candidate list for gene identification that can be screened. These genes can be assessed for coding sequence differences and expressional differences between susceptible and resistant strains. Further, functional aspects of candidate genes can be explored to indicate the functionally relevant candidates.

For analysis of gene expression, we used previously collected exon array data [245] to screen all genes in the region for differences between parental inbred strains. We then explored potential differences using the more sensitive quantitative real-time polymerase chain reaction (RT-PCR) method, in different organs collected from congenic and parental rats at an early and a late time-point in disease. Quantification was performed with SYBR green method and standardization for the amount of starting material was performed using GAPDH and β-actin. SYBR Green is a commonly used fluorescent dye that binds double–stranded DNA. As the target is amplified in each PCR cycle, the increase in fluorescence intensity is measured that is proportional to the starting amount of mRNA in the sample. Each sample is then normalized to the amount of a reference gene to give a relative quantity of expression of the target gene.

For investigation of the rat CNS, we compared histopathology between congenic and parental strains at day 35 p.i., when clinical EAE symptoms have been ongoing for approximately 3 weeks. Immunohistochemical analysis was performed on paraformaldehyde fixed paraffin embedded sections stained with hematoxylin-eosin and luxol fast blue to detect CNS inflammation and demyelination, respectively. The inflammatory index was determined from the number and size of inflammatory lesions on an average of twenty complete cross-sections of the spinal cord and the brain of each animal [246]. To assess cell infiltration, sections were stained with antibodies against CD8, CD3 and Foxp3.

Anti-MOG antibody levels in sera at day 14 p.i. (the approximate time of disease onset) were measured by ELISA. This enabled linkage study of an immunological sub-phenotype of EAE. The IgG isotype profile reflects the underlying cytokine environment in EAE, and in rat, IgG1 levels reflect the Th2-type immune response while IgG2b levels reflect the Th1-type immune response under specific circumstances [247] [248].

6 RESULTS AND DISCUSSION

The findings presented in this thesis illustrate the genetic complexity involved in inheritance of autoimmune neuroinflammation and in the mapping of complex traits, even when the focus is confined to a limited genomic region. EAE susceptibility and severity are clearly affected by multiple genes with a dissociation of effects on different aspects of disease, genes that interact with other genes and with environmental factors, and parent-of-origin effects likely to involve epigenetic regulation that mold how the genes contribute to the phenotypic expressivity.

6.1 GENETIC REGULATION OF AUTOIMMUNE NEUROINFLAMMATION

6.1.1 EAE is Polygenic (Papers I-III)

The polygenic nature of EAE shown in other studies [249] was also found in our studies (Papers I-III). The aims of Papers I and II were to fine-map susceptibility loci that had been previously identified to regulate EAE. These studies involved the hallmark steps of our classical mapping approach, described above. We refined the large original QTLs into smaller linked loci, by mapping them in the AIL which provided higher resolution and better precision. Our data show that the 58Mb region on rat chromosome four is composed of four distinct QTLs (*Eae24-Eae27)* and the 68Mb region on rat chromosome seventeen is composed of two distinct QTLs (*Eae23a* and *Eae23b*) (Figure 15). Paper III investigated regulation of autoimmune neuroinflammation on a genome-wide level in a BC, and we identified 11-16 QTLs contributing to EAE susceptibility and/or severity, depending on phenotype and cross. Each of the identified QTLs is likely to contain at least one gene that regulates EAE. These findings demonstrate that many genes are involved in regulating autoimmune neuroinflammation, and that even genetic influences from one QTL/region can depend on several genes. An example of this is a QTL identified in rat for blood pressure that harbored two closely linked genes, which also regulated blood pressure in humans, where the genes were no longer clustered [250]. These QTLs may reflect functionally related genes that are located in the vicinity of each other. Alternatively, it may be that the accumulated effect of genes was necessary to originally detect the QTL, and that these genes have similar effects in isolation to those that fall below detection/identification thresholds. The polygenic nature of MS is now well established in human genetic studies as well [112 116 119 120].

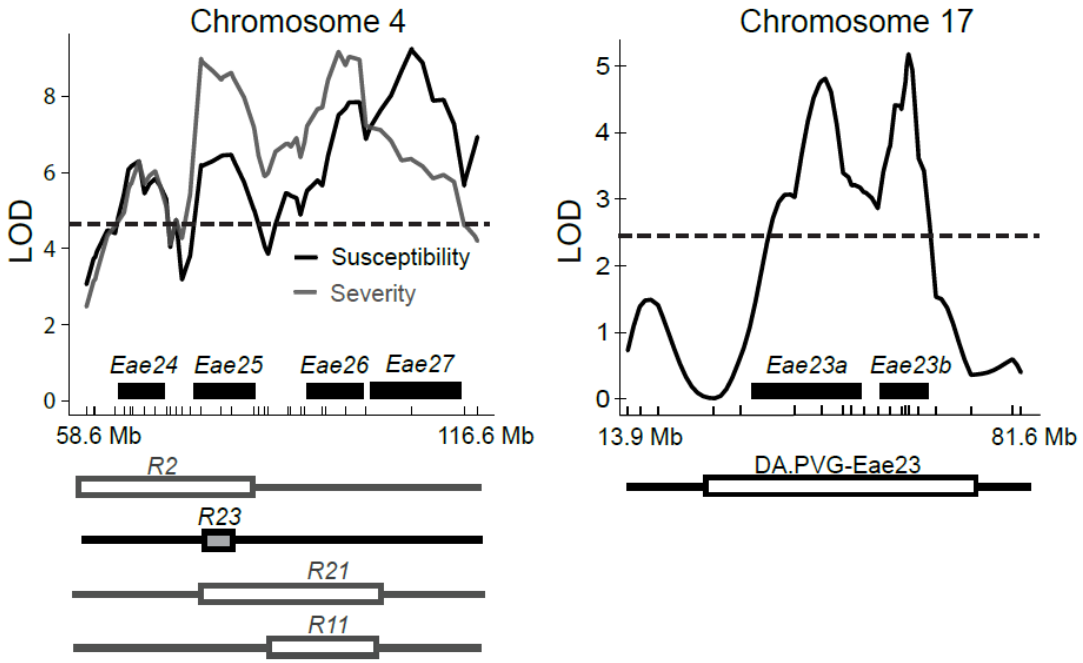

Figure 15. Linkage analysis and congenic strains support polygenic inheritance.

6.1.2 Dissociation of Genetic Effects (Papers I and III)

Although some QTLs regulate both susceptibility and severity of neuroinflammation (*Eae23a* and *Eaea23b*), one set of genes may primarily regulate susceptibility while another set of genes may modulate EAE once disease has arisen. In Paper I, we demonstrated that onset, severity and chronicity of disease are sometimes regulated by different genetic factors (Figure 15). The phenotypes capturing susceptibility in our statistical model are incidence and onset, while disease severity is captured by duration of disease and maximum and cumulative EAE scores. We demonstrated that *Eae24-Eae27* had a dissociation of genetic influence on different aspects of the disease in AIL and congenic rats (Table 1). Both *Eae24* and *Eae25* regulate susceptibility and severity. *Eae26* regulates severity but does not influence susceptibility, while *Eae27* regulates susceptibility only. Concordantly, Paper III also demonstrated several QTLs that regulated either EAE susceptibility or severity but not both. Additionally, other EAE-related phenotypes included in Paper I showed dissociation of genetic influence. Weight loss, which is often the first sign of EAE and may reflect subclinical disease [160], was regulated by *Eae25* and *Eae26*. The immune response mounted to the MOG challenge, measured by serum levels of anti-MOG IgG isotypes, was regulated by *Eae24* (IgG1 and IgG2b) and *Eae26* (IgG2b).

Table 1. QTLs regulate different aspects of disease in Paper I

	QTLs Included				Phenotypes	
	Eae24	***Eae25***	***Eae26***	***Eae27***	**Susceptibility**	**Severity**
	Suc & Sev	***Suc & Sev***	***Sev***	***Suc***	**INC & ONS**	**MAX, DUR & CUM**
	IgG	***WL***	***WL & IgG***			
R3	*X*	*X*	*X*	Partly	ns	ns
R2	*X*	*X*			ns	*
R23		*X*			*	***
R21		*X*	*X*	Partly	ns	**
R11			*X*	Partly	ns	*
R13			*X*	Partly	ns	ns

The interval-specific recombinant congenic strains investigated in Paper I contain different QTL combinations that regulate different aspects of disease. Abbreviations: Suc=susceptibility, Sev=severity, WL=weight loss, INC=incidence, ONS=onset, MAX=maximum EAE score, DUR=disease duration, CUM=cumulative EAE score.

The dissociation of genetic influence on different aspects of disease is important to acknowledge in order to uncover the molecular basis of autoimmune neuroinflammation. Hypothetically, polymorphisms within genes in these QTLs may cause an individual to pass the threshold to develop EAE, weaken resistance to the immune attack or shorten the time required to develop EAE, and thereby contribute to different aspects of disease. Furthermore, these genes could possibly regulate the severity of the EAE attack, chronic nature of the symptoms and disease course, and may also influence the recovery mechanisms either in the immune system or in the target organ [251]. Alternatively, this could reflect an unknown factor, such as environmental or genetic interactions. Genetic linkage of certain QTLs has been shown in mice to be dependant on the sex of the rat, induction protocol used and on season [177 252]. Although these factors may be altering the QTL effects in the AIL, the congenic experiments were all performed with female rats, and the same induction protocol. One advantage of the MOG-EAE model in DA rats is that disease can be induced without the use of *M. tuberculosis* or *pertussis*, alleviating the potential gene-environment interactions. The seasonal effects should be similar between strains, as we could not find significant differences between experiments for the same strain. It is therefore more likely that genetic interactions are contributing to the dissociation of

effects. This dissociation may complicate mapping efforts and contribute to inconsistent data when studies are compared. Careful analysis is required to establish which genes contribute to what aspect of disease, and how they operate in relation to one another.

6.1.3 Gene-Gene Interactions (Paper I)

Interactions of immunoregulatory genes have been suggested to play an important role in autoimmunity [253]. In Paper I, we statistically demonstrated that gene-gene interactions influence autoimmune neuroinflammation. The fact that there were no differences between the full-length congenic (*Eae24, Eae25, Eae26* and part of *Eae27*) and parental strains, while there were differences between both congenic strains that contained either the centromeric region (*Eae24* and *Eae25*) or the telomeric region (*Eae26* and part of *Eae27*) and the parental strain, further supported this finding. PVG alleles in the entire region weakened the protective effect of each QTL. Epistatic interaction, influencing susceptibility and severity, was detected between *Eae24* and *Eae25,* as well as an additive interaction between *Eae24* and *Eae27* which affected severity only (Figure 16). Additionally, the allele combinations in the region was a more important determinant for disease outcome than were the effects of each individual QTLs. *Eae25* alone influenced both susceptibility and severity and were protective during the late phase of EAE. Conversely, *Eae25* together with *Eae24* modulated severity and were protective during the acute phase of disease while *Eae25* together with *Eae26* and part of *Eae27* were protective throughout the disease course.

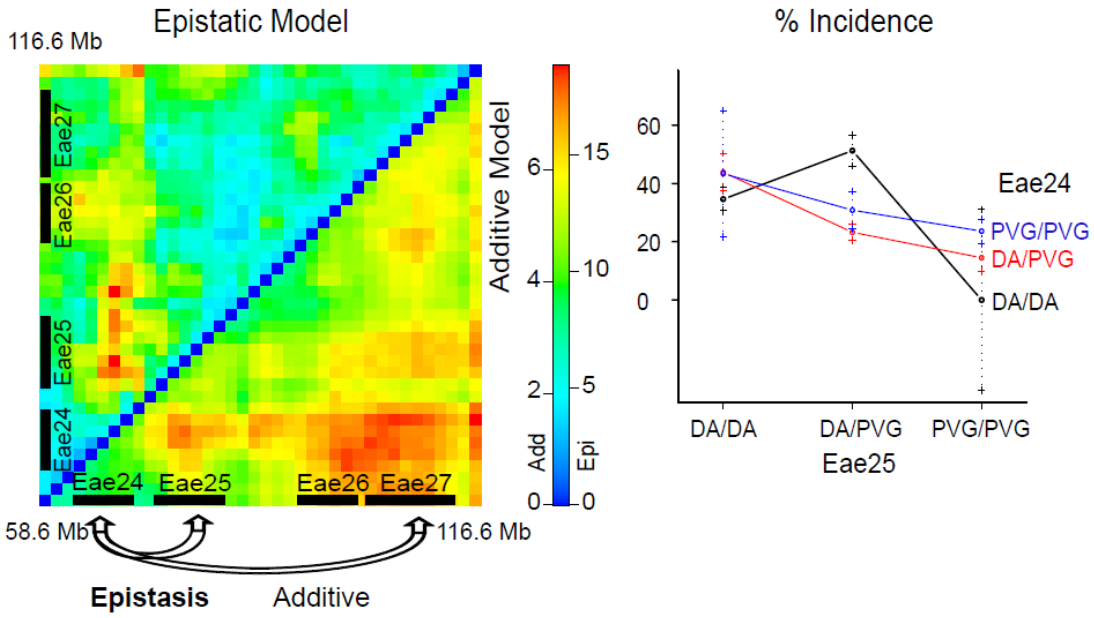

Figure 16. Gene-gene interaction model showing epistasis for *Eae24-Eae25*. Incidence is shown as a representative for all phenotypes.

These interactions can result from specific combinations of alleles at multiple loci that amplify or abrogate the independent gene effects. Such interactions have been demonstrated for the HLA and other autoimmune related systems. For example, the HLA *DRB1*14* allele abrogates the MS risk associated with *DRB1*15*, so that the risk is reduced 3-fold in *DRB1*14/*15* heterozygotes [254]. No functional explanation has been proposed for this effect. Conversely, the *DRB1*08* allele shows a modest increases in MS risk when in isolation, while it more than doubles the risk together with *DRB1*15* [254]. *DQA1*0102* increases MS risk in the presence of *DRB1*1501* while protecting against MS in the absence of *DRB1*1501* [156]. These examples illustrate how a risk variant with marginal effect when isolated, can exert a strong effect given the appropriate genetic background combinations. These epistatic effects within the HLA can explain an order of magnitude more of risk variance than the genetic effects confirmed to date with GWAS [112]. Although the HLA has a distinct genetic architecture, the epistatic interactions operating in this region are unlikely to be unique. Indeed, there is also evidence for epistasis involving non-HLA genes, but these interactions are not as well-characterized and need further validation [255]. Further, epistasis was demonstrated in a murine model for SLE, where the *Lmb3* locus would accelerate autoantibody production, T-cell activation and other features of lupus, dependent on *Fas* locus genotype [256].

Gene-gene interactions are important to consider when studying gene-disease relationships. For example, protective alleles can together with permissive interactive combinations partly abrogate each other to show a mild phenotype. On the other hand, protective alleles together with protective interactive combinations can amplify the independent effects of each other to drastically lower disease risk. This has implications both for gene identification and possibly for the treatment of disease. Genetic studies have been plagued by difficulties in corroborating results between studies. This may in part be due to interactions that modify the independent gene effects, which complicate the interpretation of results. If the contribution of a risk allele of gene A is only substantial in the presence of a certain allele of gene B, only part of the population harboring the “risk gene” A will be informative. Accounting for gene-gene interactions involved may therefore improve gene identification and the understanding of autoimmune neuroinflammation.

6.1.4 Implicated Candidate EAE Genes (Papers I and II)

Resolution to a small number of candidate genes enables identification of EAE-regulatory genes, which is a way to elucidate the underlying mechanism responsible for or contributing to autoimmune neuroinflammation. For *Eae23b*, fine-mapping produced a candidate list of 31 genes to be explored. Of these, the gene most likely to influence neuroinflammation is *ZEB1* (Paper II), which is

foremost known as an interleukin 2 (*IL2*) repressor [257 258]. This supports the involvement of the IL-2 pathway, which is already implicated in MS and EAE [111]. Interestingly, *ZEB1* splice variants were differentially expressed in our model and their regulation changed over time. In early disease, the short-form *Zfhep2* was up-regulated in the susceptible strain, but without subsequent dysregulation of downstream targets (*ZEB1* and *IL2*). Conversely, later in disease, *Zfhep2* up-regulation had led to lower levels of the full-length variant *Zfhep1,* and an up-regulation of both *ZEB1* and *IL2* in the susceptible strain. Although *Zfhep2* seem to initiate dysregulation of the system, repression of *IL2* expression was dependant on *Zfhep1* levels. This indicates that *Zfhep1* is the functional isoform and *Zfhep2* is the regulatory isoform in the setting of MOG-EAE (Figure 17). This finding probably reflects that both zinc-finger domains carried by *Zfhep1* are needed to be fully functional [259 260].

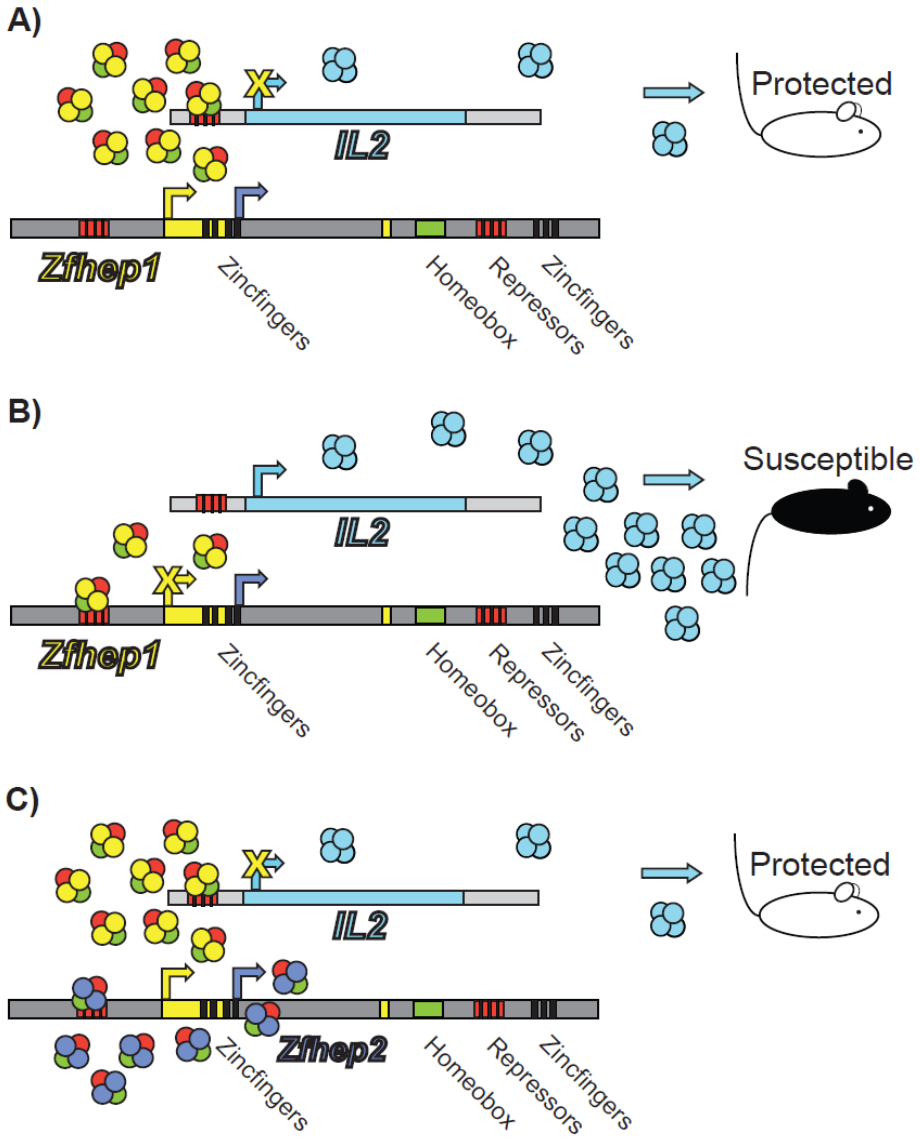

Figure 17. Schematic illustration of *ZEB1* gene regulation. A) *Zfhep1* (yellow) represses *IL-2*. B) *Zfhep1* can repress expression of its own gene, which allows *IL-2* expression. C) *Zfhep2* (blue) competitively binds *ZEB1* to allow *Zfhep1* expression, which represses *IL-2*.

We therefore postulated that the mechanism by which *Eae23b* regulates EAE is through regulation of alternative splicing of *ZEB1* which in turn regulates *IL2*. These findings indicate that expression and signaling delivered by specific splice variants may play a critical role in regulating autoimmune disease and also illustrates the importance of careful dissection of gene expression data.

The other QTLs characterized in this thesis contained a higher number of genes, but include candidates that are functionally interesting because of their effects in EAE and MS. Studies of the T-cell receptor Vß (TCRBV) cluster, located in *Eae24* (Paper I), have shown a possible involvment in MS and possible epistasis with the HLA [261-263], although conflicting data exist. Polymorphisms in these genes could be of importance in the T-cell receptor recognition of MOG peptides presented by the MHC. Conversely, the Src family associated phosphoprotein 2 (*Skap2*) mapped to *Eae25* and is known to influence B-cell function in rat MOG-EAE [264], and dendritic cell function in mouse EAE [265]. Furthermore, the known roles of TCRBV and Src family genes are consistent with the epistatic interaction between *Eae24* and *Eae25* we observed, hypothetically involving the same pathway of lymphoid cell activation [266]. Neuropeptide Y (NPY) is another potential candidate within *Eae25*, that has been shown to ameliorate MOG-EAE symptoms in mice [267]. NPY potentially has a protective role during the induction phase by inhibiting a MOG specific Th1 response. In the case of *Eae26*, the genes that encode IL-12 receptor-beta 2 [268] and IL-23 receptor [269] are interesting because of their role in the pathogenesis of neuroinflammation. Butyrophilin, of which subfamily 1 member A1 (*Btn1a1*) mapped to *Eae23a,* can ameliorate MOG$^{74-90}$-T-cell transferred EAE in DA rats[270]. The extracellular Ig-like domain of butyrophilin has been suggested to 'molecularly mimic' MOG, since butyrophillin can induce or suppress the ancephalitogenic T-cell response to MOG in these rats. Additionally, in mice butyrophilin can both prevent and suppress MOG-EAE, with subsequent suppression of IFN-γ, IL-2 and IL-12 [271]. These candidate genes can be tested in association studies of large MS case-control cohorts in parallel with functional studies performed in the experimental setting to better understand their role in disease.

6.1.5 Susceptibility and Resistance (Papers II and III)

Hypothetically, susceptibility to complex genetic disorders can be determined by an accumulation of genetic factors with weak to modest effects. Although we consider our inbred rat strains susceptible or resistant to autoimmune neuroinflammation, each strain harbours alleles that are disease-promoting and those that are disease-protective. In Paper II, we identified *Eae23a* as a transgressive QTL, with alleles from the resistant PVG strain promoting higher incidence, earlier onset, longer duration of disease and higher maximum EAE score. Conversely, *Eae23b* showed heterosis, where heterozygous DA/PVG

alleles promoted EAE with higher susceptibility and more severe disease compared to homozygous DA or PVG alleles. When we investigated allelic effects genome-wide (Paper III), alleles from the susceptible DA strain predisposed for disease at more than half of the identified loci. However, we also detected 31-36% transgressive loci where alleles from the resistant PVG strain predisposed for EAE. This indicates that every individual possesses risk genes and these are kept in balance by protective genes in those individuals who are resistant to disease. An individuals' genetic risk for developing disease (susceptibility threshold) is therefore determined in part by the combination or balance of alleles conferring risk and alleles protecting from disease, and the interactions between them. Indeed, when considering the combined effect of all four QTLs in Paper I, the allele combinations in the region were more important than the individual effect of each QTL in determining disease susceptibility and severity.

6.2 EPIGENETIC REGULATION (PAPER III)

The aim of Paper III was to determine the extent of parent-of-origin effects that contribute to the etiology of autoimmune neuroinflammation. To achieve this, we used a BC strategy designed to unequivocally identify the parental origin of disease-predisposing alleles under controlled environmental conditions. The parent-of-origin effects we observed implicate genomic imprinting, Y-chromosome and mitochondria in inheritance of EAE.

We identified several QTLs with parent-of-origin effects, with 44-73% of all, and 86% of the seven loci robustly identified in both backcrosses, displaying parental dependant transmission of risk alleles. Most of these loci were detected in offspring of F1 hybrid mothers and the majority of parent-of-origin QTLs were maternally inherited, 62-71%. Although the effect is larger, this is consistent with the maternal inheritance in MS [148-150]. This finding is also in accordance with documented and predicted higher maternal transmission of imprinted genes [272]. However, there were also QTLs with paternal transmission of disease predisposing alleles.

Parent-of-origin effect might also reflect influence of sex chromosomes. Males that had inherited Y chromosome from the susceptible DA strain significantly differed from males that had inherited Y chromosome from the resistant PVG strain. A lack of difference between females, which have inherited the same X chromosome combinations as the reciprocal males, argues against X chromosome influence. This suggests a significant contribution from the Y chromosome that influence global phenotypic expression. An example of such contribution is the BXSB male mice model for SLE, which harbors the *y-linked autoimmune accelerator* (*Yaa*) locus on the Y chromosome [273]. This locus interacts with other disease alleles to induce profound autoinflammation [274].

Although this represents a specific case of translocation of the Toll-like receptor 7 (*Tlr7)* from the X chromosome onto the Y chromosome [275], similar effects may contribute to EAE in the rat. *Yaa* has been demonstrated to suppress disease and the autoantigen-specific humoral and T cell responses in mouse models of EAE [276] and arthritis [277]. An alternative explanation could be the intrauterine positional effect described in the previous mouse EAE study, where the Y chromosome influenced EAE outcome in both males and females [276]. This effect was suggested to involve the influence of neighboring fetuses on the *in utero* hormonal environment. However, we could not detect an effect of the Y chromosome on females or a skewed female to male ratio at birth. This could reflect differences in species, cross types and disease induction.

Unlike Y chromosome, mitochondrial genome is maternally inherited and can contribute to maternal transmission. Our data imply regulation of EAE by the mitochondrial genome, with 73% of loci displaying parent-of-origin effect in the cross where mitochondria was allowed to vary compared to 44% in the cross with fixed mitochondria. The strongest mitochondrial influence was seen for the locus on chromosome 10, which displayed linkage to all EAE phenotypes in crosses that share the DA mitochondria but not in animals with mitochondria from the PVG strain. This suggests that the QTL effect is modified by the mitochondrial genome. This may potentially involve mitochondrial oxidative stress, which has been shown to interact with autosomal genes to modify neurodegeneration in optic neuritis [278]. Additionally, oxidative damage to mitochondrial DNA has been found in active MS plaques [279].

Additionally, loci on chromosomes 3, 4, 5, 6, 10, 14 and 18 displayed a pattern of inheritance resembling imprinting. The locus on chromosome 18 overlaps a mouse QTL that was previously identified to predispose for EAE in a genomic imprinting fashion [280]. The paternal transmission of EAE risk in the QTL on chromosome 6 could potentially be explained by paternally expressed genes in the *DLK1*/*DIO3* cluster [281]. *DLK1* is involved in B lymphocyte differentiation and function [282], and may influence proinflammatory cytokine expression [283]. Additionally, a SNP in the *DLK1* region associated to type-1 diabetes showed reduced paternal transmission of the protective allele [284]. The imprinting pattern may potentially indicate loci that are susceptible to gene-environment interactions. As mentioned previously, studies in rodent models have illustrated that environmental signals can remodel epigenetic marks that lead to altered gene expression and thereby influences the phenotype [66-68]. Many features of the immune response are under regulation of epigenetic mechanisms [285], that may be important in MS pathogenesis. The differentiation of naïve $CD4^+$ T-cells into effector Th1 and Th2 cells and subsequent cytokine production involve epigenetic regulation [286] of *IL4* [287 288], *IL2* [289], and IFN-γ [290]. Treg cells, which are fundamental in maintaining immunological self tolerance, are in part

directed by epigenetically regulated *FOXP3* [291] [292]. Decreased expression of F*OXP3* has been found in MS [293] and several other inflammatory diseases [294]. By using an MS-like animal model, histone deacetylase inhibitors have been shown to reduce spinal cord inflammation, demyelination and loss of neurons and axons [295]. Thus, failure to maintain epigenetic homeostasis can result in deviant gene expression that cause a loss of tolerance, which can lead to development of autoimmunity in individuals who are genetically predisposed for disease [296].

In study III, we demonstrated that stratifying the data for parental origin increases the power to identify genetic factors that contribute to disease pathogenesis. We could also explain more of the phenotypic variance by modeling the parent-of-origin contribution to disease together with the genetic contribution identified (Figure 18). In fact, accounting for these effects defines risk factors that explain 2-4 fold more of disease variance compared to the factors identified in populations confounded by the parent-of-origin. This may be a large contribution to the previously unknown part of inheritance, suggesting that parent-of-origin can explain part of the missing heritability. This significant contribution of parental origin to inheritance of autoimmune inflammation implicates a role of epigenetic factors, mitochondria and sex chromosomes in pathogenesis of inflammation. Incorporating these effects into the model of inheritance does not only enable more powerful and precise identification of risk factors but can provide a better understanding of the pathogenesis of complex inflammatory diseases.

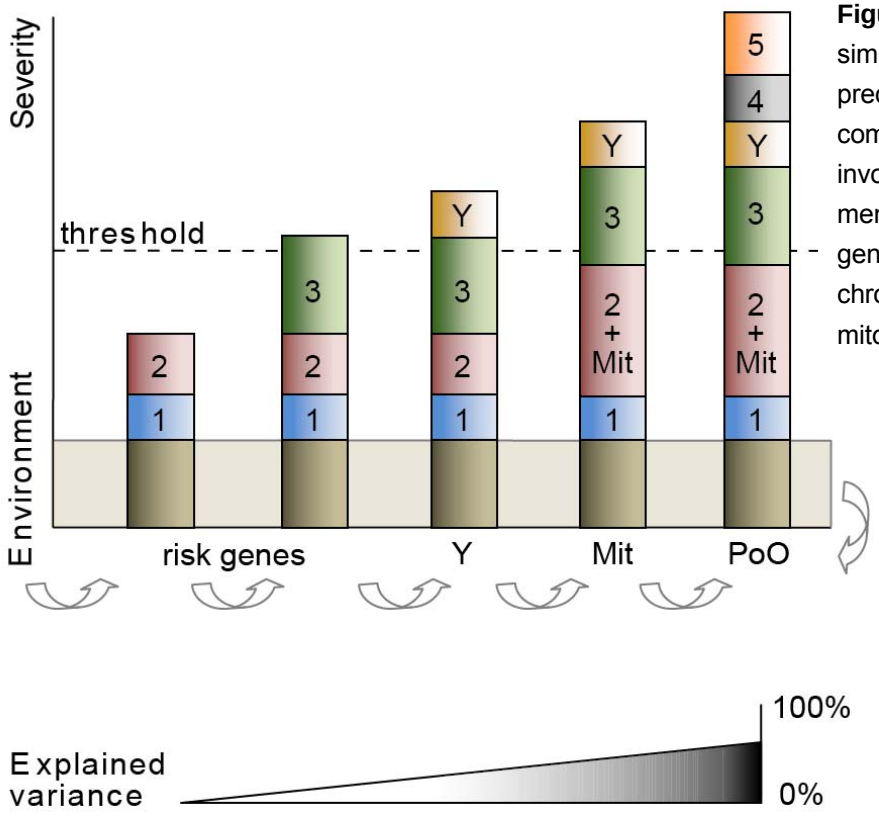

Figure 18. A simplified model of predisposition to complex diseases involves environmental factors, risk genes, epistasis, Y chromosome and mitochondria.

6.3 EAE CAN BE MAPPED IN COMPLEX POPULATIONS (PAPER IV)

To better match the complexity involved in autoimmune disease, we explored the possibility of using HS rats to map EAE. The purpose is to complete a genome-wide genetic association study in a high-resolution outbred population. We set out to establish MOG-EAE as a model in the European colony of the HS. To establish that disease could be induced in the HS rats, we performed an experiment in 50 HS females to test the effects of induction severity (incomplete or complete Freunds adjuvant with 50µg MOG/rat). A stronger disease induction and larger phenotype variation led us to select the more aggressive induction protocol, including *M. tuberculosis,* for future immunizations (Table 2, Figure 19). The pilot also showed that EAE phenotypes have a wider distribution in the HS than in traditional two-strain crosses, which indicates that the HS population allows more nuances in genetic regulation of EAE.

Table 2. Titration Experiment

	INC	ONS	EAE Score						MAX	CUM	Dis. Course		
			0	1	2	3	4	5			M	RR	C
IFA	11/25	17	14	7	2	2	0	0	0.7	37	5	1	5
CFA	18/25	15	7	3	4	3	0	8	2.4	68	2	5	3

Abbreviations: INC=incidence, ONS=onset, MAX=maximum EAE score, CUM=cumulative EAE score, M=monophasic, RR=relapsing/remitting, C=chronic.

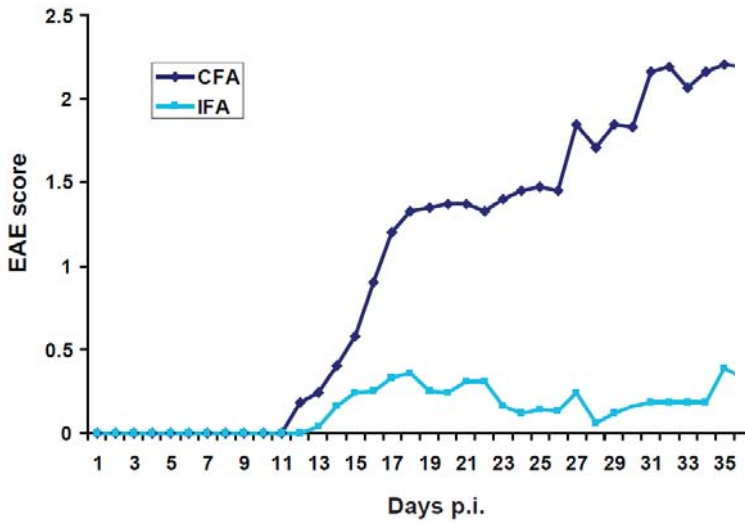

Figure 19. Rats were immunized with 50 µg MOG in Complete Freunds Adjuvant with 200 µg MT (CFA, N=25) or 50 µg MOG in Incomplete Freunds Adjuvant (IFA, N=25).

We also investigated the influence of MHC haplotypes on disease outcome in this pilot cohort. Rats with MHC types that we have experience with were

selected, AV1 and N (homozygotes and heterozygotes, respectively), and compared to rats with all other MHC types (B, L, LV1 and D). Although there were no significant differences between MHC groups (p=0.06), 72% of rats with the permissive AV1 type developed disease while only 1 of 4 rats with the highly permissive N type developed disease (Figure 20). The probable explanation for the lower disease incidence in the N group is that these rats were heterozygous for N and had only part of the susceptibility effect. The group that contained a mixture of MHC types developed disease in 40% of rats. This indicated that the expected MHC influence was present, but did not dictate disease outcome completely, suggesting that part of the influence comes from non-MHC factors.

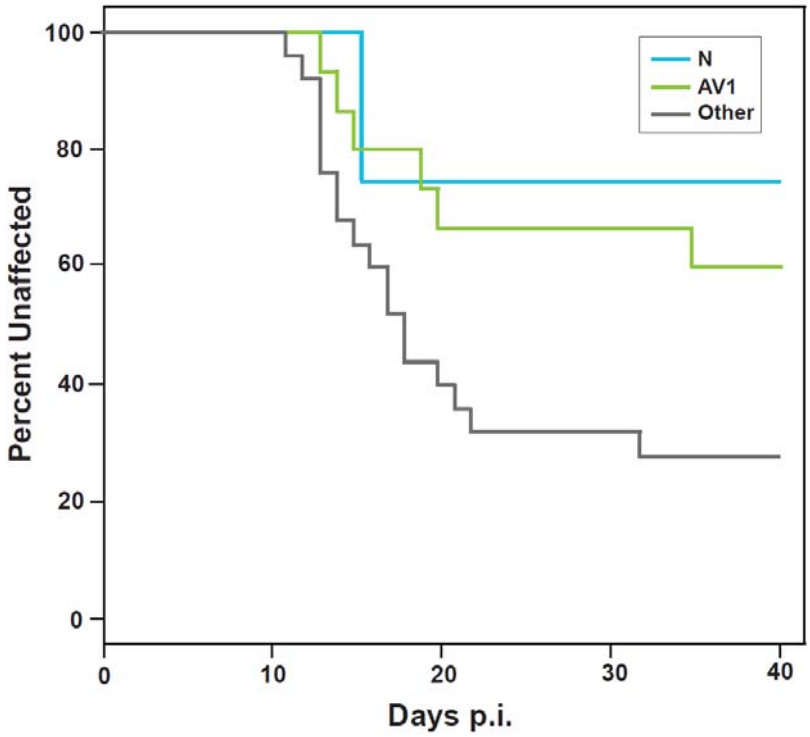

Figure 20. No differences in time to disease between MHC types.

Assuming that part of the non-MHC influence was genetic, we performed a pilot study of 25 rats to determine the variance of phenotypes within MHC types. Two homozygous MHC groups could be identified, AV1 and L, while the other group contained MHC mosaics of varying kind. The large phenotype variations within the various MHC haplotype groups strongly suggest that influence from non-MHC genes can be mapped in the HS (Table 3). This prompted us to continue phenotyping EAE in a larger HS population (Table 4), with the result that we have now phenotyped more than 2000 HS rats.

Table 3. Range of EAE phenotypes within MHC types

MHC Type	EAE Phenotype				
	INC	**ONS**	**MAX**	**DUR**	**CUM**
L (N=4)	50%	13-15	0-5	0-26	0-104
AV1 (N=5)	100%	13-18	1-5	12-23	4-99
Other (N=16)	69%	11-23	0-5	0-28	0-110

INC = Incidence of EAE, ONS = onset of EAE, MAX = maximum EAE score, DUR = duration of EAE, CUM = cumulative EAE score.

Table 4. EAE phenotype distributions in HS rats

Exp.	MHC Type	EAE Phenotype				
		INC	**ONS**	**MAX**	**DUR**	**CUM**
1	Mixed (N=217)	41%	9-40 (29±13)	0-5 (1.1±1.4)	0-20 (5±7)	0-86 (11±19)
2	Mixed (N=270)	33%	11-40 (32±12)	0-5 (0.7±1.2)	0-18 (3±5)	0-87 (5±11)
3	Mixed (N=230)	29%	10-40 (33±11)	0-5 (0.8±1.3)	0-19 (4±6)	0-90 (9±17)
4	Mixed (N=270)	39%	9-40 (30±12)	0-5 (0.9±1.2)	0-19 (5±7)	0-78 (10±15)
5	Mixed (N=240)	21%	9-40 (35±10)	0-4 (0.5±1)	0-18 (2±5)	0-59 (5±11)
6	Mixed (N=240)	21%	11-40 (36±9)	0-3 (0.5±1)	0-18 (2±5)	0-49 (4±11)
7	Mixed (N=270)	19%	10-40 (36±9)	0-4 (0.5±1)	0-17 (2±4)	0-47 (4±10)
8	Mixed (N=270)		In Progress			
	Total N = 2007	**29%**	**33±11**	**0.7±1.2**	**3±6**	**7±14**

The range of each phenotype is shown with mean and standard deviation in parenthesis. INC = Incidence of EAE, ONS = onset of EAE, MAX = maximum EAE score, DUR = duration of EAE, CUM = cumulative EAE score.

The primary aim of this study will be to identify genes that regulate autoimmune neuroinflammation in a high-resolution mosaic population, with the expectation of achieving QTL intervals of at best 1-3 genes. Additionally, other phenotypes that have been collected in the same rats will be combined to investigate additional factors that contribute to EAE. For example, blood cell counts and naïve immune response to LPS may help us characterize how the rats baseline constitution dictates its potential to respond to the MOG challenge. Further, the effects of stress on neuroinflammation, something that has been debated as a possible trigger of MS [297], can be assessed since we have obtained data on the propensity for stress and susceptibility to EAE in each animal. Additionally, tissues such as serum, spleen, spinal cord and adrenal glands have been collected for expression or other follow-up studies, which can add another dimension of investigation. We established that there is enough MHC-independent variation in disease outcome to map non-MHC QTLs in the HS, in spite of several MHC being included. This can offer a great advantage compared to our previous studies, in that gene-gene interactions with a variety of MHC haplotypes can potentially be unraveled. Further, environmental factors that may influence the results, such as season, weather, temperature and humidity, are recorded to enable the study of gene-environment interactions.

The results of our studies so far have indicated that the complexity of genetic regulation of autoimmune neuroinflammation is tremendous, even within QTLs. The HS population holds the potential to fine-map EAE in a mosaic population, which can provide the tools to tease out the complex relationships of contributing factors on a whole genome scale.

7 CONCLUSIONS

The thesis work has allowed me to draw several important conclusions about the inheritance of autoimmune neuroinflammation in rats.

First, autoimmune neuroinflammation is genetically regulated and is a polygenic trait (Papers I, II and III). Hence, susceptibility to develop disease, and the disease symptoms and course once it manifests is partly inherited from parent to offspring.

Second, it is possible to dissociate the genetic effects on different parts of the disease. Susceptibility, severity and chronicity of neuroinflammation are to some extent controlled by separate genetic loci (Papers I and III). That is important to acknowledge, in order to uncover the molecular basis of neuroinflammation and the different mechanisms contributing to different aspects of it.

Third, epistatic interactions contribute to susceptibility and progression of neuroinflammation, and can also influence a particular aspect of the disease (Paper I). These interactions can alter the main effect exerted by a risk gene in isolation and may affect the genotype-phenotype relationship. It is essential to identify and understand any interaction a risk gene is involved in or affected by in order to draw accurate conclusions about the genes' effect.

Fourth, epigenetic mechanisms, particularly parent-of-origin effects, contribute to susceptibility for neuroinflammation and may modulate disease phenotypes after disease initiation (Paper III). This adds another level of complexity to consider when studying genotype-phenotype relationships. However, epigenetic influences can also help explain the intricate interplay between genetic and environmental factors that contribute to neuroinflammation (Figure 21). This is likely to explain part of the missing heritability involved in neuroinflammation.

Fifth, the heterogeneous stock population is suitable to use for genome-wide association studies of autoimmune neuroinflammation and other complex traits. This population holds the potential to provide identification of neuro-inflammatory QTLs across the genome, each representing a small genetic interval that potentiates gene identification, resolution of closely linked QTLs, and identification of epistatic interactions genome-wide, including the MHC.

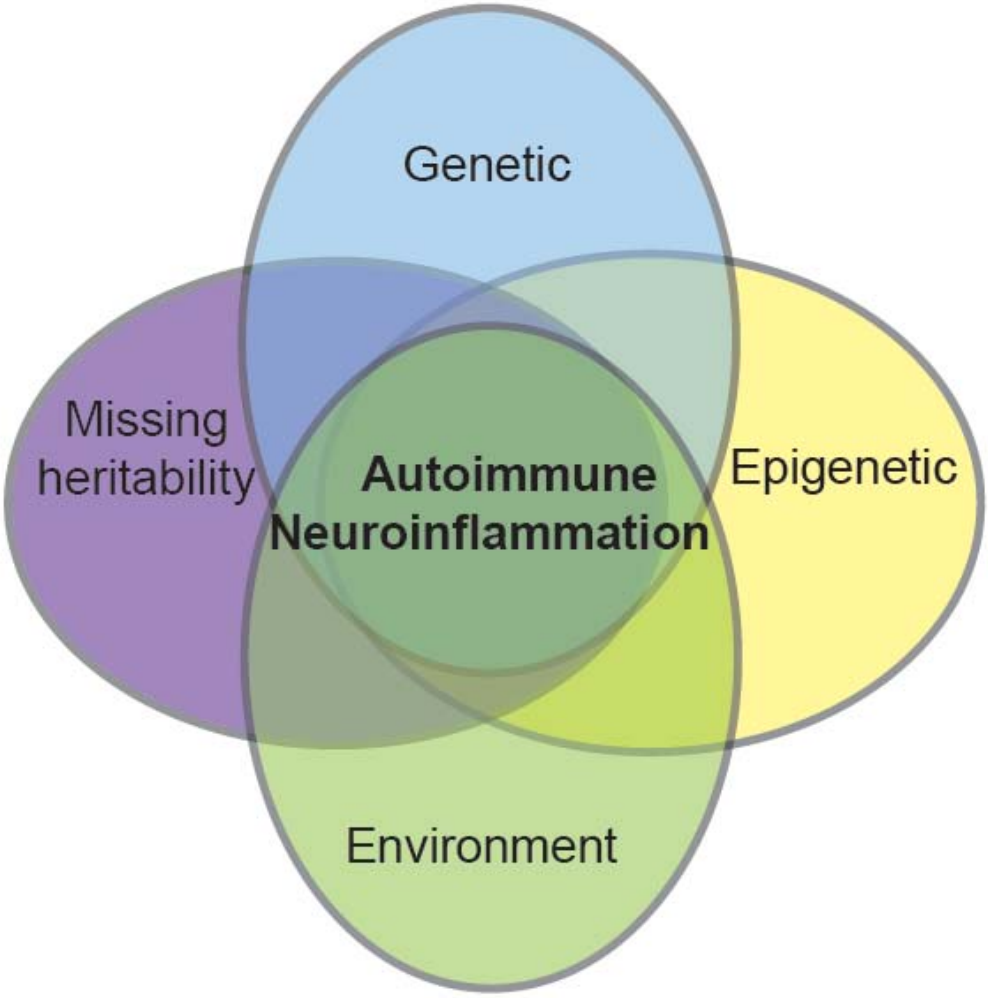

Figure 21. The proposed model of inheritance for autoimmune neuroinflammation.

8 FUTURE PERSPECTIVES

The results presented here provide new information about neuroinflammatory processes, but also generate new questions that remain to be answered.

We refined a region on chromosome four and demonstrated the existence of multiple QTLs within the region (Paper I). A number of smaller ISR congenic strains covering these QTLs, some carrying only a few genes (R23 = *Eae25*), are being functionally investigated. Smaller ISR still are under construction and candidate gene investigation in knockout mice is in progress. We also demonstrated the presence of complex interactions that amplified or abrogated the independent gene effects. One of the candidate genes located at the site for epistasis, T-cell receptor Vß, have previously shown contradictory data on the association with MS [298]. Our results can guide future studies to account for the effect of this interaction on TCRBV's influenced on autoimmune neuroinflammation. This may also help explain the conflicting associations in past studies.

The exact role played by *ZEB1* in neuroinflammation is still not defined (Paper II). Characterization of this molecule is vital to increase our knowledge about regulation of the *IL2* pathway and its role in neuroinflammation. *In vitro* and *in vivo* functional experiments on parental and congenic rats will be used to elucidate the effect of *ZEB1* splice variants on immune cells and EAE. Additionally, investigating the effects of *ZEB1* splice variants on T-cell proliferation, cellular analysis including the numbers and frequencies of pathogenic and protective populations, and cytokine profiles could elucidate the downstream effects of altered *Zfhep2* expression levels. Furthermore, histological analysis of CNS at earlier time-points could characterize the effects of *ZEB1* expression on neuroinflammation and the pathologyphysiology involved.

In addition to congenic studies, the rapidly developing strategies in gene targeting in rats, the establishment of germline-competent rat ES cells [198 199] and ZFN-mediated site-specific mutagenesis[200 201], will provide new powerful tools for functional studies of identified candidate genes.

Our data demonstrated a significant contribution of parent-of-origin effects to inheritance of autoimmune neuro-inflammation (Paper III). The epigenetic mechanisms involved need to be investigated, particularly as this may identify locations in the genome that respond to changes in the environment by gene-environment interactions. Several identified loci overlap with known clusters of imprinted genes that should be investigated in the setting of neuroinflammation. Comprehensive analysis of allele-specific expression, miRNA regulation, DNA methylation and histone modifications can be used to characterize the risk

genes and their epigenetic mechanisms, to better understand how they mediate neuro-inflammation. Incorporating these effects into future models of inheritance will not only enable more powerful and precise identification of risk factors, but will also provide a better understanding of the pathogenesis of complex inflammatory diseases.

A study of neuroinflammation in the HS is underway. The primary aim of this study will be to identify genes that regulate autoimmune neuroinflammation in this high-resolution mosaic population. The advantage is that gene identification and fine-mapping can be combined in one step, and the resolution is expected to provide QTL intervals of at best 1-3 genes (compared to a few dozen in AIL). Another advantage is that other phenotypes have been collected in the same rats, and datasets can be combined to investigate additional factors that contribute to EAE, such as naïve immune response potential and stress. Additionally, tissues and organs will be used for follow-up studies, which can add information of correlation of candidate genes with their expression as well as expression of disease markers. We established that there is enough MHC-independent variation in disease outcome to map non-MHC QTLs in the HS, in spite of several MHC being included. This can offer a great advantage compared to our previous studies in that gene-gene interactions with a variety of MHC haplotypes can potentially be unraveled. Additionally, parent-of-origin effects and environmental factors, such as season, weather, temperature and humidity, that may influence the results are recorded to enable the study of gene-environment interactions.

Technological advances in high throughput genotyping, improved statistical methods and large international efforts have enabled the identification of several genetic risk factors for common complex diseases. The rate of discovery will continue to accelerate as genome resources and mapping strategies improve. We are at the brink of deep sequencing of individual genomes on a large scale and increasingly larger GWAS studies in humans identifying a growing number of disease risk genes. An international GWAS of 10000 MS cases and controls MS, which we are part of, is underway. Rapid identification of MS genes will guide, to some extent, the focus of genetic and functional animal studies to understand how genes operate in the disease setting. As more genes are identified and provide targets for therapy, the potential for drug development will also be a focus in animal research.

9 ACKNOWLEDGEMENTS

I have had the awesome privilege of being surrounded by a group of extraordinary people during my PhD studies. I would like to express my gratitude to all colleagues, friends and family who have helped me during this time. I would particularly like to thank:

Dr. Maja Jagodic, my supervisor, a brilliant researcher and beautiful, generous person. Taking this scientific journey has been extraordinary, because I shared it with you! No words can ever capture all the thanks I owe you, so here are just a few. Thank you for sharing your love for genetics with me, for believing in me, motivating me in just the right way, and challenging me when I needed it. Thanks for giving me a lot of freedom and responsibility, always combined with lots of encouragement and help. Thanks for all the late nights around your kitchen table (and other places) with mind-blowing analyses that all made sense in the end. Most of all, thanks for your enthusiasm, friendship and support!

Dr. Kristina Becanovic, my co-supervisor and amazing energizer bunny. First, thank you for entrusting me with your projects, I know it was like handing over your first-born! Thanks for all of your advice, encouragement, support and inspiration. And thanks for all of our discussions from experimental detail to the scientific world at large (and the other one too!). Thanks for always making time for last-minute help and advice, for your cool perspective and your friendship!

Professor Tomas Olsson, my co-supervisor, thank you for your enthusiasm, for always finding time for big issues and small, for giving me a lot of independence, and for giving me the HS project! Thanks for providing a unique research environment with open dialogue, a friendly spirit and creative solutions. And thank you for investing so much generosity and opportunity in the next generation of researchers!

Dr. Erik Wallström, my co-supervisor, thank you for always being ready to answer questions and sharing your knowledge and experiences. Also, thanks for being so dependable and giving level-headed advice!

Professor Lars Klareskog, my mentor, thank you for always making time and giving very useful advice regarding life in science. Thanks for giving such a human perspective on a sometimes chaotic and hectic existence!

Dr. Bob Harris, my ‘boll-plank’ and fix-it-man, thanks for all the brain storming sessions, for contemplating (and sometimes sharing) my crazy ideas, and for

encouraging creative thought in the sometimes rigid world of science. Thanks for always taking the calls, all your smiles, words of encouragement, can-do attitude, and for your friendship!

Dr. Margarita Diez, my most extraordinary person! Your knowledge is vast, your determination remarkable, and your persistence and perseverance awe-inspiring. Thank you for sharing the HS experience with me and exploring Barcelona together! Thank you for picking me up when I am down, for slowing me down when I am frantic, and most of all for always listening and caring! These five years would not have been the same without you!

The EAE group, it's been my great pleasure to share my PhD life with you! *Monica Marta*, thanks for exploring the world of epistasis with me, for all the Punnett's squares, days of discussions, and good wine and food while doing it. Thanks for your friendship, your smiles and teaching me the tricks of the trade! And so much more… *Johan Öckinger*, thanks for all the amazingly fun times in the lab, not in the lab, and on trips around the world. For all your support, help, discussions – I can always count on you! Particularly thanks for critically reading my thesis! *Ame Beyeen*, thanks for all the fun times and shared experiences – and the Spannish! For your never-ceasing help, kind smiles, can-do attitude, words of encouragement, and for always staying calm in times of crisis. *Melanie Thessen Hedreul*, thanks for making long hours and late nights by the bench and in the animal house fun! For your amazing energy, catchy laughs, and French head-snaps! And for always challenging me to think twice! *Alan Gillett*, thanks for being my polar-opposite! For always questioning my assumptions and scrutinizing my thinking (and writing), always with a smile on your face! Thanks for all your feedback - I am much better prepared because of you! *Petra Bergman*, thanks for your vivaciousness and energy, and for always lending a hand!

All present and former colleagues at the Neuroimmunology Unit, thank you for making work fun! *Karin Harnesk*, thanks for making me laugh (even when you didn't mean to) and brightening so many days! For all the science talks, cakes, laughs, inspirational cartoons, make-up and hours of sti(bi)thching with Maggan and Maria. And a special thanks for critically reading my thesis! *Maria Swanberg*, thanks for your friendship and S&B! *Rux Covacu*, thanks for all the discussions, your kindness, so many answers to numerous questions, and for indulging my alcohol habits in the cell room! *Milena Zeitelhofer Adzemovic*, thanks for being so patient and answering my hundreds of questions regarding histopathology and microscopy – over and over again! *Ingrid Kockum*, thanks for fun discussions on statistics and for help with the HLA! *Nada Abdelmagid*, thanks for your kindness and coolness! *Mohsen Khademi*, for always helping and answering questions! *Fredrik Piehl, Lou Brundin, Clas Johansson, Mikael*

Ström, Cynthia Perez Estrada, Rickard Lindblom, Faiez Al Nimer, Cecilia Domingnes, Roham Parsa, Sohel Mia, Magda Linden, Emelie Sundqvist , Ellen Iacobeeus and *Biborka Bereczky Veress*, thanks for all the ideas, knowledge and discussions you are always willing to share! Previous lab members *Rita Nohra*, *Babbo Holm*, *Åsa Andersson, Olle Lidman, Anna Lobell, Sasja Danilov, Maja Wållberg* and *Judit Wefer*, thanks for showing the way! The new lab members, *Sevi Flytzani, Pernilla Andresen, Andre Ortlieb* and *Manuel Zeitelhofer* – what great additions to the team you are!

Britt Dahlgren, *Maine Blomstedt*, *Ann-Marie Olsson*, *Mimmi Wernman*, *Anki Larsson* and *Gullan Ryden*, for managing the administrative tasks throughout the years and always being so kind and helpful!

Thanks to all summer and project students for help.

All present and former colleagues at the Rheumatology Unit, thank you for always being so kind and helpful! A special thanks to *Therese Östberg*, for making my life beautiful (literally!) by prettyfying, getting me out of the lab to do fun things, exploring China with me, being my shopping therapist, making me laugh (at myself) and being a wonderful friend! Also for always helping at work, being there to solve problems, and give advice on immunology (I am past the buffer stage!). *Alex Espinosa*, thanks for changing my life! *Nånnis Brauner*, your positive attitude brightens the lab. My roommates, *Maria Seddighzadeh*, *Patrick LeClerc* and *Petra Amoudruz* for advice, lending an ear, good discussions and great support. *Helena Erlandsson Harris*, for all the fun! Marie Wahren Herlenius, for sharing your knowledge and experience. *Vivianne Malmström*, for FoxP3 discussions. *Lasse Ottoson* and *Daddy Mac Rönninger*, for help with more things than I can remember. *Emeli Lundström* for sharing the thesis process.

Collaborators within EURATools, thanks for sharing knowledge and fun discussions on rat genetics! *Jonathan Flint*, for being a great facilitator and being so generous with your time, knowledge and experience. It's a privilege to learn from you! *Martina Johannesson*, thank you for being a wonderful friend and most helpful colleague! For trauma-bonding over the HS and learning HAPPY – and making it all fun! For exploring Barcelona and Oxford together, so much help with Marie Curie, the list is endless… *Alberto Fernandez Teruel*, thank you for opening your lab and your heart to the European invasion – and taking such good care of us! *Esther Martinez*, thanks for your eagerness and enthusiasm for learning EAE, and for working so hard. *Esther, Gloria Blazquez, Regina Lopez, Toni Cañete*, and *Elia Vincens*, thank you for being such great friends and hosts, always taking care of us and helping with practical matters, taking us on trips, showing us the Barcelona-fun, good food and shopping, for

making long hours in the lab endurable and for making the entire HS experience so wonderful. Thank you, thank you, thank you! *Jonatan Tuncel, Ulrika Norin* and *Polinka Hernandez Pliego,* for sharing it all! Erik Werner, thank you for all of your help and for your patience!

Co-authors and collaborators: *Rikard Holmdahl, William Valdar, Hans Lassmann, Johnny Lorentzen, Richard Mott, Delyth Graham, Richard Copley, Cristina Fernández Santamaría, Percio Gulko, Max Brenner, Adolf Tobeña, Marc Guitart Masip, Lydia Giménez Llort, Anna Dominiczak, Dominique Gauguier* and *Eva E. Redei*, thank you for input and good discussions!

The AKM staff for managing the animal breeding and care.

I would like to thank my former supervisors and colleagues at the Idaho Neurological Institute. *Chris Zimmerman*, thank you for all of the opportunities! Thanks for teaching me the importance of medical research and patient care and compassion. You give future scientists a chance to blossom! *Pennie Seibert*, you changed the course of my life! Thank you for sharing your love for science with me and teaching me the intricacies of exploring the unknown. The balance between experimental design, statistical control and practical feasibility that I learned from you still guides my work. Thanks for all the experiences, encouragement, support and challenges. Most of all, thanks to you and *Greg* for our amazing friendship!

I would like to thank my big, beautiful family, for all of your love and support, both the inherited and the acquired, old members and new, past and present. You may not have seen or heard much of me lately, but you mean a lot to me!

Cilla, thanks for growing up and experiencing the world with me! *Emma*, thanks for bringing me home to Sweden! *Shele*, thanks for making my American life all that it was! Thank you all for a life shared that I cannot imagine without you! Thanks for being true friends, for always being there for me, supporting me and accepting me just as I am. Thanks for all late nights and early mornings, all the fun, the laughs and tears. If I didn't know any better, I would think you we're sporting wings on your backs! *Martin* and *Graham*, thanks for pampering and spoiling me, just when I need it! You have recharged my batteries more times than I can count. Thank you for your kind hearts, generous spirits, and for your beautiful perspectives on life. And *Graham*, thanks for my beautiful cover!

Thanks to my kids for making my life special: *Aaron & Jessie, Alma, Adrian, Emil & Agnes*.

I would especially like to thank *Mamma* and *Pappa*, for being amazing, loving, supportive parents! For teaching me to aim high without fear of failure, for respecting my choices and allowing me my own winding path, for giving 100% support and always backing me up when things don't work out. This thesis would never have happened without your help along the way! Thanks *Patrik*, the best big brother in the world, for being my greatest supporter and critic all wrapped up in one fabulous role model. Thank you for always giving me a soft place to fall and always, always, always being there for me! *Annika*, thank you for all the fun and relaxing times you bring, your effectiveness makes it all possible! *Axel, Nore, Viggo* and *Saga*, thanks for all the moments of pure joy I get to share with you - you give me great perspective on life! *Bibbi* and *Tore*, Gitta and KG, thanks for your support!

Jan, Gerd, Maria, Ville, Alfred, Agnes and *Ella*, thank you for welcoming me with open arms and for all the fun I get to share with you!

Daniel, thank you for your amazing love and support – you are my everything! Thanks for making our life together extraordinary! In regards to this book, thank you for your help with everything from carrying cages, watering sick rats, bringing lunches and dinners, to scripting code for statistical analyses, formatting figures for submission and editing papers and the thesis! Last, but absolutely not least, thanks to my constant companion during the writing of this thesis, Lillan, for sharing many late nights and tubs of Ben & Jerry's with me. We are eagerly awaiting your arrival!

10 REFERENCES

1. Monaghan F, Corcos A. On the origins of the Mendelian laws. *J Hered* 1984;75(1):67-9.
2. Moore DP, Orr-Weaver TL. Chromosome segregation during meiosis: building an unambivalent bivalent. *Curr Top Dev Biol* 1998;37:263-99.
3. Satzinger H. Theodor and Marcella Boveri: chromosomes and cytoplasm in heredity and development. *Nat Rev Genet* 2008;9(3):231-8.
4. Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant...or not? *Hum Mol Genet* 2002;11(20):2417-23.
5. Cooper DN, Ball EV, Krawczak M. The human gene mutation database. *Nucleic Acids Res* 1998;26(1):285-7.
6. Sawcer S, Ban M, Wason J, Dudbridge F. What role for genetics in the prediction of multiple sclerosis? *Ann Neurol* 2009;67(1):3-10.
7. Ramos RG, Olden K. Gene-environment interactions in the development of complex disease phenotypes. *Int J Environ Res Public Health* 2008;5(1):4-11.
8. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661-78.
9. Nagel RL. Epistasis and the genetics of human diseases. *C R Biol* 2005;328(7):606-15.
10. Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature* 1995;377(6545):150-1.
11. Mumford CJ, Wood NW, Kellar-Wood H, Thorpe JW, Miller DH, Compston DA. The British Isles survey of multiple sclerosis in twins. *Neurology* 1994;44(1):11-5.
12. Sadovnick AD, Ebers GC, Dyment DA, Risch NJ. Evidence for genetic basis of multiple sclerosis. The Canadian Collaborative Study Group. *Lancet* 1996;347(9017):1728-30.
13. Baxter AG. Immunogenetics and the cause of autoimmune disease. *Autoimmunity* 1997;25(3):177-89.
14. Wray NR, Goddard ME. Multi-locus models of genetic risk of disease. *Genome Med*;2(2):10.
15. Haegert DG. Analysis of the threshold liability model provides new understanding of causation in autoimmune diseases. *Med Hypotheses* 2004;63(2):257-61.
16. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 2009;106(23):9362-7.
17. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest* 2008;118(5):1590-605.
18. Lander ES. The new genomics: global views of biology. *Science* 1996;274(5287):536-9.
19. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001;17(9):502-10.
20. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40(8):955-62.
21. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. *Nat Genet* 2008;40(2):204-10.
22. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 2001;69(1):124-37.
23. Cazier JB, Tomlinson I. General lessons from large-scale studies to identify human cancer predisposition genes. *J Pathol*;220(2):255-62.
24. Nakhla M, Polychronakos C. Monogenic and other unusual causes of diabetes mellitus. *Pediatr Clin North Am* 2005;52(6):1637-50.

25. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994;266(5182):66-71.
26. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995;378(6559):789-92.
27. Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, et al. Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *JAMA* 2006;295(12):1379-88.
28. Lappalainen M, Halme L, Turunen U, Saavalainen P, Einarsdottir E, Farkkila M, et al. Association of IL23R, TNFRSF1A, and HLA-DRB1*0103 allele variants with inflammatory bowel disease phenotypes in the Finnish population. *Inflamm Bowel Dis* 2008;14(8):1118-24.
29. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol*;8(1):e1000294.
30. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev* 2009;19(3):212-9.
31. Schork NJ, Wessel J, Malo N. DNA sequence-based phenotypic association analysis. *Adv Genet* 2008;60:195-217.
32. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461(7265):747-53.
33. Maher B. Personal genomes: The case of the missing heritability. *Nature* 2008;456(7218):18-21.
34. Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. *Nat Genet* 2008;40(1):17-22.
35. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008;9(5):356-69.
36. Zeggini E, Rayner W, Morris AP, Hattersley AT, Walker M, Hitman GA, et al. An evaluation of HapMap sample size and tagging SNP performance in large-scale empirical and simulated data sets. *Nat Genet* 2005;37(12):1320-2.
37. Holgate ST, Davies DE, Powell RM, Howarth PH, Haitchi HM, Holloway JW. Local genetic and environmental factors in asthma disease pathogenesis: chronicity and persistence mechanisms. *Eur Respir J* 2007;29(4):793-803.
38. Peters A, Schneider A, Greven S, Bellander T, Forastiere F, Ibald-Mulli A, et al. Air pollution and inflammatory response in myocardial infarction survivors: gene-environment interactions in a high-risk group. *Inhal Toxicol* 2007;19 Suppl 1:161-75.
39. Eder W, Klimecki W, Yu L, von Mutius E, Riedler J, Braun-Fahrlander C, et al. Opposite effects of CD 14/-260 on serum IgE levels in children raised in different environments. *J Allergy Clin Immunol* 2005;116(3):601-7.
40. Neel JV. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *Am J Hum Genet* 1962;14:353-62.
41. Strachan DP. Hay fever, hygiene, and household size. *BMJ* 1989;299(6710):1259-60.
42. Phipatanakul W, Celedon JC, Raby BA, Litonjua AA, Milton DK, Sredl D, et al. Endotoxin exposure and eczema in the first year of life. *Pediatrics* 2004;114(1):13-8.
43. Ball TM, Castro-Rodriguez JA, Griffith KA, Holberg CJ, Martinez FD, Wright AL. Siblings, day-care attendance, and the risk of asthma and wheezing during childhood. *N Engl J Med* 2000;343(8):538-43.
44. Xue F, Michels KB. Intrauterine factors and risk of breast cancer: a systematic review and meta-analysis of current evidence. *Lancet Oncol* 2007;8(12):1088-100.
45. Trichopoulos D. Hypothesis: does breast cancer originate in utero? *Lancet* 1990;335(8695):939-40.
46. Schernhammer ES. In-utero exposures and breast cancer risk: joint effect of estrogens and insulin-like growth factor? *Cancer Causes Control* 2002;13(6):505-8.

47. Xue F, Colditz GA, Willett WC, Rosner BA, Michels KB. Parental age at delivery and incidence of breast cancer: a prospective cohort study. *Breast Cancer Res Treat* 2007;104(3):331-40.
48. Noakes PS, Holt PG, Prescott SL. Maternal smoking in pregnancy alters neonatal cytokine responses. *Allergy* 2003;58(10):1053-8.
49. Sawcer S. Bayes factors in complex genetics. *Eur J Hum Genet*.
50. Albrechtsen A, Sand Korneliussen T, Moltke I, van Overseem Hansen T, Nielsen FC, Nielsen R. Relatedness mapping and tracts of relatedness for genome-wide data in the presence of linkage disequilibrium. *Genet Epidemiol* 2009;33(3):266-74.
51. Besnier F, Carlborg O. A general and efficient method for estimating continuous IBD functions for use in genome scans for QTL. *BMC Bioinformatics* 2007;8:440.
52. Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet* 2002;11(20):2463-8.
53. Reik W, Walter J. Genomic imprinting: parental influence on the genome. *Nat Rev Genet* 2001;2(1):21-32.
54. Glaser RL, Jabs EW. Dear old dad. *Sci Aging Knowledge Environ* 2004;2004(3):re1.
55. Crow JF. The origins, patterns and implications of human spontaneous mutation. *Nat Rev Genet* 2000;1(1):40-7.
56. Sapienza C, Peterson AC, Rossant J, Balling R. Degree of methylation of transgenes is dependent on gamete of origin. *Nature* 1987;328(6127):251-4.
57. Chaillet JR, Vogt TF, Beier DR, Leder P. Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. *Cell* 1991;66(1):77-83.
58. Delaval K, Feil R. Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 2004;14(2):188-95.
59. Feil R, Berger F. Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet* 2007;23(4):192-9.
60. Cohen MM, Jr. Beckwith-Wiedemann syndrome: historical, clinicopathological, and etiopathogenetic perspectives. *Pediatr Dev Pathol* 2005;8(3):287-304.
61. Hark AT, Tilghman SM. Chromatin conformation of the H19 epigenetic mark. *Hum Mol Genet* 1998;7(12):1979-85.
62. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000;405(6785):486-9.
63. Delaval K, Wagschal A, Feil R. Epigenetic deregulation of imprinting in congenital diseases of aberrant growth. *Bioessays* 2006;28(5):453-9.
64. Weksberg R, Shen DR, Fei YL, Song QL, Squire J. Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. *Nat Genet* 1993;5(2):143-50.
65. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 2003;23(15):5293-300.
66. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 1999;23(3):314-8.
67. Wolff GL, Kodell RL, Moore SR, Cooney CA. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. *FASEB J* 1998;12(11):949-57.
68. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, et al. Epigenetic programming by maternal behavior. *Nat Neurosci* 2004;7(8):847-54.
69. Adams JM. Ways of dying: multiple pathways to apoptosis. *Genes Dev* 2003;17(20):2481-95.
70. Petronis A. Epigenetics and twins: three variations on the theme. *Trends Genet* 2006;22(7):347-50.
71. Handel AE, Ebers GC, Ramagopalan SV. Epigenetics: molecular mechanisms and implications for disease. *Trends Mol Med* 2009.
72. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006;7(1):21-33.

73. Rhee I, Bachman KE, Park BH, Jair KW, Yen RW, Schuebel KE, et al. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* 2002;416(6880):552-6.
74. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002;16(1):6-21.
75. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, et al. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* 1998;19(2):187-91.
76. Mutskov VJ, Farrell CM, Wade PA, Wolffe AP, Felsenfeld G. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev* 2002;16(12):1540-54.
77. Campos EI, Reinberg D. Histones: annotating chromatin. *Annu Rev Genet* 2009;43:559-99.
78. Vaissiere T, Sawan C, Herceg Z. Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutat Res* 2008;659(1-2):40-8.
79. Matzke MA, Primig M, Trnovsky J, Matzke AJ. Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. *EMBO J* 1989;8(3):643-49.
80. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. Non-coding RNAs: regulators of disease. *J Pathol*;220(2):126-39.
81. Brodersen P, Voinnet O. Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol* 2009;10(2):141-8.
82. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005;102(30):10604-9.
83. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008;40(5):638-45.
84. *Multiple Sclerosis: Diagnosis, Medical Management, and Rehabilitation*. 1 ed. New York: Demos Medical Publishing, 2000.
85. Koch-Henriksen N. Multiple sclerosis in Scandinavia and Finland. *Acta Neurol Scand Suppl* 1995;161:55-9.
86. Sundstrom P, Nystrom L, Forsgren L. Prevalence of multiple sclerosis in Vasterbotten County in northern Sweden. *Acta Neurol Scand* 2001;103(4):214-8.
87. Svenningsson A, Runmarker B, Lycke J, Andersen O. Incidence of MS during two fifteen-year periods in the Gothenburg region of Sweden. *Acta Neurol Scand* 1990;82(3):161-8.
88. Rodriguez M, Siva A, Ward J, Stolp-Smith K, O'Brien P, Kurland L. Impairment, disability, and handicap in multiple sclerosis: a population-based study in Olmsted County, Minnesota. *Neurology* 1994;44(1):28-33.
89. Berg J, Lindgren P, Fredrikson S, Kobelt G. Costs and quality of life of multiple sclerosis in Sweden. *Eur J Health Econ* 2006;7 Suppl 2:S75-85.
90. Lublin FD, Reingold SC. 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* 1996;46(4):907-11.
91. McDonald WI. Relapse, remission, and progression in multiple sclerosis. *N Engl J Med* 2000;343(20):1486-7.
92. Trapp BD, Peterson J, Ransohoff RM, Rudick R, Mork S, Bo L. Axonal transection in the lesions of multiple sclerosis. *N Engl J Med* 1998;338(5):278-85.
93. Prineas JW, Barnard RO, Revesz T, Kwon EE, Sharer L, Cho ES. Multiple sclerosis. Pathology of recurrent lesions. *Brain* 1993;116 (Pt 3):681-93.
94. Miller DH, Leary SM. Primary-progressive multiple sclerosis. *Lancet Neurol* 2007;6(10):903-12.
95. Lucchinetti CF, Parisi J, Bruck W. The pathology of multiple sclerosis. *Neurol Clin* 2005;23(1):77-105, vi.

96. Baron JL, Madri JA, Ruddle NH, Hashim G, Janeway CA, Jr. Surface expression of alpha 4 integrin by CD4 T cells is required for their entry into brain parenchyma. *J Exp Med* 1993;177(1):57-68.
97. Fehervari Z, Sakaguchi S. CD4+ Tregs and immune control. *J Clin Invest* 2004;114(9):1209-17.
98. Piccirillo CA, d'Hennezel E, Sgouroudis E, Yurchenko E. CD4+Foxp3+ regulatory T cells in the control of autoimmunity: in vivo veritas. *Curr Opin Immunol* 2008;20(6):655-62.
99. Hillert J, Olerup O. Multiple sclerosis is associated with genes within or close to the HLA-DR-DQ subregion on a normal DR15,DQ6,Dw2 haplotype. *Neurology* 1993;43(1):163-8.
100. Wallstrom E, Khademi M, Andersson M, Weissert R, Linington C, Olsson T. Increased reactivity to myelin oligodendrocyte glycoprotein peptides and epitope mapping in HLA DR2(15)+ multiple sclerosis. *Eur J Immunol* 1998;28(10):3329-35.
101. Tomassini V, De Giglio L, Reindl M, Russo P, Pestalozza I, Pantano P, et al. Anti-myelin antibodies predict the clinical outcome after a first episode suggestive of MS. *Mult Scler* 2007;13(9):1086-94.
102. Davidson A, Diamond B. Autoimmune diseases. *N Engl J Med* 2001;345(5):340-50.
103. Lalive PH. Autoantibodies in inflammatory demyelinating diseases of the central nervous system. *Swiss Med Wkly* 2008;138(47-48):692-707.
104. Costantino CM, Baecher-Allan C, Hafler DA. Multiple sclerosis and regulatory T cells. *J Clin Immunol* 2008;28(6):697-706.
105. Sadovnick AD, Armstrong H, Rice GP, Bulman D, Hashimoto L, Paty DW, et al. A population-based study of multiple sclerosis in twins: update. *Ann Neurol* 1993;33(3):281-5.
106. Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: updated risks for relatives. *Am J Med Genet* 1988;29(3):533-41.
107. Olerup O, Hillert J. HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. *Tissue Antigens* 1991;38(1):1-15.
108. A meta-analysis of whole genome linkage screens in multiple sclerosis. *J Neuroimmunol* 2003;143(1-2):39-46.
109. Gregory SG, Schmidt S, Seth P, Oksenberg JR, Hart J, Prokop A, et al. Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nat Genet* 2007;39(9):1083-91.
110. Lundmark F, Duvefelt K, Iacobaeus E, Kockum I, Wallstrom E, Khademi M, et al. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet* 2007;39(9):1108-13.
111. Weber F, Fontaine B, Cournu-Rebeix I, Kroner A, Knop M, Lutz S, et al. IL2RA and IL7RA genes confer susceptibility for multiple sclerosis in two independent European populations. *Genes Immun* 2008;9(3):259-63.
112. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007;357(9):851-62.
113. Rubio JP, Stankovich J, Field J, Tubridy N, Marriott M, Chapman C, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. *Genes Immun* 2008;9(7):624-30.
114. Hoppenbrouwers IA, Aulchenko YS, Janssens AC, Ramagopalan SV, Broer L, Kayser M, et al. Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis. *J Hum Genet* 2009.
115. De Jager PL, Baecher-Allan C, Maier LM, Arthur AT, Ottoboni L, Barcellos L, et al. The role of the CD58 locus in multiple sclerosis. *Proc Natl Acad Sci U S A* 2009;106(13):5264-9.
116. Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. *Hum Mol Genet*;19(5):953-62.
117. The expanding genetic overlap between multiple sclerosis and type I diabetes. *Genes Immun* 2009;10(1):11-4.

118. Ban M, Goris A, Lorentzen AR, Baker A, Mihalova T, Ingram G, et al. Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. *Eur J Hum Genet* 2009;17(10):1309-13.
119. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* 2009;41(7):824-8.
120. De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* 2009;41(7):776-82.
121. Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2009;18(4):767-78.
122. Barton A, Woolmore JA, Ward D, Eyre S, Hinks A, Ollier WE, et al. Association of protein kinase C alpha (PRKCA) gene with multiple sclerosis in a UK population. *Brain* 2004;127(Pt 8):1717-22.
123. Saarela J, Kallio SP, Chen D, Montpetit A, Jokiaho A, Choi E, et al. PRKCA and multiple sclerosis: association in two independent populations. *PLoS Genet* 2006;2(3):e42.
124. Aulchenko YS, Hoppenbrouwers IA, Ramagopalan SV, Broer L, Jafari N, Hillert J, et al. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. *Nat Genet* 2008;40(12):1402-3.
125. Hafler JP, Maier LM, Cooper JD, Plagnol V, Hinks A, Simmonds MJ, et al. CD226 Gly307Ser association with multiple autoimmune diseases. *Genes Immun* 2009;10(1):5-10.
126. Letourneau S, Krieg C, Pantaleo G, Boyman O. IL-2- and CD25-dependent immunoregulatory mechanisms in the homeostasis of T-cell subsets. *J Allergy Clin Immunol* 2009;123(4):758-62.
127. Maier LM, Anderson DE, Severson CA, Baecher-Allan C, Healy B, Liu DV, et al. Soluble IL-2RA levels in multiple sclerosis subjects and the effect of soluble IL-2RA on immune responses. *J Immunol* 2009;182(3):1541-7.
128. Greenberg SJ, Marcon L, Hurwitz BJ, Waldmann TA, Nelson DL. Elevated levels of soluble interleukin-2 receptors in multiple sclerosis. *N Engl J Med* 1988;319(15):1019-20.
129. Kurtzke JF. Geography in multiple sclerosis. *J Neurol* 1977;215(1):1-26.
130. Alter M, Kahana E, Loewenson R. Migration and risk of multiple sclerosis. *Neurology* 1978;28(11):1089-93.
131. Dean G, Kurtzke JF. On the risk of multiple sclerosis according to age at immigration to South Africa. *Br Med J* 1971;3(5777):725-9.
132. Hammond SR, English DR, McLeod JG. The age-range of risk of developing multiple sclerosis: evidence from a migrant population in Australia. *Brain* 2000;123 (Pt 5):968-74.
133. Willer CJ, Dyment DA, Sadovnick AD, Rothwell PM, Murray TJ, Ebers GC. Timing of birth and risk of multiple sclerosis: population based study. *BMJ* 2005;330(7483):120.
134. Acheson ED, Bachrach CA, Wright FM. Some comments on the relationship of the distribution of multiple sclerosis to latitude, solar radiation, and other variables. *Acta Psychiatr Scand Suppl* 1960;35(147):132-47.
135. Rigby WF, Waugh M, Graziano RF. Regulation of human monocyte HLA-DR and CD4 antigen expression, and antigen presentation by 1,25-dihydroxyvitamin D3. *Blood* 1990;76(1):189-97.
136. Skjodt H, Hughes DE, Dobson PR, Russell RG. Constitutive and inducible expression of HLA class II determinants by human osteoblast-like cells in vitro. *J Clin Invest* 1990;85(5):1421-6.
137. Ramagopalan SV, Maugeri NJ, Handunnetthi L, Lincoln MR, Orton SM, Dyment DA, et al. Expression of the multiple sclerosis-associated MHC class II Allele HLA-DRB1*1501 is regulated by vitamin D. *PLoS Genet* 2009;5(2):e1000369.
138. Hawkes CH. Smoking is a risk factor for multiple sclerosis: a metanalysis. *Mult Scler* 2007;13(5):610-5.
139. Sundstrom P, Nystrom L. Smoking worsens the prognosis in multiple sclerosis. *Mult Scler* 2008;14(8):1031-5.

140. Hedstrom AK, Baarnhielm M, Olsson T, Alfredsson L. Tobacco smoking, but not Swedish snuff use, increases the risk of multiple sclerosis. *Neurology* 2009;73(9):696-701.
141. Carlens C, Hergens MP, Grunewald J, Ekbom A, Eklund A, Olgart Hoglund C, et al. Smoking, Use of Moist Snuff and Risk of Chronic Inflammatory Diseases. *Am J Respir Crit Care Med*.
142. Kanai Y. Alterations of DNA methylation and clinicopathological diversity of human cancers. *Pathol Int* 2008;58(9):544-58.
143. Compston DA, Vakarelis BN, Paul E, McDonald WI, Batchelor JR, Mims CA. Viral infection in patients with multiple sclerosis and HLA-DR matched controls. *Brain* 1986;109 (Pt 2):325-44.
144. Sundstrom P, Juto P, Wadell G, Hallmans G, Svenningsson A, Nystrom L, et al. An altered immune response to Epstein-Barr virus in multiple sclerosis: a prospective study. *Neurology* 2004;62(12):2277-82.
145. Ebers GC. Environmental factors and multiple sclerosis. *Lancet Neurol* 2008;7(3):268-77.
146. Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci U S A* 2003;100(22):12877-82.
147. Dean G, Elian M. Age at immigration to England of Asian and Caribbean immigrants and the risk of developing multiple sclerosis. *J Neurol Neurosurg Psychiatry* 1997;63(5):565-8.
148. Ebers GC, Sadovnick AD, Dyment DA, Yee IM, Willer CJ, Risch N. Parent-of-origin effect in multiple sclerosis: observations in half-siblings. *Lancet* 2004;363(9423):1773-4.
149. Herrera BM, Ramagopalan SV, Lincoln MR, Orton SM, Chao MJ, Sadovnick AD, et al. Parent-of-origin effects in MS: observations from avuncular pairs. *Neurology* 2008;71(11):799-803.
150. Ramagopalan SV, Herrera BM, Bell JT, Dyment DA, Deluca GC, Lincoln MR, et al. Parental transmission of HLA-DRB1*15 in multiple sclerosis. *Hum Genet* 2008;122(6):661-3.
151. Chao MJ, Ramagopalan SV, Herrera BM, Lincoln MR, Dyment DA, Sadovnick AD, et al. Epigenetics in multiple sclerosis susceptibility: difference in transgenerational risk localizes to the major histocompatibility complex. *Hum Mol Genet* 2009;18(2):261-6.
152. Kantarci OH, Barcellos LF, Atkinson EJ, Ramsay PP, Lincoln R, Achenbach SJ, et al. Men transmit MS more often to their children vs women: the Carter effect. *Neurology* 2006;67(2):305-10.
153. Herrera BM, Ramagopalan SV, Orton S, Chao MJ, Yee IM, Sadovnick AD, et al. Parental transmission of MS in a population-based Canadian cohort. *Neurology* 2007;69(12):1208-12.
154. Morison IM, Ramsay JP, Spencer HG. A census of mammalian imprinting. *Trends Genet* 2005;21(8):457-65.
155. Mastronardi FG, Noor A, Wood DD, Paton T, Moscarello MA. Peptidyl argininedeiminase 2 CpG island in multiple sclerosis white matter is hypomethylated. *J Neurosci Res* 2007;85(9):2006-16.
156. Lincoln MR, Ramagopalan SV, Chao MJ, Herrera BM, Deluca GC, Orton SM, et al. Epistasis among HLA-DRB1, HLA-DQA1, and HLA-DQB1 loci determines multiple sclerosis susceptibility. *Proc Natl Acad Sci U S A* 2009;106(18):7542-7.
157. Schmidt H, Williamson D, Ashley-Koch A. HLA-DR15 haplotype and multiple sclerosis: a HuGE review. *Am J Epidemiol* 2007;165(10):1097-109.
158. Serrano-Fernandez P, Ibrahim SM, Zettl UK, Thiesen HJ, Godde R, Epplen JT, et al. Intergenomic consensus in multifactorial inheritance loci: the case of multiple sclerosis. *Genes Immun* 2004;5(8):615-20.
159. Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, et al. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nat Genet* 2005;37(5):486-94.

160. Wallström E OT. *Rat Models of Experimental Autoimmune Encephalomyelitis*. Totowa, NJ: Humana Press Inc, 2007.
161. Linker RA, Lee DH. Models of autoimmune demyelination in the central nervous system: on the way to translational medicine. *Exp Transl Stroke Med* 2009;1:5.
162. Becanovic K, Jagodic M, Sheng JR, Dahlman I, Aboul-Enein F, Wallstrom E, et al. Advanced intercross line mapping of Eae5 reveals Ncf-1 and CLDN4 as candidate genes for experimental autoimmune encephalomyelitis. *J Immunol* 2006;176(10):6055-64.
163. Gregersen JW, Holmes S, Fugger L. Humanized animal models for autoimmune diseases. *Tissue Antigens* 2004;63(5):383-94.
164. Bettelli E, Baeten D, Jager A, Sobel RA, Kuchroo VK. Myelin oligodendrocyte glycoprotein-specific T and B cells cooperate to induce a Devic-like disease in mice. *J Clin Invest* 2006;116(9):2393-402.
165. Krishnamoorthy G, Lassmann H, Wekerle H, Holz A. Spontaneous opticospinal encephalomyelitis in a double-transgenic mouse model of autoimmune T cell/B cell cooperation. *J Clin Invest* 2006;116(9):2385-92.
166. Pollinger B, Krishnamoorthy G, Berer K, Lassmann H, Bosl MR, Dunn R, et al. Spontaneous relapsing-remitting EAE in the SJL/J mouse: MOG-reactive transgenic T cells recruit endogenous MOG-specific B cells. *J Exp Med* 2009;206(6):1303-16.
167. Eng LF, Ghirnikar RS, Lee YL. Inflammation in EAE: role of chemokine/cytokine expression by resident and infiltrating cells. *Neurochem Res* 1996;21(4):511-25.
168. Naparstek Y, Ben-Nun A, Holoshitz J, Reshef T, Frenkel A, Rosenberg M, et al. T lymphocyte lines producing or vaccinating against autoimmune encephalomyelitis (EAE). Functional activation induces peanut agglutinin receptors and accumulation in the brain and thymus of line cells. *Eur J Immunol* 1983;13(5):418-23.
169. Cannella B, Cross AH, Raine CS. Upregulation and coexpression of adhesion molecules correlate with relapsing autoimmune demyelination in the central nervous system. *J Exp Med* 1990;172(5):1521-4.
170. Myers KJ, Dougherty JP, Ron Y. In vivo antigen presentation by both brain parenchymal cells and hematopoietically derived cells during the induction of experimental autoimmune encephalomyelitis. *J Immunol* 1993;151(4):2252-60.
171. Traugott U, McFarlin DE, Raine CS. Immunopathology of the lesion in chronic relapsing experimental autoimmune encephalomyelitis in the mouse. *Cell Immunol* 1986;99(2):395-410.
172. Chen Y, Langrish CL, McKenzie B, Joyce-Shaikh B, Stumhofer JS, McClanahan T, et al. Anti-IL-23 therapy inhibits multiple inflammatory pathways and ameliorates autoimmune encephalomyelitis. *J Clin Invest* 2006;116(5):1317-26.
173. Jager A, Dardalhon V, Sobel RA, Bettelli E, Kuchroo VK. Th1, Th17, and Th9 effector cells induce experimental autoimmune encephalomyelitis with different pathological phenotypes. *J Immunol* 2009;183(11):7169-77.
174. Stromnes IM, Cerretti LM, Liggitt D, Harris RA, Goverman JM. Differential regulation of central nervous system autoimmunity by T(H)1 and T(H)17 cells. *Nat Med* 2008;14(3):337-42.
175. Weissert R, Wallstrom E, Storch MK, Stefferl A, Lorentzen J, Lassmann H, et al. MHC haplotype-dependent regulation of MOG-induced EAE in rats. *J Clin Invest* 1998;102(6):1265-73.
176. Fillmore PD, Brace M, Troutman SA, Blankenhorn EP, Diehl S, Rincon M, et al. Genetic analysis of the influence of neuroantigen-complete Freund's adjuvant emulsion structures on the sexual dimorphism and susceptibility to experimental allergic encephalomyelitis. *Am J Pathol* 2003;163(4):1623-32.
177. Teuscher C, Doerge RW, Fillmore PD, Blankenhorn EP. eae36, a locus on mouse chromosome 4, controls susceptibility to experimental allergic encephalomyelitis in older mice and mice immunized in the winter. *Genetics* 2006;172(2):1147-53.

178. Storch MK, Stefferl A, Brehm U, Weissert R, Wallström E, Kerschensteiner M, et al. Autoimmunity to Myelin Oligodendrocyte Glycoprotein in Rats Mimics the Spectrum of Multiple Sclerosis Pathology *Brain Pathol* 1998;8(4):681-94.
179. Lucchinetti C, Bruck W, Parisi J, Scheithauer B, Rodriguez M, Lassmann H. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Ann Neurol* 2000;47(6):707-17.
180. Breij EC, Brink BP, Veerhuis R, van den Berg C, Vloet R, Yan R, et al. Homogeneity of active demyelinating lesions in established multiple sclerosis. *Ann Neurol* 2008;63(1):16-25.
181. Olsson T. Immunology of multiple sclerosis. *Curr Opin Neurol Neurosurg* 1992;5(2):195-202.
182. Steinman L. Multiple sclerosis: a coordinated immunological attack against myelin in the central nervous system. *Cell* 1996;85(3):299-302.
183. Jagodic M, Kornek B, Weissert R, Lassmann H, Olsson T, Dahlman I. Congenic mapping confirms a locus on rat chromosome 10 conferring strong protection against myelin oligodendrocyte glycoprotein-induced experimental autoimmune encephalomyelitis. *Immunogenetics* 2001;53(5):410-5.
184. Dahlman I, Jacobsson L, Glaser A, Lorentzen JC, Andersson M, Luthman H, et al. Genome-wide linkage analysis of chronic relapsing experimental autoimmune encephalomyelitis in the rat identifies a major susceptibility locus on chromosome 9. *J Immunol* 1999;162(5):2581-8.
185. Dahlman I, Wallstrom E, Weissert R, Storch M, Kornek B, Jacobsson L, et al. Linkage analysis of myelin oligodendrocyte glycoprotein-induced experimental autoimmune encephalomyelitis in the rat identifies a locus controlling demyelination on chromosome 18. *Hum Mol Genet* 1999;8(12):2183-90.
186. Becanovic K, Wallstrom E, Kornek B, Glaser A, Broman KW, Dahlman I, et al. New loci regulating rat myelin oligodendrocyte glycoprotein-induced experimental autoimmune encephalomyelitis. *J Immunol* 2003;170(2):1062-9.
187. Aitman TJ, Critser JK, Cuppen E, Dominiczak A, Fernandez-Suarez XM, Flint J, et al. Progress and prospects in rat genetics: a community view. *Nat Genet* 2008;40(5):516-22.
188. Jacob HJ. Functional genomics and rat models. *Genome Res* 1999;9(11):1013-6.
189. Jacob HJ, Kwitek AE. Rat genetics: attaching physiology and pharmacology to the genome. *Nat Rev Genet* 2002;3(1):33-42.
190. Thomas MA, Chen CF, Jensen-Seaman MI, Tonellato PJ, Twigger SN. Phylogenetics of rat inbred strains. *Mamm Genome* 2003;14(1):61-4.
191. Saar K, Beck A, Bihoreau MT, Birney E, Brocklebank D, Chen Y, et al. SNP and haplotype mapping for genetic analysis in the rat. *Nat Genet* 2008;40(5):560-6.
192. Kurtz TW, Montano M, Chan L, Kabra P. Molecular evidence of genetic heterogeneity in Wistar-Kyoto rats: implications for research with the spontaneously hypertensive rat. *Hypertension* 1989;13(2):188-92.
193. Kurtz TW, Morris RC, Jr. Biological variability in Wistar-Kyoto rats. Implications for research with the spontaneously hypertensive rat. *Hypertension* 1987;10(1):127-31.
194. Voigt B, Serikawa T. Pluripotent stem cells and other technologies will eventually open the door for straightforward gene targeting in the rat. *Dis Model Mech* 2009;2(7-8):341-3.
195. Mashimo T, Serikawa T. Rat resources in biomedical research. *Curr Pharm Biotechnol* 2009;10(2):214-20.
196. Kitada K, Ishishita S, Tosaka K, Takahashi R, Ueda M, Keng VW, et al. Transposon-tagged mutagenesis in the rat. *Nat Methods* 2007;4(2):131-3.
197. Lu B, Geurts AM, Poirier C, Petit DC, Harrison W, Overbeek PA, et al. Generation of rat mutants using a coat color-tagged Sleeping Beauty transposon system. *Mamm Genome* 2007;18(5):338-46.
198. Buehr M, Meek S, Blair K, Yang J, Ure J, Silva J, et al. Capture of authentic embryonic stem cells from rat blastocysts. *Cell* 2008;135(7):1287-98.
199. Li P, Tong C, Mehrian-Shai R, Jia L, Wu N, Yan Y, et al. Germline competent embryonic stem cells derived from rat blastocysts. *Cell* 2008;135(7):1299-310.

200. Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 2009;325(5939):433.
201. Geurts AM, Cost GJ, Remy S, Cui X, Tesson L, Usal C, et al. Generation of gene-specific mutated rats using zinc-finger nucleases. *Methods Mol Biol*;597:211-25.
202. Johannesson M, Lopez-Aumatell R, Stridh P, Diez M, Tuncel J, Blazquez G, et al. A resource for the simultaneous high-resolution mapping of multiple quantitative trait loci in rats: the NIH heterogeneous stock. *Genome Res* 2009;19(1):150-8.
203. Hansen C, Spuhler K. Development of the National Institutes of Health genetically heterogeneous rat stock. *Alcohol Clin Exp Res* 1984;8(5):477-9.
204. Jagodic M, Colacios, C., Nohra, R., Dejean, A.S., Beyeen, A.D., Khademi, M., Casemayou, A., Lamouroux, L., Duthoit, C., Papapietro, O., Sjöholm, L., Bernard, I., Lagrange, D., Dahlman, I., Lundmark, F., Oturai, A.B., Soendergaard, H.B., Kemppinen, A., Saarela, J., Tienari, P.J., Harbo, H.F., Spurkland, A., Ramagopalan, S.V., Sadovnick, D.A., Ebers, G.C., Seddighzadeh, M., Klareskog, L., Alfredsson, L., Padyukov, L., Hillert, J., Clanet, M., Edan, G., Fontaine, B., Fournie, G.J., Kockum, I., Saoudi, A., Olsson, T. A Role for VAV1 in Experimental Autoimmune Encephalomyelitis and Multiple Sclerosis. *Science Translational Medicine* 2009;1(10):10-21.
205. Serikawa T, Mashimo T, Takizawa A, Okajima R, Maedomari N, Kumafuji K, et al. National BioResource Project-Rat and related activities. *Exp Anim* 2009;58(4):333-41.
206. Wakeland E, Morel L, Achey K, Yui M, Longmate J. Speed congenics: a classic technique in the fast lane (relatively speaking). *Immunol Today* 1997;18(10):472-7.
207. Cicila GT, Lee SJ. Identifying candidate genes for blood pressure quantitative trait loci using differential gene expression and a panel of congenic strains. *Hypertens Res* 1998;21(4):289-96.
208. Lorentzen JC, Flornes L, Eklow C, Backdahl L, Ribbhammar U, Guo JP, et al. Association of arthritis with a gene complex encoding C-type lectin-like receptors. *Arthritis Rheum* 2007;56(8):2620-32.
209. Hospital F. Selection in backcross programmes. *Philos Trans R Soc Lond B Biol Sci* 2005;360(1459):1503-11.
210. Zou F. QTL mapping in intercross and backcross populations. *Methods Mol Biol* 2009;573:157-73.
211. Cui Y, Cheverud JM, Wu R. A statistical model for dissecting genomic imprinting through genetic mapping. *Genetica* 2007;130(3):227-39.
212. Darvasi A, Soller M. A simple method to calculate resolving power and confidence interval of QTL map location. *Behav Genet* 1997;27(2):125-32.
213. Darvasi A, Soller M. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics* 1995;141(3):1199-207.
214. Ockinger J, Stridh P, Beyeen AD, Lundmark F, Seddighzadeh M, Oturai A, et al. Genetic variants of CC chemokine genes in experimental autoimmune encephalomyelitis, multiple sclerosis and rheumatoid arthritis. *Genes Immun* 2009.
215. Valdar W, Solberg LC, Gauguier D, Burnett S, Klenerman P, Cookson WO, et al. Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat Genet* 2006;38(8):879-87.
216. Mott R, Flint J. Simultaneous detection and fine mapping of quantitative trait loci in mice using heterogeneous stocks. *Genetics* 2002;160(4):1609-18.
217. Caballero A, Toro MA. Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genet Res* 2000;75(3):331-43.
218. Solberg LC, Valdar W, Gauguier D, Nunez G, Taylor A, Burnett S, et al. A protocol for high-throughput phenotyping, suitable for quantitative trait analysis in mice. *Mamm Genome* 2006;17(2):129-46.

219. Huang GJ, Shifman S, Valdar W, Johannesson M, Yalcin B, Taylor MS, et al. High resolution mapping of expression QTLs in heterogeneous stock mice in multiple tissues. *Genome Res* 2009;19(6):1133-40.
220. Stefferl A, Brehm U, Storch M, Lambracht-Washington D, Bourquin C, Wonigeit K, et al. Myelin oligodendrocyte glycoprotein induces experimental autoimmune encephalomyelitis in the "resistant" Brown Norway rat: disease susceptibility is determined by MHC and MHC-linked effects on the B cell response. *J Immunol* 1999;163(1):40-9.
221. Levine S, Wenk EJ. A Hyperacute Form of Allergic Encephalomyelitis. *Am J Pathol* 1965;47:61-88.
222. Goldmuntz EA, Wilder RL, Goldfarb Y, Cash JM, Zha H, Crofford LJ, et al. The origin of the autoimmune disease-resistant LER rat: an outcross between the buffalo and autoimmune disease-prone Lewis inbred rat strains. *J Neuroimmunol* 1993;44(2):215-9.
223. Stevens DB, Gold DP, Sercarz EE, Moudgil KD. The Wistar Kyoto (RT1(l)) rat is resistant to myelin basic protein-induced experimental autoimmune encephalomyelitis: comparison with the susceptible Lewis (RT1(l)) strain with regard to the MBP-directed CD4+ T cell repertoire and its regulation. *J Neuroimmunol* 2002;126(1-2):25-36.
224. Sun D, Whitaker JN, Wilson DB. Regulatory T cells in experimental allergic encephalomyelitis. III. Comparison of disease resistance in Lewis and Fischer 344 rats. *Eur J Immunol* 1999;29(4):1101-6.
225. Yalcin B, Willis-Owen SA, Fullerton J, Meesaq A, Deacon RM, Rawlins JN, et al. Genetic dissection of a behavioral quantitative trait locus shows that Rgs2 modulates anxiety in mice. *Nat Genet* 2004;36(11):1197-202.
226. Flint J. The genetic basis of neuroticism. *Neurosci Biobehav Rev* 2004;28(3):307-16.
227. Mott R, Talbot CJ, Turri MG, Collins AC, Flint J. A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc Natl Acad Sci U S A* 2000;97(23):12649-54.
228. Valdar W, Solberg LC, Gauguier D, Cookson WO, Rawlins JN, Mott R, et al. Genetic and environmental effects on complex traits in mice. *Genetics* 2006;174(2):959-84.
229. Jagodic M, Marta M, Becanovic K, Sheng JR, Nohra R, Olsson T, et al. Resolution of a 16.8-Mb autoimmunity-regulating rat chromosome 4 region into multiple encephalomyelitis quantitative trait loci and evidence for epistasis. *J Immunol* 2005;174(2):918-24.
230. Jagodic M, Becanovic K, Sheng JR, Wu X, Backdahl L, Lorentzen JC, et al. An advanced intercross line resolves Eae18 into two narrow quantitative trait loci syntenic to multiple sclerosis candidate loci. *J Immunol* 2004;173(2):1366-73.
231. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. *Nature* 2004;429(6990):446-52.
232. Broman KW. Review of statistical methods for QTL mapping in experimental crosses. *Lab Anim (NY)* 2001;30(7):44-52.
233. Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 1989;121(1):185-99.
234. Haley CS, Knott SA. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 1992;69(4):315-24.
235. Kao CH, Zeng ZB, Teasdale RD. Multiple interval mapping for quantitative trait loci. *Genetics* 1999;152(3):1203-16.
236. Zeng ZB. Precision mapping of quantitative trait loci. *Genetics* 1994;136(4):1457-68.
237. Kao CH, Zeng ZB. General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. *Biometrics* 1997;53(2):653-65.
238. Zeng ZB, Kao CH, Basten CJ. Estimating the genetic architecture of quantitative traits. *Genet Res* 1999;74(3):279-89.
239. Ludowyk PA, Hughes W, Hugh A, Willenborg DO, Rockett KA, Parish CR. Astrocytic hypertrophy: an important pathological feature of chronic

experimental autoimmune encephalitis in aged rats. *J Neuroimmunol* 1993;48(2):121-34.
240. Voskuhl RR, Pitchekian-Halabi H, MacKenzie-Graham A, McFarland HF, Raine CS. Gender differences in autoimmune demyelination in the mouse: implications for multiple sclerosis. *Ann Neurol* 1996;39(6):724-33.
241. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. *Genetics* 1994;138(3):963-71.
242. Sillanpaa MJ, Corander J. Model choice in gene mapping: what and why. *Trends Genet* 2002;18(6):301-7.
243. Broman KW, Speed TP. A model selection approach for the identification of quantitative trait loci in experimental crosses. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2002;64:641-56.
244. Ball RD. Bayesian methods for quantitative trait loci mapping based on model selection: approximate analysis using the Bayesian information criterion. *Genetics* 2001;159(3):1351-64.
245. Gillett A, Maratou K, Fewings C, Harris RA, Jagodic M, Aitman T, et al. Alternative splicing and transcriptome profiling of experimental autoimmune encephalomyelitis using genome-wide exon arrays. *PLoS One* 2009;4(11):e7773.
246. Storch MK, Stefferl A, Brehm U, Weissert R, Wallstrom E, Kerschensteiner M, et al. Autoimmunity to myelin oligodendrocyte glycoprotein in rats mimics the spectrum of multiple sclerosis pathology. *Brain Pathol* 1998;8(4):681-94.
247. Gracie JA, Bradley JA. Interleukin-12 induces interferon-gamma-dependent switching of IgG alloantibody subclass. *European Journal of Immunology* 1996;26(6):1217-21.
248. Saoudi A, Kuhn J, Huygen K, de Kozak Y, Velu T, Goldman M, et al. TH2 activated cells prevent experimental autoimmune uveoretinitis, a TH1-dependent autoimmune disease. *Eur J Immunol* 1993;23(12):3096-103.
249. Jagodic M, Olsson T. Combined-cross analysis of genome-wide linkage scans for experimental autoimmune encephalomyelitis in rat. *Genomics* 2006;88(6):737-44.
250. Glorioso N, Herrera VL, Bagamasbad P, Filigheddu F, Troffa C, Argiolas G, et al. Association of ATP1A1 and dear single-nucleotide polymorphism haplotypes with essential hypertension: sex-specific and haplotype-specific effects. *Circ Res* 2007;100(10):1522-9.
251. Mix E, Ibrahim S, Pahnke J, Koczan D, Sina C, Bottcher T, et al. Gene-expression profiling of the early stages of MOG-induced EAE proves EAE-resistance as an active process. *Journal of neuroimmunology* 2004;151(1-2):158-70.
252. Teuscher C, Bunn JY, Fillmore PD, Butterfield RJ, Zachary JF, Blankenhorn EP. Gender, age, and season at immunization uniquely influence the genetic control of susceptibility to histopathological lesions and clinical signs of experimental allergic encephalomyelitis: implications for the genetics of multiple sclerosis. *Am J Pathol* 2004;165(5):1593-602.
253. Subramanian S, Wakeland EK. The importance of epistatic interactions in the development of autoimmunity. *Novartis Foundation symposium* 2005;267:76-88; discussion 88-93.
254. Ramagopalan SV, Ebers GC. Epistasis: multiple sclerosis and the major histocompatibility complex. *Neurology* 2009;72(6):566-7.
255. Motsinger AA, Brassat D, Caillier SJ, Erlich HA, Walker K, Steiner LL, et al. Complex gene-gene interactions in multiple sclerosis: a multifactorial approach reveals associations with inflammatory genes. *Neurogenetics* 2007;8(1):11-20.
256. Kong PL, Morel L, Croker BP, Craft J. The centromeric region of chromosome 7 from MRL mice (Lmb3) is an epistatic modifier of Fas for autoimmune disease expression. *J Immunol* 2004;172(5):2785-94.
257. Williams TM, Moolten D, Burlein J, Romano J, Bhaerman R, Godillot A, et al. Identification of a zinc finger protein that inhibits IL-2 gene expression. *Science* 1991;254(5039):1791-4.

258. Yasui DH, Genetta T, Kadesch T, Williams TM, Swain SL, Tsui LV, et al. Transcriptional repression of the IL-2 gene in Th cells by ZEB. *J Immunol* 1998;160(9):4433-40.
259. Cabanillas AM, Darling DS. Alternative splicing gives rise to two isoforms of Zfhep, a zinc finger/homeodomain protein that binds T3-response elements. *DNA Cell Biol* 1996;15(8):643-51.
260. Manavella PA, Roqueiro G, Darling DS, Cabanillas AM. The ZFHX1A gene is differentially autoregulated by its isoforms. *Biochem Biophys Res Commun* 2007;360(3):621-6.
261. Epplen C, Jackel S, Santos EJ, D'Souza M, Poehlau D, Dotzauer B, et al. Genetic predisposition to multiple sclerosis as revealed by immunoprinting. *Ann Neurol* 1997;41(3):341-52.
262. Gran B, Gestri D, Sottini A, Quiros Roldan E, Bettinardi A, Signorini S, et al. Detection of skewed T-cell receptor V-beta gene usage in the peripheral blood of patients with multiple sclerosis. *J Neuroimmunol* 1998;85(1):22-32.
263. Bennetts BH, Teutsch SM, Buhler MM, Heard RN, Stewart GJ. HLA-DMB gene and HLA-DRA promoter region polymorphisms in Australian multiple sclerosis patients. *Hum Immunol* 1999;60(9):886-93.
264. Togni M, Swanson KD, Reimann S, Kliche S, Pearce AC, Simeoni L, et al. Regulation of in vitro and in vivo immune functions by the cytosolic adaptor protein SKAP-HOM. *Molecular and cellular biology* 2005;25(18):8052-63.
265. Reinhold A, Reimann S, Reinhold D, Schraven B, Togni M. Expression of SKAP-HOM in DCs is required for an optimal immune response in vivo. *J Leukoc Biol* 2009;86(1):61-71.
266. Salmond RJ, Filby A, Qureshi I, Caserta S, Zamoyska R. T-cell receptor proximal signaling via the Src-family kinases, Lck and Fyn, influences T-cell activation, differentiation, and tolerance. *Immunol Rev* 2009;228(1):9-22.
267. Bedoui S, Miyake S, Lin Y, Miyamoto K, Oki S, Kawamura N, et al. Neuropeptide Y (NPY) suppresses experimental autoimmune encephalomyelitis: NPY1 receptor-specific inhibition of autoreactive Th1 responses in vivo. *J Immunol* 2003;171(7):3451-8.
268. Abiola O, Angel JM, Avner P, Bachmanov AA, Belknap JK, Bennett B, et al. The nature and identification of quantitative trait loci: a community's view. *Nat Rev Genet* 2003;4(11):911-6.
269. Li J, Gran B, Zhang GX, Ventura ES, Siglienti I, Rostami A, et al. Differential expression and regulation of IL-23 and IL-12 subunits and receptors in adult mouse microglia. *Journal of the neurological sciences* 2003;215(1-2):95-103.
270. Stefferl A, Schubart A, Storch M, Amini A, Mather I, Lassmann H, et al. Butyrophilin, a milk protein, modulates the encephalitogenic T cell response to myelin oligodendrocyte glycoprotein in experimental autoimmune encephalomyelitis. *J Immunol* 2000;165(5):2859-65.
271. Mana P, Goodyear M, Bernard C, Tomioka R, Freire-Garabal M, Linares D. Tolerance induction by molecular mimicry: prevention and suppression of experimental autoimmune encephalomyelitis with the milk protein butyrophilin. *Int Immunol* 2004;16(3):489-99.
272. Luedi PP, Hartemink AJ, Jirtle RL. Genome-wide prediction of imprinted murine genes. *Genome Res* 2005;15(6):875-84.
273. Izui S. Autoimmune accelerating genes, lpr and Yaa, in murine systemic lupus erythematosus. *Autoimmunity* 1990;6(1-2):113-29.
274. Croker BP, Gilkeson G, Morel L. Genetic interactions between susceptibility loci reveal epistatic pathogenic networks in murine lupus. *Genes Immun* 2003;4(8):575-85.
275. Subramanian S, Tus K, Li QZ, Wang A, Tian XH, Zhou J, et al. A Tlr7 translocation accelerates systemic autoimmunity in murine lupus. *Proc Natl Acad Sci U S A* 2006;103(26):9970-5.
276. Teuscher C, Noubade R, Spach K, McElvany B, Bunn JY, Fillmore PD, et al. Evidence that the Y chromosome influences autoimmune disease in male and female mice. *Proc Natl Acad Sci U S A* 2006;103(21):8024-9.

277. Jansson L, Holmdahl R. The Y chromosome-linked "autoimmune accelerating" yaa gene suppresses collagen-induced arthritis. *Eur J Immunol* 1994;24(5):1213-7.
278. Qi X, Lewin AS, Sun L, Hauswirth WW, Guy J. Suppression of mitochondrial oxidative stress provides long-term neuroprotection in experimental optic neuritis. *Invest Ophthalmol Vis Sci* 2007;48(2):681-91.
279. Lu F, Selak M, O'Connor J, Croul S, Lorenzana C, Butunoi C, et al. Oxidative damage to mitochondrial DNA and activity of mitochondrial enzymes in chronic active lesions of multiple sclerosis. *J Neurol Sci* 2000;177(2):95-103.
280. Encinas JA, Lees MB, Sobel RA, Symonowicz C, Weiner HL, Seidman CE, et al. Identification of genetic loci associated with paralysis, inflammation and weight loss in mouse experimental autoimmune encephalomyelitis. *Int Immunol* 2001;13(3):257-64.
281. Hagan JP, O'Neill BL, Stewart CL, Kozlov SV, Croce CM. At least ten genes define the imprinted Dlk1-Dio3 cluster on mouse chromosome 12qF1. *PLoS One* 2009;4(2):e4352.
282. Raghunandan R, Ruiz-Hidalgo M, Jia Y, Ettinger R, Rudikoff E, Riggins P, et al. Dlk1 influences differentiation and function of B lymphocytes. *Stem Cells Dev* 2008;17(3):495-507.
283. Abdallah BM, Boissy P, Tan Q, Dahlgaard J, Traustadottir GA, Kupisiewicz K, et al. dlk1/FA1 regulates the function of human bone marrow mesenchymal stem cells by modulating gene expression of pro-inflammatory cytokines and immune response-related factors. *J Biol Chem* 2007;282(10):7339-51.
284. Wallace C, Smyth DJ, Maisuria-Armer M, Walker NM, Todd JA, Clayton DG. The imprinted DLK1-MEG3 gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. *Nat Genet*;42(1):68-71.
285. Backdahl L, Bushell A, Beck S. Inflammatory signalling as mediator of epigenetic modulation in tissue-specific chronic inflammation. *Int J Biochem Cell Biol* 2009;41(1):176-84.
286. Wilson CB, Makar KW, Shnyreva M, Fitzpatrick DR. DNA methylation and the expanding epigenetics of T cell lineage commitment. *Semin Immunol* 2005;17(2):105-19.
287. Grogan JL, Mohrs M, Harmon B, Lacy DA, Sedat JW, Locksley RM. Early transcription and silencing of cytokine genes underlie polarization of T helper cell subsets. *Immunity* 2001;14(3):205-15.
288. Mullen AC, Hutchins AS, Villarino AV, Lee HW, High FA, Cereb N, et al. Cell cycle controlling the silencing and functioning of mammalian activators. *Curr Biol* 2001;11(21):1695-9.
289. Bruniquel D, Schwartz RH. Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat Immunol* 2003;4(3):235-40.
290. Young HA, Ghosh P, Ye J, Lederer J, Lichtman A, Gerard JR, et al. Differentiation of the T helper phenotypes by analysis of the methylation state of the IFN-gamma gene. *J Immunol* 1994;153(8):3603-10.
291. Vignali DA, Collison LW, Workman CJ. How regulatory T cells work. *Nat Rev Immunol* 2008;8(7):523-32.
292. Polansky JK, Kretschmer K, Freyer J, Floess S, Garbe A, Baron U, et al. DNA methylation controls Foxp3 gene expression. *Eur J Immunol* 2008;38(6):1654-63.
293. Huan J, Culbertson N, Spencer L, Bartholomew R, Burrows GG, Chou YK, et al. Decreased FOXP3 levels in multiple sclerosis patients. *J Neurosci Res* 2005;81(1):45-52.
294. Genre J, Errante PR, Kokron CM, Toledo-Barros M, Camara NO, Rizzo LV. Reduced frequency of CD4(+)CD25(HIGH)FOXP3(+) cells and diminished FOXP3 expression in patients with Common Variable Immunodeficiency: a link to autoimmunity? *Clin Immunol* 2009;132(2):215-21.
295. Camelo S, Iglesias AH, Hwang D, Due B, Ryu H, Smith K, et al. Transcriptional therapy with the histone deacetylase inhibitor trichostatin A ameliorates experimental autoimmune encephalomyelitis. *J Neuroimmunol* 2005;164(1-2):10-21.

296. Strickland FM, Richardson BC. Epigenetics in human autoimmunity. Epigenetics in autoimmunity - DNA methylation in systemic lupus erythematosus and beyond. *Autoimmunity* 2008;41(4):278-86.
297. Mitsonis CI, Potagas C, Zervas I, Sfagos K. The effects of stressful life events on the course of multiple sclerosis: a review. *Int J Neurosci* 2009;119(3):315-35.
298. Herrera BM, Cader MZ, Dyment DA, Bell JT, Ramagopalan SV, Lincoln MR, et al. Follow-up investigation of 12 proposed linkage regions in multiple sclerosis. *Genes and immunity* 2006;7(5):366-71.