UPTEC 07 013
FEB 2007

ISSN 1401-2138

MIKAEL STRÖM

Characterization of gene regions regulating nerve injury induced expression of MHC class II in the rat

Master's degree project

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC 07 013	**Date of issue 2007-02**

Author

Mikael Ström

Title (English)

Characterization of gene regions regulating nerve injury induced expression of MHC class II in the rat

Title (Swedish)

Abstract

In neurodegenerative diseases, microglial activation and inflammation are important components. These are mainly complex diseases with both a genetic and environmental component, both in many cases equally unknown. In this study, a disease model in rats has been applied to study the genetic influence in the inflammatory and neurodegenerative process with focus and MHC class II upregulation on microglia. The gene *MhcIIta* is shown to regulate MHC class II expression in the early inflammatory response after nerve injury.

Keywords

Neurodegeneration, Inflammation, MHC class II, Complement, QTL, Linkage.

Supervisors

Fredrik Piehl and Margarita Diez

Center for Molecular Medicine, Department of Clinical Neuroscience, Neuroimmunology Unit, Karolinska Institutet

Scientific reviewer

Marta E. Alarcón-Riquelme

Department of Genetics and Pathology, Uppsala University

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **47**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala

Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

Characterization of gene regions regulating nerve injury induced expression of MHC class II in the rat

Mikael Ström

Populärvetenskaplig sammanfattning

Neurodegenerativa sjukdomar är mycket vanliga och saknar dessutom i många fall effektiv behandling. Flera av dem blir vanligare med stigande ålder, t.ex. Alzheimers och Parkinssons sjukdomar. Antalet drabbade kan därför förväntas öka med en stigande genomsnittlig livslängd. Ofta är orsaken till sjukdomarna en kombination av genetiska faktorer och miljöpåverkan. Detta gäller även för Multiple Skleros, vilken oftast utbryter i tidigare ålder. Gemensamt för neurodegenerativa sjukdomar är att det nedbrytande förloppet är kopplat till en inflammatorisk process i det central nervsystemet som bidrar till nervcellsdöd.

De senaste årens framsteg inom de genetiska teknikerna har lett till ökade möjligheter att identifiera och studera geners inverkan på risken att utveckla dessa sjukdomar. I detta projekt har geners inverkan på inflammation i nervsystemet studerats med hjälp av en nervskademodell i råtta. Två olika råttstammar har använts som inflammatoriskt reagerar olika kraftigt på samma skada. Dessutom skiljer de sig i antalet nervceller som överlever skadan. Genom att korsa stammarna i flera generationer sker överkorsningar mellan genomen och mindre fragment från varje ursprungsstam bygger upp genomet i de senare generationerna. Vilka av dessa små fragment som finns med i de individer som får stor påverkan av skadan har därefter studerats med statistiska analysmetoder. Dessa fragment kan därmed anses innehålla de gener som bidrar till det inflammationen.

I denna rapport visas på den stora effekt som genen *Mhc2ta* har har 5 dagar efter skadan på uttrycket av MHC klass II, en molekyl som presenterar främmande proteiner på cellytan och är en viktig del i att aktivera ett immunsvar. Dessutom påvisas en genregion som ger upphov till högt uttryck av komplementmolekylen C1q.

Examensarbete 20 p

Civilingenjörsprogrammet i Molekylär Bioteknik
Uppsala universitet

Februari 2007

Abbreviations

AIL	Advanced Intercross Line
BBB	Blood Brain Barrier
BN	Brown Norway
cDNA	complementary DNA
cM	centi Morgan
CNS	Central Nervous System
DA	Dark Agouti
EAE	Experimental Autoimmune Encephalomyelitis
EM	Expectation Maximization
ER	Endoplasmatic Reticulum
EtBr	Ethidium Bromide
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFAP	Glial Fibrillary Acidic Protein
HKG	Housekeeping Gene
HMBG	High Mobility Group Box Chromosomal protein 1
HMM	Hidden Markov Model
HPRT	Hypoxanthine ribosyltransferase
Ii	Invariant chain
IL	Interleukin
LEW	Lewis
LOD	Logarithm of odds
MHC	Major Histocompatibility Complex
MRF	Microglial Response Factor
mRNA	Messenger RNA
MS	Multiple Sclerosis
PCR	Polymerase Chain Reaction
PNK	Polynucleotide kinase
PNS	Peripheral Nervous System
PVG	Piebald Virol Glaxo
QTL	Quantitative Trait Locus
RT-PCR	Real Time PCR
SSLP	Simple sequence length polymorphism
TBE	Tris/Borate/EDTA
TNF	Tumor Necrosis Factor
VRA	Ventral Root Avulsion

Table of contents

Page nr:

1 Introduction 3
1.1 CNS and disease 3
1.1.1 The neuron and its networks 3
1.1.2 Glial cells 4
1.1.3 Inflammation 5
1.1.4 Diseases with inflammatory components 6
1.2 The disease model 7
1.2.1 Advantages of using models 7
1.2.2 Different rat strains 8
1.2.3 Advanced Intercross Line 8
1.2.4 Ventral Root Avulsion 9
1.2.5 Identified VRA loci 10
1.3 Specific background to this study 11
1.3.1 Previous results 11
1.3.2 The analyzed phenotypes 12

2 Aim of the project 13

3 Material and methods 14
3.1 Sample preparation 14
3.1.1 Breeding and lesion 14
3.1.2 DNA extraction 14
3.1.3 mRNA extraction and cDNA preparation 15
3.2 Genotyping 16
3.2.1 Genetic markers 16
3.2.2 Labelling with ^{33}P 16
3.2.3 PCR reaction 17
3.2.4 Electrophoreses 17
3.3 Phenotyping 17
3.3.1 Housekeeping genes 18
3.3.2 Test of primers 18
3.3.3 RT-PCR with SYBR GREEN 19
3.3.4 Expression analysis 21
3.4 Statistical analysis 22
3.4.1 T-test and non parametric test 22
3.4.2 Linkage analysis 23

4 Results 25
4.1 Evaluation of Housekeeping genes 25
4.2 Strain differences in studied phenotypes 28
4.3 Linkage analysis 29
4.4 HKG influence on LOD score 35

5 Discussion.....36
5.1 Methodology..... 36
5.2 Identified linkage..... 38
5.3 Conclusion and perspectives..... 39
5.4 Primer design error..... 40

Acknowledgements 41

References.....42

Appendix.....45
A1 Markers..... 45
A2 Additional linkage diagrams..... 46

1 Introduction

1.1 CNS and disease

1.1.1 The neuron and its networks

The nervous system is a complex network of connected cells that innervate all parts of the body. The centre for this communication is the brain that processes sensory information and emotional responses, guides motor functions and enables memory and learning. The system operates through fast electric signaling. The speed of this electric signal can reach up to 150 meters per second [Purves et al. 3rd ed] and the signal is transported through extensions from and to the nerve cell body, axons and dendrites. At the end of these extensions the nerve cells can be coupled through chemical or electric synapses, where the signal is transported to the next neuron. Chemical synapses are the most common and here the electrical signal is transferred to the next neuron as a chemical signal by a transmitter substance.

Axons are long extensions that can be up to a meter in length in humans (connecting the foot and spinal cord) [Campbell et al. 6th ed]. Dendrites are shorter extensions, receiving impulses from other neurons, that can be highly branched and connect many different axons, enabling the complex network of signaling in the nervous system. However, the amount of branching differs a lot between nerve cells. Some have extremely branched dendrites, like cerebellar Purkinje cells reflecting the amount of input to cell, while other have a very low amount of branching, for example Retinal bipolar cells. A schematic presentation of a nerve cell is presented in Figure 1.1.

The nervous system can be divided into the central nervous system, CNS, and the peripheral nervous system, PNS. Cells of the nervous system are grouped into neurons (nerve cells) and supporting cells, referred to as glial cells. The CNS includes the brain and the spinal cord while the PNS comprises all other nerves that are divided into cranial and spinal nerves. Further, the CNS can be divided into white matter, mainly consisting of myelinated axons and grey matter consisting of nerve cell bodies. The CNS is in many ways a complex system and it can be estimated that about half of the approximately 20,000-25,000, genes in the human genome are expressed in the CNS [Sandberg et al. 2000, Lincoln 2004, Swanberg et al. 2006].

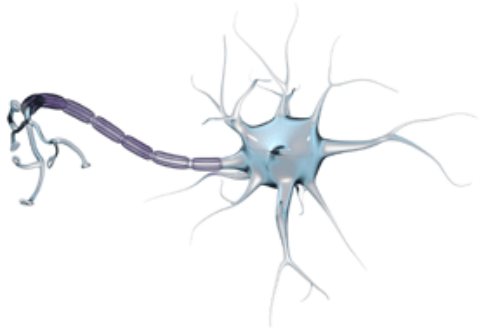


Figure 1.1: Schematic illustration of a nerve cell with branched dendrites and a myelinated axon ending in synaptic contacts. The illustration was used with permission from Anders Sandelin at Sandelinanimations AB.

The speed of the neuronal signaling is highly dependent of the insulating layer around the neurons called myelin. Myelin increases the speed of the signal by reducing the leakage and inducing salutatory conduction and thereby it enables the signal to be transported a longer distance before it has to be regenerated. One important difference between the CNS and the PNS is that in the CNS oligodendrocytes produce the myelin while in the PNS Schwann cells are responsible for this process.

1.1.2 Glial cells

Previously, glial cells were considered to play only a passive role in the function of the CNS, and were therefore given their name from the Greek word for glue. However, today we know that glia are not only supporting cells but also provide functions important for signaling in the CNS. Oligodendrocytes are one of the four different types of glial cells which are present in the CNS. The oligodendrocyte myelinates the axon by wrapping around a laminated layer that consists of 80% lipids and 20% proteins. In contrast to Schwann cells, one oligodendrocyte can form myelin around many different axons.

Second, astrocytes are star shaped glial cells that have a broad range of functions. These include to maintain the structure in the brain, release and uptake of transmitter substances and to provide nutrients for the neurons. Astrocytes also play a role in the formation of scar tissue that impedes regeneration of neurons in the CNS after injury. Furthermore, astrocytes are proposed to produce a number of neurotrophic factors [reviewed in Lidman 2003]. Glial fibrillary acidic protein, GFAP, which is expressed in the CNS almost exclusively by astrocytes [Brenner et al. 1994], can be used as marker for activation of astrocytes.

A third type of glial cells is the microglia that accounts for approximately 10% of the cells in the brain [Neumann 2001]. They act as facultative macrophages of the CNS and remove cellular debris at injury sites. Microglia are sensible to changes in the environment and in response to injury they get activated leading to a number of cellular changes including secretion of pro-inflammatory cytokines, increased capability of phagocytosis, up regulated expression of MHC class II molecules and migration to sites of injury.

Finally, ependymal cells are involved in homeostasis of the brain by maintaining circulation of cerebrospinal fluid in the ventricles of the brain that they line. It has also been proposed that ependymal cells may act as neuronal stem cells that can differentiate into astrocytes in response to spinal chord injury [Johansson et al.1999], a theory that is controversial [Spassky et al. 2005].

1.1.3 Inflammation

Inflammation is an important consequence of activation of the immune system. It is a process that is initiated by the release of cytokines and chemokines from macrophages and by activation of the complement system. Inflammation has basically three main functions: 1) to deliver effector molecules and cells to the site of infection 2) to create a physical barrier to prevent the infection to spread into the blood stream by microvascular coagulation and 3) to promote tissue repair. However, if the molecules eliciting the immune response cannot be cleared from the site of inflammation it can result in a chronic inflammation that causes damage to the tissue. Autoimmunity is a state where an immune response is activated in response to a self antigen. This can result in a chronic inflammatory response that in turn results in more tissue damage and release of more autoantigens. If this process develops into an uncontrolled process it becomes pathological and results in an autoimmune disease.

The CNS has traditionally been considered to be an immunologically privileged organ, and under normal healthy conditions, components of the immune system have a very limited ability to enter the CNS. The blood brain barrier, BBB, is a diffusion barrier that is formed by tight junctions in the epithelia and normally prevents cells and molecules of the immune system from entering the CNS. However, the BBB is nowadays considered to be a dynamic organ and diffusion is to some extent existing [de Boer et al. 2006]. Under pathogenic conditions the milieu in the CNS is changed through activation of genes that enables antigen-specific immune responses and inflammatory processes.

In CNS inflammation, activation of microglia is a key component where the activated cells through MHC class II expression turn into antigen presenting cells (APCs), and migrates to the site of injury. The activated cells also start to secrete signaling molecules including pro-inflammatory cytokines such as TNF and IL-1β [Neumann 2001]. Further, microglia have been shown to express and secrete complement component 3, C3 [Haga et al. 1993] and the components C1 and C1q. Resting microglia also secrete the complement molecules but the expression is highly up-regulated by activation [Aldskogius et al. 1999].

Interestingly, the effect of microglial activation and inflammation is at times a beneficial and at other time, a destructive process. Many studies have shown that enhanced microglial activation and inflammation is associated with greater loss of neurons [reviewed in Block et al. 2007]. There is a clear correlation between the amount of neuronal death and microglial activation after ventral root avulsion, VRA [Lundberg et al. 2001]. Experimental autoimmune encephalomyelitis, EAE, is an animal model of MS

where an autoimmune response is elicited through injection of myelin oligodendrocyte protein (MOG) that leads to degeneration of the myelin. However, rats subjected to both VRA and EAE have shown less loss of axotomized neurons compared to individuals subjected only to VRA [Hammarberg et al. 2000], indicating a positive effect of the inflammatory response. There are also studies that show that neuroinflammation and activation of microglia is crucial for regeneration and it has even been proposed that loss of microglial viability is the cause of neurodegeneration in the aging brain [Streit 2006]. The effect of neuroinflammation and microglial activation is complex and far from well understood. It is most likely a dual effect and the question is when is it is beneficial and when does it cause harm.

1.1.4 CNS diseases with inflammatory components

Neurons can die due to necrosis, apoptosis or a combination of both. Pure apoptosis occurs in the developing brain, but rarely in the adult brain. Inflammation is an important component in neuronal cell death in the adult brain. This makes inflammation an important feature of diseases such as Alzheimer's (AD) and Parkinson's (PD) diseases, the most and second most common neurodegenerative diseases of the CNS [Purves et al. 3rd ed.]. The prevalence of AD in Sweden is around 120 000, and Parkinson's affect between 15 000 and 20 000. The risk of developing dementia is highly increased for people diagnosed with PD [Lökk J et al. 2006]. Microglial activation is shown to be associated with progression of both diseases [reviewed in Block et al. 2007].

Multiple Sclerosis (MS) is another major disease in which inflammation and microglial activation are key components. In this disease myelin is attacked in a chronic autoimmune inflammation which leads to degeneration of the myelin and damage to the axons. MS has a high prevalence in northern Europe and North America. About 400 are newly diagnosed every year in Sweden and around 12 000 are affected by the disease [Rydberg 2006]. Multiple sclerosis, a disease usually diagnosed in early to middle adult life, does not have a known cause but is widely presumed to be a of genetic and environmental factors. Due to this, MS and other autoimmune and non-familial neurodegenerative diseases are referred to as complex diseases.

Furthermore, inflammatory reactions are likely to affect functional outcome after traumatic injuries and stroke as a result of the secondary degeneration that occurs in the surrounding tissue due to the inflammatory processes. Secondary degeneration contributes to a large extent to outcome of traumatic brain injury and reducing the elicited inflammatory process can reduce the loss of neurons [Liu et al. 2006].

1.2 The disease model

1.2.1 Advantages of using models

Identification of genes that are involved in the inflammatory processes of complex neurological diseases is important for our insights into the cause and progression of the diseases. Hopefully, such discoveries can also contribute to development of future drugs to better manage these diseases. In order to find genes involved in the inflammatory process in CNS disease a model has been set up in rats to find so called Quantitative Trait Loci (QTLs). A QTL is a genomic region that is statistically associated with an observed phenotype. The QTL can contain one or more genes involved in the observed pattern. By studying differences in response to an induced injury between inbred strains and defining genetic differences in parallel, these variations can be associated to genetic influences of certain regions.

There are several reasons why an animal model is the most suitable model for these types of genetic studies. First of all, a large number of subjects is needed to determine genetic influence by statistical methods. Injuries to the human nervous system due to accidents are so divergent that it is difficult to establish individual differences between outcomes that are not due to differences in the lesions. This coupled with the need for a large number of subjects makes it extremely difficult, if not pragmatically impossible, to perform studies of this type in humans. In complex diseases, it is often possible to gather larger number of comparable samples, but the complex nature and considerable degree of heterogeneity still makes it hard to identify the genetic susceptibility factors. Furthermore, the absence of environmental conditions that contribute to the development of a disease may mask the precise genetic factors. Unaffected individuals may be genetically susceptible, thereby reducing the genetic differences between groups of affected individuals and healthy controls.

Therefore, relevant animal models of disease are very valuable for finding candidate regions involved in disease development. Homologues to the identified regions can be studied in the human genome in samples from patients affected by disease.

1.2.2 Different rat strains

Inbreeding is characterized by mating of genetically related individuals. If inbreeding is continued for many generations it results in inbred strains, having a minimized genomic variation between individuals. Thereby, all individuals are considered to be gnomically identical. Many different laboratory rat strains exist that are well characterized for polymorphic genetic markers, such as Simple sequence length polymorphisms (SSLPs), throughout their genomes.

In this study two rat strains DA(*RT1*^{av1}) and PVG.1AV1 have been used where the PVG.1AV1 is a congenic strain with the same MHC haplotype as the DA rat. In parallel, the strains BN and LEW.1N have been used where LEW.1N is a congenic with the same

MHC haplotype as the BN rat. Different MHC complexes can have an effect on the inflammatory process and by using these congenic strains, phenotypic variations resulting from differences in non MHC genes can be studied.

1.2.3 Advanced Intercross Line

The use of an Advanced Intercross Line (AIL), is a method that enables fine mapping of QTLs. When two inbred strains are crossed, recombinations will occur in the parental chromosomes during meiosis resulting in a mix of fragments from the parental chromosomes in the chromosomes of the new offspring. That is, the F_2 generation will have chromosomes containing a mixture of the chromosomes from the two inbred strains. The F_2 generation can then be intercrossed which will result in more recombinations and thus shorter fragments combined from the parental generation. F_3, F_4, F_5, F_6….etc generations are thereafter produced by random intercrossing, avoiding brother-sister mating, leading to an accumulation of recombination throughout the genome as illustrated in Figure 1.2. Initially, QTLs spanning larger intervals are regularly detected in F_2 generations and thereafter fine mapped in later generations to reduce the number of genes within the QTL.

The AIL strategy is based on the fact that repeated intercrossing will reduce the linkage disequilibrium for an unlinked locus so that it finally will approach 0.5. How many generations that are required are dependent on the recombination frequency, according to the equation in Figure 1.2. It is necessary that the AIL is set up with a minimum effective number of 100 individuals per generation. The expected proportion of recombinants in a generation, $r_{t,}$ is a function of the recombination in the previous generation, r_{t-1}, plus the net increase of recombinants as described by equation 1. This net increase is dependent of recombination between double heterozygotes that generate new recombinants minus the recombination between heterozygotes that are recombinant haplotypes and result in regeneration of the parental haplotype.

$$r_t = r_{t-1} + r\frac{(1 - r_{t-1})^2}{2} - r\frac{r_{t-1}^{\ 2}}{2} \qquad \text{Equation 1}$$

An F_{10} generation has a fivefold reduced 95% confidence interval for the QTL location compared to an F_2 [Darvasi et al. 1995], thereby considerably reducing the length of the fragment that contain the regulating gene

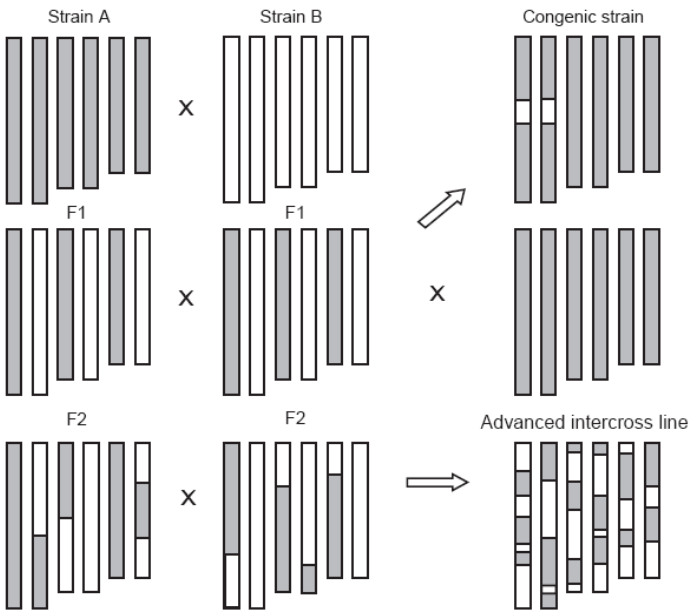


Figure 1.2: Illustration of the breeding method behind an AIL. Included is also the backcross method that is used to make congenic strains where only one segment of interest from strain B is included in the genome of stain A. This also takes many generations with active selection of individuals in each generation having the right segment for further breeding. This is what has been done with the PVG.1AV1 and LEW.1N strains. The figure was kindly provided by Olle Lidman.

1.2.4 Ventral Root Avulsion

The nerve lesion model that has been used in this study is called ventral root avulsion (VRA). VRA is performed by exposing the lumbar ventral roots L3-L5, which are subsequently avulsed by being pulled caudally. This injury results in axotomy of axons at the CNS-PNS border without any damage to the cord itself. The lesion results in degeneration of a restricted population of nerve cells, spinal cord motoneurons within the ventral horn and it leads to an activation of surrounding microglia and astrocytes (see Figure 1.3). The loss of neurons is very limited one week after surgery but the degeneration thereafter progresses so that a majority of the lesioned cells degenerate during the second and third week after surgery [Piehl 1999, Lundberg 2001]. Activation of microglia is a process that starts rapidly after VRA and can be detected through increased levels of microglial response factor (MRF)-1 [Lundberg et al. 2001]. The activation of astrocytes occurs later than the microglial response and can be detected as early as one week after VRA through increase in GFAP expression. By three weeks after a lesion is created, a more than 10-fold increase in expression can be detected on the injured side compared to the uninjured [Olsson et al. 2005]. Leukocytes are present in the scar tissue, but the amount of infiltrating leukocytes in the grey matter is low [Lundberg et al. 2001]. This is important when MHC class II expression is studied, showing that increased MHC class II levels after VRA are due to up regulation in microglia and not dependent on infiltrating cells. An advantage with the model is that exactly the same lesion can be induced in a large number of individuals.

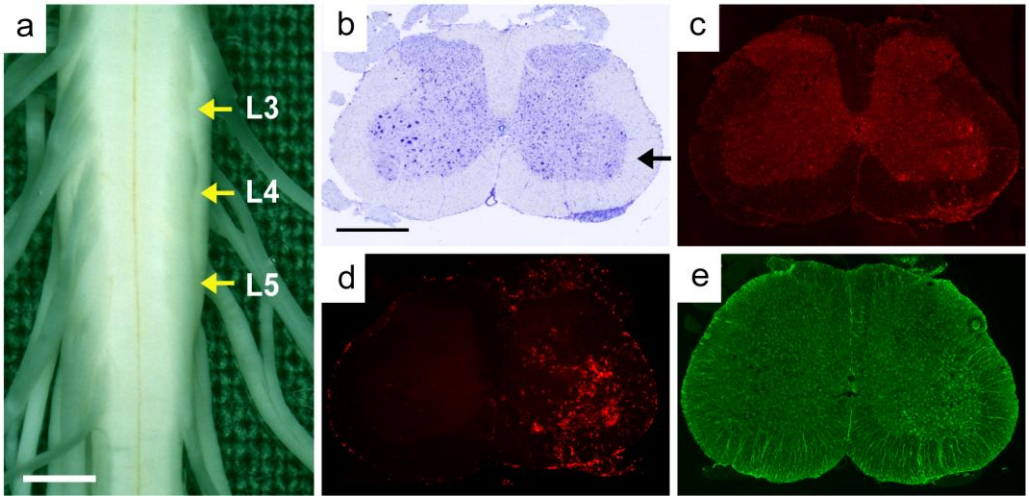


Figure 1.3: Photos of the spinal chord and section taken two weeks after avulsion. (a), spinal chord with the avulsed roots marked with arrows and scale bar 2 mm. Staining with cresyl violet indicating loss of motoneurons (b), scale bar 1 mm and arrow indicating the avulsed side in the ventral horn. Immunolabelling of sections (c-e) with CD11b/c indicating microglial activation (c), OX-6 for MHC class II expression (d), and GFAP (e) for astrocyte activation. Photos adapted from Lidman 2003 and modified.

1.2.5 Identified VRA loci

The methodology described has been used for several years and has resulted in identification of gene regions affecting the outcome of VRA. A former study reported variation in nerve cell death and inflammatory response in the strains further studied here, DA and PVG.1AV1, together with other included strains. This variation was independent of differences in MHC haplotypes [Lundberg et al. 2001]. The two strains DA and PVG.1AV1 showed the highest degree of difference in nerve cell death, microglial and astrocyte activation, changes in C3 and MHC class II expression levels (data presented in Figure 1.4). A more intermediate response was detected between the strains BN and LEW.1N. The same pattern was also detected one and three weeks after VRA, with one exception. BN clearly showed the highest expression of MHC class II one week after VRA, but together with LEW.1N a more intermediate response after 3 weeks (Figure 1.4).

Previously four different QTLs, *Vra1-4*, have been identified by the methodology used in this project. These QTLs are linked to neurodegeneration (*Vra1* and *Vra2)*, T-cell infiltration (*Vra2* and *Vra3)* and expression of MHC class II on microglia (*Vra4*) [Lidman et al. 2003].

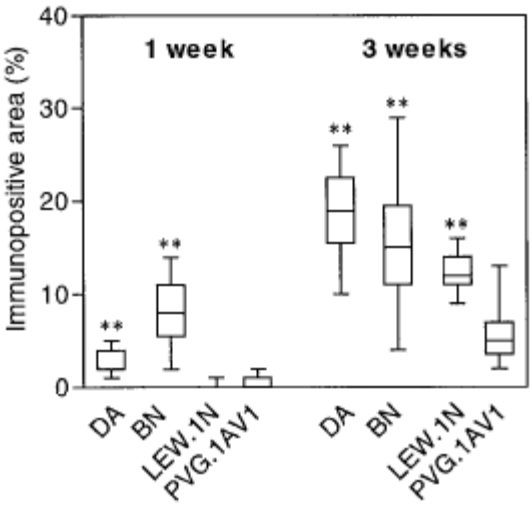


Figure 1.4: The expression of MHC class II on microglia in different strains obtained by Ox6 immunolabelling. The BN clearly has highest expression of MHC class II one week after VRA. DA and PVG.1AV1 show the largest difference in expression three weeks after VRA, while BN and LEW.1AV1 has a more intermediate response at this timepoint. Figure adapted from Lundberg et al. 2001 and modified.

The difference between DA and PVG.1AV1 in expression of MHC class II on microglia 3 weeks after VRA was defined as a locus on chromosome 10, *Vra4.* Recently this locus was fine mapped and the gene encoding the MHC class II transactivator, *Mhc2ta*, located within the *Vra4* locus could be identified as a key regulator of the MHC class II expression. This was done in an AIL experiment in F_8 generation between DA and PVG.1AV1, and studies of the human analog, *MHC2TA*, showed that this gene is associated with a variety of diseases including MS, rheumatoid arthritis, and myocardial infarction [Swanberg et al. 2005]. However BN and LEW.1N were demonstrated to be identical at the *Mhc2ta* gene. Although, BN has a higher expression of MHC class II on microglia in response to VRA. This suggests additional genetic influence located elsewhere in the genome.

1.3 Specific background to this study

1.3.1 Previous results

As previously described, haplotype maps demonstrated that BN and LEW.1N have similar or identical *Mhc2ta* alleles, but still display a difference in the early upregulation of MHC class II expression one week after VRA [Lundberg et al. 2001]. Therefore, other genes than *Mhc2ta* are probably responsible for this difference seen early between BN and LEW.1N. In order to define these regulating genetic regions an F_2 (BNxLEW.1N) cross was created and the response studied 5 days after VRA. Possibly, other genes than *Mhc2ta* cause the difference in early upregulation of MHC class II also between DA and PVG.1AV1, and possibly this is caused by the same genes as for BN and LEW.1N.

The F_2 (BNxLEW.1N) cross has been undergoing a whole-genome scan study. To start with, linkage for MHC class II expression was studied together with expression of molecules of the complement system, C3 and C1q. In the F_2 generation the amount of

recombination is rather low, and this study will therefore render larger gene regions. Samples from an F_{12} AIL between DA and PVG.1AV1 have previously been collected. Therefore it might be possible to verify linkage found between in BNxLEW.1N, by studying the identified regions in the F_{12} cross between DA and PVG.1AV1. If possible, that would able fine mapping of the regions without setting up more generation in a BNxLEW.1N AIL. Furthermore, this F_{12} study could answer if *Mhc2ta* possibly has a role in the early up regulation of MHC class II.

1.3.2 The analyzed phenotypes

In total, four different phenotypes were analyzed. Expression of three different genes encoding CD74, C3 and C1q and weight loss, were studied five days after VRA. MHC class II expression was measured by studying the expression of invariant chain (Ii), *CD74*. CD74 has two main functions in the cell. First, it binds to the newly synthesized MHC class II heterodimers when it, as all cell surface glycoproteins, are translocated to the endoplasmatic reticulim, ER. The ER has a high concentration of immature proteins that must be prohibited from binding to the MHC molecule. CD74 acts by binding to the peptide-binding groove of the MHC class II molecule and thereby hindering other proteins from binding. The second function is to target the delivery of the MHC molecules to vesicles where the CD74 molecule is cleaved so the MHC class II molecules can bind antigens. CD74 is thereby crucial for making functional molecules on the cell surface and is thereby a relevant measure of the MHC class II expression. Previously, an extremely good correlation between expression of CD74 and levels of MHC class II has been reported [Swanberg et al. 2005]

C3 and C1q are molecules of the complement system that are part of innate immunity. The complement system is made up of many different plasma proteins that react in a series of cleavage reactions through three different pathways that ultimately result in inflammatory responses and opsonization of invading pathogens. The expression of the chosen molecules is important for the early stages in complement activation. C1q is together with C1r and C1s part of the C1-complex that initiates activation through the classical pathway. In the alternative pathway, C3 is spontaneously activated. All pathways leads to the formation of the enzyme C3 convertase that cleaves the molecule C3 into C3a and C3b. C3a mediates local inflammation and C3b is involved in opsonisation and formation of the C5 convertase that produce the C5a and C5b molecules. C5a also triggers inflammation and C5b is responsible for the assembly of molecules (C5d-C9) that forms the MAC, membrane attack complex, which make a hydrophilic pore damaging the membrane of certain pathogens [Janeway et al.]. C1q was recently found to be differentially regulated between DA and PVG.1AV1 in an Affymetrix oligonucleatide array study of 278 regulated genes were the levels correlated to the degree of neurodegenaration [Swanberg et al. 2006].

2 Aim of this project

This study aims to verify gene regions that show linkage to regulation of gene expression in the BNxLEW.1N study in an F_{12} material between DA and PVG.1AV1. In addition, the study will examine the role of *Mhc2ta* in the early inflammatory response five days after VRA. Another gene, or genes, other than *Mhc2ta* is responsible for the difference between BN and LEW.1N may also influence the early response in DA and PVG.1AV1. If regions are identified and verified in the F_{12} material, these can simultaneously be reduced to a shorter genomic region. In addition, the effect of new QTLs in addition to the *Mhc2ta* influence can be examined and possible additative effects studied.

Main goals of this project:

- Investigate a possible role of *Mhc2ta* in MHC class II up-regulation on microglia 5 days after VRA
- Search for other QTLs involved in nerve injury-induced inflammation through MHC class II expression and complement activation.
- Determine if it is possible to verify linkage found in F_2 BNxLEW.1N in F_{12} DAxPVG.1AV1, and thereby fine map the QTLs of interest.
- Evaluate the major housekeeping gene GAPDH and the effect on linkage analysis by using other and multiple housekeeping genes in expression analysis.

3 Materials and Methods

3.1 Sample preparation

An AIL in the twelfth generation takes many years to set up and is therefore not possible to include in a masters degree project. At the time of the initiation of this project some ground work had previously been done. The F_{12} breeding, lesion, DNA extraction, mRNA extraction and cDNA synthesis had been performed. However, all these steps except from F_{12} breeding have by the author instead been performed in material from other ongoing studies at the lab.

3.1.1 Breeding and lesion

The animals were kept in a 12 hour light/dark cycle under pathogen free condition in the breeding facility at CMM. Originally, the PVG.1AV1 strain was obtained from Harlan UG, Ltd and the DA(*RT*1^{av1}) was provided by H.Hedrich (Mediziniche Hochschule, Hannover, Germany). A total of 191 animals were subjected to VRA and samples were collected five days after the lesion. Of the 191 individuals 163 were from the F_{12} AIL and 28 were pure DA and PVG.1AV1 individuals. The avulsed animals were killed using carbon dioxide and perfused with cold PBS. The spinal chord was thereafter carefully removed and studied to verify that the L3-L5 roots were avulsed without damage to the spinal chord. Segment L3 was cut out and used for mRNA extraction to be used in phenotype analysis. The tip of the tail was used for DNA extraction for the genotyping. Segments L4 and L5 are usually used for morphologic studies which however were not included here.

All experiments had been approved by the Stockholm Committee for ethical animal experimentation.

3.1.2 DNA extraction

DNA for genotyping was extracted from 2mm tail tip by mixing with 0.5 ml lysis buffer (100mM Tris HCl, 5mM EDTA, 0.2% SDS, 200mM NaCl) and 5µl Proteinase K (10µg/ml) per sample. The tube was shaken for 2 hours at 55°C and thereafter spun at 14 000 rpm at 4°C for 10 minutes. The supernatant was transferred to a tube containing 0.5 ml isopropanol (4°C) and shaken to precipitate the DNA. This was centrifuged at 14 000 rpm at 4°C for 10 minutes and the resulting pellet dissolved in 100µl dH2O. This

was then finally incubated at 4°C over night to dissolve the DNA. The extracted DNA was kept at 4°C and used for the genotyping.

3.1.3 mRNA extraction and cDNA preparation

Complementary DNA is synthesized from an mRNA strand and is made up of a DNA which is complementary to the RNA strand. This conversion can be made by using a special transcription enzyme, reverse transcriptase. Because mRNA is very unstable and sensitive to degradation, cDNA is preferred for storage and analysis of mRNA levels.

Tissue for mRNA extraction was taken from the ipsilateral quadrant of the L3 segment, illustrated in Figure 3.1, and kept in -80°C until extraction. The tissue was homogenized by using FastPrep, (Qbiogene), and thereafter extracting mRNA using a Qiagen Total RNA extraction kit (Qiagen, Holden, Germany)

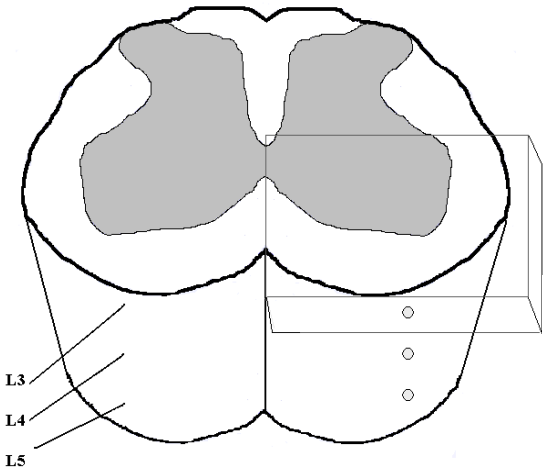


Figure 3.1. Illustration of the piece of the spinal chord section taken out from the rats with L3-L5 avulsed nerve roots. The box indicates the ipsilateral quadrant that was taken out for mRNA extraction.

cDNA was synthesized by mixing 10µl RNA sample with 7µl of SRM, synthesis reaction mix (4µl 5xFirst strand buffer(BRL), 2µl 10mM dNTPs in DEPC H2O, 1µl primer – random hexamer 0.1µg/µl) and incubated for 5 min at 70°C with following transfer upon ice. To the sample 3µl Enzyme mix (1µl 0.1M DTT, 1µl RNA guard (Invitrogen), 1µl SuperScript 200U/µl (Invitrogen) centrifuged for 10 sec at 13 000 rpm) was added and samples were incubated at room temperature for 10 min. Thereafter, samples were incubated at 42°C for 60 min, followed by 10 min at 70°C and finally a short centrifugation at high speed. The samples were added 80µl of dH_2O and stored at -20°C until used for quantitative PCR.

3.2 Genotyping

Genotyping is the process of determining the genotype of different individuals. In the AIL animals this means that it is determined from which parental strain a certain segment of the genome in the F_{12} generations animal stems from.

3.2.1 Genetic markers

A genetic marker is region that differs between strains and thereby can be used to keep track of a region in breeding experiments. Simple sequence length polymorphisms, SSLPs, are sequences that differ in length between strains and thereby it can be identified from which strain that particular gene segment stems from. Here, SSLPs were selected based on earlier studies and information from available databases on the internet (Rat Genome Database, RGD, http://rgd.mcw.edu/, Ensembl genome browser, http://www.ensembl.org/). Primers were ordered online from Sigma-Proligo (www.proligo.com).

The choice of markers was based on peak markers from the F_2 (BNxLEW.1N) single marker regression analysis indicating markers on chromosomes 1 and 7 as highly interesting. On chromosome 10, the peak marker for *Mhc2ta* and surrounding markers were chosen together with a distant marker based on BNxLEW.1N. Totally, 39 markers (8 previously genotyped in the lab) spread on the chromosomes 1, 4, 7, 8, 10 11, 12, 13, 15, 16 and 18 were successfully genotyped with focus on chromosome 1 and 7 that showed clearest linkage tendency with single marker regression in BNxLEW.1N. Markers were also added during the project based on the ongoing BNxLEW.1N study and preliminary results from the first performed linkage analysis.

Chromosome 8 had previously been genotyped in the region for *Vra1* in order study the effect of complement molecule C3. Here it was extended with more markers due to findings in BNxLEW.1N and a tendency to linkage to another phenotype. All markers with their primer sequences and genomic positions are presented in Appendix 1.

3.2.2 Labeling with ^{33}P

For purposes of detection, the PCR product is labeled with a radioisotype. By labeling the forward primers, every second synthesized DNA chains will contain a radioactive molecule which can be detected when the product is analyzed. Labeling was done by mixing 0.136µl forward primer with 0,064µl 10*kinase mix (0,024µl 10*kinase buffer, 0.0136µl 20µM ATP, 0.0264µl of 18µM spermidine), 0,014µl (10U/µl) T4 PNK (BioLabs) and 0.026µl ^{33}P-ATP (PerkinElmer). The mix was incubated at 37°C in a water bath for a minimum of 60 minutes and thereafter the PNK enzyme was inactivated by heating to 95°C for 90s on a PCR machine.

3.2.3 PCR reaction

Polymerase chain reaction (PCR), performed on 96 well plates was used to amplify the fragments to be analyzed. DNA from pure DA and PVG.1AV1 individuals were included as reference on the plates together with controls containing no DNA. For each sample 0.240µl labeled forward primer was mixed with 0.136µl reverse primer together with 3.6µl PCR mix (1.2µl dH_2O, 0.8µl 10*PCR buffer (15mM $MgCl_2$), 0.8µl dNTP (2mM)), 0.8µl Tween 20 (2mM)) and 0.032µl AmpliTaq Gold. Into the wells 4µl of DNA from the different individuals was added and finally a drop of mineral oil was used to cover the reaction mix in order to prevent evaporation.

The plate was run on according to the following program: 13 min at 94ºC, thereafter 30 cycles with 30s 94ºC, 1 min 55ºC, 90s 72ºC and finally 72ºC for 7min. The PCR product was kept in -20ºC until electrophoresis was run.

3.2.4 Electrophoresis

The PCR product was subjected to electrophoresis on a 6% polyacrylamid gel that separates the product with respect to size. Smaller fragments migrate faster in the gel than larger ones and if the markers are polymorphic, a difference can be detected in how long distance the samples have gone after a certain amount of time.

To the PCR product 2.5µl loading buffer were added and the samples were denatured for 5 minutes in 95ºC on a PCR machine. 2.5µl of the final mixture was loaded on the gel and run at 2000V, 75W, 225mA for around two hours, depending on the expected fragment size. Thereafter the gel was transferred to paper and dried for around two hours at 80ºC in a vacuum drier.

The dried gel was transferred to a light resistant cassette together with a BioMax MR Film (Kodak) in a dark room and left to be exposed for approximately two days, depending on the strength of the radioactivity. Thereafter the film was developed in a Kodak developer. The film was manually analyzed by two individuals separately. The genotype results for all markers were gathered in a Microsoft Excel file to be used later together with phenotypes for linkage analysis.

3.3 Phenotyping

The phenotype is characterized by features of the individual that is influenced by the interplay between genotype and the environment. The phenotype can be anything from colour of the fur to blood group or amount of molecules relevant for inflammation. The last description is in this project studied by investigating the level of expression of some genes that produce molecules of the innate and adaptive immune system. This has been

done with a technique called Real Time PCR (RT-PCR), where the amount of mRNA in the cells is studied.

3.3.1 Housekeeping genes

A housekeeping gene (HKG) is a gene whose expression is unaffected by the injury or other experimental manipulation and thereby has the same amount of expression in both affected and unaffected tissue. This is important because the values from the expression of the studied gene are dependent of how many cells the mRNA was extracted from. Even though visibly identical equally-sized pieces of tissue are used, the actual cell counts will differ. The results from a housekeeping analysis will thereby give a measurement of how many cells each sample contained and relating the studied gene to this will result in a value that resemble the relative expression per cell.

There are many HKGs described in the literature. Unfortunately there is no standard housekeeping gene suitable for all experiments [Schmittgen et al. 2000, Kok et al. 2005]. Genes that work in one experimental setup might show increased expression in another type of pathological process. Each experimental setup will gain by verifying the appropriate HKGs for that particular analysis, and generally using multiple HKGs is considered the best option [Kok et al. 2005]. To start with two different well known HKGs were used here, Glyceraldehyde-3-phosphate dehydrogenase [GAPDH], an enzyme involved in glycolysis, and ß-actin one of the proteins that make up the cytoskeleton.

However, these two HKGs gave varying results raising the suspicion that ß-actin was a less appropriate choice (see Results). Therefore, primers for two additional HKGs, Hypoxanthine ribosyltransferase [HPRT], and the ribosomal subunit 18s, were included in the study. The choice was based on other published studies that have shown HPRT to be an efficient HKG [Peinnequin et al. 2004: Chen et al. 2006: Kok et al. 2005] and the sequences for the primer pair from the papers were used. 18s RNA is a widely used HKG that previously had been unsuccessful in the lab. Therefore new primers were designed for 18s using Beacon designer version 5.1 with the intention of finding primers that span an exon-intron interval. This lowers the risk of amplification due to contamination of genomic DNA.

3.3.2 Test of primers

Prior to initiating the expression studies with RT-PCR, all primers were validated for specificity and optimal annealing temperature. This was done in order to make sure that the primers were efficient and gave unique amplification products. Another purpose of this test was to examine which annealing temperature that was most efficient for each primer. This differs between primers depending on length and GC content. It is important with a high efficiency to get accurate results, but too low temperature can result in unspecific binding. Therefore 8 samples with a gradient between 55 and 65°C were run

on the plate in a three step cycle (1x 95.0 °C 3 min, 45x 95°C for 10s 55°C -65°C for 30s [well A: 65.0°C, B: 64.5°C, C: 63.3°C, D: 61.4°C, E: 58.9°C, F: 57.1°C, G: 55.8°C, H: 55.0°C] finally 71x 55.0°C-90.0°C for 30s).

Finally the PCR products were run on a gel to check the products. This is regularly done with an agarose gel containing EtBr which incorporates into double stranded DNA. Incorporated EtBR can thereafter be detected with UV light. A 2% agarose E-gel (Invitrogen), was used together with Loading buffer green (Fermentas). The gel was run with 7µl PCR product, 3µl loading buffer and 10µl dH2O together with a ladder for 50bp at 400mA and 70V during 15-30 minutes.

Due to rather weak signal, an additional and more sensitive test with silver staining was used. For the silver stain a 20% TBE gel (Invitrogen), was used with 3µl PCR product, 3µl loading buffer and 14µl dH_2O and a ladder for 50bp and run at 150V/1.4W for 1.5 hour. Silver stain was performed with a manual PlusOne^{TM} DNA Silver Staining Kit (Amersham Biosciences), according to the producer's instruction.

Primer	**Fw primer**	**Rew primer**
CD74	GTGATGCACCTGCTTACGAAGT	CTCCGGGAAGCTCCCCT
C3	TGCGGCTGGAGAGTGAAG	TTACTGGCTGGAATCTTGATGG
C1q	GACCCAGTACAGCTGCTTTGG	TCATAGAACACGAGGATTCCATACA
18s*	AGTCCGTCAAGCCAATCTAC	CAGCAGTATCCCTCCCATTAG
HPRT	CTCATGGACTGATTATGGACAGGAC	GCAGGTCAGCAAAGAACTTATAGCC
GAPDH	TCAACTACATGGTCTACATGTTCCAG	TCC CAT TCT CAG CCT TGA CTG
β-actin	CGTGAAAAGATGACCCAGATCA	AGAGGCATACAGGGACAACACA

Table 3.1: Primers used in RT-PCR analysis with their respective forward and reverse sequences. The first three are studied phenotypes and the others are HKGs. The 18s primer sequences is the one spanning the exon-intron border that was found to be most specific. 18s was the only primer that was newly designed within this study. All others were previously designed and used at the lab, except for HPRT (sequence from Peinnequin et al. 2004).

*A mistake was later revealed and this primer actually amplifies the mitochondrial ribosomal subunit S18A. See section 5.4 for explanation.

3.3.3 RT-PCR with SYBR GREEN

Real Time PCR was carried out on an IQ5 machine from BioRad Laboratories. SYBR Green is a molecule that binds to the minor groove of double stranded DNA. When the molecule binds it increases the emission of light. SYBR Green is therefore not sequence specific and no special labeled primers need to be designed. However, because it is non-sequence specific it binds to all amplified fragments and thereby it is extremely important to have pure amplification with no non-specific products that otherwise will increase the signal.

A dilution series containing five samples in duplicates with ten times dilution between each sample was included on all plates as a standard curve. Because the targets that are studied have rather low expression it is not possible to dilute the cDNA 10 000 times and therefore dilution series of amplicon was used. The standard curve is used for calculation of the efficiency of the amplification which is used in the calculation of the relative quantity.

A threshold value in the exponential phase is set and from this so-called Ct values are achieved which reflect how many cycles that are needed to reach the threshold value of fluorescence for the different samples. The Ct value is proportional to the initial cDNA concentration. The Ct values are used for calculating the standard curve and differences in the unknowns are used for relative expression calculations. By this the variation in expression between individuals is detected giving a relative value of the expression. Quantitative analysis can be performed with RT-PCR if a standard curve with determined initial concentration is used. In most studies this is however not beneficial for interpreting the result and the amount of difference in expression is what is valuable to investigate.

The RT-PCR was run with the same programme as in the primer test but without temperature gradient. An annealing temperature of 60ºC was found to be suitable for all targets. Cycle two is the amplification cycle where the cDNA is amplified. In cycle 3 a melting curve is calculated by the decrease of the signal as the double stranded amplicon gets denatured by increasing temperature. Denaturing results in release of SYBR-green molecules that thereby will stop signaling. The melting temperature with highest signaling decrease can be detected. If the product is pure it will denature at a single temperature and give a single peak in the diagram. If peaks at two temperatures are detected, it can be due to unspecific amplification.

All samples were run in duplicates. An average value was used in the analysis when the samples correlated well. In cases where a big difference was seen one or both results were excluded from the study.

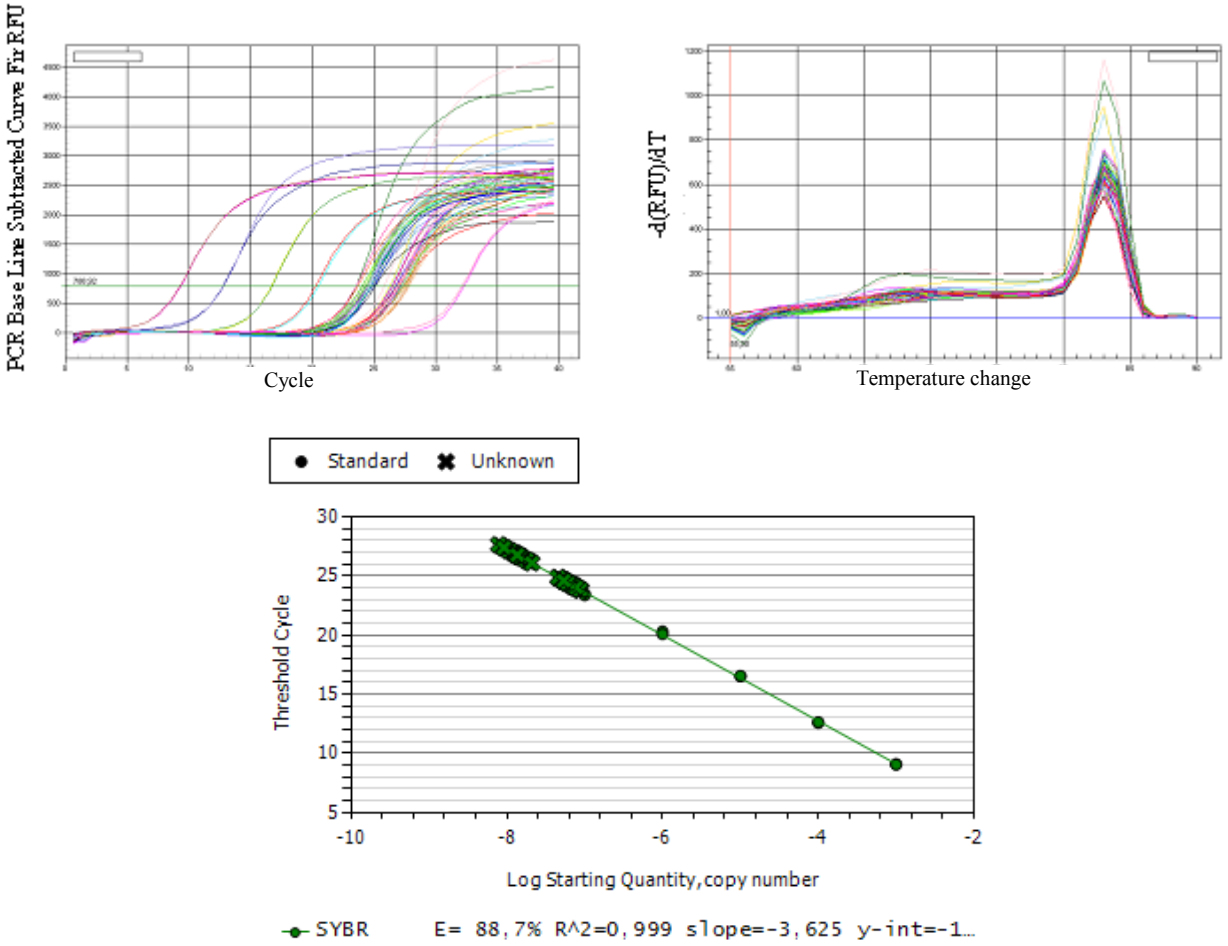


Figure 3.2: Examples of amplification (top right), melt peak (top left) standard curve (bottom) charts taken from one plate of the analysis of HPRT in this study. For each target totally 5 plates was run to fit all individuals in duplicates into the study.

3.3.4 Expression analysis

The iQ^{TM}5 Optical Software can calculate the normalized expression of the studied gene using single or multiple housekeeping genes. Normalized expression (equation 3) is defined as the relative quantity (equation 2) of the studied target divided by the relative quantity of the HKG(s).

$$\Delta C_{T\ sample(Gene\_X)} = E_{Gene\_X}^{(C_{T\,(ctrl)} - C_{T\,(sample)})} \qquad \text{Equation 2}$$

ΔC_T = Relative Quantity
E = Efficiency (100% efficiency = 2 , 0% efficiency =1)
$C_{T\ (ctrl)}$ = Average C_T for the sample assigned as control for Gene x
$C_{T\ (sample)}$ = Average C_T for the sample

$$\text{Normalized Expression}_{sample(Gene\_X)} = \frac{\Delta C_{T\ sample(Gene\_X)}}{(\Delta C_{T\ sample(HKG\_1)} * \Delta C_{T\ sample(HKG\_2)} * \ldots * \Delta C_{T\ sample(HKG\_n)})^{1/n}}$$

Equation 3

Due to the number of individuals the results from each target was spread on five plates and runs. To be able to compare individuals from different plates an inter run calibrator has to be set. This is an identical sample that is included on every plate for each target. The differences in Ct value for this identical sample between plates can therefore be assumed to be due to variations between the runs. This difference is then used by iQ™5 Optical Software in an inter-run calibration algorithm to counteract effects on inter variation on the unknowns. As inter-run calibrator, one sample from the standard curve was chosen.

3.4 Statistical analysis

3.4.1 T-test and non parametric test

Parental phenotypes were tested for statistical significance in expression of the targets between mean values for all analyzed phenotypes. This was done by performing a two tailed T-test using GraphPad Prism version 3.00 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com).

A T-test calculates the so called t value which is defined as the difference between the mean values of the two groups divided by the variability of the groups, expressed in equation 4. The variability is the square root of the sum of the variance in the groups divided by the number of observations in the group, where the variance is the squared standard deviation (SD), equation 5.

$$T = \frac{X_{mean} - Y_{mean}}{\sqrt{\frac{var_X}{n_X} + \frac{var_Y}{n_Y}}}$$

Equation 4

$$SD = \sqrt{\frac{\sum (X - X_{mean})^2}{(n-1)}}$$

Equation 5

X = each value
X_{mean} = the mean value
n =number of values

$var = SD^2$
Degrees of freedom = $(n_x + n_y) - 2$

How high the T-value needs to be to indicate a significant difference between the subsets of values depend not only on the T-value but also on the degrees of freedom and the alpha value. The degrees of freedom equal the number of observations minus two. The alpha value, significance, is set by the researcher and is often 0.05, meaning that 95 times out of hundred the difference will represent a true difference (i.e. not false positive), is considered a relevant value. The value that T has to be above is defined in certain tables and of course automatically given in programs such as GraphPad Prism.

However, the t test requires that the studied groups show normal distribution. This can not always be assumed for the expression of the phenotypes studied and therefore mainly a non parametric test has been used. The Mann-Whitney U-test can be seen as a non parametric T-test. In the test, all values are ranked and the ranking is used for calculating the U value according to equation 6. It is tabulated what value the U has to be above and is also dependent on the degrees of freedom and the significance set.

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \qquad \text{Equation 6}$$

R_1 = Sum of the rankings for values in group 1
n_1 = Number of samples in group 1
n_2 = Number of samples in group 2

For both tests when shown in graphs one star (*) indicate 95% confidence interval, two stars (**) a 99% confidence interval and three stars (***) a 99.9% confidence interval.

3.4.2 Linkage analysis

Linkage analysis was performed using the software R version 2.2.1, free software for statistical computing. Together with the software the package R/qtl was used [Broman et al. 2003] which is designed for linkage analysis of F_2 generations in experimental crosses.

The software calculates the genetic map dependent of the marker order that is put in and the amount of recombination. It is also possible to investigate the marker order by using a sliding window with a likelihood method. Likelihood analysis can also be used to assume genotyping errors. Individuals indicated with genotyping errors for certain markers can thereafter be checked on the gels and corrected in the file if it is typed wrong, or be excluded if it is unclear whether it is wrong or not. However, unlikely recombinations can occur and will therefore be indicated as likely genotyping errors. These can be valuable for the analysis and it is therefore important not to exclude individuals based on the LOD for genotyping error but to check the certain individual and marker again.

As an indication of linkage the so called LOD score is regularly used. LOD stands for logarithm of odds, and the odds is the ratio between the likelihood that a locus is linked and the likelihood that it is not linked to the observed phenotype (equation 7). LOD

scores are always calculated with logarithms using the power of 10. The LOD score thereby in the simplest form looks as follows:

$$\mathrm{LOD} = \mathrm{LOG}_{10} \frac{\mathrm{L\,(locus\,is\,linked)}}{\mathrm{L\,(locus\,is\,unlinked)}}$$

Equation 7

The genome is scanned for QTLs and there are many methods using different algorithms to choose between. These algorithms are different methods to estimate the likelihoods used to calculate the LOD according to the equation above. These include marker regression, Haley-Knott regression, expectation maximization (EM) algorithm and multiple imputation. Marker regression calculates the linkage only at the selected markers, while Haley-Knot regression [Haley, Knott 1992] use flanking markers for calculations between the genotyped markers. The EM algorithm [Dempster et al. 1977] uses maximum likelihood estimation applied on intercrosses for interval mapping [Lander & Botstein 1989]. In the imputation method a Monte Carlo algorithm is used for genome wide scans [Sen et al. 2001].

The LOD scores for the different phenotypes are calculated with the chosen algorithm and plotted in diagrams, and the LOD scores are also given in numerical format. LOD threshold has also been calculated with 95% confidence interval using 1000 permutations. The threshold value for 95 % confidence interval is included in the graphs as a dotted line.

Multiple genotypes, underlying the observed phenotype, can also be calculated by the software. This is done by using a Hidden Markov Model, HMM, which generates interaction plots where likelihood for two genotypes underlying an observed phenotype is seen. So called effect plots generates the expression value for chosen phenotypes at a specific marker and can also be generated for interacting QTLs.

R/qtl is programmed for F2 crosses. One effect of that is that the estimation of the map is based on the assumed amount of recombination in an F2 cross. The dense amount of recombination in the F_{12} AIL is therefore interpreted as that there is longer distance between the markers than what actually is the case. However the calculated relative distance between them is accurate. Therefore, the graphs here are presented without information about estimated position of the markers. All markers are presented in appendix 1 with their genomic position.

4 Results

4.1 Evaluation of Housekeeping genes

The analysis of the PCR products showed an interesting difference when analyzed with an agarose gel with ethidium bromide compared to a silver stained TBE gel. The regularly used method for this test has been EtBr and according to that method the products were pure with no clear smear or extra bands around the actual product. This was true for all targets and housekeeping primers. However, the silver staining method is here shown to be much more sensitive and detects extra bands for many of the products as can be seen in Figure 4.1.

Both for C3 and C1q extra bands that were detected on the silver stained gel were not seen on the EtBr gel. The extra band on C3 is clear but undetectable on EtBr, and therefore likely to have much lower concentration than the correctly amplified product. All extra bands detected are longer molecules and may therefore be at least partly due to contamination of genomic DNA. Another possibility is that the unspecific products are amplifications due to unspecific binding of primers. Two different primers for 18s was designed, one spanning an intron-exon border. At the silver staining gel, this was shown to increase the specificity of the primers and that primer could be established as more specific. No amplification of longer products could be detected when studying the melting curves from RT-PCR analysis and therefore these probably affect the result very little, due to the low concentration.

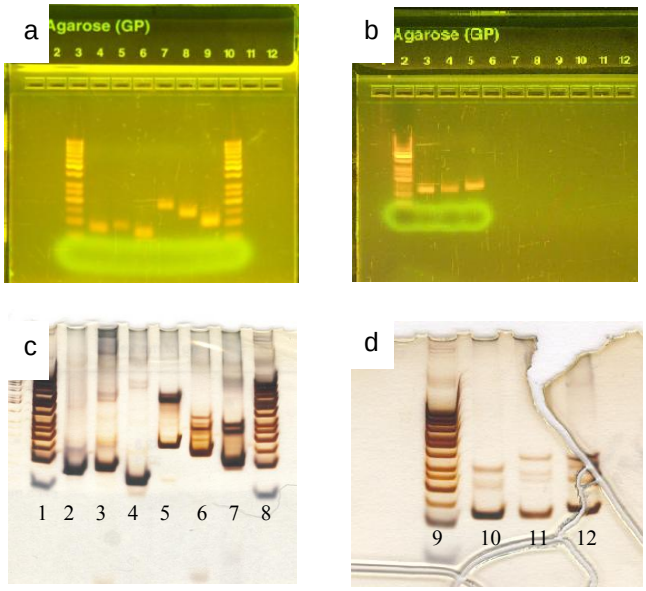


Figure 4.1: Photos from the gels used for test of primers. The upper panel is from the Agarose gels containing EtBr and the lower panel show silver stained gels. *a,c* PCR products from the following primers, starting from left: Ladder (1), GAPDH (2), ß-actin (3), CD74 (4), C3 (5), C1q (6), HMBG1 (7), Ladder (8). Photos *b,d* show ladder (9), HPRT (10), 18s spanning an intron-exon (11) and finally a second 18s primer without intron-exon span (12). The size of the ladder is 50 bp. (gel d broke whilst drying and did therefore not affect the result)

The RT-PCR run for the housekeeping gene ß-actin was problematic. First of all the melting curve for the standard and the unknowns did not show the same temperature with an increase of 4°C for the melting curve (see Figure 4.2). Second, studies of the expression of *CD74* in the parentals showed more spread within the groups compared to *18s* and *GAPDH*. Problems with ß-actin had been seen for more studies within the lab, and for all these reasons ß-actin was not used as HKG in the analysis.

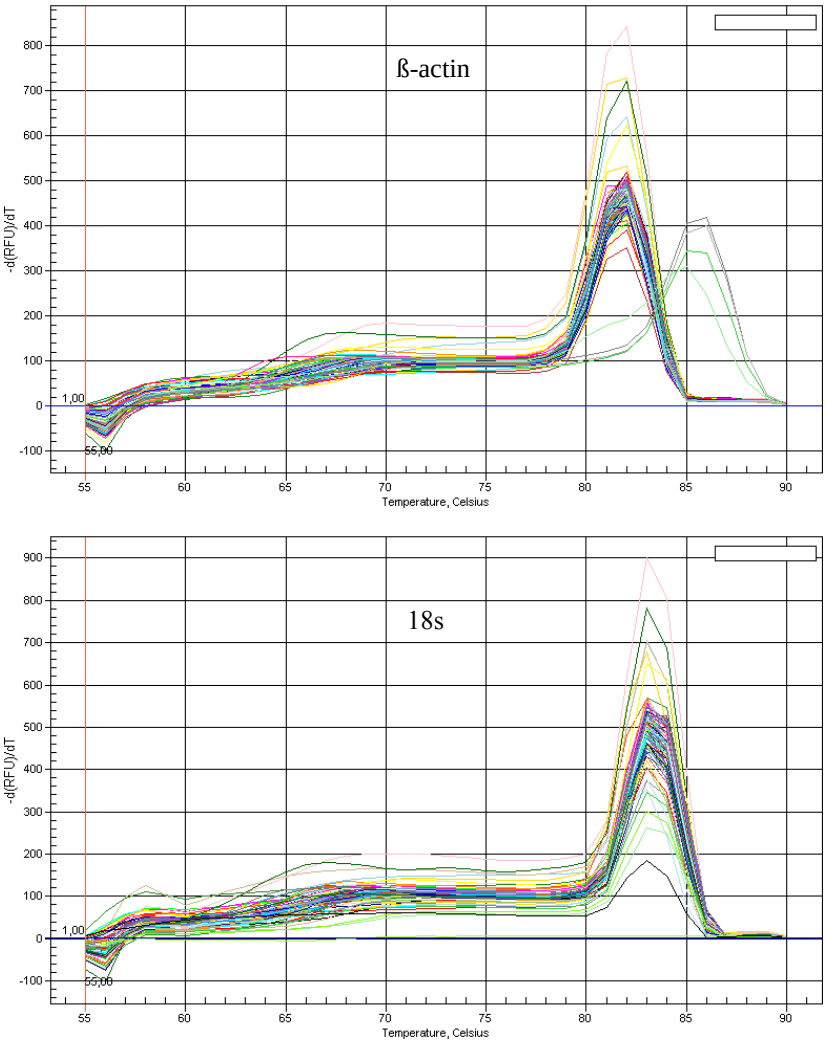


Figure 4.2: Melt peak curve charts from the same plate for the ß-actin (top) and 18s (lower) runs, showing the change in fluorescence at different temperatures. Some curves in ß-actin (left) show a shift of around 5°C with higher melting temperature. All these are samples from the standard curve samples. Increased temperature was not seen for any of the unknowns.

All three HKGs 18s, HPRT and GAPDH gave rather similar results when studied in the parentals (14 PVG.1AV1 and 11 DA rats). However, the final choice was restricted to the use of 18s and GAPDH. The choice was mainly based on calculations of the coefficient of variation for each group. This was used as a measure of how well the groups cluster using the different HKGs. For DA this coefficient was 28.76% using HPRT, 24.57% GAPDH, 24.15% 18s and 26.93% with ß-actin. The variation within the PVG.1AV1

group was (with one outlier excluded) 28.51% HPRT, 25.10% GAPDH, 24.85% 18s, 27.72% ß-actin. These results indicated that 18s and GAPDH was the most appropriate HKGs. With the hypothesis that individuals of the pure parental strains should have similar response to the injury, better clustering of the groups reasonably could be due to more stable expression of the HKG. When both GAPDH and 18s were used the coefficient of variation got ever lower (23.76 for the DA and 23.83 for the PVG.1AV1 group) than when calculations were made with the HKGs separately, indicating an advantage of using multiple HKGs in expression analysis.

Another reason for not including HPRT in the final analysis was that there was a problem with a low contamination in the negative controls which might affect the result. This problem could probably easily have been eliminated by rerunning the plates but this was not done because of lack of time. The parental plates were run at another time point and within this plate there was no problem with contamination and the results used in the parental comparison is therefore unaffected by this. However, linkage analysis was run with phenotype results for all the three HKGs to evaluate if better clustering within parental groups also had a positive effect on the LOD score.

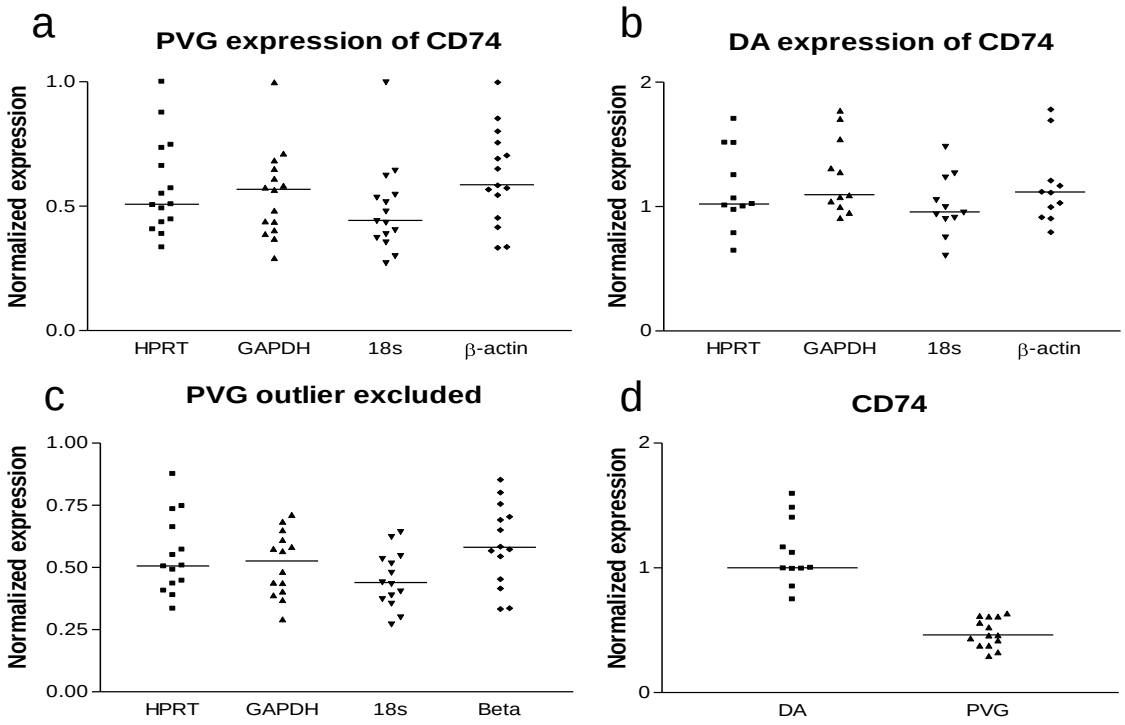


Figure 4.3: The expression of CD74 in the parentals calculated with the different HKGs tested and with multiple HKGs. In PVG.1AV1 there is one individual that is a clear outlier and in diagram a. In diagram c and d that individual is excluded. In Figure d, results using both 18s and GAPDH are presented. GAPDH and 18s together gave lower coefficient of variation than all other possible combinations of HKGs (data not shown).

4.2 Strain differences of studied phenotypes

The strain differences between the phenotypes were studied by measuring mRNA expression values from the parentals calculated with GAPDH and 18s as HKGs, Figure 4.4. All phenotypes showed significant difference, with Mann-Whitney U test, with a P value below 0.001. For all three phenotypes the DA rat is the high responder to injury. These results correspond well with the difference in nerve cell death that is detected between DA and PVG.1AV1 which shows an equally significant difference [Lundberg et al. 2001]

During the experiment the weight loss at 5 days after VRA was also measured. For weight loss there was an equally significant difference between the strains with the PVG.1AV1 being almost unaffected with a median weight loss of 0.5 grams compared with the DA rat with a median weight loss of 10 grams.

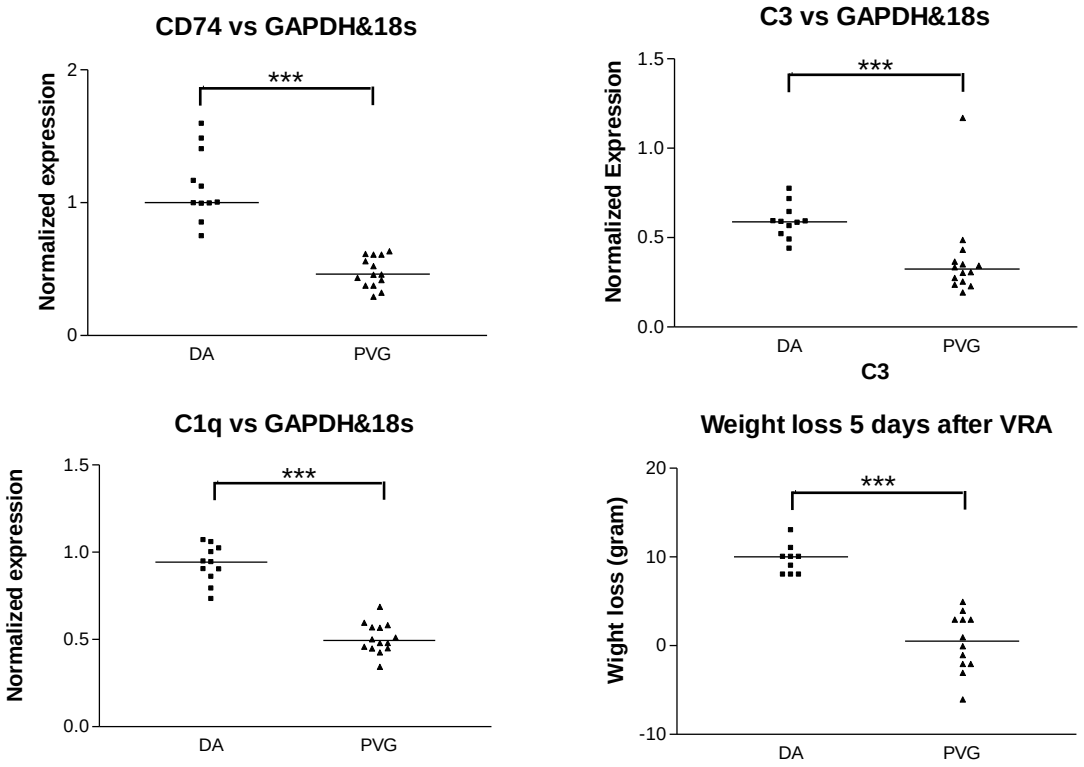


Figure 4.4: All studied phenotypes displayed a statistically significant difference in mRNA expression between the stains. For all phenotypes the P-value is below 0.001 using a Mann-Whitney U test. All groups also give equally significant difference using the t test, except from C3 that only get a P-value below 0.05, showing the great impact of outliers in a parametric test. Median values for DA/PVG.1AV1: CD74 1/0.4268, C3 0.5879/0.3426, C1q 0.9432/0.4940, Weight loss 10/0.5. All values calculated with multiple HKGs using 18s and GAPDH.

Observe that calculations of normalized expression presented in the graphs are made with relative quantities calculated relative to one and same individual (reference sample) in each phenotype group separately. This means that all values are relative to the mRNA expression in that reference sample of the particular molecule studied. Therefore, comparisons of the normalized expression between the different phenotype groups can not be made.

4.3 Linkage analysis

Of the 163 animals in the AIL, totally 156 were used in the linkage analysis. The 9 individuals that were not analyzed failed in either dissection of the spinal cord or mRNA/cDNA preparation. On average for all markers 94.8% of the individuals were successfully genotyped (Figure 4.5) for the 39 analyzed markers. The rather poor outcome of the markers on chromosome 4, *D4Rat78*, and 8, *D8Rat76*, is responsible for many of the missing genotypes, lowering the percentage down to 94.8%. However, no indication of linkage was seen for the marker on chromosome 4 and therefore it was not genotyped again. *D8Rat76* would ideally be tested again but was not done due the time limit. Dropping these markers increases the percentage genotyped to 95.9%.

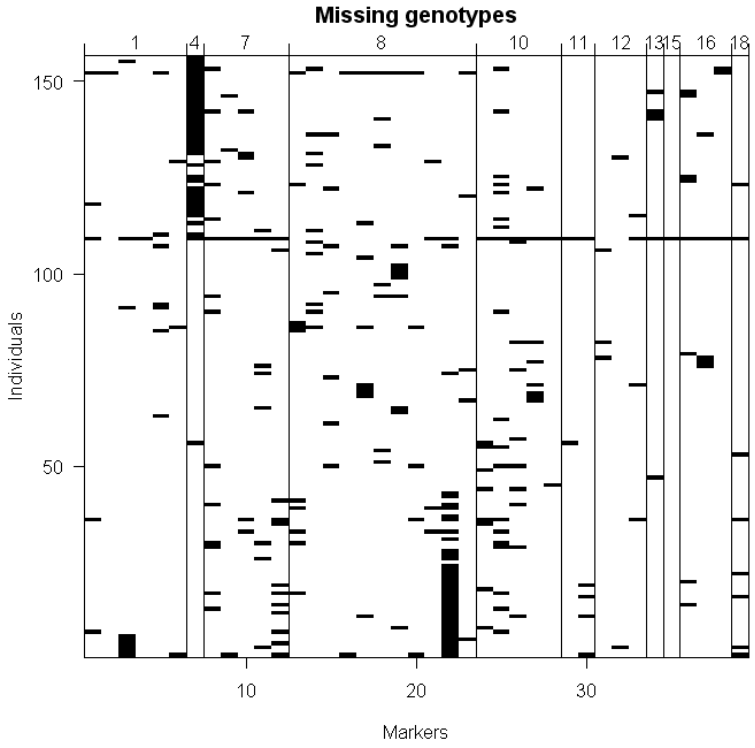


Figure 4.5: Plot indicating the missing genotypes for each marker separately, returned by R/qtl.

As mentioned previously, the software calculates the distance between the markers based on the frequency of recombination assuming an F_2 cross is studied. Figure 4.6 presents a graphic illustration of the genetic map with the calculated distances. The real distance is approximately one fifth of the calculated, based on the distance noted at the databases used to identify the markers. On chromosome 8, markers in two regions of interest have been genotyped, one with a very long distance to next marker. All markers except from the last one are within the *Vra1* region while the other was added on basis of single marker regression in BNxLEW.1N and this result in a rather long distance between two markers (*D8Rat76* and *D8Rat2*).

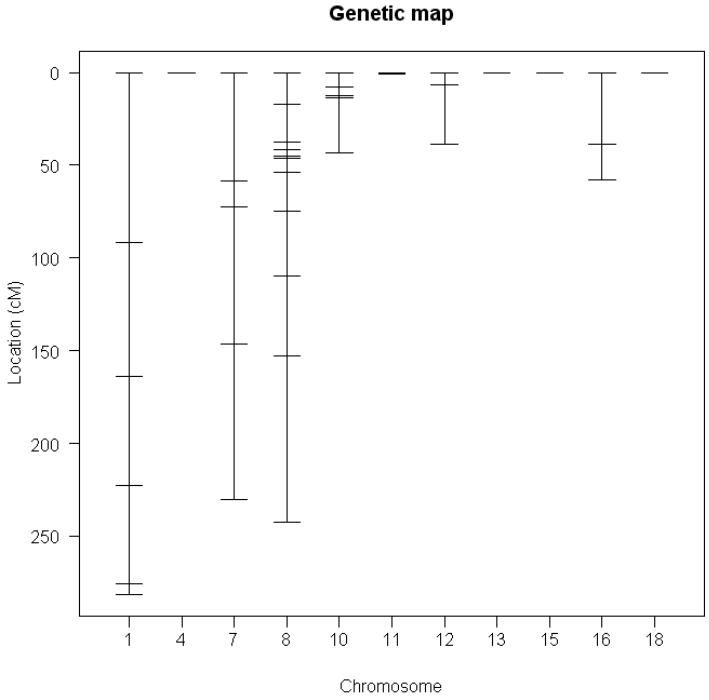


Figure 4.6: Genetic map calculated by the software. Each line indicates a genotyped marker. The distance is relative to the first marker genotyped on the chromosome. This does not mean that it is located at the beginning of the chromosomes. Calculations are based on an F_2 cross and therefore the indicated distance in cM is far too long. The correct distance is approximately one fifth of the calculated.

Linkage analysis was performed for the whole set of analyzed markers using the multiple imputation method and calculation of the level for 0.05 significance threshold was performed using 1000 permutations and plotted as a dotted line. This established a significant linkage of CD74 to the max marker for *Mhc2ta*, *D10Mgh25*, on chromosome 10 with LOD 18.48 and a linkage to C1q on chromosome 8 close to the *Vra1* region with a LOD peak of 3.82 (Figure 4.7)

The linkage analysis did however not verify many of the peaks from the BNxLEW.1N study. On chromosome 1 there are small peaks for both C3 and CD74 that however not reach the 95% significance threshold. *D1Rat234* shows a LOD of 1.585 and D1Rat94 2.16 for CD74. The C1q linkage on chromosome 8 is not within a region selected on basis of findings in the BNxLEW.1N cross, but close to the peak for the previously detected *Vra1* locus (peak around D8Rat75) for nerve cell death. Linkage for C1q has previously not been tested for this region. For the other markers on chromosomes 4, 7, 11, 12, 13, 15, 16, 18 no confirmed or suggestive linkage could be established for any of the studied phenotypes. The regions with most interesting results, chromosomes 1, 8 and 10 are presented with separate more detailed figures (Figure 4.8, 4.9, 4.10)

Linkage to weight loss was also analyzed (data not shown) but no significant peaks were detected for any of the regions tested. There was an indication at a position on chromosome 6 in the other cross. However, that marker was not polymorphic between DA and PVG.1AV1 and efforts to test other close markers was unsuccessful.

Test for 95 and 90 percent significance threshold using 1000 permutations resulted in threshold values for CD74 at 3.02 (95%), 2.65 (90%), for C3 at 3.09 (95%), 2.70 (90%) and finally C1q at 2.97 (95%), 2.71 (90%). A peak with LOD above the 95% threshold is considered to be confirmed and a peak above the 90% is considered suggestive.

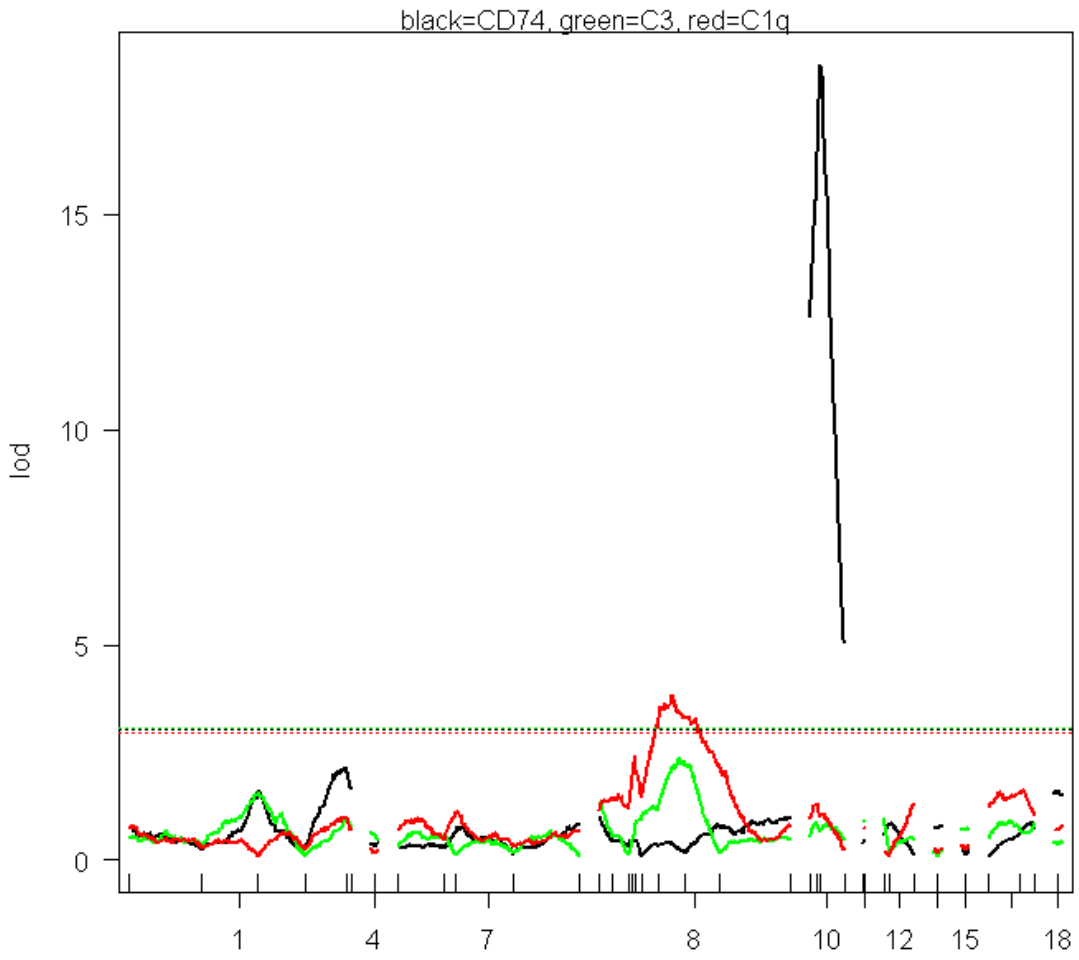


Figure 4.7: Result from linkage analysis using GAPD and 18s as HKGs. The dotted lines show the result from 1000 permutations for 95% significance, for CD74 3.02, C3 3.09 and C1q 2.97. The numbers on the x-axis represent the chromosomes, and each vertical line is a genotyped marker.

Chromosome 1

On chromosome 1 there is a tendency of linkage that however not reaches above the 0.05 significance threshold. There are two small peaks for CD74 on markers D1Rat234, LOD 1.67 and D1Rat94 with LOD 2.13. According to the plots in Figure 4.8, the peak at D1Rat234 the lowest expression is seen in heterozygotes and the difference between the DA and PVG.1AV1 allele is not very large. For the second peak the DA homozygotes show clearly higher expression than the PVG.1AV1 homozygotes, indicating that this region might actually be involved in up regulation. For C3 there is an overlapping peak on D1Rat234 with LOD 1.53. The pattern between the alleles is very similar to that for CD74 (graph not shown) with small differences between the different homozygotes.

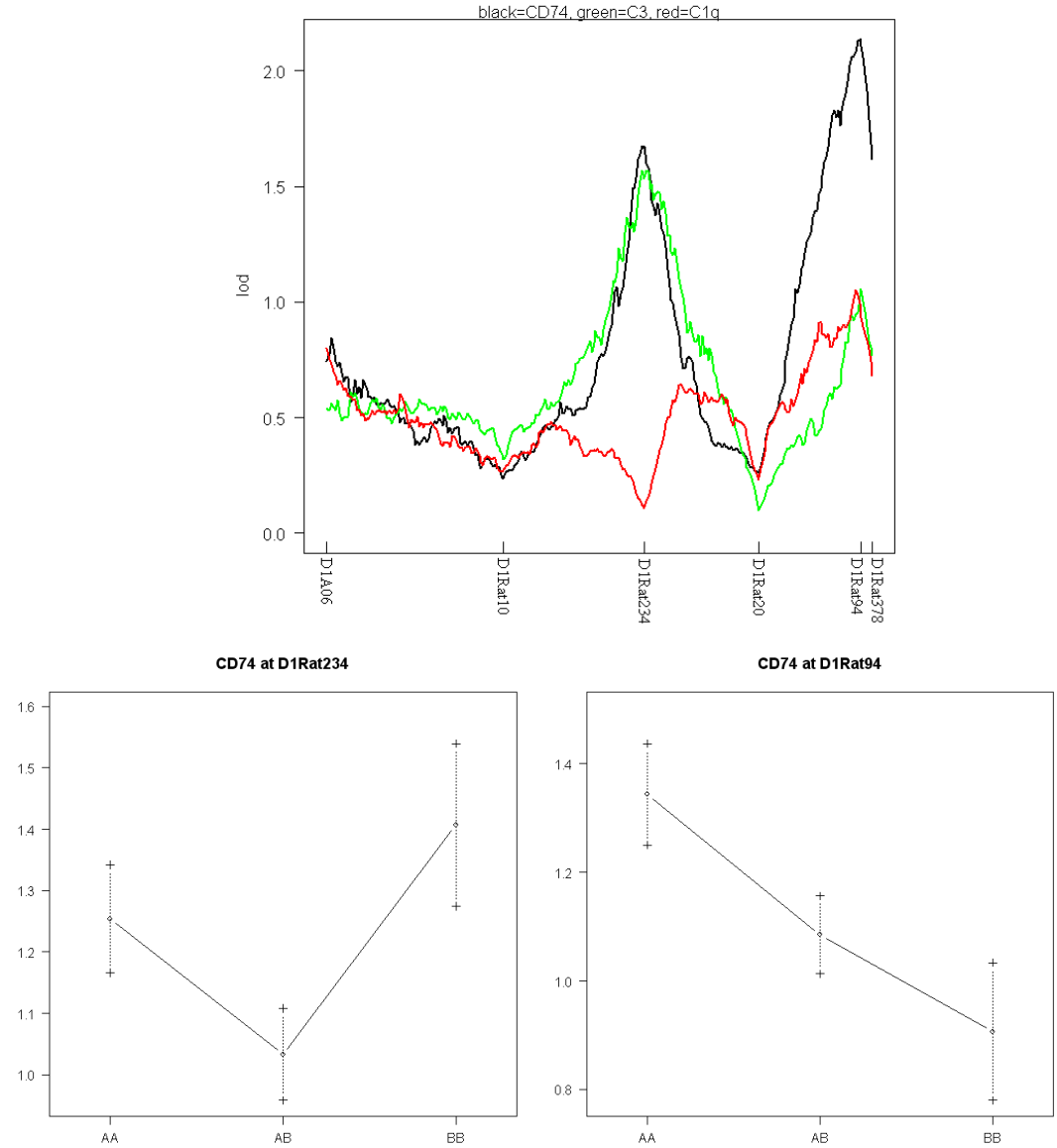


Figure 4.8: Top graph presenting the linkage analysis result from chromosome 1 and below two graphs showing the allelic effect at the two suggestive markers for CD74. Only the second peak can be a region contributing to the difference seen between DA and PVG.1AV1. D1rat234 only seem to have a lowering effect in heterozygotes. With a LOD as low as 1.67 this heterozygote effect might most reasonably be considered as noise. A=DA and B=PVG.1AV1

Chromosome 8

On chromosome 8 a previously undetected peak for C1q expression was identified with maximum LOD score at markers *D8rat129* (LOD 3.07) and on *D8rat14* (LOD 3.30). This is both above the threshold for significant linkage. However the peak is calculated to be between with maximum LOD of 3.82. The distance between these markers are approximately 6 cM and most likely the actual peak lies in between with a possibly higher LOD score than the calculated. There is also a tendency of linkage to C3 at marker *D8Rat14* with LOD 2.21 (peak calculated upstream with maximum LOD 2.36) but not even over 90% threshold.

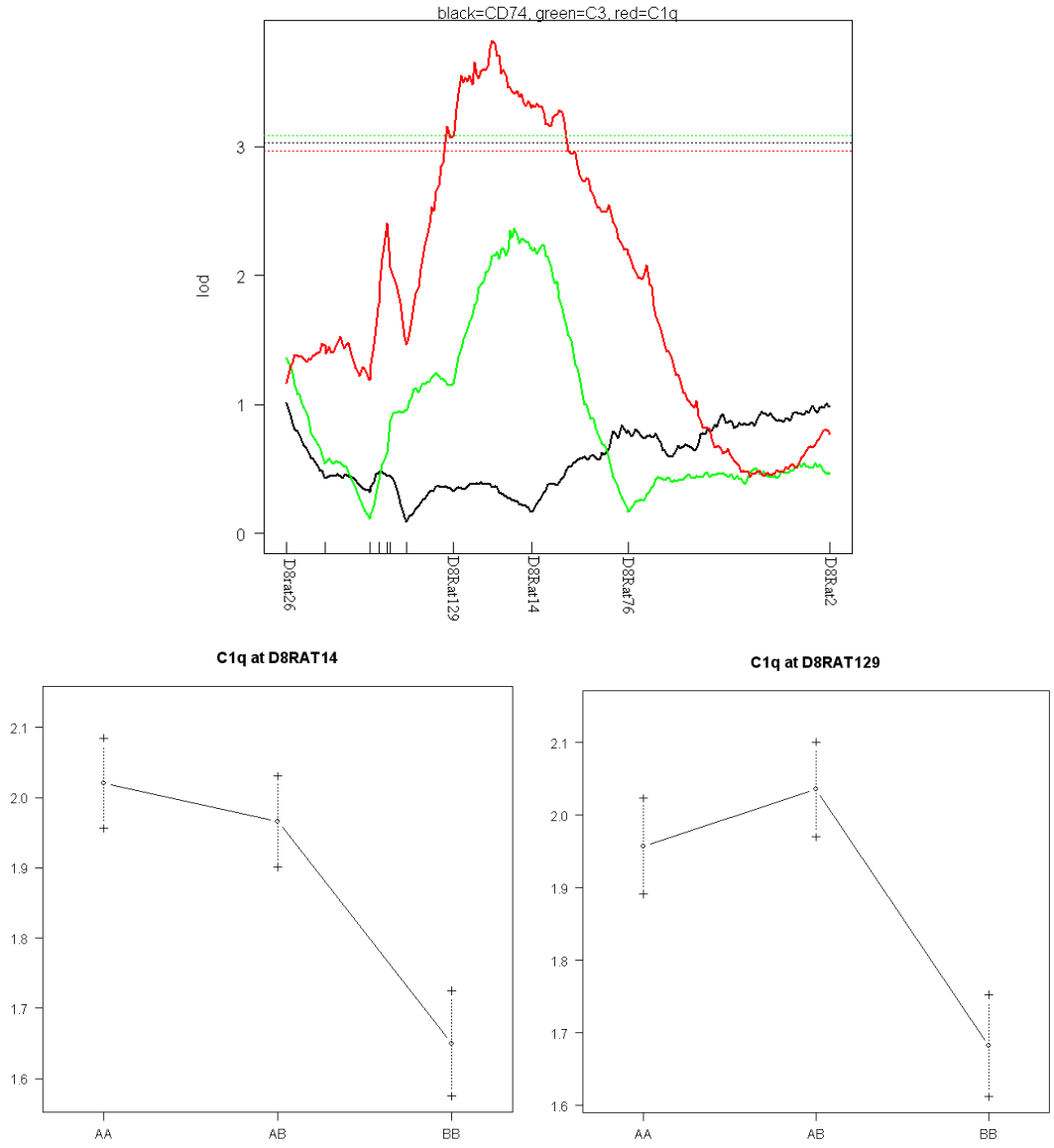


Figure 4.9: Top, the pattern on chromosome 8 with significant linkage for C1q. Bottom, plots of the allelic effects on the two markers with highest LOD for C1q. As seen clearly on the top graph there is no actual peak but more like a plateau. It is therefore likely that a peak can be found if a couple of markers are added between *D8Rat129* and *D8Rat14*. PVG homozygotes have much lower expression but there is only a small difference between DA homozygotes and the heterozygote individuals.

Chromosome10

An effect on MHC class II expression by *Mhc2ta* 5 days after VRA is not surprising, but with the result from chromosome 10 analyses the importance of this gene is definitely established. The LOD score for CD74 on *D10Mgh25* get as high as 18.48. Interestingly the *D10Mgh25* DA allele homozygotes that have the highest expression are those who also are DA allele homozygotes also for the detected marker on chromosome 1. Similarly the PVG.1AV1 homozygotes at *D10Mgh25* that have lowest expression of CD74 are also PVG.1AV1 homozygotes on *D1Rat94* (Figure 4.10), indicating that there might be some relevance of the suggestive peak on chromosome 1.

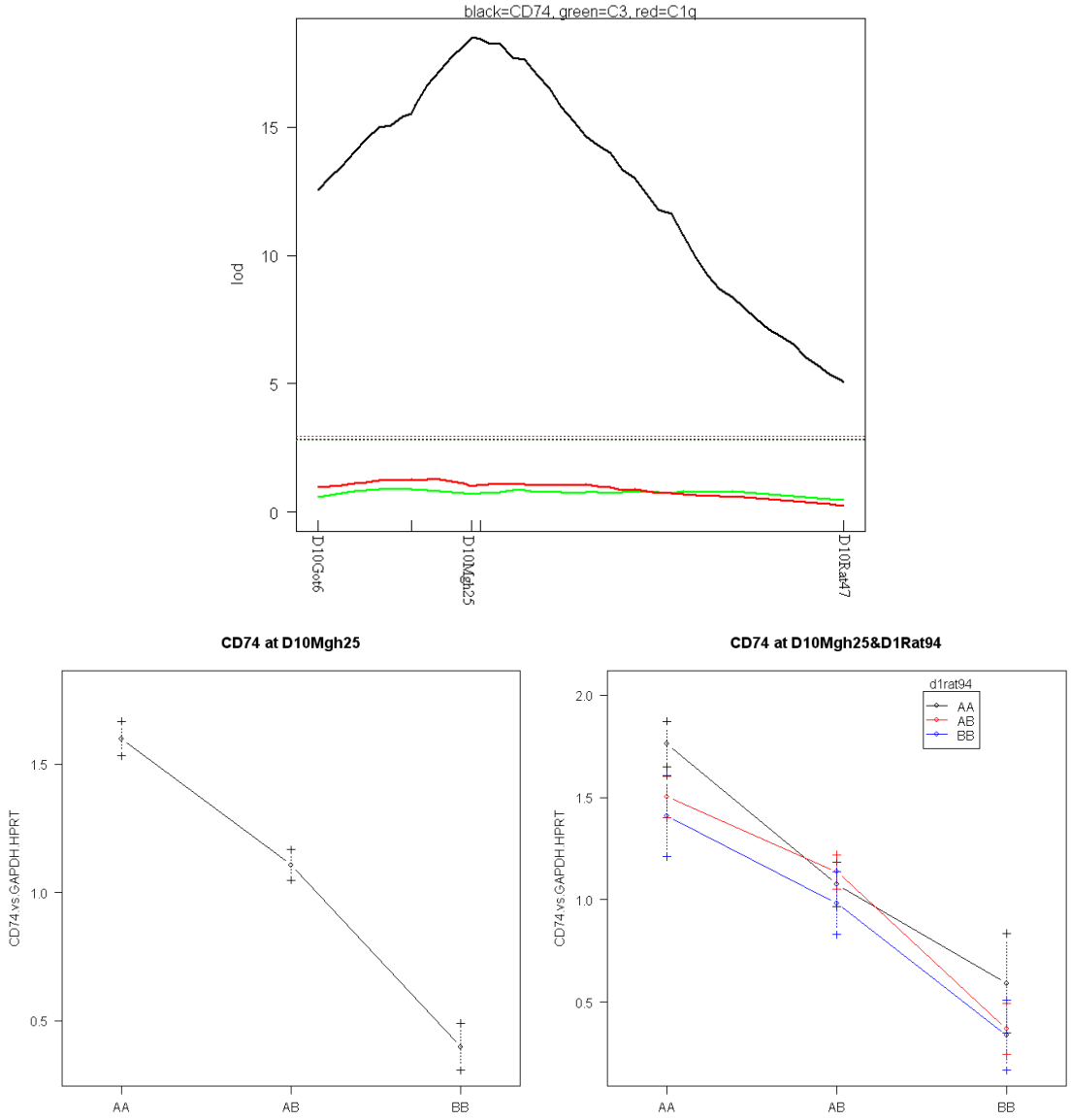


Figure 4.10: Result from linkage analysis with *D10Mgh25* as max marker. Dotted lines indicate the same threshold established with 1000 permutations as presented for the whole set of markers. The lower plots indicate the relative expression of CD74 for individuals DA homozygotes (AA), heterozygotes (AB) and PVG.1AV1 homozygotes (BB) at the position for the *Mhc2ta* marker (*D10Mgh25*). In the plot to the right, the individuals are grouped depending on the genotype at *D1Rat94*, which showed suggestive linkage on to CD74 expression on chromosome 1.

4.4 HKG influence on LOD score

The choice of HKG was shown to have an effect on the linkage analysis, when analysis was performed with expression values calculated separately with the HKGs HPRT, 18s and GAPDH. Graphs for all these calculations are presented in Appendix 2. LOD scores for CD74 on *D10MGH25* for GAPDH, HPRT and 18s respectively was calculated to 17.11, 14.64 and 16.95. Interestingly, when using multiple HKGs as for all results presented above (GAPDH and18s) the LOD increased to 18.48, higher than for both HKGs separately.

For the second confirmed peak, C1q at *D8Rat14*, where the difference between the calculations was bigger the highest LOD was achieved with HPRT. It was close to the same using 18s, but rather much lower with GAPDH. The final result using both GAPDH and 18s ended close to, although lower than, the highest LOD using 18s separately. The suggestive linkage for C3 at the same marker is not affected in the same way and show very stable results.

The results from the linkage analyses using different HKGs for expression calculations show that both 18s and GAPDH give a very similar pattern. This is also true also for HPRT, although it does not give as high linkage for e.g. *D10Mgh25 (Mhc2ta)*. Results from the clustering in the parental groups indicate that 18s and GAPDH are the best HGKs which also turned out in the linkage analysis with respect to LOD of *Mhc2ta*.

5 Discussion

5.1 Methodology

The aim of this study has been to map QTLs regulating the expression of MHC class II and complement components in the nervous system after a standardized nerve injury in intercrosses of inbred rat strains. These types of QTLs can be referred to as expression QTLs, eQTLs. The strategy has been to use data from a whole genome scan in an F_2 cross between BN and LEW.1N that differs considerably in the expression of the studied transcripts. Regions of interest that had been detected between these strains has in this second step been studied in a previously collected F_{12} DAxPVG.1AV1 and provided the possibility of testing the degree of conservation of QTLs between the different strain combinations as well as fine mapping of conserved QTLs. The outcome of the F_2 BNxLEW.1N study turned out to demonstrate linkage to multiple regions. Unfortunately only one of these, on chromosome 1, and only providing suggestive linkage, was reproduced in this F_{12} DAxPVG.1AV1 study. There are several potential explanations for this, and one likely cause is that DA and PVG.1AV1 does not differ at the same genomic regions that revealed linkage in the other cross. Thus, it is likely that DA and PVG.1AV1, the strains that differ the most in inflammatory response, contain additional gene regions that regulate the studied phenotypes.

When this study was initiated and the regions were selected, selection was based on single marker regressions and preliminary results. With the progress of the BNxLEW.1N study it is apparent that the responses are polygenic with regard to all phenotypes. There are many peaks with modest LOD scores observed and no single dominating peak as with the case of *Mhc2ta*. To identify gene regions with modest impact on MHC class II expression between DA and PVG.1AV1 might be difficult due to the dominating impact of the *Mhc2ta* also 5 days after VRA. One possible solution would be to study only individuals that are heterozygotes at the *Mhc2ta* locus. A linkage analysis with these subjects was performed that did not result in any new regions. However, the number of subjects was then reduced to fewer than 80. This is far too low to find a reliable and significant linkage. The number of individuals in the study should probably have needed to be almost the double to have the necessary statistical power.

More surprising is that none of the linkages for the complement components was reproduced in DAxPVG.1AV1. There is an identified linkage for C1q on chromosome 8, but that did not show up in BNxLEW.1N and the region was chosen on basis of the earlier findings of the *VRA1* locus for nerve cell death. It seems unlikely that none of the regions relevant for the difference between BN and LEW.1N should affect the variation between DA and PVG.1AV1. One explanation could be that there might be a stronger, dominating genetic effect underlying differences between DA and PVG.1AV1, but where alleles between BN and LEW.1N are identical. It also raises a question if the step between F_2 and F_{12} is too large to define more modest, although significant, peaks

directly at F_{12} based on the peak marker for F_2. In F_2 the distance between studied markers has been approximately 10cM. If the QTL in fact is a few cM away from the indicated peak it might not be seen at all in F_{12} if there are many small peaks with low LOD scores underlying the phenotypic difference. Maybe it was not enough only to check the peak marker from F_2 in F_{12} and that a larger region always should have been covered. An answer to that question could possibly have been achieved by adding at least one marker extra on each side of the indicated peak markers on the chromosomes where only a single marker was checked.

The use of multiple HKGs turned out to increase the LOD score for *Mhc2ta*. The overall pattern was very similar with the three HKGs (18s, GAPDH and HPRT), indicating that the expression of each of them is not affected to any large degree by the VRA lesion. The reason for the increase in LOD is not necessarily due to the use of different HKG but possibly also is because this means that each sample has 4 replicates instead of two for the HKGs. It is possible that the same increase would have been seen if one and the same HKG had been run twice. Definitely, it here is an advantage to use both 18s and GAPDH also further on. However, because of the limited amount of cDNA that can be extracted from the L3 segment of the cord, it may not be worthwhile to run multiple HKGs when we established that both 18s and GAPDH give similar results. A compromise may be to run one of them in triplicate instead of duplicate. This would likely increase the quality of the important housekeeping results, while consuming less sample, time and money than running multiple HKGs.

The choice of HKG mainly affected the modest peaks, where tendency to peaks both appeared and disappeared with choice of HKG. This might say more about the reliability of small suggestive peaks than importance of HKG choice. However, multiple HKGs will most likely lower the number of false positive linkage results, and save the scientist from the thought of adding more markers around it in order to possibly find a confirmed linkage.

What turned out to be more important than the choice of HKG, was the sample selected as an inter run calibrator. Here, one sample in the standard curve was chosen. However, this may not be an ideal solution. Sometimes samples in the standard curve are excluded to provide a better fit of the standard curve. This can not be done on any of the samples used as inter run calibrator, because that will result in only one sample left for this calculation. Therefore as a result, one has to accept a less well-performed curve fit in order to save both calibrator samples. Ideally in the future, a large batch of samples with cells from many individuals and different tissues should be made, and this same calibrator used for all targets and by all members working in the VRA project. This will make it possible to compare all possible plates within a gene study, even plates made by separate workers.

The initial choice was to use of booth EtBr gels to check the PCR product. However, because the amplicon was hard to detect with this technique silverstaining was performed in addition. The problem with the EtBr gels was solved by increasing the amount of sample added by more than the double, and therefore this technique consume a lot more

PCR product than the silverstaining gel. Furthermore, the silverstaining method also turned out to be so sensitive that extra bands, that could not be detected by EtBr or melting curves on the iQ5, was seen. These quantities are probably extremely low, but may have a small effect on the expression result, especially when the targets are expressed in low amounts. This extra sensitivity made it possible to determine that the primer for 18s that spanned an intron-exon border was more specific than the other primers that did not. Therefore I find silverstaining to be preferable in testing primers, being more sensitive and consuming less sample.

5.2 Identified linkage

It has previously been shown that the difference in expression of MHC class II, 14 days after VRA is due to *Mhc2ta.* Therefore, it is not surprising that it also here turned out to affect expression 5 days after VRA. However, it was not obvious due to the fact the difference in MHC class II expression at the early time point is smaller between DA and PVG.1AV1, having different at the *Mhc2ta* alleles, than between BN and LEW.1N, having the same (DA) *Mhc2ta* allele. With a LOD of 18.48 it can be concluded that this gene also is responsible for the difference between these strains after 5 days. This study only includes a limited number of gene regions detected in an F_2 study in another set of animals. This means that it can not be excluded that also other loci can be involved in the regulation of MHC class II expression, which do not differ between BN and LEW.1N and therefore are not possible to study in the F12 DAxPVG.1AV1. On the other hand, the study with the BNxLEW.1N strain combination showed that there are multiple smaller genetic influences that affect the MHC class II pathway. *Mhc2ta* has been reproduced in human association studies and showed that the effect on disease is mediated through an effect on the MHC class II pathway. There might thus in addition be multiple smaller genetic influences. However, with their modest effect it is likely that these regions will be very difficult to map in detail.

The second finding in this study includes the linkage on chromosome 8 to C1q. Parts of this region were previously genotyped and tested for linkage to C3 expression. However, initial findings could not be reproduced. C1q has never been analyzed before and this linkage has therefore not been seen before. However, the peak is probably not detected here, because two markers with a distance of about 6 Mb have approximately the same LOD score. It is likely that the actual peak lies between these two markers and possibly has a higher LOD score. It would be interesting to add more markers into this region to fine map it and find the peak marker. This could possibly reduce the interval so a low number of candidate genes can be identified. Interestingly, linkage to EAE has been mapped to this region both in F_2 PVGxLEW [Becanovic et al. 2003] and F_7 DAxPVG [Nohra, Becanovic, Olsson et al, unpublished observation].

None of the tested regions showed any significant linkage to weight loss. However, an indicated region on chromosome 6 was found recently in the BNxLEW.1N study. That marker was not polymorphic between DA and PVG.1AV1 and neither was a close

marker that was tested. Unfortunately, there was not time enough left to search for and test other markers during this project. However, due to the clear difference between the parental individuals there is definitely worth to try to continue with searching for a weight loss QTL.

5.3 Conclusion and perspectives

The difference seen in mRNA expression of MHC class II between BN and LEW.1N in the early response after VRA turned out to be of polygenic nature. Thus, there is no single dominating gene responsible for this difference as in the case for the *Mhc2ta* between DA and PVG.1AV1. It is therefore not surprising that these peaks do not show up in the DAxPVG.1AV1 when *Mhc2ta* in this intercross was shown to have a large impact also at this time point. Regions responsible for the phenotypic difference between BN and LEW.1N may not differ between DA and PVG.1AV1. For one to identify the genes that underlie the linkage, it will probably be necessary to set up an AIL in the BNxLEW.1N strains and study the difference within these allelic combinations. This will also enable fine mapping of the regions identified for the other studied phenotypes.

Another possible extension of the work on the complement expression would be to set up an F_2 between DA and PVG.1AV1 to study the response at 5 days after VRA. An F_2 generation with samples collected it this time point has not been previously established. With a whole genome scan in that material, other regions with relevance to C3 and C1q expression could be detected and thereafter fine mapped in this F_{12} material. Because the mRNA for phenotyping is a limiting factor it is also important to consider which other phenotypes are of interest to such a study. When expression has been studied by RT-PCR, more genotypes can thereafter be added in the future. So if new regions will be detected in a future F_2 study with relevance to *CD74*, *C3* or *C1q* expression, it will be easy to fine map these in this F_{12} material using the already achieved expression results and add just the new markers for the genotyping.

Other molecules that might be of importance to study with relevance to the inflammatory process are different co-stimulatory molecules like B7-1 and B7-2. As mentioned earlier, weight loss also is a parameter well worth more effort. The focus here has been to study the microglial activation. It could also be of interest to add molecules involved in the process of astrocytic activation.

In summary, two independent strain combinations have been used to analyze expression of key molecules in the innate immune response of the central nervous system. The results show that RT-PCR is a powerful technique to characterize eQTLs, but also that the methodology is sensitive to parameters such as choice of housekeeping genes and internal calibrator samples. In F_{12} (DAxPVG.1AV1) differences in MHC class II expression are large due to one single strong genetic effect of the *Mhc2ta.* In contrast, the difference between BN and LEW.1N are caused by multiple smaller genetic influences that cannot be confirmed in the F_{12} (DAxPVG.1AV1) intercross. Likewise, differences in

the complement molecules C3 and C1q expressions are regulated by multiple genetic influences that mostly not are conserved between the strains. In the F_{12} DAxPVG.1AV1 intercross linkage to C1q expression was demonstrated at a region previously associated to neuronal cell death and clinical disease in the VRA and EAE model respectively. The results obtained suggest that differences in inflammatory response are due to many smaller, and a few stronger, genetic influences. Probably this resembles the genetic make up of human complex diseases, where many small genetic effects contribute to disease just as well as a few stronger influences.

5.4 Primer design error

After this project was finalized an error in the 18s design was revealed. By mistake I had used the sequence for the mitochondrial ribosomal subunit instead of the regular ribosomal subunit when designing 18s primers. Therefore, the HKG referred to as 18s is actually the mitochondrial ribosomal protein S18A. This is not a gene regularly used as HKG and therefore, no conclusions about the reliability of 18s as HKG should be made from the results obtained here. However, according to the results the expression of the gene used here (S18A) seems to be stable in this model and gave similar and as good results as GAPDH with respect to strain variation, and LOD scores. Therefore, I find the observation that multiple HKGs are preferable not to be affected by this mistake. Linkage analysis has also been performed with expression values calculated with all HKGs separately and found to vary to a low extent. I therefore expect the results to be very little affected by rerunning with the real HKG 18s, a step that nevertheless has to be carried out in order to achieve results that can be published.

Acknowledgements

It has been ha pleasure to work together with all people at the neuroimmunology unit at Centre for Molecular Medicine. I am very grateful to Fredrik Piehl for giving me the opportunity to do this project in his group and for planning the project. Further, I would very much like to thank Margarita Diez for supervising, describing all used techniques and patience with all asked questions. Olle Lidman has also been very helpful with answers to frequently asked questions and suggestions to improvements of the report. Finally, thank you to all other members of the VRA group that all have been helping me with the progression of this project.

References

Aldskogius H, Liu L, Svensson M. Glial Responses to Synaptic Damage and Plasticity. Journal of Neuroscience Research 1999;58:33–41.

Becanovic K, Wallstrom, Kornek B, Glaser A, Broman K W, Dahlman I, Olofsson P, Holmdahl R, Luthman H, Lassmann H, Olsson T. New Loci Regulating Rat Myelin Oligodendrocyte Glycoprotein-Induced Experimental Autoimmune Encephalomyelitis. The Journal of Immunology 2003;170(2):1062.

Block M L, Zecca L, Hong J-S. Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. Nature Reviews Neuroscience 2007;8:57-69.

de Boer A G, Gaillard P J. Blood-brain barrier and dysfunction and recovery. Journal of Neural Transmission. 2006;113:455-462.

Brenner M, Kisseberth W C, Su Y, Besnard F, Messing A. GFAP Promoter Directs Astrocyte-specific Expression in Transgenic Mice. The Journal of Neuroscience, 1994;14(3):1030-1037.

Broman KW, Wu H, Sen Ś, Churchill GA. R/qtl: QTL mapping in experimental crosses. Bioinformatics 2003;19:889-890.

Campbell NA, Reece J B, Biology (6th ed) p1024, ISBN 0-201-75054-6.

Chen J, Rider D A, Ruan R, Identification of Valid Housekeeping Genes and Antioxidant Enzyme Gene Expression Change in the Aging Rat Liver. Journal of Gerontology. 2006;61(1):20–27.

Darvasi A, Soller M. Advanced Intercross Lines, an Experimental Population for Fine Genetic Mapping. Genetics 1995;141:1199-1207.

Dempster A P, Laird N M, Rubin D B. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society. 1977;39(1):1-38.

Haga S, Ikeda K, Sato M, Ishii T. Synthetic Alzheimer amyloid beta/A4 peptides enhance production of C3 component by cultured microglial cells. Brain Res. 1993;601:88–94.

Haley C S, Knott S A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 1992;69(4):315-324.

Janeway et.al., Immunobiology (6th ed) p57, ISBN 0-4430-7310-4
Johansson CB, Momma S, Clarke DL, Risling M, Lendahl U, Frisen J. Identification of a neural stem cell in the adult mammalian central nervous system. Cell 1999;96:25-34.

Kok J B, Roelofs R W, Giesendorf B A, Pennings J L, Waas E T, Feuth T, Swinkels D W, Span P N. Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Laboratory Investigation 2005;85**,** 154-159.

Lander E S, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP Linkage Maps. Genetics 1989;121:185-199.

Lidman O, Swanberg M, Horvath L, Broman K W, Olsson T, Piehl F. Discrete Gene Loci Regulate Neurodegeneration, Lymphocyte Infiltration, and Major Histocompatibility Complex Class II Expression in the CNS. The Journal of Neuroscience. 2003;23(30):9817-9823.

Lidman O. Genetics and Inflammation in Nerve Injury Induced Neurodegeneration, 2003, ISBN 91-7349-654-5.

Lincoln D S. Human genome: End of the beginning. Nature 2004;431:915-916.

Liu F, You S-W, Yao L-P, Liu H-L, Jiao X-Y, Shi M, Zhao Q-B, Ju G. Secondary degeneration reduced by inosine after spinal cord injury in rats. Spinal Cord 2006;44(7):421-6.

Lundberg C, Lidman O, Holmdahl R, Olsson T, Piehl F. Neurodegeneration and glial activation patterns after mechanical nerve injury are differentially regulated by non-MHC genes in congenic inbred rat strains. The Journal of comparative neurology 2001;431:75-87.

Lökk J, Winblad B. Demens – vanlig, svårbehandlad komplikation vid parkinson. Läkartidningen 2006;7:451-455.

Neumann H. Control of glial immune function by neurons. GLIA 2001;36:191-199.

Olsson T, Piehl F, Swanberg M, Lidman O. Genetic dissection of nerodegeneration and CNS inflammation. Journal of the neurological sciences. 2005;233:99-108.

Peinnequin A, Mouret C, Birot O, et al. Rat pro-inflammatory cytokine and cytokine related mRNA quantification by real-time polymerase chain reaction using SYBR green. BMC Immunol [serial online]. 2004;5. Available at:http://www.biomedcentral.com/1471-2172/5/3. Accessed October 15, 2006.

Piehl F, Lundberg C, Khademi M, Bucht A, Dahlman I, Lorentzen J C, Olsson T. Non-MHC gene regulation of nerve root injury induced spinal chord inflammation and neuron death. Journal of Neuroimmunology 1999;101:87-89.

Purves D et al. Neuroscience (3rd ed) p63, ISBN 0-87893-725-0.

Rydberg Björn. « Multiple Skleros (MS) » , internetmedicin.se 2006. http://www.internetmedicin.se/dyn_main.asp?page=717 (2 Jan 2007).

Sandberg R, Yasuda R, Pankratz D G, Carter T A, Del Rio J A, Wodicka L, Mayford M, Lockhart D J, Barlow C**.** Regional and strain-specific gene expression mapping in the adult mouse brain. PNAS 2000;97(20):11038-11043.

Schmittgen T D, Zakrajsek B A. Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. Journal of Biochemical and Biophysical Methods 2000;46:69-81.

Sen S, Churchill G A. A Statistical Framework for Quantitative Trait Mapping. Genetics 2001;159:371-387.

Spassky N, Merkle F T, Flames Nuria, Tramontin A D, García-Verdugo J M, Arturo Alvarez-Buylla. Adult Ependymal Cells Are Postmitotic and Are Derived from Radial Glial Cells during Embryogenesis. The Journal of Neuroscience. 2005;25(1):10-18.

Streit, W J, Microglial senescensce: does the brain´s immune system have an expiration date? Trends in Neuroscience 2006;29(9):506-510.

Swanberg M, Lidman O, Padyukov L, Eriksson P, Åkesson E, Jagodic M, Lobell A, Khademi M, Börjesson O, Lindgren C, Lundman P, Brookes AJ, Kere J, Luthman H, Alfredsson L, Hillert J, Klareskog L, Hamsten A, Piehl F, Olsson T. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nature Genetics 2005;37(5):486-494.

Swanberg M, Duvefelt K, Diez M, Hillert J, Olsson T, Piehl F, Lidman O. Genetically determined susceptibility to neurodegeneration is associated with expression of inflammatory genes. Neurobiology of disease 2006;24:67-88.

Appendix 1, Markers

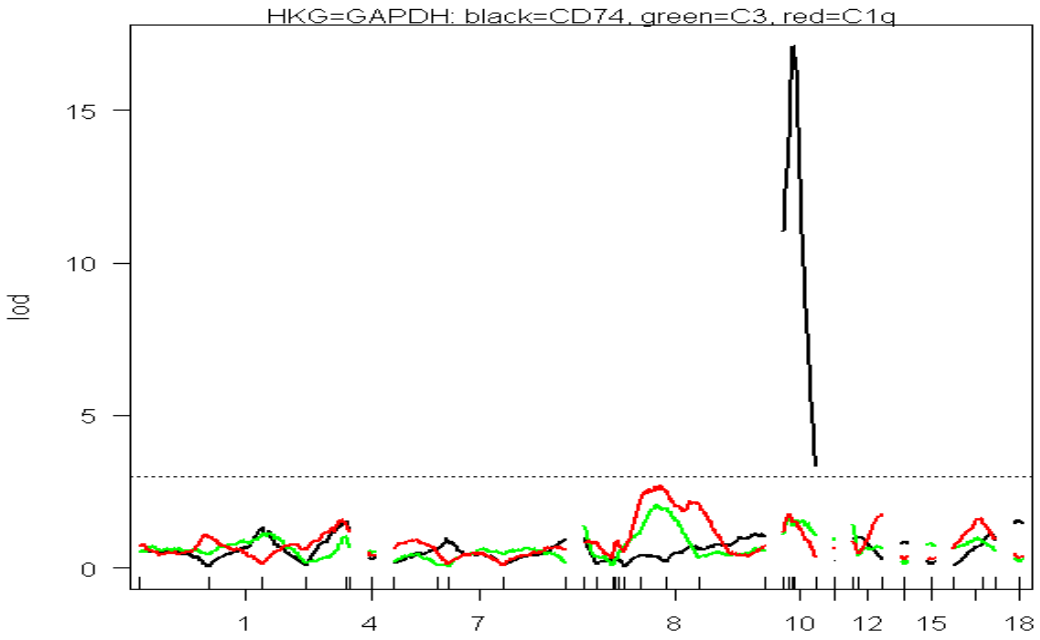
Ch	Marker name	Pos. (Mb)	FHHx ACI (cM)	Fw primer (5´> 3´)	Rew primer (5´> 3´)
1	D1a06	-	-	GGTTTGTTACTTTGTTATCC	AGATAGAGAGAGAAAGAAAGG
	D1Rat10	25.27	21.77	TGTTATCAGCCCTGTCCTCC	TGAAAAGGCTCTGGTTGTGA
	D1Rat234	41.23	27.87	GCTGCCATTTAGTCTGGCTT	GGGTACACTGGACTGGGAAA
	D1Rat20	48.78	30.32	GTTCGGAGGACAAACCAGAA	GGCCCTCACACTGTCTTCTT
	D1Rat94	65.48	35.70	TGTCCACACAAGAGGGAACA	AAGGTCAGAGGTCAATATCAGTTG
	D1Rat378	65.49	35.75	TTTGAGGTGTAACATGATTTTTGAA	GAGATACGCTAAAAAGCAGGG
4	D4Rat78	118.48	66.84	GTTTTCTCTGTGCGCCTGTT	TTGATATGAATGCTTGTTTTGCTT
7	D7Rat103	44.48	24.58	CTGGTGCTTTTGGGTCTGTT	GTGTCAAACTGTGGGGATCC
	D7Rat58	-	33.42	AATGCAAACCTGGTCAAAGC	TAAGCATTTCCCAGAGACGC
	D7Rat191	58.66	34.34	AATGACACAACAAATGGGCA	CCACTGGTTTTCGTGGTACC
	D7Rat86	88.15	40.04	TGTAGAAATCCTTCACAAACTTGAA	AAAGACACTTTTCAATCCCCC
	D7Rat129	119.20	-	CCAAACCACTGTCTTACAGATGC	AACCCACAGCTTCAGGACAC
8	D8Rat26*	79.37	53.5	GCAAGACAACCAGGCATGTA	AGCCTTGGCACTGAACACTT
	D8Rat24*	82.26	-	ACCGTTATCTTGTCCCCCTC	CCCCAGTGTAACGAACCTCT
	D8Got128*	85.85	-	CCAGGAACCTGAAACAGGAA	AAGGTAGACATTACACATGCACA
	D8Rat205*	86.30	-	CATGTAGGAACTTTCTATCACCCA	TGTGTATCTCTGTCTGTCTCTGTTG
	D8Rat138*	89.08	-	CCTGAGTGATTGTCTCAGAGGA	CCAAAAGGGAAACGTTTTGA
	D8Rat75*	89.58	59.09	CACCTCACCTTAAGAACAATGAAA	CAGTGGGAGGCTCATTTTATAGA
	D8Got145*	91.35	-	CCATCAACCCTCTGGAAAAA	CTAAGAGCATCAAAGTGCTTACACA
	D8Rat129*	98.97	-	TCAAACTTCTGAGTTTCTAAAAGGA	ATCTACAACGAAGCCACCCA
	D8Rat14	104.7	70.19	GGCCGGTCTAATTATTTCTTCA	GCCCATACGTTGCATCAAGT
	D8Rat76	-	79.05	GTGGGAGCAGATGGTCAAGT	TCCACCAATTCTTACCTGCC
	D8Rat2	123.44	86.77	TGTATAGCATGTGTGTGTGCG	GTACCTGTCATCCACATGCG
10	D10gGot6	3.22	-	GGTGTGGCTCAATGTGCC	CATTTGCCAAGGATTTCACAT
	D10Got4	4.20	-	GGCCATTTGGAGTTAAGAGCA	CAGTTTGAGTGGCTCATCTGTAATT
	D10Mgh25	5.09	-	TTCCAGTGAGATTACTGCTCCA	AAGTTCCAGACCAGTATGGGG
	D10Got7	5.13	-	CCAAGGCTGGTGAGAATCTC	ACACTGATGCCTAACATTTTT
	D10Rat47	16.08	10.07	TTTTTCCCCTTCTTTCTGACTG	TCACAATTCTCCACGTGAGG
11	D11Rat83	63.97	44.02	ATAACCAGAAATCCCCACCC	AATCCTAGTGGCAGGGAGGT
	D11Rat34	-	44.02	ATCGTTGTAACCACTTCCCG	ATTGCATTTAACCAGGTCGC
12	D12Mit2	20.92	9.01	GTGGCTCTTTTCCTTAGGGA	TCGGCTTCTGAATGTATTGG
	D12Got46	23.23	-	GCTCAAGGCCAGGATTCATA	TGCAACACACCCCACCAT
	D12Rat23	25.23	11.22	CTAAGACTTCCTGCATGGGG	TTTCTGAATCATTCCTGGAGAA
13	D13Rat141	103.29	58.82	CAATATGTTGACACTTAAAAATTGC	CCAATCACACACACAGGTGAG
15	D15Rat28	-	74.04	TGGTCTCCCTGCTAATCCTG	GTTTCAGGAACTGCCACAGG
16	D16Rat101	-	36.06	CGCTGCTGCACATGATATCT	AACAGGAAATCCATTACTTTGC
	D16Rat13	-	41.35	TAGGTTAGGGAGGGTGGGTC	TGACTTCTACTTTCATTCAAATGCA
	D16Rat93	-	47.02	TATTGCTGGTTGTTTGTGCA	ACACTGAGGTGCAAGATCCC
18	D18Rat60	-	45.28	CATGCATGTGTGGGTACACA	TGCTGGCTGGTAATTTCCTT

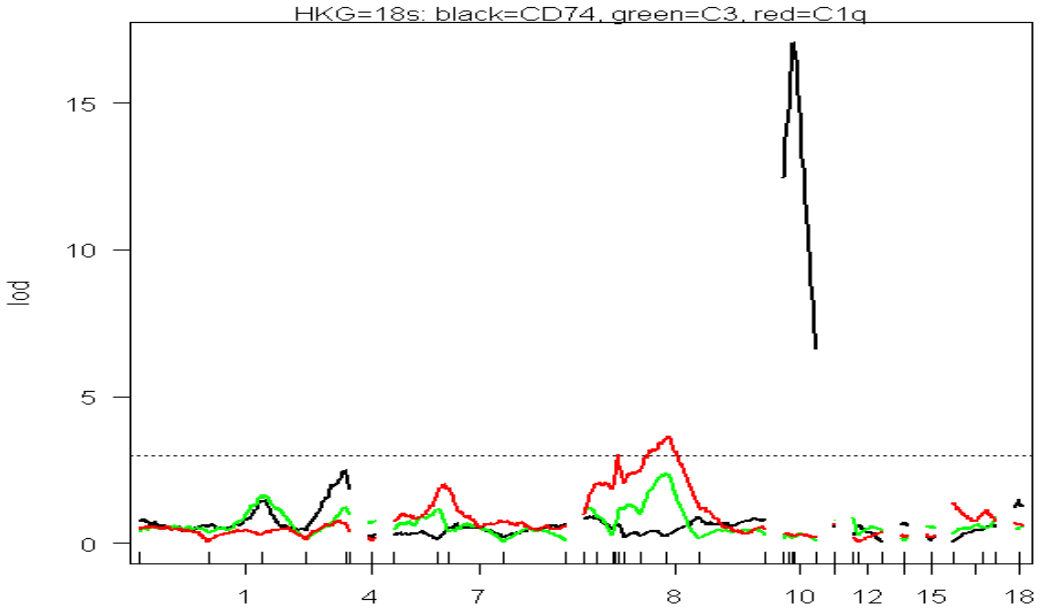
Table A1: Markers genotyped and used in linkage analysis. *D1Rat121, D6Rat81, D7Rat192* and *D13rat163* were also analyzed but turned out to be non polymorphic. Selection was based on position on the physical map FHHxACI and the genomic map and therefore some are chosen without knowledge of genomic position and vice versa. The position for markers on chromosome 10 was mainly based on an earlier study [Swanberg et al. 2005]
*Markers previously genotyped by Karin Harnesk.

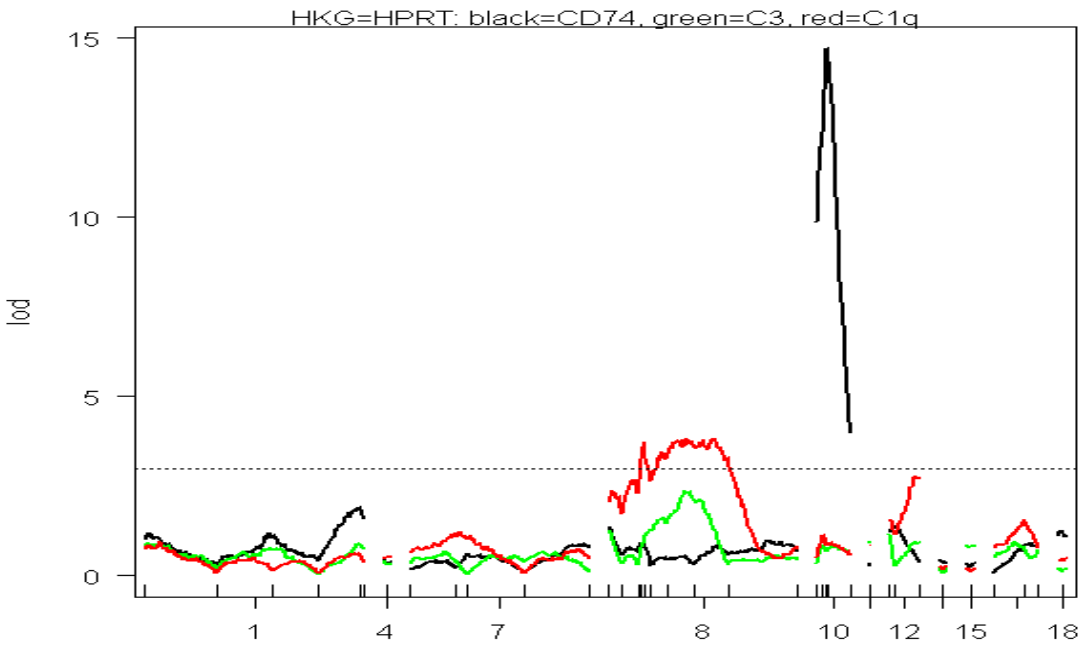
All markers were ordered based on sequences at common databases. Some already existed at the lab and others were ordered during the projects progression. The markers were mainly chosen based on their position in centimorgan on the FHHxACI map. Therefore these could be chosen without knowing the position in Mb. The marker d1a06 had which lies within the EAE 29 QTL had been designed at the lab and was at the time of this report not available at the common databases Rat Genome Database (RGD) or Ensembl Genome Browser.

Appendix 2, Effect on Linkage of HKG

Linkage analysis performed with expression values calculated with the different HKGs separately. The dotted line indicates the value of LOD 3.0 that approximately was the 95% threshold value for all phenotypes.







	Marker	GAPDH	18s	HPRT	**GAPDH&18s**
CD74	D10Mgh25	17.11	16.95	14.64	**18.48**
	D1Rat94	1.51	2.45	1.91	**2.13**
C1q	D8Rat14	2.49	3.51	3.59	**3.30**
C3	D8Rat14	1.90	2.40	2.00	**2.22**

The table above presents LOD scores for some chosen markers within each phenotype group, calculated with expression values using the different HKGs separately. Last column indicates the values from the multiple run, chosen to be used in the project. The diagrams clearly show that the difference between the HKGs is small. Corresponding diagram using multiple HGKs is presented in the Results section. 18s gives the clearest result at positions with low LOD with a few distinct peaks and the rest of the markers without linkage tendency.