The Antigen Presenting Cell Lectin-like Receptor gene Complex, APLEC, regulates susceptibility to oil induced arthritis

Doctoral thesis by

Line Mari Flornes

Institute of Basic Medical Sciences.
Department of Anatomy,
Faculty of Medicine
University of Oslo

Oslo, 2010

© Line Mari Flornes, 2010

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 996

ISBN 978-82-8072-523-3

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS.

Produced in co-operation with Unipub.
The thesis is produced by Unipub merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Contents

Acknowledgements.. 4

List of papers...6

Abbreviations...7

Introduction..9

The major arms of the immune system..10

Lectin receptors...14

Antigen presentation..17

Dendritic cell..18

Tolerance..19

Autoimmunity..25

Genetic mapping of QTL...27

Experimental animal models for autoimmunity.............................32

Aims of the study..38

Summary of results..41

General discussion... 45

NKC vs APLEC..45

APLR function..46

APLR ligands..50

APLR in autoimmune diseases...51

APLR and pathogenesis...54

QTLs mapping to APLEC...56

Future perspectives...58

References...59

Acknowledgments

The work summarized in this thesis was performed at the Department of Anatomy, Institute of Basic Medical Sciences, University of Oslo, from 2001 to 2010. I received a PhD fellowship from July 2001-Dec 2004 from VISTA (Statoil through The Norwegian National Academy of the Sciences and the Letters) and additional financial support was received from Helse SørØst through Diakonhjemmet.

I would like to thank the Department of Anatomy and the Institute of Basic Medical Sciences by present chairman Jan Bjålie for providing necessary working facilities. I am also grateful for the help I have received from friends and colleagues throughout the years. I especially wish to thank:

- my supervisor Sigbjørn Fossum for taking me into the lab and introducing me into the fields of immunology, genetics, animal models and evolution. Your engagement and optimism has been inspiring. Thank you for the time you have spent analysing, evaluating and discussing the results I have presented to you and for providing me with funding allowing me to finish my thesis.

- my co-supervisors Erik Dissen and Mike Daws. Erik: Thank you for teaching me laboratory methods and molecular biology and for your critical review of planned and executed experiments. Mike: Thank you for taking over supervision in Sigbjørns absence and for helping me finishing my thesis.

- Wendi and Marianne for expert technical assistance. Thank you Stine, Per Christian and Thorny for hours spent together at late hours in the lab and social times outside the lab. Thanks also to all other current and past members of the lab for making it a good working place. I have enjoyed the scientific and non-scientific discussions, the chats, the pub-quiz, the hiking- ski- and cycling trips we have been to, as well as the barbeques. I also need to thank my current office-mates Britt, Maria, Ana and Lise. The chats in the office and lunchroom equal my social life the past few months.

- Johnny Lorentzen for teaching me the OIA model and for a fruitful collaboration. Also thanks to Liselotte and Ulrica, I enjoyed working with you during your stay in Oslo.

- the staff at the animal facility, especially Vera and Sophia for help with the animals.

- all my co-authors for their contributions in the papers.

A special thank you to the late Bjarne Waaler, who argued for my project and persuaded Statoil to grant me a fellowship.

A big thank you to all my friends, my family and the Aurmark family for encouraging me to work on when I was demotivated and for still being my friends despite weeks of ignorance due to labwork. Thank you for all the support throughout my education and for providing me with a life outside of science.

And of course a special thank you to Liam and Kieran. Thank you Liam for all your support and love, and to Kieran: your curiosity for life is a true inspiration.

List of papers

This thesis is based on the following papers, which will be referred to by their roman numerals:

I. Flornes, LM., Nylenna, Ø., Saether, PC., Daws, MR., Dissen, E. and Fossum, S. The complete inventory of genes in the rat natural killer cell gene complex. *Manuscript (submitted).*

II. Ribbhammar, U., Flornes, LM., Bäckdahl, L., Luthman, H., Fossum, S. and Lorentzen, JC. High resolution mapping of an arthritis susceptibility locus on rat chromosome 4, and characterization of regulated phenotypes. *Hum. Mol. Genet.* 2003 12:2087-96

III. Flornes, LM., Bryceson, YT., Spurkland, A., Lorentzen, JC., Dissen, E. and Fossum, S. Identification of lectin-like receptors expressed by antigen-presenting cells and neutrophils and their mapping to a novel gene complex. *Immunogenetics*. 2004 Oct;56(7):506-17

IV. Lorentzen JC, Flornes L, Eklöw C, Bäckdahl L, Ribbhammar U, Guo JP, Smolnikova M, Dissen E, Seddighzadeh M, Brookes AJ, Alfredsson L, Klareskog L, Padyukov L, Fossum S. Association of arthritis with a gene complex encoding C-type lectin-like receptors. *Arthritis Rheum.* 2007 Jul 30;56(8):2620-2632

V. Flornes, LM., Guo, JP., Dissen, E., Daws, MR., Lorentzen, JC. and Fossum, S. APLEC receptors influence IL-17 response in oil induced arthritis. *Manuscript.*

Abbreviations

AIL	Advanced intercross lines
APC	Antigen presenting cells
APLEC	Antigen presenting cell lectin-like receptor gene complex
APLR	Antigen presenting cell lectin-like receptors
CCP	Cyclic citrullinated peptides
CTLR	C-type lectin-like receptors
DC	Dendritic cells
EAE	Experimental autoimmune encephalomyelitis
GWAS	Genome-Wide association studies
HLA	Human leukocyte antigen
IFA	Incomplete Freund's adjuvant
IL	Interleukin
ITAM	Immunoreceptor tyrosine-based activation motif
ITIM	Immunoreceptor tyrosine-based inhibitory motif
KLR	Killer cell lectin-like receptors
LD	Linkage disequilibrium
MHC	Major Histocompatibility Complex
MS	Multiple Sclerosis
mTEC	medullary thymic epithelial cells
NKC	Natural killer cell gene complex
OIA	Oil induced arthritis
PIA	Pristane induced arthritis
PAMP	Pathogen associated molecular patterns
PRR	Pattern recognition receptors
QTL	Quantitative trait loci
RA	Rheumatoid Arthritis
RF	Rheumatoid factor
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
TCR	T cell receptor
TLR	Toll like receptors
TM	Transmembrane
TRA	Tissue-restricted antigens
Treg	Regulatory T cells

Introduction

Multicellular organisms are exposed to a multitude of microbial and viral pathogens that make use of a wide variety of strategies in order to invade the host and escape host counterattacks. In the arms race between host and pathogen, the latter has the advantage of a much shorter generation time, e.g. less than 30 minutes for *E.coli* compared with 20 years for the human host. Consequently, the pathogens can quickly adjust their strategies through natural selection. The mammalian immune defence system therefore not only faces the problem of having to recognize a large number of different pathogens as foreign, but must also be able to reveal variants of pathogens not previously encountered by the host in evolution, and to launch the right kind of response, adjusted to type and severity of the attack. These requirements have resulted in the development of a flexible and complex defence system. A complex and flexible system is prone to malfunction and this may have serious consequences for the organism. Malfunctions can be divided into three major types: 1) immunodeficiency, or the failure to detect and react to intruders, 2) autoimmunity and 3) maladjusted reactions, e.g. allergies.

The focus of this thesis is autoimmunity, a condition where the defence system reacts and inflicts tissue damage to the body's own constituents. Most autoimmune diseases are relatively rare, the more common are rheumatoid arthritis (RA), multiple sclerosis (MS), systemic lupus erythematosus (SLE), autoimmune thyroid disease and type 1 diabetes. Collectively autoimmune diseases constitute the third largest category of illnesses in the industrialized world, next only to cardiac disease and cancer.

Autoimmune diseases are complex disorders where several genes and environmental factors contribute to disease development. Identification of such risk factors will be important for improved disease prophylaxis, diagnosis and therapy. The ultimate goal of the work presented is to identify genes associated with human autoimmune diseases and find out how they contribute to precipitation of disease. To this end, we have taken advantage of an experimental animal model, where autoimmune joint inflammation is induced by application of adjuvant oils to the skin. Gene candidates emerging from this work have subsequently been tested for association with RA in the human. As a background to the work, a brief introduction to the immune system and to methods for identifying genetic risk factors will be given.

The major arms of the immune system

Most of our cells contribute to defence against pathogens, the majority acting as a "home guard". In addition we have a set of specialized cells collectively referred to as leukocytes, dedicated to immune defence. To distinguish between self and non-self, leukocytes are equipped with molecules called antigen receptors. Depending on how leukocytes recognize pathogens and the immune reaction launched, the immune system can be divided into four major branches as indicated in table 1.

Table 1. The antigen receptors used by the four major arms of the immune system

	Innate	**Adaptive**
Humoral	Soluble lectins Complement system	B-cell receptors (antibodies, immunoglobelin)
Cellular	NK cell receptors Toll like receptors	T cell receptors (αβ and γδ TCR)

The table shows the types of antigen receptors employed in each arm. In the case of the innate system only a couple of examples are given from a much longer list of receptor types.

In the cellular system the antigen receptors remain cell bound, whereas in the humoral system the antigen receptors are released into the tissue fluids. The cells of the non-adaptive immune system respond vigorously upon first encounter with a particular antigen. In contrast, when cells of the adaptive system encounter antigen, the primary immune response is weak, but is enhanced upon later encounter with the same antigen.

The adaptive immune system

The hallmarks of the adaptive immune system are specificity and memory. Specificity is achieved through the generation of diverse antigen receptors by gene rearrangement. The process is independent of antigen and produces an almost infinite number of different antigen receptor specificities. The rearrangement of genes occurs at the DNA level, and is inherited stably so that all daughter cells from a B or T cell that have successfully undergone productive rearrangement will express the same antigen receptor specificity as the parental cell. Due to allelic exclusion the cells will as a rule express only one single receptor specificity. As a consequence tolerance can be induced at the level of cells by deleting clones with self reacting receptors.

The clonal division comes at a cost. As the vast majority of the cells never become engaged in immune responses, B and T cells have adopted a "resting" phenotype, consisting of little more than a condensed nucleus surrounded by a thin brim of cytoplasm. Furthermore, the sheer multitude of different clones means that the average size of each clone is tiny. Upon recognition of a particular antigen the reactive clones therefore first have to expand and the cells subsequently to differentiate into effector cells. It will therefore take several days to launch a primary immune response, which consequently will be slow and relatively weak. Following proliferation, not all the progeny differentiate into effector cells. Some are set aside as memory cells, with the same antigen specificity as the parental cells. These memory cells are of greater quantity than the parental clone and are more readily activated. Upon later challenges with the same antigen, the immune response is faster and stronger, a feature denoted as immunological memory.

The low frequency of antigen specific cells restricts the chance of antigen detection of a particular microbial intruder. In order to increase the probability of antigen detection, B- and T-cell circulate between strategically positioned checkpoints named peripheral lymphoid tissues, which include lymph nodes, splenic white pulp, tonsils and intestinal Peyer's patches. To further assist in antigen detection, professional antigen presenting cells (APC) capture antigen in the periphery and bring it to the lymphoid tissues where it is presented to T cells. In addition, the APC transmits "secondary" signals, mediated by cell bound co-receptors and secreted proteins, which contribute to instructing the T cell how to respond to the antigen presented.

The receptor repertoire generated through gene rearrangement can in principle bind any antigen. An inevitable by-product of this process is the generation of antigen receptors directed to self antigens. Because their existence may confer autoimmune reactions, tolerance to self constituents needs to be learnt during ontogeny. Due to the one cell-one receptor strategy tolerance can be obtained by eliminating or inactivating cells equipped with self-reactive receptors from the total receptor repertoire. This method of ontogenetic tolerance acquisition is probably the rationale behind the one cell-one receptor strategy, as it would be far more difficult to construct systems where cells could obtain the necessary feed-back information for differential handling of a multitude of expressed, nearly identical receptors. One great advantage with generating receptors against everything and

subtracting the self-reactive is that the remaining receptor repertoire should be able to bind all kinds of non-self antigens, including antigens previously never encountered in the phylogeny of the species. Evidently, this useful property, referred to as anticipation, cannot develop within a system dependent on phylogenetic learning.

Once activated, B cells differentiate into plasma cells which produce antibodies with the same specificity as the membrane bound antigen receptor, but now in a form secreted into the tissue fluids where they bind extracellular antigen. The binding elicits a variety of effector mechanisms, such as ingestion by phagocytic cells or the activation of the complement system, both leading to clearance of the antigen. Binding of antibodies may also neutralize invading pathogens, by immobilizing and hindering them from entering cells. Activated cytotoxic T cells, on the other hand, target infected cells and kill them by direct contact. Normally, the activation of B cells and cytotoxic T cells requires the assistance of helper T cells, provided by physical cell-to-cell contact and by secreted cytokines. In order to provide the required stimulatory signals, the helper T cells need to be activated through interaction by professional APC. As indicated above, this depends both on the recognition of antigen presented by the APC and the reception of secondary signals. Importantly, the helper T cells recognize different antigen determinants than the B cells and the cytotoxic T cells do. The requirement of independent recognition of different antigen epitopes plus stimulation by secondary, co-stimulatory signals represents an important safe-guard against inappropriate activation of immune responses.

The non-adaptive (innate) immune system

The *de novo* generation of novel protein-encoding genes through gene rearrangement during ontogenetic development of the individual is a unique property of the adaptive immune system. In contrast, the receptors used for antigen recognition by the leukocytes belonging to the non-adaptive or innate arm of the immune system are inherited in the traditional way, i.e. “hard-wired” via the germ-line DNA. This raises the question of how to avoid the problem referred to above with rapidly proliferating parasites escaping detection through natural selection. The answer is that microbes contain macromolecular structures that are hard if not impossible to alter without strongly reducing the fitness of the microbe, thus prohibiting escape by the microbe from immune detection by mutations. Examples are the complex peptidoglycan bacterial coat and the formyl group attached to

the initial methionine in all newly synthesized bacterial proteins. The germ-line encoded receptors of the innate immune system are directed against such evolutionary conserved molecular motifs. Evolutionary conservation means that these structures often are shared by a range of different microbes. They are therefore referred to as pathogen associated molecular patterns (PAMP) (1). Immune receptors recognizing PAMPs are called pattern recognition receptors (PRR) and include Toll like receptors (TLR), intracellular Nod-like receptors and C-type lectin-like receptors (CTLR, i.e. containing C-type lectin-like domains) (1-3). Extracellular plasma bound TLR detect components of the cell wall of bacteria, fungi and protozoa, whereas intracellular TLR can detect bacterial and viral nucleic acid structures. Upon recognition of pathogens, TLR have the ability to initiate intracellular signalling leading to activation of the cell (4). Many, but not all CTLR bind carbohydrates. Some CTLR contain an internalization motif and are internalized upon binding with antigen, resulting in antigen presentation. However, a number of CTLR possess signalling motifs and are involved in intracellular signalling similar to the TLR (5). Binding of PRR to pathogens or pathogen derived structures will activate host defence mechanisms such as secretion of inflammatory cytokines and chemokines as well as internalization or lysis of the microbe.

With receptor specificities completely defined by the germ-line DNA, self-tolerization is acquired phylogenetically through natural selection. There is therefore no need for tolerance acquisition ontogenetically, i.e. for the individual to “learn” the difference between self and non-self, and therefore no need for clonal division of the cells with respect to antigen receptor specificities. Each cell may therefore simultaneously carry many different antigen receptors, enabling them to recognize a wide range of different pathogens. Following intrusion of antigens, a substantial fraction, if not the majority of these cells will immediately recognize the intruder. Rather than being assigned to tedious, most of the time non-productive patrolling of control posts, like the B and T cells have to do, the cells of the innate system can therefore be directly positioned at the frontlines where the enemy is likely to attack, i.e. along the skin and the mucosal linings and in the bloodstream. Moreover, upon encounter with intruders the cells do not need to undergo rounds of proliferation to amplify the response and subsequent maturation and, but are combat-ready, pre-packed with various granules containing antimicrobial proteins, lysosomal enzymes and regulatory cytokines. This arm of the immune system is therefore acting fast even on the first encounter with antigen, hence the term “innate”, but the

response is not amplified in subsequent encounter with the same antigen, hence the term "non-adaptive".

Another innate response is the complement system which consists of a number of secreted proteins that interact with each other forming a cascade. The proteins are circulating in the blood, and upon binding to microbes the cascade is set off, cleaving the next protein in the cascade. Some of the peptides generated in the cascade act as inflammatory agents attracting phagocytic cells (neutrophils and macrophages), others create an attack complex in the target cell membrane, leading to cell lysis, or mark the targets for uptake by phagocytes (opsonization).

In addition to constituting an independent arm of the immune defence system, able to combat intruders on their own, the innate system provides support and instruction to the cells of the adaptive immune system. Of particular importance is the key role played by professional APC in stimulating helper T cells. As briefly mentioned above, professional APC sample, process, transport and present antigens to T cells, and through the expression of co-stimulatory receptors and secreted cytokines deviates the immune responses in suitable directions, adjusted to the type of attack. As will be detailed below, this includes deviation towards non-response.

Lectin receptors

The professional APC utilize a wide array of receptors and co-receptors to recognize and sample foreign antigens and to communicate via cell-to-cell contact with other immune cells. For reasons explained below, we shall here concentrate on one subset of such receptors, belonging to the C-type lectin superfamily. The C-type lectin-like receptors constitute one subgroup of a large class of proteins collectively referred to as lectins, which originally are defined as carbohydrate binding proteins. The CTLR are characterized by the possession of one or more C-type lectin-like domains, consisting of a characteristic structural motif first identified as the carbohydrate recognition domain of mannose binding lectin, where the C denotes calcium dependent binding of carbohydrates (6). Many of the proteins with this domain do not depend on calcium for saccharide binding or do not even bind carbohydrates at all, hence the term "lectin-like". In

vertebrates the C-type lectin-like domains are classified into 17 groups based on overall protein architecture (6-8).

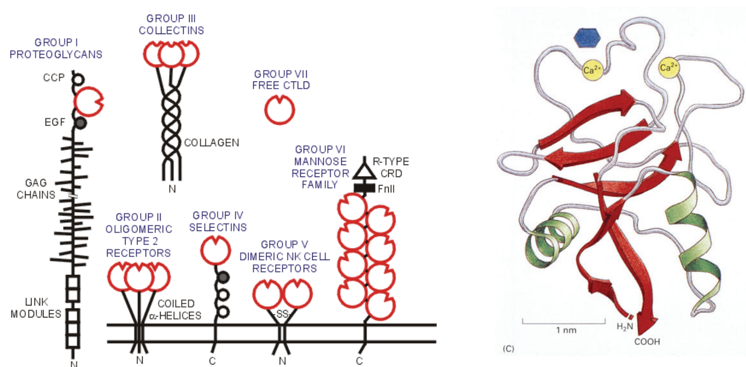

Fig. 1 C-type lectins and the C-type lectin domain

Some of the groups in the C-type lectin superfamily, including group II and group V, and diagrammatically the carbohydrate recognition domain (C-type lectin domain) of mannose binding protein, with alpha helices in green, beta strands in red, Ca^{++} -binding sites in yellow and position of saccharide binding site indicated by blue hexagon. From Kurt Drickamer: www.imperial.ac.uk/research/animallectins/

Group II and V contain membrane bound C-type lectin-like receptors of which many are expressed by leukocytes and may function as PRR. Group II and V are type 2 transmembrane (TM) proteins with a single membrane external C-type lectin domain (6). The group V receptors are mainly encoded by genes in the Natural Killer gene complex, NKC, and the receptors are referred to as killer cell lectin-like receptors, KLR. The KLRs are expressed as disulphide linked homo- or heterodimers (9,10). The characteristic amino acid motif mediating Ca^{2+} dependent carbohydrate binding has been lost, and in cases where ligands have been identified most of these are proteins (9,11). A notable exception is Dectin-1, where the ligand is a β-glucan (12). Many of the receptors contain known signalling motifs and are involved in regulation of NK cell activation. The signalling properties of the KLRs will be detailed below.

In contrast to the group V receptors, the group II receptors have retained their motifs for carbohydrate and Ca^{2+} binding, and carbohydrate ligands have been identified for many of the receptors (8,13,14). Many of the receptors contain distinct internalization motifs in the cytoplasmic tails indicating that these receptors are involved in pathogen internalization (15). One subgroup of the group II receptors are the receptors encoded by genes in the antigen presenting cell lectin-like receptor gene complex, APLEC. The APLEC receptors

(APLRs) do not contain known internalization motifs; however some of the receptors contain signalling motifs similar to the KLRs and are probably involved in regulation of cell activation (16-19). The role of the APLRs will be discussed in more detail in the general discussion.

Signalling in C-type lectin receptors

The focus of this thesis is the APLEC receptors. Some of these contain signalling motifs identical to those present in KLR. As most studies have been performed on KLR, their signalling properties will be outlined as a likely model of APLR functional properties.

The KLR come in two functional forms, activating and inhibitory. The inhibitory receptors have intracellular domains with the immunoreceptor tyrosine-based inhibitory motif (ITIM) V/I/LxYxxL/V (where x denotes any amino acid). Upon binding of ligand, the ITIM tyrosine becomes phosphorylated, followed by recruitment and activation of SH-2 domain-containing tyrosine phosphatases that mediate the downstream effect, inhibition of the killer cell (20,21). In contrast, the activating receptors do not contain any known intracellular signalling motif. As a rule, they instead contain a positively charged amino acid in their TM domain. The positive charge facilitates association with TM adapter proteins containing a negatively charged amino acid in their TM domain (fig 2). The adapter proteins Dap12 and FcRγ contain the immunoreceptor tyrosine-based activation motif (ITAM) YxxL/Ix_{6-8}YxxL/I in the cytoplasmic domain, whereas Dap10 contains a YINM motif. Upon receptor ligation the tyrosine becomes phosphorylated, recruiting tyrosine kinases which trigger an activating signalling pathway (21). Several of the genes in the NKC encode protein chains lacking signalling motifs. Some of these form heterodimers with chains containing such motifs. This is the case for human CD94, which associates with either the inhibitory NKG2A chain or the inhibitory NKG2C chain and for rat KLRE, which associates with either the inhibitory KLRI1 or activating KLRI2 (10,22).

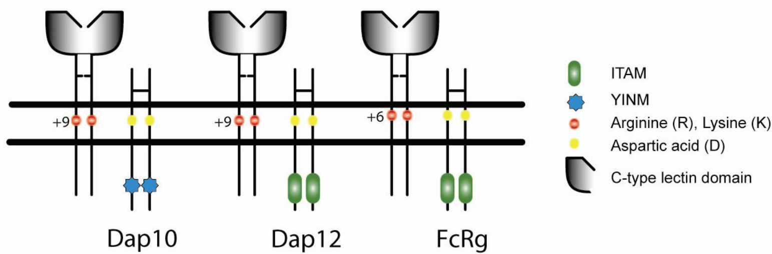

Figure 2. Relative positions of charged amino acids in TM domain of adapter proteins
The adapter proteins DAP10 and DAP 12 contain a negatively charged amino acid in the center of the TM domain, whereas FcRγ contain a negatively charged amino acid close to the extracellular border of the TM domain. For association, the receptors need an amino acid with a positive charge in the corresponding position. The numbers indicate the position of the charged residue counted from the extracellular surface (adapted from (23)).

The expression of receptors with opposing signalling function reflects complex mechanisms for regulation of cell activation, where the signals mediated by the receptors are integrated and the cellular activation status depends on the balance between activating and inhibitory signals.

Antigen presentation

Antigens are presented to T cells by specialized molecules on the cell surface. These molecules are encoded by genes residing in the Major Histocompatibility Complex, MHC. The human MHC, localized on chromosome 6, is the most gene dense genomic region known and the majority of the genes within this complex encode proteins with immunological functions. The proteins can be divided into three classes termed MHC class I, II and III, where the classical MHC class I and MHC class II genes encode proteins that present peptides to the T cells. There are several different MHC class I and class II loci, most of which are highly polymorphic (24). The classical class I molecules presents peptides derived from proteins within the cell, either a self constituent or derived from an intracellular pathogen. The MHC class I molecules present peptides to cytotoxic T cells. The majority of nucleated cells express MHC class I molecules, and can thus be screened for intracellular parasites. If a cytotoxic T cell recognizes the peptide-MHC class I complex, it can kill the cell once activated (24).

Peptides presented by MHC class II molecules are derived from exogenous proteins engulfed by the cell either by phagocytosis or by receptor mediated endocytosis. The proteins are broken down into peptides in endosomes and lysosomes and loaded onto MHC class II molecules. The complex of MHC class II molecules and peptide is then brought to the cell surface and is presented to helper T cells. MHC class II molecules are constitutively expressed by professional antigen presenting cells which include macrophages, B-cells and dendritic cells (25).

Dendritic cell

Among the professional antigen presenting cells dendritic cells (DC) are unique in their ability to activate naïve helper T cells. The cells constitute a heterogenous group of cells. Two different lineages, lymphoid and myeloid have been described, and the different lineages can be divided in different subsets and maturation stages with different phenotypes, localizations and functions. As the DC phenotypes depend on lineage and maturation stage, only a general overview of dendritic cells will be given here.

The DC typically resides in tissues in an immature state. The distribution of immature DC is especially high in the skin and along the mucosal linings, the entry point of the microbes. Immature DC has low MHC class II expression and high PRR expression. Upon recognition of pathogen, the DC internalizes the pathogen and undergoes phenotypic alterations referred to as maturation, characterized by antigen processing and up- and down-regulation of receptors. The cell is mobilized, leaving the peripheral tissue. The exit route is mainly with draining lymph, and with the chemokine receptor CCR7 (a lymph node homing receptor) up-regulated, the cell is usually destined to end up in the draining lymph node. MHC class II molecules are translocated to the plasma membrane and co-stimulatory molecules and secretion of cytokines are up-regulated. The capacity to phagocytose is diminished during maturation. Upon arrival in the lymph node, the by now mature DCs are ready to present antigen to naïve T cells (26-28).

Upon pathogen recognition DC also secrete various cytokines in the infected tissue. The cytokines serve to stimulate tissue resident macrophages and to attract blood-borne NK cells and neutrophils. These are immediately engaged in local battle against the intruders while they at the same time recruit precursor DC. The precursor DC subsequently

develops into immature DC, replacing the recently emigrated DC. As long as pathogens are present, maturing DC will continue to migrate to the draining lymph node, sustaining the activation of T cells. Thus dendritic cells are important to activate and amplify both the innate and the adaptive response (27,28). In addition, DC plays a crucial role in both central and peripheral tolerance induction.

Tolerance

The receptor repertoire generated through gene rearrangement can in principle bind any antigen, including self antigens. The presence of clones with self-reacting receptors may lead to autoimmune reactions, and tolerance to self constituents needs to be learnt. Both B- and T cells need to be tolerized, but here only the mechanisms for T cell tolerance will be described. Generally the helper T cell is the crucial cell to tolerize, as both B cells and cytotoxic T cells need assistance from helper T cells for activation. In the thymus the developing T cells, or thymocytes, are educated, i.e. they undergo positive and negative selection, so that non-functional clones are selected away and self-reacting clones deleted. The outcome is a pool of T cells capable of recognizing foreign antigen presented by self-MHC molecules. The deletion of self-reactive T cell clones during thymocyte development is referred to as central tolerance. The process, however, does not completely eliminate all potentially self-reactive T cells. Autoagressive T cells escaping the thymus are instead rendered unresponsive by control mechanisms referred to as peripheral tolerance.

Central tolerance

For their continued survival developing thymocytes depend on signals mediated through their T cell receptor (TCR) and thus on the ability of their TCR to bind self peptide-self MHC complexes and the strength of this binding (29). Thymocytes with a TCR that do not bind self MHC are useless, as they will not be able to interact with infected target cells or APC in the periphery. If the TCR is non-functional, no signal will be transmitted and the thymocytes are eliminated due to lack of stimulation - “death by neglect”. Conversely, thymocytes carrying a TCR with strong affinity for self peptide-MHC complex are potentially dangerous. High affinity TCRs mediate strong signals which induce apoptosis and thus eliminate the thymocyte by negative selection. Thymocytes with low TCR affinity for self peptide-MHC complex are unlikely to react with self in the

periphery; however, they may potentially bind strongly to non-self peptide presented by self MHC. The intermediate signalling through low affinity TCR will give a sufficiently strong signal for survival, but not so strong as to induce apoptosis. These cells are consequently positively selected for further differentiation into mature T cells (30,31). The signalling strength is not dependent on TCR affinity alone, but also on overall avidity of receptors, co-receptors and co-stimulatory molecules (29,32).

A requirement for central tolerance induction is the presentation of a diverse array of tissue-restricted antigens (TRA) in the thymus. A population of epithelial cells in the thymic medulla, mTEC are specialized to express a diverse set of non-thymic, tissue-specific molecules. This promiscuous expression of self-antigens represents virtually all organs, thereby mirroring the peripheral self (33). The expression of TRA in mTEC is in part controlled by a transcription factor called AIRE (*autoimmune regulator*) predominantly expressed in the mTEC (34,35). As the name imply AIRE is involved in the regulation of autoimmune disorders. Mutations in the human *AIRE* gene cause a multi-systemic autoimmune syndrome known as autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) (36,37). *Aire* knock-out mice suffer from autoimmune symptoms that affect multiple organs similar to the human syndrome APECED and analysis of mTEC from *Aire* -deficient mice showed reduced transcription of several tissue-specific genes associated with various autoimmune diseases (34,38). The transcription level of a number of TRA appears to be dependent on Aire to various degrees, indicating complex regulation of expression by Aire. Moreover, additional factors that regulate TRA expression exists as transcription of some genes are not influenced by Aire at all (34,39-41).

The presentation of TRA to thymocytes may be mediated by mTECs, or by thymic APC. It has been shown that TRAs expressed in mTECs can be transferred to and cross-presented by thymic DCs (42). Transfer may be a result of antigen uptake of secreted exosomes or uptake of apoptotic bodies of mTECs (43,44). In addition to regulating TRA expression, Aire induces apoptosis in mTECs, a possible mechanism to promote cross-presentation of TRA by thymic DC (31,45).

Peripheral tolerance

Despite the clonal deletion of autoreactive T cell clones during thymocyte development, some autoreactive T cell clones escape to the periphery. However, several peripheral mechanisms contribute to keeping these potentially dangerous cells in check. Self antigens are constantly taken up and processed by immature DC and presented to T cells in the draining lymph node. Immature DC have low levels of MHC class II molecules and costimulatory receptors (46). If an autoreactive T cell recognise self antigen presented by immature DC, the T cell will not receive appropriate signals to be activated. The engagement of TCR but lack of other activation signals will result in apoptosis of the T cell. This is a similar mechanism to negative selection in thymus and is called peripheral clonal deletion. In a few cases the cell is not deleted, but rendered unresponsive, a state referred to as T cell anergy. Upon later encounter with the antigen anergized T cells cannot be activated even when the proper co-stimulatory signals are present (47).

Furthermore, many self components are so scarce that the antigens are rarely presented by DC, and the chance of being detected by a T cell is minute. Because the antigens are present, but do not engage TCR and provoke a response, the immune system is said to ignore these antigens. Some self antigens are never detected by T cells because they only exist in certain anatomical sites called immunologically privileged sites. Main examples are the anterior chamber of the eye, the testis and the central nervous system. Even foreign antigens accessing these sites do not trigger an immune response. The privileged sites are usually protected by a barrier reducing infiltration of cells and large molecules. Furthermore, some cells in privileged sites express Fas ligand (FasL), which can bind Fas receptor on T cells. The ligation of Fas and FasL induce apoptosis of the T cell. In addition, the privileged sites have minimal lymphatic drainage; few tissue specific antigens will drain to lymph nodes to be presented there, resulting in little or no activation of T cells (48,49).

Subsets of helper T cells and immune deviation

Upon interaction with activated antigen-presenting cells such as DC, T helper (Th) cells can differentiate into a variety of effector subsets. The T helper cells have traditionally been divided in two subsets depending on the cytokine profile and effector functions, first described by Mosmann and Coffman in the mid-eighties (50). T helper subset 1 (Th1) produce a cytokine profile characterized by the production of interferon-γ (IFNγ) and IL-

2. In general, Th1 responses are protective against infections of intracellular pathogens. In contrast, Th2 subset produces IL-4, IL-5, IL-10, and IL-13, promoting protection against extracellular parasites, including helminths. The two subsets counter-regulates each other as IFNγ inhibit differentiation of Th2 cells and IL-4 inhibits differentiation of Th1 cells (51). The differentiation to either subsets are called polarisation. As a strong Th1 response may evoke a chronic inflammation leading to tissue damage and autoimmunity, skewing an immune response towards a Th2 response can avoid autoimmunity. A shift in Th1/Th2 response is called immune deviation and was described as a mechanism of non-deletional T cell tolerance (52).

In addition to the classical T helper cell subsets and regulatory T cells, new subsets have been proposed, based on distinct cytokine production. Only the subset called Th17, however, is generally accepted as a genuine Th subset. Th17 cells secrete the pro-inflammatory cytokine IL-17A (IL-17). Th17 cells are involved in clearance of pathogens, especially at mucosal surfaces, and are also associated with the pathogenesis of experimental autoimmune diseases and human autoimmune diseases. Tissues with chronic inflammation in Multiple Sclerosis (MS), RA and psoriasis patients are infiltrated by a high number of Th17 cells (51,53,54).

It appears that Th1 and Th2 response restrict Th17 polarization, but evidence of the opposite is, as of yet, not found. As aberrant Th1 and Th17 responses are associated with autoimmunity and Th2 responses with allergy and asthma, the balance between the subsets needs to be under tight control. It is now recognized that a population of T cells called regulatory T cells plays a major role in the modulation of the Th1, Th2 and Th17 responses and tolerance induction.

Regulatory T cells

The role of regulatory T cells (Treg) is emerging as perhaps the most important mechanism for peripheral tolerance. Tregs are specialized suppressor cells that actively suppress the activation, differentiation and proliferation of T cells as well as the effector functions of activated T cells. Tregs may also modulate the activity and function of other regulatory cells in particular dendritic cells (55,56). Naturally arising Tregs ($CD4^+CD25^+foxp3^+$) are thought to derive from autoreactive thymocyte precursors and

complete their functional maturation in the thymus. These Tregs are more readily activated upon antigen exposure and exert dominant control over naïve T cells in the periphery (56). Although Tregs are activated by specific antigen through the TCR, the Treg can suppress effector T cells with many other antigen specificities, a feature called bystander suppression.

The importance of Treg in tolerance induction is demonstrated in several rodent models. Depletion of Treg in rodents leads to spontaneous development of autoimmune diseases and the reconstitution of Treg prevents autoimmune symptoms (57-60). Humans with mutation in *FOXP3*, the main transcription factor controlling the development of Treg, develop multisystem autoimmune diseases (61).

The mechanisms of Treg suppression are complex and not very well defined. Tregs may mediate suppression through direct cell contact with effector T cells by the receptors CD39, CD73 and LAG-3 and through production of inhibitory cytokines, in particular IL-10 and TGFβ (62). It is also proposed that Tregs can mediate suppression by cytolysis of effector T cells by means of granzyme and perforin and by disruption of metabolic pathways, i.e adsorbing IL-2 needed for T-cell survival (55,62). Finally Tregs may kill DC necessary for T-cell activation, by means of granzyme and perforin or down-modulate DC in a CTLA-4 dependent mechanism (62).

Self /non-self and danger signals

The first postulate of tolerance was made by Burnet in “The production of Antibodies” from 1949. To explain the lack of an antibody response in animals exposed to foreign antigen during fetal life, as observed by Ray Owen, Burnet proposed that self/non-self discrimination is established during embryonic development. He later formulated the clonal selection theory, the deletion of self reacting clones of B cells during embryonic development. Experimentally induced tolerance was first described in 1953 by Medawar, when he demonstrated that tolerance to skin grafts can be induced by inoculation of living cells from strain A into foetal mice of the inbred CBA strain (63). In 1960 Burnet and Medawar were jointly awarded the Noble prize in medicine and physiology for their discovery of tolerance. To date the concept of tolerance largely builds on the clonal selection of B and T cells, although some modifications have been made, i.e tolerance is

not induced during embryonic development only, but rather during thymocyte and B cell development.

The role of central tolerance for the body's ability to distinguish between self and non-self has been challenged by several immunologists. The most prominent alternative theory is the “danger model” championed by Polly Matzinger. The danger model proposes that distressed or injured cells release danger signals, and these danger signals can be detected by various PRR leading to activation of the APC and initiation of an innate response. Activated APC (DC in particular) will secrete activating cytokines and up-regulate co-stimulatory molecules, facilitating the activation of helper T cells. In the absence of danger signals APC will not be activated and will fail to activate helper T cells, auto-reactive or not, thus sustaining functional tolerance. Healthy tissues, free of injury or infection and accordingly danger signals, may constantly induce tolerance to self antigens in this way (64).

Cytokines and inflammation

Cytokines are small signalling molecules with a crucial role in the differentiation and regulation of immune cells. Cytokines have distinct properties, e.g. pro-inflammatory and anti-inflammatory, and are produced and secreted upon stimuli. Depending on stimulation, activated APC will produce different cytokines and the local cytokine profile will drive the polarization of T cells and hence the immune response. Moreover, the cytokines act in a cascade like fashion, where one cytokine stimulate the production of others that will reinforce the reaction.

As mentioned above, a Th1 response is recognised as a pro-inflammatory response characterized by the classical pro-inflammatory cytokines IFNγ, tumour necrosis factor-α (TNFα) and IL-12. IFNγ is important in activating macrophages, increasing phagocytosis and antigen processing. TNFα is a potent chemo-attractant for neutrophils and stimulates phagocytosis by macrophages. IFNγ and IL-12 are positive feedback cytokines polarizing naïve Th cells to Th1 cells. In addition IL-12 enhances the cytotoxicity of NK and $CD8^+$ T cells (51).

Th17 cells secrete the pro-inflammatory cytokine IL-17A (IL-17) and IL-17F as well as TNFα and IL-6 and IL-21. IL-17 is a potent inflammatory cytokine, inducing the

expression of other inflammatory cytokines like IL-6 and IL-8 (also known as CXCL8). IL-17F binds to the same receptors as IL-17, indicating that they have similar functions, however IL-17 has a much higher affinity. Both IL-17 and IL-17F recruit and activate neutrophils. IL-21 stimulates Th17 differentiation, and thus serves as a positive feedback cytokine to maintain and amplify the Th17 subset (51,53,65).

The feedback mechanism to maintain and amplify the different subsets may lead to an unresolved inflammatory response causing chronic inflammation which may trigger tissue damage and autoimmunity. Elevated levels of pro-inflammatory cytokines are linked to various autoimmune diseases, e.g. high levels of IL-1β, IL-17 and TNFα are reported in RA patients and IL-17 and IL-6 are highly expressed in MS patients (51,53).

Autoimmunity

The many autoimmune diseases bear evidence that the mechanisms for ensuring self-tolerance often fail. These may be divided into single organ or systemic, the latter affecting multiple organs and tissues. Here I shall concentrate on human RA and an experimental animal model, oil induced arthritis (OIA), that we consider relevant to RA and which has been the main focus of the experimental work of this thesis.

Rheumatoid Arthritis (RA)

RA is a chronic inflammatory disease characterized by symmetric inflammation of peripheral joints. Infiltration of inflammatory cells in the joint synovium leads to cartilage destruction and bone erosion, usually causing severe pain and decreased mobility of affected joins. In the human RA can be divided into different clinical subtypes. The most prevalent of them has a late onset, usually at age 50-60 and is about three times more common in females than males (66,67). RA patients often develop cardiovascular diseases and have a 10-15 years shorter life expectancy than healthy individuals (68,69). The prevalence in the Caucasian population range from 0,5-1,0% and is somewhat lower in the Asian population (0,2-0,3%) (70).

Previously the diagnosis of RA was primarily based on clinical manifestations, leading to diagnosis of patients with advanced arthritis only. The most widely used serologic marker is the serum rheumatoid factor (RF), an autoantibody against the Fc portion of the

immunoglobulin G molecule. However it is only present in about 70% of RA patients, and it is not RA specific. It is also present in other autoimmune diseases such as Sjögren's syndrome as well as in up 10% of the healthy population (71). Additional molecular markers for RA have been identified; however, in general they are not present in all RA patients and may be present in other patient groups as well as in healthy individuals. They therefore serve more as disease indicators rather than as disease specific diagnostic markers. The antiperinuclear factor is highly RA specific, but has not been used as routine test due to laborious technique and lack of standardization (72,73).

Over the last decade autoantibodies against cyclic citrullinated peptides, anti -CCP, has been established as a diagnostic marker of RA. Citrulline is an amino acid formed by post-transcriptional modification of arginine residues by the enzyme peptidyl arginine deiminase (PAD) and has been linked to autoimmunity (74). Anti-CCP is present in about 80% of all RA patients, and is also present in the early stages of the disease, making it at better diagnostic marker than the RF factor (73,75-77).

There is extensive variation between RA patients with regard to time of disease onset, disease severity, progressiveness and response to treatments. The clinical heterogeneity reflects the fact that multiple genes as well as environmental factors influence precipitation of disease. The identification of these factors will be central for diagnostic, prognostic, therapeutic and preventive purposes.

Environmental risk factors

Due to the higher incidence of arthritis in females, the level of sex-hormones has been argued to influence RA susceptibility. There are indications that pregnancy and use of contraceptives prevents disease or leads to disease remission. However the data are conflicting, and may be linked to postponing the disease rather than preventing it (78,79).

Infections have been proposed as triggers of autoimmunity, however reports are inconclusive (70). Studies of animal models have demonstrated that bacterial and viral infections can trigger autoimmune responses (80,81). Several studies have reported an association of cigarette smoking and RA (82-89). There is an increased risk of developing

RA in heavy smokers (85,86,89,90). Finally, smoking is a risk factor for developing sero-positive RA (88,90).

Reports of autoimmunity and development of autoimmune symptoms following silicone implants, long term mineral oil administration, subcutaneous injections of adjuvants and accidental injections of vaccines have lead to the proposal of human adjuvant disease (HAD) (91-102). Recent reports have confirmed occupational exposure to hydraulic oils and silica as risk factors for RA (103,104). Risk factors for RA may be shared with other autoimmune diseases, indicating common pathogenic pathways. Smoking and exposure to smoking is confirmed as a risk factor for multiple sclerosis (105,106). Overt thyroid disease and systemic lupus erythematosus (SLE) has also been associated with smoking (107,108). Recently the clustering of SLE patients exposed to oil-waste reservoir was published, indicating an association between SLE and oil exposure (109). Exposure to silica is associated with an increased risk of developing SLE and anti-neutrophil cytoplasmic autoantibody-associated glomerulonephritis (110-112).

Genetic risk factors

In diseases where several genetic regions contribute to the development, severity and progression of the disease, the regions influence the phenotype of the disease in a quantitative manner. Such genetic regions are called quantitative trait loci (QTL). Identification of the QTL and the underlying genes is crucial to understand the pathogenic pathway.

Genetic mapping of QTL

There are two main approaches to identify QTL; linkage analysis and association studies. Both methods are similar in that they use genetic variation to mark genomic loci and then attempt to detect co-segregation of marker and disease. The most frequently used DNA marker is single nucleotide polymorphism, SNP i.e. a single base substitution or insertion/deletion.

Linkage analysis is based on genetic linkage between two separate genetic loci, i.e. higher than random chance of joint inheritance of alleles from the two loci (113). Alleles of loci

on different chromosomes will not be co-inherited due to random segregation of chromosomes. Alleles of loci in close proximity on the same chromosome tend to be co-inherited. They can, however, be separated due to genetic recombination, a process during meiosis where parts of the chromosome are exchanged between two homologous chromosomes. The likelihood for recombination increases with physical distance between the loci, but not linearly, as recombination does not happen at random, but at certain "hotspots" along the chromosome (114,115). Between such "hotspots" little or no recombination occurs, and the alleles of the loci will be inherited together in blocks. The combination of alleles (measured by SNPs) along such block is called a haplotype, consequently the blocks are referred to as haplotype blocks (116). In linkage studies it is tested whether a disease trait (QTL) is coinherited with a DNA marker in related subjects. The closer the SNP is located to the QTL, the more likely it is that the marker will co-segregate with disease phenotype (113).

Genetic association studies aim to detect association between one or more genetic polymorphisms and a trait in a population. It is tested if the pattern of markers for a given locus is at a different frequency in cases and matched controls (117). Association is based on linkage disequilibrium (LD) which refers to the non-random association of alleles at two or more loci. If the LD is high, the alleles are in *dis*equilibrium to the assumption of random distribution of the alleles. Every time recombination occurs between two loci in the population, the linkage disequilibrium between them is weakened. Linkage spans generally over longer regions than LD, as there is a greater chance of random recombination events in unrelated individuals. Because LD spans over shorter regions than linkage, association analysis potentially has a higher resolution than linkage studies, but requires many more markers to be examined (117,118).

Considerations in linkage and association studies

Success in identification of QTL and the underlying gene by linkage studies has been limited to loci for monogenic diseases, like cystic fibrosis and Huntington's disease (119,120). Linkage studies have limited power to detect QTL in complex diseases due to weak effect of individual QTL. Association studies have more theoretical power than linkage studies to identify QTL with modest risk (121). Aassociation studies have

identified several loci in complex diseases both within and outside MHC, however, many of them have failed to be replicated, mainly due to underpowered studies (122).

A major factor of statistical power is sample size. Linkage studies are limited by the difficulty to get large enough sample size to generate the necessary power to detect linkage. There are simply not enough cases. In order to detect linkage to a marker in complex disease, the minimum number of affected families should be several thousands (121,123), an unfeasibly large family material. Association studies do not require the vast sample size as linkage studies; however practical concerns in collection and handling samples as well as cost have hampered large studies. Recent development in genotyping technology and the corresponding drop in cost have reduced this limitation of sample size.

Sample size is not the only factor influencing statistical power. The frequency of the allele in the population and the magnitude of the genetic effect it gives also determine the power. A disease susceptibility allele with low frequency in the population, but with strong genetic effect requires a smaller sample size than an allele with high frequency and minor genetic effect in order to generate sufficient power to claim association (116,121,124).

An important consideration of the power in a study is the correction for multiple-hypothesis testing. Each SNP tested will represent one hypothesis, and in Genome-Wide association studies (GWAS) thousands to hundred thousands markers will be tested and some of them will associate with a QTL just by chance. Correction for multiple-hypothesis testing will reduce the number of false positive result, but may also result in failure to detect a true positive result (116,125).

A single QTL may harbour several independent mutations. These can be common variants with weak effect or rare variants with intermediate to strong effect. As examples, rare independent mutations in salt handling genes contribute to blood pressure variation (126), several alleles that independently influence development of prostate cancer have been identified at chromosome 8q24 and three functional variants of IFN regulatory factor 5 are independent risk factors for SLE (127,128). Independent mutations are likely to distribute randomly to a reference SNP, thereby cancelling each other and obscuring genetic association to the reference SNP. Genetic mapping in recent founder populations

such as the Tristanians (Tristan de Cunha islanders), the Icelandic and the Finns have proven successful for identification of rare mutations. Due to isolation the population is made up of distantly related individuals (extended family). A mutation in the founder population will be amplified to a high proportion during the expansion of the population, giving the disease allele a high frequency (129-131). In extended families the recombination frequency will be lower than in a more mixed population, resulting in high LD over longer regions (i.e larger haplotype blocks) facilitating association studies (118).

Standardization of diagnostic and phenotypic criteria is easier in isolated populations, thus diagnostic errors are reduced. Isolated populations often share common environment and culture. This reduces background noise from environmental risk factors compared with studies performed in non-isolated populations (131). The frequencies of alleles vary across human populations, which may cause difficulty in replicating association of SNPs to diseases among different ethnicities (132). A SNP found to associate with prostate cancer at chromosome 8q24 is at a higher frequency in African American, giving better significance of association in African Americans than in Europeans (127).

Although genetic mapping has lead to the discovery of QTL for several complex diseases, few associated genes and causal mutations have so far been identified in the human. Re-sequencing regions where associations are found is necessary to identify causal mutations.

Genetic mapping of RA QTL

Several genome scans have been performed to identify QTL involved in RA susceptibility. In the period 1998 - 2002, four genome scans performed in Japan, USA, France and UK were published (133-136). They all show significant linkage to MHC, but no other genetic region gave significant linkage. Lander proposed the term suggestive linkage for loci with $p<0,05$ in whole genome scans. Linkage cannot be claimed, but should be reported with intent to confirm the linkage in an extension study (137). By accepting a less strict significance, loci with suggestive linkage were reported, many of the suggestive loci revealed were found in more than one scan. However, extension studies of the scans in the US, UK and in France gave inconsistent results. The US and UK extension scan did not confirm the previously identified loci, but identified new loci with suggestive linkage (138,139). The French scan confirmed 10 out of 22 previously reported loci as well as identifying new suggestive loci (140). In order to increase sample

size and statistical power, meta-analyses have been performed for RA by compilation of available genome scans (141,142). These two meta-analyses are based on partly the same genome scans and both reveal several loci linked to RA susceptibility, seven of which are found in both studies. Of interest to this thesis is that both meta-analyses identify a locus with cytogenic location 12p12-13, a region syntenic to the region of rat chromosome 4 containing the QTL *Oia2.*

RA gene candidates in the MHC

For a long time the only region conclusively linked to RA was the MHC or human leukocyte antigen (HLA) on human chromosome 6 (143,144). The most important candidate genes in this complex are the HLA class II genes. The association is strongest for HLA DRB1 alleles that encode a similar amino acid sequence, termed the shared epitope. The sequence is five amino acids long and resides in the hypervariable region of the DRβ1 chain. These residues form one side of the antigen presenting binding site. The shared epitope hypothesis postulates that the shared epitope is involved in the pathogenesis of RA by presentation of arthritogenic antigens. The association of HLA DRB1 alleles with shared epitope (alleles *0101, *0102, *0404, *0405, *0408, *1001 and *1402) is linked to higher risk for disease development as well as severity (145).

RA gene candidates outside the MHC

In the past few years, a few gene candidates for RA have been reported. In 2004 Begovich *et al.* published the association of *PTPN22* to RA (146). The gene was found through candidate gene study after a locus was identified in a US genome-wide association study. Association was replicated in affected families study, and has further been replicated in caucausian populations by other groups (147-154) but not in Japanese and Korean populations (155,156). *PTPN22* is however associated with RA in South Asians in the UK and in a Colombian population (157,158). *PTPN22* encodes a cytoplasmic protein tyrosin phosphatase involved in negative regulation of T cell activation. Begovich *et al.* reported a missense single nucleotide polymorphism (C1858T) that altered the proteins function, leading to T cells with a lower threshold for activation (146).

A Japanese genome scan reported a locus spanning over a region containing all known genes encoding peptidylarginine deiminase, PAD. In a Japanese case control study using SNP markers in the PAD genes 1-4, association was found to variants of four SNPs in

PADI4 (159). Given the fact that auto-antibodies against citrullinated proteins are present in most RA patients, an association to a citrullinating enzyme is interesting. Association to *PADI4* was replicated in a Japanese study (160) and association has also been found in Chinese and Korean populations (161,162). However, association to *PADI4* in Caucasians has been conflicting; no association were found in a French, UK or Hungarian studies (163-165) and Plenge et al. reported association in a north American cohort, but not in a Swedish (166). In another UK study a two-marker haplotype for *PADI4* was weakly associated with RA, and *PADI4* were over-expressed in cases compared to controls (167).

Another candidate gene originally found in Japanese studies is the Fc receptor-like 3 gene. A SNP in the promoter region is associated with RA, SLE and autoimmune thyroid disease (168). This SNP has also been associated with RA in Caucasians (169-171). Other candidate genes reported are the T cell immunoglobulin and mucin-domain-containg receptor -1(*TIM1*), cytotoxic T lymphocyte antigen-4 (*CTLA4*), solute carrier 22A4 (*SLC22A4*), tumor necrocis factor receptor associated factor 1 (*TRAF1*), Interleukin 1β (*IL1b*) and the programmed death-1 gene (*PD1*) (172-178).

Experimental animal models for autoimmunity

Another approach to obtain gene candidates is to develop experimental autoimmune animal models and perform genome scanning to detect QTLs, and see whether the syntenic regions in the human is associated with the disease in question. Animal models provide a tool to control both environmental factors as well as the genetics. One of the great advantages is the possibility to manipulate the animal genome either through breeding (congenic strains and advanced intercrossed lines) or advanced genome technology (transgenic or knock-out strains) (179,180).

A number of different animal models exist for many different autoimmune diseases and other diseases like cancer and hypertension as well as infection models. Some inbred strains develop a disease spontaneously, other diseases can be induced. The basis for all experimentally induced models is to find two strains with different phenotype to the same stimulation. In the rat, the Dark Agouti (DA) strain is susceptible to a number of autoimmune diseases; experimental autoimmune encephalomyelitis (EAE), experimental allergic neuritis, experimental allergic uveitis, experimental autoimmune thyroiditis as well as all experimental arthritis models (181-189). Of all strains tested so far, DA is the

only strain that is susceptible to oil induced arthritis (OIA), and in other arthritis models the DA rat consistently display a more severe arthritis than other susceptible strains. Similarly, the inbred rat strain PVG is resistant to the experimental models mentioned above (185,186,188,189). Comparing the genome of the DA and PVG rat strains have contributed in isolating QTLs for several different diseases. For a complex heterogeneous disease like RA, no experimental model will display all the features of RA, but all the different models may contribute in the understanding of the pathogenesis of RA. Central in this thesis is oil induced arthritis in the susceptible DA rat and resistant congenic strains.

Oil induced arthritis (OIA)

OIA is induced by a single intradermal injection of Incomplete Freund's adjuvant (IFA) at the base of the tail. IFA is an adjuvant used in many different immunization protocols to induce experimental inflammatory diseases. It has also been used as an adjuvant in vaccination protocols for humans and livestock. Its purpose as an adjuvant (from Latin: adjuvare; to help) is to stimulate the immune system to raise a strong response against the compound in the vaccine (190). IFA is composed of 85 % mineral oils (Bayol F) and 15% of the emulsifying agent Arlacel A (Mannide monooleat).

OIA is characterized by a mild inflammation in the small joints of the hind paws and ankles 12-15 days post injection in susceptible rats. Sometimes inflammation also progresses to the front paws. The inflammation is monophasic, it spontaneously subsides by day 45, with little or no erosion of joints as reported by Cannon et al and Lorentzen et al (188,191). However, the course and severity of disease can be altered in a different environment. Compared to the mild disease development reported by Cannon et al and Lorentzen et al, DA rats housed in the animal facility in Domus Medica, Oslo, Norway almost always develop relatively severe arthritis in all four paws upon arthritis induction, and inflammation does not subside without treatment with anti-inflammatory drugs.

Adoptive transfer of T cells can transfer disease (192,193). OIA is influenced by genes both within and outside of the MHC, as shown by testing different MHC congenic strains. Compared with the DA strain, the PVG.1AV1 strain develops mild disease with incomplete penetrance. The difference in susceptibility between DA and the MHC

identical PVG.1AV1 can only be explained by influence from genes that do not reside in MHC (194).

Identifying QTLs in animal models

The strategy to identify QTLs in animal models is very much the same as identifying QTLs for disease in humans, but with one important difference: the control over genetics. A genetic linkage map is created by genome scanning of a population of rats produced by crossing different stains with a different phenotype to a challenge, in our lab we used DA (susceptible to OIA) and PVG (resistant to OIA). After a DAxPVG cross, the heterozygous F1 generation is intercrossed to produce an F2 generation. Due to Mendelian law of random segregation of chromosomes, any individual in the F2 generation will be a mix of homozygous DA (DD), heterozygous DA/PVG (D/P) or homozygous PVG (PP). Adding to this, there will be recombination events. A QTL can be mapped by testing the F2 population for the phenotype; susceptibility, and compare with the genotypes. If a genotype for any of the markers is co-inherited with susceptibility, the marker is linked to disease. In OIA, three QTLs were found by genome scanning of an F2 population: *Oia1* within the MHC on chromosome 20, *Oia2* on chromosome 4 and *Oia3* on chromosome 10 (195).

Congenic strains for fine mapping

The identification of a QTL from a F2 linkage map is crude. One strategy to fine map the QTL interval is to generate congenic strains, here illustrated with the identification of loci for OIA using DA (susceptible) and PVG (resistant) as examples. (Fig.3). To generate an *Oia2* congenic strain, F1 offspring from a cross between DA and PVG are backcrossed with the DA strain. The F2 generation is genotyped and individuals heterozygous for markers in the *Oia2* QTL are selected for further breeding and backcrossed with DA. This selective breeding and backcrossing is continued for at least 10 generations to get rid of undesired PVG genomic background. After the N generation, crossing of two heterozygous congenic animals will give rise to homozygous offspring which have the PVG alleles in the *Oia2*, but the rest of the genome is DA. The *Oia2* congenic rats are resistant to oil induced arthritis, whereas DA rats are not. Continued breeding of the congenic strain in heterozygous form has an advantage due to the potential recombination event in the congenic interval. Further backcrossing into the parental DA strain may

produce intra-congenic recombinants with a contracted congenic region compared to the original congenic strain, thus providing a more precise mapping of the QTL.

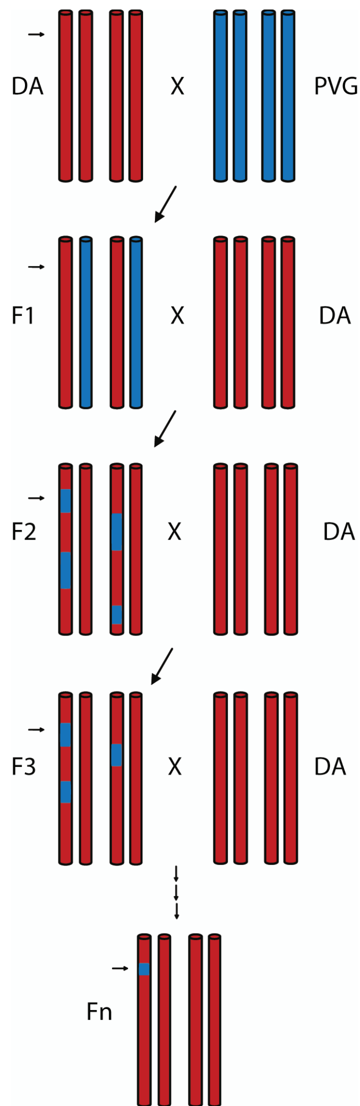

Figure 3. Generation of a congenic strain through marker selected breeding.
Two inbred rat strains are crossed to produce a F1 heterozygous population. The F1 are then backcrossed with one of the parental strains, in this case DA. Due to recombination events, PVG chromosomal regions are introduced on the DA background. Offspring heterozygous in the region of interest, *Oia2*, marked with an arrow, are further backcrossed with DA for n generations. In the end crossing offspring heterozygous for the *Oia2* locus will yield a strain with PVG genotype in the *Oia2* locus on DA background.

A narrower congenic interval will ease the search for gene candidates, as there will be fewer genes to test. Breeding of congenic strains has proven successful in narrowing down the gene candidates behind QTL in various experimental models. Previously *Ncf1* has been identified as an arthritis severity regulating gene in the rat model pristane induced arthritis (PIA) by breeding of congenic and intra-congenic recombinant strains (196). The congenic interval was narrowed to a segment of approximately 300 kb, harbouring only two genes. Sequence analysis revealed polymorphisms in the *Ncf1* gene. The polymorphism in the *Ncf1* gene has a functional outcome. The variant in the arthritis prone DA rat leads to reduced oxidative burst compared to the variant in the resistant congenic rat (196). Furthermore, in the mouse a mutation in *Ncf1* leads to an alternatively spliced mRNA and a truncated protein with loss of function. This mutation was backcrossed into another strain, allowing testing for experimental autoimmune models. Congenic mice with the mutated *Ncf1* show increased severity of collagen induced arthritis as compared with mice with normal *Ncf1* (197). In addition to high resolution mapping, the generation of congenic strains is useful as it makes a polygenic disease monogenic, which will help unraveling the pathogenic pathway of a disease contributing gene, without the contaminating effect from other QTLs (180,198).

An alternative way to fine-map QTL is to generate advanced intercross lines, AIL. Generation of AIL is not based on marker selected breeding couples, but controlled random intercross for every generation. The genetic recombination events in every intercross will accumulate for every generation. It is important to keep a large number of breeding pair to avoid brother sister mating to prevent homozygous chromosomes (hence the term controlled random intercross). After *n* generations the accumulated recombinations will result in a mosaic genome between the two parental strains. A large number of generations will provide a high resolution linkage map (179). In addition to the possibility to fine-map QTLs, AIL is well suited to study gene-gene interactions, such as epistasis, a phenomenon where one gene may influence the phenotypic expression of another. A partial AIL revealed epistatic interactions and sub-division of one locus into three loci in a mouse EAE model (199).

From isolated QTL to structural gene

Theoretically the structural gene behind a QTL can be identified by the continued production of intra-congenic strains with progressively narrower congenic region until the congenic region spans only over a single gene. There are practical limitations to this approach. The generation of intracongenic strains is dependent on recombination event within the congenic region. If the congenic region is narrow, there will be less chance for a recombination in the desired region. This necessitates a vast number of breeding pairs and increased housing expenses. The detection of a recombinant may also be hampered by lack of a genetic marker that is polymorphic between the parental strains.

Previously, genes in the congenic region had to be identified by cumbersome methods like chromosome walking or exon trapping. The release of the mouse and rat genome sequence has greatly eased the search for candidate genes. With the sequences of most or all genes in the congenic region available, together with improved software to predict protein function, researchers could make a qualified guess about gene candidates. For identification, the candidate genes have to be tested for features that can explain the difference in phenotype. This includes identification of possible differences in sequences, splice patterns, expression levels and functional properties of the encoded proteins.

From animal model to human

Through precise mapping of a QTL, the syntenic region in the human can be tested for association to the disease. Alternatively, if the QTL mapping in an animal model resulted in a concrete gene candidate, the orthologoue gene and closely related genes in the syntenic region can be tested for association. The animal model serves as a search light, identifying interesting regions/genes relevant for human disease. In 2001 Barton et al reported a linkage study on human chromosomal regions syntenic to five QTL in the rat that regulate experimental arthritis (200). Linkage was reported to human 17q21-25, syntenic to *Oia3*. The other QTLs; *Oia2*, *Pia4* (pristane induced arthritis locus 4), *Cia4* (collagen induced arthritis locus 4) and *Aia2* (adjuvant induced arthritis locus 2) did not show any linkage in this study, however the region 12p12-11 syntenic to *Oia2* locus show linkage in two meta analyses reported by Fisher *et al* and Choi *et al* (141,142).

Aims of the study

At the start of my study, Lorentzen at Karolinska Institutet, Stockholm, had just localized *Oia2* to the distal part of chromosome 4. In Oslo we had mapped the NK cell gene complex (NKC) to this genetic region. The work was based on genetic analyses of back-crossed progeny, where we found a crossing-over frequency of 0.4 %, i.e. < 0.5 cM, between the genes flanking the complex. This was identical to the crossing-over frequency described for the mouse NKC by Wayne Yokoyama, suggesting that the size of the complex and number of genes within would be relatively modest (201). We therefore embarked on making a total inventory of the rat NKC, based on a PAC library from the Wellcome Trust Centre for Human Genetics, Oxford. The work turned out to be more difficult than anticipated, as the gene complex proved to be much larger than predicted from the genetic map. We were subsequently steamrolled by the rat genome initiative, whereby the complete genome was sequenced with, for a mammalian genome, unprecedented rapidity.

The NK cells are potent immunoregulatory cells through their production of cytokines, thus the NKC was a likely candidate region for harbouring the *Oia2* locus, which revived our interest in complete characterization of the complex. The interest subsequently faded with the discovery that *Oia2* mapped outside the NKC (paper II). However, as incomplete and partly erroneous reports of the gene content of the rat NKC later were published by others, based on *in silico* gene predictions from the released genomic sequence, we considered it worthwhile to sum up our work on the characterization of the rat NKC, which includes the cDNA cloning of the majority of the predicted CTLR genes and analyses of their expression patterns (paper I). The work also clarifies why we consider it useful to categorize the genetic region containing newly detected related genes just upstream of the NKC (the APLEC) as a separate gene complex (paper III).

As a preparation for identifying functional loci linked to the NKC, we had made congenic strains for this region, based on the two parental strains, PVG and DA. Coincidentally, the two strains were the same that Lorentzen had used in his studies of oil induced arthritis. The idea that the *Oia2* locus was linked to the NKC was therefore easy to test. The first congenic strains produced in Oslo did indeed confirm linkage between *Oia2* and the NKC. However, by narrowing the transferred region through continued back-crossing

the genes in the NKC had to be excluded. A novel recombinant strain, R2, back-crossed from the OIA resistant strain R1, turned out to be OIA susceptible, although the whole NKC of R2, like R1, exhibited the PVG haplotype. Genotyping showed that the haplotype block transferred from the PVG strain had been shortened on the centromeric side of the NKC, thus positioning the QTL to this region. Incidentally, at the same time genes encoding novel CTLR receptors were reported to map within or close to the NKC in the mouse and the human. However, they exhibited different cellular expression patterns from most NKC encoded receptors. Supported by our crude rat maps, it was clear that if the novel identified genes had rat orthologues, they would be positioned proximal to the rat NKC, in the middle of the chromosome region where we had mapped *Oia2*. Based on deduced cellular expression patterns and sequence features, the novel receptors seemed to be involved in the regulation of activation of professional antigen presenting cells, including DC. They therefore seemed to be ideal candidate genes for *Oia2*. Due to limitation of resources we divided the working tasks between Stockholm and Oslo, so that they in Stockholm continued back-crossing of the DA.C4^{PVG} congenic strains to further narrow the region containing the *Oia2* gene, whereas we in Oslo should characterize the genetic region containing the novel genes, which we called APLEC.

Both working tasks were successful. By the generation of a series of recombinant strains in Stockholm (paper II and paper IV), the *Oia2* was conclusively mapped to the novel gene complex. By then we had characterized the APLEC both in the rat, mouse and human. In addition to cDNA cloning of novel genes in all three species, we had sequenced and compared gene expression patterns of the genes in the OIA susceptible DA strain and the two resistant strains PVG and LEW. Interestingly, the DA strain turned out to be a natural knock-out strain for one of the genes, *Dcar1* (dendritic cell activating receptor gene 1).

The conclusive mapping of the *Oia2* to the APLEC represented an important milestone, the end of a long search and the starting point of novel studies, fanning out in different directions. One of these is to see whether the homologous gene complex is associated with autoimmune diseases in the human (paper IV), the other to try to pinpoint the roles of the APLEC receptors in the pathogenic mechanisms leading from application of oil to the skin to development of inflammation in distant joints.

Summary of specific aims

1) Characterize the Natural Killer gene complex, NKC, in terms of gene content and gene expression programs.

2) Fine-map the chromosomal position of *Oia2*.

3) Characterize the antigen presenting cell lectin-like receptor gene complex, APLEC, in terms of gene content and gene expression programs.

4) Identify the structural gene within APLEC associated with OIA.

5) Identify pathogenic mechanisms leading from application of oil to the skin to development of inflammation in distant joints.

Summary of results

Paper I

The complete inventory of genes in the rat natural killer cell gene complex.

The rat NKC, which spans ~ 3.3 million base pairs (Mb), can be divided into four regions. Flanking the complex on the centromeric side is the ~ 0.8 Mb region I containing four genes encoding NKR-P1 receptors and 10 genes (plus a gene fragment) encoding their CLR counter-receptors. On the telomeric side is the large (~1.8 Mb) region IV containing altogether 34 *Ly49* genes (including several pseudogenes) and the *Klrh1* gene. Just centromeric to the Ly49 region is the much smaller (~ 0.3 Mb) region III containing eight genes encoding the CD94/NKG2, KLRE/KLRI and KLRK1 (NKG2D) receptors. The cDNA cloning and expression analyses of the genes from these three regions of the rat NKC have previously been reported elsewhere. Intercalated between the NKRP1/CLR and CD94/NKG2 clusters is a 0.3 Mb region which contains the *Lox1* gene encoding the oxidized LDL receptor, previously described by others. *In silico* homology screening against BAC clones and the whole genome sequence (www.ensembl.org/Rattus_norvegicus) with mouse and rat sequences as probes revealed eight other genes in region II: *Cd69, Clec2m, Clec12a, Clec12b, Clec1b, Clec9a, Clec1a* and *Dectin1(Clec7a)*, which were all cDNA cloned. Apart from the CLR genes, the genes encoded by regions I, III and IV are expressed predominantly by NK cells. In contrast, the genes within region II are expressed by other cell types, in particular myeloid, but also endothelial and testicular cells. With a total gene content of 67 genes predicted to encode group V CTL receptors, the rat NKC represents one of the largest known mammalian gene complexes consisting exclusively of paralogous genes.

Paper II

High resolution mapping of an arthritis susceptibility locus on rat chromosome 4, and characterization of regulated phenotypes.

A genome scan had previously mapped the QTL *Oia2* to the distal part of rat chromosome 4, including the natural killer cell gene complex (NKC). In this paper we fine map *Oia2* by testing a series of DA.C4^{PVG} congenic strains generated through marker assisted transfer of parts of the distal chromosome 4 from the OIA resistant PVG strain to the genetic background of the susceptible DA strain. The first results from Oslo showed that

Oia2 must reside just proximal to the NKC, but outside this gene complex. Subsequent analyses of novel recombinant strains in Stockholm reduced the genomic region containing the QTL to approximately 1.2 Mb, containing a limited number of genes. The region included the CTLR genes in the newly discovered APLEC gene complex.

Paper III

Identification of lectin-like receptors expressed by antigen-presenting cells and neutrophils and their mapping to a novel gene complex.

Independent recent reports from the mouse and the human had revealed genes encoding novel CTLR preferentially expressed by professional APC and were variably suggested to lie close to or within the NKC. From our work on the rat NKC we considered it more likely that putative rat orthologues were localized together and outside the NKC, within the region where we had mapped *Oia2*. We therefore considered these genes strong *Oia2* candidates and decided to identify the related genes in the rat. Although the work started before the genomic sequences of the mouse and the rat were released, we succeeded in cDNA cloning all the rat orthologues plus novel related genes, altogether eight genes predicted to encode type II CTL receptors, named macrophage inducible C-type lectin (*Mincle*), macrophage C-type lectin (*Mcl*), dendritic cell activating receptor (*Dcar1*) and dendritic cell inhibitory receptors 1 to 4 (*Dcir1-4*). The rat orthologue of dendritic cell associated lectin-2 (*Dectin2*), turned out to be a pseudogene (rat *Dectin2p*). We also reported the cDNA cloning of two related novel human genes, *DECTIN2* and *MCL,* and five novel mouse genes: *Dcir2-4*, *Dcar1* and *Dcar2*[1]. We showed that the genes mapped together ~ 6 Mb proximal to the NKC, differed from the KLR encoded by the NKC by the preservation of amino acids involved in saccharide binding, hence their classification as group II in the C-type lectin super family, and that they were preferentially expressed by professional APC (dendritic cells, macrophages and B cells) plus neutrophils. Based on sequence, expression and chromosomal mapping analyses we proposed that the genes defined a novel, evolutionary conserved gene complex, which we named the antigen presenting cell lectin-like receptor gene complex, APLEC.

---

[1] We originally submitted the mouse and rat *Dcar* genes as *Aplra1* and *-2*, but changed nomenclature when mouse *Dcar2* was reported by others just before our paper was published.

Paper IV

Association of arthritis with a gene complex encoding C-type lectin-like receptors

The aim with this work was to pinpoint the structural gene associated with *Oia2* and to see whether the syntenic chromosomal region is associated to RA in the human. Continued back-crossing of DA.C4^{PVG} congenics with DA eventually produced the OIA resistant recombinant strain, R17 (17th generation of back-crossing). Here, the chromosomal interval containing the protective PVG haplotype block was fully contained in the region flanked on the outside by the genetic markers D4Kir12 and D4Kir55. The distance between these two markers is only 0,8Mb. The only detectable genes in this region were the APLEC genes *Mincle*, *Mcl*, *Dcar1* and *Dcir1-3*. Quantitative PCR analyses showed significant expression differences between DA and R17 for *Mincle, Mcl, Dcar1* and *Dcir3.* cDNA sequencing of the strains, revealed a few largely conservative missense mutations in DA versus the OIA resistant PVG and LEW alleles for *Mincle*, *Dcir1* and *Dcir2*. The most striking finding was an early nonsense mutation in *Dcar1* in DA, but not in PVG or LEW. DA thus represents a natural knock-out strain for this gene. The genes in the syntenic human gene complex were tested for association to arthritis by SNP typing of ~ 2 000 RA patients and matched controls. Of 35 SNPs tested, the SNP rs1133104, lying within the human *DCIR* gene, showed statistically significant association with odds ratio of 1.27 (95% confidence interval 1.06-1.52; p= 0.0073) in patients negative for antibodies to cyclic citrullinated peptide (anti-CCP). We thus conclude that the gene complex APLEC is associated with arthritis both in the rat and in the human. In the rat, the most likely *Oia2* candidate is *Dcar1*, whereas in the human association to RA was to *DCIR*.

Paper V

Effect of APLEC receptors on cytokine expression in oil induced arthritis (OIA) in rats.

In this work we aimed to identify the role of the candidate genes in the mechanisms leading to inflammation in the joints following application of oil to the skin. Previous adoptive transfer experiments had indicated that the arthritogenic response is initiated in the lymph nodes draining the site of oil injection. Here we measured the expression levels of APLEC genes and selected cytokines by quantitative PCR in draining and non-draining lymph nodes of OIA susceptible DA rats and resistant congenic R17 rats at different time points following skin application of adjuvant oil. The most conspicuous finding was early strong up-regulation of *Dcar1* mRNA in both strains, followed by strong up-regulation of

Il17 in DA, but not in R17 rats. In R17 rats the early surge of *Dcar1* mRNA is likely to result in rapid upregulation of this receptor in the lymph node, whereas in DA the non-sense mutation precludes synthesis of the Dcar1 protein. The temporal link between Dcar1 and IL-17 synthesis suggests causal relationship between the two events. IL-17 is a pro-inflammatory cytokine, produced by the T helper subset referred to as Th17 cells, which has been linked to arthritis and other autoimmune and inflammatory diseases. We therefore propose that the Dcar1 protein is involved in helper T cell polarization, where oil stimulation in the absence of Dcar1 protein skews the response towards Th17 differentiation and IL-17 production, which in turn leads to the subsequent inflammatory response in the joints.

General Discussion

The most important discovery in this work is the identification of the gene complex APLEC and the linkage of this complex to arthritis in both rat and human. I here aim to discuss my results in the context of novel findings.

NKC vs APLEC

In papers I and III we characterize two gene complexes, the NKC and the APLEC, lying fairly close together and consisting of related receptors. In the mouse and rat NKC, the large collection of KLR genes is uninterrupted by other genes, except for single intruding non-CTLR gene, encoding GABA(A) receptor-associated protein like 1, splitting blocks II and III of the NKC. This is also observed in the human NKC. The much smaller APLEC contains nothing but CTLR genes. In both mouse and rat the two complexes are separated by a chromosomal region of 5 - 6 Mb containing a variety of non- C-type lectin superfamily genes. In the human the distance between the complexes is less; approximately ~ 1.2 Mb, also separated by non- C-type lectin superfamily genes. The group II genes cluster together at a distance from the group V genes in all four orders of mammals investigated: dog, cattle, human and rodents (202).

Within the large C-type lectin superfamily, the receptors encoded in the NKC and APLEC are classified as belonging to different subgroups, groups V and II, respectively. Evolutionary the group II C-type lectins seem older as they have retained the Ca^{++}-dependent carbohydrate binding domain, whereas group V do not have a functional form of this domain (8). Group II, but not group V lectins have been described in the teleost fish, *Fugu rubipes* (203). Clusters of group V genes were reported in the teleost *Oreochromis niloticus* (204). However, according to the authors, the *Oreochromis* KLR loci are not true orthologues to the mammalian KLR, but paralogues, signifying that the clusters have developed after the divergence of teleostean and mammalian lineages. Interestingly, a multigene family of immune related lectin-like receptors cloned from *Fugu rubipes* were reported (205). The receptors belong to the group II lectins and Mincle is among the mammalian proteins with greatest homology to these genes, suggesting common ancestry.

APLR function

So far knowledge about the functional roles of the APLR is limited and their ligands are largely unknown. In contrast, the functional properties of several KLR have been extensively studied. As detailed in the introduction, many of the KLR regulate NK cell activation by transducing inhibitory or activating signals upon ligand binding. The signalling properties are linked to characteristic sequence features of the receptors, notably ITIMs in the cytoplasmic tails of the inhibitory receptors, and positively charged amino acids (arginine or lysine) in the TM region of the activating receptors, mediating association with ITAM containing adapter proteins. ITIMs or the presence of arginine in the TM region are seen in some of the APLR, suggesting that they similarly take part in regulation of cell activation. In the case of KLR the lysis of target cells by NK cells permits unequivocal definitions of the terms “activating” and “inhibitory”. For APCs the consequences of activation is less clear-cut and activation can lead to the generation of inhibitory (tolerogenic) DC or regulatory DC that modulate activity. Thus in the following, the terms “activating” and “inhibitory” simply refers to whether the receptors mediate signals through ITAMs (contained in associated adapter proteins) or ITIMs, independent of the final cellular output.

Inhibitory receptors

Human DCIR, as well as mouse and rat Dcir1 and - 2 contain an ITIM in their cytoplasmic tail, pointing to inhibitory roles. Human DCIR has been shown to associate with SHP1 and SHP2, two tyrosine phosphatases that associate with inhibitory receptors in the Ly49 family of KLRs, corroborating that DCIR is an inhibitory receptor (206,207). The observation that human DCIR is down-modulated in activated neutrophils, has led to the proposal that DCIR represses the activation of neutrophils (208). Kanazawa et al. has shown that the ITIM of mouse Dcir1 can inhibit the Ca^{++} mobilization and tyrosin phosphorylation. Following mutation of the tyrosine in the ITIM to phenylalanine, Dcir failed to block Ca^{++} mobilisation and tyrosine phosphorylation (17). Thus human and mouse DCIR act as inhibitory receptors, dependent upon their ITIM. Mouse Dcir2 also contains an ITIM, but apart from demonstration of internalisation of the receptor no functional studies of Dcir2 regarding its signalling properties have been published (209). The ITIM is conserved in both mouse and rat and it is highly likely that Dcir2 is a functional inhibitory receptor in the two species.

Activating receptors

MINCLE, DECTIN2, BDCA-2 and Dcar2 have all been reported to associate with the ITAM-containing adapter protein FcRγ (18,210-212). In KLR, the rule is that the association to FcRγ requires a positively charged amino acid (arginine, R or lysine, K) in the TM region of the receptor close to the extracellular side to match the position of the negatively charged amino acid (aspartic acid, D) in the TM of FcRγ (see Fig.4) (23,213). The TM regions of APLRs were predicted using Smart and InterProScan and are indicated in fig.4. MINCLE has a positively charged amino acid (R) in the TM region, corresponding with the position of the negative charge in the FcRγ. Yamasaki et al demonstrated association of MINCLE and FcRγ after co-transfection and immunoprecipitation. Mutating the arginine to the neutral isoleucine abolished the association. Furthermore, endogenous MINCLE and FcRγ were co-immunoprecipitated in lysates from mouse peritoneal macrophages, confirming the association. Thus, Mincle seems to associate with FcRγ in a conventional manner (210).

In contrast, Dectin-2, BDCA-2 and Dcar1 and Dcar2 do not have a positively charged amino acid in the TM region. Kanazawa et al. reported that mouse Dcar2 associate with FcRγ. The authors suggested that an arginine located on the border of the predicted TM and cytoplasmic tail of the receptor was in part responsible for the association (18). Dectin-2 also contains an arginine and human BDCA-2 a lysine (K) in the same position (see fig.4). It should be noted, however, that single-pass transmembrane proteins often have positively charged amino acids in this position, of importance for the correct orientation of the TM in the membrane during protein synthesis (214).

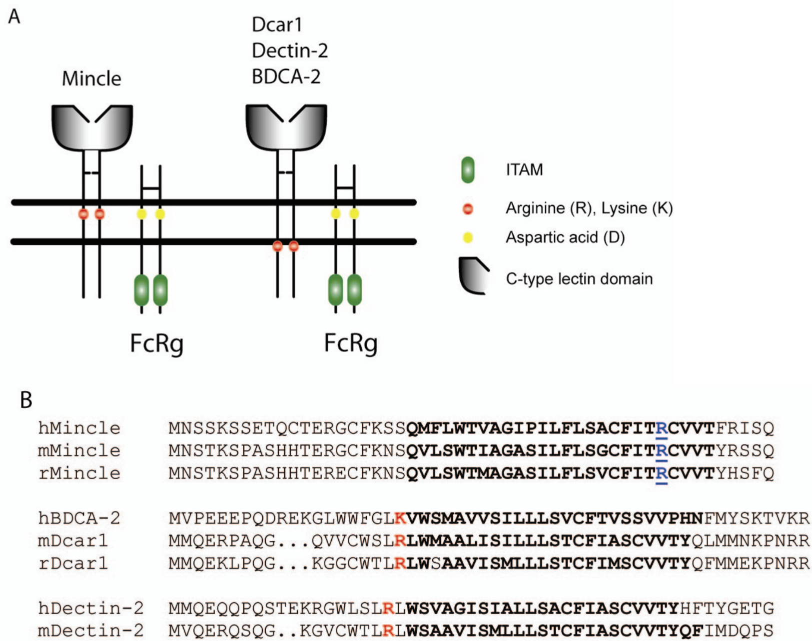

Figure 4. Relative positions of the charged amino acids of FcRγ and associating receptors.
A) The position of the positively charged amino acid in Mincle, Dcar1, Dectin-2 and DLEC/BDCA2 is indicated with a red circle. Mincle has an arginine in the position corresponding to the position of aspartic acid (D) in FcRγ. Dcar1 and Dectin-2 contains an arginine whereas BDCA-2 contains a lysine in the TM border. (Adapted from (23)). B) The predicted TM of Mincle, Dcar1, BDCA-2 and Dectin-2 are indicated in bold. In Mincle the arginine responsible for association with FcRγ is marked with blue. The positive charged amino acid suggested to facilitate adapter association in Dcar1, BDCA-2 and Dectin-2 is marked with red.

In the case of mDcar2, the data supporting the association with FcRγ are weak. Co-transfection of Dcar2-Flag and FcRγ gave only modest surface expression as detected with anti-Flag, and mutating the arginine to neutral or negatively charged residues (R/I, R/D) did not alter Dcar2-Flag surface expression in co-transfection studies with FcRγ (18). This indicates that the possible association between Dcar2 and FcRγ is not mediated by the arginine on the TM/cytoplasmic border in Dcar2. Despite several attempts to reproduce the results by co-transfection of rat Dcar1-Flag and FcRγ in our laboratory, we obtained only very weak to no surface expression of the receptor. A possible explanation is that Dcar1 needs another receptor chain to form a stable complex for expression, as is the case with the CD94/NKG2 and the KLRE/I heterodimers (10,22). It may also be that FcRγ is not the natural adapter, and that surface expression of Dcar1 requires association with a yet unidentified adapter protein.

In the human, association of BDCA-2 with FcRγ was shown in a transfectant cell system (212). Here chimeras of BDCA-2 with the TM domain from DCIR or MCL were co-transfected with FcRγ. Precipitation of FcRγ was not possible after co-transfections with the chimeras, only after co-transfection with wild type BDCA-2 (212). The TM of wildtype BDCA-2 and MCL differ only in 6 residues, and out of these three are polar in BDCA-2 and neutral in MCL. This difference in polarity may be involved in the association with the adapter protein, either directly or indirect through another unidentified protein.

In the mouse, Dectin-2 was reported to associate with FcRγ, as demonstrated by co-precipitation after co-transfection (211). Mutating the arginine on the TM/cytoplasmic border to valine (neutral) did not affect the association of FcRγ after co-transfection, indicating that the arginine residue is not responsible for association between receptor and adapter protein (211). A motif in the intracellular domain of mouse Dectin-2 appears to be important for association with FcRγ. In co-transfection studies with Dectin-2 mutants, a mutant lacking the intracellular domain (aa1-14) resulted in a poor precipitation of FcRγ, whereas transfection of Dectin-2 lacking aa1-7 did not alter the precipitation of FcRγ compared with wildtype Dectin-2. Thus amino acids 8-14 seem to be important for association to FcRγ (211). The protein sequence in the intracellular domain close to the TM domain is conserved in rodent Dectin-2 and Dcar1 and to a certain degree in human Dectin-2 and BDCA-2, but not in Mincle (figure 5).

	1 8 14
mDectin-2	MVQERQS**QG..KGVCW**
rDectin-2	MVQERHP**QG..KGVCW**
mDcar1	MMQERPA**QG..QVVCW**
rDcar1	MMQEKLP**QG..KGGCW**
mDcar2	MVQERQL**QG..KAVSW**
hBDCA-2	MVPEEEP**QDREKGLWW**
hDectin-2	MMQEQQP**QSTEKRGWL**
mMincle	MNSTKSPASHHTERGCFKNS**..QVLSW**

Figure 5. Amino acids 1-14 of intracellular domain of selected APLRs.
The aminoacid sequence 8-14 (shown in bold) is conserved between rat and mouse Dectin-2 and Dcar1 and to some degree in human BDCA-2 and Dectin-2, but not in Mincle or any other of the APLRs.

This indicates that a motif in the intracellular domain may be involved in the association of receptors and FcRγ. It should be noted, however, that except for Mincle, the association to FcRγ has been demonstrated only through co-transfection of APLR and FcRγ in cells with no endogenous APLRs or FcRγ. The artificial (over)expression of receptor and adapter protein may cause low-affinity, non-functional interactions. Association therefore needs to be verified in cells with endogenous receptor and adapter protein.

APLR ligands

Few functional studies of the receptors have been reported, however the identification of ligand may give cues about the receptors function. As detailed previously, the receptors have the conserved domains for Ca^{++} binding, they are "true" lectins and this indicates that the APLR are involved in carbohydrate binding and possibly have a role in pathogen recognition.

Mouse Dectin-2 has been shown to bind hyphae of the fungi *Candida albicans*, *Microsporum audouinii* and *Tricohophyton rubrum* and is highly specific for mannose. Furthermore, mouse Dectin-2 has been shown to recognise extracts from house dust mites and the mold *Aspergillus fumigatus*, inducing production of proinflammatory lipid mediators. The recognition of these pathogens indicates a role for Dectin-2 as a pattern recognition receptor (211,215,216).

Human and mouse MINCLE were reported to bind *C. albicans* inducing production of TNFα (217,218). Another ligand for Mincle has been identified. Yamasaki et al demonstrated that mouse Mincle bind SAP130, a subunit of nuclear ribonucleoprotein released from dead cells, indicating that Mincle can sense damaged cells and function as a "danger receptor" (210). However, the same group reported recently that mouse Mincle recognises *Malassezia sp* and that recognition were abolished when the putative mannose binding domain was mutated (219). Furthermore Mincle was recently shown to recognise the glycolipid Trehalose-dimycolate (TDM) from the cell wall of *Mycobacterium tuberculosis.* TDM stimulated macrophages produce inflammatory cytokines and nitric oxide. This production was abolished macrophages derived from Mincle deficient mice

(220). This is corroborated by another group (221). Together the data suggest that Mincle has a role as a pattern recognition receptor.

No ligands have been identified for the other APLEC receptors. Still, the main hypothesis is that these receptors are pattern recognition receptors. Human DCIR, MCL and DLEC/BDCA-2 as well as mouse Dcir-2 and Dcar-1 have been reported to be internalized upon crosslinking or binding of antibody (209,222-225). Internalization is no evidence of pattern recognition as most receptors will become internalized upon crosslinking or binding of ligand, however in the case of hDCIR, mDcir-2 and mDcar-1 it has been demonstrated that internalisation leads to delivery of antigens to the antigen processing pathway (209,224,225). Furthermore, binding of antibody to human DCIR and BDCA-2 resulted in decreased production of the major antiviral cytokine IFN-α from plasmacytoid dendritic cells (212,223,224).

The recognition of microbes and internalization of receptors upon binding with ligand/antibody, points towards a role for the APLR as pattern recognition receptors. The functional data also confirm that the receptors are involved in the regulation of the activation status of dendritic cells and mediate immune response by shaping the cytokine profile.

APLR in autoimmune diseases

In paper IV we identify *Dcar1* as a possible causative genetic factor for OIA and report the association of a SNP in APLEC close to *DCIR* with human RA. Sequence analysis of the rat genes revealed differences in *Mincle*, *Dcar*, *Dcir1* and *Dcir3* between the two strains, and expression analysis revealed differences in mRNA for *Mincle*, *MCL*, *Dcar1* and *Dcir3*. The missense mutations found in *Mincle*, *Dcir1* and *Dcir2* are not likely to have a serious influence on protein function, and the effect of the expression differences are unclear, nevertheless the genes cannot be excluded to not have any impact on arthritis development. The main gene candidate for *Oia2* is *Dcar1,* as based on the finding of the inactivating, early nonsense mutation in DA rats.

Previously, Mincle had been linked to RA in an indirect manner. Mincle is up-regulated in bone marrow derived mononuclear cells from RA patients, determined by micro array

(226). However, the up-regulation of Mincle may be a result rather than a causal factor in the precipitation of disease. A reduced number of BDCA-2 positive cells have been demonstrated in SLE and MS patients compared to healthy controls (227,228). However, since BDCA-2 is a marker for plasmacytoid DC (pDC), these reports only informs us that there are reduced numbers of pDC in the patients, and not of any involvement of BDCA-2 in the diseases. pDC are major producers of interferon-α which elevated in SLE patients, in this respect BDCA-2 triggering may be beneficial in SLE patients as BDCA-2 triggering leads to reduced production of interferons (212,223).

A report demonstrates that *Dcir1* deficiency cause autoimmune disease, as *Dcir1* knock-out mice spontaneously develop sialadenitis (inflammation in salivary gland) and enthesistis (inflammation in tendon attachment site). These mice also spontaneously produce higher levels of autoantibodies like rheumatoid factor and autoantibodies specific for Sjögrens syndrome. Furthermore, the *Dcir1* knock-out mice are more susceptible to collagen induced arthritis than wild-type mice. The *Dcir1* knock-out mice have an expansion of dendritic cells, and the authors suggest that *Dcir1* is involved in the maintenance of the DC population (229).

Both the animal model OIA and genome scans points at a susceptibility locus for arthritis in APLEC (141,142). We therefore tested a panel of RA patients and healthy controls for SNPs within APLEC, which resulted in the finding of association of *DCIR* to RA anti-CCP negative patients (paper IV). The SNPs used in the study was randomly chosen from reported SNPs in GenBank. At the start of the study, few SNPs were available, and most of them were located in intergenetic regions. Any association of arthritis to these SNPs would only indicate that a gene or genes were linked to disease. We found an association between RA in anti-CCP negative patients and SNP (rs1133104) and to a haplotype of 5 SNPs mapping to DCIR. However only one marker (rs2024301) was within the coding sequence of DCIR, one in the 3'UTR (rs1133104) and the rest were located upstream of DCIR. These polymorphisms are unlikely to be responsible for the development of disease. Their association just points to DCIR as a candidate gene. We aim to verify the association of RA to APLEC using novel, more relevant SNPs covering APLEC in a Norwegian panel of RA patients.

Sverdrup reported that oil exposure is linked to increased risk of anti-CCP positive RA (103). In paper IV we found association of *DCIR* to anti-CCP negative RA. The difference in anti-CCP status may be a result of multiple pathogenic pathways. As demonstrated in a recent publication, the effect of the APLEC genes on arthritis phenotype vary between different arthritis models in the rat (230). Interestingly, the APLEC congenic rat strain (R17) showed elevated levels of autoantibodies in CIA compared to the more susceptible DA strain, despite lower arthritis severity score. It may be that RA/autoimmunity triggered by oil exposure in humans is linked to production of autoantibodies, whereas we in paper IV may have picked up subgroup of RA patients that are anti-CCP negative.

Autoimmunity and occupational exposure to oil

In 2003 an article series in the Norwegian newspaper "Dagbladet" identified a group of workers exposed for mineral oils through occupation who now experience various symptoms of autoimmune and inflammatory diseases. After an initial report of ~30 people exposed to mineral oil, more people with similar symptoms came forward, and now the group counts more than 150 people. Some of them have got the diagnosis MS, a few are diagnosed with RA, but most have got no diagnosis at all. The symptoms they experience range from different neurological and inflammatory autoimmune symptoms to cancer and hypersensitivity. The wide varieties of symptoms make it hard to find a link between their symptoms and work, and the individuals struggle with the acceptance that their conditions are work related.

Because the workers in question have been exposed to different kinds of oil, it has been reasoned that it is not the oil in itself that triggers disease, but the additives in the oil. However, several reports of autoimmunity and autoimmune symptoms after administration of adjuvants exist, which have led to the proposal of human Adjuvant Disease (HAD) (91,93-99). Two publications confirmed occupational exposure to hydraulic oils and silica as risk factors for RA (103,104) and recently a cluster of SLE patients which had been exposed to an oil-waste reservoir was published (109). It is possible that the oil act as an adjuvant and that in susceptible individuals the oil triggers an unwanted immune response. The difference in symptoms may be explained by different composition of oils, additives in the oil and genetic background. In OIA the mineral oil triggers a pathogenic pathway where an APLEC gene is involved, and our

working hypothesis is that in certain cases of human RA and autoimmune diseases oil may function as an environmental stimulus that triggers a similar pathogenic pathway in which an APEC gene is playing a role. In parallel with the SNP typing of the Norwegian RA panel, the panel of oil exposed workers will also be typed to se whether any of SNPs in the APLEC genes show skewing compared to healthy controls.

APLR and pathogenesis

The OIA model has proven to be useful to find genetic risk factors for RA; however it may be even more valuable as a tool to determine the pathogenesis of disease. By generating an APLEC congenic strain, we are able to see differences in responses in DA versus R17 rats caused by APLEC genes after various stimuli. In paper V we show the effect of oil stimuli on the expression of APLEC genes and the consequences that had for cytokine expression. This work supports the hypothesis that *Dcar1* is the most likely candidate for *Oia2* although other APLEC genes cannot be ruled out as disease causing or contributing factors. Analysis of APLEC gene expression revealed that the receptors showed different expression in terms of fold up-regulation and pattern over time. This indicates that the genes are also regulated at the transcriptional level, and not only by influx of cells to the lymph nodes.

Analysis of cytokine expression revealed a significant increase of *Il17* expression in DA but not in R17 rats following oil injection. The fact that rats which develop arthritis express higher levels of *Il17* transcripts is not surprising. IL-17 is an inflammatory cytokine produced by the Th17 subset and has been reported to be induced in EAE (231). It is also elevated in humans with a variety of other autoimmune diseases (232-235).

Differentiation of the Th17 subset has over the last few years been extensively studied. Th17 cells are generated from $CD4^+$ Treg precursors. In presence of TGFβ these cells develop into T regulatory cells promoting tolerance. However in presence of IL-6 and TGFβ, the precursors develop into IL-17 producing cells (236-238). The development of Th17 cells and IL-17 production can be amplified by presence of IL-1β and IL-23 (239,240). IL-23 is also involved in maintaining IL-17 production (236). In our model we saw no difference in *Il6* mRNA expression between DA and R17 rats, and the only other cytokine that showed a significant difference in expression between the two strains was

Il1b. It has been reported that development of the Th17 subset in mouse could be generated independently of IL-6, and that it is in fact IL-1β that plays the dominant role in the Th17development (241). Recently it has been demonstrated that early Th17 cell differentiation in the mouse is regulated by IL-1β signalling both *in vivo* and *in vitro* (242). Similarly, human Th17cells were induced by IL-1β and amplified by IL-6 (243,244). Thus, it is possible that the elevated expression of *Il17* in DA rats is directed by *Il1b*.

The difference in *Il17* mRNA expression between the APLEC congenic strain and the DA rat was striking. These two strains only differ in the APLEC, thus one or more of these genes are clearly involved in the generation of pathogenic IL-17 producing T cells in the DA rat. The presence of *Il17* transcripts first presents at day 5p.i. in DA rats, indicating that the effect of APLEC gene(s) causing arthritis take place early in the induction phase. Prior to day 5p.i. there are only significant expression differences between DA and R17 for *Mincle* and *Dcir1* at day 2p.i. Therefore the gene candidates for *Oia2* can be narrowed down to *Mincle* and *Dcir1* based on mRNA expression and *Dcar1* and *Dcir2* based on sequence polymorphisms, with *Dcar1* being the prime candidate due to the nonsense mutation, leaving the DA rat as a natural knock-out for this gene. In addition the most striking expression profile was that of *Dcar1*. After an early strong up-regulation, the expression drops, but remains at an elevated level until arthritis presents at day 14-15p.i. In squalene induced arthritis (SIA), a similar model to OIA, a pathogenic subset of T cells develops around day 5p.i. as determined by adoptive transfer experiments. T cells harvested form rats immunised with squalene can transfer arthritis to naïve rats if the T cells are harvested after day 5p.i. (245). Given the similarities between the two models, it is possible that a pathogenic T cell subset develops in the same period in OIA susceptible rats. As we see presence of *Il17* from day 5 p.i., it is highly likely that this pathogenic subset is Th17 cells.

Exactly how the APLEC receptors are involved is not understood and needs further investigation. However, it is tempting to speculate that the increased expression of Dcar1 leads to a production of cytokines which inhibit the generation of Th17 cells in R17 rats, whereas the nonsense mutation prevent the synthesis of Dcar1 protein in the DA rat and leads to generation of Th17 cells.

It should be noted that we in our model only measured the mRNA expression. RNA expression does not always correlate with protein expression and there may be differences in the cytokines between DA and APLEC congenic rats at the protein level. We are currently investigating the difference in protein expression between DA and APLEC congenic rats for various cytokines. With a newly developed antibody against Dcar1 we are also able to verify the expression difference of Dcar1 between DA and congenic rats at the protein level.

QTLs mapping to APLEC

Rat chromosome 4 is dense in QTL for other experimental models for arthritis, like pristane induced arthritis (PIA), collagen induced arthritis (CIA), squalene induced arthritis and adjuvant induced arthritis (AIA) (189,246-250). A number of other experimental models such as the MS model experimental autoimmune encephalomyelitis (EAE), the diabetes models IDDM and NIDDM, (insulin dependent diabetes mellitus, non-insulin dependent diabetes mellitus) and experimental autoimmune uveoritinitis (EAU) also have QTLs located on rat chromosome 4, some of which have been mapped to span over APLEC (251-253). The EAE model is interesting in this regard, where a congenic interval gives down-modulating affect of arthritis and up-modulation of EAE (254).

In the human, a genome scan in MS affected families revealed a peak with a lod (logarithm of odds) score on chromosome 12, corresponding to a nominal significance of 5%, however the peak failed to be replicated in a second stage scan with higher stringency and lower number of affected families (255). The markers giving the peak cover APLEC, indicating a MS locus near or in APLEC. As mentioned in the introduction two meta-analyses based on partly the same genome scans both reveals a locus linked to RA susceptibility, with cytogenic location that spans over the human APLEC (141,142). The fact that QTLs in several experimental models and QTLs found in human linkage and association studies maps to APLEC indicate that one or more of the APLRs may be involved in the pathogenesis of several autoimmune or inflammatory diseases.

There are indications that several autoimmune diseases may share a common disease pathway. Originally identified as a gene regulating arthritis severity in the rat, *Ncf1* is

linked to both CIA and EAE in the mouse (197). In the human, the gene *PTPN22* have been associated with several autoimmune diseases; RA, type I diabetes, SLE and other inflammatory diseases 146,256-259), similarly *CTLA4* have been associated with RA, Graves disease, type 1 diabetes, and other autoimmune diseases (173,260-262). For this reason we plan to test other autoimmune and inflammatory diseases for association to APLEC in parallel with the Norwegian RA panel. The patient material consists of patients with Type1 diabetes, atherosclerosis and inflammatory bowel disease.

Future perspectives

Identifying risk factors to any autoimmune disease is important; however it is just a start. The ultimate goal is to develop better therapies; to slow down or stop the disease from progressing or even cure the disease. To develop better therapies we need to understand the pathogenesis of the disease. The association of RA to APLEC is thus only a small step forward, but the finding will instigate research on how the APLRs are involved in the development of disease.

As mentioned in the discussion, SNP typing with novel SNPs to confirm association of APLEC to RA in a Norwegian patient material and to determine whether other autoimmune and inflammatory diseases are associated with APLEC is an ongoing work. Examining the APLEC genes with respect to expression and splice patterns in RA patients compared to healthy controls may give clues to how the receptors are involved in development of autoimmune reactions, and are ongoing projects in our laboratory.

Future plans include studies of Aplr signalling, in both human and rat, hoping to delineate a pathogenic pathway. With new antibodies raised against rat Dcar1 and MCL we are better equipped to do functional studies. Biochemical analyses to reveal whether the APLEC receptors are homo- or heterodimers, as well as determining intracellular signalling pathways are planned. Examining the roles of the APLEC receptors by *in vivo* studies are intended. In particular, comparisons between DA and congenic strains and do *in vivo* blocking of the various receptors to see the effect of disease phenotype may give valuable clues to disease pathway.

Reference List

1. Medzhitov, R., and C. A. Janeway, Jr. 1997. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* 91:295.

2. Franchi, L., N. Warner, K. Viani, and G. Nunez. 2009. Function of Nod-like receptors in microbial recognition and host defense. *Immunol. Rev.* 227:106.

3. Geijtenbeek, T. B., S. J. van Vliet, A. Engering, B. A. 't Hart, and K. Y. van. 2004. Self- and nonself-recognition by C-type lectins on dendritic cells. *Annu. Rev. Immunol.* 22:33.

4. Pasare, C., and R. Medzhitov. 2004. Toll-like receptors: linking innate and adaptive immunity. *Microbes. Infect.* 6:1382.

5. van Vliet, S. J., D. J. den, S. I. Gringhuis, T. B. Geijtenbeek, and K. Y. van. 2007. Innate signaling and regulation of Dendritic cell immunity. *Curr. Opin. Immunol.* 19:435.

6. Drickamer, K. 1999. C-type lectin-like domains. *Curr. Opin. Struct. Biol.* 9:585.

7. Drickamer, K., and M. E. Taylor. 1993. Biology of animal lectins. *Annu. Rev. Cell Biol.* 9:237.

8. Zelensky, A. N., and J. E. Gready. 2005. The C-type lectin-like domain superfamily. *FEBS J.* 272:6179.

9. Yokoyama, W. M., and B. F. M. Plougastel. 2003. Immune functions encoded by the natural killer gene complex. *Nature Reviews Immunology* 3:304.

10. Lazetic, S., C. Chang, J. P. Houchins, L. L. Lanier, and J. H. Phillips. 1996. Human natural killer cell receptors involved in MHC class I recognition are disulfide-linked heterodimers of CD94 and NKG2 subunits. *J. Immunol.* 157:4741.

11. Lanier, L. L. 1998. NK cell receptors. *Annu. Rev. Immunol.* 16:359.

12. Brown, G. D., P. R. Taylor, D. M. Reid, J. A. Willment, D. L. Williams, L. Martinez-Pomares, S. Y. Wong, and S. Gordon. 2002. Dectin-1 is a major beta-glucan receptor on macrophages. *J. Exp. Med.* 196:407.

13. Stambach, N. S., and M. E. Taylor. 2003. Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells. *Glycobiology* 13:401.

14. Fadden, A. J., O. J. Holt, and K. Drickamer. 2003. Molecular characterization of the rat Kupffer cell glycoprotein receptor. *Glycobiology* 13:529.

15. Robinson, M. J., D. Sancho, E. C. Slack, S. LeibundGut-Landmann, and Reis e Sousa. 2006. Myeloid C-type lectins in innate immunity. *Nat. Immunol.* 7:1258.

16. Bates, E. E., N. Fournier, E. Garcia, J. Valladeau, I. Durand, J. J. Pin, S. M. Zurawski, S. Patel, J. S. Abrams, S. Lebecque, P. Garrone, and S. Saeland. 1999. APCs express DCIR, a novel C-type lectin surface receptor containing an immunoreceptor tyrosine-based inhibitory motif. *J. Immunol.* 163:1973.

17. Kanazawa, N., T. Okazaki, H. Nishimura, K. Tashiro, K. Inaba, and Y. Miyachi. 2002. DCIR acts as an inhibitory receptor depending on its immunoreceptor tyrosine-based inhibitory motif. *J. Invest Dermatol.* 118:261.

18. Kanazawa, N., K. Tashiro, K. Inaba, and Y. Miyachi. 2003. Dendritic cell immunoactivating receptor, a novel C-type lectin immunoreceptor, acts as an activating receptor through association with Fc receptor gamma chain. *J. Biol. Chem.* 278:32645.

19. Kanazawa, N. 2007. Dendritic cell immunoreceptors: C-type lectin receptors for pattern-recognition and signaling on antigen-presenting cells. *J. Dermatol. Sci.* 45:77.

20. Long, E. O. 2008. Negative signaling by inhibitory receptors: the NK cell paradigm. *Immunol. Rev.* 224:70.

21. Lanier, L. L. 2008. Up on the tightrope: natural killer cell activation and inhibition. *Nat. Immunol.* 9:495.

22. Saether, P. C., I. H. Westgaard, S. E. Hoelsbrekken, J. Benjamin, L. L. Lanier, S. Fossum, and E. Dissen. 2008. KLRE/I1 and KLRE/I2: a novel pair of heterodimeric receptors that inversely regulate NK cell cytotoxicity. *J. Immunol.* 181:3177.

23. Lanier, L. L. 2009. DAP10- and DAP12-associated receptors in innate immunity. *Immunol. Rev.* 227:150.

24. 1999. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* 401:921.

25. Cresswell, P. 1994. Assembly, transport, and function of MHC class II molecules. *Annu. Rev. Immunol.* 12:259.

26. Banchereau, J., F. Briere, C. Caux, J. Davoust, S. Lebecque, Y. J. Liu, B. Pulendran, and K. Palucka. 2000. Immunobiology of dendritic cells. *Annu. Rev. Immunol.* 18:767.

27. Lanzavecchia, A., and F. Sallusto. 2001. The instructive role of dendritic cells on T cell responses: lineages, plasticity and kinetics. *Curr. Opin. Immunol.* 13:291.

28. Ueno, H., E. Klechevsky, R. Morita, C. Aspord, T. Cao, T. Matsui, P. T. Di, J. Connolly, J. W. Fay, V. Pascual, A. K. Palucka, and J. Banchereau. 2007. Dendritic cell subsets in health and disease. *Immunol. Rev.* 219:118.

29. Alam, S. M., P. J. Travers, J. L. Wung, W. Nasholds, S. Redpath, S. C. Jameson, and N. R. Gascoigne. 1996. T-cell-receptor affinity and thymocyte positive selection. *Nature* 381:616.

30. Goldrath, A. W., and M. J. Bevan. 1999. Selecting and maintaining a diverse T-cell repertoire. *Nature* 402:255.

31. Kyewski, B., and L. Klein. 2006. A central role for central tolerance. *Annu. Rev. Immunol.* 24:571.

32. Alam, I., Q. Sun, L. Liu, D. L. Koller, T. Fishburn, L. G. Carr, M. J. Econs, T. Foroud, and C. H. Turner. 2006. Identification of a quantitative trait locus on rat chromosome 4 that is strongly linked to femoral neck structure and strength. *Bone* 39:93.

33. Derbinski, J., A. Schulte, B. Kyewski, and L. Klein. 2001. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nat. Immunol.* 2:1032.

34. Anderson, M. S., E. S. Venanzi, L. Klein, Z. Chen, S. P. Berzins, S. J. Turley, B. H. von, R. Bronson, A. Dierich, C. Benoist, and D. Mathis. 2002. Projection of an immunological self shadow within the thymus by the aire protein. *Science* 298:1395.

35. Heino, M., P. Peterson, J. Kudoh, K. Nagamine, A. Lagerstedt, V. Ovod, A. Ranki, I. Rantala, M. Nieminen, J. Tuukkanen, H. S. Scott, S. E. Antonarakis, N. Shimizu, and K. Krohn. 1999. Autoimmune regulator is expressed in the cells regulating immune tolerance in thymus medulla. *Biochem. Biophys. Res. Commun.* 257:821.

36. 1997. An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nat. Genet.* 17:399.

37. Nagamine, K., P. Peterson, H. S. Scott, J. Kudoh, S. Minoshima, M. Heino, K. J. Krohn, M. D. Lalioti, P. E. Mullis, S. E. Antonarakis, K. Kawasaki, S. Asakawa, F. Ito, and N. Shimizu. 1997. Positional cloning of the APECED gene. *Nat. Genet.* 17:393.

38. Ramsey, C., O. Winqvist, L. Puhakka, M. Halonen, A. Moro, O. Kampe, P. Eskelin, M. Pelto-Huikko, and L. Peltonen. 2002. Aire deficient mice develop multiple features of APECED phenotype and show altered immune response. *Hum. Mol. Genet.* 11:397.

39. Anderson, M. S., E. S. Venanzi, Z. Chen, S. P. Berzins, C. Benoist, and D. Mathis. 2005. The cellular mechanism of Aire control of T cell tolerance. *Immunity.* 23:227.

40. Derbinski, J., J. Gabler, B. Brors, S. Tierling, S. Jonnakuty, M. Hergenhahn, L. Peltonen, J. Walter, and B. Kyewski. 2005. Promiscuous gene expression in thymic epithelial cells is regulated at multiple levels. *J. Exp. Med.* 202:33.

41. Peterson, P., T. Org, and A. Rebane. 2008. Transcriptional regulation by AIRE: molecular mechanisms of central tolerance. *Nat. Rev. Immunol.* 8:948.

42. Gallegos, A. M., and M. J. Bevan. 2004. Central tolerance to tissue-specific antigens mediated by direct and indirect antigen presentation. *J. Exp. Med.* 200:1039.

43. Albert, M. L., S. F. Pearce, L. M. Francisco, B. Sauter, P. Roy, R. L. Silverstein, and N. Bhardwaj. 1998. Immature dendritic cells phagocytose apoptotic cells via alphavbeta5 and CD36, and cross-present antigens to cytotoxic T lymphocytes. *J. Exp. Med.* 188:1359.

44. Thery, C., L. Duban, E. Segura, P. Veron, O. Lantz, and S. Amigorena. 2002. Indirect activation of naive CD4+ T cells by dendritic cell-derived exosomes. *Nat. Immunol.* 3:1156.

45. Gray, D., J. Abramson, C. Benoist, and D. Mathis. 2007. Proliferative arrest and rapid turnover of thymic epithelial cells expressing Aire. *J. Exp. Med.* 204:2521.

46. Steinman, R. M., D. Hawiger, and M. C. Nussenzweig. 2003. Tolerogenic dendritic cells. *Annu. Rev. Immunol.* 21:685.

47. Redmond, W. L., and L. A. Sherman. 2005. Peripheral tolerance of CD8 T lymphocytes. *Immunity.* 22:275.

48. Ferguson, T. A., and T. S. Griffith. 2006. A vision of cell death: Fas ligand and immune privilege 10 years later. *Immunol. Rev.* 213:228.

49. Niederkorn, J. Y. 2006. See no evil, hear no evil, do no evil: the lessons of immune privilege. *Nat. Immunol.* 7:354.

50. Mosmann, T. R., H. Cherwinski, M. W. Bond, M. A. Giedlin, and R. L. Coffman. 1986. Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *J. Immunol.* 136:2348.

51. Zhu, J., and W. E. Paul. 2008. CD4 T cells: fates, functions, and faults. *Blood* 112:1557.

52. Rocken, M., and E. M. Shevach. 1996. Immune deviation--the third dimension of nondeletional T cell tolerance. *Immunol. Rev.* 149:175.

53. Korn, T., E. Bettelli, M. Oukka, and V. K. Kuchroo. 2009. IL-17 and Th17 Cells. *Annu. Rev. Immunol.* 27:485.

54. Pene, J., S. Chevalier, L. Preisser, E. Venereau, M. H. Guilleux, S. Ghannam, J. P. Moles, Y. Danger, E. Ravon, S. Lesaux, H. Yssel, and H. Gascan. 2008. Chronically inflamed human tissues are infiltrated by highly differentiated Th17 lymphocytes. *J. Immunol.* 180:7423.

55. Vignali, D. A., L. W. Collison, and C. J. Workman. 2008. How regulatory T cells work. *Nat. Rev. Immunol.* 8:523.

56. Sakaguchi, S., T. Yamaguchi, T. Nomura, and M. Ono. 2008. Regulatory T cells and immune tolerance. *Cell* 133:775.

57. Sakaguchi, S., K. Fukuma, K. Kuribayashi, and T. Masuda. 1985. Organ-specific autoimmune diseases induced in mice by elimination of T cell subset. I. Evidence for the active participation of T cells in natural self-tolerance; deficit of a T cell subset as a possible cause of autoimmune disease. *J. Exp. Med.* 161:72.

58. McKeever, U., J. P. Mordes, D. L. Greiner, M. C. Appel, J. Rozing, E. S. Handler, and A. A. Rossini. 1990. Adoptive transfer of autoimmune diabetes and thyroiditis to athymic rats. *Proc. Natl. Acad. Sci. U. S. A* 87:7618.

59. Powrie, F., and D. Mason. 1990. OX-22high CD4+ T cells induce wasting disease with multiple organ pathology: prevention by the OX-22low subset. *J. Exp. Med.* 172:1701.

60. Asano, M., M. Toda, N. Sakaguchi, and S. Sakaguchi. 1996. Autoimmune disease as a consequence of developmental abnormality of a T cell subpopulation. *J. Exp. Med.* 184:387.

61. Gambineri, E., T. R. Torgerson, and H. D. Ochs. 2003. Immune dysregulation, polyendocrinopathy, enteropathy, and X-linked inheritance (IPEX), a syndrome of systemic autoimmunity caused by mutations of FOXP3, a critical regulator of T-cell homeostasis. *Curr. Opin. Rheumatol.* 15:430.

62. Wing, K., and S. Sakaguchi. 2010. Regulatory T cells exert checks and balances on self tolerance and autoimmunity. *Nat. Immunol.* 11:7.

63. Billingham, R. E., L. Brent, and P. B. Medawar. 1953. Actively acquired tolerance of foreign cells. *Nature* 172:603.

64. Matzinger, P. 2002. The danger model: a renewed sense of self. *Science* 296:301.

65. Bettelli, E., T. Korn, and V. K. Kuchroo. 2007. Th17: the third member of the effector T cell trilogy. *Curr. Opin. Immunol.*

66. Gabriel, S. E. 2001. The epidemiology of rheumatoid arthritis. *Rheum. Dis. Clin. North Am.* 27:269.

67. Alamanos, Y., and A. A. Drosos. 2005. Epidemiology of adult rheumatoid arthritis. *Autoimmun. Rev.* 4:130.

68. Pincus, T., and L. F. Callahan. 1989. Reassessment of twelve traditional paradigms concerning the diagnosis, prevalence, morbidity and mortality of rheumatoid arthritis. *Scand. J. Rheumatol. Suppl* 79:67.

69. Gabriel, S. E. 2008. Cardiovascular morbidity and mortality in rheumatoid arthritis. *Am. J. Med.* 121:S9.

70. Silman, A. J., and J. E. Pearson. 2002. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res.* 4 Suppl 3:S265-S272.

71. Mewar, D., and A. G. Wilson. 2006. Autoantibodies in rheumatoid arthritis: a review. *Biomed. Pharmacother.* 60:648.

72. Despres, N., G. Boire, F. J. Lopez-Longo, and H. A. Menard. 1994. The Sa system: a novel antigen-antibody system specific for rheumatoid arthritis. *J. Rheumatol.* 21:1027.

73. Schellekens, G. A., H. Visser, B. A. de Jong, F. H. van den Hoogen, J. M. Hazes, F. C. Breedveld, and W. J. van Venrooij. 2000. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum.* 43:155.

74. Vossenaar, E. R., A. J. Zendman, W. J. van Venrooij, and G. J. Pruijn. 2003. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays* 25:1106.

75. Vencovsky, J., S. Machacek, L. Sedova, J. Kafkova, J. Gatterova, V. Pesakova, and S. Ruzickova. 2003. Autoantibodies can be prognostic markers of an erosive disease in early rheumatoid arthritis. *Ann. Rheum. Dis.* 62:427.

76. Schellekens, G. A., B. A. de Jong, F. H. van den Hoogen, L. B. van de Putte, and W. J. van Venrooij. 1998. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J. Clin. Invest* 101:273.

77. Visser, H., C. S. le, K. Vos, F. C. Breedveld, and J. M. Hazes. 2002. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum.* 46:357.

78. Ostensen, M. 1999. Sex hormones and pregnancy in rheumatoid arthritis and systemic lupus erythematosus. *Ann. N. Y. Acad. Sci.* 876:131.

79. Silman, A. J. 2002. Contraceptives, pregnancy, and RA. *Ann. Rheum. Dis.* 61:383.

80. Munz, C., J. D. Lunemann, M. T. Getts, and S. D. Miller. 2009. Antiviral immune responses: triggers of or triggered by autoimmunity? *Nat. Rev. Immunol.* 9:246.

81. Bach, J. F. 2005. Infections and autoimmune diseases. *J. Autoimmun.* 25 Suppl:74.

82. Vessey, M. P., L. Villard-Mackintosh, and D. Yeates. 1987. Oral contraceptives, cigarette smoking and other factors in relation to arthritis. *Contraception* 35:457.

83. Karlson, E. W., I. M. Lee, N. R. Cook, J. E. Manson, J. E. Buring, and C. H. Hennekens. 1999. A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. *Arthritis Rheum.* 42:910.

84. Silman, A. J., J. Newman, and A. J. MacGregor. 1996. Cigarette smoking increases the risk of rheumatoid arthritis. Results from a nationwide study of disease-discordant twins. *Arthritis Rheum.* 39:732.

85. Hutchinson, D., L. Shepstone, R. Moots, J. T. Lear, and M. P. Lynch. 2001. Heavy cigarette smoking is strongly associated with rheumatoid arthritis (RA), particularly in patients without a family history of RA. *Ann. Rheum. Dis.* 60:223.

86. Reckner, O. A., T. Skogh, and G. Wingren. 2001. Comorbidity and lifestyle, reproductive factors, and environmental exposures associated with rheumatoid arthritis. *Ann. Rheum. Dis.* 60:934.

87. Symmons, D. P., C. R. Bankhead, B. J. Harrison, P. Brennan, E. M. Barrett, D. G. Scott, and A. J. Silman. 1997. Blood transfusion, smoking, and obesity as risk factors for the development of rheumatoid arthritis: results from a primary care-based incident case-control study in Norfolk, England. *Arthritis Rheum.* 40:1955.

88. Uhlig, T., K. B. Hagen, and T. K. Kvien. 1999. Current tobacco smoking, formal education, and the risk of rheumatoid arthritis. *J. Rheumatol.* 26:47.

89. Criswell, L. A., L. A. Merlino, J. R. Cerhan, T. R. Mikuls, A. S. Mudano, M. Burma, A. R. Folsom, and K. G. Saag. 2002. Cigarette smoking and the risk of rheumatoid arthritis among postmenopausal women: results from the Iowa Women's Health Study. *Am. J. Med.* 112:465.

90. Stolt, P., C. Bengtsson, B. Nordmark, S. Lindblad, I. Lundberg, L. Klareskog, and L. Alfredsson. 2003. Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. *Ann. Rheum. Dis.* 62:835.

91. Gupta, P., R. P. Singh, M. V. Murali, S. K. Bhargava, and P. Sharma. 1992. Kerosene oil poisoning--a childhood menace. *Indian Pediatr.* 29:979.

92. Gupta, P., R. P. Singh, M. V. Murali, and P. P. Sharma. 1992. Prognostic score for kerosene oil poisoning. *Indian Pediatr.* 29:1109.

93. Gupta, R. K., E. H. Relyveld, E. B. Lindblad, B. Bizzini, S. Ben-Efraim, and C. K. Gupta. 1993. Adjuvants--a balance between toxicity and adjuvanticity. *Vaccine* 11:293.

94. Kumagai, Y., C. Abe, and Y. Shiokawa. 1979. Scleroderma after cosmetic surgery: four cases of human adjuvant disease. *Arthritis Rheum.* 22:532.

95. Baldwin, C. M., Jr., and E. N. Kaplan. 1983. Silicone-induced human adjuvant disease? *Ann. Plast. Surg.* 10:270.

96. Straniero, N. R., and D. E. Furst. 1989. Environmentally-induced systemic sclerosis-like illness. *Baillieres Clin. Rheumatol.* 3:63.

97. Yel, L., W. Chen, and S. Gupta. 2004. Cellular immunodeficiency and autoimmunity in long-term mineral oil administration. *Ann. Allergy Asthma Immunol.* 92:88.

98. Kammuller, M. E., N. Bloksma, and W. Seinen. 1988. Chemical-induced autoimmune reactions and Spanish toxic oil syndrome. Focus on hydantoins and related compounds. *J. Toxicol. Clin. Toxicol.* 26:157.

99. Cabral, A. R., J. cocer-Varela, R. Orozco-Topete, E. Reyes, L. Fernandez-Dominguez, and D. arcon-Segovia. 1994. Clinical, histopathological,

immunological and fibroblast studies in 30 patients with subcutaneous injections of modelants including silicone and mineral oils. *Rev. Invest Clin.* 46:257.

100. Kuroda, Y., D. C. Nacionales, J. Akaogi, W. H. Reeves, and M. Satoh. 2004. Autoimmunity induced by adjuvant hydrocarbon oil components of vaccine. *Biomed. Pharmacother.* 58:325.

101. Kuroda, Y., J. Akaogi, D. C. Nacionales, S. C. Wasdo, N. J. Szabo, W. H. Reeves, and M. Satoh. 2004. Distinctive patterns of autoimmune response induced by different types of mineral oil. *Toxicol. Sci.* 78:222.

102. Satoh, M., Y. Kuroda, H. Yoshida, K. M. Behney, A. Mizutani, J. Akaogi, D. C. Nacionales, T. D. Lorenson, R. J. Rosenbauer, and W. H. Reeves. 2003. Induction of lupus autoantibodies by adjuvants. *J. Autoimmun.* 21:1.

103. Sverdrup, B., H. Kallberg, C. Bengtsson, I. Lundberg, L. Padyukov, L. Alfredsson, and L. Klareskog. 2005. Association between occupational exposure to mineral oil and rheumatoid arthritis: results from the Swedish EIRA case-control study. *Arthritis Res. Ther.* 7:R1296-R1303.

104. Stolt, P., H. Kallberg, I. Lundberg, B. Sjogren, L. Klareskog, and L. Alfredsson. 2005. Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from the Swedish EIRA study. *Ann. Rheum. Dis.* 64:582.

105. Hawkes, C. H. 2007. Smoking is a risk factor for multiple sclerosis: a metanalysis. *Mult. Scler.* 13:610.

106. Sundstrom, P., L. Nystrom, and G. Hallmans. 2008. Smoke exposure increases the risk for multiple sclerosis. *Eur. J. Neurol.* 15:579.

107. Brix, T. H., P. S. Hansen, K. O. Kyvik, and L. Hegedus. 2000. Cigarette smoking and risk of clinically overt thyroid disease: a population-based twin case-control study. *Arch. Intern. Med.* 160:661.

108. Majka, D. S., and V. M. Holers. 2006. Cigarette smoking and the risk of systemic lupus erythematosus and rheumatoid arthritis. *Ann. Rheum. Dis.* 65:561.

109. Dahlgren, J., H. Takhar, P. nderson-Mahoney, J. Kotlerman, J. Tarr, and R. Warshaw. 2007. Cluster of systemic lupus erythematosus (SLE) associated with an oil field waste site: a cross sectional study. *Environ. Health* 6:8.

110. Parks, C. G., G. S. Cooper, L. A. Nylander-French, W. T. Sanderson, J. M. Dement, P. L. Cohen, M. A. Dooley, E. L. Treadwell, E. W. St Clair, G. S. Gilkeson, J. A. Hoppin, and D. A. Savitz. 2002. Occupational exposure to crystalline silica and risk of systemic lupus erythematosus: a population-based, case-control study in the southeastern United States. *Arthritis Rheum.* 46:1840.

111. Finckh, A., G. S. Cooper, L. B. Chibnik, K. H. Costenbader, J. Watts, H. Pankey, P. A. Fraser, and E. W. Karlson. 2006. Occupational silica and solvent exposures and risk of systemic lupus erythematosus in urban women. *Arthritis Rheum.* 54:3648.

112. Hogan, S. L., K. K. Satterly, M. A. Dooley, P. H. Nachman, J. C. Jennette, and R. J. Falk. 2001. Silica exposure in anti-neutrophil cytoplasmic autoantibody-associated glomerulonephritis and lupus nephritis. *J. Am. Soc. Nephrol.* 12:134.

113. Dawn, T. M., and J. H. Barrett. 2005. Genetic linkage studies. *Lancet* 366:1036.

114. Lichten, M., and A. S. Goldman. 1995. Meiotic recombination hotspots. *Annu. Rev. Genet.* 29:423.

115. Reich, D. E., S. F. Schaffner, M. J. Daly, G. McVean, J. C. Mullikin, J. M. Higgins, D. J. Richter, E. S. Lander, and D. Altshuler. 2002. Human genome sequence variation and the influence of gene history, mutation and recombination. *Nat. Genet.* 32:135.

116. Plenge, R., and J. D. Rioux. 2006. Identifying susceptibility genes for immunological disorders: patterns, power, and proof. *Immunol. Rev.* 210:40.

117. Cordell, H. J., and D. G. Clayton. 2005. Genetic association studies. *Lancet* 366:1121.

118. Reich, D. E., M. Cargill, S. Bolk, J. Ireland, P. C. Sabeti, D. J. Richter, T. Lavery, R. Kouyoumjian, S. F. Farhadian, R. Ward, and E. S. Lander. 2001. Linkage disequilibrium in the human genome. *Nature* 411:199.

119. Kerem, B., J. M. Rommens, J. A. Buchanan, D. Markiewicz, T. K. Cox, A. Chakravarti, M. Buchwald, and L. C. Tsui. 1989. Identification of the cystic fibrosis gene: genetic analysis. *Science* 245:1073.

120. Gusella, J. F., N. S. Wexler, P. M. Conneally, S. L. Naylor, M. A. Anderson, R. E. Tanzi, P. C. Watkins, K. Ottina, M. R. Wallace, A. Y. Sakaguchi, and . 1983. A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306:234.

121. Risch, N., and K. Merikangas. 1996. The future of genetic studies of complex human diseases. *Science* 273:1516.

122. Lohmueller, K. E., C. L. Pearce, M. Pike, E. S. Lander, and J. N. Hirschhorn. 2003. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat. Genet.* 33:177.

123. Ohashi, J., and K. Tokunaga. 2002. The expected power of genome-wide linkage disequilibrium testing using single nucleotide polymorphism markers for detecting a low-frequency disease variant. *Ann. Hum. Genet.* 66:297.

124. Altshuler, D., M. J. Daly, and E. S. Lander. 2008. Genetic mapping in human disease. *Science* 322:881.

125. Borecki, I. B., and M. A. Province. 2008. Linkage and association: basic concepts. *Adv. Genet.* 60:51.

126. Ji, W., J. N. Foo, B. J. O'Roak, H. Zhao, M. G. Larson, D. B. Simon, C. Newton-Cheh, State MW, D. Levy, and R. P. Lifton. 2008. Rare independent mutations in

renal salt handling genes contribute to blood pressure variation. *Nat. Genet.* 40:592.

127. Haiman, C. A., M. L. Le, J. Yamamato, D. O. Stram, X. Sheng, L. N. Kolonel, A. H. Wu, D. Reich, and B. E. Henderson. 2007. A common genetic risk factor for colorectal and prostate cancer. *Nat. Genet.* 39:954.

128. Graham, R. R., C. Kyogoku, S. Sigurdsson, I. A. Vlasova, L. R. Davies, E. C. Baechler, R. M. Plenge, T. Koeuth, W. A. Ortmann, G. Hom, J. W. Bauer, C. Gillett, N. Burtt, D. S. Cunninghame Graham, R. Onofrio, M. Petri, I. Gunnarsson, E. Svenungsson, L. Ronnblom, G. Nordmark, P. K. Gregersen, K. Moser, P. M. Gaffney, L. A. Criswell, T. J. Vyse, A. C. Syvanen, P. R. Bohjanen, M. J. Daly, T. W. Behrens, and D. Altshuler. 2007. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc. Natl. Acad. Sci. U. S. A* 104:6758.

129. Zamel, N., P. A. McClean, P. R. Sandell, K. A. Siminovitch, and A. S. Slutsky. 1996. Asthma on Tristan da Cunha: looking for the genetic link. The University of Toronto Genetics of Asthma Research Group. *Am. J. Respir. Crit Care Med.* 153:1902.

130. Wright, A. F., A. D. Carothers, and M. Pirastu. 1999. Population choice in mapping genes for complex diseases. *Nat. Genet.* 23:397.

131. Peltonen, L., A. Palotie, and K. Lange. 2000. Use of population isolates for mapping complex traits. *Nat. Rev. Genet.* 1:182.

132. Yamada, R., and K. Yamamoto. 2007. Mechanisms of disease: genetics of rheumatoid arthritis--ethnic differences in disease-associated genes. *Nat. Clin. Pract. Rheumatol.* 3:644.

133. Shiozawa, S., S. Hayashi, Y. Tsukamoto, H. Goko, H. Kawasaki, T. Wada, K. Shimizu, N. Yasuda, N. Kamatani, K. Takasugi, Y. Tanaka, K. Shiozawa, and S. Imura. 1998. Identification of the gene loci that predispose to rheumatoid arthritis. *Int. Immunol.* 10:1891.

134. Cornelis, F., S. Faure, M. Martinez, J. F. Prud'homme, P. Fritz, C. Dib, H. Alves, P. Barrera, V. N. de, A. Balsa, D. Pascual-Salcedo, K. Maenaut, R. Westhovens, P. Migliorini, T. H. Tran, A. Delaye, N. Prince, C. Lefevre, G. Thomas, M. Poirier, S. Soubigou, O. Alibert, S. Lasbleiz, S. Fouix, C. Bouchier, F. Liote, M. N. Loste, V. Lepage, D. Charron, G. Gyapay, A. Lopes-Vaz, D. Kuntz, T. Bardin, and J. Weissenbach. 1998. New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc. Natl. Acad. Sci. U. S. A* 95:10746.

135. Jawaheer, D., M. F. Seldin, C. I. Amos, W. V. Chen, R. Shigeta, J. Monteiro, M. Kern, L. A. Criswell, S. Albani, J. L. Nelson, D. O. Clegg, R. Pope, H. W. Schroeder, Jr., S. L. Bridges, Jr., D. S. Pisetsky, R. Ward, D. L. Kastner, R. L. Wilder, T. Pincus, L. F. Callahan, D. Flemming, M. H. Wener, and P. K. Gregersen. 2001. A genomewide screen in multiplex rheumatoid arthritis families

suggests genetic overlap with other autoimmune diseases. *Am. J. Hum. Genet.* 68:927.

136. MacKay, K., S. Eyre, A. Myerscough, A. Milicic, A. Barton, S. Laval, J. Barrett, D. Lee, S. White, S. John, M. A. Brown, J. Bell, A. Silman, W. Ollier, P. Wordsworth, and J. Worthington. 2002. Whole-genome linkage analysis of rheumatoid arthritis susceptibility loci in 252 affected sibling pairs in the United Kingdom. *Arthritis Rheum.* 46:632.

137. Lander, E., and L. Kruglyak. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* 11:241.

138. Jawaheer, D., M. F. Seldin, C. I. Amos, W. V. Chen, R. Shigeta, C. Etzel, A. Damle, X. Xiao, D. Chen, R. F. Lum, J. Monteiro, M. Kern, L. A. Criswell, S. Albani, J. L. Nelson, D. O. Clegg, R. Pope, H. W. Schroeder, Jr., S. L. Bridges, Jr., D. S. Pisetsky, R. Ward, D. L. Kastner, R. L. Wilder, T. Pincus, L. F. Callahan, D. Flemming, M. H. Wener, and P. K. Gregersen. 2003. Screening the genome for rheumatoid arthritis susceptibility genes: a replication study and combined analysis of 512 multicase families. *Arthritis Rheum.* 48:906.

139. Eyre, S., A. Barton, N. Shephard, A. Hinks, W. Brintnell, K. MacKay, A. Silman, W. Ollier, P. Wordsworth, S. John, and J. Worthington. 2004. Investigation of susceptibility loci identified in the UK rheumatoid arthritis whole-genome scan in a further series of 217 UK affected sibling pairs. *Arthritis Rheum.* 50:729.

140. Osorio, Y. F., H. Bukulmez, E. Petit-Teixeira, L. Michou, C. Pierlot, S. Cailleau-Moindrault, I. Lemaire, S. Lasbleiz, O. Alibert, P. Quillet, T. Bardin, B. Prum, J. M. Olson, and F. Cornelis. 2004. Dense genome-wide linkage analysis of rheumatoid arthritis, including covariates. *Arthritis Rheum.* 50:2757.

141. Fisher, S. A., J. S. Lanchbury, and C. M. Lewis. 2003. Meta-analysis of four rheumatoid arthritis genome-wide linkage studies: confirmation of a susceptibility locus on chromosome 16. *Arthritis Rheum.* 48:1200.

142. Choi, S. J., Y. H. Rho, J. D. Ji, G. G. Song, and Y. H. Lee. 2006. Genome scan meta-analysis of rheumatoid arthritis. *Rheumatology. (Oxford)* 45:166.

143. Dobloug, J. H., O. Forre, E. Kass, and E. Thorsby. 1980. HLA antigens and rheumatoid arthritis. Association between HLA-DRw4 positivity and IgM rheumatoid factor production. *Arthritis Rheum.* 23:309.

144. Hardwick, L. J., S. Walsh, S. Butcher, A. Nicod, J. Shatford, J. Bell, M. Lathrop, and B. P. Wordsworth. 1997. Genetic mapping of susceptibility loci in the genes involved in rheumatoid arthritis. *J. Rheumatol.* 24:197.

145. Gregersen, P. K., J. Silver, and R. J. Winchester. 1987. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum.* 30:1205.

146. Begovich, A. B., V. E. Carlton, L. A. Honigberg, S. J. Schrodi, A. P. Chokkalingam, H. C. Alexander, K. G. Ardlie, Q. Huang, A. M. Smith, J. M. Spoerke, M. T. Conn, M. Chang, S. Y. Chang, R. K. Saiki, J. J. Catanese, D. U.

Leong, V. E. Garcia, L. B. McAllister, D. A. Jeffery, A. T. Lee, F. Batliwalla, E. Remmers, L. A. Criswell, M. F. Seldin, D. L. Kastner, C. I. Amos, J. J. Sninsky, and P. K. Gregersen. 2004. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am. J. Hum. Genet.* 75:330.

147. Orozco, G., E. Sanchez, M. A. Gonzalez-Gay, M. A. Lopez-Nevot, B. Torres, R. Caliz, N. Ortego-Centeno, J. Jimenez-Alonso, D. Pascual-Salcedo, A. Balsa, P. R. de, A. Nunez-Roldan, M. F. Gonzalez-Escribano, and J. Martin. 2005. Association of a functional single-nucleotide polymorphism of PTPN22, encoding lymphoid protein phosphatase, with rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Rheum.* 52:219.

148. Lee, A. T., W. Li, A. Liew, C. Bombardier, M. Weisman, E. M. Massarotti, J. Kent, F. Wolfe, A. B. Begovich, and P. K. Gregersen. 2005. The PTPN22 R620W polymorphism associates with RF positive rheumatoid arthritis in a dose-dependent manner but not with HLA-SE status. *Genes Immun.* 6:129.

149. Hinks, A., A. Barton, S. John, I. Bruce, C. Hawkins, C. E. Griffiths, R. Donn, W. Thomson, A. Silman, and J. Worthington. 2005. Association between the PTPN22 gene and rheumatoid arthritis and juvenile idiopathic arthritis in a UK population: further support that PTPN22 is an autoimmunity gene. *Arthritis Rheum.* 52:1694.

150. Seldin, M. F., R. Shigeta, K. Laiho, H. Li, H. Saila, A. Savolainen, M. Leirisalo-Repo, K. Aho, E. Tuomilehto-Wolf, K. Kaarela, M. Kauppi, H. C. Alexander, A. B. Begovich, and J. Tuomilehto. 2005. Finnish case-control and family studies support PTPN22 R620W polymorphism as a risk factor in rheumatoid arthritis, but suggest only minimal or no effect in juvenile idiopathic arthritis. *Genes Immun.* 6:720.

151. Steer, S., B. Lad, J. A. Grumley, G. H. Kingsley, and S. A. Fisher. 2005. Association of R602W in a protein tyrosine phosphatase gene with a high risk of rheumatoid arthritis in a British population: evidence for an early onset/disease severity effect. *Arthritis Rheum.* 52:358.

152. Viken, M. K., S. S. Amundsen, T. K. Kvien, K. M. Boberg, I. M. Gilboe, V. Lilleby, L. M. Sollid, O. T. Forre, E. Thorsby, A. Smerdel, and B. A. Lie. 2005. Association analysis of the 1858C>T polymorphism in the PTPN22 gene in juvenile idiopathic arthritis and other autoimmune diseases. *Genes Immun.* 6:271.

153. Lie, B. A., M. K. Viken, S. Odegard, H. D. van der, R. Landewe, T. Uhlig, and T. K. Kvien. 2007. Associations between the PTPN22 1858C->T polymorphism and radiographic joint destruction in patients with rheumatoid arthritis: results from a 10-year longitudinal study. *Ann. Rheum. Dis.* 66:1604.

154. Zhernakova, A., P. Eerligh, C. Wijmenga, P. Barrera, B. O. Roep, and B. P. Koeleman. 2005. Differential association of the PTPN22 coding variant with autoimmune diseases in a Dutch population. *Genes Immun.* 6:459.

155. Ikari, K., S. Momohara, E. Inoue, T. Tomatsu, M. Hara, H. Yamanaka, and N. Kamatani. 2006. Haplotype analysis revealed no association between the PTPN22 gene and RA in a Japanese population. *Rheumatology. (Oxford)* 45:1345.

156. Lee, H. S., B. D. Korman, J. M. Le, D. L. Kastner, E. F. Remmers, P. K. Gregersen, and S. C. Bae. 2009. Genetic risk factors for rheumatoid arthritis differ in caucasian and Korean populations. *Arthritis Rheum.* 60:364.

157. Mastana, S., A. Gilmour, A. Ghelani, H. Smith, and A. Samanta. 2007. Association of PTPN22 with rheumatoid arthritis among South Asians in the UK. *J. Rheumatol.* 34:1984.

158. Gomez, L. M., J. M. Anaya, C. I. Gonzalez, R. Pineda-Tamayo, W. Otero, A. Arango, and J. Martin. 2005. PTPN22 C1858T polymorphism in Colombian patients with autoimmune diseases. *Genes Immun.* 6:628.

159. Suzuki, A., R. Yamada, X. Chang, S. Tokuhiro, T. Sawada, M. Suzuki, M. Nagasaki, M. Nakayama-Hamada, R. Kawaida, M. Ono, M. Ohtsuki, H. Furukawa, S. Yoshino, M. Yukioka, S. Tohma, T. Matsubara, S. Wakitani, R. Teshima, Y. Nishioka, A. Sekine, A. Iida, A. Takahashi, T. Tsunoda, Y. Nakamura, and K. Yamamoto. 2003. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat. Genet.* 34:395.

160. Ikari, K., M. Kuwahara, T. Nakamura, S. Momohara, M. Hara, H. Yamanaka, T. Tomatsu, and N. Kamatani. 2005. Association between PADI4 and rheumatoid arthritis: a replication study. *Arthritis Rheum.* 52:3054.

161. Fan, L. Y., W. J. Wang, Q. Wang, M. Zong, L. Yang, H. Zhang, L. S. Sun, T. B. Lu, and J. Han. 2008. A functional haplotype and expression of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Chinese. *Tissue Antigens* 72:469.

162. Kang, C. P., H. S. Lee, H. Ju, H. Cho, C. Kang, and S. C. Bae. 2006. A functional haplotype of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Koreans. *Arthritis Rheum.* 54:90.

163. Barton, A., J. Bowes, S. Eyre, K. Spreckley, A. Hinks, S. John, and J. Worthington. 2004. A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. *Arthritis Rheum.* 50:1117.

164. Caponi, L., E. Petit-Teixeira, M. Sebbag, F. Bongiorni, S. Moscato, F. Pratesi, C. Pierlot, J. Osorio, S. Chapuy-Regaud, M. Guerrin, F. Cornelis, G. Serre, and P. Migliorini. 2005. A family based study shows no association between rheumatoid arthritis and the PADI4 gene in a white French population. *Ann. Rheum. Dis.* 64:587.

165. Poor, G., Z. B. Nagy, Z. Schmidt, M. Brozik, K. Meretey, and P. Gergely, Jr. 2007. Genetic background of anticyclic citrullinated peptide autoantibody production in Hungarian patients with rheumatoid arthritis. *Ann. N. Y. Acad. Sci.* 1110:23.

166. Plenge, R. M., L. Padyukov, E. F. Remmers, S. Purcell, A. T. Lee, E. W. Karlson, F. Wolfe, D. L. Kastner, L. Alfredsson, D. Altshuler, P. K. Gregersen, L. Klareskog, and J. D. Rioux. 2005. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. *Am. J. Hum. Genet.* 77:1044.

167. Harney, S. M., C. Meisel, A. M. Sims, P. Y. Woon, B. P. Wordsworth, and M. A. Brown. 2005. Genetic and genomic studies of PADI4 in rheumatoid arthritis. *Rheumatology. (Oxford)* 44:869.

168. Kochi, Y., R. Yamada, A. Suzuki, J. B. Harley, S. Shirasawa, T. Sawada, S. C. Bae, S. Tokuhiro, X. Chang, A. Sekine, A. Takahashi, T. Tsunoda, Y. Ohnishi, K. M. Kaufman, C. P. Kang, C. Kang, S. Otsubo, W. Yumura, A. Mimori, T. Koike, Y. Nakamura, T. Sasazuki, and K. Yamamoto. 2005. A functional variant in FCRL3, encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. *Nat. Genet.* 37:478.

169. Newman, W. G., Q. Zhang, X. Liu, E. Walker, H. Ternan, J. Owen, B. Johnson, W. Greer, D. P. Mosher, W. P. Maksymowych, V. P. Bykerk, E. C. Keystone, C. I. Amos, and K. A. Siminovitch. 2006. Rheumatoid arthritis association with the FCRL3 -169C polymorphism is restricted to PTPN22 1858T-homozygous individuals in a Canadian population. *Arthritis Rheum.* 54:3820.

170. Thabet, M. M., J. Wesoly, P. E. Slagboom, R. E. Toes, and T. W. Huizinga. 2007. FCRL3 promoter 169 CC homozygosity is associated with susceptibility to rheumatoid arthritis in Dutch Caucasians. *Ann. Rheum. Dis.* 66:803.

171. Eike, M. C., G. B. Nordang, T. H. Karlsen, K. M. Boberg, M. H. Vatn, K. hl-Jorgensen, K. S. Ronningen, G. Joner, B. Flato, A. Bergquist, E. Thorsby, O. Forre, T. K. Kvien, D. E. Undlien, and B. A. Lie. 2008. The FCRL3 -169T>C polymorphism is associated with rheumatoid arthritis and shows suggestive evidence of involvement with juvenile idiopathic arthritis in a Scandinavian panel of autoimmune diseases. *Ann. Rheum. Dis.* 67:1287.

172. Chae, S. C., J. H. Song, S. C. Shim, K. S. Yoon, and H. T. Chung. 2004. The exon 4 variations of Tim-1 gene are associated with rheumatoid arthritis in a Korean population. *Biochem. Biophys. Res. Commun.* 315:971.

173. Rodriguez, M. R., A. Nunez-Roldan, F. Aguilar, A. Valenzuela, A. Garcia, and M. F. Gonzalez-Escribano. 2002. Association of the CTLA4 3' untranslated region polymorphism with the susceptibility to rheumatoid arthritis. *Hum. Immunol.* 63:76.

174. Tokuhiro, S., R. Yamada, X. Chang, A. Suzuki, Y. Kochi, T. Sawada, M. Suzuki, M. Nagasaki, M. Ohtsuki, M. Ono, H. Furukawa, M. Nagashima, S. Yoshino, A. Mabuchi, A. Sekine, S. Saito, A. Takahashi, T. Tsunoda, Y. Nakamura, and K. Yamamoto. 2003. An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat. Genet.* 35:341.

175. Plenge, R. M., M. Seielstad, L. Padyukov, A. T. Lee, E. F. Remmers, B. Ding, A. Liew, H. Khalili, A. Chandrasekaran, L. R. Davies, W. Li, A. K. Tan, C. Bonnard, R. T. Ong, A. Thalamuthu, S. Pettersson, C. Liu, C. Tian, W. V. Chen, J. P. Carulli, E. M. Beckman, D. Altshuler, L. Alfredsson, L. A. Criswell, C. I. Amos, M. F. Seldin, D. L. Kastner, L. Klareskog, and P. K. Gregersen. 2007. TRAF1-C5 as a risk locus for rheumatoid arthritis--a genomewide study. *N. Engl. J. Med.* 357:1199.

176. Harrison, P., J. J. Pointon, K. Chapman, A. Roddam, and B. P. Wordsworth. 2008. Interleukin-1 promoter region polymorphism role in rheumatoid arthritis: a meta-analysis of IL-1B-511A/G variant reveals association with rheumatoid arthritis. *Rheumatology. (Oxford)* 47:1768.

177. Lin, S. C., J. H. Yen, J. J. Tsai, W. C. Tsai, T. T. Ou, H. W. Liu, and C. J. Chen. 2004. Association of a programmed death 1 gene polymorphism with the development of rheumatoid arthritis, but not systemic lupus erythematosus. *Arthritis Rheum.* 50:770.

178. Prokunina, L., L. Padyukov, A. Bennet, F. U. de, B. Wiman, J. Prince, L. Alfredsson, L. Klareskog, and M. arcon-Riquelme. 2004. Association of the PD-1.3A allele of the PDCD1 gene in patients with rheumatoid arthritis negative for rheumatoid factor and the shared epitope. *Arthritis Rheum.* 50:1770.

179. Darvasi, A., and M. Soller. 1995. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics* 141:1199.

180. Darvasi, A. 1998. Experimental strategies for the genetic dissection of complex traits in animal models. *Nat. Genet.* 18:19.

181. Gasser, D. L., C. M. Newlin, J. Palm, and N. K. Gonatas. 1973. Genetic control of susceptibility to experimental allergic encephalomyelitis in rats. *Science* 181:872.

182. Dahlman, I., E. Wallstrom, H. Jiao, H. Luthman, T. Olsson, and R. Weissert. 2001. Polygenic control of autoimmune peripheral nerve inflammation in rat. *J. Neuroimmunol.* 119:166.

183. Sun, B., S. H. Sun, C. C. Chan, and R. R. Caspi. 2000. Evaluation of in vivo cytokine expression in EAU-susceptible and resistant rats: a role for IL-10 in resistance? *Exp. Eye Res.* 70:493.

184. Penhale, W. J., A. Farmer, S. J. Urbaniak, and W. J. Irvine. 1975. Susceptibility of inbred rat strains to experimental thyroiditis: quantitation of thyroglobulin-binding cells and assessment of T-cell function in susceptible and non-susceptible strains. *Clin. Exp. Immunol.* 19:179.

185. Vingsbo, C., P. Sahlstrand, J. G. Brun, R. Jonsson, T. Saxne, and R. Holmdahl. 1996. Pristane-induced arthritis in rats: a new model for rheumatoid arthritis with a chronic disease course influenced by both major histocompatibility complex and non-major histocompatibility complex genes. *Am. J. Pathol.* 149:1675.

186. Lorentzen, J. C., and L. Klareskog. 1997. Comparative susceptibility of DA, LEW, and LEW.1AV1 rats to arthritis induced with different arthritogens: mineral

oil, mycobacteria, muramyl dipeptide, avridine and rat collagen type II. *Transplant. Proc.* 29:1692.

187. Lorentzen, J. C. 1999. Identification of arthritogenic adjuvants of self and foreign origin. *Scand. J. Immunol.* 49:45.

188. Cannon, G. W., M. L. Woods, F. Clayton, and M. M. Griffiths. 1993. Induction of arthritis in DA rats by incomplete Freund's adjuvant. *J. Rheumatol.* 20:7.

189. Backdahl, L., U. Ribbhammar, and J. C. Lorentzen. 2003. Mapping and functional characterization of rat chromosome 4 regions that regulate arthritis models and phenotypes in congenic strains. *Arthritis Rheum.* 48:551.

190. Billiau, A., and P. Matthys. 2001. Modes of action of Freund's adjuvants in experimental models of autoimmune diseases. *J. Leukoc. Biol.* 70:849.

191. Lorentzen, J. C., H. Erlandsson, A. Mussener, L. Mattsson, S. Kleinau, U. Nyman, and L. Klareskog. 1995. Specific and long-lasting protection from collagen-induced arthritis and oil-induced arthritis in DA rats by administration of immunogens. *Scand. J. Immunol.* 42:82.

192. Kleinau, S., and L. Klareskog. 1993. Oil-induced arthritis in DA rats passive transfer by T cells but not with serum. *J. Autoimmun.* 6:449.

193. Svelander, L., A. Mussener, H. Erlandsson-Harris, and S. Kleinau. 1997. Polyclonal Th1 cells transfer oil-induced arthritis. *Immunology* 91:260.

194. Lorentzen, J. C., and L. Klareskog. 1996. Susceptibility of DA rats to arthritis induced with adjuvant oil or rat collagen is determined by genes both within and outside the major histocompatibility complex. *Scand. J. Immunol.* 44:592.

195. Lorentzen, J. C., A. Glaser, L. Jacobsson, J. Galli, H. Fakhrai-rad, L. Klareskog, and H. Luthman. 1998. Identification of rat susceptibility loci for adjuvant-oil-induced arthritis. *Proc. Natl. Acad. Sci. U. S. A* 95:6383.

196. Olofsson, P., J. Holmberg, J. Tordsson, S. Lu, B. Akerstrom, and R. Holmdahl. 2003. Positional identification of Ncf1 as a gene that regulates arthritis severity in rats. *Nat. Genet.* 33:25.

197. Hultqvist, M., P. Olofsson, J. Holmberg, B. T. Backstrom, J. Tordsson, and R. Holmdahl. 2004. Enhanced autoimmunity, arthritis, and encephalomyelitis in mice with a reduced oxidative burst due to a mutation in the Ncf1 gene. *Proc. Natl. Acad. Sci. U. S. A* 101:12646.

198. Flint, J., W. Valdar, S. Shifman, and R. Mott. 2005. Strategies for mapping and cloning quantitative trait genes in rodents. *Nat. Rev. Genet.* 6:271.

199. Johannesson, M., J. Karlsson, P. Wernhoff, K. S. Nandakumar, A. K. Lindqvist, L. Olsson, A. D. Cook, A. Andersson, and R. Holmdahl. 2005. Identification of epistasis through a partial advanced intercross reveals three arthritis loci within the Cia5 QTL in mice. *Genes Immun.* 6:175.

200. Barton, A., S. Eyre, A. Myerscough, B. Brintnell, D. Ward, W. E. Ollier, J. C. Lorentzen, L. Klareskog, A. Silman, S. John, and J. Worthington. 2001. High resolution linkage and association mapping identifies a novel rheumatoid arthritis susceptibility locus homologous to one linked to two rat models of inflammatory arthritis. *Hum. Mol. Genet.* 10:1901.

201. Yokoyama, W. M., J. C. Ryan, J. J. Hunter, H. R. Smith, M. Stark, and W. E. Seaman. 1991. cDNA cloning of mouse NKR-P1 and genetic linkage with LY-49. Identification of a natural killer cell gene complex on mouse chromosome 6. *J. Immunol.* 147:3229.

202. Hao, L., J. Klein, and M. Nei. 2006. Heterogeneous but conserved natural killer receptor gene complexes in four major orders of mammals. *Proc. Natl. Acad. Sci. U. S. A* 103:3192.

203. Zelensky, A. N., and J. E. Gready. 2004. C-type lectin-like domains in Fugu rubripes. *BMC. Genomics* 5:51.

204. Kikuno, R., A. Sato, W. E. Mayer, S. Shintani, T. Aoki, and J. Klein. 2004. Clustering of C-type lectin natural killer receptor-like loci in the bony fish Oreochromis niloticus. *Scand. J. Immunol.* 59:133.

205. Panagos, P. G., K. P. Dobrinski, X. Chen, A. W. Grant, D. Traver, J. Y. Djeu, S. Wei, and J. A. Yoder. 2006. Immune-related, lectin-like receptors are differentially expressed in the myeloid and lymphoid lineages of zebrafish. *Immunogenetics* 58:31.

206. Richard, M., N. Thibault, P. Veilleux, G. Gareau-Page, and A. D. Beaulieu. 2006. Granulocyte macrophage-colony stimulating factor reduces the affinity of SHP-2 for the ITIM of CLECSF6 in neutrophils: a new mechanism of action for SHP-2. *Mol. Immunol.* 43:1716.

207. Burshtyn, D. N., W. Yang, T. Yi, and E. O. Long. 1997. A novel phosphotyrosine motif with a critical amino acid at position -2 for the SH2 domain-mediated activation of the tyrosine phosphatase SHP-1. *J. Biol. Chem.* 272:13066.

208. Richard, M., N. Thibault, P. Veilleux, R. Breton, and A. D. Beaulieu. 2003. The ITIM-bearing CLECSF6 (DCIR) is down-modulated in neutrophils by neutrophil activating agents. *Biochem. Biophys. Res. Commun.* 310:767.

209. Dudziak, D., A. O. Kamphorst, G. F. Heidkamp, V. R. Buchholz, C. Trumpfheller, S. Yamazaki, C. Cheong, K. Liu, H. W. Lee, C. G. Park, R. M. Steinman, and M. C. Nussenzweig. 2007. Differential antigen processing by dendritic cell subsets in vivo. *Science* 315:107.

210. Yamasaki, S., E. Ishikawa, M. Sakuma, H. Hara, K. Ogata, and T. Saito. 2008. Mincle is an ITAM-coupled activating receptor that senses damaged cells. *Nat. Immunol.*

211. Sato, K., X. L. Yang, T. Yudate, J. S. Chung, J. Wu, K. Luby-Phelps, R. P. Kimberly, D. Underhill, P. D. Cruz, Jr., and K. Ariizumi. 2006. Dectin-2 is a

pattern recognition receptor for fungi that couples with the Fc receptor gamma chain to induce innate immune responses. *J. Biol. Chem.* 281:38854.

212. Rock, J., E. Schneider, J. R. Grun, A. Grutzkau, R. Kuppers, J. Schmitz, and G. Winkels. 2007. CD303 (BDCA-2) signals in plasmacytoid dendritic cells via a BCR-like signalosome involving Syk, Slp65 and PLCgamma2. *Eur. J. Immunol.* 37:3564.

213. Feng, J., M. E. Call, and K. W. Wucherpfennig. 2006. The assembly of diverse immune receptors is focused on a polar membrane-embedded interaction site. *PLoS. Biol.* 4:e142.

214. von, H. G. 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* 341:456.

215. McGreal, E. P., M. Rosas, G. D. Brown, S. Zamze, S. Y. Wong, S. Gordon, L. Martinez-Pomares, and P. R. Taylor. 2006. The carbohydrate-recognition domain of Dectin-2 is a C-type lectin with specificity for high mannose. *Glycobiology* 16:422.

216. Barrett, N. A., A. Maekawa, O. M. Rahman, K. F. Austen, and Y. Kanaoka. 2009. Dectin-2 recognition of house dust mite triggers cysteinyl leukotriene generation by dendritic cells. *J. Immunol.* 182:1119.

217. Bugarcic, A., K. Hitchens, A. G. Beckhouse, C. A. Wells, R. B. Ashman, and H. Blanchard. 2008. Human and mouse macrophage-inducible C-type lectin (Mincle) bind Candida albicans. *Glycobiology* 18:679.

218. Wells, C. A., J. A. Salvage-Jones, X. Li, K. Hitchens, S. Butcher, R. Z. Murray, A. G. Beckhouse, Y. L. Lo, S. Manzanero, C. Cobbold, K. Schroder, B. Ma, S. Orr, L. Stewart, D. Lebus, P. Sobieszczuk, D. A. Hume, J. Stow, H. Blanchard, and R. B. Ashman. 2008. The macrophage-inducible C-type lectin, mincle, is an essential component of the innate immune response to Candida albicans. *J. Immunol.* 180:7404.

219. Yamasaki, S., M. Matsumoto, O. Takeuchi, T. Matsuzawa, E. Ishikawa, M. Sakuma, H. Tateno, J. Uno, J. Hirabayashi, Y. Mikami, K. Takeda, S. Akira, and T. Saito. 2009. C-type lectin Mincle is an activating receptor for pathogenic fungus, Malassezia. *Proc. Natl. Acad. Sci. U. S. A* 106:1897.

220. Ishikawa, E., T. Ishikawa, Y. S. Morita, K. Toyonaga, H. Yamada, O. Takeuchi, T. Kinoshita, S. Akira, Y. Yoshikai, and S. Yamasaki. 2009. Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J. Exp. Med.* 206:2879.

221. Schoenen, H., B. Bodendorfer, K. Hitchens, S. Manzanero, K. Werninghaus, F. Nimmerjahn, E. M. Agger, S. Stenger, P. Andersen, J. Ruland, G. D. Brown, C. Wells, and R. Lang. 2010. Cutting edge: mincle is essential for recognition and adjuvanticity of the mycobacterial cord factor and its synthetic analog trehalose-dibehenate. *J. Immunol.* 184:2756.

222. Arce, I., L. Martinez-Munoz, P. Roda-Navarro, and E. Fernandez-Ruiz. 2004. The human C-type lectin CLECSF8 is a novel monocyte/macrophage endocytic receptor. *Eur. J. Immunol.* 34:210.

223. Dzionek, A., Y. Sohma, J. Nagafune, M. Cella, M. Colonna, F. Facchetti, G. Gunther, I. Johnston, A. Lanzavecchia, T. Nagasaka, T. Okada, W. Vermi, G. Winkels, T. Yamamoto, M. Zysk, Y. Yamaguchi, and J. Schmitz. 2001. BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon alpha/beta induction. *J. Exp. Med.* 194:1823.

224. Meyer-Wentrup, F., D. itez-Ribas, P. J. Tacken, C. J. Punt, C. G. Figdor, V. de, I, and G. J. Adema. 2008. Targeting DCIR on human plasmacytoid dendritic cells results in antigen presentation and inhibits IFN-alpha production. *Blood* 111:4245.

225. Kaden, S. A., S. Kurig, K. Vasters, K. Hofmann, K. S. Zaenker, J. Schmitz, and G. Winkels. 2009. Enhanced dendritic cell-induced immune responses mediated by the novel C-type lectin receptor mDCAR1. *J. Immunol.* 183:5069.

226. Nakamura, N., Y. Shimaoka, T. Tougan, H. Onda, D. Okuzaki, H. Zhao, A. Fujimori, N. Yabuta, I. Nagamori, A. Tanigawa, J. Sato, T. Oda, K. Hayashida, R. Suzuki, M. Yukioka, H. Nojima, and T. Ochi. 2006. Isolation and expression profiling of genes upregulated in bone marrow-derived mononuclear cells of rheumatoid arthritis patients. *DNA Res.* 13:169.

227. Migita, K., T. Miyashita, Y. Maeda, H. Kimura, M. Nakamura, H. Yatsuhashi, H. Ishibashi, and K. Eguchi. 2005. Reduced blood BDCA-2+ (lymphoid) and CD11c+ (myeloid) dendritic cells in systemic lupus erythematosus. *Clin. Exp. Immunol.* 142:84.

228. Sanna, A., Y. M. Huang, G. Arru, M. L. Fois, H. Link, G. Rosati, and S. Sotgiu. 2008. Multiple sclerosis: reduced proportion of circulating plasmacytoid dendritic cells expressing BDCA-2 and BDCA-4 and reduced production of IL-6 and IL-10 in response to herpes simplex virus type 1. *Mult. Scler.* 14:1199.

229. Fujikado, N., S. Saijo, T. Yonezawa, K. Shimamori, A. Ishii, S. Sugai, H. Kotaki, K. Sudo, M. Nose, and Y. Iwakura. 2008. Dcir deficiency causes development of autoimmune diseases in mice due to excess expansion of dendritic cells. *Nat. Med.* 14:176.

230. Guo, J. P., L. Backdahl, M. Marta, L. Mathsson, J. Ronnelid, and J. C. Lorentzen. 2008. Profound and paradoxical impact on arthritis and autoimmunity of the rat antigen-presenting lectin-like receptor complex. *Arthritis Rheum.* 58:1343.

231. Thakker, P., M. W. Leach, W. Kuang, S. E. Benoit, J. P. Leonard, and S. Marusic. 2007. IL-23 is critical in the induction but not in the effector phase of experimental autoimmune encephalomyelitis. *J. Immunol.* 178:2589.

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

233. Chabaud, M., J. M. Durand, N. Buchs, F. Fossiez, G. Page, L. Frappart, and P. Miossec. 1999. Human interleukin-17: A T cell-derived proinflammatory cytokine produced by the rheumatoid synovium. *Arthritis Rheum.* 42:963.

234. Wong, C. K., C. Y. Ho, E. K. Li, and C. W. Lam. 2000. Elevation of proinflammatory cytokine (IL-18, IL-17, IL-12) and Th2 cytokine (IL-4) concentrations in patients with systemic lupus erythematosus. *Lupus* 9:589.

235. Kurasawa, K., K. Hirose, H. Sano, H. Endo, H. Shinkai, Y. Nawata, K. Takabayashi, and I. Iwamoto. 2000. Increased interleukin-17 production in patients with systemic sclerosis. *Arthritis Rheum.* 43:2455.

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

237. Veldhoen, M., R. J. Hocking, C. J. Atkins, R. M. Locksley, and B. Stockinger. 2006. TGFbeta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity.* 24:179.

238. Mangan, P. R., L. E. Harrington, D. B. O'Quinn, W. S. Helms, D. C. Bullard, C. O. Elson, R. D. Hatton, S. M. Wahl, T. R. Schoeb, and C. T. Weaver. 2006. Transforming growth factor-beta induces development of the T(H)17 lineage. *Nature* 441:231.

239. Gutcher, I., and B. Becher. 2007. APC-derived cytokines and T cell polarization in autoimmune inflammation. *J. Clin. Invest* 117:1119.

240. Sutton, C., C. Brereton, B. Keogh, K. H. Mills, and E. C. Lavelle. 2006. A crucial role for interleukin (IL)-1 in the induction of IL-17-producing T cells that mediate autoimmune encephalomyelitis. *J. Exp. Med.* 203:1685.

241. Kryczek, I., S. Wei, L. Vatan, J. Escara-Wilke, W. Szeliga, E. T. Keller, and W. Zou. 2007. Cutting edge: opposite effects of IL-1 and IL-2 on the regulation of IL-17+ T cell pool IL-1 subverts IL-2-mediated suppression. *J. Immunol.* 179:1423.

242. Chung, Y., S. H. Chang, G. J. Martinez, X. O. Yang, R. Nurieva, H. S. Kang, L. Ma, S. S. Watowich, A. M. Jetten, Q. Tian, and C. Dong. 2009. Critical regulation of early Th17 cell differentiation by interleukin-1 signaling. *Immunity.* 30:576.

243. costa-Rodriguez, E. V., G. Napolitani, A. Lanzavecchia, and F. Sallusto. 2007. Interleukins 1beta and 6 but not transforming growth factor-beta are essential for the differentiation of interleukin 17-producing human T helper cells. *Nat. Immunol.* 8:942.

244. Wilson, N. J., K. Boniface, J. R. Chan, B. S. McKenzie, W. M. Blumenschein, J. D. Mattson, B. Basham, K. Smith, T. Chen, F. Morel, J. C. Lecron, R. A. Kastelein, D. J. Cua, T. K. McClanahan, E. P. Bowman, and M. R. de Waal. 2007. Development, cytokine profile and function of human interleukin 17-producing helper T cells. *Nat. Immunol.* 8:950.

245. Holm, B. C., J. C. Lorentzen, and A. Bucht. 2004. Adjuvant oil induces waves of arthritogenic lymph node cells prior to arthritis onset. *Clin. Exp. Immunol.* 137:59.

246. Nordquist, N., P. Olofsson, C. Vingsbo-Lundberg, U. Petterson, and R. Holmdahl. 2000. Complex genetic control in a rat model for rheumatoid arthritis. *J. Autoimmun.* 15:425.

247. Olofsson, P., S. Lu, J. Holmberg, T. Song, P. Wernhoff, U. Pettersson, and R. Holmdahl. 2003. A comparative genetic analysis between collagen-induced arthritis and pristane-induced arthritis. *Arthritis Rheum.* 48:2332.

248. Griffiths, M. M., J. Wang, B. Joe, S. Dracheva, Y. Kawahito, J. S. Shepard, V. R. Reese, S. Call-Vining, A. Hashiramoto, G. W. Cannon, E. F. Remmers, and R. L. Wilder. 2000. Identification of four new quantitative trait loci regulating arthritis severity and one new quantitative trait locus regulating autoantibody production in rats with collagen-induced arthritis. *Arthritis Rheum.* 43:1278.

249. Furuya, T., J. L. Salstrom, S. Call-Vining, G. W. Cannon, B. Joe, E. F. Remmers, M. M. Griffiths, and R. L. Wilder. 2000. Genetic dissection of a rat model for rheumatoid arthritis: significant gender influences on autosomal modifier loci. *Hum. Mol. Genet.* 9:2241.

250. Kawahito, Y., G. W. Cannon, P. S. Gulko, E. F. Remmers, R. E. Longman, V. R. Reese, J. Wang, M. M. Griffiths, and R. L. Wilder. 1998. Localization of quantitative trait loci regulating adjuvant-induced arthritis in rats: evidence for genetic factors common to multiple autoimmune diseases. *J. Immunol.* 161:4411.

251. Dahlman, I., J. C. Lorentzen, K. L. de Graaf, A. Stefferl, C. Linington, H. Luthman, and T. Olsson. 1998. Quantitative trait loci disposing for both experimental arthritis and encephalomyelitis in the DA rat; impact on severity of myelin oligodendrocyte glycoprotein-induced experimental autoimmune encephalomyelitis and antibody isotype pattern. *Eur. J. Immunol.* 28:2188.

252. Martin, A. M., E. P. Blankenhorn, M. N. Maxson, M. Zhao, J. Leif, J. P. Mordes, and D. L. Greiner. 1999. Non-major histocompatibility complex-linked diabetes susceptibility loci on chromosomes 4 and 13 in a backcross of the DP-BB/Wor rat to the WF rat. *Diabetes* 48:50.

253. Sun, S. H., P. B. Silver, R. R. Caspi, Y. Du, C. C. Chan, R. L. Wilder, and E. F. Remmers. 1999. Identification of genomic regions controlling experimental autoimmune uveoretinitis in rats. *Int. Immunol.* 11:529.

254. Becanovic, K., L. Backdahl, E. Wallstrom, F. boul-Enein, H. Lassmann, T. Olsson, and J. C. Lorentzen. 2003. Paradoxical effects of arthritis-regulating chromosome 4 regions on myelin oligodendrocyte glycoprotein-induced encephalomyelitis in congenic rats. *Eur. J. Immunol.* 33:1907.

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

256. Bottini, N., L. Musumeci, A. Alonso, S. Rahmouni, K. Nika, M. Rostamkhani, J. MacMurray, G. F. Meloni, P. Lucarelli, M. Pellecchia, G. S. Eisenbarth, D. Comings, and T. Mustelin. 2004. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nat. Genet.* 36:337.

257. Kyogoku, C., C. D. Langefeld, W. A. Ortmann, A. Lee, S. Selby, V. E. Carlton, M. Chang, P. Ramos, E. C. Baechler, F. M. Batliwalla, J. Novitzke, A. H. Williams, C. Gillett, P. Rodine, R. R. Graham, K. G. Ardlie, P. M. Gaffney, K. L. Moser, M. Petri, A. B. Begovich, P. K. Gregersen, and T. W. Behrens. 2004. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am. J. Hum. Genet.* 75:504.

258. Criswell, L. A., K. A. Pfeiffer, R. F. Lum, B. Gonzales, J. Novitzke, M. Kern, K. L. Moser, A. B. Begovich, V. E. Carlton, W. Li, A. T. Lee, W. Ortmann, T. W. Behrens, and P. K. Gregersen. 2005. Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. *Am. J. Hum. Genet.* 76:561.

259. Vang, T., A. V. Miletic, N. Bottini, and T. Mustelin. 2007. Protein tyrosine phosphatase PTPN22 in human autoimmunity. *Autoimmunity* 40:453.

260. Heward, J. M., A. Allahabadia, M. Armitage, A. Hattersley, P. M. Dodson, K. Macleod, J. Carr-Smith, J. Daykin, A. Daly, M. C. Sheppard, R. L. Holder, A. H. Barnett, J. A. Franklyn, and S. C. Gough. 1999. The development of Graves' disease and the CTLA-4 gene on chromosome 2q33. *J. Clin. Endocrinol. Metab* 84:2398.

261. Nistico, L., R. Buzzetti, L. E. Pritchard, A. B. Van der, C. Giovannini, E. Bosi, M. T. Larrad, M. S. Rios, C. C. Chow, C. S. Cockram, K. Jacobs, C. Mijovic, S. C. Bain, A. H. Barnett, C. L. Vandewalle, F. Schuit, F. K. Gorus, R. Tosi, P. Pozzilli, and J. A. Todd. 1996. The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Belgian Diabetes Registry. *Hum. Mol. Genet.* 5:1075.

262. Brand, O., S. Gough, and J. Heward. 2005. HLA , CTLA-4 and PTPN22 : the shared genetic master-key to autoimmunity? *Expert. Rev. Mol. Med.* 7:1.

I

The complete inventory of receptors encoded by the rat natural killer cell gene complex

Brief title: Rat natural killer cell gene complex

L. M. Flornes, Ø. Nylenna, P. C. Saether, M.R. Daws, E. Dissen and S. Fossum

Institute of Basic Medical Sciences, Department of Anatomy, University of Oslo

Brief title: The rat natural killer cell gene complex

Key words: NK cell gene complex, lectin-like receptors, phylogenesis

Correspondence: Sigbjørn Fossum, Institute of Basic Medical Sciences,

University of Oslo, P.O. Box 1105 Blindern, N-0317 Oslo, Norway

(Phone + 47 22 85 12 78; fax: + 47 22 85 12 78; e-mail: sigbjorn.fossum@medisin.uio.no)

Abbreviations: **APLEC**: antigen-presenting cell lectin-like receptor gene complex, **CLEC**: C-type lectin-like, **CLSF**: C-type lectin superfamily, **sDC**: intestinal lymph dendritic cells, **ITAM**: immunoreceptor tyrosine-based activation motif , **ITIM**: immunoreceptor tyrosine-based inhibition motif, **NKC**: natural killer cell gene complex, **pMΦ**: peritoneal macrophages, **TM**: transmembrane

Abstract

The natural killer cell gene complex (NKC) encodes receptors belonging to the C-type lectin superfamily expressed primarily by NK cells and other leukocytes. In the rat the chromosomal region which starts with the *Nkrp1a* locus and ends with the *Ly49i8* locus, is predicted to contain 67 group V C-type lectin superfamily genes, making it one of the largest congregation of paralogous genes in vertebrates. Based on physical proximity and phylogenetic relationships between these genes, the rat NKC can be divided into four major parts. We have previously reported the cDNA cloning of the majority of the genes belonging to the centromeric *Nkrp1/Clr* cluster and the two telomeric groups, the *Klre1* - *Klri2* and the *Ly49* clusters. Here we close the gap between the *Nkrp1/Clr* and the *Klre1* - *Klri2* clusters by presenting the cDNA-cloning and expression patterns of eight genes spanning from *Cd69* to *Dectin1*, including the novel *Clec2m* gene. The definition, organization and evolution of the rat NKC are discussed.

Introduction

In 1991 Yokoyama and co-workers showed that genes encoding the C-type lectin-like (CLEC) receptors NKR-P1 and Ly49 were genetically closely linked in the mouse and postulated the existence of a genetic region encoding functionally related C-type lectin superfamily (CLSF) members. As these receptors primarily were expressed by natural killer (NK) cells, it was named the NK cell gene complex (NKC) (Yokoyama et al. 1991). In the pursuit of identifying structural gene(s) associated with regulation of NK cell mediated alloreactivity in the rat, we subsequently identified the equivalent rat gene complex (Dissen et al. 1996). Remarkably, both in the mouse and the rat the genetic distance between the *Nkrp1a* and the *Ly49* genes was estimated to ~ 0.5 cM, based on analyses of recombinant inbred strains (Yokoyama et al. 1991) and parental to F_1 backcrosses (Dissen et al. 1996),

respectively. Assuming normal crossing-over frequencies this would correspond to a physical distance of less than one Mb. However, the first physical map of the NKC, based on pulsed field gel electrophoresis, indicated the region to be much larger (Dissen et al. 1996). In the rat other CLSF genes were rapidly mapped to the intervening region (Berg et al. 1998a), (Berg et al. 1998b), (Dissen et al. 1997), followed by still more genes in the mouse and, with the identification of a human NKC, in the human (Renedo et al. 1997), (Suto et al. 1997), (Hamann et al. 1997), (Colonna et al. 2000), (Sobanov et al. 2001), (Boles et al. 1999), (Plougastel et al. 2001). The releases of the sequences amassed by the international genome projects finally exposed the real size of this gene complex in different species. Thus, aided by the Rat Genome project, we could report that the genetic region spanning from the most centromeric *Nkrp1* to the most telomeric *Ly49* gene covered 3.3 Mb, with a predicted content of 67 CLSF genes (Nylenna et al. 2005a). However, although the conservation of sequence features among the many intraspecific paralogs and interspecific orthologs has greatly facilitated prediction of novel CLSF genes, *in silico* prediction is error-prone and uninformative as to whether the predicted genes are expressed.

Functional traits such as resistance to cytomegalovirus (rev. in (Webb et al. 2002) and fungal infection (rev. in(Sun and Zhao 2007) and association with celiac disease have been mapped to discrete genes in the NKC (Hue et al. 2004), and association to experimentally induced arthritis was mapped to the neighbouring gene complex, APLEC, also encoding CLSF receptors (Flornes et al. 2004), (Lorentzen et al. 2007). For other quantitative trait loci (QTL) mapped to this chromosomal region, including loci controlling experimentally induced inflammatory responses in the rat, the associated structural genes await identification. An accurate and complete inventory of the genetic content of this genetic region would represent a useful tool in their ultimate identification, and thus of importance for studying pathogenetic

mechanisms behind the human inflammatory and autoimmune diseases. We have previously described the cDNA cloning of the genes constituting the *Klre1* - *Klri2* cluster (Berg et al. 1998a), (Berg et al. 1998b), (Dissen et al. 1997),(Saether et al. 2005;Westgaard et al. 2003) and, barring pseudogenes, most of the genes belonging to the *Ly49* gene cluster (Naper et al. 2002), (Nylenna et al. 2005b). Together the two clusters make up the telomeric part of the rat NKC. From the centromeric end of the complex we recently reported the cDNA cloning of the *Nkrp1* genes and most of the *Clr* genes (Kveberg et al. 2009), leaving a gap between the most distal *Clr* gene and the *Lox1* (*Olr1*) gene, previously reported by others (Nagase et al. 1998) and situated immediately centromeric to the *Klre1* - *Klri2* cluster. Here we close this gap by presenting the cDNA cloning and expression patterns of eight genes spanning from *Cd69* to *Dectin1*, including a novel gene which we have called *Clec2m*. We also include expression analyses of the *Nkrp1* and *Clr* genes and present the genomic organization of the 67 predicted group V CLSF localized in the rat NKC.

Materials and Methods

cDNA cloning

As the genes reported here were cDNA cloned before the availability of the rat genome sequence, they were identified as described in (Flornes et al. 2004) by searching 1) the GenBank rat Trace archive and the EST database for sequences homologous to the published mouse and human receptors, using the NCBI BLAST program, 2) performing pairwise BLAST on recently released partially or fully sequenced rat BAC clones. Gene specific (nested) primers in the 5′- and 3′-untranslated regions (UTRs) were generated from the predicted sequences (primers shown in Supplemental Table 1). The 3′- UTR of *Clec12a* could not be identified and this gene was therefore cDNA cloned using RACE with nested 3’ primers from the GeneRacer™ kit (Invitrogen.com).

RNA from mouse and rat lymph node, bone marrow or testis were isolated with Dynabeads mRNA Direct kit (Invitrogen) and first strand cDNA synthesis was carried out with M-MLV reverse transcriptase RNase H^- (Promega) using 1 μg total RNA in a 20-μl reaction volume. PCR were performed on first strand cDNA using gene specific primers and *PfuTurbo* DNA polymerase (Stratagene) and the products were cloned into pCR2.1-TOPO vector (Invitrogen). For every gene three independent clones were sequenced. The sequences were analyzed using software supplied by the Norwegian EMBNet node at the Biotechnology Centre in Oslo, Norway.

Quantitative PCR

A panel of RNA from purified cells and tissues from PVG rats was prepared previously (Dissen et al. 1997). Briefly, cells and tissues were extracted by lysis in isothiocyanate (Gibco BRL) followed by ultracentrifugation on a cesium chloride (Gibco BRL) gradient and phenol/chloroform extractions. Dendritic cells were isolated by spontaneous migration from split ear-halves floating on RPMI for 48 hours, testis were dissected from rats at day 4 and week 7,5 and 15. RNA from dendritic cells and testis were isolated with RNeasy (Qiagen) and first strand cDNA synthesis for all samples were carried with Superscript II reverse transcriptase (Invitrogen) and random hexamers using 1μg total RNA in a 30 μl reaction. Amplification was performed on an Applied Biosystem 7900HT Fast Real Time PCR system (PerkinElmer) using the 5'nuclease assay and qPCR SuperMix with Rox (Invitrogen). Analyses were done with the SDS2.1 software (Applied Biosystems, PerkinElmer). Primers and probes were designed using PrimerExpress 3.0 (Applied Biosystems, PerkinElmer) and purchased from Eurofins MWG Operon (Ebersberg, Germany). Individual samples were run in triplicate, and the relative quantity of RNA for each target was normalised to RNA from the reference gene plasma membrane calcium ATPase 4 (PMCA4).

Bioinformatics

Sequence similarity searches were performed with BLAST programs, running on the NCBI or Ensembl websites. The phylograms were constructed with NJ plot based on alignments with the ClustalX program, with the pileup program in the GCG package (Accelrys Inc., San Diego, CA), and Bayesian inference methods (MrBayes – http://mrbayes.csit.fsu.edu)

Results and discussion

Sequences and expression patterns

Eight genes spanning the gap between the *Nkrp1/Clr* genes and the *Lox1* locus were cDNA cloned. In Fig. 1 the predicted amino acid sequences are presented together with their mouse and human counterparts, with closest sequence similarity (Fig. 2) and conserved chromosomal positions and orientation between the three species as orthology criteria.

Cd69* and *Clec2m. Whereas the majority of class V CLSF genes consist of six coding exons, *Cd69* and *Clec2m* have only five (Fig.1). Rat *Cd69* shows relatively high expression levels , as measured by RT-PCR, in dendritic cells derived from skin (sDC), peritoneal macrophages (pMΦ) and resting B and T cells as well as ConA activated lymphocytes (Fig. 3). The expression in resting lymphocytes is noteworthy. *Cd69* was first cloned in the human (Ziegler et al. 1993), where it originally was reported as expressed primarily on activated lymphocytes and on NK cells, neutrophils and platelets. The CD69 receptor has consequently been widely used as an *in vivo* as a marker for leukocyte activation and inflammatory responses. It has been extensively studied in the human and the mouse, where *in vitro* studies originally indicated proinflammatory function, but where more recent *in vivo* results indicate that it may act as a regulatory molecule, modulating inflammatory responses (rev. in (Sancho et al.

2005)). *Clec2m* has to our knowledge not previously been cDNA cloned. Hao et al. (Hao et al. 2006) labeled the *in silico* predicted rat and mouse gene *Clec15a* and *Clec15ap*, respectively, with the mouse gene predicted to be a pseudogene. We have cDNA cloned full-length versions of the gene both in the rat and the mouse (Fig.1), but were unable to identify it in the human. As the term *Clec15a* is an established synonym for *Klrg1* (*Mafa*), we propose the name *Clec2m*, denoting close kinship to the *Clr* (*Clec2d*) genes and high expression in macrophages (see below). The gene shows highest expression levels in sDC and pMΦ, and among tissues bone marrow and testis (Fig. 3).

Clec12a (Micl) and Clec12b Clec12a possesses a single immunoreceptor tyrosine-based inhibition motif (ITIM) sequence (IxYxxL) in the cytoplasmic tail. Rat Clec12b contains the ITIM-like sequence AxYxxL, whereas mouse and human Clec12b both have the consensus ITIM VxYxxL (Fig. 1). For CLEC12A the inhibitory function suggested by the ITIM has been demonstrated both in the human, where it was first identified as myeloid inhibitory C-type lectin-like receptor (MICL) (Marshall et al. 2004), and later in the mouse (rev. in (Huysamen and Brown 2009)). In the human and the mouse (Pyz et al. 2008) *Clec12a* is expressed predominantly on DC, macrophages and granulocytes, and down-regulated following stimulation with select TLR (toll-like receptors) ligands. In the rat, the highest expression levels of *Clec12a* were found in sDC and pMΦ, and among tissues in bone marrow and brain. Less is known about the closely related *Clec12b* (rev. in (Huysamen and Brown 2009). Its expression profile in the rat is distinct from that of *Clec12a*, with moderate expression levels in sDC and neutrophils, no detectable expression in pMΦ, and with strong expression in the testis (Fig. 3).

Clec1b*, *Clec9a* and *Dectin1 These receptors share the atypical immunoreceptor tyrosine-based activation motif (ITAM) D/ExYxxL in their cytoplasmic tail (Fig.1). In the human and the mouse all three have been shown to recruit and mediate activation signals via spleen tyrosine kinase (Syk) (Rogers et al. 2005), (Fuller et al. 2007), (Huysamen et al. 2008). Clec1b, first identified in the human and named CLEC-2 by Colonna et al. (Colonna et al. 2000), is expressed on various cell types including megakaryocytes and platelets, where it triggers platelet activation and aggregation (Suzuki-Inoue et al. 2006). Various exogenous and one endogenous ligand, the sialoglycoprotein podoplanin, have been identified (Suzuki-Inoue et al. 2007). In the rat, Clec1b expression levels are moderate in neutrophils, low in lymphocytes, undetectable in sDC and pMΦ and high in spleen and liver (Fig.3). Clec9a has recently been identified both in the human and the mouse as an activating receptor (Huysamen et al. 2008), (Caminschi et al. 2008), expressed primarily by a subset of monocytes and the rare $CD141^+$ subset of DC (Huysamen et al. 2008). In the rat, it shows high expression in the spleen, moderate to low expression levels of several other tissues and cell types, but no detectable expression in sDC or $CD4^+$ T cells (Fig.3). Whereas the physiological roles of Clec1b and Clec9a remain unknown, much insight has been gained about roles of Dendritic cell-associated C-type lectin 1 or Dectin1. It was first identified in the mouse (Ariizumi et al. 2000) and has been shown both in the mouse and in the human to function as the main leukocyte β-glucan binding receptor, with a major role in antifungal immunity (Brown 2006), (rev. in (Sun and Zhao 2007)). In the mouse Dectin1 is most strongly expressed on monocytes, macrophages, neutrophils and microglia, weakly on a T cell subset, and in the human also on B cells, mast cells and eosinophils (Brown 2006). In the rat, Dectin1 is expressed in all tissues and cell types tested (except the RNK16 cell line), with particularly strong expression in sDC, pMΦ, bone marrow and the thymus (Fig.3).

Clec1a Finally, although the human Clec1a receptor was first published ten years ago, under the name CLEC-1 (Colonna et al. 2000), little is known about its functional properties and roles. In the human, the cytoplasmic tail contains one tyrosine residue close to the N-terminus. This is also present in the mouse and the rat sequences, which have an additional tyrosine residue forming the pattern $YxxTx_{13}YxxT$(Fig. 1). Whether the tyrosines are subject to phosphorylation is not known. A notable sequence feature of Clec1 is the two cysteines immediately preceding the cysteines predicted to form the disulphide bond labeled 2 in fig. 1. This is a characteristic shared with the Ly49 receptors, where the two extra cysteines form an additional SS-bond between the α1 chain and the last β chain of the lectin domain. In the rat Clec1a is expressed in all cell types and tissues tested, with strongest expression in ConA-stimulated blasts, the spleen and the kidneys.

Comments on expression patterns

The presence of ITIM motifs in Clec12a (and possibly Clec12b) and (atypical) ITAM motifs in Clec1b and Clec9a suggests inhibitory and activating functions, as shown for the human and mouse orthologs. The occurrence of closely related receptors with opposite signaling properties is suggestive of opposing pairs, but this seems contradicted by their widely different expression profiles (Fig. 3). In Fig. 3 we have also included expression analyses of the *Clr* and the *Nkrp1* genes, as this information has not previously been published in the rat and is therefore needed for complete comparison of the expression programs of the rat NKC genes. The pattern of strong NK cell expression exhibited by NKR-P1A, -B and -F receptors is the rule for the majority of the rat NKC receptors previously reported. In contrast, most of the genes reported here, as well as most of the *Clr* genes, show modest to low or no expression in NK cells, demonstrating a genetic content beyond the original definition of this large congregation of genes as an NK cell gene complex.

Phylogenetic analyses

A striking feature of the eight receptors reported is the wide range of sequence divergences between the rat, mouse and human orthologs (Fig.2). The most extreme examples are *Clec1a*, which is highly conserved, and *Clec12a*, which is extensively changed between the three species (Figs 1 and 2). For Clec1a the rat and mouse protein per cent identity is 95.5 and the rodent versus human identities 71.3/69.4 (human-rat/human-mouse). For Clec12a the corresponding figures are 73.3 and 47.7/50.6. The dissimilarities between the three Clec12a sequences suggest positive selection, a notion further strengthened by Ka/Ks analyses of the cDNA sequences. Ka/Ks analysis compares the rate of non-synonymous to synonymous substitutions, with a higher value indicating possible positive selection. For *Clec1a* the mouse versus rat Ka/Ks ratio is 0.10, and for *Clec12a* 1.00. Splitting the *Clec12a* sequences into exons 1-3, encoding the cytoplasmic tail/transmembrane/neck domains, and exons 3 – 6, encoding the the lectin domain, i.e. the presumed ligand binding part of the receptor, gives Ka/Ks values of 0.57 and 1.42, respectively. The ligand of these receptors are unknown, but the analysis suggests that the Clec1a ligand is phylogenetically conserved, whereas the Clec12a receptors may be chasing a more rapidly evolving ligand.

Phylogenetic analyses indicate that CD69 and Clec2m belong to a different subfamily than the other six receptors (Fig. 2). On inclusion of the other CLSF receptors encoded by the rat NKC, they exhibit closest sequence similarity with the Clr subfamily (Fig. 4). Although sensitive to gap parameters and not significant according to bootstrap analyses, the association seems reasonable considering the physical localization of *Cd69* and *Clec2m* next to the *Clr* genes (Fig. 5) combined with the fact that they consist of only five coding exons, lacking a separate exon encoding the external membrane-proximal stalk. This is a property

they share with the *Clr* genes, and contrasts with all the other NKC CLSF genes, which generally consist of six coding exons (apart from *Lox1* which have eight coding exons).

As for the other six genes, the *Clec12* and *Clec1* genes tended to cluster together (Fig. 2). The shape of the tree is, however, highly sensitive to gap parameters and to inclusion of other NKC genes (Fig. 4). The tree shape in Fig. 2 is similar to that reported by Hao et al, except for *Clec9a*, which was assigned to an entirely different clade of CLSF genes (Hao et al. 2006). Arguments in favor of including *Clec9a* in the *Clec12 – Lox1* clade is the sharing of chromosomal localization and sequence features such as the atypical ITAM motif. Furthermore, a conserved feature of the NKC encoded CLSF receptors is the $Cx_{10}C$ loop of amino acids clamped by an SS-bridge between the flanking cysteines labeled 1 in Fig. 1. In front of these do Clec1a and –b, Clec12a and - b and Clec9a have two additional cysteines, the first encoded by a codon near the end of exon 3 (Fig. 1) and restricted to these receptors , whereas the second, encoded by the third codon in exon 4 (Fig. 1) is also present in CD94, Klrk1, the Nkg2 and the Klri receptors. The atypical ITAM motif and the two extra cysteines are likely to represent derived (apomorphic) rather than ancestral (plesiomorphic) states, so that the assigning of *Clec9a* to a different clade would implicate convergent evolution or sequence transfer through gene conversion. The simpler explanation is that their shared presence stems from phylogenetic kinship. Whether the extra cysteines are involved in inter- or intrachain disulphide bonding is not known.

The organization of genes in the rat NKC

The phylogenetic tree shown in Fig. 4 includes group II and group V CLSF members localized on rat chromosome 4 (for simplicity only eight Ly49 genes are shown, as all 34 cluster together in a single clade). In addition to the genes depicted in Fig.5 the tree includes

Mafa1 (*Klrg1*), *Mdl1* (*Clec5a*) and the seven genes grouped together in the Aplec cluster. The tree is seen to consist of six major branches. The exact branching is highly sensitive to gap parameters, but apart from the placement of *Clec9a* discussed above is identical to the tree presented by Hao et al (Hao et al. 2006) with respect to major branches. The interesting feature of the tree is that the gene content of the major branches correlates exactly with the chromosomal clustering of the genes (apart from *Mafa1* and *Mdl1*). In addition to supporting the soundness of the tree, the observation indicates that this large congregation of paralogous genes has evolved by local gene duplications without major reshufflings. At the local level chromosomal inversion events have occurred, as evidenced by the changes in gene orientations, as well as segmental duplications involving more than one gene in the case of the *Ly49* gene cluster, as previously discussed (Nylenna et al. 2005b). A notable feature is the conservation of orientation of the 41 moste telomeric genes of the complex, from *Klrk1* to *Ly49i8*. The evolutionary events leading to the gene arrangement of the *Nkrp1*/*Clr* cluster seem particularly complex. In addition to shifting gene orientations, the cluster consist of two separate blocks containing *Nkrp1* genes with opposite signaling function next to series of *Clr* genes with incongruent patterns of topographical versus sequence neighbourness.

The demarcation of the NKC

The chromosomal region starting with *Nkrp1a* and ending with *Ly49i8* contains almost all known group V CLSF genes in the rat. Exceptions are *Cd72*, localized on chromosome 5, and *Klrg1* (*Mafa1*), *Mdl1* (*Clec5a*) and *Clec2l*, all three on rat chromosome 4, at distances 7, 97 and 99 Mb centromeric to *Nkrp1a*, respectively. Furthermore, the region contains nothing but group V CLSF genes, with *Gabarapl1* (Gamma-aminobutyric acid receptor-associated protein-like 1) as the single known exception. Although heterogeneous in gene content it displays conserved gene organization and persistence of orthologous gene lineages across

four major mammalian orders, as described by Hao et al. (Hao et al. 2006). These authors also included in the NKC the *Mincle – Dcir* gene cluster (Fig. 5) which in all four orders occupy a chromosomal region close to, but clearly separate from the *Nkrp1a - Ly49i8* region. When we first described the genetic region containing the *Mincle – Dcir* gene cluster, which we named the antigen-presenting cell lectin-like receptor gene complex (APLEC), we presented arguments for and against including these genes in the NKC. Counterarguments were 1) The evolutionary distance between the two groups of genes. Whereas the APLEC genes are classified as group II CLSF (http://www.imperial.ac.uk/research/animallectins/), with conserved amino acid residues involved in calcium dependent saccharide binding (Weis et al. 1991), the group V CLSF have lost these amino acid motifs (Weis et al. 1998). 2) The distance separating the APLEC genes from the NKC. In the mouse and the rat it is 5 - 6 Mb, in the human, cow, horse and dog only ~ 1.5 Mb. In all these species the intervening region is packed with non-CLSF genes, which would be included in the NKC if defined as encompassing the APLEC, with loss of communicative precision and a potential source of confusion in the future mapping of functional traits to this chromosomal region. Even when narrowly defined the NKC represent one of the largest congregation of paralogous genes known in vertebrates.

Acknowledgements

This work was supported by The Norwegian Research Council, The Norwegian Cancer Society and Bergljot and Sigurd Skaugen’s fund. The expert technical assistance of W. Jensen and M. Lauritzen is gratefully acknowledged. The authors have no conflicting financial interests.

Supplementary Table 1: Primers used for cDNA cloning.

Gene	*Forward*	*Reverse*
rClr2	CATGTCTCCCTTGAATCCTAGGGAC	GTAGATAGCATGTTCTGTGTTCCTCGTG
rClr4	GAGATCTAACCATGTCTCCCTTGAATCCT	GACAAACTCTCTACCCAGGACACAAAAC
rClr6	GATCTAACCATGCCTCCCTTGAATCCT	GGACAAACTCTCTACCCAGAACACGAA
rClr7	CCAGGATTAGACAAAATTCCCTGTC	GTTAATGAGGCAGCTACTGTAGAGA
rClr10	TCTTCCTCTCAGCTCAGACATGGGT	TCTCACGCATGCTTTGGCACATTCC
rClr11	CCTCCTGTAGCAGAACCCAGCTT	GGAAAGAGTAGCACCCATAGATAGCAGA
rCD69	AGGGCACACCATAGACCACT	TGCTATGGCACAGTCACCTATACC
rClec2m	CTCAACCTCTGTAGTCACTGAAGCG	AGTTTAATTGCTGGCTTCTTAGTAAC
rClec12a	GACGGTTTCTGGAGCTTTCTCTGAAT	*
rClec12b	TGGGCCTGAAGAAATTGCAACA	GCTTCCTCCTTTCTTGCATGGGT
rClec1b	GAGCTTTGACCTCCTGGCAGTTCAAG	TTCTCTGACATCCTGCCGGTCCTT
rClec9a	GTGTTGGGCTCCTATGTTGAC	TGCTCTCTGACCTCTTTCGAAC
rClec1a	TCTCACCGGCATTTCTGCAGGAATC	CCTTCACGTTGCCCTCCCTTGACA
rDectin-1	ACATTCAAGTGCTCTGCCTACGTAGA	ACACGCATTCCACCTTTACACTCTC
mClec2m	CCTTTCCGCTTAAGAATTTTCTCAATCT	TACACATCACAGACAAAACGAGACACTG

Only the innermost pair of nested primers. * Clec12a was cloned using nested 3' primers from GeneRacer^{TM} kit (Invitrogen.com).

Supplementary Table2: Primers and probes used for qPCR.

Gene	Forward	Reverse	Probe (5'-FAM, 3'-TAMRA)
Nkrp1a	TCAGTGTGTATCACCTCCATCTCTTC	TTCAAAGCCAACCTGTGTGAAC	CCCAGATGCCTGTCGGTGCCC
Nkrp1b	GAGCCACCTCCATCTCTTTCC	CAAGCACAGCCGAGTTTCAG	CCCGCGCTGGCATCGATTG
Nkrp1f	CATTCCCATCTCTCTCCACAGAT	CAGCCAAGTTTCAGAGCCAAA	CTTGTCGGTGCCCACACTGGCA
Nrkp1g	TTTGGAAGTGGATGAATGGTTCT	GAGATGAGGGCACAGCTGTTT	ACGCATCACCGGCAAGAACGAAG
Clr2	GATCGGTCTGCACAGAGAGTCA	CGATGGGAACCAAATTGTTATATG	CATCACAGCCTTGGAAGTGGACAGACAAC
Clr4	TAACTGGACATTCAGCCAGGACTT	CCCTTTGTGTCTCTTTAGGAAATTCA	AGGCCGAGCTAGCACGATTTGACACC
Clr6	GCCAGACCCTCTGTAAGGAACA	CGGACTCCCTTTGTATCTCCTTAA	AGGCCGAGCTAGCACGATTTGACACC
Clr7	GAGAAACAGGTGAAATGCTCCAT	TGAGGACTAGGATCACTGCATAACA	TCTCAGAATCTTCTCCCCTGAATCTCCTGC
Clr10	ACGTGGGAAACTGGACATTCA	TCCTCCAGGCTGTCAAACAGA	CCTCGTGCATGGCATTAGATTCCCATC
Clr11	GGAGGGAGAAATGGGTAAAATGT	CCATAGCAGCAGTAAAGCTTAGCA	AAATGTCTCAGAATCGTCTCCCCCGAA
CD69	CGTTTCTCAGCCCAATCCA	GCTGTTATTGATTCGGAAAAGGA	TGTTTCAGTCCACCAGCATATCGCTTCA
Clec2m	GACCGAAAGCTCCCTCAGACT	AACACCAATTACCAGGATCCAAA	TCTCAGTCAAGCCACCACCGGTTAACA
Clec12a	AGCCAGAAGTCTGACAAATGTGGG	AAGATGCATCCTAAGGCCACCACT	AGTACCTTCTGCTGGTTCCTATTCACAACA
Clec12b	TGGCTGTGGGAGGATGGT	GATCCATTGAAACTGGCAAGCT	CGTTGCCCCCTCCAACCCTATTTAATG
Clec2	AAGCCCCGGAAACAAGCT	AGCCCCATGGTTGAGATCAG	TCAGCTGAGCCTGCCTCTTCTTGGTG
Clec9a	CCATTTTCTTGGGCATCAAGTT	TCCTGTCGGATGAGCCTTTC	TCCCAGGTGTCCTCTCTTGTAATGGAGCA
Clec1a	TGTATTCTGGAACATCCACTGTCA	TGGGCGCCAAACTGAAGA	ACGAGCAGAGCCCAGGCACTCAGA
Dectin1	CCTCCAAGGCATCCCAAA	TGGCATGCATGATCCAATTG	ACAGGCGTCTTTTCTGGACCTTGCC
Lox1	CCTATCTTTAGATGCCCAGTTACTAC	CCGATGTAATCCCATCCAAAA	TCTTACAAGCAACTTCCCATTCCACCTCC
Pmca4	TGGCGTTGCCCATCGT	GGCCTACTCTGTAAAGAAAATGATGA	CACAGGCGTCCAAGTGCCGGAC

Figure legends

Figure 1. Amino acid sequence alignments of the CLSF receptors.

r – rat, m – mouse, h – human. Dashes (-) means identity with top sequence. Points (.) indicate gaps. Ex1, Ex2, etc. on top of alignment show start site of corresponding exons. Transmembrane regions underlined. ITAMs and ITIMs indicated. Potential N-glycosylation sites boxed. Membrane external cysteines highlighted in yellow. Numbers connected with black lines on top of the CD69 and Dectin-1 sequences indicate SS-bonds as demonstrated for crystal structures of human CD69 and mouse Dectin-1 (http://www.ncbi.nlm.nih.gov/Genbank/). For the other genes the corresponding cysteines are similarly numbered. X above cysteines in Clec1a indicates probable extra disulphide bridge, which in addition to the bridge between the cysteines labelled 2, links the first α helix with the last β strand of the CTL domain. Genbank accession numbers: rat CD69 - GU357488, rat Clec2m - GU357487, mouse Clec2m - GU553093, rat Clec12a - GU357484, rat Clec12b - GU357483, rat Clec1b - GU357482 rat Clec9a - GU357486, rat Clec1a - GU357481, rat Dectin-1 - GU357485

Figure 2. Dendrogram of the novel receptors (plus Lox1) and their human and mouse orthologs (based on amino acid sequences). Numbers indicate bootstrap values (1000 reiterations) and Bayesian clade credibility values. Values shown only for branching points with bootstrap values > 900 or credibility values > 90.

Figure 3. Gene expression as measured by quantitative PCR.

The expressions of genes were calculated as the ratio of copy number of target gene to the copy number of endogenous reference gene, PMCA4. Thresholds for intermediate and high expression were set to 1/100 and 1 relative to PMCA4. Not detected: no PCR product

detected after 40 cycles. DC - dendritic cells, PM - peritoneal macrophages, RNK16 - NK cell line, LAK - Il2-stimulated NK cells. Similar results were obtained using the HPRT1 gene (Hypoxanthine guanine phosphoribosyl transferase 1) as endogenous reference gene.

Figure 4. Dendrogram based on deduced protein sequences of all the CLSF genes encoded by the rat NKC (only eight of the predicted 34 Ly49 receptors shown). Also included are Mafa-1/Klrg-1 and Mdl-1/Clec5a, plus the seven predicted functional Aplec genes. Vertical black lines indicate major branches of the NKC receptors. The genes containing five coding exons indicated with 5 ex. Only bootstrap values > 900 from 1000 reiterations shown.

Figure 5. Map of the rat NKC. Arrows indicate gene orientation. Numbers below the line indicate distance from the starting point (the Nkrp1 gene at 165.637 Mb according to the current release of the rat genome). The four major clusters indicated with colour codes: green: *Nkrp1/Clr* gene cluster (including *Cd69* and *Clec2m*), magenta: *Clec12a – Lox1* cluster, red: *Klre1 – Klri2* cluster, blue: *Ly49* cluster. The topography of the chromosomal map with respect to major clusters is seen to be congruent with the major branches of the phylogenetic tree in Fig. 4.

Reference List

Ariizumi K, Shen GL, Shikano S, Xu S, Ritter R, III, Kumamoto T, Edelbaum D, Morita A, Bergstresser PR, Takashima A (2000) Identification of a novel, dendritic cell-associated molecule, dectin- 1, by subtractive cDNA cloning. J. Biol. Chem. 275:20157-20167

Berg SF, Dissen E, Westgaard IH, Fossum S (1998a) Molecular characterization of rat NKR-P2, a lectin-like receptor expressed by NK cells and resting T cells. Int. Immunol. 10:379-385

Berg SF, Dissen E, Westgaard IH, Fossum S (1998b) Two genes in the rat homologous to human NKG2. Eur. J. Immunol. 28:444-450

Boles KS, Barten R, Kumaresan PR, Trowsdale J, Mathew PA (1999) Cloning of a new lectin-like receptor expressed on human NK cells. Immunogenetics 50:1-7

Brown GD (2006) Dectin-1: a signalling non-TLR pattern-recognition receptor. Nat. Rev. Immunol. 6:33-43

Caminschi I, Proietto AI, Ahmet F, Kitsoulis S, Shin TJ, Lo JC, Rizzitelli A, Wu L, Vremec D, van Dommelen SL, Campbell IK, Maraskovsky E, Braley H, Davey GM, Mottram P, van d, V, Jensen K, Lew AM, Wright MD, Heath WR, Shortman K, Lahoud MH (2008) The dendritic cell subtype-restricted C-type lectin Clec9A is a target for vaccine enhancement. Blood 112:3264-3273

Colonna M, Samaridis J, Angman L (2000) Molecular characterization of two novel C-type lectin-like receptors, one of which is selectively expressed in human dendritic cells. European Journal of Immunology 30:697-704

Dissen E, Berg SF, Westgaard IH, Fossum S (1997) Molecular characterization of a gene in the rat homologous to human CD94. Eur. J. Immunol. 27:2080-2086

Dissen E, Ryan JC, Seaman WE, Fossum S (1996) An autosomal dominant locus, Nka, mapping to the Ly-49 region of a rat natural killer (NK) gene complex, controls NK cell lysis of allogeneic lymphocytes. J. Exp. Med. 183:2197-2207

Flornes LM, Bryceson YT, Spurkland A, Lorentzen JC, Dissen E, Fossum S (2004) Identification of lectin-like receptors expressed by antigen presenting cells and neutrophils and their mapping to a novel gene complex. Immunogenetics 56:506-517

Fuller GL, Williams JA, Tomlinson MG, Eble JA, Hanna SL, Pohlmann S, Suzuki-Inoue K, Ozaki Y, Watson SP, Pearce AC (2007) The C-type lectin receptors CLEC-2 and Dectin-1, but not DC-SIGN, signal via a novel YXXL-dependent signaling cascade. J. Biol. Chem. 282:12397-12409

Hamann J, Montgomery KT, Lau S, Kucherlapati R, van Lier RA (1997) AICL: a new activation-induced antigen encoded by the human NK gene complex. Immunogenetics 45:295-300

Hao L, Klein J, Nei M (2006) Heterogeneous but conserved natural killer receptor gene complexes in four major orders of mammals. Proc. Natl. Acad. Sci. U. S. A 103:3192-3197

Hue S, Mention JJ, Monteiro RC, Zhang SL, Cellier C, Schmitz J, Verkarre V, Fodil N, Bahram S, Cerf-Bensussan N, Caillat-Zucman S (2004) A direct role for NKG2D/MICA interaction in villous atrophy during celiac disease. Immunity 21:367-377

Huysamen C, Brown GD (2009) The fungal pattern recognition receptor, Dectin-1, and the associated cluster of C-type lectin-like receptors. FEMS Microbiol. Lett. 290:121-128

Huysamen C, Willment JA, Dennehy KM, Brown GD (2008) CLEC9A is a novel activation C-type lectin-like receptor expressed on BDCA3+ dendritic cells and a subset of monocytes. J. Biol. Chem. 283:16693-16701

Kveberg L, Dai KZ, Westgaard IH, Daws MR, Fossum S, Naper C, Vaage JT (2009) Two major groups of rat NKR-P1 receptors can be distinguished based on chromosomal localization, phylogenetic analysis and Clr ligand binding. Eur. J. Immunol. 39:541-551

Lorentzen JC, Flornes L, Eklow C, Backdahl L, Ribbhammar U, Guo JP, Smolnikova M, Dissen E, Seddighzadeh M, Brookes AJ, Alfredsson L, Klareskog L, Padyukov L, Fossum S (2007) Association of arthritis with a gene complex encoding C-type lectin-like receptors. Arthritis Rheum. 56:2620-2632

Marshall AS, Willment JA, Lin HH, Williams DL, Gordon S, Brown GD (2004) Identification and characterization of a novel human myeloid inhibitory C-type lectin-like receptor (MICL) that is predominantly expressed on granulocytes and monocytes. J. Biol. Chem. 279:14792-14802

Nagase M, Abe J, Takahashi K, Ando J, Hirose S, Fujita T (1998) Genomic organization and regulation of expression of the lectin-like oxidized low-density lipoprotein receptor (LOX-1) gene. J. Biol. Chem. 273:33702-33707

Naper C, Hayashi S, Lovik G, Kveberg L, Niemi EC, Rolstad B, Dissen E, Ryan JC, Vaage JT (2002) Characterization of a novel killer cell lectin-like receptor (KLRH1) expressed by alloreactive rat NK cells. Journal of Immunology 168:5147-5154

Nylenna O, Flornes LM, Westgaard IH, Woon PY, Naper C, Vaage JT, Gauguier D, Ryan JC, Dissen E, Fossum S (2005a) Killer cell lectin-like receptors and the natural killer cell gene complex. Glycobiology and Medicine 564:23-24

Nylenna O, Naper C, Vaage JT, Woon PY, Gauguier D, Dissen E, Ryan JC, Fossum S (2005b) The genes and gene organization of the Ly49 region of the rat natural killer cell gene complex. European Journal of Immunology 35:261-272

Plougastel B, Dubbelde C, Yokoyama WM (2001) Cloning of Clr, a new family of lectin-like genes localized between mouse Nkrp1a and Cd69. Immunogenetics 53:209-214

Pyz E, Huysamen C, Marshall AS, Gordon S, Taylor PR, Brown GD (2008) Characterisation of murine MICL (CLEC12A) and evidence for an endogenous ligand. Eur. J. Immunol. 38:1157-1163

Renedo M, Arce I, Rodriguez A, Carretero M, Lanier LL, Lopez-Botet M, Fernandez-Ruiz E (1997) The human natural killer gene complex is located on chromosome 12p12- p13. Immunogenetics 46:307-311

Rogers NC, Slack EC, Edwards AD, Nolte MA, Schulz O, Schweighoffer E, Williams DL, Gordon S, Tybulewicz VL, Brown GD, Reis e Sousa (2005) Syk-dependent cytokine induction by Dectin-1 reveals a novel pattern recognition pathway for C type lectins. Immunity. 22:507-517

Saether PC, Westgaard IH, Flornes LM, Hoelsbrekken SE, Ryan JC, Fossum S, Dissen E (2005) Molecular cloning of KLRI1 and KLRI2, a novel pair of lectin-like natural killer-cell receptors with opposing signalling motifs. Immunogenetics 56:833-839

Sancho D, Gomez M, Sanchez-Madrid F (2005) CD69 is an immunoregulatory molecule induced following activation. Trends Immunol. 26:136-140

Sobanov Y, Bernreiter A, Derdak S, Mechtcheriakova D, Schweighofer B, Duchler M, Kalthoff F, Hofer E (2001) A novel cluster of lectin-like receptor genes expressed in monocytic, dendritic and endothelial cells maps close to the NK receptor genes in the human NK gene complex. European Journal of Immunology 31:3493-3503

Sun L, Zhao Y (2007) The biological role of dectin-1 in immune response. International Reviews of Immunology 26:349-364

Suto Y, Yabe T, Maenaka K, Tokunaga K, Tadokoro K, Juji T (1997) The human natural killer gene complex (NKC) is located on chromosome 12p13.1-p13.2. Immunogenetics 46:159-162

Suzuki-Inoue K, Fuller GL, Garcia A, Eble JA, Pohlmann S, Inoue O, Gartner TK, Hughan SC, Pearce AC, Laing GD, Theakston RD, Schweighoffer E, Zitzmann N, Morita T, Tybulewicz VL, Ozaki Y, Watson SP (2006) A novel Syk-dependent mechanism of platelet activation by the C-type lectin receptor CLEC-2. Blood 107:542-549

Suzuki-Inoue K, Kato Y, Inoue O, Kaneko MK, Mishima K, Yatomi Y, Yamazaki Y, Narimatsu H, Ozaki Y (2007) Involvement of the snake toxin receptor CLEC-2, in podoplanin-mediated platelet activation, by cancer cells. J. Biol. Chem. 282:25993-26001

Webb JR, Lee SH, Vidal SM (2002) Genetic control of innate immune responses against cytomegalovirus: MCMV meets its match. Genes and Immunity 3:250-262

Weis WI, Kahn R, Fourme R, Drickamer K, Hendrickson WA (1991) Structure of the Calcium-Dependent Lectin Domain from A Rat Mannose-Binding Protein Determined by Mad Phasing. Science 254:1608-1615

Weis WI, Taylor ME, Drickamer K (1998) The C-type lectin superfamily in the immune system. Immunological Reviews 163:19-34

Westgaard IH, Dissen E, Torgersen KM, Lazetic S, Lanier LL, Phillips JH, Fossum S (2003) The Lectin-like Receptor KLRE1 Inhibits Natural Killer Cell Cytotoxicity. J. Exp. Med. 197:1551-1561

Yokoyama WM, Ryan JC, Hunter JJ, Smith HR, Stark M, Seaman WE (1991) cDNA cloning of mouse NKR-P1 and genetic linkage with LY-49. Identification of a natural killer cell gene complex on mouse chromosome 6. J Immunol 147:3229-3236

Ziegler SF, Ramsdell F, Hjerrild KA, Armitage RJ, Grabstein KH, Hennen KB, Farrah T, Fanslow WC, Shevach EM, Alderson MR (1993) Molecular characterization of the early activation antigen CD69: a type II membrane glycoprotein related to a family of natural killer cell activation antigens. Eur. J. Immunol. 23:1643-1648

Figure 1

CD69

```
    Ex1                  Ex2
r   MNSEECSITENSSSHLERGQ RDHGTSVHFEKHREGSIQVPIPCAVLVVVLITSLIIALFALSV
m   -D--N--------------- K-----------H------S--W---------------I--N-
h   -S--N-F-A----L-P-S-- ENDA--P--STRH---F----L---MN--F--I-----I----

     Ex3                 1          1                2
r   GKYNCPGFYENLESFDHHAASCKNEWFSYNGKCYFFSTTTKTWALAQKSCSEDDATLAVIDSEKDM
m   -------L--K---S---V-T-----I--KRT---------S-----R-----A------------
h   -Q-----Q-TFSMPS-S-VS--SED-VG-QR----I--VKRS-TS--NA---HG------------

    Ex4                                 Ex5     3            3       2
r   AFLKRYAGGLKHWIGLRNEASQTWKWANGKEFNSW FNVTGSKKCVSLNHTDVASVDCEANLHWICSKASL
m   T-----S-E-E-----K---N-------------- --L---GR---V--KN-TA------F------P-R
h   N-------REE-----KK-PGHP---S------N- ------D---F-KN-E-S-ME--K--Y---N-PYK
```

Clec2m

```
    Ex1                      Ex2
r   MQTESSLRLSQSS....HHRLTFD LKKVVTLWILVIGVIAVLLWGFFSFPKKF
m   ---DT-P--P-T-HERR----M-- A---L---T-LG-A------------R--

    Ex3              1          1                2
r   TVTRQTKNEVCSDGVKICLHGWNKLNRNCFTHFSHANSWFTAKETCKFHDATLAVFNDQTEL
m   ----TAR-KT-D-E--L-PKD----RQ---SRIQ-K---L--ND--EL--------I-K--V

    Ex4                                   Ex5     3            3       2
r   DIVMKQMEDIQTFWIGLYKKDFRGPWVWTNGSKYNNW YEVQDYGHCAFLHKSGIDSTNCNDLKEYICTQEGHCP
m   E-L-N-IQEMK------HRQNLL-I-----------L H----H-Q---V-QK-----S-E-Q------R--Q--
```

Clec12a (Micla, Clec1c)

```
    Ex1 ITIM                        Ex2                               Ex3
r   MSEEIVYADLKFQDSDKKEESQKSDKCGRKV PSAGSYSQHKTVLILILLCLLLFIGVVALGCIF YTTLETEMIKSNQLQRVKEE
m   --------N--I--P-----T-------G-- SADA-H--Q----------------MGV--G-- ----A-----------A---
h   ----VT----Q--N-SEM-KIPEIG-F-E-A -P-P-HVWRPAA-F-T------L--LGV-ASM- HV--KI--K-M-K--N-S--

                                                Ex4  1          1                2
r   LQENVSLQLMHNLNNSKKIKTLSAMLQNIATQLCQELSKKEPG HKCKPCPKASDWYKDSCYSRFQKYATWQESVEFCSARNASLLKVKNKDEL
m   ---------K----S-----N------ST-----R--YS---E --------G-E--------QLNQ-G------MA---------------V-
h   --R-------S-M-I-N--RN--TT--T---K--R--YS--QE -------RRWI-H-----FLSDDVQ-----KMA-A-Q-------N--NA-

    Ex5                                   Ex6         3            3       2
r   EFIKSK.ELYNYWLALPPSKMYRSYELLSEKMFLSEG FKRSTYDITKMSCGFIRGEYVYYTNCDEEKYTMCEETASKVQVESVLSDLPEGSIL
m   ----Y-.K-RY-----L-R-DRTQ-P.---------E SE---D--D-KY----DRVN----Y-TD-NNII---------L----NG---D-R
h   -----QSRS-D---G-S-EEDSTRGMRVDNIINS-AW VI-NAP-LNN-Y----NRL--Q-YH-TYK-RMI--KM-NP--LG-TYFREA
```

Clec12b

```
    Ex1 ITIM                        Ex2                               Ex3
r   MSDEATYTTLMLQDSAGVRGNQDANNLRKEE CPAQSPLWRGAALSLMTLCLMLVTGLVTLATMF LQVSNDINSDSEKLSELQKIIHPQQDNLSESL
m   ----V--A--------R----R-G------G H----S--------------V------------ ---------------Q---S------------
h   --E-V--A--TF-----A-N-R-G-----RG H--P--I--H---G-V------LI-----GM-- ---------------Q---T-QQ------QQ-

                                              Ex4  1          1                 2
r   NNSRK.GLTEESLQSQISALLERQGQMATKLCKEFLIHASD HKCNPCPKTWQWHGNSCYYFSANEEKSWRDSRKDCTDKNATLVKIDSIEER
m   -S---.S-----------------E---------------- ------------Y--------I------S------I-----------T---
h   G--NNLSME--F-K----SV-K--E---I---Q-LI--T-- -R------M---YQ------TT----T-AN-----I---S-------L--K

    Ex5                                     Ex6           3            3       2
r   DLLQSQLSLTFSFFWLGLSWDSSSRNWLWEDGSLPPPTL FNEKELASFNGSRECAYFERGNIYVSRCSAEISWICEKTASLVKIEDLE
m   --------------------N--G---------F----- LSD---------------------T---R---P-----R---------D
h   -F-M--PL-M-------------G-S-F-----V-S-S- -ST---DQI---KG----QK------------F-------AP--T---D
```

Clec1b

```
   Ex1 ITAM               Ex2
r  MQDEDGYITLNIKPRKQALSSA EPA.SSWWRLTALILLISTMGLVAGLVALGIMS
m  ---------------------- ---.-----VM-------S----V---------
h  -------------T--P--I-V GS-S-----VM------LCV-M-V-------W-

   Ex3                                        Ex 4  1          1                 2
r  VTQQKYLLAEKENLSATLQQLAKKFCQELIRQSEIKTKSSFE HKCSPCATKWRYHGDSCYGFFRRNLTWEDSKLFCSEQNATLVKTASQSTL
m  ---------------------------------------T-- ---------------------------E--Q-Y-T---------------
h  -M-RN--QG-N--RTG-------R---YV-K---L-GTFKG. -----D-N---Y---------H-----E--Q-Y-TDM----L-IDNRNIV

   Ex5                                   Ex6         3            3       2
r  KYITDRITSVRWIGLSRQNSKKDWMWEDSSVLHNNS IDLSRNTEENMNCAYLHNGKIHPASCTERHYLICERNAALTRVEQLL
m  D-IAE---------------------------RK-G -N--G-------------------K-----------GM---D---
h  E-IKA-THLI--V-----K-NEV-K---G--ISE-M FEFLEDGKG------F----M--TF-ENK---M---K-GM-K-D--P
```

Clec9a

```
   Ex1 ITAM                        Ex2
r  MHEEEIYTSLQWDIPTSEASQKCPSLSKCPG TWCIVTVISCVVCVGLLAASIFLGIKF
m  --A--------------------Q-P---S- A--V--M------M----T--------
h  -------------S-APDTY---L-SN--S- AC-L-M-----F-M---T--------L

   Ex3                                              Ex4  1          1                 2
r  SQVSSLVMEQRERLIRQDTALLNLTEWQRNHTLQLKSCQASLQRSLRSG SNCNPCPPNWIQNGKSCYYAFDRWETWNNSKKSCLKEGDSLLQIDSKEEM
m  F------L--Q----Q-----V---Q---KY--EY..---L-----H-- -D-S---H-----------V-E---M--I---------A--F--------
h  L---TIA-Q-Q-K--Q-ER----F---K-SCA--M-Y---FM-N--S-A H-SS---N-----RE----VSEI-SI-HT-QEN-----ST----E-----

   Ex5                                        Ex6         3            3      2
r  FINLSIWKLKGGYEYWVGVFQDGPSGSWFWEDGSSPLSDL LPTDRQLSASQICGYLKDHTLISDNCSNWKYFICEKKAFGSCI
m  EFIS--G-----NK---------I---------------- --AE--R--G--------S-----K-DS--------------
h  D--TG-LR-I--S-D----LS---H--R-L-Q-----SPG --AE-SQ--N-V---V-SNS-L-S---T--------Y-LR-SV
```

Clec1a

```
   Ex1                                    Ex2                                Ex3
r  MQAKYSSTRDMLDDD.DTTISLYSGTSTVTRRAEPRHSE NGTPSSVWRPVALTLLTLCLVLLVGLAALGLVF FQFYQLSNTQQDSITEKDER
m  ---------------.----------------------- --------------------------------- --------I----------K
h  ---------------G---M--H-QA-AT--HP---RT- HRA---T--------------------W--L- ---------G--T-SQME--

                                          Ex4  1          1          X   2
r  LGNMSRQLQSLQTQNRKLTETLQQVAVKLCRELYNKSGG HRCSPCPEKWKWHGNKCYQFYKESKTWQSCEYFCLADNATMLKISTQEEL
m  ------------A-----I-------------------- ------------Y-D----------N-------------------------
h  ---T-QE-----V--I--AGS--H--E---------A-A ------T-Q-----DN------D--S-ED-K----SE-S-----NK--D-

   Ex5                                     Ex6          3           3     X 2
r  DFAMPQSYSEFFYSYWTGLSRNGSGKAWLWTDGTPYSFEL FEIIIDPTNLRSRDCMTIFNGKAYSKDCKELRRCACERVAGRVVPEELQ
m  ---------------------------------------- -Q-----N---N--------------------------------G---
h  E--AS--------------L-PD-------M-----TS-- -H----V-SP-----VA-L--MI--------K--V---R--M-K--S-HVPPETLGEGD
```

Dectin1

```
           Ex1 ITAM Ex2                                                Ex3
r          MDEDGY TQLDFGTRNIHKRPVKSEKGSPAPSSRWRSIAVALGILCLLTVVVAAVLGALA FRRFNSGRYPEEKDNF
m  MKYHSHIENL----- -----S-QD-----RG-----R----P--P---G-----FVV----------- -W-H----N-------
h  MEYHPDLENL----- ---H-DSQSNTRIA-V-----C-A-PP--L---I------VIL---V---TM- IW-S---SNTL-NGY-

                                     Ex4  1         1              2
r  PSRNKENH.KPTEPSLDEKVAPSKASQTTG VFSGPCLPNWIMHAKSCYLFSFSENSWYGSRRHCSQLGAHLLKIDNAKEF
m  L-------.----S----------------- G--QS--------G---------G------K----------------S---
h  L-------SQ--QS--EDS-T-T--VK--- -L-S--P----IYE-------M-L---D--K-Q-W---SN-----SSN-L

   Ex5                                    Ex6             3           3        2
r  EFIESQTSSHRVNSFWIGLSRNQSEGPWFWEDGSAFTPNS FQVRNTAPQESLPHNCVWIHGSEVYNQMCIASSFTICEKEL
m  -------------A-----------------------F--- ------V-----L--------------I-NT---S------
h  G--VK-V--QPD---------P-T-V--L-----T-SS-L ----T--T--NPSP------V-V--D-L-SVP--S----KFSM
```

Fig. 2

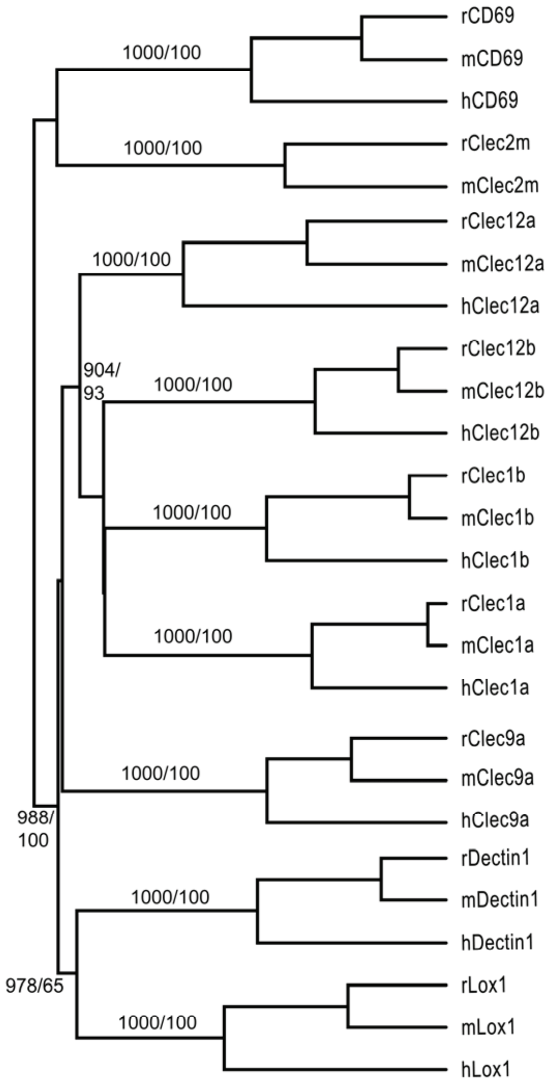

Fig. 3

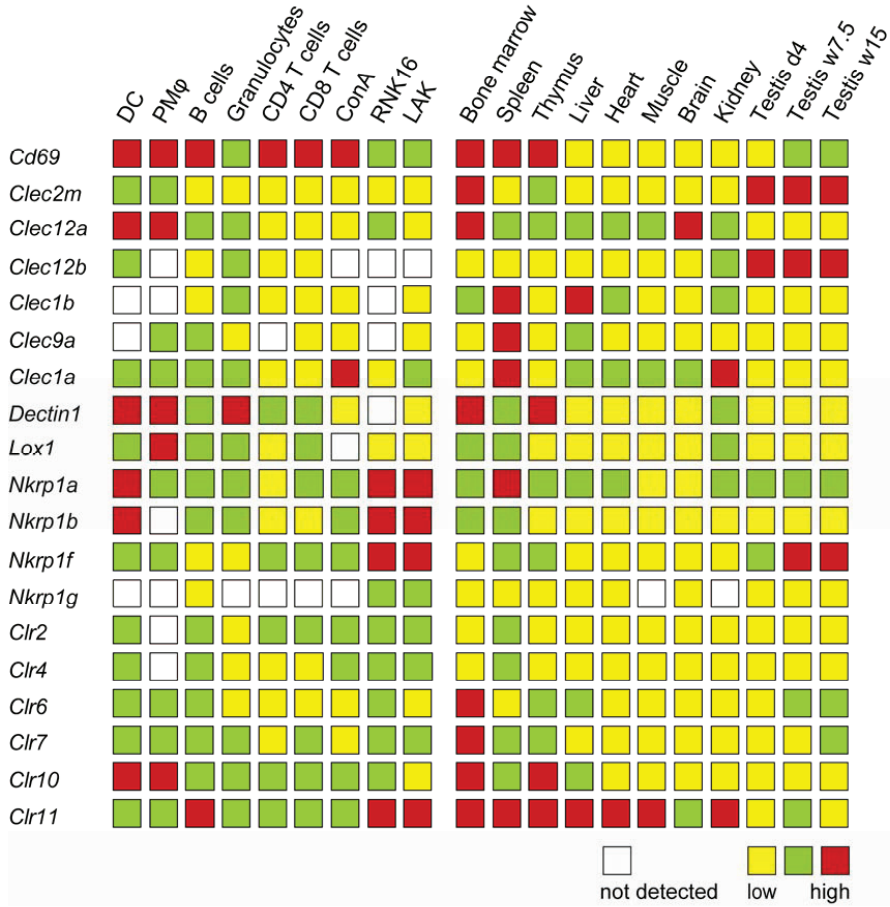

Fig. 4

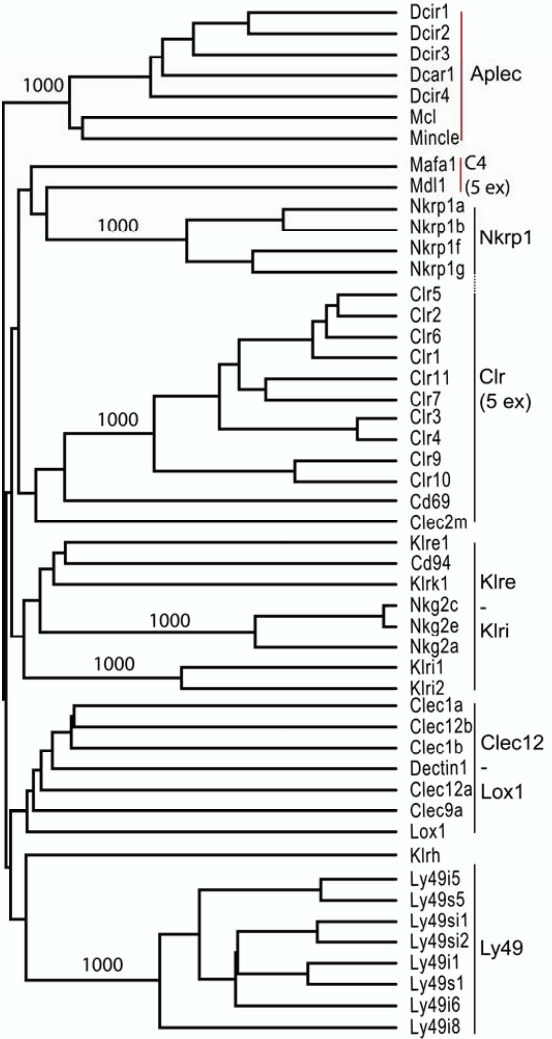

Fig. 5