Päivi Jokinen

Department of Veterinary Biosciences,
Department of Medical Genetics,
Research program for Molecular Medicine and
Folkhälsan Research Center
University of Helsinki
Helsinki, Finland

Identifying Genetic Risk Factors in Canine Autoimmune Disorders

Päivi Jokinen

ACADEMIC DISSERTATION

To be publicly discussed, with the permission of the Faculty of Veterinary Medicine of the University of Helsinki, for public examination in Auditoria Arenan, Folkhälsan building, Topeliuksenkatu 20, Helsinki on 28 January 2011, at 12 noon.

Helsinki 2011

Cover photo: Janne Penttinen

ISBN 978-952-92-8475-7 (pbk.)
ISBN 978-952-10-6788-4 (PDF)
http://ethesis.helsinki.fi/

Unigrafia Oy
Helsinki 2011

Päivi Jokinen

Supervisor	Professor Hannes Lohi, PhD Department of Veterinary Biosciences, Department of Medical Genetics, Research program for Molecular Medicine and Folkhälsan Research Center University of Helsinki Helsinki, Finland
Reviewers	Professor Johanna Schleutker, PhD Laboratory of Cancer Genetics Institute of Medical Technology University of Tampere and Docent Hanna Jarva, MD, PhD Department of Bacteriology and Immunology Haartman Institute University of Helsinki
Opponent	Dr. Catherine André, PhD Canine Genetics and Genomics Institute of Genetics and Development University of Rennes

To my loved ones

Abstract

Autoimmune diseases are more common in dogs than in humans and are already threatening the future of some highly predisposed dog breeds. Susceptibility to autoimmune diseases is controlled by environmental and genetic factors, especially the major histocompatibility complex (MHC) gene region. Dogs show a similar physiology, disease presentation and clinical response as humans, making them an excellent disease model for autoimmune diseases common to both species. The genetic background of canine autoimmune disorders is largely unknown, but recent annotation of the dog genome and subsequent development of new genomic tools offer a unique opportunity to map novel autoimmune genes in various breeds. Many autoimmune disorders show breed-specific enrichment, supporting a strong genetic background. Furthermore, the presence of hundreds of breeds as genetic isolates facilitates gene mapping in complex autoimmune disorders. Identification of novel predisposing genes establishes breeds as models and may reveal novel candidate genes for the corresponding human disorders. Genetic studies will eventually shed light on common biological functions and interactions between genes and the environment.

This study aimed to identify genetic risk factors in various autoimmune disorders, including systemic lupus erythematosus (SLE)-related diseases, comprising immune-mediated rheumatic disease (IMRD) and steroid-responsive meningitis arteritis (SMRA) as well as Addison's disease (AD) in Nova Scotia Duck Tolling Retrievers (NSDTRs) and chronic superficial keratitis (CSK) in German Shepherd dogs (GSDs). We used two different approaches to identify genetic risk factors. Firstly, a candidate gene approach was applied to test the potential association of MHC class II, also known as a dog leukocyte antigen (DLA) in canine species. Secondly, a genome-wide association study (GWAS) was performed to identify novel risk loci for SLE-related disease and AD in NSDTRs.

We identified DLA risk haplotypes for an IMRD subphenotype of SLE-related disease, AD and CSK, but not in SMRA, and show that the MHC class II gene region is a major genetic risk factor in canine autoimmune diseases. An elevated risk was found for IMRD in dogs that carried the DLA-DRB1*00601/DQA1*005011/DQB1*02001 haplotype (OR = 2.0, 99% CI = 1.03-3.95, p = 0.01) and for ANA-positive IMRD dogs (OR = 2.3, 99% CI = 1.07-5.04, p-value 0.007). We also found that DLA-DRB1*01502/DQA*00601/DQB1*02301 haplotype was significantly associated with AD in NSDTRs (OR = 2.1, CI = 1.0-4.4, *P* = 0.044) and the DLA-DRB1*01501/DQA1*00601/DQB1*00301 haplotype with the CSK in GSDs (OR=2.67, CI=1.17-6.44, p= 0.02). In addition, we found that homozygosity for the risk haplotype increases the risk for each disease phenotype and that an overall homozygosity for the DLA region predisposes to CSK and AD. Our results have enabled the development of genetic tests to improve breeding practices by avoiding the production of puppies homozygous for risk haplotypes.

We also performed the first successful GWAS for a complex disease in dogs. With less than 100 cases and 100 controls, we identified five risk loci for SLE-related disease and AD and found strong candidate genes involved in a novel T-cell activation pathway. We show that an inbred dog population has fewer risk factors, but each of them has a stronger

genetic risk. Ongoing studies aim to identify the causative mutations and bring new knowledge to help diagnostics, treatment and understanding of the aetiology of SLE-related diseases.

Contents

List of original publications

This thesis is based on the following publications:

I MHC class II risk haplotype associated with Canine Chronic Superficial Keratitis in German Shepherd Dogs. **Jokinen P,** Rusanen E, Kennedy LJ and Lohi H. Veterinary Immunology and Immunopathology. In press.

II Association of a dog leukocyte antigen class II haplotype with hypoadrenocorticism in Nova Scotia Duck Tolling Retrievers. Hughes AM*, **Jokinen P***, Bannasch DL, Lohi H, Oberbauer AM. Tissue Antigens. 2010 Jun;75(6):684-90.

III MHC class II polymorphism is associated with a canine SLE-related disease complex. Wilbe M, **Jokinen P**, Hermanrud C, Kennedy LJ, Strandberg E, Hansson-Hamlin H, Lohi H, Andersson G. Immunogenetics. 2009 Aug;61(8):557-64.

IV Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex.Wilbe M, **Jokinen P***, Truvé K*, Seppala EH, Karlsson EK, Biagi T, Hughes A, Bannasch D, Andersson G, Hansson-Hamlin H, Lohi H#, Lindblad-Toh K#. Nature Genetics. 2010 Mar;42(3):250-4.

*These authors contributed equally to the study.

#co-directed and corresponding authors

The publications are referred to in the text by their roman numerals.

Abbreviations

ab	antibody
ACTH	adenocorticotropic hormone
AD	Addison's disease
AF	anal furunculosis
AI	autoimmune
AIRE	autoimmune regulator
ANA	antinuclear antibody
APC	antigen -presenting cell
APECED	autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy
APS	autoimmune polyendocrine syndrome
BCR	B-cell receptor
CDV	canine distemper virus
CFA	canine chromosome
CFS	cerebrospinal fluid
CI	confidence interval
CIDD	Canine Inheritance Disorders Database
CLT	canine lymphocytic thyroiditis
CMH	Cochran-Mantel-Haenszel
CNV	copy number variant
CRA	canine rheumatoid arthritis
CS	Cocker Spaniel
CSK	chronic superficial keratitis
D	aspartic acid
DLA	dog leukocyte antigen
DLE	discoid lupus erythematosus
e.g.	exempli gratia
EPI	exocrine pancreatic insuffiency
GSD	German Shepherd Dog
GWAM	genome-wide association mapping
GWAS	genome-wide association study
H	heavy (chain)
HLA	human leukocyte antigen
HVR	hypervariable (region)
IBS	identity by state
IDID	Inherited Diseases in Dogs Database
IFN-γ	interferon-γ
Ig	immunoglobulin
IL	interleukin
IMHA	immune-mediated haemolytic anaemia
IMRD	immune-mediated rheumatic disease
IMTP	immune-mediated thrombocytopenia
LD	linkage disequilibrium
LE	lupus erythematosus

MDS	multidimensional scaling
MG	myasthenia gravis
MHC	major histocompatibily complex
MIT	Massachusetts Institute of Technology
mtDNA	mitochondrial DNA
NCBI	National Center for Biotechnology Information (database)
NF-AT	nuclear factor of activated T-cells
NF-ATc2	calcineurin-dependent transcription factor
NIH	National Institute of Health
NK	natural killer cell
NME	necrotizing meningoencephalitis
NOD	non-obese diabetic (mouse)
NSDTR	Nova Scotian Duck Tolling Retriever
OMIA	Online Mendelian Inheritance in Animal (database)
OR	odds ratio
PAMPS	pathogen-associated molecular patterns
PPR	pathogen recognition receptor
PTPN22	protein-tyrosine phosphatase, non-receptor-type 22
Q	glutamine
R	arginine
RA	rheumatoid arthritis
SCLE	subacute cutaneous lupus erythematosus
SLE	systemic lupus erythematosus
SLU	Swedish University of Agricultural Sciences
SNP	single-nucleotide polymorphism
snRNP	small nuclear ribonucleoprotein complex
SNRPE	small nuclear ribonucleoprotein polypeptide E
SRMA	steroid-responsive meningitis arteritis
T1D	type 1 diabetes
T_C	cytotoxic T-cell
TCR	T-cell receptor
T_H	T-helper cell
TNF-β	tumour necrosis factor β
T_{reg}	regulatory T-cell
UCSC	University of California, Santa Cruz (database)
V	variable (region)
VHK-like	Vogt-Koyanagi-Harada-like syndrome
VRK1	vaccinia-related kinase 1

1 Introduction

Autoimmune diseases occur when an adaptive immune response develops against self-antigens, causing inflammation that may lead to tissue damage. Expression of autoimmunity can be organ-specific, as in type 1 diabetes mellitus affecting pancreatic islets, or systemic, as in systemic lupus erythematosus (SLE), which affects multiple tissues [1]. Susceptibility to autoimmune diseases is controlled by environmental and genetic factors, especially major histocompatibility complex (MHC) class II alleles [2]. More than 60% of canine inherited diseases are shared with humans and the coding sequences of dogs and humans show ~90% similarity. Dogs are large animals, share a living environment with humans and show similar physiology, disease presentation and clinical response, making them an excellent disease model for disorders common to both species [3-6].

Tight bottle necks in the population history of a domestic dog, such as breed creation, World War II, infection outbreaks and modern breeding practices relying on popular sires and tight inbreeding, have accumulated different genetic risk factors and diseases in dog breeds. Dog breeds consist of genetically similar individuals and resemble isolated human populations, such as Finns and Icelanders, that are widely used in genetic studies [7,8]. Observed as a group, dogs show the same extensive genetic diversity as humans, or ancient wolves. At the genome level, this can be seen as long haplotype blocks within a breed and short across breed. Extensive linkage disequilibrium within a breed enables the use of genetic markers, such as single-nucleotide polymorphisms (SNPs) in genome-wide association studies (GWAS) with a small number of samples and markers. Ancient mutations may have segregated into related breeds showing the same disease phenotype, and as the haplotype blocks are short between the breeds, related breeds can be used to narrow down (fine-mapping) and verify the associated loci between a marker and a phenotype [9,10]. This two-stage strategy has been successfully used to identify several Mendelian traits such as white coat colour [11], the hair ridge that causes predisposition to dermoid sinus [12], recessive cone-rod dystrophy [13] and ectodermal dysplasia [14].

Several breeds are highly susceptible to autoimmune diseases. Nova Scotia Duck Tolling Retrievers (NSDTRs) have been recognized to have a strong genetic predisposition to several autoimmune diseases, including immune-mediated rheumatic disease (IMRD) [15], steroid-responsive meningitis arteritis (SRMA) [16,17], hypoadrenocorticism (Addison's disease, AD) [18] and canine lymphocytic thyroiditis (CLT) [19]. IMRD and SRMA may be a part of the same disorder, canine systemic lupus erythematosus (SLE)-related disease complex. German Shepherd dogs (GSDs) are over-represented with chronic superficial keratitis (CSK) [20] and reported to show also congenital focal alopecia areata [21], SLE [22], exocrine pancreatic insufficiency (EPI) [23], anal furunculosis (AF) [24,25] and myasthenia gravis (MG) [26].

Regardless of the high prevalence of autoimmune disorders in dogs, the genetic background remains largely unknown. Previous studies have associated MHC II genes in canine diabetes [27], hypothyroiditis [28,29], AF [25], canine primary immune-mediated haemolytic anaemia (IMHA) [30] and canine rheumatoid arthritis (CRA) [31]. In this study, we focused on the characterization of the genetic risk factors in particular autoimmune diseases in two breeds of dog including SLE-related diseases, comprising IMRD and

SRMA, AD and CSK. No previous genetic studies have been reported in any of these diseases, although an autoimmune origin has been suspected in each disorder. Utilizing novel genomic tools and candidate and genome-wide approaches, we mapped several new genetic risk loci. This study establishes novel canine models for human autoimmune disorders, reveals novel candidate genes and pathways and provides new genetic tests for breeders.

2 Review of the literature

2.1 Autoimmune disorders

2.1.1 Overview of innate and adaptive immunology

The role of the immune system is to protect the host from invading pathogens. The innate immune system is present at birth and lacks of memory and strict recognition of antigen. In its simplest form, the innate immune system comprises anatomical and physiological barriers, such as skin, mucous membranes, and temperature, pH and oxygen levels. Soluble components of the innate immune system include digestive enzymes, such as lysozyme in tears, peptides that bind essential nutrients, such as iron-binding lactoferrin in a mammary gland, and the complement system, an enzymatic protein cascade that produces various chemoattractants, inflammatory mediators, opsonins and a hole-punching complex capable of disrupting membrane structures. Cytokines and chemokines secreted by many cells, including the epithelial cells, mediate intercellular communication and initiate a variety of signalling pathways [1,32].

The cellular components of the innate immune system include cytotoxic cells, neutrophils, eosinophils, basophils and mast cells. Neutrophils and eosinophils are also phagocytic cells. Macrophages are phagocytic cells in the tissues and dendritic cells in tissues and lymphatic organs. Macrophages and dendritic cells serve as antigen-presenting cells (APCs). Antigenic peptides are presented in association with major histocompatibility complex (MHC) class I or II molecules to cytotoxic or helper T-lymphocytes, respectively. Natural killer (NK) cells are lymphocytes without specific antigen recognition capability and induce apoptosis in altered cells. NK cells have recently been shown to have a memory, which suggests that they may be an evolutionary bridge between the innate and adaptive immune systems [33]. Although phagocytic cells lack specific recognition of antigen, they do identify certain pathogen-associated molecular patterns (PAMPS) not found in higher organisms through pathogen recognition receptors (PRRs) [1,32].

The adaptive immune system develops after birth and possesses a memory, enabling a heightened immune response to previously encountered antigens. Lymphocytes specifically recognize the foreign antigen and are divided into different types based on the mechanism of antigen recognition and effector functions. Specificity is achieved during lymphocyte development through a gene rearrangement, a somatic DNA recombination of gene segments encoding the variable (V) region of the antigen receptor. B-cells develop in bone marrow and produce a great variety of antigen receptors called immunoglobulins (Igs). All Igs are identical in a single cell and recognize a specific antigen. Igs expressed on a B-cell surface are called B-cell receptors (BCRs) and Igs with the same antigen specificity that are secreted by plasma cells as soluble form are called antibodies (abs). Isotype or class of the ab is determined by the heavy (H) chain and in part directs the function following the activation [1,32].

T-lymphocytes also develop in bone marrow, but mature in the thymus. The antigen receptors on T-cells are always membrane-bound and called T-cell receptors (TCRs). The structure and generation of antigen specificity are identical to that of B-cells and their function is to signal activation. The major difference is the recognition of antigen. TCRs can only bind antigens associated with the self MHC molecules. T-lymphocytes are divided into two subtypes based on function and cell surface markers. Cytotoxic T-cells (T_C) express CD8 glycoprotein and induce apoptosis in altered cells. T-helper lymphocytes (T_H) express CD4 cell marker and modulate immune response primarily through cytokine secretion. T_H cells can further be divided into T_H1, T_H2, T_H17 and regulatory T-cells (T_{reg}). T_H1 cells secrete interferon-γ (IFN-γ) and tumour necrosis factor β (TNF-β). They promote elimination of intracellular pathogens, and cell-mediated and delayed-type hypersensitivity responses. T_H2 cells secrete interleukins (IL) IL-4 and IL-5, which contribute to allergic responses and the clearance of extracellular pathogens, such as worms, and promote humoral response. T_H17 cells secrete IL-17 and IL-22, which are important cytokines in fighting extracellular bacteria and fungi. T_{reg} cells express CD25 as well as CD4 cell marker and are suppressive mediators of immune responses as well as important in maintaining peripheral tolerance [34].

The lymphocyte-antigen encounter takes place in secondary lymphoid tissue and leads to activation through changes in gene expression, proliferation and differentiation into effector cells. After encountering an antigen, B-cells differentiate into plasma and memory cells. Plasma cells secrete antibodies, which neutralize extracellular pathogens by coating, agglutinating and opsonizing them. Perhaps most importantly, they activate the complement cascade. T_C cells induce apoptosis in target cells by releasing the content of cytoplasmic granules and/or by expressing a transmembrane protein Fas-ligand. T_H cells direct the immune response towards humoral or cell-mediated response by secreting cytokines. T_{reg} cells secrete cytokines that modulate the function of dendritic cells and lymphocytes and even induce apoptosis in the latter [34,35].

Many vital organs regarding survival and reproduction, which possess limited capacity for regeneration, are considered to be immune privilege body sites. These organs include the brain, cornea, testes and the pregnant uterus. However, recent evidence suggests that immune privilege is not a global suppression of all immune responses but in fact an active and closely regulated adaptation of the immune system with the objective of protecting organs from immune-mediated damage. Most harmful immune responses are down-regulated, while others, less harmful, are preserved. Anterior chamber-associated immune deviation (ACAID) is an example of this kind of regional immunity [36]. After encountering an antigen, the APCs travel directly to the spleen, where they interact with other cells of the immune system, resulting in activation of T_H1-suppressing T_{reg} cells [37]. Brain-associated immune deviation (BRAID) resembles ACAID, but has not been as thoroughly characterized [36].

2.1.2 Overview of autoimmunity

The ability to differentiate self from foreign is an essential basis in avoiding immune-mediated damage to self-tissue. A central tolerance is introduced during foetal development in bone marrow and thymus by negative selection, resulting in apoptosis of strongly self-reactive lymphocytes. Autoimmune diseases occur when the self-tolerance is lost and an adaptive immune response develops against self-antigens [1,38]. Autoimmune regulator (AIRE) is a transcription factor that promotes expression of tissue-specific antigens in thymic medullary cells, enabling the formation of self-peptide-MHC complexes. The CD4+CD8+ (double-positive, DP) thymocytes, derived from bone marrow haematopoietic precursors, interact with these cortical epithelial cells, enabling the negative selection of too strongly binding T-cells. Thymocytes that interact with appropriate affinity with peptide-MHC class I complexes become CD8+ T-cells, while those that interact with peptide- MHC class II complexes become CD4+ T-cells [39]. As important are mechanisms maintaining peripheral tolerance, which eliminate or inactivate the potentially autoreactive T-cells that have escaped negative selection. These include the loss of suppression of Tregs [40].

Organ-specific autoimmune pathogenesis has primarily been associated with T_H1, but not T_H2 cells. In some systemic autoimmune diseases, like SLE, T_H2 cells have been shown to have an influence, but they are not considered the driving force. Recently identified T_H17 cells have been demonstrated to have a major role in autoimmunity and T_{reg} cells in preventing immune-mediated damage. The balance and interplay of all of these T-cell subtypes with each other are critical for developing autoimmune diseases [40].

A constant concentration of autoantigens and the lack of their eradication them makes autoimmune diseases chronic. Chronic inflammation gives positive feedback by attracting macrophages and neutrophils by secreted cytokines and chemokines and by revealing new autoantigens from damaged tissues, a phenomenon called epitope spreading. Epitope spreading may explain the relapses common to many autoimmune diseases [1,38].

The autoimmune diseases may be organ-specific, affecting limited tissues, or systemic, with autoimmunity being expressed in several tissues. In systemic autoimmune diseases, such as in SLE, non-organ specific autoantibodies attack ubiquitous self-molecules. In SLE, the main target is chromatin. In organ-specific autoimmune diseases, the target antigens are found in one or a few organs and the tissue destruction is limited to these organs, although there may be symptoms affecting the whole body, such as fever [1]. Canine organ-specific autoimmune disorders include several diseases affecting the haematologic system, such as IMHA, immune-mediated thrombocytopenia (IMTP) and immune-mediated neutropenia [38]. Also several autoimmune diseases of the endocrine system have been characterized, such as autoimmune thyroiditis, autoimmune diabetes mellitus and AD. Autoimmune diseases affecting the skin include discoid lupus and bullous skin diseases and those affecting the the musculoskeletal system, MG and CRA. Ocular autoimmune diseases are e.g. canine uveodermatologic syndrome or Vogt-Koyanagi-Harada –like syndrome (VKH-like) and CSK [38,41].

2.1.3 Genetic background of autoimmune diseases

A few autoimmune syndromes exist where a single gene is a causative risk factor, such as autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) [42,43]. In APECED, the transcription factor gene, AIRE, is defective, causing the destruction of multiple endocrine tissues. Still, most of the autoimmune disorders are thought to be polygenic, and several genes and pathways have already been identified in humans. These susceptibility genes are often involved in autoantigen availability and clearance, apoptosis, signalling, cytokine gene expression and expression of co-stimulatory molecules. Genetics studies in canine autoimmune diseases have to date revealed several associations with the MHC class II locus. As yet, few other genetic risk factors outside the MHC class II region have been identified in canine autoimmune diseases [44]. This is not the case in human autoimmune diseases, where several genes have been identified. Some of these are presented in Table 1. New array and sequencing technology is likely to reveal new genes and pathways also behind canine autoimmune diseases in the near future.

Copy number variations (CNVs) are structural variations in a genome from one kilobase to several megabases in length. CNVs are rarer than SNPs, but are often located in gene areas, causing more likely changes in gene expression levels, disruption of gene dosage, unmasking of recessive alleles or regulatory polymorphism and loss of regulatory elements [45-48]. Several CNVs are known to be associated to common diseases in humans, including cancer, neuropsychiatric diseases, infectious diseases and autoimmune diseases, SLE being one of them [46,49]. DNA structural variation has been mapped in dogs, and it is likely that CNV variation contributes to the genetic basis of complex diseases in dogs as well [45].

Epigenetic modifications describe inherited changes in the expression of DNA that result from reasons other than what is coded in a DNA sequence. These include DNA methylation, chromatin remodelling, such as post-translational modifications of the histone proteins and RNA interference [50]. Several acetylated proteins, have been associated with rheumatoid arthritis alone and methylation is particularly associated with autoimmune diseases [51].

2.1.3.1 Major histocompatibility complex (MHC)

MHC is a multigene family found in all vertebrates studied to date. The human MHC region, also known as human leucocyte antigen (HLA) region, is located on chromosome 6p21 and extends over 3.6 Mb. MHC region is divided into three subregions, MHC classes I, II and III. Canine MHC or DLA is a 3.9 Mb gene cluster mainly located on chromosome 12. The MHC in carnivore species was split perhaps over 55 million years ago into two pieces within the TRIM (member of the tripartite motif) gene family found in HLA. DLA class II, III, and I regions were situated in a pericentromeric region of chromosome 12, whereas the remaining region was located in a subtelomeric region of chromosome 35. In addition, two class I genes are found on chromosomes 7 and 18 [52] (Figure 1). Comparing mammalian species, it can be seen that chromosome breaks, inversion and/or centromere invasion have occurred in the MHC region during the evolution of each species [52,53]

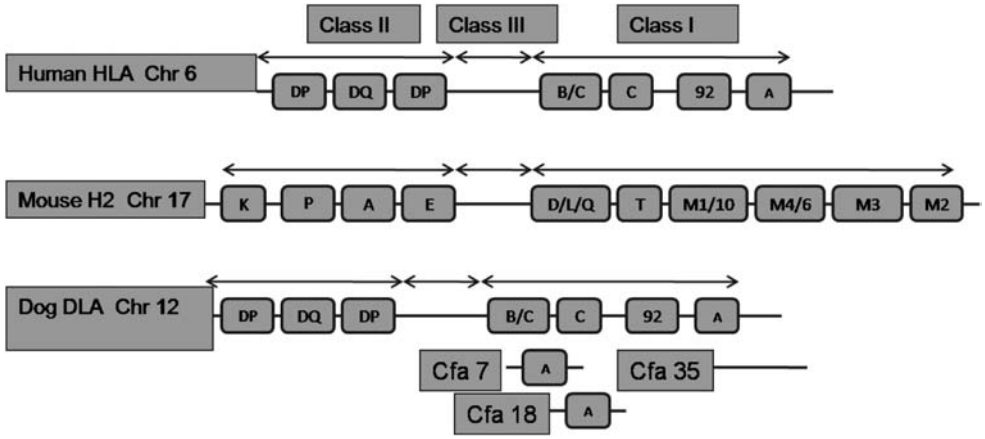

Figure 1 *Genomic structure of human, mouse and dog major histocompatibility complexes. Picture modified from* [52]

The MHC region encodes several genes involved in both the innate and adaptive immune system. The primary function of the MHC is to recognize, bind and transport antigens to the surfaces of APCs, where they are presented to T-cells [1,54]. MHC class III encodes the complement pathway genes, cytokines TNF-α and –β and heat shock proteins. MHC class I and II encode genes that recognize, bind and present antigen peptides to cytotoxic CD8+ and helper CD4+ T-cells, respectively. This interaction between the APC and T-cell initiates the cellular and humoral immune response. MHC class I is expressed on all nucleated cells and binds endogenously produced peptides. Class II is expressed on APCs, such as macrophages, B-lymphocytes and dendritic cells, and presents exogenous material that was endo- or phagocytosed. Under inflammation, also fibroblasts and vascular endothelial cells may express MCH class II molecules. MHC class I and II are polygenic and polymorphic, and alleles are co-dominantly expressed, yielding a high molecular diversity [32,52,55].

The DLA class II region includes four loci, DLA-DRB1, -DRA, -DQA1 and -DQB1, with one functional gene at each locus. All DLA class II genes, except DRA, are highly polymorphic. The polymorphism in the DLA region is genetically maintained by point mutations, genetic recombination and gene conversion. Research for new variants is ongoing, and to date 148 DLA-DRB1, 70 DLA-DQA1 and 26 DLA-DQB1 alleles have been identified (Dr. LJ Kennedy, personal communication). Many of the alleles are breed-specific and most breeds have a very limited diversity of alleles. In comparison, the human HLA-DRB1 gene has over 600 alleles. MHC class II molecules are composed of two heterodimeric transmembrane glycoprotein chains α and β, each consisting of two domains. The α domains are encoded by DLA-DQA1 and DLA-DRA1 genes, and the β domains by DLA-DQB1 and DLA-DRB1 genes. The α_1 and β_1 subunits form the peptide binding groove (Figure 2).

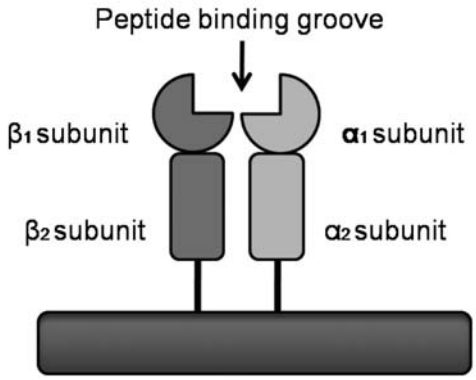

Figure 2 *MHC class II protein. The α_1 and β_1 subunits form the peptide binding groove.*

SNP differences in second exons of the genes DLA-DRB1, DLA-DQA1 and DLA-DQB1 create changes mostly in the hypervariable (HVR) regions of the peptide binding cleft, therefore altering the specificity of peptide recognition, binding and T-cell presentation. The DRB1 alleles are usually seen only with one combination with DQA1- and DQB1- alleles as the DQ alleles may be seen in combination with several different DRB1 alleles [56]. These allele combinations or haplotypes may act epistaticly and provide some biological advantage as has been shown in human studies [57].

There are several suggested mechanisms to maintain the polymorphisms in MHC region. These can be divided into two main models, the disease-based and reproductive mechanism. The disease-based model operates through balancing selection between host and pathogen and is based on their co-evolution. The *heterozygote advantage hypothesi*s suggests that heterozygosity is favoured, as heterozygotes are able to present antigens more broadly. This hypothesis is also known as the *overdominance/dominance hypothesis*; the hetorozygote in the overdominance hypothesis would be fitter than the fittest homozygote and in the dominant theory, the heterozygote would be fitter that the homozygotes on average, but no more than the fittest homozygote. The *negative frequency-dependent selection hypothesis*, also known as the *rare-allele advantage hypothesis*, proposes that parasites evolve to exploit the defects in the most common host genotype, and the host therefore benefits from the rare alleles. *Fluctuationg selection* proposes that the spatial and temporal diversity and the amount of pathogens are the driving force, rather than co-evolution.

The reproductive model is based on sexual selection and also has two different hypotheses. The first suggests that disease-based fitness differences between MHC genotypes favour reproductive mechanisms that would produce offspring with high fitness genotypes. This MHC-dependent mating might enhance parasite resistance in two ways, either by providing advantageous heterozygotes or by racing with the evolution of mutating parasites. The latter is known as the *moving target hypothesis*. The second reproductive model is based on *inbreeding avoidance hypothesis,* and the aim is to avoid

the negative consequences of inbreeding, such as accumulation of recessive deleterious mutations [58].

MHC has been associated with almost every autoimmune disease, although the causal variants are not being characterized in most cases due to extensive linkage disequilibrium (LD) in the region [54]. Some of the predisposing HLA alleles and haplotypes have been listed along with other susceptibility genes in Table 1.

2.1.4. Environmental background of autoimmune diseases

Autoimmune diseases have a strong genetic influence, but often environmental factors are needed to trigger the disease in genetically predisposed individuals. There are several mechanisms by which pathogens trigger the autoimmune diseases. Firstly, autoreactive T-cells can be activated via molecular mimicry by cross-reactive recognition of an infectious antigen that has similarity to self-antigen. Secondly, infection causes tissue damage, revealing self-antigens that are normally not exposed. This together with the secreted inflammatory mediators may activate bystander lymphocytes not specific to the pathogen. Self-antigens can then be taken up by activated APCs, processed and presented to autoreactive T-cells in a process known as bystander activation. Tissue destruction may also cause epitope spreading. Thirdly, microbial superantigens may activate a large subset of T-cells, some of which are specific to self-antigens. In most cases, the autoimmune reaction ends as the pathogen is eradicated, but may sustain in genetically predisposed individuals. Drugs and toxins may react chemically with self-proteins and form compounds foreign to the immune system. These haptenated proteins may activate the immune response, leading to autoimmune reactions [1]. It has also been suggested that exposure to environmental toxins during early development causes inherited epigenetic modifications [59].

The hygiene hypothesis proposes that the decreasing incidence of infections in developed countries is the cause of the increasing incidence of both autoimmune and allergic diseases. This cannot be explained only by different genetic background. For example, the incidence of type 1 diabetes (T1D) with the same genetic background is close to six-fold higher in Finland than in the adjacent Karelian Republic of Russia [60]. On the other hand, environmental risk factors alone do not explain this difference, as evidenced by the high concordance of T1D in monozygotic twins. The best support for the theory is provided by different animal models, such as the non-obese diabetic (NOD) mouse [61]. NOD mice bred in 'conventional' facilities show little or no diabetes, whereas close to 100% of the female NOD mice bred in specific pathogen-free conditions develop the disease. In addition, a protective effect of probiotics and bacterial extracts was reported at the onset of diabetes. The proposed underlying mechanism is a T_H1–T_H2 deviation, caused by antigenic lymphocyte competition for cytokines, recognition for MHC-self-peptide and growth factors necessary for the activation of B- and T-cells. Also T_{reg} cells or antigen-independent stimulation through TLRs may be involved. Two immune responses caused by different antigens are known to inhibit each other, and therefore, a strong immune response to a pathogen might inhibit a weak immune response to an autoantigen [62].

2.1.5. Shared autoimmune disorders in humans and dogs

The Online Mendelian Inheritance in Animal (OMIA) database lists a total of 506 inherited diseases in dogs (20.09.2010), from which at least 235 are considered as potential disease models for human diseases. Table 1 lists the autoimmune diseases thought to be shared with humans and identified genes according to OMIA [63]. Canine and human clinical diagnostics vary and dog diseases are usually not divided into as many sub-phenotypes as diseases in humans. Therefore, in the Table 1, human diseases are referred to using more general nomenclature.

Table 1 *Common autoimmune diseases and identified genes in human and dog* [63] *24.11.2010.*

	Human disease	Genes	Dog disease	Genes
1	Alopecia areata	*HLA-DQ3, 18p11.3-p11.2*	Alopecia	
2	Diabetes mellitus - all forms	*HNF4A, GCK, GPD2, NEUROD1, IRS1, PPARG, FOXP3, IGF2BP2, WFS1, NIDDM4, AL1, ENPP1, IL6, GCK, PAX4, SLC30A8, TCF7L2, ABCC8, KCNJ11, MAPK8IP1, HNF1A, IPF1, IRS2, LIPC, SLC2A4, HNF1B, GCGR, RETN, AKT2, HNF4A, NIDDM3, PTPN1, PTPN22, ITPR3, IDDM1, IL6, HNF1A, OAS1, HLA-DRB1*04-DQB1*0302 and HLA-DRB1*03, several mitochondrial genes, including MTTL1, MTTE and MTTK,*	Diabetes mellitus	*DLA, IL-4, IL-12b, PTPN22, IL-10, IL-6, TNFá, IFNã, insulin variable number tandem repeat*
3	Epidermolysis bullosa	*KRT5, KRT14, ITGB4, COL7A1, PLEC1, ITGA6, ITGB4, Xq27.3-qter*	Epidermolysis bullosa	*COL7A1, LAMA3*
4	Haemolytic anaemia, autoimmune		Haemolytic anaemia, autoimmune	
5	Hypoadrenocorticism	*AIRE, PTPN22, HLA-A1, -B8 and DR3*	Hypoadrenocorticism	
6	Hypothyroidism	*Nkx2.5, FKHL15, UBR1, TRH, HLA-DR3*	Hypothyroidism	
7	Myasthenia gravis	*HLA, CHAT, 10q11.2, 17p13*	Myasthenia gravis	
8	Narcolepsy	*HLA, HCRT, NLC1A, NRCLP2, 4p13-q21, NRCLP3, 21q11.2*	Narcolepsy	*HCRTR2*
9	Anonychia-onychodystrophy		Onychodystrophy	
10	Pancreatic insuffiency, combined exocrine		Pancreatic insuffiency, exocrine. Sub-type: Pancreatic acinar atrophy	
11	Pemphigus	*ATP2C1, HLA, DESMOGLEIN 3*	Pemphigus	
12	Polyglandular autoimmune syndrome, type II		Polyglandular autoimmune syndrome, type II	
13	Systemic lupus erythematosus	*PTPN22, FCGR2B, FCGR3A, TNFSF6, SLEB1, PDCD1, TREX1, SLEB3, BANK1, C4A, SLEH1, SLEB4, SLEB5, DNASE1, HLA-DR2 and DR3*	Systemic lupus erythematosus	
14	Thrombocytopenia	*SALL4, MASTL, contiguous gene deletion of 11q23, CFHR1, CFHR3, HF1, WAS, ADAMTS13, WAS, WAS, GATA1*	Thrombocytopenia	

Epidemiological studies are scarce compared to human autoimmune diseases, but some dog breeds are clearly overrepresented with immunological disorders. Breed-specific prevalence estimates were reported only in 21 from a total of 312 disorders based on a PubMed search [64]. The few examples available of prevalence estimates in autoimmune diseases are: haemolytic anemia 11-25% in eighteen different breeds, Sebaceous adenitis 24% in Akitas. The Orthopedic Foundation for Animals lists breed statistics based on laboratory testing. These numbers are only indicative since the individuals may not present random sampling of a breed, although some numbers of evaluated animals are high enough to make suggestive conclusions. In Table 2 are listed some examples of the breeds predisposed to hypothyroidism. A recent study in Swedish Giant Schnauzers gave a prevalence of 16% for the canine autoimmune lymphocytic thyroiditis (CLT) as it is here 6.6% [65]. A similar phenotype in human is Hashimoto's thyroiditis which has been listed as a rare disease by Orphanet, the portal for rare diseases and orphan drugs (www.orpha.net).

Table 2 *Top 25 breeds affected with hypothyroidism according to Orthopedic Foundation for Animals database. Breeds with over 50 evaluations are listed. Modified from www.offa.org, 24.22.2010.*

Breed	Rank	Evaluations No.	Autoimmune Thyroiditis %	Idiopathic Hypothyroidism %
ENGLISH SETTER	1	515	27.4	0.4
TIBETAN MASTIFF	2	55	12.7	1.8
KUVASZ	3	248	13.7	0.0
SHETLAND SHEEPDOG	4	690	13.3	0.3
GERMAN WIREHAIRED POINTER	5	223	11.7	0.9
TIBETAN TERRIER	6	51	11.8	0.0
BOXER	7	718	10.9	0.4
WELSH SPRINGER SPANIEL	8	437	9.4	1.4
DALMATIAN	9	347	10.7	0.0
RHODESIAN RIDGEBACK	10	2610	9.8	0.7
NOVA SCOTIA DUCKTOLLING RETRIEVER	11	452	9.1	0.0
IRISH SETTER	12	342	9.1	0.0
BEAGLE	13	110	9.1	0.0
AMERICAN PIT BULL TERRIER	14	79	8.9	0.0
AKITA	15	415	8.4	0.2
AFGHAN HOUND	16	151	6.6	2.0
HAVANESE	17	283	8.1	0.0
LEONBERGER	18	551	7.3	0.7
AMERICAN STAFFORDSHIRE TERRIER	19	157	7.6	0.0
FLAT-COATED RETRIEVER	20	106	5.7	1.9
GIANT SCHNAUZER	21	290	6.6	0.7
CANAAN	22	84	7.1	0.0
AUSTRALIAN TERRIER	23	57	7.0	0.0
GOLDEN RETRIEVER	24	2207	6.6	0.3
BORZOI	25	657	6.2	0.6

2.1.5.1 Systemic lupus erythematosus (SLE)

Lupus erythematosus (LE) in humans is a heterogeneous autoimmune disease with varying immune responses and clinical course. LE can be divided into two subclasses according to characteristic clinical features, Systemic lupus erythematosus (SLE) and

cutaneous lupus erythematosus (CLE). CLE can be subdivided to discoid lupus erythematosus (DLE), acute cutaneous lupus erythematosus and subacute cutaneous lupus erythematosus (SCLE), which usually manifest solely in skin lesions, but may sometimes show extracutaneous signs. SLE is a multisystemic disease with variable symptoms such as skin manifestations, arthritis, serositis, proteinuria and neurological disorders. Autoantibodies are infrequent in DLE, but are practically always present in SLE in multiple specifities and almost always in SCLE. One of these autoantibodies is antinuclear antibody (ANA) [66,67]. The prevalence of SLE varies between different populations from 3/100 000 in Iceland and Japan to 91/100 000 in Spain [68]. The CLE prevalence has not been as widely studied, but is estimated to be two to three-fold more common than SLE [69]. Both genetic and environmental factors are thought to contribute to the aetiology, and selected genes underlying SLE have been listed in Table 1.

2.1.5.2 Hypoadrenocortisism (Addison's disease, AD)

The clinical and pathological features of Addison's disease (AD) were first described by Thomas Addison in 1855. AD is caused by insufficient production of corticosteroids and mineralocorticoids due to autoimmune destruction of the adrenal cortex [70]. Autoantibodies in AD are directed against the enzymes involved in steroid synthesis, and they have predictive use, which is exploited by the ACTH stimulation test, detecting subclinical adrenocortical dysfunction with a high sensitivity. Isolated AD cases are rare and usually accompanied by other endocrinopathies such as autoimmune thyroid disease, pernicious anaemia and diabetes mellitus. AD is a part of the autoimmune polyendocrine syndrome (APS) in 100% of APS II cases and in 72% of APS I (APECED) cases. The prevalence of APECED is increased in Finland and is included in the "Finnish heritage of disease" with a prevalence of 1/25 000 [71] The prevalence of AD in the general population is from 30/1 000 000 to 60/1 000 000 [72]. The aetiology of AD is not fully understood, but, as a part of the APS I, it has been associated with the HLA-DRB1*03 allele, whereas other symptoms of APS I are associated with over 60 different mutations in the AIRE gene and other MHC class II variants [71]. AD in the isolated form and in the context of APS II has been associated with HLA-A1, -B8 and -DR3 [72].

2.1.5.3 Autoimmunity in the eye: an immune-privileged site

Immune-mediated diseases are relatively common in humans and dogs and present mechanistically interesting autoimmune conditions without reactive lymphoid tissues. Ocular disorders may affect the eye globe as a whole or individual structures such as the cornea, conjunctiva and eye lids, sclera and episclera, optic neuron, retina and extraocular muscles [73] (Figure 3). This might seem surprising because of the absence of lymphatic drainage, except for the conjunctiva, and selective blood-ocular barriers, which restrict the access of antigens to potentially reactive lymphoid tissue. The cornea has for long even been considered an immune-privileged organ because of the lack of lympatics and blood vessels [74].

Autoimmune uveitis in humans comprises a group of potentially blinding ocular inflammatory diseases, with an annual incidence of over 150 000 persons in the United States. Anterior uveitis is less destructive to the vision and affects mainly the front of the lens. Posterior uveitis (uveoretinitis) is more likely to result in blindness due to irreversible damage to the neural retina and adjacent structures. Posterior uveitis may involve only the eye, as in sympathetic ophthalmia and birdshot retinochoroidopathy or be a part of a systemic syndrome, such as of Behcet's disease, sarcoidosis and Vogt-Koyanagi Harada disease [75].

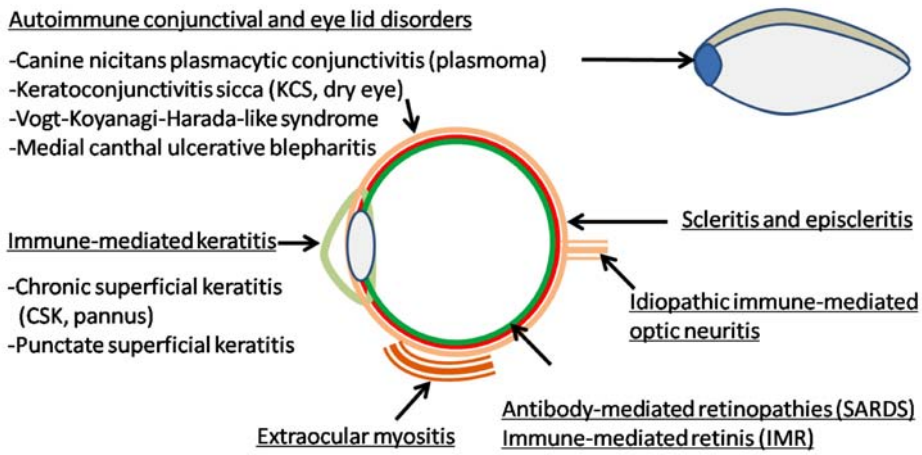

Figure 3 *Ocular immune-mediated diseases in dogs. Many of these diseases are poorly described and the pathogenesis is mostly unknown.*

2.1.6 Autoimmune disorders in Nova Scotia Duck Tolling Retrievers

Nova Scotia Duck Tolling Retrievers (NSDTRs) are highly susceptible to several autoimmune diseases, including IMRD, SRMA and AD. Other autoimmune diseases, such as AIHA and hypothyroidism, have been reported in lower frequency by the Finnish breed club. IRMD and SRMA may be a part of the same immune disease syndrome SLE-related disease [15,17,18]. We focused on IMRD, SRMA and AD, which will be described in more detail below.

2.1.6.1 SLE-related disease

IMRD and SRMA are hypothesized to be a part of the same autoimmune disorder, SLE-related disease syndrome, based on the breed predisposition and segregation of both phenotypes in the same breeding lines [15].

2.1.6.1.1 Immune mediated rheumatic disease (IMRD)

Typical immune-mediated rheumatic disease (IMRD) -affected dogs show similar clinical signs as patients with human SCLE and SLE; 84% of SCLE patients and 96% of SLE patients show antinuclear antibodies, while IMRD-affected dogs show ANA positivity in 70% of cases. Another typical clinical sign is polyarthritis, which is seen in 16% of SCLE and 68% of SLE patients, whereas all IMRD dogs display arthritis. Other common symptoms include skin manifestations, fever and kidney and liver problems [15,76,77]. The median age of disease onset is three years, and the disease frequency is clearly elevated compared with other breeds [15]. The prevalence of SLE has not been studied in dogs, but it is a rare disease in dogs in general. In eleven-year period (1991-2001), 83 dogs had been diagnosed with noninfectious, nonerosive, immune-mediated polyarthritis, from which only seventeen dogs had been confirmed to have SLE, in the veterinary teaching hospital at the Western College of Veterinary Medicine. The total number of canine patients from this period was 23 661 [78]. In a five-year period (2002-2007), 121 SLE affected dogs had been tested ANA positive in Clinical Pathology Laboratory of the University Animal Hospital in Uppsala and 26% of these dogs were NSDTRs, suggesting that this breed is highly susceptible for IMRD.

2.1.6.1.2 Steroid-responsive meningitis arteritis (SRMA)

Dogs have two forms of steroid-responsive meningitis arteritis (SRMA). In the acute form of SRMA, acute neck pain is typical, which manifests in a reluctance to turn the head and lowering it while walking. Other clinical signs are fever, stiff gait, hunched back while walking, depression and anorexia, probably because of difficulties in lowering the head and opening the mouth. Often dogs pant excessively due to severe pain. Cerebrospinal fluid (CFS) shows a significant neutrophilic pleocytosis and an elevated protein concentration. CSF is not a widely used diagnostic method, and therefore, exclusion of the main potential differential diagnoses, such as disc herniation and polyarthritis of the cervical facet joints, as well as breed disposition, are used to confirm the diagnosis. The more protracted form shows severe neurological signs such as ataxia, paresis, tetraparesis or paraplegia, mild to moderate mixed cell pleocytosis in CSF and possible protein elevation. Some dogs develop polyarthritis, which is always seen in IMRD-affected dogs. The dogs develop signs of SRMA at the age of 4-19 months and a lifelong therapy with corticosteroids may be needed to avoid relapses, which occur in 50% of cases. Unresponsiveness to medical treatment may lead to euthanasia. The estimated prevalence of SRMA according to a Norwegian study is around 2.5%, which may be underestimated because of strict inclusion criteria [17].

2.1.6.2 Hypoadrenocortisism (Addison's disease, AD)

Dogs as well as humans affected with AD often present with a variety of non-specific signs, including vomiting, diarrhea, lethargy, anorexia, muscular weakness and depression [79-81]. The age of onset in NSDTRs is around four years, and the dogs as well as humans

are treated by supplementing the missing hormones [80]. The diagnosis is confirmed with an ACTH stimulation test.

2.1.7 Autoimmune disorders in German Shepherd dogs

From all breeds, German Shepherd dogs have been reported to have the most inherited defects, total of 77 [64]. Even if conformation-related defects are excluded, the number of familial disorders is 58. It is therefore not surprising that also several autoimmune diseases are included, such as congenital focal alopecia areata [21], SLE [22], exocrine pancreatic insufficiency (EPI) [23], AF [24,25], CSK [20] and MG [26]. Our focus in this study is on CSK, which will be described here in more detail.

2.1.7.1 Canine chronic superficial keratitis (CSK)

Canine chronic superficial keratitis (CSK) is a progressive autoimmune ocular disease often leading to blindness if left untreated. Characteristic for CSK is progressive, bilateral vascularisation, fibrous tissue formation and pigmentation of the anterior corneal stroma [20] (Figure 4). Although CSK is found in many breeds, it is most prevalent in GSDs [74,82-86] (Table 3).

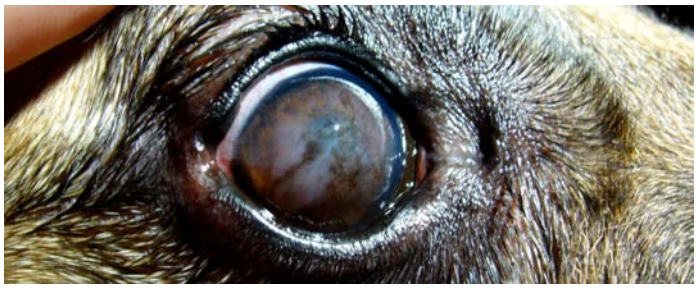

Figure 4 *Progressive, bilateral vascularization, fibrosis and pigmentation of the anterior corneal stroma in the eye of German Shepherd Dog affected with chronic superficial keratitis. Photo: Elina Rusanen.*

Table 3 *Dog breeds reported with chronic superficial keratitis in the literature.*

Breed	Reference	Breed	Reference
Akitas	Slatter 1977	Golden retrievers	Slatter 1977
Australian shepherds	Andrew 2008	Great Danes	Slatter 1977
Belgian tervuerens	Chavkin 1994	Greyhounds	Slatter 1977
Border collies	Slatter 1977	Kelpies	Stanley 1988
Boxers	Überreiter 1959	Labrador retrievers	Slatter 1977
Bull mastiffs	Andrew 2008	Poodles	Slatter 1977
Collies	Slatter 1977	Shetland sheepdogs	Slatter 1977
Dachshunds	Überreiter 1959	Siberian huskies	Slatter 1977
Dalmatians	Nell et al. 2005	Weimareners	Slatter 1977
German shepherd dogs	Überreiter 1959	Vizlas	Andrew 2008

The initial phase of CSK mainly involves IFN-γ -producing CD4+ T-lymphocytes that infiltrate from the temporal region of the limbus into the superficial corneal stroma. The next phase involves invading macrophages, plasma cells and neutrophils [87]. Increased expression of MHC class II proteins has been observed in the central cornea. This aberrant MHC class II expression has been proposed to be associated with the IFN-γ secretion of the invading T-helper cells [41]. The presence of CD4+ T-cells is typical for ocular autoimmune diseases [88]. In addition, CSK is responsive to topical steroids and cyclosporine, further indicating an autoimmune origin [89].

2.2. The dog as a model species for human inherited disorders

2.2.1 Origin of the domestic dog

Dogs were domesticated from wolves less than 16 300 years ago [90] (Figure 5). The latest study based on genomes of mitochondrial DNA (mtDNA) suggest that the existing canine breeds have a common origin, most probably in south-eastern Asia, south of the Yangtze river. It is estimated that at least 51 female wolves with different mtDNA haplotypes and a total of several hundred individuals were domesticated and that the modern domestic dogs originate from these wolves [90]. Domestication has been estimated to have a 5% reduction in nucleotide diversity, whereas breed formation resulted in an average reduction of 35% [91]. Therefore, the domestication event itself has not affected the genetic diversity as extensively as modern breeding practices.

Figure 5 *Illustration of the bottlenecks in the history of the domestic dog.*

2.2.2 Breed creation

Most of the over 400 modern dog breeds have been created over the last 400 years. The founder effect is very strong in pure-bred dogs [4]. Each pure breed represents a group of genetically very similar animals that have descended from only a few ancestors. Some breeds have gone through several severe bottlenecks during the World Wars or depression and infectious disease breakouts, reducing the effective breeding population to only a few dogs. Modern breeding practices have also extensively narrowed the genetic diversity. Tight inbreeding accumulates recessive disease alleles and the frequency of these alleles is further increased by a use of ‘‘popular sires’’. Popular sires are dogs successful in dog shows or in competition events or otherwise considered superior representatives of the breed, and they may produce >100 litters in their lifetime.

It is clear that in-breeding accumulates recessive mendelian diseases, but the effect of ‘in-breeding depression’ on polygenic, late-onset diseases is more complex. Firstly, the combined effect of individual risk factors for complex disease is multiplicative rather than additive. Secondly, late-onset diseases may not be selected against as they are first visible after reproductive age. Thirdly, the negative effects of rare homozygous mutations may be more severe than effects of common risk variants, as they usually do not exist in out-bred populations where homozygotes are rare. Fourthly, in-breeding affects the response to environmental factors such as infections. Fifthly, if the theory of heterozygote advantage

exists, inbreeding reduces clearly the polymorphism and even if homozygosity itself is not harmful in some genes or gene areas, these benefits are lost [92]. As an example of inbreeding effects, the study of sea lions in California showed increased bacterial and helminth infections and longer recovery time [93].

As a result of the founder effect and tight inbreeding, breed-specific physical features, behaviour and over 500 diseases, such as epilepsies, cancers, allergies and autoimmune disorders, have been accumulated in pure-bred dogs [4]. Only humans have been identified with more known genetic diseases. Today the OMIA database identifies 506 inherited canine diseases and more than 60% of these are thought to be shared with humans with very similar physiology, disease presentation and clinical response. Even among the top ten most common inherited canine diseases, there are several that serve as a disease model for humans [3-5]. Mutations behind several monogenic diseases have already been identified, such as the hair ridge, which causes predisposition to dermoid sinus [12], recessive cone-rod dystrophy [13] and ectodermal dysplasia [14].

2.1.2.1 Nova Scotia Duck Tolling Retriever

The NSDTR was developed in the Yarmouth region of Nova Scotia in the early 1800s as a gundog to assist hunters to lure and retrieve ducks. Canine Distemper Virus (CDV) outbreaks were reported to occur twice in 1908 and 1912 reducing the population to only a few individuals. The first NSDTRs registered with the Canadian Kennel Club (1945) were derived from the stock that survived these distemper outbreaks, and the first NSDTRs were imported to the Scandinavian countries as late as in the middle of the 1980s [94] (Figure 6).

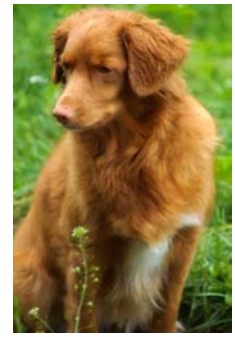

Figure 6 *Novascotian Duck Tolling Retriever. Photo Jarno Nevalainen.*

2.2.2.2 German Shepherd Dog

The GSDs originate from the herding and farm dogs in southern- and central Germany and have been systematically bred since the breed club ”Verein für Deutche Schäferhunde” was founded in 1899. The purpose was to create a versatile working dog to serve humans and the breed remains the most popular working dog worldwide. The first GSDs were imported to Finland in the 1910s [95,96] (Figure 7).

Figure 7 *German Shepherd Dog. Photo Eila Kärkkäinen.*

2.2.3 Dog genome and genomic tools

The fundamental element in identifying the genes for a particular characteristic or disease is a founder effect. The study cohort would ideally consist of individuals with a common ancestry such as Icelanders and Finns in humans, and inbred laboratory animals or pure-bred breeds of pet dogs. When the population descends from a small group of ancestors, they are more likely to share the same founder mutations, and genetic heterogeneity is much lower. In a study of 85 dog breeds, humans and dogs were shown to have essentially the same level of nucleotide heterozygosity when all the dog breeds were considered as one species. As stated earlier, the genetic diversity did not markedly diminish during domestication of the dog from wolves. Within a dog breed, the homogeneity is much greater than within distinct human populations, 94.6 % and 72.5%, respectively, which is supported by the history of severe bottlenecks in breed creation. Therefore, much of the total genetic variation comes from differences between dog breeds and is nearly 6-fold the variation between human populations [97].

2.2.3.1 Dog genome structure provides advantages in gene mapping

The dog genome consists of 76 acrocentric autosomes, and two sex chromosomes. A female chromosome is X and a male chromosome Y, giving a total diploid number of 78. The abbreviation CFA is used here for canis familiaris chromosomes. The dog genome project, which was completed in 2005, provided a high-quality dog genome sequence and a dense single-nucleotide polymorphism (SNP) map containing 2.5 million SNPs. In addition, a detailed haplotype analysis was performed on the whole boxer sequence and a 6% sequence of the genome from 10 additional dogs representing different breeds [9]. LD in domestic dog breeds correlates with breed history. Breeds like the akita, Bernese mountain dog and Pekingese, which have experienced severe bottlenecks in the last 100 years, have LD blocks that extent over 3 Mb. In golden and Labrador retrievers, which are popular breeds without severe bottlenecks in the past, the LD blocks are around 1 Mb. By comparison, in humans, LD blocks are < 100 kb. Extensive LD blocks are an advantage in genome-wide association studies where fewer markers are needed to map the genomic location of the association. The disadvantage is that the identified loci are usually several megabases long. Strong inbreeding of dogs has also resulted in low haplotype diversity in regions of extensive LD, and dog breeds, although highly differentiated genetically, share haplotypes with each other to a high degree. Shared haplotypes are short (<10 kb) and enable the use of related breeds with the same phenotype, and presumably with the same founder mutation, to narrow down the associated loci [9,10] (Figure 8).

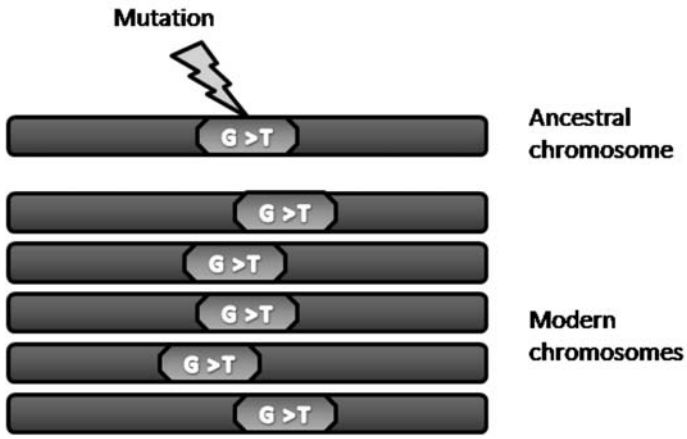

Figure 8 *Haplotype surrounding the mutation in ancestral wolf haplotype and in haplotypes on chromosomes of modern dog breeds.*

Power calculations estimate that genome-wide association analyses may be performed on dogs with 15 000 SNPs. To identify alleles for a simple recessive trait, 20 affected and 20 healthy control dogs are needed. To map complex traits, one would need at least 100 cases and 100 controls for traits that have a fivefold increased risk. For complex traits with only a twofold increased risk, 500 cases and controls are estimated to be needed [9].

There are several other reasons why the dog is an excellent model for complex diseases, in which both genes and environmental risk factors contribute to the development of the disease. Dogs share most of the environmental risk factors with humans. As companion animals, dogs are exposed to smoking, environmental pollution, toxins, radon and sometimes even the same diet. The health of companion dogs is well taken care of and documented. The insurance companies and breed and kennel clubs keep records of pedigrees, and diseases and there are several databases and computing tools available for genetic research. Also full post-mortem tissues are available. Most importantly, the coding sequences of dogs and humans are more similar to each other than to mice and they have a similar physiology, histology and clinical course of the disease [6,98].

2.2.3.2 Dog genetic resources and genomic tools available

New high-throughput technology and dog genome sequence, enable studies at a genome-wide level within a reasonable time and relatively low costs. The dog genome sequence and related resources are available in the databases of the University of California, Santa Cruz (UCSC) http://genome.ucsc.edu/ and the National Center for Biotechnology Information (NCBI) http://www.ncbi.nlm.nih.gov/projects/genome/guide/dog/. Moreover, several databases on canine inherited disorders are available, including the Online Mendelian Inheritance in Animals (OMIA) database at the NCBI site http://omia.angis.org.au/, Canine Inherited Disorders Database (CIDD)

http://www.upei.ca/cidd/intro.htm, which is a joint initiative of the Sir James Dunn Animal Welfare Centre at the Atlantic Veterinary College, University of Prince Edward Island, and the Canadian Veterinary Medical Association and the Inherited Diseases in Dogs Database (IDID) http://server.vet.cam.ac.uk/index.html compiled by David Sargan at the University of Cambridge. In addition, there are several web pages related to genetic research, including the FHCRC Dog Genome Project at the National Human Genome Research Institute, which is a part of the National Institutes of Health (NIH) in Bethesda, Maryland, the Animal Healthtrust in Newmarket, Suffolk http://www.aht.org.uk/ and the LUPA project, a collaborative research project funded by the European commission under the 7th research framework programme http://www.eurolupa.org/.

Traditional sequencing and microsatellite markers are still used, but the new SNP-, CNV- and sequencing array technologies have revolutionized genetic research. The first genome-wide SNP genotyping microarray generated by collaboration of the Broad Institute and Affymetrix contained ~27 000 markers. The improved version with ~50 000 markers and the first Illumina array with ~22 000 SNPs have been subsequently launched [11].

3 Aims of the study

The purpose of this study was to identify genetic loci and variants influencing the susceptibility to different autoimmune diseases in dogs. The main approaches used in this thesis were demonstrating association through known MHC class genes and demonstrating SNP association through a genome-wide association study.

Specific aims were as follows:

1. To investigate whether MHC class II genes DRB1, DQA1 and DQB1 are associated
    a. with IMRD , SRMA and Addisons' disease in NSDTRs.
    b. with CSK in GSDs.
2. To identify novel susceptibility loci for IMRD, SRMA and Addisons' disease in NSDTRs by a genome-wide association study.

4 Materials and methods

4.1 Research site

This study was carried out in Professor Hannes Lohi's research group at the Department of Veterinary Biosciences, the Department of Medical Genetics, the Program in Molecular Medicine, University of Helsinki, and The Folkhälsan Institute of Genetics, Department of Molecular Genetics, Biomedicum I, Helsinki.

SLE project was conducted in collaboration with the research groups of Professor Kerstin Lindblad-Toh at the Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), USA, and the Department of Medical Biochemistry and Microbiology, Uppsala University, Sweden, Associate Professor Helene Hansson-Hamlin at the Department of Clinical Sciences, Swedish University of Agricultural Sciences (SLU) and Professor Göran Andersson at the Department of Animal Breeding and Genetics, SLU. The Addison's disease project was conducted in collaboration with Associate Professor Danika Bannasch at the Department of Population Health and Reproduction and Angela Hughes, DVM at the Department of Medicine and Epidemiology, University of California, USA. Lorna Kennedy, PhD of the University of Manchester has assisted with MHC haplotyping in all three MHC studies.

4.2 Study population

A total of 222 NSDTRs, 65 GSDs, six Cocker Spaniels, twelve boxers and four Petit Basset Griffon Vendeens were included as cases in our studies. In addition, 203 NSDTRs, 39 GSDs, four Cocker Spaniels, twenty boxers and six Petit Basset Griffon Vendeens served as healthy controls. The case control association analysis was performed with 44 SRMA dogs, 37 IMRD and 57 controls. Table 4 describes the study population in detail.

Affected dogs and population controls, unrelated at the grandparent level, were used in all studies, except the AD study, where we used discordant sib-pairs. A partial pedigree of the Finnish SLE-affected NSDTRs is presented in Figure 9.

Table 1 *Number of affected dogs and healthy controls in studies I - IV.*

	Study I		Study II		Study III			Study IV							
	DLA genotyping		DLA genotyping		DLA genotyping			Genome-wide genotyping			Fine-mapping				
BREED	CSK cases	Controls	Hypo-adreno-cortisism cases	Controls	IMRD cases (ANA+)	SRMA cases	Controls	IMRD cases (ANA+)	SRMA cases	Controls	IMRD cases (ANA+)	SRMA cases	Controls	Hypo-adreno-cortisism cases	Canine Lympho-cytic Thyroiditis cases
NSDTR Finnish					12 (6)	22	30	12 (6)	24	29	17 (7)	26	40	3	
NSDTR Swedish					39 (27)	27	48	25 (16)	20	28	54 (25)	49	104	1	20
NSDTR USA			29	34							10	3	59	39	
German Shepherd	30	25									5		14		
Cocker Spaniel											6		4		
Boxers												12	20		
Petite Basset Griffon Vendeen												4	6		

NSDTR = Nova Scotian Duck Tolling Retriever, CSK = chronic superficial keratitis, IMRD = immune-mediated rheumatic disease, SRMA = steroid responsive meningitis arteritis, ANA = antinuclear autoantibody, DLA = Dog leukocyte antigen

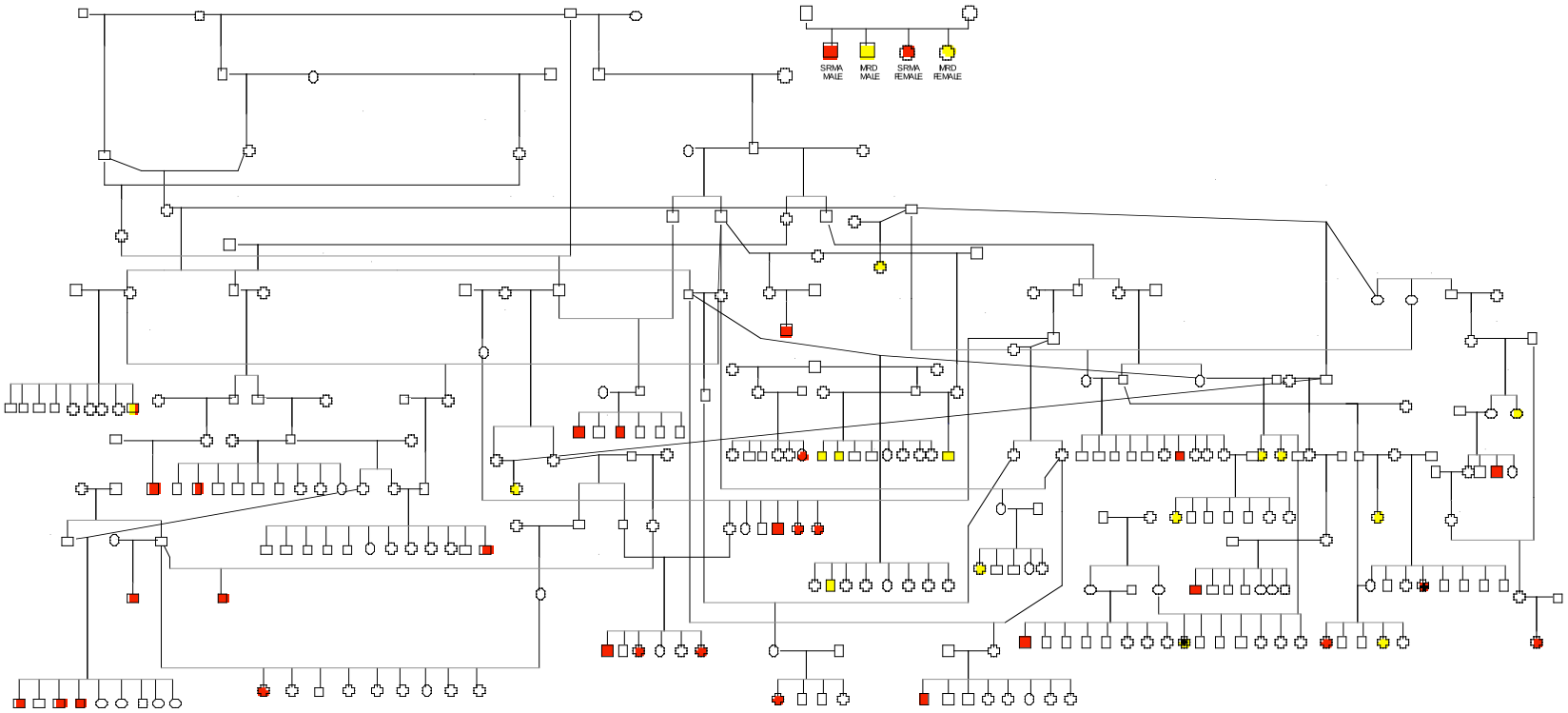

Figure 1 *A partial pedigree of the Finnish Nova Scotian Duck Tolling Retrievers (NSDTRs) affected with Systemic lupus erythematosus (SLE) - related disease. Steroid responsive meningitis arteritis (SRMA) dogs are shown in red and immune-mediated rheumatic disease (IMRD) dogs in yellow.*

4.3 Diagnostic procedures

Dogs were chosen based on strict inclusion and exclusion criteria. To be classified as affected by IMRD, the dogs had to have displayed musculoskeletal disorders consistent with symmetrical polyarthritis, suffered from pain affecting several joints of the extremities and displayed stiffness, mainly after rest. Signs needed to be apparent for at least 14 days. The presence of ANA was tested by using indirect immunofluorescence (IIF-ANA). The test was considered positive at a titre of ≥1:100. 33 out of 51 study dogs tested IIF-ANA positive. Positive ANA tests were repeated 2-3 months later, with the same positive result and the same IIF-ANA pattern. Sera from all healthy controls were negative on the IIF-ANA test. Dogs classified as SRMA-affected displayed high fever and strong neck pain and responded to corticosteroid treatment. The diagnostics procedure is described in more detail by Hansson-Hamlin [15].

For a dog to be classified as affected by AD, an adrenocorticotropic hormone (ACTH) stimulation test must have been performed with pre- and post-ACTH stimulation serum cortisol concentrations <2.5 µg/dl (68 nmol/l). Exclusion criteria were dogs with no clinical signs of any autoimmune disease and >7 years of age.

The dogs diagnosed with CSK, as well as the healthy controls, had undergone a thorough ophthalmic eye examination by an experienced ophthalmologist to ensure their recent ocular health status. We also developed a detailed health questionnaire that was sent to all participating dog owners. Besides the demographic information and specific questions about CSK, the questionnaire collected information about the dog's relatives. All dogs affected with any other eye or autoimmune diseases were excluded.

4.4 Blood samples and DNA isolation

All animals in our study were privately owned pets. A DNA sample was donated for use in a genetic study. EDTA- blood (3-5 ml) was collected at various veterinary clinics and dog shows with the owners' consent. DNA was isolated using standard procedures (studies II, III).

4.5 Sequencing for MHC class II and allele assignment (I-III)

To identify the MHC class II haplotypes, we first sequenced the purified PCR products of exon 2 from DLA locus genes DRB1, DQA1 and DQB1 using an ABI 3730 or 3730xl sequencer. The sequences were analysed manually and compared with a consensus sequence with the MatchToolsNavigator program. The program MatchTools was used to assign the DRB1-, DQA1- and DQB1 alleles by comparing the sequences against a large sequence library (http://www.ebi.ac.uk/ipd/mhc/index.html). MHC-II haplotypes were built using information from previous studies (Dr. Lorna Kennedy, personal data).

4.6 Genome-wide genotyping (IV)

Genome-wide association (GWA) genotyping was performed at Biomedicum Helsinki using 22,000 validated SNPs in the CanineSNP20 BeadChip panel and the Illumina's Infinium HD DNA Analysis instrument. Automated genotype calling was performed using the Illuminus software [99].

4.7 Fine-mapping of the associated regions (IV)

A total of 822 SNPs at ≈1 SNP/10 kb density were genotyped by the iPLEX SEQUENOM MassARRAY platform at the Broad Institute. Fine-mapping was performed with samples used in GWA, additional samples of NSDTRs and dogs with the same phenotype from other breeds. Additional breeds with the IMRD or SRMA phenotype were included to identify shared haplotypes across breeds (Figure 8). Also NSDTRs with other autoimmune diseases, AD and CLT, were included to identify loci predisposing to autoimmune disorders in general. The total sample size at this point was 416 NSDTRs and included 81 IMRD cases (32 ANA-positive), 78 SRMA cases, 43 AD cases, 20 CLT cases and 203 healthy controls. The details of the samples are presented in Table 4. The normal procedure requires an independent association in each breed, but due to low sample sizes we performed the haplotype analysis without this prior knowledge.

4.8 Statistical analysis (I-IV)

In studies I-III, the cases and controls were divided into separate groups in various ways based on the presence or absence of an allele, genotype or haplotype. In each group the number of allele, genotype or haplotype was calculated and a 2x2 Contingency Table was created to display the proposition of each variable in a matrix format. The significance of the frequencies between cases and controls was assessed by χ^2 statistics and odds ratios (OR) and relative risks with 95% or 99% confidence intervals and p-values were calculated. While calculating the OR for the risk haplotype homozygosity in CSK and AD where the number of homozygous controls was zero, we used a pseudocount (1) to be able to perform the calculations. This approach means adding number one to each observed number of counts [100].

In study IV, all SNP and haplotype associations were analyzed with the software package PLINK [101]. For marker quality control, SNPs with a minor allele frequency of <5% were excluded, as were SNPs with call rates <75%. Our GWA samples were collected from two different countries in which the NSDTR breed may have been divided into separate subpopulations causing a false association due to differentiated allele frequencies between the subpopulations (=population stratification, PS). To test for the presence of PS in our sample we used multidimensional scaling (MDS) analysis to construct multidimensional scaling plots, where each spot corresponds to a specific individual. PS was adjusted by IBS clustering with two groups, which reflects the population structure in our sample. We opted for a multi-population analysis approach,

and chose Cochran-Mantel-Haenszel (CMH) conditional on clustering as the primary association analysis method [102]. The robust and generally accepted $p<5 \times 10\text{-}6$ was chosen as the limit for significance, with further interest for follow-up studies paid to loci showing with multiple SNPs within close proximity with p-values between $1 \times 10\text{-}5$ - $5 \times 10\text{-}6$. A quantile-quantile plot of the CMH analysis and overall inflation factor ($\lambda = 1.2$) were used as final quality control measures. The genome-wide significance was assured by using 100 000 permutations. The analysis was performed for the IMRD and SRMA sub-phenotypes separately and all cases combined.

At fine-mapping stage, for marker quality control, SNPs with a minor allele frequency of <1% were excluded, as were SNPs with call rates <80%. The same settings were used to analyse two to eight SNP haplotypes generated with a sliding window approach provided by PLINK.

4.9 Ethical issues

We have a license authorized by the Animal Experiment Committee of the County Administrative Board of Southern Finland (ESLH-2009-07827/Ym-23, valid until 16.10.2012). No genetically modified dogs are produced in our Finnish dog genetics programme. We do not breed dogs for research purposes or own any of the dogs studied.

5 Results

5.1 MHC class II candidate gene studies (I-III)

MHC class II gene region has been associated in most, if not all, autoimmune disorders in humans and in several autoimmune disorders in dogs. As there are no previous genetic studies reported in CSK, IMRD, SRMA or AD, we investigated whether specific DLA-DRB1*DQA1*DQB1 – risk haplotypes exist in any of these disorders.

5.1.1 DLA class II polymorphism in Finnish, Swedish and North-American NSDTRs

We genotyped altogether 241 NSDTRs, 64 from Finland, 114 from Sweden and 63 from North America (Table 4) and identified five DLA-DRB, four DLA-DQA1 and five DLA-DQB1 alleles, which formed seven different haplotypes (Table 5). Two of the most common haplotypes were seen in very high frequencies, the third, fourth and fifth most common haplotypes were seen in moderate frequencies and the rest were extremely rare, seen in only a few dogs. A clear difference existed in haplotype one and five frequencies between North-American and Scandinavian dogs, with p-values of 1.97^{e-25} and 1.076^{e-07}, respectively. Finnish and Swedish dogs showed a similar distribution of haplotypes with each other, as did the US and Canadian dogs.

Table 5 *Haplotype frequencies in Scandinavian and North-American Nova Scotia Duck Tolling Retrievers.*

	Haplotype			Scandinavian		American	
Haplotype no.	DRB1 Allele	DQA1 Allele	DQB1 Allele	Haplotype frequency (no. + %)		Haplotype frequency (no. + %)	
1	00601	005011	02001	142	40.3	31	22.8
2	01502	00601	02301	120	34.1	43	31.6
3	01501	00601	00301	53	15.1	26	19.1
4	02301	00301	00501	34	9.7	10	7.4
5	01501	00601	02301			22	16.2
6	00401	00201	01501	3	0.9	2	1.5
7	01502	00601	00301			2	1.5
Total	5	4	5	352	100.0	136	100.0

5.1.2 DLA class II polymorphism in Finnish GSDs

We genotyped 55 GSDs and identified eight DLA-DRB1, five DLA-DQA1 and eight DLA-DQB1 alleles, which formed eleven different haplotypes (Table 6). The haplotypes were unevenly distributed and the two most common ones were seen in very high frequencies, comprising over 75% of all haplotypes in the study population. A third haplotype was seen at a moderate frequency of 9.1%. All of the other haplotypes were

rare, seen only in one to three dogs. Three of the dogs carried a double DLA-DQB1 allele on one haplotype, consisting of alleles DQB1*01303 and DQB1*01701; this was named DLA-DQB1*013017. One of the identified alleles in the DLA-DRB1 locus was new and was officially named DLA-DRB1*01104 and submitted to the NCBI database (accession number FN995992). One DLA-DQB1 allele was also new and has now officially been named DLA-DQB1*05901 (accession number FN995993).

Table 6 *Haplotype frequencies in the Finnish German Shepherd Dogs.*

	Haplotype			CSK		Controls		All	
Haplo-type no.	DRB1 Allele	DQA1 Allele	DQB1 Allele	Haplotype frequency (no. + %)		Haplotype frequency (no. + %)		Haplotype frequency (no. + %)	
1	01101	00201	01302	20	34.4	26	53.3	46	41.8
2	01501	00601	00301	26	43.1	11	22.2	37	33.6
3	00101	00101	00201	5	8.6	5	11.1	10	9.09
4	00201	00901	00101	1	1.7	2	2.2	3	2.7
5	00102	00101	00201	2	3.4	1	0	3	2.7
6	01201	00401	013017	1	1.7	2	4.4	3	2.7
7	01502	00601	02301	2	3.4			2	1.8
8	01104	00201	01302			2	4.4	2	1.8
9	01502	00601	05901	2	3.4			2	1.8
10	01101	00201	01303	1	1.7			1	0.9
11	01501	00601	02301			1	2.2	1	0.9
Total	8	5	8	60	100	50	100	110	100

CSK=chronic superficial keratitis

5.1.3 DLA class II haplotype association with CSK in GSDs

We genotyped thirty affected CSK dogs and twenty-five healthy population controls to search for an association with the MHC class II region. All of the dogs participating in the study had been examined by an experienced ophthalmologist to confirm the diagnosis, to exclude other eye diseases and to confirm the recent health status of control dogs. The DLA-DRB1*01501/DQA1*00601/DQB1*00301 haplotype was significantly associated with the CSK in GSDs (OR=2.67, 95% CI=1.17-6.44, p = 0.02) (Table 7).

Table 7 *Association of the risk haplotype DRB1*01501/DQA1*00601/DQB1*00301with chronic superficial keratitis in German Shepherd Dogs.*

Association with risk haplotype DRB1*01501/DQA1*00601/DQB1*00301				
CSK % (no.)	Control % (no.)	Odds ratio	95% CI	p-value
43.1 (26)	22.2 (11)	2.67	1.17-6.44	0.02

CSK=chronic superficial keratitis

5.1.4 DLA class II haplotype association with hypoadrenocortisism in NSDTRs

To test for an association between the MHC class II locus and AD, we genotyped twenty-nine AD-affected NSDTRs, twenty-one from the USA and eight from Canada. The diagnosis was based on the ACTH stimulation test. In addition, five AD-suspected NSDTRs were included, three from the US and two from Canada. The diagnosis of these dogs was based on clinical signs and findings, and response to treatment. Eleven unaffected full- and half-siblings of the affected NSDTRs, nine from the US and two from Canada, and twenty-three country-matched NSDTRs were included as controls. We found that the DLA-DRB1*01502/DQA*00601/DQB1*02301 haplotype was significantly associated with AD in NSDTRs (OR = 2.1, 95% CI = 1.0-4.4, p = 0.044) (Table 8). The association was even stronger when examining the US population alone (OR = 2.8, 95% CI = 1.1-7.1, p = 0.025).

Table 8 *Association of the risk haplotype DLA-DRB1*01502/DQA*00601/DQB1*02301 with hypoadrenocortisism in American Nova Scotian Duck Tolling Retrievers.*

	Association with risk haplotype DLA-DRB1*01502/DQA*00601/DQB1*02301				
	Addison's disease % (no)	Control % (no.)	Odds Ratio	95% CI	p-value
USA dogs	39.6 (24)	18.8 (24)	2.8	1.1-7.1	0.044
Canadian dogs	40.0 (10)	35.0 (10)			
All dogs	38.6 (34)	22.9 (34)	2.1	1.0-4.3	0.047

5.1.5 DLA class II haplotype association with IMRD in NSDTRs

A total of 176 dogs were genotyped in this project. We studied 51 IMRD dogs, 33 dogs of which tested ANA positive on the IIF-ANA test. The sample also included 49 SRMA dogs and 78 healthy controls. Two dogs were affected by both SRMA and IMRD. We found an elevated risk for IMRD in dogs that carried the DLA-DRB1*00601/DQA1*005011/DQB1*02001 haplotype (OR = 2.0, 99% CI = 1.03-3.95, p = 0.01) and for ANA-positive IMRD dogs (OR = 2.3, 99% CI = 1.07-5.04, p = 0.007) (Table 9). No association was observed with the SRMA.

Table 9 *Association of the risk haplotype DLA-DRB1*00601 /DQA1*005011 /DQB1*02001 with immune-mediated rheumatic disease in Scandinavian Nova Scotian Duck Tolling Retrievers.*

Association with risk haplotype DLA-DRB1*00601 /DQA1*005011 /DQB1*02001				
Disease % (no.) affected	Control % (no.)	Odds ratio	99% CI	p-value
IMRD				
51.0 (52)	34.6 (53)	2.0	1.0-4.0	0.01
ANA +				
54.5 (33)	34.0 (53)	2.3	1.1-5.0	0.007

IMRD=immune-mediated rheumatic disease, ANA=antinuclear antibody.

5.1.6 Association of MHC class II homozygosity with autoimmunity

Homozygosity for the risk haplotype increased the disease risk significantly in all studies. In CSK and AD, none of the control dogs were homozygous for the risk haplotype, whereas in IMRD 11.5% of the controls carried two copies of the risk haplotype. The DLA-DRB1*01501/DQA1*00601/DQB1*00301 haplotype was homozygous in 8 CSK-affected dogs and none of the controls ($OR_{estimate}$ > 8.5, 95% CI = 1.4-224, p_{fisher} = 0.017) (Table 10).

Table 10 *Association of homozygosity in DRB1*01501/DQA1*00601/DQB1*00301 risk haplotype with chronic superficial keratitis in Finnish German Shepherd Dogs. Pseudocount (1) has been added to the numbers.*

Homozygosity for risk haplotype DRB1*01501/DQA1*00601/DQB1*00301				
CSK % (no.)	Control % (no.)	Odds ratio	95% CI	p-value
36.0 (9)	4 (1)	8.5	1.4-224	0.017

CSK=chronic superficial keratitis

The DLA-DRB1*01502/DQA*00601/DQB1*02301 haplotype was homozygous in six AD-affected NSDTRs and in none of the controls ($OR_{estimate}$ > 8.9, 95% CI = 1.4-237.7, p_{fisher} = 0.02) (Table 11).

Table 11 *Association of homozygosity in the DLA-DRB1*01502/DQA*00601/DQB1*02301 risk haplotype with hypoadrenocortisism in American Nova Scotian Duck Tolling Retrievers. Pseudocount (1) has been added to the numbers.*

Homozygosity for risk haplotype DLA-DRB1*01502/DQA*00601/DQB1*02301				
Addison's disease % (no.)	Control % (no.)	Odds ratio	95% CI	p-value
20.6 (7)	3.5 (1)	8.9	1.4-237.7	0.02

The twofold risk present in heterozygous IMRD dogs increased up to fivefold if the dogs were homozygous for the haplotype (OR = 4.9, 99% CI = 1.52-16.0, p = 0.0005). ANA-positive dogs were even more susceptible, with a sevenfold risk (OR = 7.2, 99% CI = 2.0-25.9, $p \leq 0.0001$) (Table 12).

Table 12 *Association of homozygosity in the risk haplotype DLA-DRB1*00601/DQA1*005011 /DQB1*0200 haplotype with immune-mediated rheumatic disease in Scandinavian Nova Scotian Duck Tolling Retrievers.*

	Homozygosity for risk haplotype DLA-DRB1*00601 /DQA1*005011 /DQB1*02001			
Disease % (no.) affected	Controls % (no.)	Odds ratio	99% CI	p-value
IMRD 39.2 (20)	11.5 (9)	4.9	1.52-16.0	0.0005
ANA + 48.5 (16)	11.5 (9)	7.2	2.0-25.9	<0.0001

IMRD=immune-mediated rheumatic disease, ANA=antinuclear antibody.

In addition, in CSK and AD, an overall homozygosity, regardless of the haplotype was shown to increase the risk, and in AD it was also associated with early onset of the disease. Fourteen out of 30 CSK cases versus four out of 25 controls were homozygous for the MHC class II haplotypes in general (OR=4.37, 95% CI=1.27-18.46, p = 0.02) (Figure 10). Ten AD-affected dogs were homozygous for MHC class II compared with only two control dogs (OR = 6.7, 95% CI = 1.5-29.3, p = 0.011) (Figure 11).

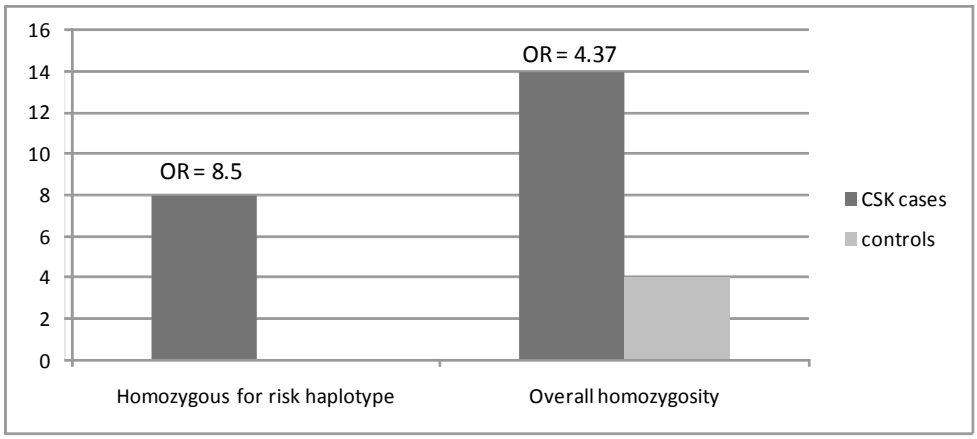

Figure 10 *Number of dogs homozygous for major histocompatibility complex (MHC) class II region: comparison of dogs with and without chronic superficial keratitis (CSK). CSK dogs indicated in dark gray and control dogs in light gray.*

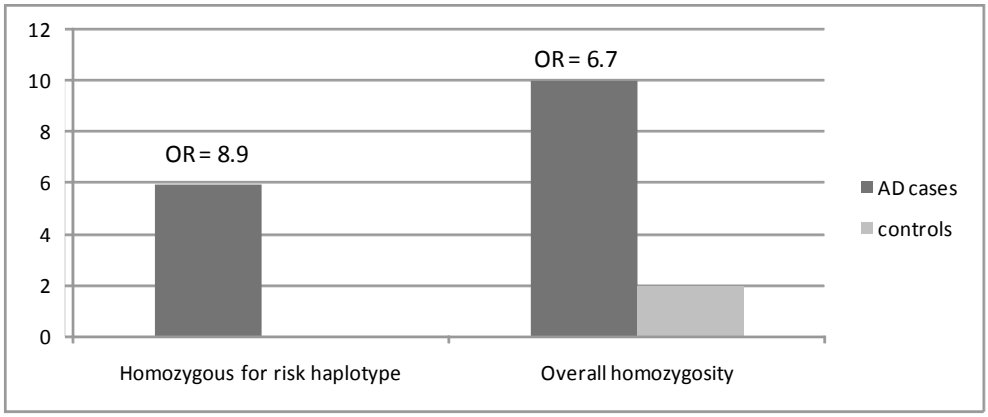

Figure 11 *Number of dogs homozygous for the major histocompatibility complex (MHC) class II region: Comparison of dogs with and without Addison's disease (AD). AD dogs indicated in dark grey and control dogs in light gray.*

5.2 Genome-wide association and fine-mapping studies in dogs (IV)

In study IV, we performed the first successful genome-wide association study for a complex disease in dogs identifying IMRD and SRMA loci in NSDTRs . Genotyping on 81 SLE-related disease affected dogs and 57 healthy control dogs was performed using the Illumina's canine-specific 22k SNP chip arrays. GWA was followed by fine-mapping the associated regions with additional samples and markers.

5.2.1 GWAS

Based on the MDS analysis, we observed some stratification when all dogs were analysed together and when SRMA-affected dogs were analysed alone (Figure 12).

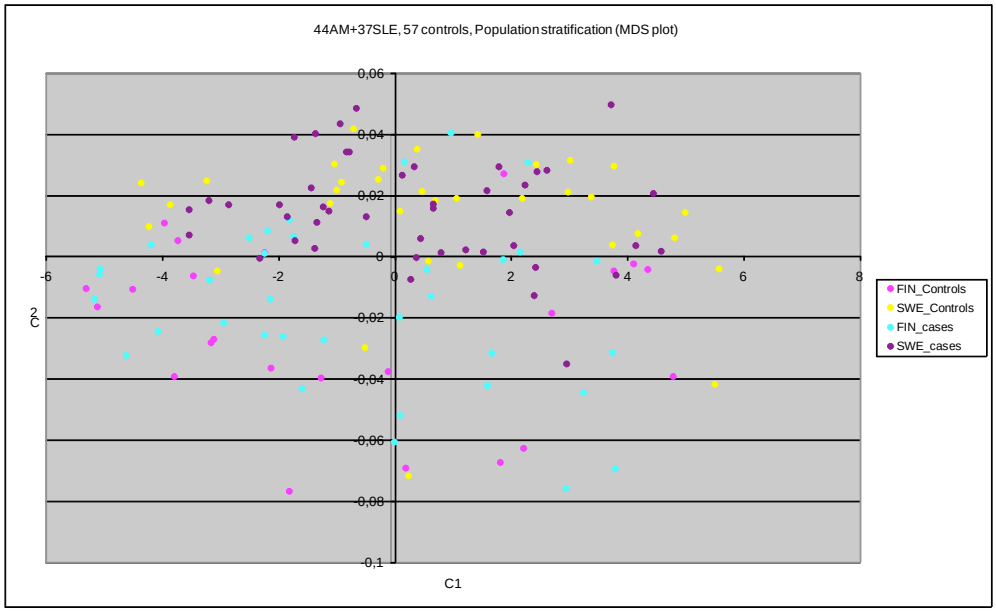

Figure 12 *Multidimensional scaling plot of the genome-wide SNP data showing minor population stratification (PS). When spots are totally overlapping each other, there is no PS.*

We identified five loci associated with the SLE–related disease complex, from which three loci reached genome–wide significance after correction for multiple tests (p-value = 0.02-0.04). All results are combined in Table 13.

We found a large region containing multiple associated SNPs on canine chromosome (CFA) 32 at ≈25 Mb (p_{raw}= 1.5 x 10^{-5} and p_{genome}= 0.12) when all cases were analysed as one group. After correction for stratification, the region showed an even stronger association (p_{raw}= 7.9 x 10^{-6} and p_{genome}= 0.06).

Analysis of the ANA-positive IMRD sub-phenotype separately revealed four highly associated regions in CFA 3, 8, 11 and 24. The strongest associations were found for a single SNP on CFA 8 at ≈69 Mb (p_{raw}= 1.5 x 10^{-6} and p_{genome}= 0.02) and on multiple SNPs on CFA 24 at ≈38 Mb (p_{raw}= 3.2 x 10^{-6} and p_{genome}= 0.04). Both associations reached genome-wide significance. We also found multiple associated SNPs on CFA 11 at ≈66 Mb (p_{raw}= 7.4 x 10^{-6} and p_{genome}= 0.08) and on CFA 3 at ≈57 Mb (p_{raw}= 2.2 x10^{-5} p_{genome}= 0.18), which did not reach genome-wide significance. Very minor population stratification (IR= 1.2) did not need to be corrected for.

SRMA-affected dogs showed stronger stratification by nationality and were analysed in three different ways, as one group and as Swedish and Finnish dogs separately. When analysing all SRMA-affected dogs together, we identified two regions with multiple associated SNPs, one on CFA 28 at ≈14 Mb (p_{raw}= 6.5 x 10^{-5} and p_{genome}= 0.37) and another on CFA 32 at ≈25 Mb (p_{raw}= 7.4 x 10^{-5} and p_{genome}= 0.40). The latter overlapped the peak seen in the analysis of both phenotypes together and reached genome-wide significance after correction for stratification (p_{raw}= 7.10 x 10^{-6} and p_{genome}= 0.04), while the former association was lost. The association on CFA 32 persisted when the Swedish

dogs were analysed alone, but did not reach genome-wide significance (p_{raw}= 2.1 x 10^{-4} and p_{genome}= 0.69) (data not shown). When analysing the Finnish SRMA dogs separately, genome-wide significance was found for a single SNP on CFA 30 at ≈29 Mb (p_{raw}= 6.7 x 10^{-6} and p_{genome}= 0.03).

Table 13 *Results of the genome-wide association and finemapping studies using 81 affected and 57 healthy control dogs. Modified from* [103].

Chr	Disease phenotype		GWA p-value	Fine-mapping p-value	OR	Size (kb)	Most relevant candidate genes
3	ANA+	IMRD	2.2 x 10 $^{-5}$	2.2 x 10 $^{-11}$	4.4	113	*AK126887, AP3B2, SCARNA15, FSD2, RPL23A, WHDC1L1 and HOMER2*
8	ANA+	IMRD	1.5 x 10 $^{-6}$	5.2 x 10 $^{-5}$	3.0	895	*SNRPE and VRK1*
8	SRMA		n.s.	2.5 x 10 $^{-7}$	2.4	78	*between SNRPE and VRK1*
11	ANA+	IMRD	7.4 x 10 $^{-6}$	5.5 x 10 $^{-13}$	8.1	127	*EPB41L4B, C9orf4 and PTPN3*
24	ANA+	IMRD	3.2 x 10 $^{-6}$	1.5 x 10 $^{-12}$	5.1	127	*AK128395, WFDC10B, WFDC13, AY372174, WFDC1 and DNTTIP1*
32	ANA+	IMRD	7.3 x 10 $^{-5}$	2.9 x 10 $^{-7}$	3.4	1.600	*DAPP1, PPP3CA and BANK1*
32	SRMA		7.1 x 10 $^{-6}$	8.8 x 10 $^{-8}$	0.3	723	*DAPP1*
32	ANA+	IMRD	7.9 x 10 $^{-6}$	1.1 x 10 $^{-6}$	3.1	1.500	*DAPP1, PPP3CA and BANK1*

GWA=genome-wide association analysis. The sizes of the associated regions are based on the fine-mapping data within the breed complemented with the areas of haplotype-sharing across breeds.

5.2.2 Fine-mapping

To replicate and narrow down the associated regions, we performed fine-mapping with additional samples and 822 SNPs. The average density of SNPs was one SNP/10 Kb. We used a total of 425 NSDTRs, 81 IMRD-affected, of which 32 tested ANA-positive, 78 SRMA-affected, 43 AD-affected, 20 CLT-affected and 203 healthy controls. Additional breeds were used to identify shared haplotypes between the breeds, which are expected to be much shorter than within a breed. To analyse ANA-positive IMRD, GSDs and cocker spaniels were included, and to analyse SRMA Boxers and Petite Basset Griffon Vendeen were included (Table 4).

Fine-mapping data obtained from IMRD- and SRMA-affected dogs and their healthy controls were analysed in three groups: dataset 1 (dogs included in GWA, n = 138), dataset 2 (additional NSDTRs, including related individuals, n = 186) and a combined dataset (datasets 1 and 2, n = 324).

Two of the loci associated with both sub-phenotypes, CFA 8 and CFA 32. P-values and odds ratios are presented in more detail in Table 13, but were in the range of 10^{-5}–10^{-8} and 2.4–3.4, respectively. The CFA 32 locus remained at approximately the same strength and is still relatively large with three signals across a 1.6-Mb region. The CFA 32 locus contains three relevant candidate genes (*DAPP1*, *PPP3CA* and *BANK1*); *BANK1* has previously been associated with human SLE [104]. Other genes within the CFA 32 locus

include *MAP2K1IP1*, *DNAJB14*, *H2AFZ*, *RPS23*, *DDIT4L*, *EMCN*, *RPS24* and *RPS17*. The locus in CFA 8 contains the genes *SNRPE* and *VRK1* and four SNPs with a 97-kb haplotype were shared between Boxers and Petit Basset Griffon Vendeens affected with SRMA.

When analysing the ANA-positive dogs alone, three of the four peaks showed an even stronger association on chromosomes 3, 11 and 24, with p-values of 10^{-11}–10^{-13} for the combined dataset and with higher, but nevertheless significant, *P*-values for datasets 1 and 2 (10^{-4}–10^{-11}). The strongest association was found on CFA 11, which contains a highly associated 124-kb five-SNP haplotype in NSDTR and a 127-kb shared haplotype between NSDTR, GSD and Cocker Spaniel (CS) cases. This locus contains the genes *EPB41L4B*, *C9orf4* and *PTPN3*. The second best association is found on CFA 24, containing a seven-SNP 96-kb haplotype in NSDTR and a 127-kb shared haplotype between NSDTR, GSD and CS cases. This candidate locus contains six genes, *AK128395*, *WFDC10B*, *WFDC13*, *AY372174*, *WFDC1* and *DNTTIP1*. The association on chromosome 3 includes a two-SNP 113-kb haplotype in NSDTR and a 256-kb shared haplotype between NSDTRs, GSDs and CSs. This locus contains seven genes, *AK126887*, *AP3B2*, *SCARNA15*, *FSD2*, *RPL23A*, *WHDC1L1* and *HOMER2*.

6 Discussion

Characterization of the genetic background behind canine SLE-related disease, AD and CSK, will likely reveal novel genes and pathways and help to elucidate the aetiology of the diseases. It will also establish clinically relevant large animal models for human studies in canine breeds, and gene discovery would enable development of new gene tests to assist in breeding plans.

Given the major role of the MHC II in human autoimmune disorders, we strongly suspected an influence of the DLA region behind the aetiology of these canine diseases and chose to perform a candidate gene study to see if this was the case. We also aimed to prove the dog's power in genome-wide studies since it had been hypothesized that the unique breed and genome structure of dog would simplify association studies also in complex diseases. The dog genome sequence was published in 2005 and the first microarrays were launched in 2007, and we were fortunate to be among the first researchers to whom these tools were available.

No genetic risk factors have previously been identified behind any of the canine diseases in this study. Our findings support the suspected autoimmune origin of all diseases investigated, as the associations were identified with genes strongly linked with autoimmune diseases in humans and dogs and with genes functioning in potential immunological pathways.

6.1 Genetic diversity and population structure indicate narrow genetic diversity in GSDs and NSDTRs

The distribution and frequencies of the DLA-DRB1, -DQA1 and -DQB1 alleles and haplotypes vary between breeds. Some alleles may be breed-specific or limited to a few breeds, while some breeds show only a very limited distribution of alleles due to breed history [56,105].

Considering the population history, we observed an expected number of haplotypes in NSDTRs, which were unevenly distributed. The level of genetic diversity is lower in this breed than that found in many other dog breeds [106]. We identified five DLA-DRB, four DLA-DQA1 and five DLA-DQB1 alleles, which formed seven different haplotypes. Two of the most common haplotypes were seen in very high frequencies in both Scandinavian and American study populations, consisting together of 74.4% and 54.4% of all haplotypes, respectively. The third and fourth common haplotypes were seen in moderate frequencies in both populations, and the fifth in American dogs. The rest were extremely rare and seen in only a few dogs. There was a clear difference in the haplotype frequencies between North-American and Scandinavian dogs. Finnish and Swedish dogs showed a more similar distribution of haplotypes. The samples had been selected for a disease association study and were not therefore ideal for population diversity studies. In addition, in AD study we used also discordant sib-pairs as healthy controls. Larger number of unrelated individuals should be genotyped to evaluate the DLA polymorphism in American NSDTRs more accurately. Both the history and pedigree information suggest

that today's NSDTRs are one population, but during the last three decades, the subpopulations in different countries have diverged in different directions. MHC studies show that MHC haplotypes do not differ between countries, but a slight difference exists in haplotype frequencies. In addition, the MDS plot based on genome-wide genotyping indicates that breeds on different continents have become differentiated to some degree.

The situation was better for GSDs, but as it is one of the most common breeds in the world, one might have expected to see more genetic diversity. The study population was, however, much smaller, but our objective was to explore the association with a disease phenotype, not to study the diversity in the breed. Unrelated population controls were used, giving a wider perspective to the diversity than could be seen by using discordant sib-pairs. Still, the Finnish data is biased by half of the dogs being selected for being affected with CSK. The data provided by Lorna Kennedy show 10 haplotypes considered typical for the breed and an additional 21 haplotypes common in other breeds but seen in only a few GSDs and could therefore be descendant from breed-crosses (Table 14). The haplotypes were unevenly distributed, which is commonly seen in pure-bred dogs and reflects the narrowed gene pool.

Table 14 *Haplotype frequencies of 322 German Shepherd Dogs with mainly European origin (Lorna Kennedy, personal communication).*

Haplotypes considered specific for GSDs									
								Homozygous dogs	
DRB1	DQA1	DQB1	n	No. of haplos	Haplotype frequency	No. of dogs	% Dogs	No. +	%
01101	00201	01302	332	245	36,9	196	59,0	49	14,8
01501	00601	00301	332	141	21,2	119	35,8	22	6,6
00101	00101	00201	332	70	10,5	66	19,9	4	1,2
00102	00101	00201	332	46	6,9	43	13,0	3	0,9
01501	00601	02301	332	43	6,5	39	11,7	4	1,2
01201	00401	013017	332	25	3,8	23	6,9	2	0,6
00201	00901	00101	332	14	2,1	13	3,9	1	0,3
01502	00601	02301	332	14	2,1	13	3,9	1	0,3
01502	00601	00301	332	10	1,5	10	3,0	0	0,0
00601	005011	00701	332	8	1,2	7	2,1	1	0,3
Haplotypes considered to present GSD crossbreeds									
								Homozygous dogs	
DRB1	DQA1	DQB1	n	No. of haplos	Haplotype frequency	No. of dogs	% Dogs	No. +	%
00202	00901	00101	332	5	0,8	5	1,5	0	0,0
02001	00401	01303	332	5	0,8	5	1,5	0	0,0
00401	00201	01501	332	4	0,6	4	1,2	0	0,0
01201	00401	01303	332	4	0,6	2	0,6	2	0,6
00102	00101	00802	332	3	0,5	3	0,9	0	0,0
01101	00201	01303	332	3	0,5	3	0,9	0	0,0
01104	00201	01302	332	3	0,5	3	0,9	0	0,0
04001	01001	01901	332	3	0,5	3	0,9	0	0,0
01801	00101	00802	332	3	0,5	2	0,6	1	0,3
00601	00401	01303	332	2	0,3	2	0,6	0	0,0
01601	00101	00201	332	2	0,3	2	0,6	0	0,0
01501	00601	02201	332	2	0,3	1	0,3	1	0,3
00801	00301	00401	332	1	0,2	1	0,3	0	0,0
00901	00101	008011	332	1	0,2	1	0,3	0	0,0
01201	00101	00201	332	1	0,2	1	0,3	0	0,0
01501	00901	00101	332	1	0,2	1	0,3	0	0,0
01501	00601	01901	332	1	0,2	1	0,3	0	0,0
01501	00601	05401	332	1	0,2	1	0,3	0	0,0
02001	00401	01302	332	1	0,2	1	0,3	0	0,0
02301	00301	00501	332	1	0,2	1	0,3	0	0,0
04801	00402	02301	332	1	0,2	1	0,3	0	0,0

6.2 MHC class II is a major genetic risk factor also in canine autoimmune diseases, proving the autoimmune origin

MHC class II gene region is the major risk locus in human autoimmune disorders and has been associated with several autoimmune disorders also in domestic dogs. The DLA associations of IMRD, CSK and AD combined with previous reports provide strong evidence for DLA class II being the major risk locus in canine autoimmunity. Table 15 lists all reported MHC class II associations with canine autoimmune disorders.

The risk haplotype for IMRD in NSDTRs is present in 45% in the overall study population, being the most common haplotype. The frequency among the control dogs is slightly over 30% in each country. The risk haplotype has also been found in six other

breeds: Alaskan malamute, Labrador retriever, Newfoundlander, golden retriever, Cavalier King Charles spaniel and Cocker spaniel [106]. A similar haplotype, differing only by a DLA-DQB1 allele, has been shown to predispose to immune-mediated haemolytic anaemia in a combined dataset of different breeds [30]. The DLA-DRB1 allele in IMRD risk haplotype has been found in 100% of the Dobermans affected with Doberman hepatitis [107]

The risk haplotype for AD is the second most common haplotype in NSDTRs worldwide, present in over 40% of the dogs. The AD risk haplotype has recently been shown to be protective against NME in pug dogs [108] and shares a DQA1 allele with the CSK risk haploype (Study I). In addition, it has previously been associated with a slightly increased risk for diabetes and a similar haplotype carrying the same DRB1- and DQA1 alleles has been associated with IMHA, although this association was lost when using breed-matced controls [27,30].

It remains to be seen whether DR or DQ or both are the actual genetic risk factors for IMRD and AD. The high frequency of both risk haplotypes in NSDTRs is problematic because removal of breeding animals carrying this haplotype would create a serious genetic bottleneck, which would severely threaten the viability of the breed. A careful planning in breeding, including exclusion of the affected individuals from breeding and avoiding production of litters with puppies homozygous for the risk haplotype, is the only solution, unless breed-crosses are considered. Increasing the frequency of the rare haplotypes might be an option, but there is of course no knowledge of either their protective or disease-causing associations. The MHC class II haplotypes are not the only genetic risk factors behind these autoimmune disorders, although the DLA gene region is probably the major predisposing locus. The GWAS results will be discussed later in this section.

Table 15 *All reported DLA-DRB1, DQB1 and DQA1 associations with canine autoimmune disorders, including our results.*

		Associated allele or haplotype			Heterozygotes		Homozygotes		
Disease	**Breed**	**DRB1**	**DQA1**	**DQB1**	**OR**	**P**	**OR**	**P**	**Comments**
CSK	GSD	01501	00601	00301	2.67	0.02	8.5	0.017	Elevated risk in dogs homozygous regardless of the haplotype
AF	GSD	00101			5.01	1×10^{-8}			Homozygosity associated with early onset
Diabetes		009	001	008	2.05	0.0002			
Diabetes		015	006	023	1.52	0.0006			
Diabetes		002	009	001	1.51	0.03			
Diabetes			004	013	0.66	<0.005			Protective
Diabetes			Arg55+		1.82	$<5 \times 10^{-5}$			
IMRD	NSDTR	00601	005011	02001	2.0	0.05	4.9	0.0025	
ANA+	NSDTR	00601	005011	02001	2.3	0.035	7.2	$<5 \times 10^{-4}$	
AD	Portuguese water dogs					0.00022-0.00079			Association with microsatellite markers
AD	NSDTR	01502	00601	02301	2.1	0.047	8.9	0.02	Homozygosity associated with early onset, elevated risk in dogs homozygous regardless of the haplotype
CLT	Doberman	01201	00101	00201	2.43	0.02			
CLT	GS	01201	GS		6.5	0.048			
CLT	GS	01301	00301	00501	0.3	0.017			Protective
IMHA-warm		00601	005011	00701					Significant association in breed-matched analysis
IMHA-cold		015	00601	00301					No significant association in breed-matched analysis !
CRA		002			3.5	0.03			Shared epitope
CRA					2.4	0.04			Association for the precense of any shared epitope bearing DRB1-allele
NME	Pugs	010011	00201	01501	1.13	<0.0001	12.75		Observed frequency in population over 75%
NME	Pugs	01502	00601	02301	0.13	<0.001	0		Protective
NME	Pugs	01501	00601	02301	0.11	<0.019			Protective
VKH-like syndrome	Akita		00201		15.99	0.013			
JGD	Mastiff, Mixed breed				5	<0.05			Association with microsatellite markers
Hepatitis	Doberman	00601	00401	01303			14.9	$<5 \times 10^{-5}$	
Hepatitis	Doberman	01501	00901	00101	15.8	0.001			Protective
SLO	Gordon Setter	01801	00101	00802	2.1	0.0004	5.4	0.0003	Homozygous dogs compared to dogs with no risk allele
SLO	Gordon Setter	02001	00401	01303	0.03	0.0001			Protective

CSK=chronic superficial keratitis, AF=anal furunculosis, IMRD=immune-mediated rheumatic disease, ANA=antinuclear antibody, AD=Addison's disease, CLT=canine lymphocytic thyroiditis, IMHA=immune mediated hemolytic anemia, CRA=canine rheumatic arthritis, NME=necrotizing meningoencephalitis, VHK-like syndrome=Vogt-Koyanaki-Harada-like syndrome, JGD=juvenile generalized demodicosis, SLO=symmetrical lupoid onychodystrophy, GSD=german shepherd, NSDTR=Nova Scotia Duck Tolling Retriever, GS=giant schnauzer.

The MHC class II risk haplotype for CSK in GSDs is the second most common haplotype in the breed. This genetic association with the DLA gene region supports the previous clinical, histological and pharmacological studies of CSK as an immune-mediated disease [20,41,87,89]. The third common haplotype has previously been associated with AF in GSDs and has no common alleles with the CSK risk haplotype. Table 14 displays the haplotype frequencies of 322 mostly European GSDs, and the risk haplotype is present in over 20% of the dogs. The frequency of the haplotype is over 30% in the Finnish study population, which reflects the study design, half of the dogs being affected with the associated disease. The risk haplotype has been found in 18 other breeds, but at a much lower frequency [106]. Any breed can be affected with CSK, but whether this risk haplotype contributes to the disease risk in any other breed has not been examined.

The DLA-DRB1*015/DQA1*006/DQB1*023 haplotype has previously been associated with canine diabetes in the Samoyed, Cairn terrier and Tibetan terrier [27]. The CSK risk haplotype differs from the diabetes haplotype at the DLA-DQB1 locus, although DLA-DRB1 was not characterized in as much detail as in our study. The three DLA-DRB1*015 alleles differ from each other by one nucleotide. The DLA-DRB1*015/DQA1*00601/DQB1*00301 haplotype has been associated with canine IMHA in English Springer Spaniels [30]. As in the diabetes study, the first allele was not characterized in detail and the association was lost when using breed-matched controls. The AD risk haplotype in NSDTRs has the same DLA-DQA1 allele, while there is a difference of one amino acid in the DLA-DRB1. In addition, the DRB1 allele in CSK risk haplotype is shared with the protective Doberman hepatitis haplotype, and together with the DQA1 allele with the protective NME haplotype in pugs.

6.3 Homozygosity of the MCH class II risk haplotype increases the risk for autoimmune diseases – mechanism?

Dogs carrying two copies of the risk haplotype were observed to have an elevated risk of developing the disease. In IMRD dogs, the risk was nearly fivefold, in ANA positive IMRD dogs over sevenfold, in AD dogs nearly ninefold and in CSK-affected dogs 8.5-fold. These ORs are among the highest risks reported for an autoimmune disease and MHC class II [54]. The risks calculated for AD and CSK dogs are estimates, since there were no control dogs homozygous for the risk haplotype. Low number of samples may also affect the OR estimate calculated by using pseudocount. Increased risk with homozygosity for risk haplotype has previously been shown with AF in GSDs and with necrotizing meningoencephalitis **(**NME**)** in pug dogs [25,108]. In addition, two recent studies have shown a stronger susceptibility for a homozygosity for a DLA risk haplotype, one with Doberman hepatitis [107] and the other with symmetrical lupoid onychodystrophy (SLO) in Gordon Setters [109]. In humans, a dose effect of HLA-DRB1*1501 has been observed in susceptibility of multiple sclerosis in two different studies and an HLA-DQB1*02 dose effect in coeliac disease [110-112]. Also, the HLA-DRB1 gene has been associated with susceptibility, severity and progression of disease in rheumatoid arthritis [113].

The mechanism of dose effect remains to be solved, but different theories have been proposed. As there is an extensive LD in the MHC region, other genes in LD within or

outside the MHC region might contribute to the disease risk. The MHC region is full of immunologically important genes, which might affect, for example the cytokine profile, thereby creating an advantageous environment for loss of peripheral tolerance.

A homozygous disadvantage is a disputed theory, as many believe that it is the presence rather than the number of an MHC allele that causes the effect. However, a homozygous disadvantage of MHC alleles has been observed in both human and canine autoimmune diseases. For example, human T1D susceptibility displays unusual patterns of inheritance. Protective MHC alleles have been suggested to express a dominant nature and susceptibility alleles a recessive inheritance. This difference may be due to the binding affinity of the MHC molecules for the β-cell antigens so that products of protective alleles bind very strongly, thus requiring fewer MHC molecules. MHC molecules coded by susceptibility alleles have a low affinity, requiring more molecules to effectively compete for binding [39]. The binding of the MHC molecules coded by the susceptibility alleles then leads to the activation of autoreactive T-cells that have escaped negative selection in the thymus, which again has been assisted by MHC molecules, and possible onset of autoimmune disease. The MHC-assisted T-cell variety could by definition be responsible for susceptibility to autoimmune disease together with the peptide presentation properties of MHC class II. Therefore, higher levels of surface expression cannot be ruled out as a causative factor. Increased MHC class II expression has been observed in CSK, but it could also be a secondary effect due to IFN-γ production and be related to the inflammation itself and be a part of the pathogenesis by elongating it rather than causing it. Elevated MHC class II expression has also been observed in endothelial cells in several human autoimmune diseases, including RA, SLE, multiple sclerosis and Crohn´s disease [114], and it has been shown to correlate with the severity of Doberman hepatitis [115].

The effect of MHC molecules are usually allele or haplotype specific and may be predisposing or protective. However, we discovered not only an elevated risk with homozygosity for the MHC class II risk haplotype in two of the studies, but also an increased risk with homozygous individuals regardless of the DLA haplotype. In our knowledge, this phenomenon has not previously been reported in dogs or humans. Even isolated human populations are much more heterozygous than dogs, and it is very unlikely that this kind of homozygosity effect could be shown in any human population. One possible mechanism to explain an increased risk in general DLA-haplotype homozygosity is extensive LD in the MHC region, and therefore, homozygosity might uncover recessive non-beneficial alleles. The higher risk for homozygousity for risk haplotype indicates that it has itself an impact on pathology, as overall homozygosity might reflect the influence of neighbouring genes. Overall homozygosity might also reflect the situation in other parts of the genome, as the MHC locus is the most polymorphic region in the genome. When there is no heterozygosity left in the DLA region, the other parts of the genome are hardly any different. Especially when an inbred population is in question, this is likely the situation and recessive alleles are uncovered.

Another mechanism might also be considered. As previously described in the context of the hygiene hypothesis, pathogens may mount up protective immune responses, and, as the MHC homozygous dogs are able to recognize fewer antigens, it is plausible to conclude that this protection is impaired. This would predispose the dogs not only to infections, but also to autoimmunity [62]. This raises a question, if maximum heterozygosity

in MHC region is optimal within an individual. If this was true, in the evolutionary point of view, we would probably see more duplicated genes in MHC region or maybe wider repertoire of differently expressed MHC peptides. One could think that maximal MCH diversity within individual would be beneficial in fighting infections, but could it also mediate an over-reactive immune response, harmful for the individual or increase the number of potentially self-reactive T-cells? Some T-cells may also be cross-reactive with self antigens and it may not be beneficial to have excess in activated T-cells. Increase in MHC diversity within individual might actually also affect the ability to fight infections, if an elevated number of expressed MHC molecules decrease the variety of T-cells by negative selection. In conclusion, it is likely that the present level of diversity is optimal, but the alleles and haplotypes providing best fitness for an individual vary through time and place, that is for example changes in the environmental risk factors. At a presence of certain pathogen, particular allele in homozygous form might be beneficial, but predispose to an autoimmune disease later in life.

6.4 The shared epitope in DLA-DRB1 allele is an indication of rheumatic autoimmune disease

The DLA-DRB1*00601 allele in the IMRD risk haplotype contains a five-amino-acid-long epitope RARAA at amino acid positions 70-74. These amino acids are part of the DNA segment that encodes the third hypervariable region (HVR-3). HVRs are mainly responsible for the peptide recognition and binding properties of the MHC class II molecule peptide binding groove. This group of amino acids is called RA shared epitope and it is found in all RA risk alleles both in humans and dogs and it is the most significant genetic risk factor for RA [31,114]. A similar epitope (QARAA) is found in the human SLE-associated HLA-DRB1*1501 allele [117]. The mechanistic basis of the shared epitope is unknown, but effects on arthritogenic antigen presentation and T-cell repertoire selection have been proposed. Dose effects of the shared epitope on penetrance and disease severity of RA have been shown [118].

Current evidence suggests that glutamine (Q) or arginine (R) at position 70 is critical for RA risk, and aspartic acid (D) at that position confers protection [118]. The epitope in the IMRD-associated DLA-DRB1 allele has Q at position 70, therefore fulfilling the criteria of being a similar predisposing variant as those increasing the risk for human RA and SLE. This leads further supports for a common MHC class II-associated mechanism of rheumatoid disease in both humans and dogs.

6.5 The first successful GWAS in complex diseases of dogs identifies several risk loci for autoimmune diseases in NSDTRs

This was a proof-of-principal study to show the power of the unique canine breed structure in GWAS of genetic risk loci in complex diseases. An extensive LD within a breed enables the mapping with approximately 15 000 markers and with less than 100

cases and 100 healthy controls. Several traits have been mapped in canine Mendelian diseases such as the hair ridge, which causes predisposition to dermoid sinus [12], recessive cone-rod dystrophy [13] and ectodermal dysplasia [14], but this was the first successful GWAS regarding polygenic disorders. We identified five loci associated with SLE-related disease using only 81 affected dogs and 57 healthy controls. Futher mapping with additional genetic markers (822) and in additional samples (n=487) and phenotypes, also from other breeds, verified the results and suggests common risk factors as well as disease-specific variants that influence the disease course in a systemic or organ-specific direction.

NSDTRs are strongly predisposed to several autoimmune diseases, such as SLE-related disease comprising well-characterized sub-phenotypes IMRD and SRMA, and AD [15,17,18]. The breed history of NSDTRs has had severe genetic bottlenecks, the ultimate being the canine distemper virus outbreaks in the early 1900s. Survival of only a few breeding individuals has likely accumulated rare recessive risk variants, and the surviving dogs may have had over-reactive immune responses capable of overcoming the repeated infections. As today's dogs descend from these survivors, it is clear that they carry the same genetic make-up.

We identified two loci that were associated with more than one phenotype with GWAM. The large CFA 32 region is shared between IMRD and SRMA, and it is plausible that common autoimmune predisposing loci containing risk factors leading to interruption of maintainance of self-tolerance exist. On the other hand, it is possible that the 1.6-Mb region contains several loci and variants contributing to different phenotypes. Three excellent candidate genes are found in the CFA 32 region: *DAPP1*, *PPP3CA* and *BANK1*. *BANK1* encodes a B-cell–specific scaffold protein and LYN tyrosine kinase substrate that promotes tyrosine phosphorylation of inositol 1,4,5-trisphosphate receptors. A non-synonymous substitution in the *BANK1* gene causes alternative splicing and has been associated with human SLE [104]. Ca(2+)/calmodulin-regulated protein phosphatise (calcineurin) is a heterodimer of a Ca(2+)-binding protein (calcineurin B) and a calmodulin-binding catalytic subunit (calcineurin A). There are several isoforms of the catalytic subunit, derived from alternative splicing of gene products of at least two genes, one of which is *PPP3CA* [118]. *PPP3CA* has been shown differential expression in human SLE patients compared with controls [119]. Calcineurin has also been reported to be the target of two important immunosuppressive drugs: cyclosporine A and FK506 [120]. The third of the top three candidates in the CFA 32 locus is *DAPP1* gene, expressed in both T- and B-cells, where it leads to an indirect dose-dependent inhibition of TCR- and BCR-induced activation of the NF-AT pathway [121,122].

The other locus shared with more than one phenotype was on CFA 8. Both SRMA and ANA-positive IMRD subphenotypes were associated separately, but not jointly with the locus. This might reflect the fact that there are two adjacent loci. When ANA positive dogs were analysed alone, the association was seen on the small nuclear ribonucleoprotein polypeptide E (*SNRPE)* gene, and when SRMA dogs were analysed alone, the association was between the genes *SNRPE* and vaccinia-related kinase 1 (*VRK1).* SNRPE is one of the small nuclear ribonucleoprotein complexes (snRNPs) recognized by circulating autoantibodies in human SLE. This protein is one of four 'core' proteins associated with all known snRNAs in the U family (U1, U2, U4, U5 and U6). *VRK1* is a novel serine-threonine kinase that regulates several transcription factors, including p53. The gene p53

has been shown to be one the most consistently under-expressed genes in autoimmune diseases, causing differential expression of other genes involved in autoimmune disease, and is therefore suggested to be central to autoimmunity [123].

ANA-positive dogs were associated with three more loci on CFA 3, 11 and 24 when analysed alone. The strongest association was found on CFA 11, which contains several genes, including protein-tyrosine phosphatise, non-receptor-type 3 (*PTPN3)*. *PTPN3* is involved in T-cell activation and belongs to the same family as protein-tyrosine phosphatase, non-receptor-type 22 (*PTPN22)*. *PTPN22 is* probably the second most common genetic risk factor after MHC for many autoimmune diseases in human, including SLE [124]. *PTPN3* has not been reported to be involved in autoimmunity, but it is thought to inhibit T-cell activation by de-phosphorylating targets involved in TCR signalling. The expression of *PTPN3* has been shown to reduce activation of reporter genes driven by NF-AT [125].

The second best association was found with CFA 24, containing six genes, and the third associated region on CFA 3 contains seven genes, *HOMER2* appearing the most interesting, as it has been reported to act as negative regulator of T-cell activation. In addition, Homer -deficient mice are shown to develop an autoimmune-like pathology with lymphocyte infiltration and hyperplasia in lymph nodes [126].

The ANA-positive associated loci on CFA 3, 11 and 24 showed higher p-values (10^{-11}–10^{-13}) after fine-mapping and odds ratios of 4.5–8. Both loci that were shared between IMRD and SRMA, CFA 8 and CFA 32, show *P*-values of 10^{-5}–10^{-8} and odds ratios of 2.4–3.4. According to power calculations, reliable detection of risk factors contributing a two- to fourfold increased risk cannot be expected with our sample size, but the loci with the highest odds ratios are within the detectable range. The associated regions contain relevant candidate genes based on biological function and are all worthy of follow-up studies.

The MHC locus previously associated with IMRD did not show an association here (Study III). We had three SNPs in GWAS in this region, from which two were totally homozygous in our NSDTR sample and were therefore not informative and the third showed very little heterogeneity. This region in general shows extensive polymorphism with for instance, 148 DLA-DRB1, 70 DLA-DQA1, and 26 DLA-DQB1 alleles, and it is therefore hard to tag with informative markers.

Based on our results, we hypothesize that IMRD and SRMA are a part of the same disease complex and have common as well as sub-phenotype specific risk factors. Therefore we examined the affected dogs both as one group and the two sub-phenotypes separately. Shared loci with AD suggest a common locus predisposing to autoimmunity.

6.6 New immunological pathway in SLE

Three of the associated loci we found contain four genes that regulate the nuclear factor of the activated T-cell (NF-AT) signaling pathway. Antigen-specific immune responses are initiated by the interaction of the TCR with antigenic peptide bound to MHC proteins on the surface of antigen-presenting cells. The shared locus on CFA 32 contains two genes, *PPP3CA* and *DAPP1,* involved in the NF-AT pathway. *PPP3CA* encodes the catalytic

subunit of calcineurin. Calcineurin is the downstream target of intracellular Ca^{2+} signalling that follows the T-cell-MHC contact. Also increased translocation of calcineurin-dependent transcription factor **(**NF-ATc2**)** to the nucleus in the early stages of cell activation has been observed in human SLE patients [127]. In addition to their role in T-cell activation, NFATc transcription factors also generate peripheral tolerance against self-antigens mediated by controlling activation-induced cell death and clonal anergy of T-helper cells and the activity of regulatory T-cells [128]. *DAPP1* is expressed in both T- and B-cells, where it leads to an indirect dose-dependent inhibition of TCR- and B-cell receptor (BCR) -induced activation of NF-AT [121,122]. The *PTPN3* gene on ANA-positive CFA 11 is thought to inhibit this pathway as well, although it has not been associated with autoimmunity [125]. The third inhibitor, scaffold protein competing for binding to NF-AT with calcineurin, is found in the ANA-positive locus CFA 3.

A typical feature of SLE is clinical heterogeneity, resulting in variable disease manifestations. When we started in 2007, nine human lupus susceptibility genes had been convincingly identified, compared with the more than 30 convincing genetic associations known today [129]. Both adaptive and innate immune systems are involved, producing different subphenotypes. Our results indicate strongly involvement of T-cell activation. We hypothesize that predisposition to autoimmune diseases in the modern NSDTR may be a result of the early NSDTRs' ability to survive the outbreaks of canine distemper virus in the early 1900s. In addition, modern breeding practices may have accumulated certain risk factors. All in all, it is clear that there are several predisposing genes or regulatory elements that together contribute to development of these autoimmune diseases. The causing variants remain to be identified, as does the way they interact and accumulate the increasing risk. There may also be other variants with a smaller risk that have escaped our study due to the small sample size. In the future, the development of individual drug treatments, *e.g*. cyclosporine plus corticosteroids based on a dog's particular risk genotype, as well as diagnostic tools and gene tests to aid in planning of breeding strategies might be possible.

6.7 The dog is an excellent model for complex genetic studies

Dogs have been used as models to understand many diseases, as in epilepsy research by cloning the first canine epilepsy gene, *Epm2b* [130]. Mutations in the same gene were found in humans to cause a fatal form of epilepsy, lafora disease. The dog has also been used as a model to develop therapies before they are tested in humans, as is being done in the canine model of Duchenne muscular dystrophy in golden retrievers [131].

Results of our successful GWAS of complex diseases in dogs proves the value of the domestic dog as a disease model for polygenic disorders. Predisposing high-risk mutations have accumulated and become relatively common in dog breeds due to several historical bottlenecks. While tens of thousands of samples may be needed to map human diseases with rare, low-risk variants due to clinical and genetic heterogeneity, this can now be performed in a domestic dog with less than 100 cases and 100 controls. Ancient mutations originating from ancient dog have been spread in several breeds after breed creation.

These mutations are still surrounded by very similar DNA segments, and advantage can be taken in GWASs of these shared haplotypes surrounding the causative variants.

Dogs are exposed to the same environmental risk factors and often even share the same diet with humans. Therefore dog is an excellent disease model for autoimmune diseases common in both species and the findings may help to understand both common biological functions and interactions between genes and environment. Identification of new genes and pathways involved in canine autoimmune diseases increases our understanding of the pathogenesis of immunological disorders both in human and dog. This may open doors for novel diagnostics, gene tests, treatment and drug development in both species. In dogs, identification of mutations will help to reduce the incidence of the disease in the breed. Complex diseases are not easy to eradicate from the breed, but identification of risk variants at least provides more tools to accomplish this goal. Autoimmune disorders are already so common in some dog breeds that they threaten the existence of the whole breed.

7 Conclusions and future perspectives

When I began my work in autumn 2006, I was full of enthusiasm and started literally from a clean slate. The canine genetics group was new and I was the first to be employed. I am a passionate dog lover, so it was a dream come true to be able to work towards improving dogs' health, not to mention the potential benefits for human welfare.

We found the first genetic evidence for canine SLE-related disease, AD and CSK. The MHC class II was a strong candidate locus based on both human and dog studies, and we were able to confirm an association with CSK, AD and IMRD phenotypes. Our results suggest in each of these phenotypes that the DLA-DRB1, -DQB1, -DQA haplotypes are a major genetic risk factor, but that they are all polygenic disorders with possibly several other genes contributing to the development of the disease.

The development of technology has been rapid in these four years, with all the new dog arrays and high-throughput sequencing technologies becoming available. We performed the first successful GWAS in canine complex diseases and identified five loci that predispose to a SLE-related disease and AD in NSDTRs. Many of these loci are involved in a novel T-cell activation pathway. This study proves the strength of disease mapping in dogs. The power of the two-stage strategy is based on extensive LD and long haplotypes within a breed, and short haplotypes between dog breeds. In addition, a very small number of founders during breed creation and subsequent inbreeding have accumulated ancient mutations and diseases. Causative mutations may be detected with the new genotyping technologies and information provided by the full genome sequence.

In addition, we have observed a narrow genetic diversity in NSDTRs and GSDs, and a clear increase in disease risk in dogs homozygous for the risk haplotypes. Surprisingly, we found that overall homozygosity for the MHC class II region is associated with increased risk in CSK and AD. The mechanisms behind the involvement of homozygosity in the DLA locus remains unknown, but may reflect either the T-cell repertoire or antigen presentation or be a consequence of other predisposing genes in LD within the DLA region.

Our results provide tools and information to assist breeding practices and the fight against diseases in these breeds as well as important knowledge of the inheritance. Hopefully, our future work will provide more accurate knowledge of the pathogenesis by detecting the exact causative variants. These findings would open doors for new diagnostics, gene tests and drug development in both dogs and humans.

Our future plans include identification of additional risk factors for CSK in GSDs by GWAS. We already have preliminary results connecting the aetiology of AF, EPI and CSK based on an association with common SNPs and haplotypes in these three autoimmune diseases. As in NSDTRs, it seems that there are indeed common risk factors for autoimmune diseases, as well as disease-specific risk factors. We are also interested in examining related breeds to see whether the predisposing MHC haplotype is an universal risk factor or specific only to CSK in GSDs.

Future plans for identification of the disease-causing genetic variants in NSDTRs and in breeds sharing the associated haplotypes include a hybrid capture and targeted resequencing of the associated haplotypes. The full sequence data will provide thousands of variants and potential candidate mutations will be prioritized for further evaluation

based on two criteria: presence on haplotypes shared by individuals of multiple breeds and localization within the conserved elements.

Acknowledgements

The work for this thesis was carried out during 2007-2010 at Department of Veterinary Biosciences in the Faculty of Veterinary Medicine, the Department of Medical Genetics in Haartman Institute, and the Program in Molecular Medicine in the Faculty of Medicine, University of Helsinki and also in the Folkhälsan Institute of Genetics, at the Department of Molecular Genetics, Biomedicum I, Helsinki. I wish to express my gratitude to the numerous people who participated in this project.

For providing excellent research facilities, I thank the previous heads of the Program in Molecular Medicine, Academy Professor Leena Peltonen-Palotie and Professor Kimmo Kontula and the present head, Professor Hannes Lohi, the head of the Folkhälsan Institute of Genetics, Professor Anna-Elina Lehesjoki, the head of the Department of Veterinary Biosciences, Professor Airi Palva, and the head of the Department of Medical Genetics Professor Irma Järvelä. Professor Antti Sukura is warmly thanked for accepting the important post of custos at the dissertation.

My sincere gratitude is owed to my supervisor, Professor Hannes Lohi, for giving me the opportunity to work in his group on a subject close to my heart. I have rarely seen a group leader so enthusiastic about his work, thank you for giving me the opportunity to apply state-of-art research methods in the study of canine genetics.

I am grateful to the official reviewers, Docents Hanna Jarva and Professor Johanna Schleutker, for their constructive and critical evaluation of this thesis. The manuscript submitted for preliminary examination was revised by Carol Ann Pelli, Hon.B.Sc. and I thank her for skilful language revision.

I owe my deepest gratitude to the co-authors of my papers, especially Maria Wilbe, MSc and Angela Hughes, DVM. We have had inspiring conversations and lots of fun during conference meetings. My warmest thanks go to Lorna Kennedy, PhD, who taught me sequence analysis and whose help has been indispensable in several projects over the years. You have always kindly helped with every problem I have had. You are a true friend. For scientific collaboration and fruitful conversations, I thank Professor Kerstin Lindblad-Toh, Docent Helene Hansson-Hamlin, Elinor Karlsson, PhD and Professor Göran Andersson for scientific collaboration and inspiring conversations. The contribution of other co-authors, namely Professor Anita Oberbauer, MSc Katarina Truvé, MSc Christina Hermanrud, Professor Erling Strandberg, Tara Biagi, B.C. and Assistant Professor Danika Bannasch, is warmly acknowledged.

My gratitude is due to the staff of the University Teaching Hospital, University of Helsinki, and especially to Elina Rusanen who performed the ophtalmic examinations on the dogs participating in our study. It has been educational working with you and despite the distance and the haste also very pleasant.

A very special thanks is owed to my co-workers and friends in the canine genetics groups: Eija Seppälä, Ranja Eklund, Minna Virta, Kaisa Kyöstilä, Saija Ahonen, Lotta Koskinen, Elisa Nevalainen, Sini Karjalainen, Anna But, Anja Ruusulehto, Marjo Hytönen, Saija Lampinen and Hanna Dyggve. The team spirit has been exceptional and I consider all of you more like friends than co-workers. Thank you for support, help, advice and enjoyable coffee and lunch breaks, not to mention the invigorating retreats with our dogs. Eija, I especially thank you for all the help and support over the years. You have

taught me much about the work of a scientist. I highly appreciate your expertise, encouragement, help and friendship. Anja, my warmest thanks to you also for giving valuable comments on my thesis and for offering help whenever needed.

All co-workers at the Folkhälsan Institute of Genetics, Program in Molecular Medicine, Department of Molecular Genetics and Department of Veterinary Biosciences are thanked for smooth co-operation and for providing an inspiring working environment. Special thanks go to my former supervisor Maija Wessmann for generous support and encouragement over the years and for valuable and constructive feedback while finishing this dissertation. Päivi Tikka-Kleemola, it was nice sharing an office with you and Maija again, even if only part-time.

I also wish to thank everyone at the Institute for Molecular Medicine Finland FIMM Technology Center and especially Pekka Ellonen in the DNA Sequencing Laboratory, who always patiently provided help and education.

I thank all participating dog owners as well as the Finnish and Swedish breed clubs of Nova Scotia Duck Tolling Retrievers and the Finnish breed of German Shepherd dogs for their invaluable contributions.

My warmest thanks to my friends and relatives for endlessly supporting and encouraging me! I would not be here without my wonderful support group. You have balanced my life and have always been there when needed.

Rakkaimmat kiitokset vanhemmilleni Eilalle ja Ilmarille avustanne ja myötäelämisestä hyvinä ja huonoina päivinä. Olen aina voinut luottaa teihin asiassa kuin asiassa ja olen onnekkain ihminen maan päällä omatessani teidänlaisenne vanhemmat. Yhtälailla rakkaat kiitokset pikkusiskolleni Merville, tädilleni Maisalle ja enolleni Joukolle vilpittömästä tuestanne, joka on auttanut minua ponnistelemaan eteenpäin niin monessa asiassa. Santulle, omalle ihmiselleni, lämmin halaus ja suukko sekä lämpimimmät ja rakkaimmat kiitokset, että olet ihmeellisellä tavalla jaksanut jakaa ilot ja surut kanssani vuosien mittaan. Olen aina voinut luottaa sinuun kaikessa ja tukesi on ollut täysin korvaamatonta.

This study was financially supported by the grants awarded to the supervisor from the Sigrid Juselius Foundation, Biocentrum Helsinki, the Academy of Finland, the Jane and Aatos Erkko Foundation and the University of Helsinki Research Funds. Parts of this work were funded by the grants of Dr. Leena Peltonen and the Center of Excellence of Complex Disease Genetics of the Academy of Finland and the “väitöskirjan loppuunsaattamisraha” by the University of Helsinki. All support is gratefully acknowledged.

References

1. Janeway C.A., Travers P., Walport M., Shlomchik M. **Immunobiology**Garland Science Publishing, 2005.

2. Todd JA, Bell JI, McDevitt HO.HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 1987; 329:599-604.

3. Nicholas FW. Online mendelian inheritance in animals (OMIA): A comparative knowledgebase of genetic disorders and other familial traits in non-laboratory animals. Nucleic Acids Res 2003; 31:275-7.

4. Ostrander EA, Kruglyak L.Unleashing the canine genome. Genome Res 2000; 10:1271-4.

5. Sargan DR. IDID: Inherited diseases in dogs: Web-based information for canine inherited disease genetics. Mamm Genome 2004; 15:503-6.

6. Kirkness EF, Bafna V, Halpern AL, et al.The dog genome: Survey sequencing and comparative analysis. Science 2003; 301:1898-903.

7. Kristiansson K, Naukkarinen J, Peltonen L.Isolated populations and complex disease gene identification. Genome Biol 2008; 9:109.

8. Sutter NB, Ostrander EA.Dog star rising: The canine genetic system. Nat Rev Genet 2004; 5:900-10.

9. Lindblad-Toh K, Wade CM, Mikkelsen TS, et al.Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 2005; 438:803-19.

10. Sutter NB, Eberle MA, Parker HG, Pullar BJ, Kirkness EF, Kruglyak L, Ostrander EA.Extensive and breed-specific linkage disequilibrium in canis familiaris. Genome Res 2004; 14:2388-96.

11. Karlsson EK, Baranowska I, Wade CM, et al.Efficient mapping of mendelian traits in dogs through genome-wide association. Nat Genet 2007; 39:1321-8.

12. Salmon Hillbertz NH, Isaksson M, Karlsson EK, et al.Duplication of FGF3, FGF4, FGF19 and ORAOV1 causes hair ridge and predisposition to dermoid sinus in ridgeback dogs. Nat Genet 2007; 39:1318-20.

13. Wiik AC, Wade C, Biagi T, Ropstad EO, Bjerkas E, Lindblad-Toh K, Lingaas F.A deletion in nephronophthisis 4 (NPHP4) is associated with recessive cone-rod dystrophy in standard wire-haired dachshund. Genome Res 2008; 18:1415-21.

14. Drogemuller C, Karlsson EK, Hytonen MK, et al.A mutation in hairless dogs implicates FOXI3 in ectodermal development. Science 2008; 321:1462.

15. Hansson-Hamlin H, Lilliehook I.A possible systemic rheumatic disorder in the nova scotia duck tolling retriever. Acta Vet Scand 2009; 51:16.

16. Redman J. Steroid-responsive meningitis-arteritis in the nova scotia duck tolling retriever. Vet Rec 2002; 151:712.

17. Anfinsen KP, Berendt M, Liste FJ, Haagensen TR, Indrebo A, Lingaas F, Stigen O, Alban L.A retrospective epidemiological study of clinical signs and familial predisposition associated with aseptic meningitis in the norwegian population of nova scotia duck tolling retrievers born 1994-2003. Can J Vet Res 2008; 72:350-5.

18. Burton S, DeLay J, Holmes A, Somerville C, Eye J, Shaw D, Wack O, Hanna P.Hypoadrenocorticism in young related nova scotia duck tolling retrievers. Can Vet J 1997; 38:231-4.

19. Lucke VM, Gaskell CJ, Wotton PR.Thyroid pathology in canine hypothyroidism. J Comp Pathol 1983; 93:415-21.

20. Bedford PG, Longstaffe JA.Corneal pannus (chronic superficial keratitis) in the german shepherd dog. J Small Anim Pract 1979; 20:41-56.

21. Tobin DJ, Gardner SH, Luther PB, Dunston SM, Lindsey NJ, Olivry T.A natural canine homologue of alopecia areata in humans. Br J Dermatol 2003; 149:938-50.

22. Hubert B, Teichner M, Fournel C, Monier JC.Spontaneous familial systemic lupus erythematosus in a canine breeding colony. J Comp Pathol 1988; 98:81-9.

23. Westermarck E. The hereditary nature of canine pancreatic degenerative atrophy in the german shepherd dog. Acta Vet Scand 1980; 21:389-94.

24. Day MJ, Weaver BMQ.Pathology of surgically resected tissue from 305 cases of anal furunculosis in the dog. J Small Anim Pract 1992; 33:583-9.

25. Kennedy LJ, O'Neill T, House A, et al.Risk of anal furunculosis in german shepherd dogs is associated with the major histocompatibility complex. Tissue Antigens 2008; 71:51-6.

26. Shelton GD, Willard MD, Cardinet GH,3rd, Lindstrom J.Acquired myasthenia gravis. selective involvement of esophageal, pharyngeal, and facial muscles. J Vet Intern Med 1990; 4:281-4.

27. Kennedy LJ, Davison LJ, Barnes A, et al.Identification of susceptibility and protective major histocompatibility complex haplotypes in canine diabetes mellitus. Tissue Antigens 2006; 68:467-76.

28. Kennedy LJ, Huson HJ, Leonard J, Angles JM, Fox LE, Wojciechowski JW, Yuncker C, Happ GM.Association of hypothyroid disease in doberman pinscher dogs with a rare major histocompatibility complex DLA class II haplotype. Tissue Antigens 2006; 67:53-6.

29. Kennedy LJ, Quarmby S, Happ GM, et al.Association of canine hypothyroidism with a common major histocompatibility complex DLA class II allele. Tissue Antigens 2006; 68:82-6.

30. Kennedy LJ, Barnes A, Ollier WE, Day MJ.Association of a common dog leucocyte antigen class II haplotype with canine primary immune-mediated haemolytic anaemia. Tissue Antigens 2006; 68:502-8.

31. Ollier WE, Kennedy LJ, Thomson W, et al.Dog MHC alleles containing the human RA shared epitope confer susceptibility to canine rheumatoid arthritis. Immunogenetics 2001; 53:669-73.

32. Kennedy MA. A brief review of the basics of immunology: The innate and adaptive response. Vet Clin North Am Small Anim Pract 2010; 40:369-79.

33. Sun JC, Lanier LL.Natural killer cells remember: An evolutionary bridge between innate and adaptive immunity? Eur J Immunol 2009; 39:2059-64.

34. Vignali DA, Collison LW, Workman CJ.How regulatory T cells work. Nat Rev Immunol 2008; 8:523-32.

35. Mosmann TR, Coffman RL.TH1 and TH2 cells: Different patterns of lymphokine secretion lead to different functional properties. Annu Rev Immunol 1989; 7:145-73.

36. Niederkorn JY. See no evil, hear no evil, do no evil: The lessons of immune privilege. Nat Immunol 2006; 7:354-9.

37. Biros D. Anterior chamber-associated immune deviation. Vet Clin North Am Small Anim Pract 2008; 38:309,21, vi-vii.

38. Gershwin LJ. Autoimmune diseases in small animals. Vet Clin North Am Small Anim Pract 2010; 40:439-57.

39. Ballotti S, Chiarelli F, de Martino M.Autoimmunity: Basic mechanisms and implications in endocrine diseases. part I. Horm Res 2006; 66:132-41.

40. Jager A, Kuchroo VK.Effector and regulatory T-cell subsets in autoimmunity and tissue inflammation. Scand J Immunol 2010; 72:173-84.

41. Williams DL. Major histocompatibility class II expression in the normal canine cornea and in canine chronic superficial keratitis. Vet Ophthalmol 2005; 8:395-400.

42. Ahonen P, Myllarniemi S, Sipila I, Perheentupa J.Clinical variation of autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) in a series of 68 patients. N Engl J Med 1990; 322:1829-36.

43. Aaltonen J, Bjorses P, Sandkuijl L, Perheentupa J, Peltonen L.An autosomal locus causing autoimmune disease: Autoimmune polyglandular disease type I assigned to chromosome 21. Nat Genet 1994; 8:83-7.

44. Karlsson EK, Lindblad-Toh K.Leader of the pack: Gene mapping in dogs and other model organisms. Nat Rev Genet 2008; 9:713-25.

45. Chen WK, Swartz JD, Rush LJ, Alvarez CE.Mapping DNA structural variation in dogs. Genome Res 2009; 19:500-9.

46. Fanciulli M, Norsworthy PJ, Petretto E, et al.FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. Nat Genet 2007; 39:721-3.

47. Reymond A, Henrichsen CN, Harewood L, Merla G.Side effects of genome structural changes. Curr Opin Genet Dev 2007; 17:381-6.

48. Stankiewicz P, Lupski JR.Structural variation in the human genome and its role in disease. Annu Rev Med 2010; 61:437-55.

49. Fanciulli M, Petretto E, Aitman TJ.Gene copy number variation and common human disease. Clin Genet 2009; .

50. Gibney ER, Nolan CM.Epigenetics and gene expression. Heredity 2010; 105:4-13.

51. Jungel A, Ospelt C, Gay S.What can we learn from epigenetics in the year 2009? Curr Opin Rheumatol 2010; 22:284-92.

52. Yuhki N, Beck T, Stephens R, Neelam B, O'Brien SJ.Comparative genomic structure of human, dog, and cat MHC: HLA, DLA, and FLA. J Hered 2007; 98:390-9.

53. Debenham SL, Hart EA, Ashurst JL, Howe KL, Quail MA, Ollier WE, Binns MM.Genomic sequence of the class II region of the canine MHC: Comparison with the MHC of other mammalian species. Genomics 2005; 85:48-59.

54. Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM, Vyse TJ, Rioux JD.Defining the role of the MHC in autoimmunity: A review and pooled analysis. PLoS Genet 2008; 4:e1000024.

55. It V, Barrientos L, Lopez Gappa J, Posik D, Diaz S, Golijow C, Giovambattista G.Association of canine juvenile generalized demodicosis with the dog leukocyte antigen system. Tissue Antigens 2010; .

56. Kennedy LJ, Barnes A, Happ GM, Quinnell RJ, Courtenay O, Carter SD, Ollier WE, Thomson W.Evidence for extensive DLA polymorphism in different dog populations. Tissue Antigens 2002; 60:43-52.

57. Hirayama K, Matsushita S, Kikuchi I, Iuchi M, Ohta N, Sasazuki T.HLA-DQ is epistatic to HLA-DR in controlling the immune response to schistosomal antigen in humans. Nature 1987; 327:426-30.

58. Bernatchez L, Landry C.MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? J Evol Biol 2003; 16:363-77.

59. Bollati V, Baccarelli A.Environmental epigenetics. Heredity 2010; 105:105-12.

60. Kondrashova A, Reunanen A, Romanov A, et al.A six-fold gradient in the incidence of type 1 diabetes at the eastern border of finland. Ann Med 2005; 37:67-72.

61. Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. N Engl J Med 2002; 347:911-20.

62. Okada H, Kuhn C, Feillet H, Bach JF.The 'hygiene hypothesis' for autoimmune and allergic diseases: An update. Clin Exp Immunol 2010; 160:1-9.

63. Reprogen, Faculty of Veterinary Science, University of Sydney. Online mendelian inheritance in animals, OMIA. 2010:.

64. Summers JF, Diesel G, Asher L, McGreevy PD, Collins LM.Inherited defects in pedigree dogs. part 2: Disorders that are not related to breed standards. Vet J 2010; 183:39-45.

65. Ferm K, Bjornerfeldt S, Karlsson A, Andersson G, Nachreiner R, Hedhammar A.Prevalence of diagnostic characteristics indicating canine autoimmune lymphocytic thyroiditis in giant schnauzer and hovawart dogs. J Small Anim Pract 2009; 50:176-9.

66. Koskenmies S, Widen E, Kere J, Julkunen H.Familial systemic lupus erythematosus in finland. J Rheumatol 2001; 28:758-60.

67. Hochberg MC. Updating the american college of rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997; 40:1725.

68. Danchenko N, Satia JA, Anthony MS.Epidemiology of systemic lupus erythematosus: A comparison of worldwide disease burden. Lupus 2006; 15:308-18.

69. Tebbe B, Orfanos CE.Epidemiology and socioeconomic impact of skin disease in lupus erythematosus. Lupus 1997; 6:96-104.

70. Antal Z, Zhou P.Addison disease. Pediatr Rev 2009; 30:491-3.

71. Halonen M, Eskelin P, Myhre AG, et al.AIRE mutations and human leukocyte antigen genotypes as determinants of the autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy phenotype. J Clin Endocrinol Metab 2002; 87:2568-74.

72. Martin Martorell P, Roep BO, Smit JW.Autoimmunity in addison's disease. Neth J Med 2002; 60:269-75.

73. Davey MP, Rosenbaum JT.The human leukocyte antigen complex and chronic ocular inflammatory disorders. Am J Ophthalmol 2000; 129:235-43.

74. Andrew SE. Immune-mediated canine and feline keratitis. Vet Clin North Am Small Anim Pract 2008; 38:269,90, vi.

75. Caspi RR. Ocular autoimmunity: The price of privilege? Immunol Rev 2006; 213:23-35.

76. Koskenmies S, Jarvinen TM, Onkamo P, Panelius J, Tuovinen U, Hasan T, Ranki A, Saarialho-Kere U.Clinical and laboratory characteristics of finnish lupus erythematosus patients with cutaneous manifestations. Lupus 2008; 17:337-47.

77. Hansson-Hamlin H, Lilliehook I, Trowald-Wigh G.Subgroups of canine antinuclear antibodies in relation to laboratory and clinical findings in immune-mediated disease. Vet Clin Pathol 2006; 35:397-404.

78. Stull JW, Evason M, Carr AP, Waldner C.Canine immune-mediated polyarthritis: Clinical and laboratory findings in 83 cases in western canada (1991-2001). Can Vet J 2008; 49:1195-203.

79. Lifton SJ, King LG, Zerbe CA.Glucocorticoid deficient hypoadrenocorticism in dogs: 18 cases (1986-1995). J Am Vet Med Assoc 1996; 209:2076-81.

80. Kintzer PP, Peterson ME.Treatment and long-term follow-up of 205 dogs with hypoadrenocorticism. J Vet Intern Med 1997; 11:43-9.

81. Peterson ME, Kintzer PP, Kass PH.Pretreatment clinical and laboratory findings in dogs with hypoadrenocorticism: 225 cases (1979-1993). J Am Vet Med Assoc 1996; 208:85-91.

82. Nell B, Walde I, Billich A, Vit P, Meingassner JG.The effect of topical pimecrolimus on keratoconjunctivitis sicca and chronic superficial keratitis in dogs: Results from an exploratory study. Vet Ophthalmol 2005; 8:39-46.

83. Chavkin MJ, Roberts SM, Salman MD, Severin GA, Scholten NJ.Risk factors for development of chronic superficial keratitis in dogs. J Am Vet Med Assoc 1994; 204:1630-4.

84. Slatter DH, Lavach JD, Severin GA, Young S.Uberreiter's syndrome (chronic superficial keratitis) in dogs in the rocky mountain area--a study of 463 cases. J Small Anim Pract 1977; 18:757-72.

85. Stanley RG. Superficial stromal keratitis in the dog. Aust Vet J 1988; 65:321-3.

86. Überreiter O. Conjunctivo-keratitis superficialis chronic beim hunde und ihre operative behandlung. Wien. Tierärztl. Mschr. 1959; 48:.

87. Williams DL. Histological and immunohistochemical evaluation of canine chronic superficial keratitis. Res Vet Sci 1999; 67:191-5.

88. de Smet MD, Chan CC.Regulation of ocular inflammation--what experimental and human studies have taught us. Prog Retin Eye Res 2001; 20:761-97.

89. Williams DL, Hoey AJ, Smitherman P.Comparison of topical cyclosporin and dexamethasone for the treatment of chronic superficial keratitis in dogs. Vet Rec 1995; 137:635-9.

90. Pang JF, Kluetsch C, Zou XJ, et al.mtDNA data indicate a single origin for dogs south of yangtze river, less than 16,300 years ago, from numerous wolves. Mol Biol Evol 2009; 26:2849-64.

91. Gray MM, Granka JM, Bustamante CD, Sutter NB, Boyko AR, Zhu L, Ostrander EA, Wayne RK.Linkage disequilibrium and demographic history of wild and domestic canids. Genetics 2009; 181:1493-505.

92. Rudan I, Campbell H.Five reasons why inbreeding may have considerable effect on post-reproductive human health. Coll Antropol 2004; 28:943-50.

93. Acevedo-Whitehouse K, Gulland F, Greig D, Amos W.Inbreeding: Disease susceptibility in california sea lions. Nature 2003; 422:35.

94. Strang A. and MacMillan G. The nova scotia duck tolling retrieverAlpine Publications, 1996.

95. Saksanpaimenkoiraliitto. Saksanpaimenkoiran jalostuksen tavoiteohjelma 2009-2013. 2010; .

96. Verein für Deutche Schäferhunde. Verein für deutche schäferhunde, home page.

97. Parker HG, Kim LV, Sutter NB, et al.Genetic structure of the purebred domestic dog. Science 2004; 304:1160-4.

98. Starkey MP, Scase TJ, Mellersh CS, Murphy S.Dogs really are man's best friend--canine genomics has applications in veterinary and human medicine! Brief Funct Genomic Proteomic 2005; 4:112-28.

99. Teo YY, Inouye M, Small KS, Gwilliam R, Deloukas P, Kwiatkowski DP, Clark TG.A genotype calling algorithm for the illumina BeadArray platform. Bioinformatics 2007; 23:2741-6.

100. Durbin R, Eddy S, Krogh A, Mitchison G, eds. Biological sequence analysis, probabilistic models of protein and nucleic acids. Cambridge University Press, The Edinburgh building, Cambridge: PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE, 1998.

101. Purcell S, Neale B, Todd-Brown K, et al.PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81:559-75.

102. Mantel N, HAENSZEL W.Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959; 22:719-48.

103. Wilbe M, Jokinen P, Truve K, et al.Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex. Nat Genet 2010; .

104. Kozyrev SV, Abelson AK, Wojcik J, et al.Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet 2008; 40:211-6.

105. Kennedy LJ, Barnes A, Happ GM, et al.Extensive interbreed, but minimal intrabreed, variation of DLA class II alleles and haplotypes in dogs. Tissue Antigens 2002; 59:194-204.

106. Kennedy LJ, Barnes A, Short A, et al.Canine DLA diversity: 1. new alleles and haplotypes. Tissue Antigens 2007; 69 Suppl 1:272-88.

107. Dyggve H, Kennedy LJ, Meri S, Spillmann T, Lohi H, Speeti M.Association of doberman hepatitis to canine major histocompatibility complex II. Tissue Antigens 2010; .

108. Greer KA, Wong AK, Liu H, Famula TR, Pedersen NC, Ruhe A, Wallace M, Neff MW.Necrotizing meningoencephalitis of pug dogs associates with dog leukocyte antigen class II and resembles acute variant forms of multiple sclerosis. Tissue Antigens 2010; .

109. Wilbe M, Ziener ML, Aronsson A, et al.DLA class II alleles are associated with risk for canine symmetrical lupoid onychodystropy (SLO). PLoS One 2010; 5:e12332.

110. Barcellos LF, Oksenberg JR, Begovich AB, et al.HLA-DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. Am J Hum Genet 2003; 72:710-6.

111. Barcellos LF, Sawcer S, Ramsay PP, et al.Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Hum Mol Genet 2006; 15:2813-24.

112. Nenna R, Mora B, Megiorni F, Mazzilli MC, Magliocca FM, Tiberti C, Bonamico M.HLA-DQB1*02 dose effect on RIA anti-tissue transglutaminase autoantibody levels and clinicopathological expressivity of celiac disease. J Pediatr Gastroenterol Nutr 2008; 47:288-92.

113. Gonzalez-Gay MA, Garcia-Porrua C, Hajeer AH.Influence of human leukocyte antigen-DRB1 on the susceptibility and severity of rheumatoid arthritis. Semin Arthritis Rheum 2002; 31:355-60.

114. Turesson C. Endothelial expression of MHC class II molecules in autoimmune disease. Curr Pharm Des 2004; 10:129-43.

115. Speeti M, Stahls A, Meri S, Westermarck E.Upregulation of major histocompatibility complex class II antigens in hepatocytes in doberman hepatitis. Vet Immunol Immunopathol 2003; 96:1-12.

116. Tsuchiya N, Kawasaki A, Tsao BP, Komata T, Grossman JM, Tokunaga K.Analysis of the association of HLA-DRB1, TNFalpha promoter and TNFR2 (TNFRSF1B)

polymorphisms with SLE using transmission disequilibrium test. Genes Immun 2001; 2:317-22.

117. Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. Curr Opin Rheumatol 2010; 22:293-8.

118. Wang MG, Yi H, Guerini D, Klee CB, McBride OW.Calcineurin A alpha (PPP3CA), calcineurin A beta (PPP3CB) and calcineurin B (PPP3R1) are located on human chromosomes 4, 10q21-->q22 and 2p16-->p15 respectively. Cytogenet Cell Genet 1996; 72:236-41.

119. Olsen NJ, Moore JH, Aune TM.Gene expression signatures for autoimmune disease in peripheral blood mononuclear cells. Arthritis Res Ther 2004; 6:120-8.

120. Guerini D. Calcineurin: Not just a simple protein phosphatase. Biochem Biophys Res Commun 1997; 235:271-5.

121. Marshall AJ, Niiro H, Lerner CG, Yun TJ, Thomas S, Disteche CM, Clark EA.A novel B lymphocyte-associated adaptor protein, Bam32, regulates antigen receptor signaling downstream of phosphatidylinositol 3-kinase. J Exp Med 2000; 191:1319-32.

122. Sommers CL, Gurson JM, Surana R, et al.Bam32: A novel mediator of erk activation in T cells. Int Immunol 2008; 20:811-8.

123. Olsen N, Sokka T, Seehorn CL, Kraft B, Maas K, Moore J, Aune TM.A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. Ann Rheum Dis 2004; 63:1387-92.

124. Chung SA, Criswell LA.PTPN22: Its role in SLE and autoimmunity. Autoimmunity 2007; 40:582-90.

125. Han S, Williams S, Mustelin T.Cytoskeletal protein tyrosine phosphatase PTPH1 reduces T cell antigen receptor signaling. Eur J Immunol 2000; 30:1318-25.

126. Huang GN, Huso DL, Bouyain S, et al.NFAT binding and regulation of T cell activation by the cytoplasmic scaffolding homer proteins. Science 2008; 319:476-81.

127. Kyttaris VC, Wang Y, Juang YT, Weinstein A, Tsokos GC.Increased levels of NF-ATc2 differentially regulate CD154 and IL-2 genes in T cells from patients with systemic lupus erythematosus. J Immunol 2007; 178:1960-6.

128. Serfling E, Klein-Hessling S, Palmetshofer A, Bopp T, Stassen M, Schmitt E.NFAT transcription factors in control of peripheral T cell tolerance. Eur J Immunol 2006; 36:2837-43.

129. Moser KL, Kelly JA, Lessard CJ, Harley JB.Recent insights into the genetic basis of systemic lupus erythematosus. Genes Immun 2009; 10:373-9.

130. Lohi H, Young EJ, Fitzmaurice SN, et al.Expanded repeat in canine epilepsy. Science 2005; 307:81.

131. Bartlett RJ, Stockinger S, Denis MM, et al.In vivo targeted repair of a point mutation in the canine dystrophin gene by a chimeric RNA/DNA oligonucleotide. Nat Biotechnol 2000; 18:615-22.