# Discovery of a Novel Pathway for an SLE-related Disease Complex in the Canine Breed Nova Scotia Duck Tolling Retriever

Maria Wilbe
*Faculty of Veterinary Medicine and Animal Science*
*Department of Animal Breeding and Genetics*

Doctoral Thesis
Swedish University of Agricultural Sciences
Uppsala 2013

Acta Universitatis agriculturae Sueciae
2013:16

Cover: Yrish, a healthy Nova Scotia duck tolling retriever included in present studies.
(photo: R. Meijer)

ISSN 1652-6880
ISBN 978-91-576-7773-0

Print: SLU Service/Repro, Uppsala 2013

# Discovery of a Novel Pathway for an SLE-related Disease Complex in the Canine Breed Nova Scotia Duck Tolling Retriever


## Abstract

The dog is an excellent model to study inherited complex diseases, due to its unique population history and haplotype structure. In this thesis, dogs from the breed Nova Scotia duck tolling retriever (NSDTR) have been used as a model for defining the genetic factors controlling a systemic lupus erythematosus (SLE)-related disorder called immune-mediated rheumatic disease (IMRD) and a steroid responsive meningitis-arteritis (SRMA).

IMRD is characterized by stiffness, mainly after resting, and pain from several joints of extremities and/or muscle pain. The majority of the affected dogs show anti-nuclear antibody (ANA)-positivity that can be divided either into a speckled ($ANA^S$) or homogenous ($ANA^H$) staining pattern. Dogs affected by SRMA display severe neck pain, fever and stiffness and an increased infiltration of immune cells in cerebrospinal fluid in the acute phase of disease. SRMA dogs show a negative ANA result.

We performed a candidate gene study to investigate if dog leukocyte antigen (DLA) class II is associated with the canine SLE-related disease. An increased risk for $ANA^S$ dogs was observed for a homozygous risk haplotype and a general homozygosity at DLA class II gives $ANA^H$ dogs an increased risk for developing disease.

Genome-wide association mapping identified additional susceptibility loci for the SLE-related disease on canine chromosomes (CFA) 3, 8, 11, 24 and 32. Further analysis revealed that most $ANA^S$ dogs homozygous for the DLA risk haplotype also have the genetic risk factors at CFA 11 and 32. Re-sequencing of the five associated regions was performed to identify specific genes and genetic variants involved in the disease.

Expression studies of the candidate genes for $ANA^S$ dogs revealed that the *PTPN3* (CFA 11) gene is downregulated and that *DDIT4L* and *BANK1* (CFA 32) is upregulated in dogs with the risk haplotype. The identified genes may be important in T-cells, B-cells and possibly macrophages.

This thesis describes the first successful mapping of a complex trait in the dog and shows that MHC class II together with two other identified genetic risk factors contribute to development of systemic autoimmune disease.



*Keywords:* dog, systemic lupus erythematosus, antinuclear antibodies, DLA class II, NF-AT, genome-wide association mapping, model organism



*Author's address:* Maria Wilbe, SLU, Department of Animal Breeding and Genetics, P.O. Box 597, 751 24 Uppsala, Sweden
*E-mail:* Maria.Wilbe@slu.se

# Contents

# List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I **Wilbe M,** Jokinen P, Hermanrud C, Kennedy LJ, Strandberg E, Hansson-Hamlin H, Lohi H, Andersson G (2009). MHC class II polymorphism is associated with a canine SLE-related disease complex. *Immunogenetics* 61(8): 557-564.

II **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 (2010). Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex. *Nature Genetics* 42(3): 250-254.

III **Wilbe M*,** Kozyrev SV*, Farias F, Hedlund A, Pielberg G, Gustafson U, Carlborg Ö, Andersson G, Lindblad-Toh K, Hansson-Hamlin H (2013). Risk of rheumatic disease with speckled ANA phenotype in dogs is associated with DLA class II and risk factors on Cfa11 and 32 conferring altered expression of *PTPN3, DDIT4L* and *BANK1*. Submitted.

Papers I-II are reproduced with the permission of the publishers.

# Related Work by the Author

(Not included in the thesis)

**Wilbe M,** Sundberg K, Hansen IR, Strandberg E, Nachreiner RF, Hedhammar Å, Kennedy LJ, Andersson G, Björnerfeldt S (2010). Increased genetic risk or protection for canine autoimmune lymphocytic thyroiditis in Giant Schnauzers depends on DLA class II genotype. *Tissue Antigens* 75(6): 712-719.

**Wilbe M,** Lund Ziener M, Aronsson A, Harlos C, Sundberg K, Norberg E, Andersson L, Lindblad-Toh K, Hedhammar Å, Andersson G, Lingaas F (2010). DLA class II alleles are associated with risk for canine Symmetrical Lupoid Onychodystrophy (SLO). *PLoS ONE* 5(8): e12332. doi:10.1371/journal.pone.0012332 (2010).

Truvé K, Eriksson O, Norling M, **Wilbe M,** Mauceli E, Lindblad-Toh K, Bongcam-Rudloff E (2011). SEQscoring: a tool to facilitate the interpretation of data generated with next generation sequencing technologies. *EMBnet journal* 17(1): 38-45.

**Wilbe M,** Andersson G (2012). MHC class II is an important genetic risk factor for canine systemic lupus erythematosus (SLE)-related disease: implications for reproductive success. *Reproduction in Domestic Animals* 47(1): 27-30. doi: 10.1111/j.1439-0531.2011.01962.x.

# Abbreviations

| | |
|---|---|
| ANA | antinuclear antibodies |
| $ANA^{H}$ | antinuclear antibodies with homogenous staining pattern |
| $ANA^{S}$ | antinuclear antibodies with speckled staining pattern |
| BCR | B-cell receptor |
| bp | base pair |
| CFA | *Canis familiaris* chromosome |
| CNV | copy number variant |
| DLA | dog leucocyte antigen |
| DNA | deoxyribonucleic acid |
| GWAM | genome-wide association mapping |
| GWAS | genome-wide association study |
| HLA | human leukocyte antigen |
| IBD | identical-by-descent |
| IMRD | immune-mediated rheumatic disease |
| InDel | insertion or deletion |
| Kb | kilo bases |
| LD | linkage disequilibrium |
| MAF | minor allele frequency |
| Mb | mega bases |
| MHC | major histocompatibility complex |
| NSDTR | Nova Scotia duck tolling retriever |
| OR | odds ratio |
| PBMC | peripheral blood mononuclear cell |
| PCR | polymerase chain reaction |
| RNA | ribonucleic acid |
| RNAseq | RNA sequencing |
| SLE | systemic lupus erythematosus |
| SNP | single nucleotide polymorphism |

| | |
|---|---|
| SRMA | steroid-responsive meningitis-arteritis |
| TCR | T-cell receptor |
| UTR | untranslated region |

# 1 Introduction

## 1.1 The domestic dog as a model

The exact time and location of the domestication process of the dog from its grey wolf ancestors is debated and several efforts have been made to identify its origin. Fossil records of dog-like skeletons have been dated back 33,000 years ago (Ovodov *et al.*, 2011), where mitochondrial DNA (deoxyribonucleic acid) analysis and single nucleotide polymorphism (SNP) data dates the domestication process ~15,000 years back (Larson *et al.*, 2012; Pang *et al.*, 2009; Savolainen *et al.*, 2002; Vila *et al.*, 1997). How the dog was domesticated is also unclear (Niskanen *et al.*, 2013; Ding *et al.*, 2012). One hypothesis is that human selected wolfs to fulfill their purposes (*e.g.* a friendly behavior and good hunting skills). Both behavioral and morphological differences occur between dog and wolf and efforts have been made to identify genetic differences responsible for such traits. A genome-wide association study (GWAS) performed in several dog breeds and 12 wolfs distributed world-wide was performed and identified strong selection for genes important in brain function, starch digestion and fat metabolism, which may underlie an adaptation to a starch rich diet (Axelsson *et al.*, 2013).

During the dog domestication process a first bottleneck occurred, decreasing the genetic pool, since it is believed that only a limited number of wolfs were founders of the dog population. A second bottleneck occurred recently, during breed creation, creating around 400 different breeds worldwide (Fogle *et al.*, 2000). Most breeds were shaped less than 200 years ago. The dog has been selected and bred for its unique characteristics, such as morphological traits (skull shape, body size, coat color *etc*.) as well as behavioral traits (hunting, herding, guarding, tolling and retrieving) creating a unique population structure, with each breed arising from a limited number of founders. This has resulted in a unique population structure within the dog.

Short ancestral haplotype blocks are shared between breeds, from the domestication of the wolf population and long haplotype blocks occur within a dog breed (Karlsson & Lindblad-Toh, 2008; Lindblad-Toh *et al.*, 2005; Sutter *et al.*, 2004).

Many breeds share recent common ancestors meaning that they also likely share common disease-causing alleles, offering a unique opportunity for genetic studies (Parker *et al.*, 2004). Dogs suffer from many of the same diseases as humans, such as cancer, epilepsy, heart diseases, allergies and autoimmune diseases. The different diseases have in both dog and human a similar frequency and the same genetic factors are often involved (Karlsson & Lindblad-Toh, 2008).

Specific diseases occur in certain breeds, suggesting a common genetic component. The genetic variation is limited within breeds compared to between breeds, which gives canine studies unique opportunities to understand the genetic mechanisms of natural variation in mammals, including disease susceptibility (Ostrander, 2012; Shearin & Ostrander, 2010).

There are several advantages for using the dog as a genetic model:

- Similarities between the canine and human genome
- Each breed represents a closed breeding pool
- Homogeneity within a dog breed
- Trait-causing mutations are identical-by-descent (IBD) within a breed
- Same spontaneously occurring diseases as in humans
- Human and dog share environment
- Medical records and pedigrees accessible
- Easier to identify associated loci of interest than in humans
- Less cases and controls are needed

Challenges when using the dog as a model include:

- Difficult to obtain sufficient number of samples needed
- Population stratification
- High linkage disequilibrium (LD) can make it difficult to get from associated locus to mutation
- Difficult to identify additional breeds that share the same phenotype for narrowing down associated regions

### 1.1.1 Nova Scotia duck tolling retriever

In this thesis, dogs from the breed Nova Scotia duck tolling retriever (NSDTR) have been used as a model for understanding the genetic factors controlling complex immune-mediated disease, as this breed has been shown to be

overrepresented to develop such diseases. The history of the NSDTR is interesting and important to remark since it influences the genetic studies and its interpretations of the results presented in this thesis.

The NSDTR is bred for its special behavior– tolling. A tolling dog plays, runs and jumps around the shore, sometimes disappearing for creating a curiosity from birds, luring them to swim closer to the shore to appear within shooting distance to the hunter. After the hunter shoots the bird, the dogs are sent out to retrieve them (Figure 1) (Strang & MacMillan, 1996).

*Figure 1.* Pancho, a Nova Scotia duck tolling retriever (NSDTR). These dogs are bred for their fox-like appearance that is used to attract and retrieve birds for the hunter (photo: R. Meijer).

The origin of the NSDTR is slightly unclear and several speculations occur. Both the appearance and their special behavior (tolling) have been used in attempts to trace the breed's background.

The NSDTR as a breed may be traced to Europe before the sixteenth century where dogs with similar behaviors and appearance were described. It is believed that the NSDTR derives from the Dutch breed Kooikerhondje. These dogs were used by the Dutch people to lure ducks into traps – a similar behavior as the NSDTR, but without the retrieving qualities. From Holland these dogs were imported to England and France for their fox-like resemblance to lure ducks into the hunters nests. During the French colonization of Acadia (nowadays Nova Scotia, Canada), it is hypothesized that the French people brought these dogs over when they established their first permanent settlement. The NSDTR was developed in the Yarmouth region of Nova Scotia and first described in the early 1800s. At that time it was known as Little River Duck Dog or the Yarmouth Toller (Strang & MacMillan, 1996).

Little is known of what happened during the coming 200 years. But crosses with different breeds have been proposed for the development of the modern NSDTR breed. Potential crosses with Labrador retriever, Chesapeake Bay retriever, Brittany spaniel, Golden retriever and perhaps some small farm collie

have been suggested (Perrin, 2012; Strang & MacMillan, 1996). It is important to understand the background of the NSDTR to ease the genetic studies and break down the high degree of LD within a dog breed by using related breeds.

An important part of the history affecting the development of NSDTR is that canine distemper virus (CDV) outbreaks were reported to occur twice in history; in 1908 and 1912. This reduced the population size to only a few individuals. A theory is that the predisposition to autoimmune diseases in the modern NSDTR may be a result of the early NSDTRs ability to survive these outbreaks of CDV (Strang & MacMillan, 1996). A highly efficient immune response towards a virus or other pathogens can lead to that the body starts to react against a self antigen and treats it as foreign and initiates an autoimmune reaction directed towards the self antigen or cells expressing the antigen. NSDTRs are reported to have an increased incidence of a systemic lupus erythematosus (SLE)-related disease (immune-mediated rheumatic disease (IMRD)) (Hansson-Hamlin & Lilliehook, 2009), steroid-responsive meningitis-arteritis (SRMA) (Anfinsen *et al.*, 2008; Redman, 2002) and Addison's disease (Hughes *et al.*, 2010; Hughes *et al.*, 2007; Burton *et al.*, 1997). In our Scandinavian study population we most often see IMRD and SRMA (Hansson-Hamlin & Lilliehook, 2009), whereas in USA, Addison's disease seems to be more common. One possible explanation to this is that different risk alleles segregate in these populations. Sometimes NSDTRs are also affected by hypothyroidism (unpublished observation).

The first 15 NSDTR dogs were registered in the Canadian Kennel Club in 1945, but these dogs were still a quite well kept secret of southwestern Nova Scotia. Breeding problems almost led to extinction of the population once again in the mid 20$^{th}$ century. Breeder enthusiasts saved the breed by using 13 dogs that are the founders of the current NSDTR population. The first NSDTR was imported to Sweden in 1984 (Strang & MacMillan, 1996). It is currently a popular breed and around 400 dogs are registered in Sweden each year (Tollarklubben, 2007).

### 1.1.2 Autoimmune and immune-mediated diseases

Developments of autoimmune and immune-mediated diseases are dependent on both genetic and environmental risk factors. Combined, these factors will affect the overall reactivity of the immune system and control which antigen to be targeted. Autoimmune diseases are classified into two groups, organ-specific and systemic autoimmune diseases. This is dependent on where the target antigen is expressed. In organ-specific autoimmune diseases the autoantigen only occurs in a certain tissue or cell type, whereas in systemic

autoimmune disease the response is directed against autoantigens expressed throughout the body.

Autoimmune diseases occur in up to 3–5% in the general human population (Jacobson *et al.*, 1997) and most organs in the body have an autoimmune disease connected (Marrack *et al.*, 2001).

Similar or sometimes the same genes are often involved in several autoimmune diseases. SLE may exemplify this, where the same genes or gene families are reported to be associated with other systemic rheumatic diseases or sometimes even organ-specific autoimmune diseases. One hypothesis is that there are genetic risk factors that are common to many autoimmune diseases and others that are specific for each disease (Shamim & Miller, 2000).

Autoimmune diseases can be mediated by T-cells, where both $CD4^+$ T-helper cells and sometimes $CD8^+$ T-killer cells have an important role and the organ damage is mediated by T-cells (*e.g*. type I diabetes) (Homann & von Herrath, 2004). In SLE, damage is initiated by autoantibodies and $CD4^+$ T-helper cells and an increased level of antibodies are observed. These antibodies form immune complexes and can cause glomerulonephritis (Kotzin, 1996).

The dog share many autoimmune and immune-mediated diseases with humans, such as SLE-related disease vs. SLE in humans (the disease studied in this thesis). A few of them are described below (Table 1).

*Table 1*. Examples of similar autoimmune or immune-mediated diseases that occur spontaneously in the dog and human.

| Dog disease | Human disease |
|---|---|
| Canine diabetes mellitus | Latent autoimmune diabetes of adults |
| Hypoadrenocorticism | Addison's disease |
| Primary immune-mediated haemolytic anaemia | Autoimmune haemolytic anaemia |
| Canine systemic lupus erythematosus (SLE)/ SLE-related rheumatic disease (ANA-positive) | Systemic lupus erythematosus (SLE)/ SLE-related rheumatic disease (ANA-positive) |
| Canine rheumatoid arthritis | Rheumatoid arthritis |
| Symmetrical lupoid onychodystrophy | Several keratin disorders |
| Canine lymphocytic thyroiditis | Hashimoto's thyroiditis |
| Necrotizing meningoencephalitis | Acute forms of multiple sclerosis |
| Uveodermatologic (UV) syndrome | Vogt – Koyanagi – Harada syndrome |

### 1.1.3 SLE and SLE-related disease complex

In humans, SLE is a heterogeneous autoimmune disease with a wide range of clinical signs. The classification is based on 11 different criteria defined by the

American College of Rheumatology (ACR) (Hochberg, 1997; Tan *et al.*, 1982), ranging from mild to severe symptoms. Four groups include dermatological signs (such as butterfly rash, discoid rash, photosensitivity and oral ulcerations) and another four groups include systemic criteria (arthritis, serositis, renal and neurological disorders). Laboratory findings include haematologic disorders and immunologic disorders (such as groups based on ANA-positivity). Four (or more) of the criteria are required for an SLE diagnosis (Hochberg, 1997; Tan *et al.*, 1982).

Attempts to create a similar list of criteria for SLE in dogs have been made, but no consistent list exists. Therefore it is often more difficult to diagnose canine SLE. Moreover, ANA positive dogs usually show clinical signs with focus on stiffness and musculoskeletal signs and more seldom impact of other organs. Thus, the disease in these dogs is often called SLE-related disease (Hansson *et al.*, 1996).

NSDTRs have been shown to be overrepresented to develop immune-mediated disease complexes; mainly the SLE-related disorder IMRD and SRMA. These disorders have been genetically investigated in this thesis.

When the genetic studies started in 2007 it was unclear if these diseases had a correlation or if they were separate disorders. One of the aims of this project was to define whether unique or combined genetic risk factors occur within the disease complex/complexes.

IMRD is characterized by stiffness, mainly after being inactive, pain from several joints of extremities and muscle pain. Sometimes fever and skin problems are observed, while concurrent liver and kidney changes are rare. Initial signs of disease are usually shown between 2-6 years of age. Most affected dogs have high serum concentrations of anti-nuclear antibodies (ANA) and are therefore termed ANA-positive. Treatment usually involves corticosteroids and the results may differ. For some dogs, treatment eventually may be withdrawn while in a minority of cases corticosteroid-treatment is insufficient. Most dogs need a lifelong corticosteroid-treatment with a low dose (Hansson-Hamlin & Lilliehook, 2009). During 2010-2011, 16 NSDTR newly diagnosed dogs with an ANA-positive phenotype were reported to us at the University Animal Hospital, Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden. ANA tests may be performed at other clinics, but the majority of the cases are sent to the clinic at SLU. Around 400 NSDTRs are registered each year, suggesting an incidence of at least 2% in the Swedish population. However, a longer time period is needed to obtain conclusive results. The incidence of 2% is most likely an underestimate, but it gives an indication of how common this disease is in the Scandinavian NSDTR population. In 2003, another study estimated the prevalence of IMRD in the

Swedish population by sending out a health form to all dogs born 1998. 35% of the dogowner responded and the prevalence of IMRD was estimated to 2.3% (Tollarklubben, 2007).

Dogs affected by SRMA display severe neck pain, fever and stiffness. The age of onset is usually between 4-19 months. An increased infiltration of immune cells can be seen during investigation of cerebrospinal fluid in the acute phase of disease. SRMA dogs always show a negative ANA result. Corticosteroids are often recommended for these dogs as well, usually with a remarkably good response within one or two days (Anfinsen *et al.*, 2008). In the study from 1998, the prevalence of SRMA in Sweden was predicted to 3.0% (Tollarklubben, 2007).

Our hypothesis was that IMRD and SRMA represent two separate disorders with some common but mostly distinct genetic risk factors. In our genetic studies we therefore considered the IMRD and SRMA as two separate disorders but also analyzed them for shared genetic risk factors. Clinically the phenotypes look very different, although they both seem to have an immune-mediated background.

### 1.1.4 Antinuclear antibodies (ANA)

During an immune response against extracellular pathogens, professional antigen-presenting cells process the antigens and present peptides on major histocompatibility complex (MHC) class II molecules to the T-cell receptor (TCR) on $CD4^+$ T helper cells. After such an antigen-presentation of the extracellular antigens, antibodies are produced by B-lymphocytes. In autoimmune diseases autoantibodies are produced against self-antigens. Such autoantibodies may be used as a hallmark for several autoimmune diseases in both humans and domestic animals.

ANA are autoantibodies directed against different nuclear antigens and are found in patients with certain systemic rheumatic diseases, including SLE, in both humans and dogs (Kavanaugh *et al.*, 2000; Hansson *et al.*, 1996). ANA can be subdivided according to their specificity. In humans, certain autoantibodies have been shown to produce distinct patterns of staining when they react with specific antigens *e.g.* dsDNA (double-stranded DNA) (Kavanaugh & Solomon, 2002), histone (Burlingame & Rubin, 1991), Sm (Smith antigen, a complex of ribonucleic acid (RNA) and protein) (Tan & Kunkel, 1966), RNP (ribonucleoprotein, react with proteins present in the U1 snRNP complex) (Benito-Garcia *et al.*, 2004), SSA/Ro (two proteins, Ro60 localized to the nucleus and nucleolus and Ro52 localized in the cytoplasm) (Chan *et al.*, 1991) and SSB/La (RNA-binding protein important in transcription mediated by RNA polymerase III) (Chambers *et al.*, 1988).

In general, anti-dsDNA and -histone antibodies produce a homogenous staining pattern ($ANA^H$), whereas anti-Sm, -RNP, -SSB/La and -SSA/Ro produces a speckled pattern ($ANA^S$). $ANA^H$ is associated with antibodies directed against chromosomal antigens, whereas $ANA^S$ is associated with antibodies against non-chromosomal antigens. Moreover, autoantibodies directed towards certain specific antigens are in humans correlated to different specific disorders and some specific autoantibodies occur in several disorders (Tan, 1989). A specific marker for SLE is Anti-Smith (Anti-Sm) antibodies and anti-dsDNA (Kavanaugh & Solomon, 2002; Tan & Kunkel, 1966).

When conducting the immunofluorescence (IIF) ANA-test on canine sera, these patients usually display either a homogenous or a speckled staining pattern. The specific autoantigens may be difficult to determine with traditional methods in SLE-related disease in dogs (Hansson & Karlsson-Parra, 1999). Anti-RNP and anti-dsDNA antibodies have been reported in several dog breeds displaying ANA reactivity (Lin *et al.*, 2006; Monier *et al.*, 1992). The specific ANA-reactivity in IIF ANA-positive NSDTR dogs is so far unknown.

Dogs from several different breeds have been reported to be affected by an SLE-like disease (including a positive ANA-test) such as the German shepherd (Hansson *et al.*, 1996; Thoren-Tolling & Ryden, 1991).

A study conducted using dogs from 27 different breeds affected by SLE or SLE-related disease showed that 25% had a homogenous ($ANA^H$) staining and had clinical signs from multiple organs. The other 75% showed a speckled pattern ($ANA^S$) with musculoskeletal disorders, fatigue and fever (Hansson-Hamlin *et al.*, 2006)(Figure 2). Similar results were obtained in a study using 33 NSDTRs where 39% showed a homogenous staining pattern and 61% showed a speckled pattern (Hansson-Hamlin & Lilliehook, 2009).

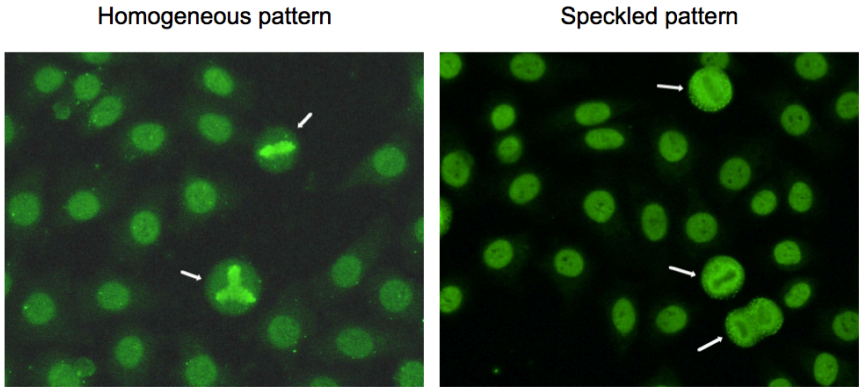


*Figure 2.* ANA-staining patterns in dogs shown by the indirect immunofluorescence (IIF) ANA-test. Homogenous staining shows reactivity in the chromosomal regions whereas speckled pattern lack chromosomal activity. (Modified with permission from (Hansson-Hamlin *et al.,* (2006))

Besides SLE and SLE-related diseases, a positive ANA-test in the dog has also been observed in a few other autoimmune diseases. Gordon setter dogs with symmetrical lupoid onychodystrophy (Ovrebo Bohnhorst *et al.*, 2001) and beagles with autoimmune thyroiditis (Vajner, 1997) may display ANA-positivity (30-35% vs. 10%).

### 1.1.5 Similarities to human SLE

Some symptoms are shared between human SLE and SLE-related disease complex in the dog, such as arthritis and a positive ANA test (Hansson-Hamlin & Lilliehook, 2009) (Table 2). Other symptoms that occur in both species are skin symptoms, fever, and liver problems. Kidney problem is rarely seen in the dog but do occur. MHC class II has also been identified as a genetic risk factor in both humans and the dog (Paper I and III) (Fernando *et al.*, 2008).

*Table 2.* Shared clinical symptoms between human SLE and the SLE-related disease affecting NSDTR.

| | ANA + | Joint | Skin | Fever | Liver | Kidney | MHC class II |
|---|---|---|---|---|---|---|---|
| Dog | Yes | Yes | Sometimes | Sometimes | Sometimes | Rarely | Yes |
| Human | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Women in childbearing age are more often affected than men (80-90%) (Danchenko *et al.*, 2006), which is a difference compared to the SLE-related disease in NSDTR dogs, where a similar distribution is observed between gender (Hansson-Hamlin & Lilliehook, 2009). In general, there are no gender

differences observed in autoimmune diseases in the dog and the reason why is unknown.

Both genetic components and environmental factors are involved as risk factors for development of the disease. More than 40 different genetic risk factors have been identified and confirmed in humans, many of importance for B- and T-cell activation, type I interferon, immune-complexes and toll-like receptors (Cui *et al.*, 2013).

Genetic data identified in the SLE-related disease and presented in this thesis implicates the importance of a T cell-specific pathway (Paper I-III). Several genes important for T cell-specific signaling and activation have also been identified in humans with SLE, including the human leukocyte antigen (HLA) class II region, *PTPN22*, *TNFSF4, STAT4* and *CD44* (Table 3). Interestingly, all of these genes have also been denoted as risk factors in several other autoimmune diseases (Cui *et al.*, 2013).

*Table 3.* Genetic risk factors associated with T-cell signaling or activation identified in human SLE patients.

| Chromosome | Gene | SNP type | Reference |
|---|---|---|---|
| 1p13 | *PTPN22* | exonic | (Criswell *et al.*, 2005) |
| 1q25.1 | *TNFSF4* | intergenic | (Cunninghame Graham *et al.*, 2008) |
| 2q32.3 | *STAT4* | intron | (Remmers *et al.*, 2007) |
| 6p21.32-33 | *HLA-DQA1, HLA-DQA2* | intergenic | (Chung *et al.*, 2011) |
| 6p21.32-33 | *HLA-DR3* | intron | (Chung *et al.*, 2011) |
| 6p21.32-33 | *HLA-DRB1* | intergenic | (Gateva *et al.*, 2009; Han *et al.*, 2009) |
| 11p13 | *CD44* | intergenic | (Lessard *et al.*, 2011) |

PTPN22 (protein tyrosine phosphatase, non-receptor type 22) is associated with the inhibition of T-cell receptor and development of autoantibody production (Criswell *et al.*, 2005). Tumor necrosis factor superfamily 4 (TNFSF4) is a co-stimulatory molecule for $CD4^+$ T helper cells and is expressed on activated antigen-presenting cells. TNFSF4 show correlation to increased expression in patients with SLE (Cunninghame Graham *et al.*, 2008). STAT4 (Signal transducer and activator of transcription 4) transmits signals from cytokines, type I interferons and interleukins (-12 and -23) and stimulates transcription of genes important for T-cell differentiation (Th1) (Nishikomori *et al.*, 2002). CD44 is a cell-surface glycoprotein and is important in lymphocyte activation. T-cells from SLE patients have shown overexpressed CD44 (Li *et al.*, 2007). The HLA class II region is the strongest genetic risk factor for SLE (Chung *et al.*, 2011; Gateva *et al.*, 2009; Han *et al.*, 2009) and has been associated to a number of autoimmune diseases in humans like

multiple sclerosis, rheumatoid arthritis and type I diabetes (Fernando *et al.*, 2008).

Most of the associated SNPs in the studies described above occur in non-coding regions. There are infrequently SNPs occurring in regions with a known function, such as in coding region, splice site mutation or 3' or 5' untranslated regions (UTR). The mutations are more likely to be important for gene regulation, transcription factor binding sites, enhancers or promoters.

The genes identified as risk factors for human SLE are only explaining a small part of the inheritance of the disease, only showing odds ratios (OR) with a median effect (~1.25) (Manolio *et al.*, 2009; Pawitan *et al.*, 2009). A higher concordance rate in monozygotic twins than in dizygotic twins or siblings has been observed, but lack complete concordance (Alarcon-Segovia *et al.*, 2005). The low OR and missing heritability suggests an important role for the environmental factors in the pathogenesis of SLE.

There is also extensive clinical heterogeneity of SLE. By using the SLE-related disease in NSDTR as a model, we may discover new genes and pathways involved in the development of SLE and provide improved subdivision of patients dependent on the same immunological subgroup.

## 1.2 Mapping the first complex trait in the dog

In 1989, DNA sequencing identified the first genetic variant for an inherited canine disorder, haemophilia B. A single missense mutation was identified in the canine Factor IX gene, which encodes a glycoprotein that is required for blood coagulation (Evans *et al.*, 1989). The rapid developments of high-density SNP arrays, nucleotide sequencing technologies and the sequencing of genomes have made it easier and possible to identify loci also for a complex trait.

The results presented in this thesis describe the first successful genetic mapping of a complex trait in the domestic dog. Previously, genetic mapping of several monogenic diseases or traits have been described in the dog such as the gene encoding the sleep disorder narcolepsy (Lin *et al.*, 1999). Two proof-of-principle GWAS were previously performed on monogenic traits, the white spotting locus and the hair ridge in Rhodesian ridgebacks (Karlsson *et al.*, 2007) predisposing dermoid sinus (Salmon Hillbertz *et al.*, 2007). The results in this thesis have used a combination of three different strategies, a candidate gene approach to investigate the dog MHC class II, a genome-wide association mapping (GWAM) to unbiased search for risk loci and next generation sequencing technologies to identify genetic variants important for mechanisms involved in disease development. The sequencing of the dog genome and

techniques used to identify associated regions are described in the following sections.

### 1.2.1 The dog genome and genetic variation

The report of a draft sequence of the human genome in 2001 (Lander *et al.*, 2001; Venter *et al.*, 2001) opened up new doors to study the genome in depth. In 2005, the dog genome sequence of a female boxer, Tasha, was published, covering 99% of the genome (Lindblad-Toh *et al.*, 2005). This dog was chosen based on tests for a low rate of heterozygosity, which makes it easier to assembly the sequences and generate a better quality of the genome sequence.

The dog has 38 autosomal chromosomes and the sex chromosomes X and Y. The genome structure is 2.4 Gb in size and contains 20,657 protein-coding genes with RNA-sequencing evidence (M. Grabherr, personal communication). The dog genome is smaller than the human and mouse genomes, partly because of a lower amount of some repeat insertions including endogenous retroviruses in the dog genome (Barrio *et al.*, 2011). The mechanisms underlying this difference are unknown but could be caused by a combination of canids being more efficient in purging repetitive DNA sequences and having a more restricted number of canid-specific infectious retroviruses. Overall, 94% of the dog genome is in conserved synteny to the human and mouse genomes (Lindblad-Toh *et al.*, 2005). Some genes that only appear in the dog genome are genes encoding G protein-coupled olfactory receptors. This is a result of positive selection for duplications in this gene family leading to canid-specific expansion of such genes.

In parallel to the sequencing of the dog genome, SNP discovery was performed using dogs from 11 different breeds, which identified 2.5 million SNPs (on average 1 SNP per 1,000 base pair (bp)) (Lindblad-Toh *et al.*, 2005). Types of variation that occur within the genome are SNPs, insertions or deletions (InDels) and other structural variations such as copy number variants (CNV) - deleting or duplicating a large region of the DNA and inversions - a chromosome rearrangement.

Within a dog breed LD is extensive with haplotypes extending 0.5-1.0 mega bases (Mb) in size, but between breeds LD is only around 10 kilo bases (Kb). Some breeds share longer haplotypes within the shorter ones, suggesting that shared genetic risk factors can occur, due to the two bottlenecks described in chapter 1.1 (Karlsson & Lindblad-Toh, 2008; Sutter *et al.*, 2004).

### 1.2.2 Candidate gene approach

The genes studied in a candidate gene approach are chosen based on a hypothesis-biased search. The genes can be located in a place in the genome

previously associated with linkage studies, the genes may be involved in the same pathway as other genes implicated to be important for the disease, genes may be expressed in cells or tissues important and relevant for the disease or associated in other model organisms with same/similar disease.

Most commonly, SNPs are genotyped to tag associated haplotypes, but sequencing can also be used for part of a gene, the entire gene, promoters and conserved elements. The candidate gene approach is useful, because less markers and samples are needed to generate sufficient statistical power and less multiple testing corrections are required. It is also of lower cost than more high-throughput methods.

The candidate gene approach was used for Paper I and III for sequencing the polymorphic exon 2 of MHC class II, because it is known as an important genetic risk factor in SLE and other autoimmune diseases in humans, mice and other mammals. The biological function of MHC class II molecules as antigen-presenting molecules is also of relevance for the disease development.

### 1.2.3 Genome-wide association mapping

Genome-wide association mapping (GWAM) was used in Paper II to perform an unbiased scan of the entire genome in controls and cases affected by the SLE-related disease in NSDTR. The dog-sequencing project (Lindblad-Toh *et al.*, 2005) identified a large number of SNPs used to create canine SNP genotyping arrays used for GWAM. The first SNP array contained 27,000 markers (Karlsson *et al.*, 2007). Subsequently, a 22K array (that was used in Paper II in this thesis) and 50K were developed and finally the high density 173K genotyping array (Vaysse *et al.*, 2011) is now available.

When performing a GWAS, the entire genome is scanned in all individuals from a case-control population in an attempt to find regions where the affected individuals are genetically identical and differ from the healthy controls, defining an associated haplotype. The associations can either be a single marker in LD with the causative mutation, an associated haplotype or the actual causative mutation. In dogs, like other domestic animals, this approach is particularly powerful since dog breed creation occurred recently, few recombination events have occurred and long haplotype blocks can be identified that potentially share the trait of interest, an IBD region. The advantage of using the dog is that fewer markers, compared to *e.g.* humans are needed. It has been estimated that 10,000 to 15,000 SNPs are sufficient in the dog (Lindblad-Toh *et al.*, 2005; Sutter *et al.*, 2004) compared to 300,000 to 1,000,000 in humans (Gabriel *et al.*, 2002).

Fewer samples are also needed in the dog compared to humans. It has been estimated that to map a Mendelian recessive trait in the dog, only 20 cases and

20 controls are needed and that for a complex disease an allele conferring a 5-fold increased risk requires 100 cases and 100 controls to have 98% chance of being detected (Karlsson & Lindblad-Toh, 2008; Karlsson *et al.*, 2007; Lindblad-Toh *et al.*, 2005). Compared to humans, larger study populations are needed *e.g*. 6,000 cases and 6,000 controls would generate 94% power to detect disease susceptibility variants with an odds ratio of 1.3 and minor allele frequency (MAF) of 0.1 (Wang *et al.*, 2005).

The study population used is a case/control cohort based on strict inclusion and exclusion criteria. In the dog it can be easier to collect random unrelated cases and controls than to collect complete families (as previously used in linkage studies). To avoid stratification it is important to use unrelated individuals from the same population structure, otherwise false positives may occur due to the heterogeneity effect.

We used a 2-stage mapping strategy. In the first stage, GWAM, we used one breed with the disease or trait of interest. A homogenous study population is required to avoid false positives. The hypothesis is that the mutation arose before breed-creation. Allele frequencies between cases and controls are compared and long (around 1Mb) regions are identified. To narrow down associated regions the second stage is applied, fine-mapping. More SNPs are added to the associated regions and preferably samples from a related breed sharing the same disease or trait of interest. Since haplotypes are short (around 10 Kb) across breeds the associated haplotype can effectively be narrowed down (Karlsson & Lindblad-Toh, 2008; Karlsson *et al.*, 2007).

Many studies have been successfully performed to map monogenic traits and in this thesis we present the first complex trait successfully mapped (Paper II).

## 1.3 The role of MHC class II in canine autoimmune diseases

Major histocompatibility complex (MHC) was described in the 1930s by Peter A. Gorer (Gorer, 1937). The organization of the region is highly conserved between species and consists of three classes:

- MHC class I: expressed by all nucleated cells except neurons, present intracellular antigens and mediates $CD8^+$ T killer cells to destroy infected cells.
- MHC class II: expressed by professional antigen-presenting cells, presents extracellular antigens to $CD4^+$ T helper cells, which leads to activation of B-cells to produce antibodies to the specific antigen.

- MHC class III: encode genes in the complement system (innate immune system).

The data presented in this thesis have focused on the polymorphic genes of the dog MHC class II because of their well-established role as genetic risk factors for development of autoimmune disease.

MHC class II genes are expressed as cell surface attached alpha/beta heterodimers on specialized leukocytes (professional antigen presenting cells such as dendritic cells, macrophages and B-cells) and act mainly as cell surface receptors for processed antigens of extracellular origin. These peptide antigens are bound to the polymorphic antigen-presenting cleft of the class II molecule (Bjorkman *et al.*, 1987). The size of peptides bound by class II molecules are usually between 13-25 amino acids and are presented by MHC class II molecules to T-cell receptors on $CD4^+$ T helper cells. In the event of activation, helper T-cells will release cytokines and interleukins, which can alter various types of immune reactions, such as the regulation of antibody production. This process is important for initiating the immune responses against infectious organisms but is also likely to be involved in immune-mediated or autoimmune diseases.

A wide diversity and polymorphism exist within the MHC to ensure that individuals can respond to a large variety of antigens. The MHC class II genes encode extremely polymorphic class II antigens (Bontrop *et al.*, 1999; Bach, 1985). In humans there are > 1,800 different alleles identified at the HLA class II complex and the DRB1 locus is the most polymorphic (Robinson *et al.*, 2013). There are several systemic rheumatic diseases in both dogs and humans associated with the MHC, including SLE (Fernando *et al.*, 2008; Smerdel-Ramoya *et al.*, 2005; Teichner *et al.*, 1990).

The classical MHC class II region in humans is called the HLA region and consists of DP, DQ and DR genes. The genes produce two glycoprotein chains, an α-chain and a β-chain that together form a heterodimeric structure with extracellular domains (α1+α2 and β1+β2), where α1 and β1 forms the peptide binding cleft. DRβ forms heterodimers with DRα and DQα with DQβ. At least one class II molecule from each class II sub region (DR and DQ in dog) composed of an alpha-beta heterodimer is expressed on the cell surface of antigen presenting cells (Albert *et al.*, 1985). The highest polymorphism (except DRα) is found in exon 2 of the α- and β-chains, which are involved in peptide binding (Little & Parham, 1999).

The HLA complex is located on human chromosome 6p21.31 and extend approximately 4 Mb (Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium, 1999). Almost

half of the expressed genes within this region (>120) are involved in the immune defense (Stewart *et al.*, 2004).

The MHC class II region differs markedly between species. Humans have one monomorphic HLA-DRA gene, four expressed HLA-DRB genes (DRB1, DRB3, DRB4 and DRB5) (Andersson *et al.*, 1998) (and five DRB pseudogenes), one expressed HLA-DQA1 and DQA2 gene and finally one expressed HLA-DQB1 and DQB2 gene (Robinson *et al.*, 2013). In comparison, the feline MHC class II only consists of 44 genes, with an expansion of three DRA genes, seven different DRB genes and a lack of the entire DQ region (Yuhki *et al.*, 2003).

The dog MHC is called the dog leukocyte antigen (DLA) and consists of three classes. The DLA region is located on canine chromosome (CFA) 12, and the class II region is around 711 Kb in length. There are 45 loci of which 29 are complete functional genes (five are unprocessed pseudogenes and 10 are processed pseudogenes) (Debenham *et al.*, 2005). The dog has one DLA-DRA1, DLA-DRB1, DLA-DQA1, DLA-DQB1 gene, respectively and one DLA-DQB2 pseudogene. DLA-DQB2 gene shares 98% identity with DLA-DQB1 over 2562 bp (intron 2 to the end of exon 6) (Debenham *et al.*, 2005; Wagner *et al.*, 1999).

In 2008, 67 DLA-DRB1 alleles, 21 DLA-DQA1 alleles and 54 DLA-DQB1 alleles had been identified in the dog (Kennedy, 2013; Kennedy *et al.*, 2001; Kennedy *et al.*, 2000). Currently there are 300 DLA-DRB1, 51 DLA-DQA1 and 150 DLA-DQB1 alleles identified, but all are not yet confirmed. 206 DLA-DRB1, 37 DLA-DQA1 and 100 DLA-DQB1 alleles have been confirmed (L.J Kennedy, personal communication). The DLA-DRA1 locus appears to be monomorphic. There is a large distribution of polymorphism between the class II loci. DRB is most polymorphic followed by DQB and DQA whereas DRA is largely monomorphic in all mammals studied. The functional reason for why DR is more polymorphic than DQ is not known but could be related to requirement for different antigen presentation.

There is a large inter-breed variety in DLA class II both regarding the observed allele frequencies and the number of identified alleles in each breed. A general frequency of approximately 30% homozygotes has been observed at all three loci. Number of alleles within a dog breed varies from 2 – 22 and often there are certain specific alleles that dominate in frequency in a breed (Kennedy *et al.*, 2002). This interbreed variation could partly explain the fact that some breeds are susceptible to certain autoimmune or immune-mediated diseases.

### 1.3.1 MHC is involved in several canine immune-mediated diseases

The association between HLA and autoimmune disease in humans has been extensively reviewed during the past decades and several associations have been identified (Fernando *et al.*, 2008). Below follows a brief description of what is currently known within the field of canine autoimmune or immune-mediated diseases and its association to MHC class II (summarized in Table 4).

Canine diabetes mellitus occurs spontaneously in various breeds (Catchpole *et al.*, 2005). Association between diabetes mellitus and DLA class II has been reported by Kennedy et al. who studied the disease in various breeds and found three haplotypes associated with the disease (DRB1*009/DQA1*001/ DQB1*008, DRB1*015/DQA1*0061/DQB1*023 and DRB1*002/DQA1*009/ DQB1*001) and one protective haplotype (DQA1*004/DQB1*013). A trend was observed between the breeds with DLA-DQA1*00101 as a shared risk factor between breeds (Kennedy *et al.*, 2006b; Catchpole *et al.*, 2005).

Hypoadrenocorticism, also called Addison's disease, is a deficiency of corticosteroids and mineralocorticoids produced by the adrenal gland. A large number of breeds are predisposed to the disease. Several closely related alleles of the DLA-DRB1/DQA1/DQB1 in six different breeds show association to hypoadrenocorticism (Springer spaniel; DRB1*01501/DQA1*00601/ DQB1*02301, Labrador retriever; DLA-DRB1*00101/DQA1*00101/ DQB1*00201, Bearded collie; DLA-DRB1*00901/DQA1*00101/DQB1* 00802 and Cocker spaniel DLA-DRB1*00901/DQA1*00101/ DQB1*008011) (Massey *et al.*, 2013a). Another study conducted in NSDTR found a haplotype related to the previously identified as the risk haplotype, only differing at DLA-DRB1 (DLA-DRB1*01502/DQA*00601/ DQB1*02301) (Hughes *et al.*, 2010). Portuguese water dogs affected by hypoadrenocorticism are associated to a microsatellite marker located within the DLA region (Chase *et al.*, 2006).

Primary immune-mediated haemolytic anaemia is an autoimmune disease where red blood cells are destroyed, giving rise to signs of anaemia. Two haplotypes (DRB1*01501/DQA1*00601/DQB1*00301 and DLA-DRB1*00601/DQA1*005011/DQB1*00701) was found to be significantly associated with the disease in several breeds (Kennedy *et al.*, 2006a).

Chronic inflammatory Hepatitis in Dobermann is a chronic and progressive inflammatory liver disease with a suggested autoimmune etiology (Speeti *et al.*, 2003; Meyer *et al.*, 1980). Dobermanns homozygous for the risk haplotype DLA-DRB1*00601/DQA1*00401/DQB1*01303, especially with homozygosity for DLA-DRB1*00601 are susceptible to hepatitis. A protective effect was also identified (DLA-DQA1*00901/DQB1*00101 and DLA-DRB1*01501) (Dyggve *et al.*, 2011). Another study including English springer

spaniel identified two risk alleles (DRB1*00601 and DQB1*00701) and two protective alleles (DRB1*00501 and DQB1*00501) (Bexfield *et al.*, 2012).

Canine rheumatoid arthritis (CRA) or canine polyarthritis is described as symmetrical spontaneously occurring inflammatory polyarthritis (Bari *et al.*, 1989). The DLA-DRB1*002 allele was found associated in several breeds (Ollier *et al.*, 2001). This allele contains the shared five amino acid epitope RARAA at amino acid positions 70–74. This is known as the shared epitope for rheumatoid arthritis in human (Gregersen *et al.*, 1987).

Symmetrical Lupoid Onychodystrophy is an immune-mediated disease in dogs and is described as separation and sloughing of several claws from claw beds and ultimately affecting all claws. The disease occurs in many breeds but Gordon setter has been reported to have a particularly high prevalence (Ovrebo Bohnhorst *et al.*, 2001). We have reported that dogs carrying haplotype DRB1*01801/DQA1*00101/DQB1*00802 were at increased risk for developing the disease and that the risk factor was even stronger in homozygous form. A protective haplotype (DRB1*02001/DQA1*00401/ DQB1*01303) was also identified and the effect of the protective haplotype was clearly stronger than the risk haplotype (Wilbe *et al.*, 2010b).

Canine Lymphocytic Thyroiditis (CLT) is characterized by autoimmune destruction of the thyroid gland with increased levels of thyroid-stimulating hormone (TSH) and sometimes presence of circulating autoantibodies against thyroglobulin (TgAA) (Dixon & Mooney, 1999; Nachreiner *et al.*, 1998). A risk haplotype (DLA-DRB1*01201/DQA1*00101/DQB1*00201) have been identified in giant schnauzers for developing CLT and the same DQA1 risk allele has been observed in Dobermanns and in a variety of additional breeds (Wilbe *et al.*, 2010a; Kennedy *et al.*, 2006c; Kennedy *et al.*, 2006d). In our study, protection for the disease was also identified in dogs carrying another haplotype (DRB1*01301/DQA1*00301/DQB1*00501) (Wilbe *et al.*, 2010a).

Canine anal furunculosis affects German shepherd dogs. This is an inflammatory disease of the perianal tissues. The disease is believed to be immune-mediated with evidence of response to immunosuppressive drugs (Hardie *et al.*, 2005). Kennedy et al. found strong genetic association to the DLA-DRB1*00101 allele and that homozygosity for DRB1*00101 had an earlier onset of anal furunculosis (Barnes *et al.*, 2009; Kennedy *et al.*, 2008).

Necrotizing meningoencephalitis in dogs is an autoimmune disease. It occurs in many different breeds but the Pug dogs share clinical similarities with several acute forms of multiple sclerosis in humans. MHC class II was shown to be a genetic risk factor where the haplotype (DRB1*0100110/DQA1*00201/DQB1*01501) was found to be associated with increased risk. Homozygosity gave an even higher risk for disease and a

protective effect was observed (DRB1*01502/DQA1*00601/DQB1*02301) (Greer *et al.*, 2010).

Uveodermatologic syndrome is an immune-mediated disease characterized by inflammatory infiltrates and loss of melanocytes in affected tissues (eye and skin) that ultimately causes blindness. The disease is highly predisposed in the American Akita breed. Angles et al found that DLA-DQA1*00201 was significantly associated with the syndrome (Angles *et al.*, 2005).

Canine chronic superficial keratitis is an inflammatory ocular disease of autoimmune nature. A risk haplotype (DLA-DRB1*01501/DQA1*00601/

DQB1*00301) was associated with disease in German shepherd dogs. There was also an increased risk for dogs homozygous for the risk haplotype and an increased risk associated with general homozygosity of the MHC class II (Jokinen *et al.*, 2011).

Polymyositis is frequently observed in the Hungarian Vizsla. An immune-mediated etiology has been suggested. A risk haplotype (DLA-DRB1*02001/DQA1*00401/DQB1*01303) was identified and the risk increased for homozygotes (Massey *et al.*, 2013b).

*Table 4.* Identified DLA associations with canine autoimmune or immune-mediated diseases.

| Condition | Comment | Breeds | References |
|---|---|---|---|
| Canine diabetes mellitus | Suggested immune-mediated disease | Range of breeds; Samoyed, Cairn and Tibetan terrier | (Kennedy *et al.*, 2007; Kennedy *et al.*, 2006b; Catchpole *et al.*, 2005) |
| Hypoadrenocorticism | Suggested autoimmune disease | Nova Scotia duck tolling retriever | (Massey *et al.*, 2013a; Hughes *et al.*, 2010; Chase *et al.*, 2006) |
| Primary immune-mediated haemolytic anaemia | Suggested autoimmune disease | Range of breeds | (Kennedy *et al.*, 2006a) |
| SLE related immune-mediated rheumatic disease | Suggested autoimmune disease | Nova Scotia duck tolling retriever | Paper I and III |
| Chronic inflammatory hepatitis | Suggested autoimmune disease | Dobermann | (Bexfield *et al.*, 2012; Dyggve *et al.*, 2011) |
| Canine rheumatoid arthritis | Autoimmune disease. RA shared epitope | Range of breeds | (Ollier *et al.*, 2001) |
| Symmetrical lupoid onychodystrophy | Immune-mediated disease | Gordon setter<br>Bearded collie<br>Giant schnauzer | (Wilbe *et al.*, 2010b) |
| Canine Lymphocytic thyroiditis | Autoimmune disease | Range of breeds;<br>Dobermann<br>Giant schnauzer | (Wilbe *et al.*, 2010a; Kennedy *et al.*, 2006c; Kennedy *et al.*, 2006d) |
| Canine anal furunculosis | Immune-mediated disease | German shepherd | (Barnes *et al.*, 2009; Kennedy *et al.*, 2008) |
| Necrotizing meningoencephalitis | Suggested autoimmune disease | Pug | (Greer *et al.*, 2010) |
| Uveodermatologic (VKH-like) syndrome | Immune-mediated disease | Akita | (Angles *et al.*, 2005) |
| Canine chronic superficial keratitis | Suggested autoimmune disease | German shepherd | (Jokinen *et al.*, 2011) |
| Polymyositis | Suggested immune-mediated disease | Hungarian Vizsla | (Massey *et al.*, 2013b) |

Thus, many canine autoimmune or immune-mediated diseases have shown DLA class II association. It is important to remark that although DLA-DRB1/DQA1/DQB1 genes are encoding proteins that are functionally confirmed to control immune responses, they are only markers for the disease risk until functional experiments have been performed to confirm the risk. This is due to the high LD within MHC in all species and in particular in dogs where the extensive inbreeding further increases LD that occurs within a dog

breed. Many other genes important for immune function is located within the DLA region and can carry the functional mutation/s.

However, functional studies using transgenic mice have shown that MHC class II is sufficient for inducing autoimmune diseases. *E.g*. transgenic mice expressing HLA-DQ8, a gene in linkage with DR4 associated with susceptibility with rheumatoid arthritis, were produced. The mice showed a collagen-induced arthritis and became an important animal model to study human autoimmune arthritis and suggesting that HLA-DQ molecule might have an important role in determining susceptibility to rheumatoid arthritis (Nabozny *et al.*, 1996).

In this thesis we have added one more breed to the studies of autoimmune or immune-mediated diseases and its association to MHC class II, the SLE-related disease complex in NSDTR.

### 1.3.2 General risk in homozygosity

The selective breeding practice has caused a reduced genetic variation and an increase in homozygosity (Lindblad-Toh *et al*., 2005). This can also be observed in the MHC class II of the dog, where several autoimmune or immune-mediated diseases show an increased risk if the associated haplotype occurs in homozygous form. A homozygous risk haplotype have been identified in dogs affected by symmetrical lupoid onychodystrophy (Wilbe *et al*., 2010b), canine chronic superficial keratitis (Jokinen *et al.*, 2011), chronic inflammatory hepatitis in Dobermann (Dyggve *et al*., 2011), necrotizing meningoencephalitis (Greer *et al*., 2010), polymyositis (Massey *et al.*, 2013b), earlier onset of canine anal furunculosis (Kennedy *et al*., 2008) and the SLE-related disease in NSDTRs, presented in this thesis (Paper I and III).

A general risk in homozygosity at DLA class II have been observed, independent of haplotype, in canine chronic superficial keratitis (Jokinen *et al.*, 2011) and hypoadrenocorticism (Massey *et al*., 2013a) and in $ANA^{H}$ dogs with the SLE-related disease presented in this thesis (Paper III). Dogs with hypoadrenocorticism tend to have an earlier onset of disease if homozygous (Hughes *et al*., 2010). No protective homozygous haplotype have been observed. The fact that homozygosity at DLA class II leads to differential risk for developing immune-mediated disease is expected to be caused by the ability of presentation of autoantigens by the risk DLA but not by the protective DLA class II molecules. MHC class II heterozygotes can express a wider variety of class II molecules than homozygotes, thereby being able to respond to larger diversity of pathogens. The finding is in agreement with rodent models for experimentally induced autoimmune disease where high-

and low responder strains (differing only at MHC) develop disease at different levels upon challenge with the inducing antigens (Holmdahl, 2003).

Heterozygosity at MHC appears advantageous in most natural populations. The prevailing hypothesis is that the selection for high degree of MHC class II polymorphism is pathogen-driven (Borghans *et al.*, 2004). Polymorphism is maintained because of heterozygote advantage in the individual of being able to present a larger repertoire of peptide antigens and thus increase the chance of mounting an immune response towards pathogenic microorganisms. Thus, MHC polymorphism is maintained because of selective pressure imposed by different pathogens occurring in a population's environment (Havlicek & Roberts, 2009). There are two classical examples for this. The first was described in chicken where Marek's disease was shown to be caused by a particular herpes-like virus and that the chicken with certain MHC types were susceptible and chickens with other MHC types were resistant leading to frequency-dependent selection for the chickens with resistant MHC types (Longenecker *et al.*, 1977). The other example is malaria in humans, where humans with HLA class II haplotype DRB1*1302/DQB1*0501 are able to defend the malaria-inducing trypanosomal parasite leading to increased frequency of the protective MHC type in West Africa where malaria is common (Hill *et al.*, 1991). It has also been shown that certain HLA class I alleles influence HIV progression (Carrington & O'Brien, 2003).

The exceptions to the general observation of high MHC polymorphism in natural mammalian populations are found either among solitary species or in populations that have experienced bottle-necks due to high hunting pressure or to other environmental factors (O'Brien & Yuhki, 1999; Mikko & Andersson, 1995; Ellegren *et al.*, 1993) Last, MHC variation has shown to be preserved because of sexual selection, where individuals prefer to mate with a partner with dissimilar MHC genotype to their own (Wedekind & Furi, 1997; Potts *et al.*, 1991).

# 2 Aims of this Thesis

The overall aim of this research project was to identify genetic risk factors for the immune-mediated diseases IMRD, which is an SLE-related disease complex, and SRMA, aseptic meningitis, in the canine breed NSDTR. In particular, the IMRD disease type characterized by a positive ANA phenotype was studied. Identification of novel genes important for this SLE-related disease may have major importance for both dog and human health.

The specific aims were to:

- Perform a candidate gene study to investigate if DLA class II is associated with the canine SLE-related disease and SRMA.

- Investigate if genome-wide association mapping identifies additional susceptibility loci for the SLE-related disease and SRMA.

- Perform regional re-sequencing of associated regions to identify specific genes and genetic variants involved in the disease and perform expression studies of positional candidate genes.

- Evaluate whether differential gene expression of the gene variants carried on risk versus control haplotypes is associated with IMRD with the speckled ANA phenotype.

- Characterize the relationship between ANA type (speckled and homogenous) and the genetic constitution at DLA class II and other identified risk loci for IMRD.

- Identify novel genes and pathways associated with the canine SLE-related disease complex.

# 3 Present Studies

## 3.1 Introduction

The studies (Paper I-III) presented in this thesis aims to identify genetic risk factors for an SLE-related disease complex in the dog NSDTR. The IMRD complex has previously been described (Hansson-Hamlin & Lilliehook, 2009) and since the NSDTR breed is overrepresented for the disease, genetic components are suggested to be involved in development of this disease complex. No specific mode of inheritance has been reported, but it appears to be a multifactorial disease. DLA class II has previously been reported to be involved in several autoimmune and immune-mediated disease in dogs and humans (Bexfield *et al.*, 2012; Dyggve *et al.*, 2011; Jokinen *et al.*, 2011; Greer *et al.*, 2010; Hughes *et al.*, 2010; Wilbe *et al.*, 2010a; Wilbe *et al.*, 2010b; Barnes *et al.*, 2009; Kennedy *et al.*, 2008; Kennedy *et al.*, 2006b; Kennedy *et al.*, 2006d; Smerdel-Ramoya *et al.*, 2005; Graham *et al.*, 2002; Ollier *et al.*, 2001), Therefore, a candidate gene approach was performed to investigate whether the DLA class II is a genetic risk factor for this SLE-related disease complex.

We also performed an unbiased search of the entire genome in an effort to identify additional genetic risk factors by using a GWAS. The DLA class II region contains several highly polymorphic genes and all alleles were not covered in the canine 22K SNP chip array used in these studies. Therefore, the optimal way to explore whether DLA class II is a genetic risk factor for the disease complex was by performing a separate sequence-based candidate gene study.

## 3.2 Methods and results

### 3.2.1 Dog samples

All dogs included in the studies described in Paper I-III were purebred and privately owned. All owners approved their dogs to participate in the study by filling out an owner's consent form. Ethical permission was granted by the Ethical board for experimental animals in Uppsala, Sweden (Dnr C138/6). Samples were obtained by a close contact with different veterinary clinics throughout Sweden, by visiting NSDTR competitions and shows, attending breed club meetings and by direct contact with dog owners with interest to participate in the study. Samples were collected during 2002-2013.

Veterinarians selected dogs to be included in the studies based on strict inclusion and exclusion criteria following examination. To be classified as affected by IMRD the dogs had to display musculoskeletal signs consistent with symmetrical polyarthritis and suffer from pain affecting several joints of extremities and display stiffness, mainly after rest. Signs needed to be apparent for at least 14 days. A positive IIF-ANA titre further strongly supported the diagnosis and classified the dogs as ANA-positive IMRD. Dogs classified as SRMA-affected display high fever and strong pain mainly from the neck and respond to corticosteroid treatment. The diagnosis is supported by cerebrospinal fluid showing a significant neutrophilic pleocytosis. These dogs have a negative ANA-test. Dogs fulfilling the exclusion criteria were the dogs with no clinical signs of any autoimmune disease and >7 years of age.

EDTA blood and serum samples were collected for all dogs included in the study population (Paper I-III). Genomic DNA was purified from 200 µl of blood using a commercially available kit, Qiagen QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. For Paper III, we also collected peripheral blood from 165 healthy NSDTR dogs in Tempus Blood RNA tubes (Applied Biosystems, Foster City, CA). Samples were kept on ice during transportation and RNA was purified from total blood using the Tempus Spin RNA Isolation Reagent kit (Applied Biosystems) according to the manufacturer's instructions.

### 3.2.2 IIF ANA procedure

Serum samples were used for Indirect Immunofluorescence (IIF)-ANA tests. The diagnostic procedures were performed at the University Animal Hospital, SLU, Uppsala, Sweden using monolayers of HEp-2 cells fixed on glass slides (Immuno Concepts, Sacramento, CA). Slides were examined by fluorescence microscopy and the nuclear fluorescence pattern was used to subdivide a positive ANA result into either a homogenous or a speckled pattern (Hansson-

Hamlin *et al.*, 2006). The test was considered positive at a titre of ≥1:100. Positive ANA tests were divided into either a speckled ($ANA^S$) or homogenous ($ANA^H$) staining pattern and tests were repeated after two-three months. The ANA phenotype was assigned to a dog if two subsequent tests performed on samples taken from the same dog were consistent. Sera from all healthy controls were negative on the IIF-ANA test. Only one IIF-ANA test was performed on healthy dogs, because no dogs without clinical signs have ever showed a positive test.

### 3.2.3 DLA class II in ANA-positive dogs

In Paper I, we investigated the possible role of MHC class II as a genetic risk factor in IMRD and SRMA by sequence-based typing of the polymorphic exon 2 of the DLA-DRB1, -DQA1 and -DQB1 loci in the DLA class II region. In paper III we further correlated MHC class II genotype with the ANA-staining pattern in IMRD dogs. PCR (polymerase chain reaction) fragments containing DLA-DRB1, -DQA1 and -DQB1 exon 2 sequences were amplified, purified and sequenced using capillary-electrophoresis. Data analysis of the exon 2 nucleotide sequences for DLA-DRB1, -DQA1 and -DQB1 were performed with MatchTools and MatchToolsNavigator (Applied Biosystems). 2x2 Contingency Tables were used to calculate OR, 99% Confidence Interval and p-values. Cases and controls were divided into presence or absence of each allele, haplotype or genotype and the total number of individuals in each group was calculated. A comparison was made between the cases and controls between total number of alleles, haplotypes and genotypes in each group.

In Paper I, a total of 176 dogs including 51 IMRD (33 ANA-positive), 49 SRMA cases and 78 healthy controls (two dogs were affected by both IMRD and SRMA) were included in the analysis.

A total of five DLA-DRB1, four DLA-DQA1 and four DLA-DQB1 alleles were identified forming five different haplotypes and 13 genotypes. Haplotype 2 (DLA-DRB1*00601/DQA1*005011/DQB1*02001) occurred in increased frequency in all IMRD ANA-positive cases together compared to controls (54.5% vs. 34.6%, OR= 2.3). An even stronger association was found in all ANA-positive dogs homozygous for haplotype 2 (genotype 2) (48.5% vs. 11.5%, OR= 7.2). No association to DLA class II was observed for SRMA-affected dogs.

In the subsequent study (Paper III), we increased the number of NSDTR with IMRD and a positive ANA test. A total of 64 cases were ANA-positive (26 cases were classified as $ANA^H$ and 32 cases as $ANA^S$). Six of the 64 clinical cases were not analyzed for ANA subgroup due to lack of more serum. All 78 healthy control sera were negative for IIF ANA.

We investigated for DLA class II association and further dissected the phenotype by dividing the ANA-positive cases into $ANA^H$ and $ANA^S$ to determine whether dogs with different types of ANA pattern has differential genetic association to DLA. The same number of alleles, haplotypes and genotypes were observed as in our previous study (Paper I). A similar association in this larger sample set was observed for all ANA-positive dogs jointly for haplotype 2 (50.8% vs. 34.6%, OR= 2.0) and genotype 2 (45.3% vs. 11.5%, OR= 6.4). Among the $ANA^S$ dogs, 30 of 32 were either heterozygous or homozygous for haplotype 2 and 26 of those were homozygous for haplotype 2 (81.3%, OR= 33.2) implicating that homozygosity for the risk DLA class II haplotype results in increased risk of developing IMRD with the $ANA^S$ phenotype. No significant association was observed between the different haplotypes or genotypes of MHC class II and the cases with $ANA^H$ pattern.

We next removed the $ANA^S$ risk genotype and analyzed that data based on homozygosity. An increase in homozygosity was seen in $ANA^H$ cases compared to controls (62.5% vs. 13.0% OR= 11.1), implicating a general disadvantage due to homozygosity at DLA class II for $ANA^H$ dogs (Table 5).

*Table 5.* Homozygous frequencies show an association to both $ANA^S$ and $ANA^H$ dogs. Genotype 2 frequencies show an increased risk in $ANA^S$ dogs. By removing the risk genotype, a significant association is identified for $ANA^H$ dogs in general homozygosity at DLA. Significant associations are indicated in bold.

| | Tot pop % *n=142* | Controls % *n=78* | All cases % *n=64* | $ANA^S$ % *n=32* | $ANA^H$ % *n=26* |
|---|---|---|---|---|---|
| Homozygous | 45.8 (65) | 23.1 (18) | 73.4 (47) | **84.4 (27)** | **65.4 (17)** |
| Genotype 2 | 26.8 (38) | 11.5 (9) | 45.3 (29) | **81.3 (26)** | 7.7 (2) |
| | Tot pop % *n=104* | Controls % *n=69* | All cases % *n=35* | $ANA^S$ % *n=6* | $ANA^H$ % *n=24* |
| Homozygous no risk | 26.0 (27) | 13.0 (9) | 51.4 (18) | 16.7 (1) | **62.5 (15)** |

S= Speckled

H= Homogenous

### 3.2.4 GWAS identifies five additional loci

To perform a systematic and unbiased search of the entire dog genome for additional genetic risk factors a GWAS was conducted (Paper II). A case-control association analysis of 138 NSDTR from the Swedish and Finnish populations (44 SRMA dogs, 37 IMRD (where 22 were ANA-positive) and 57 healthy controls) was performed. All dogs used for the GWAS were unrelated

at the grandparental level. Genotyping was performed with the CanineSNP20 BeadChip panel that contains 22,000 validated SNPs.

We hypothesized that IMRD and SRMA might have both common and unique genetic risk factors and therefore examined the diseased dogs both as one group and for the classified sub-phenotypes, ANA-positive IMRD and SRMA separately. To test for the presence of stratification in our sample population we used PLINK (Purcell *et al.*, 2007) to make multidimensional scaling plots. We then produced scatter plots for two dimensions in which each spot corresponds to a specific individual. Some stratification was observed when all diseased dogs were analyzed together and for the SRMA-affected dogs alone. We therefore used PLINK to correct for population stratification based on clustering by identity-by-state, thereby accounting for population substructure to lower false positive rates and increase power.

Analysis of all cases (IMRD and SRMA) as one group identified a large region containing multiple associated SNPs on CFA 32 ($p_{raw}$= 1.5 x $10^{-5}$ and $p_{genome}$= 0.12). After correction for stratification, the region showed even stronger association ($p_{raw}$= 7.9 x $10^{-6}$ and $p_{genome}$= 0.06) (Figure 3).

Four highly associated regions were identified when analyzing ANA-positive IMRD dogs separately. The strongest association was found for one SNP on CFA 8 ($p_{raw}$= 1.5 x $10^{-6}$ and $p_{genome}$= 0.02) and a region containing multiple SNPs on CFA 24 ($p_{raw}$= 3.2 x $10^{-6}$ and $p_{genome}$= 0.04), both reaching genome-wide significance. We also found multiple associated SNPs on CFA 11 ($p_{raw}$= 7.4 x $10^{-6}$ and $p_{genome}$= 0.08) and on CFA 3 ($p_{raw}$= 2.2 x$10^{-5}$ $p_{genome}$= 0.18) (Figure 3). The associated peaks were the same when all IMRD-affected dogs were included (data not shown) as when only ANA-positive IMRD-affected dogs were used, but with a weaker degree of association.

When analyzing all SRMA-affected dogs separately, one region with multiple associated SNPs was identified on CFA 32 ($p_{raw}$= 7.10 x $10^{-6}$ and $p_{genome}$= 0.04) reaching genome-wide significance (Figure 3).

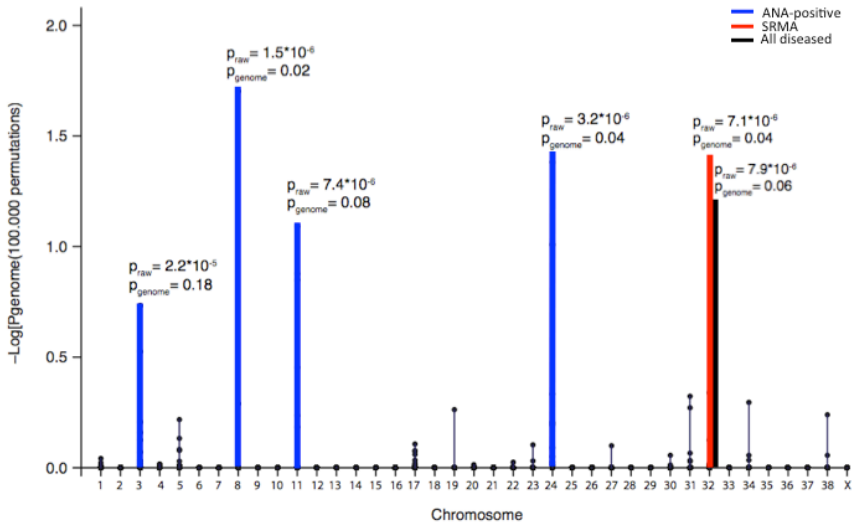


*Figure 3*. Five strongly associated regions were detected in the study population. One locus on CFA 32 was identified in all diseased dogs (black). Four loci on CFA 3, 8, 11 and 24 was associated with ANA-positive dogs (blue) and one loci (CFA 32) was identified in SRMA dogs (red).

The associated regions had large haplotype sizes (mostly ≈1 Mb in size), as expected within a canine breed (Lindblad-Toh *et al.*, 2005). However, a couple of regions were larger suggesting the possibility that genetic risk is caused by multiple nearby loci or extended LD within this region.

### 3.2.5 Fine-mapping verifies the loci

To validate associated loci from the GWAS, we performed fine-mapping (Paper II). Additional NSDTR and dogs from other breeds were included to reach a total number of 405 dogs. The sample consisted of 82 IMRD cases (32 ANA-positive, not divided into speckled and homogenous staining), 78 SRMA cases and 176 controls. Additional breeds, with individuals affected by corresponding diseases and control dogs, were included to allow identification of shared haplotypes across breeds. ANA-positive dogs were represented by 19 German shepherds (five cases and 14 controls) and 10 cocker spaniels (six cases and four controls), while for SRMA-affected dogs we used 30 boxers (10 cases and 20 controls) and 10 petite basset griffon Vendeen (four cases and six controls).

822 SNPs at ≈1 SNP/10 Kb density were genotyped. SNPs were analyzed with PLINK with MAF > 0.1 and call rate > 80%. The same settings were used to analyze haplotypes generated with a sliding window approach provided by PLINK.

Due to the low number of samples, examination for shared haplotypes was performed without prior association in the other breeds. However, they could be used for haplotype sharing across breeds.

For ANA-positive dogs, three of the four peaks showing association in the GWA analysis also showed increased p-values for association with the disease at the fine-mapping level, whereas one locus showed weaker association (CFA 8) in the fine-mapping analysis. The SNP showing the strongest genetic association was found on CFA 11 ($p_{raw}$= 8.7 x $10^{-13}$, OR= 7.9) and resides within a highly associated haplotype ($p_{raw}$= 1.3 x $10^{-12}$). The candidate region contains the *EPB41L4B*, *C9orf4* and *PTPN3* genes. The second best association results was a haplotype on CFA 24 ($p_{raw}$= 1.6x$10^{-12}$, OR= 5.0), including six genes, *AK128395*, *WFDC10B*, *WFDC13*, *AY372174*, *WFDC1* and *DNTTIP1*. The association on CFA 3 includes a haplotype ($p_{raw}$= 2.2 x $10^{-11}$ and OR= 4.5) containing seven genes, *AK126887*, *AP3B2*, *SCARNA15*, *FSD2*, *RPL23A*, *WHDC1L1* and *HOMER2*. The GWA hit on CFA 8 became lower after fine-mapping for ANA-positive dogs and was therefore not considered validated for ANA-positivity alone.

The region on CFA 32 also showed association to ANA-positive dogs alone ($p_{raw}$= 3.5 x $10^{-7}$, OR= 3.3). This is a large region (1.6 Mb) that includes several genes, *DAPP1*, *MAP2K1IP1*, *DNAJB14, DDIT4L, EMCN, PPP3CA* and *BANK1*.

For SRMA, the region on CFA 32 was validated ($p_{raw}$= 2.4 x $10^{-7}$, OR= 0.3) and contains the genes *DAPP1*, *MAP2K1IP1*, *DNAJB14* and *H2AFZ*. A novel association for SRMA dogs alone was identified on CFA 8 ($p_{raw}$= 3.2 x $10^{-7}$, OR= 2.6). The peak resides in a gene desert between the *SNRPE* and *VRK1* genes.

When all cases were analyzed together, the main peak of association on CFA 32 remained at approximately the same strength ($p_{raw}$=4.4 x $10^{-6}$, OR= 3.0). The region remains large and appears to comprise three peaks, positioned at the genes *DAPP1*, *PPP3CA* and *BANK1*. Since the association is present in multiple phenotypes, it is possible that it contains one or more risk alleles.

### 3.2.6 Targeted re-sequencing of candidate loci

We next performed targeted re-sequencing of associated regions (Paper III). The five regions (CFA 3, 8, 11, 24 and 32), spanning approximately 5Mb, were re-sequenced in seven individuals (four ANA-positive cases and three healthy

controls), using hybrid capture and 400-600 X coverage, with Illumina sequencing, in order to identify candidate mutations. The sequencing data were aligned with Burrows-Wheeler Aligner (BWA) (Li & Durbin, 2010). SAMtools (Li *et al.*, 2009) was used for SNP and InDel calling, BEDTools (Quinlan & Hall, 2010) for calculations of coverage statistics and finally SEQscoring (Truve *et al.*, 2011) to visualize the data and discover differences in the genomic sequence, such as SNPs, InDels and CNVs, between ANA-positive IMRD and healthy control dogs.

A total of 13,084 potential SNPs were detected and of these, 426 SNPs were located within or close (±5 bp) to a conserved element (Lindblad-Toh *et al.*, 2011; Garber *et al.*, 2009). 2780 possible InDels were detected, among those, 88 occurred within or close to a conserved element (±5 bp). To identify structural variations, such as larger insertions, CNV or deletions SEQscoring (Truve *et al.*, 2011) was used to calculate coverage differences between cases and controls. We did not identify any structural variants that differed between cases and controls.

### 3.2.7 $ANA^S$ association to DLA class II, CFA 11 and 32

We next combined the results from DLA class II association study with the additional identified risk loci (Paper III). Twenty-five of the $ANA^S$-positive NSDTRs with MHC class II genotype 2 were included for additional analysis as well as 145 healthy controls and a total of 63 ANA-positive dogs (regardless of ANA-staining pattern). 384 SNPs for the five identified loci (CFA 3, 8, 11, 24 and 32) were chosen from the next generation sequencing data set. The SNPs were chosen based on the following criteria:

- Difference in allele frequency in cases compared to controls
- Positioned in either protein coding regions, 5' UTR or 3' UTR
- Located within non-coding conserved elements

Conserved elements were identified using comparative sequence analysis based on the analysis of 29 mammals using SiPhy (Lindblad-Toh *et al.*, 2011; Garber *et al.*, 2009). We examined the GWA risk loci on CFA 3, 8, 11, 24 and 32 for association to the speckled phenotype ($ANA^S$).

Strong associations to CFA 11 and 32 were observed in DLA haplotype 2 homozygous $ANA^S$ dogs. None of the other chromosomal regions showed a significant association. The region on CFA 11 covering the *PTPN3* gene comprised a haplotype of 15 SNPs that was highly associated to $ANA^S$ dogs (most associated SNP, $p= 8.2 \times 10^{-07}$, OR= 5.7). The most associated region on CFA 32 ($p= 1.5 \times 10^{-08}$-$4.7 \times 10^{-05}$ and OR= 3.4-6.8) is 1.3 Mb in size and 23

SNPs show significant p-values. The region contains several genes, including *DAPP1*, *MAP2K1IP1*, *DNAJB14*, *H2AFZ*, *DDIT4L*, *EMCN, PPP3CA* and *BANK1* and the most associated SNP (32:24556037) is located upstream of *DAPP1*. One highly associated haplotype comprising two SNPs in high LD was also identified (32:24556037 and 32:25485961).

Similar levels of association were observed between the CFA 11 region and $ANA^S$ dogs, implying that this region is more important for $ANA^S$ dogs since fewer dogs were used to achieve the same statistical significance for the association. The region on CFA 32 shows more significant P-values in the $ANA^S$ group compared to all ANA, suggesting that this region is most important for IMRD $ANA^S$–positive dogs.

### 3.2.8 Multi locus analysis

A multi locus analysis was performed to analyze the combined contribution of other risk factors at MHC class II, CFA 11 and 32. In total we had 56 healthy controls and 26 IMRD dogs with $ANA^S$–positivity with genotype data on all three loci (MHC class II, CFA 11 and 32). Among the 56 healthy controls, nine dogs were homozygous for the MHC risk genotype 2 and the other had either one copy of it (23 dogs) or none (24 dogs). 44.4% lacked additional risk loci, 33.3% had one extra risk locus, 22.2% had two extra risk loci and none had three additional risk loci.

Among IMRD $ANA^S$-positive dogs, 23 out of 26 were homozygous for haplotype 2, one was heterozygous and the other two had no MHC class II risk. Among the dogs homozygous for haplotype 2, 39.1% had zero additional risk factors, 17.4% had one additional risk factor, 34.8% had two additional risk factors and 8.7% had three additional risk factors.

The nine control dogs with MHC risk had a total of seven risk haplotypes for the additional loci, which gives an additional contribution of 19.4%. There were 26 additional risk haplotypes out of the 23 $ANA^S$ homozygous dogs which gives and additional contribution of 28.3%.

### 3.2.9 Expression studies of candidate genes

Peripheral blood mononuclear cells (PBMC) were collected from 165 healthy dogs to prepare RNA to be used to measure the mRNA expression levels for all candidate genes within associated regions on CFA 11 (*PTPN3*) and CFA 32 (*DAPP1*, *MAP2K1IP1*, *DNAJB14, DDIT4L, EMCN, PPP3CA* and *BANK1*). These genes were selected because they are located at the highest associated SNP or within the most associated haplotype. The dogs were genotyped for risk and non-risk SNPs/haplotypes and correlated to the level of gene expression.

Three SNPs (11:67,516,041, 11:67,538,032 and 11:67,538,806) in the highly associated 15 SNP haplotype on CFA 11, spanning the *PTPN3* gene, were genotyped in the healthy dogs. No dogs homozygous for the whole risk haplotype on CFA 11 were present in the healthy group. Among the $ANA^S$ patients in the genetic study, only four dogs were homozygous for the risk haplotype on CFA 11 and all of them showed very severe IMRD that eventually led to death. Therefore, a SNP in the close vicinity of the haplotype (also associated $p= 7.5 \times 10^{-4}$) where all three genotypes occur (position 11:67583604) was genotyped. This four SNP haplotype was used for correlation with expression. An 8-fold downregulation of *PTPN3* mRNA levels (p= 0.013) was identified in dogs carrying the risk haplotype.

All genes on the large region (1.3 Mb) on chromosome 32 were also analyzed for differential expression and correlated to genetic association. Two of them, *DDIT4L* (3-fold, p=0.0002) and *BANK1* (1-fold, p=0.006), show an upregulation of mRNA expression related to the two SNP haplotype (32:24556037 and 32:25485961). A modest upregulation of *DAPP1* and *PPP3CA* is correlated with the SNP (32:24890208), which is included in a risk haplotype but not associated alone.

## 3.3 Discussion

In Paper I we performed a case/control candidate gene approach by sequencing the polymorphic exon 2 of the DLA-DRB1, -DQB1 and –DQA1 class II genes. We also (Paper II) performed a GWAS followed by fine-mapping in an effort to identify additional genetic risk factors. MHC class II was identified as a risk factor for both $ANA^S$ and $ANA^H$ IMRD dogs but not for SRMA dogs. Four additional risk factors for the ANA-positive phenotype were identified by a GWAS (CFA 3, 11, 24 and 32), where CFA 11 and 32 were associated to the speckled phenotype (Paper III). Expression studies revealed that *PTPN3*, *DDIT4L* and *BANK1* have different expression according to risk/non risk haplotype identified for $ANA^S$ dogs (Paper III) and that several different cell types, such as T-cells, B-cells and possibly macrophages are involved.

We demonstrated that the dog is an excellent model for studying complex genetic disease and that only few cases and controls are needed to identify genetic risk factors underlying a complex genetic disease.

### 3.3.1 DLA class II is a genetic risk factor

In Paper I and III, we identified MHC class II as an important genetic risk factor for ANA-positive IMRD dogs. A particular homozygous risk haplotype was identified in $ANA^S$ dogs and a general homozygous disadvantage was

found in $ANA^H$ dogs. The observed association with an OR at 33 is one of the highest reported for an autoimmune disease and MHC class II, suggesting that MHC class II is probably of major importance for development of IMRD with the $ANA^S$ phenotype in dogs.

Only two $ANA^S$ dogs lack haplotype 2 and both are slightly atypical IMRD cases. One of the cases showed clinical signs from multiple organ systems in a way not usually observed in NSDTRs diagnosed with IMRD. Clinical signs shown by this dog were skin necrosis on ear flaps and around the mouth, weight loss and muscle atrophy of the head muscles, apart from the joint pain and stiffness indicating a rheumatic disease. Finally this patient was euthanized due to breathing problems. This dog had MHC class II genotype 6 and only the risk factor at chromosome 11. The other dog was diagnosed with SRMA at 10 months of age and it was treated with corticosteroids and recovered completely. Later on, at two years of age, the dog was diagnosed with IMRD and had become ANA-positive. In this material, this is the only dog suffering from both SRMA and IMRD (data not shown). This dog was homozygous for MHC class II (genotype 3) and did not carry the risk factors at chromosome 11 and 32.

Five DLA-DRB1, four DLA-DQA1 and four DLA-DQB1 alleles forming a total of five different haplotypes was observed in the Scandinavian NSDTR population. In the US and Canada the same number of alleles have been detected forming a total of seven different haplotypes (Hughes *et al.*, 2010). The two extra haplotypes identified only differ in DQB1 (DLA-DRB1*01501/DQA1*00601/DQB1*02301 and DLA-DRB1*01502/ DQA1*00601/DQB1*00301 (Hughes *et al.*, 2010) compared to the haplotypes previously identified in the Scandinavian population (Paper I) suggesting that a recombination event has occurred in the North American population of NSDTR.

A similar haplotype to the $ANA^S$-positive risk, where the DLA-DQB1 allele differs (DLA-DRB1*00601/DQA1*005011/DQB1*00701), has been shown to predispose to another immune-mediated disease, immune-mediated hemolytic anemia in a number of different breeds (Kennedy *et al.*, 2006a). This DLA-DRB1*00601 allele contains the five amino acid epitope RARAA known as the shared epitope for rheumatoid arthritis in human (Gregersen *et al.*, 1987). Dogs affected by rheumatoid arthritis have an increased risk for developing disease when displaying the RARAA epitope from various DLA-DRB1 alleles (Ollier *et al.*, 2001). In human SLE patients, a similar epitope as the RA-shared epitope (QARAA) is found to be associated with SLE (Tsuchiya *et al.*, 2001).

The identification of MHC class II as genetic risk factor provides further support that IMRD is an autoimmune disease and that it represents an

important model for human systemic rheumatic autoimmune diseases, like SLE since MHC class II is involved also as a risk factor for development of human SLE (Smerdel-Ramoya *et al.*, 2005; Graham *et al.*, 2002).

The general homozygous disadvantage observed may be due to competitive binding actions of MHC class II molecules, but homozygosity at DLA by itself is not sufficient to disease development.

### 3.3.2 Additional loci identified genetically

In Paper II, we demonstrated the power of mapping a complex disease in the dog by using less than 100 cases and 100 controls, which was the first complex disease mapped in the dog. Five candidate loci were identified (four for IMRD, one for SRMA and one shared locus). Three of these were strongly validated by fine-mapping in IMRD dogs (CFA 3, 11 and 24) and CFA 32 were validated both as separate risk factors for the diseases and combined. We also showed that ANA-positive dogs with a speckled phenotype ($ANA^S$) are associated to the loci in CFA 11 and 32 and that all ANA-positive dogs, regardless of ANA phenotype are associated to CFA 3 and 24.

Haplotypes are extended over more than Mb lengths within a dog breed and in order to break down the long haplotypes, an additional breed with the same phenotype is used (Karlsson & Lindblad-Toh, 2008). This is however a challenge when using the NSDTR, because to our knowledge, no related breeds share such a strong predisposition for the SLE-related disease complex, which makes it difficult to identify the causative mutation and shorten down the long haplotype blocks. A closely related breed could instead be used to break down long associated haplotypes by identifying shared non-risk loci.

The region on CFA 11 contains the *PTPN3* gene only and the region on CFA 32 is large (1.3 Mb in size), spanning several genes. However, three excellent candidate genes occur on the CFA 32 region, *DAPP1*, *PPP3CA* and *BANK1*, of which *BANK1* has already been associated with human SLE (Kozyrev *et al.*, 2008).

A multi locus analysis was performed to determine whether $ANA^S$ cases homozygous for the risk MHC class II in general also carried more risk loci at the other identified chromosomes (CFA 11 and 32). There was a trend towards an accumulation of risk haplotypes in cases vs. controls (28.3% and 19.4%). Since sample material is small, no statistically significant results were obtained, however, an accumulation towards more risk genotypes in total is observed in $ANA^S$ dogs compared to controls. This supports the finding that these three loci jointly contribute to the disease risk and that all loci should be taken into account and not only MHC class II. This implicates a multifactorial

disease with incomplete penetrance where environmental triggers are likely to play a crucial role.

Further analyses are needed to explore whether the dogs that carry risk alleles at all three loci develop a more severe form of disease than dogs only carrying risk alleles at a few of the loci. Further studies are also required to understand how risk factors interact with unknown environmental risk factors for this immune-mediated disease complex. This study has, however, provided insight to potential multi-locus contributions to IMRD development.

### 3.3.3 A novel pathway suggested

Gene expression studies correlated to genetic association to $ANA^S$ dogs revealed that the *PTPN3* (protein tyrosine phosphatase nonreceptor 3) gene on CFA 11 is downregulated in dogs with the risk haplotype. This gene belongs to the same family as *PTPN22*, which is a major genetic risk factor for many different autoimmune diseases including SLE (Chung & Criswell, 2007).

It has been suggested that the phosphatase encoded by the *PTPN3* gene participate in TCR-signaling as a negative regulator by dephosphorylating the TCR, which downstream inhibits activation of nuclear factor of activated T-cells (NF-AT) (Sozio *et al.*, 2004; Han *et al.*, 2000).

The CFA 32 region contains several genes whose mRNA levels are upregulated in dogs carrying the risk haplotype. Because of the canine distemper virus outbreak (CDV) that occurred in the beginning of 1900s (Strang & MacMillan, 1996), there might have been a selection for this whole region in dogs that survived the CDV outbreaks, since this region contain several genes important for immune function and regulation.

The top SNP from the genotyping (32:24556037) correlates with differential mRNA expression of *DDIT4L* and *BANK1*. The function of the protein encoded by *DDIT4L* (DNA-damage-inducible transcript 4-like) is relatively unknown but it may be involved in negative regulation of mammalian target of rapamycin pathway (mTOR), which has a fundamental role in cell growth control (Corradetti *et al.*, 2005).

*BANK1* (B-cell scaffold protein with ankyrin repeats) is an interesting gene extensively studied in human SLE. *BANK1* was found to be associated to SLE-development in several independent case-control data sets (Yang *et al.*, 2010; Chang *et al.*, 2009; Gateva *et al.*, 2009; Guo *et al.*, 2009; Kozyrev *et al.*, 2008).

A modest upregulation was identified for *DAPP1* and *PPP3CA* mRNA expression in $ANA^S$–positive dogs. These genes encode proteins that are essential for both TCR- and B-cell receptor (BCR)-mediated immune responses and cell proliferation and may act via the NF-AT pathway.

*DAPP1* (dual adapter for phosphotyrosine and 3-phosphotyrosine and 3-phosphoinositide) is expressed in both T- and B-cells (Sommers *et al.*, 2008) and indirectly inhibit both TCR- and BCR-signaling (Ortner *et al.*, 2011). Consequently, it is plausible that they may influence activation of NF-AT (Marshall *et al.*, 2000).

*PPP3CA* (protein phosphatase 3, catalytic subunit, alpha isozyme) encodes the catalytic subunit of calcineurin. This gene is an important and recognized target of two important immunosuppressive drugs (cyclosporine A and FK506) (Guerini, 1997; Clipstone & Crabtree, 1992) and has been reported to be differentially expressed in patients with SLE (Kyttaris *et al.*, 2007).

NF-AT proteins constitute a family of transcription factors with a key regulatory function of immunological tolerance and can cause development of autoimmune diseases if they are regulated incorrectly. The NF-AT family consists of five members NFATc1, NFATc2, NFATc3, NFATc4, and NFAT5. NFATc1 through NFATc4 are regulated by calcium signaling by calcineurin (Serfling *et al.*, 2006). Several of our associated genes show correlation to the NF-AT pathway. Importantly, several of the identified genetic risk factors are crucial for T-cell activation, B-cell activation and professional antigen-presenting cells like macrophages. It is thus conceivable that patients with a more severe form of IMRD carry multiple of these genetic risk factors or particular disadvantageous combinations of risk factors. Our findings also implicate that the NF-AT pathway may be a critical pathway for IMRD. Future studies may reveal a possibility for NF-AT involvement also in a sub-type of human SLE.

### 3.3.4 Same disease complex or not?

When this study was initiated it was unclear whether IMRD and SRMA belongs to the same disease complex or not. Firstly, we showed that DLA class II is only a risk factor for IMRD but not SRMA (Paper I). Secondly, we identified a shared genetic risk locus on CFA 32 between the diseases and additional three that were specific for IMRD (Paper II). Based on our investigation, it is clear that there is one common genetic risk locus (CFA 32) and a number of different genes involved as specific risk factors for the different diseases. All of the identified risk loci contain genes, which are known to be important for immune function and in particular for the NF-AT pathway. It is therefore clear that it is not the same disease, but still unclear if the diseases belong to the same complex or whether they should be considered as largely different diseases.

# 4 Conclusions

The results of this thesis contribute to an overall increased knowledge concerning the genetic background for the development of canine SLE-related disease. It also highlights the dog as an excellent model organism for mapping genetic risk factors for complex traits as the deeper understanding obtained might be of major importance for human rheumatic disease as well. The main conclusions were as follows:

- DLA class II is associated with the canine SLE-related disease IMRD but not with SRMA. More specifically, homozygosity for a specific DLA class II haplotype confers increased risk for dogs with IMRD with $ANA^{S}$ phenotype, whereas a general homozygosity at DLA class II gives $ANA^{H}$ dogs an increased risk for developing disease.

- Genome-wide association mapping followed by fine-mapping and regional re-sequencing identified five additional risk loci for the SLE-related IMRD (CFA 3, 8, 11, 24 and 32). The loci on CFA 3, 11 and 32 were associated with ANA-positive dogs and the locus on 32 was shared between the two phenotypes.

- Most $ANA^{S}$ dogs homozygous for the DLA risk haplotype also carried the genetic risk factors at CFA 11 and 32.

- mRNA expression studies revealed that the *PTPN3* gene (CFA 11) is significantly downregulated and that *DDIT4L* and *BANK1* mRNA expression (CFA 32) is significantly upregulated.

- Several of the identified genetic risk factors are crucial for T-cell activation, B-cell activation and professional antigen-presenting cells like macrophages.

# 5 Future prospects

## 5.1 DLA class II

The degree of LD in the MHC class II region needs to be defined in the NSDTR population. It cannot be ruled out that other genes in close proximity to this region are of importance for IMRD development. However, as previously described, many other studies of autoimmune diseases in both human and animals indicate that MHC class II is directly involved as a major genetic risk factor for such diseases.

The autoantigens that bind to the MHC class II molecules expressed from the risk haplotype remains to be defined. Functional studies remain to determine whether MHC class II molecules observed in dog patients with autoimmune rheumatic diseases present a common autoantigen that binds MHC class II molecules containing the RA epitope.

## 5.2 Identification of mutations and downstream targets

Further genetic studies are needed, especially to explore the $ANA^{H}$ phenotype and shared risk loci for all ANA dogs (regardless of staining). The associated regions on CFA 3 and 24 seem to belong to all ANA dogs (regardless of staining), but due to a low sample number in $ANA^{H}$ dogs, it is possible that we have missed to detect separate risk factors for this phenotype. A study using more dogs should be conducted.

One of the highest priority aims is to identify the exact disease-causing mutations. Because the mutations could be either coding or regulatory they will have different functional consequences and accordingly different functional studies will be needed to define their consequences and involvement in disease development. They can be coding (synonymous or non-synonymous), affect RNA stability, be located in a promoter or enhancer region, in a conserved

element or in transcription factor binding sites. Different techniques are needed depending on the nature of identified mutations. If identified mutations occur in a coding region, the functional effect will be studied using cell culture methods by overexpressing the genes with the mutation and study the functional consequences. Regulatory mutations can be studied by quantitative PCR (to evaluate the expression) and allele-specific PCR. Mutations in promoter and enhancer regions can be evaluated for functional consequence on transcription using transient transfections in relevant cell culture systems with reporters carrying either the wild-type or mutated regulatory regions.

Pathway analysis is another informative way to study the downstream effects of mutations. We are currently collecting fresh PBMC from healthy dogs and dogs in the acute phase of the disease, before treatment is started, to perform mRNA pathway studies. The goal is to identify downstream targets by RNA sequencing (RNAseq). We also aim to sort PBMC and perform RNAseq in different cell types (such as B-cells, Th1 and Th2 cells and dendritic cells) to evaluate if certain cell types have differentially expressed genes.

We have shown that interactions between three loci occur in the $ANA^{S}$ phenotype. It would therefore be interesting to elucidate these three interactions in a cell model and study the mechanisms of *e.g.* NF-AT and T-cell activation.

## 5.3 Characterize the ANA autoantigen and cytokine changes

Canine ANA specificity determination is currently usually restricted to a speckled (non-chromosomal) or homogenous (chromosomal) staining pattern. A future goal is to identify subgroups of ANA-reactivity by identifying specific autoantigens. A novel autoantigen can be used as a biomarker in diagnostic procedures and also help us to further understand the disease mechanism by correlating it to genetic studies.

An imbalance in cytokine production and cytokine levels relates with SLE development in humans (Su *et al.*, 2012; Lee *et al.*, 2010). Identification of cytokines and possible alterations will therefore be examined in blood samples from NSDTRs with immune-mediated diseases. The ultimate ambition is to improve treatment and help breeding of healthier dogs in general and Nova Scotia duck tolling retrievers in particular.

## 5.4 Genes and pathways identified in dogs in human SLE patients

Our plan is to study the genes and pathways found in the canine SLE-related disease, in a cohort of human SLE patients divided into sub-phenotypes. We have designed two arrays, a smaller and a larger. The smaller array is 5.7 Mb and covers the genes identified in our dog candidate regions, NF-AT pathway and other known "SLE-genes" of interest. The larger array covers 29.4 Mb and consists of 1300 genes important for our different dog-immune projects, including diabetes, Addison's disease, Shar-Pei fever, Lymphocytic thyroiditis, and atopic dermatitis. Specific genes involved in the NF-AT pathway, autoimmunity in general, T-cell activation, B-cell activation, interferons, interleukins, tumor necrosis factors, toll-like receptors, genes in the hyaluronic pathway and the complement system are covered. Both designs capture ± 2 Kb relative of identified transcription start sites, ± 20 bp of exon borders and 3' UTR, and all conserved elements located ±100 Kb upstream and downstream of the genes and within the gene.

A pilot trial was performed using the smaller array and a total of 124 patients and 17 controls. Patients were divided into nine different pools (11-17 individuals/pool) depending on their clinical or immunological sub-phenotype and analyzed accordingly. We selected distinctive patient groups (such as patients with nephritis, skin manifestations or presence of anti-dsDNA antibodies) and tested if rare or common gene variants or pathways occur more frequently in specific subgroups. A large phenotypic diversity occurs in human SLE patients and it would therefore be helpful to sub-categorize SLE patients. We have identified NF-AT as a new major pathway for the canine SLE-like disease and this pathway may also play a role in the development of SLE or other rheumatoid human diseases.

The pools had an average coverage of 3775 X and less than 0.3% of the tiled region missed sequence data. We identified a large number of novel SNPs, with many variations that only occurred in the case pools, suggesting that these might represent novel genes important for specific phenotypes of SLE. However, data need to be confirmed in a larger sample cohort as well as SNP genotyping as this pilot trial only gives us a hint of novel pathways and genes.

The ultimate goal is to develop new diagnostic and treatment regimes for patients based on identification of new important genetic risk factors by subdividing patients based on both different clinical and genetic appearances.

# References


Alarcon-Segovia, D., Alarcon-Riquelme, M.E., Cardiel, M.H., Caeiro, F., Massardo, L., Villa, A.R. & Pons-Estel, B.A. (2005). Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. *Arthritis Rheum* 52(4), 1138-47.

Albert, E.D., Bauer, M.P. & Mayr, W.R. (1985). *Histocompatibility Testing 1984*. Berlin, Heidelberg: Springer Verlag.

Andersson, G., Svensson, A.C., Setterblad, N. & Rask, L. (1998). Retroelements in the human MHC class II region. *Trends Genet* 14(3), 109-14.

Anfinsen, K.P., Berendt, M., Liste, F.J., Haagensen, T.R., Indrebo, A., Lingaas, F., Stigen, O. & Alban, L. (2008). 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* 72(4), 350-5.

Angles, J.M., Famula, T.R. & Pedersen, N.C. (2005). Uveodermatologic (VKH-like) syndrome in American Akita dogs is associated with an increased frequency of DQA1*00201. *Tissue Antigens* 66(6), 656-65.

Axelsson, E., Ratnakumar, A., Arendt, M.L., Maqbool, K., Webster, M.T., Perloski, M., Liberg, O., Arnemo, J.M., Hedhammar, A. & Lindblad-Toh, K. (2013). The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 495, 360-4.

Bach, F.H. (1985). The HLA class II genes and products: the HLA-D region. *Immunol Today* 6(3), 89-94.

Bari, A.S., Carter, S.D., Bell, S.C., Morgan, K. & Bennett, D. (1989). Anti-type II collagen antibody in naturally occurring canine joint diseases. *Br J Rheumatol* 28(6), 480-6.

Barnes, A., O'Neill, T., Kennedy, L.J., Short, A.D., Catchpole, B., House, A., Binns, M., Fretwell, N., Day, M.J. & Ollier, W.E. (2009). Association of canine anal furunculosis with TNFA is secondary to linkage disequilibrium with DLA-DRB1*. *Tissue Antigens* 73(3), 218-24.

Barrio, A.M., Ekerljung, M., Jern, P., Benachenhou, F., Sperber, G.O., Bongcam-Rudloff, E., Blomberg, J. & Andersson, G. (2011). The first sequenced carnivore genome shows complex host-endogenous retrovirus relationships. *PLoS One* 6(5), e19832.

Benito-Garcia, E., Schur, P.H. & Lahita, R. (2004). Guidelines for immunologic laboratory testing in the rheumatic diseases: anti-Sm and anti-RNP antibody tests. *Arthritis Rheum* 51(6), 1030-44.

Bexfield, N.H., Watson, P.J., Aguirre-Hernandez, J., Sargan, D.R., Tiley, L., Heeney, J.L. & Kennedy, L.J. (2012). DLA class II alleles and haplotypes are associated with risk for and protection from chronic hepatitis in the English Springer spaniel. *PLoS One* 7(8), e42584.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. & Wiley, D.C. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329(6139), 506-12.

Bontrop, R.E., Otting, N., de Groot, N.G. & Doxiadis, G.G. (1999). Major histocompatibility complex class II polymorphisms in primates. *Immunol Rev* 167, 339-50.

Borghans, J.A., Beltman, J.B. & De Boer, R.J. (2004). MHC polymorphism under host-pathogen coevolution. *Immunogenetics* 55(11), 732-9.

Burlingame, R.W. & Rubin, R.L. (1991). Drug-induced anti-histone autoantibodies display two patterns of reactivity with substructures of chromatin. *J Clin Invest* 88(2), 680-90.

Burton, S., DeLay, J., Holmes, A., Somerville, C., Eye, J., Shaw, D., Wack, O. & Hanna, P. (1997). Hypoadrenocorticism in young related Nova Scotia duck tolling retrievers. *Can Vet J* 38(4), 231-4.

Carrington, M. & O'Brien, S.J. (2003). The influence of HLA genotype on AIDS. *Annu Rev Med* 54, 535-51.

Catchpole, B., Ristic, J.M., Fleeman, L.M. & Davison, L.J. (2005). Canine diabetes mellitus: can old dogs teach us new tricks? *Diabetologia* 48(10), 1948-56.

Chambers, J.C., Kenan, D., Martin, B.J. & Keene, J.D. (1988). Genomic structure and amino acid sequence domains of the human La autoantigen. *J Biol Chem* 263(34), 18043-51.

Chan, E.K., Hamel, J.C., Buyon, J.P. & Tan, E.M. (1991). Molecular definition and sequence motifs of the 52-kD component of human SS-A/Ro autoantigen. *J Clin Invest* 87(1), 68-76.

Chang, Y.K., Yang, W., Zhao, M., Mok, C.C., Chan, T.M., Wong, R.W., Lee, K.W., Mok, M.Y., Wong, S.N., Ng, I.O., Lee, T.L., Ho, M.H., Lee, P.P., Wong, W.H., Lau, C.S., Sham, P.C. & Lau, Y.L. (2009). Association of BANK1 and TNFSF4 with systemic lupus erythematosus in Hong Kong Chinese. *Genes Immun* 10(5), 414-20.

Chase, K., Sargan, D., Miller, K., Ostrander, E.A. & Lark, K.G. (2006). Understanding the genetics of autoimmune disease: two loci that regulate late onset Addison's disease in Portuguese Water Dogs. *Int J Immunogenet* 33(3), 179-84.

Chung, S.A. & Criswell, L.A. (2007). PTPN22: its role in SLE and autoimmunity. *Autoimmunity* 40(8), 582-90.

Chung, S.A., Taylor, K.E., Graham, R.R., Nititham, J., Lee, A.T., Ortmann, W.A., Jacob, C.O., Alarcon-Riquelme, M.E., Tsao, B.P., Harley, J.B., Gaffney, P.M., Moser, K.L., Petri, M., Demirci, F.Y., Kamboh, M.I., Manzi, S., Gregersen, P.K., Langefeld, C.D., Behrens, T.W. & Criswell, L.A. (2011). Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production. *PLoS Genet* 7(3), e1001323.

Clipstone, N.A. & Crabtree, G.R. (1992). Identification of calcineurin as a key signalling enzyme in T-lymphocyte activation. *Nature* 357(6380), 695-7.

Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium (1999). *Nature* 401(6756), 921-3.

Corradetti, M.N., Inoki, K. & Guan, K.L. (2005). The stress-inducted proteins RTP801 and RTP801L are negative regulators of the mammalian target of rapamycin pathway. *J Biol Chem* 280(11), 9769-72.

Criswell, L.A., Pfeiffer, K.A., Lum, R.F., Gonzales, B., Novitzke, J., Kern, M., Moser, K.L., Begovich, A.B., Carlton, V.E., Li, W., Lee, A.T., Ortmann, W., Behrens, T.W. & Gregersen, P.K. (2005). Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. *Am J Hum Genet* 76(4), 561-71.

Cui, Y., Sheng, Y. & Zhang, X. (2013). Genetic susceptibility to SLE: Recent progress from GWAS. *J Autoimmun*.

Cunninghame Graham, D.S., Graham, R.R., Manku, H., Wong, A.K., Whittaker, J.C., Gaffney, P.M., Moser, K.L., Rioux, J.D., Altshuler, D., Behrens, T.W. & Vyse, T.J. (2008). Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. *Nat Genet* 40(1), 83-9.

Danchenko, N., Satia, J.A. & Anthony, M.S. (2006). Epidemiology of systemic lupus erythematosus: a comparison of worldwide disease burden. *Lupus* 15(5), 308-18.

Debenham, S.L., Hart, E.A., Ashurst, J.L., Howe, K.L., Quail, M.A., Ollier, W.E. & Binns, M.M. (2005). Genomic sequence of the class II region of the canine MHC: comparison with the MHC of other mammalian species. *Genomics* 85(1), 48-59.

Ding, Z.L., Oskarsson, M., Ardalan, A., Angleby, H., Dahlgren, L.G., Tepeli, C., Kirkness, E., Savolainen, P. & Zhang, Y.P. (2012). Origins of domestic dog in southern East Asia is supported by analysis of Y-chromosome DNA. *Heredity (Edinb)* 108(5), 507-14.

Dixon, R.M. & Mooney, C.T. (1999). Canine serum thyroglobulin autoantibodies in health, hypothyroidism and non-thyroidal illness. *Res Vet Sci* 66(3), 243-6.

Dyggve, H., Kennedy, L.J., Meri, S., Spillmann, T., Lohi, H. & Speeti, M. (2011). Association of Doberman hepatitis to canine major histocompatibility complex II. *Tissue Antigens* 77(1), 30-5.

Ellegren, H., Hartman, G., Johansson, M. & Andersson, L. (1993). Major histocompatibility complex monomorphism and low levels of DNA fingerprinting variability in a reintroduced and rapidly expanding population of beavers. *Proc Natl Acad Sci U S A* 90(17), 8150-3.

Evans, J.P., Brinkhous, K.M., Brayer, G.D., Reisner, H.M. & High, K.A. (1989). Canine hemophilia B resulting from a point mutation with unusual consequences. *Proc Natl Acad Sci U S A* 86(24), 10095-9.

Fernando, M.M., Stevens, C.R., Walsh, E.C., De Jager, P.L., Goyette, P., Plenge, R.M., Vyse, T.J. & Rioux, J.D. (2008). Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet* 4(4), e1000024.

Fogle, B., Morgan, T. & Fogle, B.-. (2000). *The New Encyclopedia of the Dog*. New York: Dorling Kindersley.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S.N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E.S., Daly, M.J. & Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *Science* 296(5576), 2225-9.

Garber, M., Guttman, M., Clamp, M., Zody, M.C., Friedman, N. & Xie, X. (2009). Identifying novel constrained elements by exploiting biased substitution patterns. *Bioinformatics* 25(12), i54-62.

Gateva, V., Sandling, J.K., Hom, G., Taylor, K.E., Chung, S.A., Sun, X., Ortmann, W., Kosoy, R., Ferreira, R.C., Nordmark, G., Gunnarsson, I., Svenungsson, E., Padyukov, L., Sturfelt, G., Jonsen, A., Bengtsson, A.A., Rantapaa-Dahlqvist, S., Baechler, E.C., Brown, E.E., Alarcon, G.S., Edberg, J.C., Ramsey-Goldman, R., McGwin, G., Jr., Reveille, J.D., Vila, L.M., Kimberly, R.P., Manzi, S., Petri, M.A., Lee, A., Gregersen, P.K., Seldin, M.F., Ronnblom, L., Criswell, L.A., Syvanen, A.C., Behrens, T.W. & Graham, R.R. (2009). A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet* 41(11), 1228-33.

Gorer, P.A. (1937). The genetic and antigenic basis of tumour transplantation. *The Journal of Pathology and Bacteriology* 44(3), 691-7.

Graham, R.R., Ortmann, W.A., Langefeld, C.D., Jawaheer, D., Selby, S.A., Rodine, P.R., Baechler, E.C., Rohlf, K.E., Shark, K.B., Espe, K.J., Green, L.E., Nair, R.P., Stuart, P.E., Elder, J.T., King, R.A., Moser, K.L., Gaffney, P.M., Bugawan, T.L., Erlich, H.A., Rich, S.S., Gregersen, P.K. & Behrens, T.W. (2002). Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. *Am J Hum Genet* 71(3), 543-53.

Greer, K.A., Wong, A.K., Liu, H., Famula, T.R., Pedersen, N.C., Ruhe, A., Wallace, M. & Neff, M.W. (2010). Necrotizing meningoencephalitis of Pug dogs associates with dog leukocyte antigen class II and resembles acute variant forms of multiple sclerosis. *Tissue Antigens* 76(2), 110-8.

Gregersen, P.K., Silver, J. & Winchester, R.J. (1987). The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum* 30(11), 1205-13.

Guerini, D. (1997). Calcineurin: not just a simple protein phosphatase. *Biochem Biophys Res Commun* 235(2), 271-5.

Guo, L., Deshmukh, H., Lu, R., Vidal, G.S., Kelly, J.A., Kaufman, K.M., Dominguez, N., Klein, W., Kim-Howard, X., Bruner, G.R., Scofield, R.H., Moser, K.L., Gaffney, P.M., Dozmorov, I.M., Gilkeson, G.S., Wakeland, E.K., Li, Q.Z., Langefeld, C.D., Marion, M.C., Williams, A.H., Divers, J., Alarcon, G.S., Brown, E.E., Kimberly, R.P., Edberg, J.C., Ramsey-Goldman, R., Reveille, J.D., McGwin, G., Jr., Vila, L.M., Petri, M.A., Vyse, T.J., Merrill, J.T., James, J.A., Nath, S.K., Harley, J.B. & Guthridge, J.M. (2009). Replication of the BANK1 genetic association with systemic lupus erythematosus in a European-derived population. *Genes Immun* 10(5), 531-8.

Han, J.W., Zheng, H.F., Cui, Y., Sun, L.D., Ye, D.Q., Hu, Z., Xu, J.H., Cai, Z.M., Huang, W., Zhao, G.P., Xie, H.F., Fang, H., Lu, Q.J., Xu, J.H., Li, X.P., Pan, Y.F., Deng, D.Q., Zeng, F.Q., Ye, Z.Z., Zhang, X.Y., Wang, Q.W., Hao, F., Ma, L., Zuo, X.B., Zhou, F.S., Du, W.H., Cheng, Y.L., Yang, J.Q., Shen, S.K., Li, J., Sheng, Y.J., Zuo, X.X., Zhu, W.F., Gao, F., Zhang, P.L., Guo, Q., Li, B., Gao, M., Xiao, F.L., Quan, C., Zhang, C., Zhang, Z., Zhu, K.J., Li, Y., Hu, D.Y., Lu, W.S., Huang, J.L., Liu, S.X., Li, H., Ren, Y.Q., Wang, Z.X., Yang, C.J., Wang, P.G., Zhou, W.M., Lv, Y.M., Zhang, A.P., Zhang, S.Q., Lin, D., Li, Y., Low, H.Q., Shen, M., Zhai, Z.F., Wang, Y., Zhang, F.Y.,

Yang, S., Liu, J.J. & Zhang, X.J. (2009). Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet* 41(11), 1234-7.

Han, S., Williams, S. & Mustelin, T. (2000). Cytoskeletal protein tyrosine phosphatase PTPH1 reduces T cell antigen receptor signaling. *Eur J Immunol* 30(5), 1318-25.

Hansson, H. & Karlsson-Parra, A. (1999). Canine antinuclear antibodies: comparison of immunofluorescence staining patterns and precipitin reactivity. *Acta Vet Scand* 40(3), 205-12.

Hansson, H., Trowald-Wigh, G. & Karlsson-Parra, A. (1996). Detection of antinuclear antibodies by indirect immunofluorescence in dog sera: comparison of rat liver tissue and human epithelial-2 cells as antigenic substrate. *J Vet Intern Med* 10(4), 199-203.

Hansson-Hamlin, H. & Lilliehook, I. (2009). A possible systemic rheumatic disorder in the Nova Scotia duck tolling retriever. *Acta Vet Scand* 51, 16.

Hansson-Hamlin, H., Lilliehook, I. & Trowald-Wigh, G. (2006). Subgroups of canine antinuclear antibodies in relation to laboratory and clinical findings in immune-mediated disease. *Vet Clin Pathol* 35(4), 397-404.

Hardie, R.J., Gregory, S.P., Tomlin, J., Sturgeon, C., Lipscomb, V. & Ladlow, J. (2005). Cyclosporine treatment of anal furunculosis in 26 dogs. *J Small Anim Pract* 46(1), 3-9.

Havlicek, J. & Roberts, S.C. (2009). MHC-correlated mate choice in humans: a review. *Psychoneuroendocrinology* 34(4), 497-512.

Hill, A.V., Allsopp, C.E., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A., Bennett, S., Brewster, D., McMichael, A.J. & Greenwood, B.M. (1991). Common west African HLA antigens are associated with protection from severe malaria. *Nature* 352(6336), 595-600.

Hochberg, M.C. (1997). Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 40(9), 1725.

Holmdahl, R. (2003). Dissection of the genetic complexity of arthritis using animal models. *J Autoimmun* 21(2), 99-103.

Homann, D. & von Herrath, M. (2004). Regulatory T cells and type 1 diabetes. *Clin Immunol* 112(3), 202-9.

Hughes, A.M., Jokinen, P., Bannasch, D.L., Lohi, H. & Oberbauer, A.M. (2010). Association of a dog leukocyte antigen class II haplotype with hypoadrenocorticism in Nova Scotia Duck Tolling Retrievers. *Tissue Antigens* 75(6), 684-90.

Hughes, A.M., Nelson, R.W., Famula, T.R. & Bannasch, D.L. (2007). Clinical features and heritability of hypoadrenocorticism in Nova Scotia Duck Tolling Retrievers: 25 cases (1994-2006). *J Am Vet Med Assoc* 231(3), 407-12.

Jacobson, D.L., Gange, S.J., Rose, N.R. & Graham, N.M. (1997). Epidemiology and estimated population burden of selected autoimmune diseases in the United States. *Clin Immunol Immunopathol* 84(3), 223-43.

Jokinen, P., Rusanen, E.M., Kennedy, L.J. & Lohi, H. (2011). MHC class II risk haplotype associated with canine chronic superficial keratitis in German Shepherd dogs. *Vet Immunol Immunopathol* 140(1-2), 37-41.

Karlsson, E.K., Baranowska, I., Wade, C.M., Salmon Hillbertz, N.H., Zody, M.C., Anderson, N., Biagi, T.M., Patterson, N., Pielberg, G.R., Kulbokas, E.J., 3rd, Comstock, K.E., Keller, E.T., Mesirov, J.P., von Euler, H., Kampe, O., Hedhammar, A., Lander, E.S.,

Andersson, G., Andersson, L. & Lindblad-Toh, K. (2007). Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat Genet* 39(11), 1321-8.

Karlsson, E.K. & Lindblad-Toh, K. (2008). Leader of the pack: gene mapping in dogs and other model organisms. *Nat Rev Genet* 9(9), 713-25.

Kavanaugh, A., Tomar, R., Reveille, J., Solomon, D.H. & Homburger, H.A. (2000). Guidelines for clinical use of the antinuclear antibody test and tests for specific autoantibodies to nuclear antigens. American College of Pathologists. *Arch Pathol Lab Med* 124(1), 71-81.

Kavanaugh, A.F. & Solomon, D.H. (2002). Guidelines for immunologic laboratory testing in the rheumatic diseases: anti-DNA antibody tests. *Arthritis Rheum* 47(5), 546-55.

Kennedy, L.J. *IPD-MHC Database Canine (DLA)*. [online] Available from: http://www.ebi.ac.uk/ipd/mhc/dla/index.html. [Accessed 2013-03-01].

Kennedy, L.J., Altet, L., Angles, J.M., Barnes, A., Carter, S.D., Francino, O., Gerlach, J.A., Happ, G.M., Ollier, W.E., Polvi, A., Thomson, W. & Wagner, J.L. (2000). Nomenclature for factors of the dog major histocompatibility system (DLA), 1998: first report of the ISAG DLA Nomenclature Committee. *Anim Genet* 31(1), 52-61.

Kennedy, L.J., Angles, J.M., Barnes, A., Carter, S.D., Francino, O., Gerlach, J.A., Happ, G.M., Ollier, W.E., Thomson, W. & Wagner, J.L. (2001). Nomenclature for factors of the dog major histocompatibility system (DLA), 2000: Second report of the ISAG DLA Nomenclature Committee. *Tissue Antigens* 58(1), 55-70.

Kennedy, L.J., Barnes, A., Happ, G.M., Quinnell, R.J., Bennett, D., Angles, J.M., Day, M.J., Carmichael, N., Innes, J.F., Isherwood, D., Carter, S.D., Thomson, W. & Ollier, W.E. (2002). Extensive interbreed, but minimal intrabreed, variation of DLA class II alleles and haplotypes in dogs. *Tissue Antigens* 59(3), 194-204.

Kennedy, L.J., Barnes, A., Ollier, W.E. & Day, M.J. (2006a). Association of a common dog leucocyte antigen class II haplotype with canine primary immune-mediated haemolytic anaemia. *Tissue Antigens* 68(6), 502-8.

Kennedy, L.J., Barnes, A., Short, A., Brown, J.J., Seddon, J., Fleeman, L., Brkljacic, M., Happ, G.M., Catchpole, B. & Ollier, W.E. (2007). Canine DLA diversity: 3. Disease studies. *Tissue Antigens* 69 Suppl 1, 292-6.

Kennedy, L.J., Davison, L.J., Barnes, A., Short, A.D., Fretwell, N., Jones, C.A., Lee, A.C., Ollier, W.E. & Catchpole, B. (2006b). Identification of susceptibility and protective major histocompatibility complex haplotypes in canine diabetes mellitus. *Tissue Antigens* 68(6), 467-76.

Kennedy, L.J., Huson, H.J., Leonard, J., Angles, J.M., Fox, L.E., Wojciechowski, J.W., Yuncker, C. & Happ, G.M. (2006c). Association of hypothyroid disease in Doberman Pinscher dogs with a rare major histocompatibility complex DLA class II haplotype. *Tissue Antigens* 67(1), 53-6.

Kennedy, L.J., O'Neill, T., House, A., Barnes, A., Kyostila, K., Innes, J., Fretwell, N., Day, M.J., Catchpole, B., Lohi, H. & Ollier, W.E. (2008). Risk of anal furunculosis in German shepherd dogs is associated with the major histocompatibility complex. *Tissue Antigens* 71(1), 51-6.

Kennedy, L.J., Quarmby, S., Happ, G.M., Barnes, A., Ramsey, I.K., Dixon, R.M., Catchpole, B., Rusbridge, C., Graham, P.A., Hillbertz, N.S., Roethel, C., Dodds, W.J., Carmichael,

N.G. & Ollier, W.E. (2006d). Association of canine hypothyroidism with a common major histocompatibility complex DLA class II allele. *Tissue Antigens* 68(1), 82-6.

Kotzin, B.L. (1996). Systemic lupus erythematosus. *Cell* 85(3), 303-6.

Kozyrev, S.V., Abelson, A.K., Wojcik, J., Zaghlool, A., Linga Reddy, M.V., Sanchez, E., Gunnarsson, I., Svenungsson, E., Sturfelt, G., Jonsen, A., Truedsson, L., Pons-Estel, B.A., Witte, T., D'Alfonso, S., Barizzone, N., Danieli, M.G., Gutierrez, C., Suarez, A., Junker, P., Laustrup, H., Gonzalez-Escribano, M.F., Martin, J., Abderrahim, H. & Alarcon-Riquelme, M.E. (2008). Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. *Nat Genet* 40(2), 211-6.

Kyttaris, V.C., Wang, Y., Juang, Y.T., Weinstein, A. & Tsokos, G.C. (2007). Increased levels of NF-ATc2 differentially regulate CD154 and IL-2 genes in T cells from patients with systemic lupus erythematosus. *J Immunol* 178(3), 1960-6.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kaspryzk, A.,

Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S. & Chen, Y.J. (2001). Initial sequencing and analysis of the human genome. *Nature* 409(6822), 860-921.

Larson, G., Karlsson, E.K., Perri, A., Webster, M.T., Ho, S.Y., Peters, J., Stahl, P.W., Piper, P.J., Lingaas, F., Fredholm, M., Comstock, K.E., Modiano, J.F., Schelling, C., Agoulnik, A.I., Leegwater, P.A., Dobney, K., Vigne, J.D., Vila, C., Andersson, L. & Lindblad-Toh, K. (2012). Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci U S A* 109(23), 8878-83.

Lee, H.M., Sugino, H. & Nishimoto, N. (2010). Cytokine networks in systemic lupus erythematosus. *J Biomed Biotechnol* 2010, 676284.

Lessard, C.J., Adrianto, I., Kelly, J.A., Kaufman, K.M., Grundahl, K.M., Adler, A., Williams, A.H., Gallant, C.J., Anaya, J.M., Bae, S.C., Boackle, S.A., Brown, E.E., Chang, D.M., Criswell, L.A., Edberg, J.C., Freedman, B.I., Gregersen, P.K., Gilkeson, G.S., Jacob, C.O., James, J.A., Kamen, D.L., Kimberly, R.P., Martin, J., Merrill, J.T., Niewold, T.B., Park, S.Y., Petri, M.A., Pons-Estel, B.A., Ramsey-Goldman, R., Reveille, J.D., Song, Y.W., Stevens, A.M., Tsao, B.P., Vila, L.M., Vyse, T.J., Yu, C.Y., Guthridge, J.M., Bruner, G.R., Langefeld, C.D., Montgomery, C., Harley, J.B., Scofield, R.H., Gaffney, P.M. & Moser, K.L. (2011). Identification of a systemic lupus erythematosus susceptibility locus at 11p13 between PDHX and CD44 in a multiethnic study. *Am J Hum Genet* 88(1), 83-91.

Li, H. & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26(5), 589-95.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16), 2078-9.

Li, Y., Harada, T., Juang, Y.T., Kyttaris, V.C., Wang, Y., Zidanic, M., Tung, K. & Tsokos, G.C. (2007). Phosphorylated ERM is responsible for increased T cell polarization, adhesion, and migration in patients with systemic lupus erythematosus. *J Immunol* 178(3), 1938-47.

Lin, L., Faraco, J., Li, R., Kadotani, H., Rogers, W., Lin, X., Qiu, X., de Jong, P.J., Nishino, S. & Mignot, E. (1999). The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 98(3), 365-76.

Lin, T.Y., Chan, L.C., Fan, Y.H., Lin, C.H., Chow, K.C., Lin, S.L., Lan, J.L., Lin, F.J. & Chiou, S.H. (2006). Use of a recombinant protein containing major epitopes of hnRNP G to detect anti-hnRNP G antibodies in dogs with systemic lupus erythematosus. *Res Vet Sci* 81(3), 335-9.

Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M.F., Parker, B.J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., Ward, L.D., Lowe, C.B., Holloway, A.K., Clamp, M., Gnerre, S., Alfoldi, J., Beal, K., Chang, J., Clawson, H., Cuff, J., Di Palma, F., Fitzgerald, S., Flicek, P., Guttman, M., Hubisz, M.J., Jaffe, D.B., Jungreis, I., Kent,

W.J., Kostka, D., Lara, M., Martins, A.L., Massingham, T., Moltke, I., Raney, B.J., Rasmussen, M.D., Robinson, J., Stark, A., Vilella, A.J., Wen, J., Xie, X., Zody, M.C., Baldwin, J., Bloom, T., Chin, C.W., Heiman, D., Nicol, R., Nusbaum, C., Young, S., Wilkinson, J., Worley, K.C., Kovar, C.L., Muzny, D.M., Gibbs, R.A., Cree, A., Dihn, H.H., Fowler, G., Jhangiani, S., Joshi, V., Lee, S., Lewis, L.R., Nazareth, L.V., Okwuonu, G., Santibanez, J., Warren, W.C., Mardis, E.R., Weinstock, G.M., Wilson, R.K., Delehaunty, K., Dooling, D., Fronik, C., Fulton, L., Fulton, B., Graves, T., Minx, P., Sodergren, E., Birney, E., Margulies, E.H., Herrero, J., Green, E.D., Haussler, D., Siepel, A., Goldman, N., Pollard, K.S., Pedersen, J.S., Lander, E.S. & Kellis, M. (2011). A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478(7370), 476-82.

Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., 3rd, Zody, M.C., Mauceli, E., Xie, X., Breen, M., Wayne, R.K., Ostrander, E.A., Ponting, C.P., Galibert, F., Smith, D.R., DeJong, P.J., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C.W., Cook, A., Cuff, J., Daly, M.J., DeCaprio, D., Gnerre, S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K.P., Parker, H.G., Pollinger, J.P., Searle, S.M., Sutter, N.B., Thomas, R., Webber, C., Baldwin, J., Abebe, A., Abouelleil, A., Aftuck, L., Ait-Zahra, M., Aldredge, T., Allen, N., An, P., Anderson, S., Antoine, C., Arachchi, H., Aslam, A., Ayotte, L., Bachantsang, P., Barry, A., Bayul, T., Benamara, M., Berlin, A., Bessette, D., Blitshteyn, B., Bloom, T., Blye, J., Boguslavskiy, L., Bonnet, C., Boukhgalter, B., Brown, A., Cahill, P., Calixte, N., Camarata, J., Cheshatsang, Y., Chu, J., Citroen, M., Collymore, A., Cooke, P., Dawoe, T., Daza, R., Decktor, K., DeGray, S., Dhargay, N., Dooley, K., Dooley, K., Dorje, P., Dorjee, K., Dorris, L., Duffey, N., Dupes, A., Egbiremolen, O., Elong, R., Falk, J., Farina, A., Faro, S., Ferguson, D., Ferreira, P., Fisher, S., FitzGerald, M., Foley, K., Foley, C., Franke, A., Friedrich, D., Gage, D., Garber, M., Gearin, G., Giannoukos, G., Goode, T., Goyette, A., Graham, J., Grandbois, E., Gyaltsen, K., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Healy, C., Hegarty, R., Honan, T., Horn, A., Houde, N., Hughes, L., Hunnicutt, L., Husby, M., Jester, B., Jones, C., Kamat, A., Kanga, B., Kells, C., Khazanovich, D., Kieu, A.C., Kisner, P., Kumar, M., Lance, K., Landers, T., Lara, M., Lee, W., Leger, J.P., Lennon, N., Leuper, L., LeVine, S., Liu, J., Liu, X., Lokyitsang, Y., Lokyitsang, T., Lui, A., Macdonald, J., Major, J., Marabella, R., Maru, K., Matthews, C., McDonough, S., Mehta, T., Meldrim, J., Melnikov, A., Meneus, L., Mihalev, A., Mihova, T., Miller, K., Mittelman, R., Mlenga, V., Mulrain, L., Munson, G., Navidi, A., Naylor, J., Nguyen, T., Nguyen, N., Nguyen, C., Nguyen, T., Nicol, R., Norbu, N., Norbu, C., Novod, N., Nyima, T., Olandt, P., O'Neill, B., O'Neill, K., Osman, S., Oyono, L., Patti, C., Perrin, D., Phunkhang, P., Pierre, F., Priest, M., Rachupka, A., Raghuraman, S., Rameau, R., Ray, V., Raymond, C., Rege, F., Rise, C., Rogers, J., Rogov, P., Sahalie, J., Settipalli, S., Sharpe, T., Shea, T., Sheehan, M., Sherpa, N., Shi, J., Shih, D., Sloan, J., Smith, C., Sparrow, T., Stalker, J., Stange-Thomann, N., Stavropoulos, S., Stone, C., Stone, S., Sykes, S., Tchuinga, P., Tenzing, P., Tesfaye, S., Thoulutsang, D., Thoulutsang, Y., Topham, K., Topping, I., Tsamla, T., Vassiliev, H., Venkataraman, V., Vo, A., Wangchuk, T., Wangdi, T., Weiand, M., Wilkinson, J., Wilson, A., Yadav, S., Yang, S., Yang, X., Young, G., Yu, Q., Zainoun,

J., Zembek, L., Zimmer, A. & Lander, E.S. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438(7069), 803-19.

Little, A.M. & Parham, P. (1999). Polymorphism and evolution of HLA class I and II genes and molecules. *Rev Immunogenet* 1(1), 105-23.

Longenecker, B.M., Pazderka, F., Gavora, J.S., Spencer, J.L., Stephens, E.A., Witter, R.L. & Ruth, R.F. (1977). Role of the major histocompatibility complex in resistance to Marek's disease: restriction of the growth of JMV-MD tumor cells in genetically resistant birds. *Adv Exp Med Biol* 88, 287-98.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F., McCarroll, S.A. & Visscher, P.M. (2009). Finding the missing heritability of complex diseases. *Nature* 461(7265), 747-53.

Marrack, P., Kappler, J. & Kotzin, B.L. (2001). Autoimmune disease: why and where it occurs. *Nat Med* 7(8), 899-905.

Marshall, A.J., Niiro, H., Lerner, C.G., Yun, T.J., Thomas, S., Disteche, C.M. & Clark, E.A. (2000). A novel B lymphocyte-associated adaptor protein, Bam32, regulates antigen receptor signaling downstream of phosphatidylinositol 3-kinase. *J Exp Med* 191(8), 1319-32.

Massey, J., Boag, A., Short, A.D., Scholey, R.A., Henthorn, P.S., Littman, M.P., Husebye, E., Catchpole, B., Pedersen, N., Mellersh, C.S., Ollier, W.E. & Kennedy, L.J. (2013a). MHC class II association study in eight breeds of dog with hypoadrenocorticism. *Immunogenetics*.

Massey, J., Rothwell, S., Rusbridge, C., Tauro, A., Addicott, D., Chinoy, H., Cooper, R.G., Ollier, W.E.R. & Kennedy, L.J. (2013b). Association of an MHC Class II Haplotype with Increased Risk of Polymyositis in Hungarian Vizsla Dogs. *PLoS One* 8(2), e56490.

Meyer, D.J., Iverson, W.O. & Terrell, T.G. (1980). Obstructive jaundice associated with chronic active hepatitis in a dog. *J Am Vet Med Assoc* 176(1), 41-4.

Mikko, S. & Andersson, L. (1995). Low major histocompatibility complex class II diversity in European and North American moose. *Proc Natl Acad Sci U S A* 92(10), 4259-63.

Monier, J.C., Ritter, J., Caux, C., Chabanne, L., Fournel, C., Venet, C. & Rigal, D. (1992). Canine systemic lupus erythematosus. II: Antinuclear antibodies. *Lupus* 1(5), 287-93.

Nabozny, G.H., Baisch, J.M., Cheng, S., Cosgrove, D., Griffiths, M.M., Luthra, H.S. & David, C.S. (1996). HLA-DQ8 transgenic mice are highly susceptible to collagen-induced arthritis: a novel model for human polyarthritis. *J Exp Med* 183(1), 27-37.

Nachreiner, R.F., Refsal, K.R., Graham, P.A., Hauptman, J. & Watson, G.L. (1998). Prevalence of autoantibodies to thyroglobulin in dogs with nonthyroidal illness. *Am J Vet Res* 59(8), 951-5.

Nishikomori, R., Usui, T., Wu, C.Y., Morinobu, A., O'Shea, J.J. & Strober, W. (2002). Activated STAT4 has an essential role in Th1 differentiation and proliferation that is independent of its role in the maintenance of IL-12R beta 2 chain expression and signaling. *J Immunol* 169(8), 4388-98.

Niskanen, A.K., Hagstrom, E., Lohi, H., Ruokonen, M., Esparza-Salas, R., Aspi, J. & Savolainen, P. (2013). MHC variability supports dog domestication from a large number of wolves: high diversity in Asia. *Heredity (Edinb)* 110(1), 80-5.

O'Brien, S.J. & Yuhki, N. (1999). Comparative genome organization of the major histocompatibility complex: lessons from the Felidae. *Immunol Rev* 167, 133-44.

Ollier, W.E., Kennedy, L.J., Thomson, W., Barnes, A.N., Bell, S.C., Bennett, D., Angles, J.M., Innes, J.F. & Carter, S.D. (2001). Dog MHC alleles containing the human RA shared epitope confer susceptibility to canine rheumatoid arthritis. *Immunogenetics* 53(8), 669-73.

Ortner, D., Grabher, D., Hermann, M., Kremmer, E., Hofer, S. & Heufler, C. (2011). The adaptor protein Bam32 in human dendritic cells participates in the regulation of MHC class I-induced CD8+ T cell activation. *J Immunol* 187(8), 3972-8.

Ostrander, E.A. (2012). Franklin H. Epstein Lecture. Both ends of the leash--the human links to good dogs with bad genes. *N Engl J Med* 367(7), 636-46.

Ovodov, N.D., Crockford, S.J., Kuzmin, Y.V., Higham, T.F., Hodgins, G.W. & van der Plicht, J. (2011). A 33,000-year-old incipient dog from the Altai Mountains of Siberia: evidence of the earliest domestication disrupted by the Last Glacial Maximum. *PLoS One* 6(7), e22821.

Ovrebo Bohnhorst, J., Hanssen, I. & Moen, T. (2001). Antinuclear antibodies (ANA) in Gordon setters with symmetrical lupoid onychodystrophy and black hair follicular dysplasia. *Acta Vet Scand* 42(3), 323-9.

Pang, J.F., Kluetsch, C., Zou, X.J., Zhang, A.B., Luo, L.Y., Angleby, H., Ardalan, A., Ekstrom, C., Skollermo, A., Lundeberg, J., Matsumura, S., Leitner, T., Zhang, Y.P. & Savolainen, P. (2009). mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol Biol Evol* 26(12), 2849-64.

Parker, H.G., Kim, L.V., Sutter, N.B., Carlson, S., Lorentzen, T.D., Malek, T.B., Johnson, G.S., DeFrance, H.B., Ostrander, E.A. & Kruglyak, L. (2004). Genetic structure of the purebred domestic dog. *Science* 304(5674), 1160-4.

Pawitan, Y., Seng, K.C. & Magnusson, P.K. (2009). How many genetic variants remain to be discovered? *PLoS One* 4(12), e7969.

Perrin, T. *The Nova Scotia Duck Tolling Retriever*. [online] Available from: http://www.moderndogmagazine.com/breeds/nova-scotia-duck-tolling-retriever. [Accessed 2013-02-27].

Potts, W.K., Manning, C.J. & Wakeland, E.K. (1991). Mating patterns in seminatural populations of mice influenced by MHC genotype. *Nature* 352(6336), 619-21.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. & Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3), 559-75.

Quinlan, A.R. & Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6), 841-2.

Redman, J. (2002). Steroid-responsive meningitis-arteritis in the Nova Scotia duck tolling retriever. *Vet Rec* 151(23), 712.

Remmers, E.F., Plenge, R.M., Lee, A.T., Graham, R.R., Hom, G., Behrens, T.W., de Bakker, P.I., Le, J.M., Lee, H.S., Batliwalla, F., Li, W., Masters, S.L., Booty, M.G., Carulli, J.P., Padyukov, L., Alfredsson, L., Klareskog, L., Chen, W.V., Amos, C.I., Criswell, L.A., Seldin, M.F., Kastner, D.L. & Gregersen, P.K. (2007). STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med* 357(10), 977-86.

Robinson, J., Halliwell, J.A., McWilliam, H., Lopez, R., Parham, P. & Marsh, S.G. (2013). The IMGT/HLA database. *Nucleic Acids Res* 41(Database issue), D1222-7.

Salmon Hillbertz, N.H., Isaksson, M., Karlsson, E.K., Hellmen, E., Pielberg, G.R., Savolainen, P., Wade, C.M., von Euler, H., Gustafson, U., Hedhammar, A., Nilsson, M., Lindblad-Toh, K., Andersson, L. & Andersson, G. (2007). Duplication of FGF3, FGF4, FGF19 and ORAOV1 causes hair ridge and predisposition to dermoid sinus in Ridgeback dogs. *Nat Genet* 39(11), 1318-20.

Savolainen, P., Zhang, Y.P., Luo, J., Lundeberg, J. & Leitner, T. (2002). Genetic evidence for an East Asian origin of domestic dogs. *Science* 298(5598), 1610-3.

Serfling, E., Klein-Hessling, S., Palmetshofer, A., Bopp, T., Stassen, M. & Schmitt, E. (2006). NFAT transcription factors in control of peripheral T cell tolerance. *Eur J Immunol* 36(11), 2837-43.

Shamim, E.A. & Miller, F.W. (2000). Familial autoimmunity and the idiopathic inflammatory myopathies. *Curr Rheumatol Rep* 2(3), 201-11.

Shearin, A.L. & Ostrander, E.A. (2010). Leading the way: canine models of genomics and disease. *Dis Model Mech* 3(1-2), 27-34.

Smerdel-Ramoya, A., Finholt, C., Lilleby, V., Gilboe, I.M., Harbo, H.F., Maslinski, S., Forre, O., Thorsby, E. & Lie, B.A. (2005). Systemic lupus erythematosus and the extended major histocompatibility complex--evidence for several predisposing loci. *Rheumatology (Oxford)* 44(11), 1368-73.

Sommers, C.L., Gurson, J.M., Surana, R., Barda-Saad, M., Lee, J., Kishor, A., Li, W., Gasser, A.J., Barr, V.A., Miyaji, M., Love, P.E. & Samelson, L.E. (2008). Bam32: a novel mediator of Erk activation in T cells. *Int Immunol* 20(7), 811-8.

Sozio, M.S., Mathis, M.A., Young, J.A., Walchli, S., Pitcher, L.A., Wrage, P.C., Bartok, B., Campbell, A., Watts, J.D., Aebersold, R., Hooft van Huijsduijnen, R. & van Oers, N.S. (2004). PTPH1 is a predominant protein-tyrosine phosphatase capable of interacting with and dephosphorylating the T cell receptor zeta subunit. *J Biol Chem* 279(9), 7760-9.

Speeti, M., Stahls, A., Meri, S. & Westermarck, E. (2003). Upregulation of major histocompatibility complex class II antigens in hepatocytes in Doberman hepatitis. *Vet Immunol Immunopathol* 96(1-2), 1-12.

Stewart, C.A., Horton, R., Allcock, R.J., Ashurst, J.L., Atrazhev, A.M., Coggill, P., Dunham, I., Forbes, S., Halls, K., Howson, J.M., Humphray, S.J., Hunt, S., Mungall, A.J., Osoegawa, K., Palmer, S., Roberts, A.N., Rogers, J., Sims, S., Wang, Y., Wilming, L.G., Elliott, J.F., de Jong, P.J., Sawcer, S., Todd, J.A., Trowsdale, J. & Beck, S. (2004). Complete MHC haplotype sequencing for common disease gene mapping. *Genome Res* 14(6), 1176-87.

Strang, A. & MacMillan, G. (1996). *The Nova Scotia Duck Tolling Retriever*. Loveland, Colorado: Alpine Publications, Inc.

Su, D.L., Lu, Z.M., Shen, M.N., Li, X. & Sun, L.Y. (2012). Roles of pro- and anti-inflammatory cytokines in the pathogenesis of SLE. *J Biomed Biotechnol* 2012, 347141.

Sutter, N.B., Eberle, M.A., Parker, H.G., Pullar, B.J., Kirkness, E.F., Kruglyak, L. & Ostrander, E.A. (2004). Extensive and breed-specific linkage disequilibrium in Canis familiaris. *Genome Res* 14(12), 2388-96.

Tan, E.M. (1989). Antinuclear antibodies: diagnostic markers for autoimmune diseases and probes for cell biology. *Adv Immunol* 44, 93-151.

Tan, E.M., Cohen, A.S., Fries, J.F., Masi, A.T., McShane, D.J., Rothfield, N.F., Schaller, J.G., Talal, N. & Winchester, R.J. (1982). The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 25(11), 1271-7.

Tan, E.M. & Kunkel, H.G. (1966). Characteristics of a soluble nuclear antigen precipitating with sera of patients with systemic lupus erythematosus. *J Immunol* 96(3), 464-71.

Teichner, M., Krumbacher, K., Doxiadis, I., Doxiadis, G., Fournel, C., Rigal, D., Monier, J.C. & Grosse-Wilde, H. (1990). Systemic lupus erythematosus in dogs: association to the major histocompatibility complex class I antigen DLA-A7. *Clin Immunol Immunopathol* 55(2), 255-62.

Thoren-Tolling, K. & Ryden, L. (1991). Serum auto antibodies and clinical/pathological features in German shepherd dogs with a lupuslike syndrome. *Acta Vet Scand* 32(1), 15-26.

Tollarklubben *RAS Rasspecifika avelsstrategier Nova Scotia Duck Tolling Retriever*. [online] Available from: http://www.skk.se/Global/Dokument/RASdokument/RAS-nova-scotia-duck-tolling-retriever.pdf?epslanguage=sv. [Accessed 2013-02-28].

Truve, K., Eriksson, O., Norling, M., Wilbe, M., Mauceli, E., Lindblad-Toh, K. & Bongcam-Rudloff, E. (2011). SEQscoring: a tool to facilitate the interpretation of data generated with next generation sequencing technologies. *EMBnet journal* 17(1), 38-45.

Tsuchiya, N., Kawasaki, A., Tsao, B.P., Komata, T., Grossman, J.M. & Tokunaga, K. (2001). Analysis of the association of HLA-DRB1, TNFalpha promoter and TNFR2 (TNFRSF1B) polymorphisms with SLE using transmission disequilibrium test. *Genes Immun* 2(6), 317-22.

Vajner, L. (1997). Lymphocytic thyroiditis in beagle dogs in a breeding colony: findings of serum autoantibodies. *Vet Med (Praha)* 42(11), 333-8.

Vaysse, A., Ratnakumar, A., Derrien, T., Axelsson, E., Rosengren Pielberg, G., Sigurdsson, S., Fall, T., Seppala, E.H., Hansen, M.S., Lawley, C.T., Karlsson, E.K., Bannasch, D., Vila, C., Lohi, H., Galibert, F., Fredholm, M., Haggstrom, J., Hedhammar, A., Andre, C., Lindblad-Toh, K., Hitte, C. & Webster, M.T. (2011). Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet* 7(10), e1002316.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R.,

Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigo, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A. & Zhu, X. (2001). The sequence of the human genome. *Science* 291(5507), 1304-51.

Vila, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J. & Wayne, R.K. (1997). Multiple and ancient origins of the domestic dog. *Science* 276(5319), 1687-9.

Wagner, J.L., Burnett, R.C. & Storb, R. (1999). Organization of the canine major histocompatibility complex: current perspectives. *J Hered* 90(1), 35-8.

Wang, W.Y., Barratt, B.J., Clayton, D.G. & Todd, J.A. (2005). Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 6(2), 109-18.

Wedekind, C. & Furi, S. (1997). Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? *Proc Biol Sci* 264(1387), 1471-9.

Wilbe, M., Sundberg, K., Hansen, I.R., Strandberg, E., Nachreiner, R.F., Hedhammar, A., Kennedy, L.J., Andersson, G. & Bjornerfeldt, S. (2010a). Increased genetic risk or

protection for canine autoimmune lymphocytic thyroiditis in Giant Schnauzers depends on DLA class II genotype. *Tissue Antigens* 75(6), 712-9.

Wilbe, M., Ziener, M.L., Aronsson, A., Harlos, C., Sundberg, K., Norberg, E., Andersson, L., Lindblad-Toh, K., Hedhammar, A., Andersson, G. & Lingaas, F. (2010b). DLA class II alleles are associated with risk for canine symmetrical lupoid onychodystrophy [corrected](SLO). *PLoS One* 5(8), e12332.

Yang, W., Shen, N., Ye, D.Q., Liu, Q., Zhang, Y., Qian, X.X., Hirankarn, N., Ying, D., Pan, H.F., Mok, C.C., Chan, T.M., Wong, R.W., Lee, K.W., Mok, M.Y., Wong, S.N., Leung, A.M., Li, X.P., Avihingsanon, Y., Wong, C.M., Lee, T.L., Ho, M.H., Lee, P.P., Chang, Y.K., Li, P.H., Li, R.J., Zhang, L., Wong, W.H., Ng, I.O., Lau, C.S., Sham, P.C. & Lau, Y.L. (2010). Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. *PLoS Genet* 6(2), e1000841.

Yuhki, N., Beck, T., Stephens, R.M., Nishigaki, Y., Newmann, K. & O'Brien, S.J. (2003). Comparative genome organization of human, murine, and feline MHC class II region. *Genome Res* 13(6A), 1169-79.

# Acknowledgments

This work was performed at the Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences (SLU). I would like to thank everyone I have had the opportunity to work with during these years, in particular I would like to acknowledge:

**Göran Andersson**, my main supervisor for choosing me as your PhD student. I would like to thank you for interesting scientific discussions, all of our fun trips, valuable teaching experiences and for letting me be a part of so many interesting projects. I especially appreciate that you have always believed in me, trusted me, giving me the freedom to develop as a scientist and all kind of support!

**Kerstin Lindblad-Toh**, for introducing me to all the collaborators, your scientific support and your great commitment to our research. Your knowledge and enthusiasm have been a great inspiration for me during these years.

**Helene Hansson Hamlin** for teaching me all clinical aspects of the disease, for your endless support and all the fun times. This research would have been impossible without your clinical expertise.

**Erik Bongcam-Rudloff** for introducing me to the world of bioinformatics. Thanks also for the fun teaching experience in Kenya.

I would like to acknowledge all the people who have worked in the dog SLE group, **Sergey Kozyrev** and **Fabiana Farias** for doing an excellent work on the expression studies, **Katarina Truvé** for enjoyable collaboration, especially on the GWAS study. **Gerli Pielberg** for being such a nice and skilled

supervisor in the lab and **Ulla Gustafson** for all the help with sample preparation, genotyping and teaching me how to work in the lab!

**Leif Andersson** for many valuable and interesting discussions and support in several projects.

**Åke Hedhammar** for nice collaborations in several projects and your contacts with the dog community!

Thanks to all past and present members in the dog group at the BROAD Institute for interesting and valuable discussions. In particular, I would like to acknowledge **Tara Biagi**, **Hyun Ji**, **Elinor Karlsson**, **Michele Perloski**, **Ross Swofford** and **Claire Wade**.

Our international collaborators **Dannika Bannasch**, **Angela Hughes**, **Päivi Jokinen**, **Hannes Lohi**, and **Eija Seppala** for excellent work on the GWAS paper.

**Lorna J Kennedy** for introducing me to DLA class II!

**Erling Strandberg** for accepting me as a PhD student at the department and statistical help.

**Gudrun Wieslander** and later **Helena Pettersson**, thank you for all help with administrative questions!

Thanks to all the **co-author** and **collaborators** during these years!

The **Swedish dog group**, thanks for interesting and valuable meetings and discussions and especially **Cecilia Johansson**, who is doing such a nice job coordinating us!

Thanks to the students performing an excellent effort on some of the studies in this thesis, **Christina Hermanrud**, **Charlotte Harlos**, **Nomingerel Munkhtsooj** and **Anna Hedlund**.

I would like to thank the group at Akademiska sjukhuset, **Kerstin Ahlgren**, **Olle Kämpe** and **Anna Lobell** for nice collaborations!

Thanks to all **past** and **present** members of the lab!

My past roomies, **Anders**, **Jonas** and **Lisa**, this time would not have been so nice without you! Also thank to **Yulna** and **Andreas** (later roomies) for interesting discussions and fun times.

Thanks for lots of energy and music guidelines to the Monday training group: **Rebecca**, **Andreas**, **Matt** and **Mia** (last but not least ☺)!

All contributing **dog owners** and **veterinarians** who allowed us to take samples and provided us with important dog information during many years and at any time. **Gun**, thanks for being a backup so many times during sample collections!

Special thanks to my colleagues and friends for your incredible support! I specially would like to thank **Jeanette**, **Lisa** and **Mia**.

A special thank you to **Tollarklubben** for your sample support, valuable collaboration and interest in conducting this work!

To **my family** for all kinds of support and encouragement during this time!

A final thank goes to all the **dogs** participating in these studies!