Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science

Functional Analysis of *PTPN3* associated with Immune-Mediated Rheumatic Disease in Nova Scotia Duck Tolling Retrievers

Pedro Sousa

Examensarbete / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics,

421

Uppsala 2013

Master's Thesis, 30 HEC

Erasmus Mundus Programme – European Master in Animal Breeding and Genetics

Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics

Functional Analysis of *PTPN3* associated with Immune-Mediated Rheumatic Disease in Nova Scotia Duck Tolling Retrievers

Pedro Sousa

Supervisors:
Sergey Kosyrev, UU, Department of Medical Biochemistry and Microbiology
Göran Andersson, SLU, Department of Animal Breeding and Genetics
Hans Sölkner, BOKU, Vienna, AT
Examiner:
Erling Strandberg, SLU, Department of Animal Breeding and Genetics

Credits: 30 HEC
Course title: Degree project in Animal Science
Course code: EX0556
Programme: Erasmus Mundus Programme - European Master in Animal Breeding and Genetics
Level: Advanced, A2E

Place of publication: Uppsala
Year of publication: 2013
Name of series: Examensarbete / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, 421
On-line publication: http://epsilon.slu.se

Key words: lupus, gene expression, genetic association, cloning, luciferase assay

Functional Analysis of *PTPN3* associated with Immune-Mediated Rheumatic Disease in Nova Scotia Duck Tolling Retrievers

Pedro Sousa

THESIS ANIMAL BREEDING AND GENETICS
June 2013

SUPERVISORS
Sergey Kosyrev
Göran Andersson
Hans Sölkner

EXAMINER
Erling Strandberg

Education and Culture
Erasmus Mundus

Table of Contents

1 RESEARCH FRAMEWORK

This study investigates whether risk haplotypes of highly associated regions in *PTPN3* gene in dogs suffering from immune mediated rheumatic disease, might affect the regulatory potential and expression levels of this gene influencing the phenotype.

1.2 Aim of the study

The overall aim of this thesis is to identify causative mutations that may alter *PTPN3* gene expression in NSDTRs with risk for IMRD. Risk haplotypes are associated with dogs expressing ANA positivity and phenotypic signs of illness. It was previously shown that dogs from the breed Nova Scotia duck tolling retriever (NSDTR) reveal a high degree of susceptibility to autoimmune rheumatic diseases.

Specific aims:

1. Bioinformatics analysis of 3'UTR and the associated region located in intron 18 of *PTPN3* gene and DNA sequencing analysis of regions of interest.

2. Determine functional differences between risk and protective haplotype variants of 3´UTR and intron 18 regions of *PTPN3* gene in NSDTRs, using recombinant DNA techniques, molecular cloning, transient transfection of Jurkat T-cells and MDCK cells, reporter gene assays and RNA degradation analysis will be performed to develop this study.

1.3 Introduction

In normal individuals the immune system should be self tolerant. This is achieved by a number of mechanisms that prevent B and T lymphocytes from becoming self-reactive. Autoimmunity starts developing when an adaptive immune response tolerance breaks down and reacts to self-antigens, causing an inflammatory reaction and eventual tissue damage. Autoimmunity can have systemic or regional consequences, as it is seen in systemic lupus erythematosus (SLE), affecting multiple tissues and organs (Janeway *et al.*, 2005).

NSDTRs have been shown to be associated with a strong genetic predisposition to different autoimmune diseases, including immune-mediated rheumatic disease (IMRD) and steroid-responsive meningitis arteritis (SRMA). IMRD and SRMA are considered to be part of the same disorder, canine SLE-related disease complex (Jokinen, 2011). Regardless of the high prevalence of autoimmune disorders in dogs, the genetic background seems to be shared by several associated regions containing different genes, making the identification of causative ones a difficult task. A recent genome-wide association study (GWAS) revealed that, four of the strongest candidate genes (*PPP3CA, HOMER2, DAPP1* and *PTPN3*) are highly associated with dogs suffering from IMRD. Apparently, these same genes seem to be also

involved in regulation of nuclear factor of activated T cells (NFAT) pathway (Wilbe *et al*., 2009).
The strongest associated region is located on chromosome 11 and contains only one gene, *PTPN3*. Targeted re-sequencing of this gene in healthy and diseased dogs helped to define the risk haplotype. Encoded *PTPN3* protein seems to play an important role in TCR-signaling inhibition and transduction and is one of the most associated genes in NSDTRs with IMRD (Wilbe *et al.*, 2010). Analysis of gene expression in peripheral blood mononuclear cells (PBMCs) revealed that dogs carrying risk haplotypes express much lower *PTPN3* mRNA levels (up to eight times fold) compared to dogs with protective haplotype (Wilbe *et al*., 2013). This evidence suggests that the major effect might be associated with the regulation of gene expression. Associated variants from risk haplotype seem to be spread out in a few regions with regulatory potential, including the *PTPN3* promoter, intronic enhancers and mRNA stability in the 3'UTR. SNPs in introns and 3'UTRs may have a strong influence in transcription or translation as well as a direct effect on mRNA stability or translation efficiency (Wilbe *et al*., 2013). Besides, *PTPN3* seems to be involved in a novel T-cell activation pathway, NFAT as a negative regulator of TCR (Wilbe *et al*., 2013).

Although many genes have been associated with SLE, functional variants for most of these genes have not yet been identified, and that is why the biology underlying them is still unknown. It is of extreme importance to understand not only how they function and interact with each other, but also what happens exactly if one of them is missing or overrepresented.
This study aimed to identify causative mutations and determine the molecular effects leading to down-regulation of *PTPN3* gene expression in NSDTRs with the risk haplotype.

Transient transfection analyses in MDCK and Jurkat T cells were performed using a series of constructs containing dog *PTPN3*-intron 18 and 3'-UTR ligated to a luciferase reporter.
We have developed functional analysis using *in silico* tools and recombinant DNA techniques like PCR, molecular cloning, plasmid DNA purification, transformation in E. coli and transfection of two different mammalian cell types, luciferase expression analysis and RNA degradation assays. This study was undertaken to study genetic impact of associated variants of *PTPN3* with susceptibility to IMRD and to assess the functional implications of this association.

2 REVIEW OF LITERATURE

2.1 The Dog as experimental model for Human Inherited disorders

Dogs have been proven to be a very important model organism for mapping human complex diseases (Lindblad-Toh *et al*., 2005). Because a similar spectrum of diseases that occur spontaneously over the course of their lives as well as advantageous population structure (Ostrander *et al.*, 2000). Ancient dogs were initially domesticated from wolves since around 15,000 years ago, most likely through multiple domestication events. Its genome has short-range Linkage Disequilibrium (LD) and short haplotype blocks that

would be expected due to its large size and long time period since domestication bottleneck (Karlsson and Lindblad-Toh., 2008).

The modern dog seems to be the domestic species with the most diversified phenotypic variations. There are currently more than 400 different breeds and each one of these seem to be characterized by unique behavioral and physical traits that have evolved in sequence of two main important population bottlenecks, the first during breed creation and second during the World Wars or depression as well as infectious disease breakouts, reducing the effective breeding population to only a few dogs. Strong artificial selection and tight inbreeding had also accumulated recessive disease alleles (Karlsson and Lindblad-Toh, 2008). These events had lead to unexpected health impairments, especially amongst purebred dog populations. Enrichment of risk alleles due to random fixation during bottlenecks, hitchhiking of unwanted mutations together with desirable traits and pleiotropic effects of selected variants had definitely contributed to disease incidence in modern dogs (Karlsson and Lindblad-Toh., 2008)).

The majority of canine breeds are quite recent (less than 200 years old) revealing extensive degree of linkage disequilibrium and long haplotype blocks within them, this due to the fact that only a small subset of chromosomes was selected from a given group of domestic dogs. The long-range patterns that were carried on these chromosomes became common within different breeds, creating long-range LD estimated at several megabases, 40–100 times longer than in humans (Lindblad-Toh *et al*., 2005). These patterns make canine models excellent for GWAS and mapping, where fewer markers and smaller number of individuals within a population are needed compared with human studies. In 2005, not only the dog genome was completely sequenced but also the development of resources needed for association mapping. Understanding canine haplotype structure and definition of dense SNP map are currently powerful tools used for high-throughput SNP genotyping of the dog. Besides, dogs share basically the same number of genes as humans, most of which are 1:1 orthologues (Lindblad-Toh *et al*., 2005). Identification of genes and causative loci in dogs can help to better understand pathogenesis and new pathways associated to the same disease in humans, opening windows to development of novel diagnostic tools, gene tests and new treatments and drugs. In addition, dissecting the network of interacting loci that cause diseases with complex, polygenic inheritance in dogs, and defining their severity, will give insight into the considerably more complex web underlying human diseases.

IMRD has a high degree of similarity with its homologue human disorder, and shows a complex pattern. Furthermore in NSDTRs there is a high prevalence of immune-mediated diseases including a disease-complex that resembles human systemic lupus erythematosus.

2.2 Shared Autoimmune disorders between dogs and humans

According to the Online Mendelian Inheritance in Animal (OMIA) database the information retrieved in 11.03.2013, showed that there are 587 inherited traits/diseases in dogs, from which 303 are considered as potential disease models for human diseases. The list of autoimmune diseases in dogs shared with humans is relatively long and includes alopecia, diabetes mellitus, epidermolysis bullosa, autoimmune hemolytic anemia, hypothyroidism,

myasthenia gravis, pancreatic insufficiency, pemphigus, systemic lupus erythematosus among others (Table 1). Pure-bred dogs seem to be the most affected.

Human autoimmune disease	Canine autoimmune disease
Alopecia areata	Alopecia
Anonychia-onycodystrophy	Onychodystrophy
Diabetes mellitus	Diabetes mellitus
Epidermolysis bullosa	Epidermolysis bullosa
Haemolytic anaemia, autoimmune	Haemolytic anaemia, autoimmune
Hypoadrenocorticism (Addisson)	Hypoadrenocorticism
Hashimoto's thyroiditis	Lymphocytic thyroiditis
Myastenia gravis	Myastenia gravis
Narcolepsy	Narcolepsy
Pancreatic insufficiency, Exocrine	Pancreatic insufficiency, exocrine
Pemphigus foliaceus	Pemphigus
Polyglandular autoimmune syndrome, type II	Polyglandular autoimmune syndrome, type II
Systemic lupus erythematosus	Systemic lupus erythematosus

Table 1 Common autoimmune diseases shared between human and dogs (http://omia.angis.org.au, accessed March, 2013).

2.3 Systemic Lupus Erythematosus in Humans and Dogs

Systemic lupus erythematosus (SLE) or lupus, is a human autoimmune disease characterized by a Type III hypersensitivity reaction caused by systemic deposition of autoantibody-immune complexes, where potentially any part of the body might be affected. Similarly to other autoimmune disorders, the immune system in an SLE-affected patient directs its attack to the body's self-tissues, cells and molecules. Localization of an inflammatory reaction followed by tissue damage characterizes the different disease phenotypes (James *et al.*, 2005). The progression of the disease cannot be predicted and is characterized by alternating periods of illness and remissions.

In SLE, the immune system starts producing antibodies against self-proteins and chromatin present in the cell nucleus. This event may eventually be triggered by different environmental factors (ultraviolet light, drugs, and viruses). These environmental stimuli may cause destruction of cell membranes and expose nuclear material like chromatin and nuclear proteins. In some individuals the immune system directs its defense against these nuclear-related proteins producing antibodies against them (Kindt *et al*., 2007).

There is no current treatment for lupus, and ongoing research is being conducted worldwide towards identification of causes (both genetic and environmental) and effective treatments in order to increase life quality of affected human and animal patients.

Similar conditions are also observed in other mammals including the dog. SLE-related diseases like immune-mediated rheumatic disease can be observed in German Shepherds, English Springer Spaniel, Boxer and Petit Basset Griffon Vendeen and Nova Scotia Duck Tolling Retrievers. NSDTRs were initially selected and developed in the end of 19th century in the Canadian region of Nova Scotia as a duck hunting and retrieving dog. During

initial stages of breed selection a couple of important Canine Distemper Virus outbreaks significantly reduced size of the NSDTR population to only a few individuals. There is some evidence supporting that progenitors that perpetuated this breed derived from the stock that survived this devastating canine viral outbreak resulting in a genetic bottleneck and a strong founder effect for the genetic risk factors of SLE (Strang *et al*., 1996). NSDTR seems to be amongst the most susceptible dog breeds to a specific set autoimmune disorders designated SLE-related diseases. Immune-Mediated Rheumatic Disease (IMRD), Steroid-Responsive Meningitis Arteritis (SRMA) and Hypoadrenocorticism.

No epidemiological studies have to date been reported in the scientific literature regarding SLE-related disease prevalence in dogs, however this condition seems to be relatively rare in dogs in general.

A study developed by Hansson-Hamlin, SLU and colleagues while working at Clinical Pathology department of the Uppsala University for ANA positive tests suggest that the canine breed NSDTRs are much more susceptible to IMRD with 26% incidence among total cases from different breeds (Jokinen, 2011).
SLE-related disease complex includes both IMRD and SRMA. Epidemiological and clinical studies suggest the prevalence of symptoms associated with chronic musculoskeletal signs with stiffness and pain from different joints, as well as aseptic meningitis, reporting high frequency in this breed (Hansson *et al*., 2009). One of the most important diagnostic criteria of SLE is the presence of high levels of circulating antinuclear antibodies (ANA) that can be detected by direct immunofluorescence-antinuclear antibody test *IIF-ANA*. Similarly to humans, it has already been attempted to identify a clear set or list of uniform and consistent phenotypes for SLE-related conditions in the dog but currently no such phenotypic criteria have been successfully defined (Hansson *et al*., 2009). Clinical signs may range from polymyositis, fever, neurological disorders, anemia, skin disorders and thrombocytopenia (Chabane *et al*., 1999).

2.4 Immune-mediated rheumatic disease (IMRD) and Steroid-responsive meningitis arteritis (SRMA) in NSDTRs

In human patients, rheumatic conditions develop secondary to autoimmune factors where clinical problems affect connective tissues, soft tissues and articular joints (Hansson et al, 2009). Similarly, affected dogs seem to share the same clinical signs revealing ANA positivity in 70% of cases and signs of polyarthritis in 100% of cases (Jokinen, 2011). Other commonly detected signs may include fever, dermatological disorders as well as liver and kidney problems. Blood analysis testing for the presence of rheumatoid factor can be performed, however, rheumatoid factor is in some cases present in dogs without rheumatoid arthritis, and not all dogs with rheumatoid arthritis test positive for the factor. Radiographic evaluation of the affected joints can be used for diagnostic purposes and typically show the morphological changes in bones - The joint surfaces are frequently irregular and bony spurs coming off of the bone are commonly seen. Another commonly used diagnostic method is to analyze the synovial fluid and determine whether inflammatory cells are present in large numbers, and whether the synovial fluid is less viscous than normal and may appear cloudy

(Hansson *et al.*, 2009).
NSDTRs reveal the onset of this condition between 2 months and 3 years of age and comparing to other breeds the frequency is markedly higher (Hansson *et al.*, 2009). In another SLE-related disease, SRMA, clinical signs are caused by a combination of meningitis and arteritis of leptomeningeal vessels. Two different forms of SRMA may develop in dogs. The acute, the most common and typical one, characterized by episodes of hyperesthesia along vertebral column and cervical rigidity, resulting in difficulties to turn and lower the head, stiff gait, pyrexia, depression and anorexia (Tipold *et al.*, 1994) mimicking intervertebral disc disease. Cerebrospinal fluid analysis reveals high numbers of polymorphonuclear pleocytosis and elevated protein concentration and red blood cells secondary to leakage from damaged vessels. Cerebrospinal fluid analysis regarding bacterial presence is routinely negative although it is not a commonly used diagnostic method. In some dogs, meningitis may also affect the meninges of the brain and the choroid plexus. Computed tomography scan or magnetic resonance imaging might reveal high contrast enhancement of the meninges. A second, more chronic relapsing form of SRMA has more reserved prognoses, with meningeal fibrosis secondary to initial inflammation that may obstruct cerebrospinal fluid flow or occlude the vasculature and potentially causing secondary hydrocephalus or ischemia of nervous system parenchyma (Summers *et al.*, 1995). Other neurological signs include variable degrees of paresis and ataxia, menace deficits, anisocoria or strabismus and mild to moderate levels of mononuclear cells or mixed cell population in cerebrospinal fluid and eventual raise in total protein (Tipold, 1995). Most of the affected dogs have elevated IgA levels in both blood serum and cerebrospinal fluid, a finding that is most likely associated to immune system deregulation (Felsburg *et al.*, 1992). In some situations, dogs may develop signs of polyarthritis, which is also present in all cases of IMRD-diseased dogs. The onset of the first signs of SRMA may usually develop at 4 to 19 months of age, and a lifelong corticosteroid therapy may be necessary to inhibit recurrence, which might happen in around 50% of all affected dogs. An epidemiological study developed in a Norwegian population of NSDTRs revealed an estimated prevalence of SRMA of around 2,5%. This value however, is most likely underestimated (Anfinsen *et al.*, 2008). In some cases, due to complications or consequences of corticosteroid long-term therapy, some dogs have to be euthanized.

2.5 Genetics of SLE-related disease in NSDTRs

The genetics of SLE is still not yet still completely understood. Developing genetic studies and subsequent functional assays are making efforts to acquire a better understanding of the etiology of this disorder. Although human studies reveal that SLE is pedigree related, no single unique causative gene has yet been identified. Instead, multiple genes may have a strong influence on individual's chances of starting to develop SLE when triggered by environmental factors (Martens *et al.*, 2009). Most frequently, associated allelic variants seem to be linked to both increase in disease risks and protection as well (Erman, 2001). Besides, studies focusing on genetic susceptibility to any given autoimmune disease are also complicated by the presence of protective haplotypes that seem to mask the susceptibility and completely inhibit the risk imposed by the susceptibility or risk haplotypes (Morel *et al.*, 1999). Monogenic autoimmune diseases are very rare (Ahonen, 1990). Several genes having small effects characterize polygenic nature of autoimmunity, and some of these

genes and pathways have already been identified. Most of the susceptibility genes are strongly related to apoptosis, signaling, auto antigen availability and clearance, and cytokine and co-stimulatory molecules expression. Major histocompatibility complex MHC class II locus, T cell receptor (TCR) regulation and a few others for example, have been strongly associated with autoimmunity in canine models (Wilbe *et al*., 2010).

As indicated above, incidence of autoimmune disorders, in particular SRMA and IMRD are overrepresented in the breed NSDTR (Hansson-Hamlin, 2009). A recent study developed by Wilbe aiming to determine whether these are indeed two different diseases sharing common genetic risk factors or on the other hand, represent two separate disorders was performed using GWAS, to identify genetic risk factors associated with disease. Five loci were successfully identified as associated with SLE-related disease complex using a two-stage mapping strategy. All of the associated regions contained several genes making identification of causative ones very difficult. Four very strong candidate genes were identified (*PPP3CA*, *HOMER2*, *DAPP1* and *PTPN3*) that seem to be involved in regulation of nuclear factor of activated T cells (NFAT) pathway (Fig. 1).

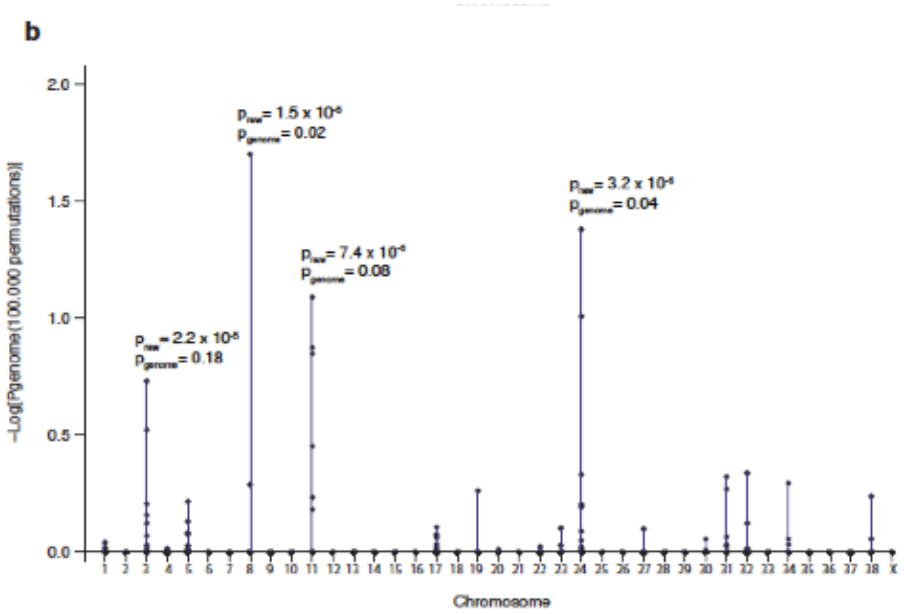

Figure 1. *Associated regions located on four chromosomes for ANA-positive SRMA and IMRD dogs (CFA 3, 8, 11 and 24). From: Wilbe 2009, Nature Genetics. 42(3):250-254*

Fine-mapping was also performed with the purpose to validate associated regions by addition of more NSDTR cases and controls. Three loci located on chromosomes 3, 11 and 24 associated with ANA-positivity were strongly validated, with p-values in the range 10^{-11}-10^{-13} (Wilbe *et al.*, 2010).

Expectedly, a few loci with a strong effect each may affect the development of this immune-mediated disease complex in NSDTRs and a combination of these genetic risk factors might be sufficient to predispose to this disease (Wilbe *et al.*, 2010). Further analysis associating GWA risk loci on CFA 3, 8, 11, 24 and 32 with ANA speckled phenotype dogs revealed strong associations to CFA 11 and 32. The region on chromosome 11 covering the *PTPN3* gene comprised a haplotype highly associated to ANA speckled (ANA^S) dogs (Wilbe *et al.*, 2013).

2.6 *Protein Tyrosine Phosphatase Non-Receptor type* 3 (*PTPN3)* Gene

In Humans, *PTPN3* is located in Chromosome 9: 112,137,974-112,260,593 reverse strand. So far, 7 alternatively spliced transcript variants encoding different isoforms have been found for this gene (www.ensembl.org, accessed April 2013), and comprises 122,620 bases. In the dog, and according to Ensembl Canis lupus familiaris version 70.31 (CanFam3.1) *PTPN3* is also encoded on reverse strand of chromosome 11 and comprises 122,2 Kb, 26 exons from which 25 are coding exons and 913 amino acids (www.ensembl.org, accessed March 2013). *PTPN3* gene is part of the same family as *PTPN22*, which was previously associated as a major genetic risk factor for SLE and other autoimmune diseases in humans (Chung, 2007).

2.7 Genetic variants in *PTPN3* gene in NSDTRs

A four SNP haplotype out of the 15 most highly associated SNPs (11:67,516,041, 11:67,538,032, 11:67,538,806 and 11:67,583,604) across the *PTPN3* gene in CFA 11 was used for correlation of mRNA expression levels between healthy and severely diseased IMRD and ANA^{S} dogs, and an 8-fold down regulation of mRNA levels (p=0.013) was identified in dogs carrying the risk haplotype (Wilbe *et al.*, 2013). A four SNP haplotype was generated across *PTPN3* (fig. 2) and most severe cases of IMRD that lead dogs to death were homozygous for the risk haplotype G/G-T/T-C/C-A/A (11:67,516,041, 11:67,538,032, 11:67,538,806 and 11:67,583,604, respectively). The protective haplotype for the same SNPs was defined by A/A-G/G-T/T-G/G, whereas the only risk haplotype present in healthy dogs was G/A-T/G-T/C-A/A and this haplotype was not correlated with any significant down-regulation of relative mRNA levels in *PTPN3* expression (Wilbe *et al.*, 2013).

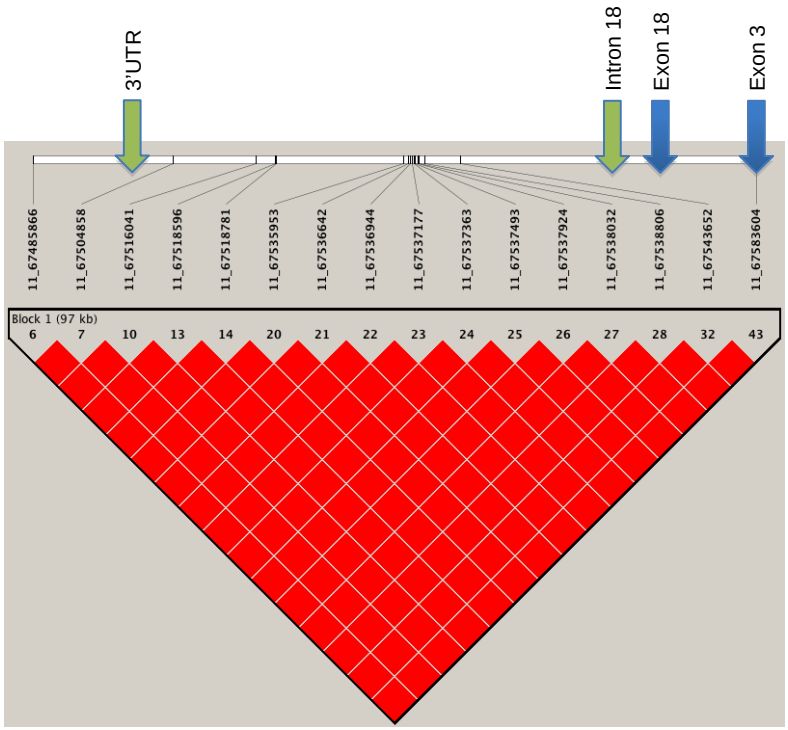

Figure 2*. A 16 SNP 97 Kb haplotype overlapping PTPN3 gene. The indicated four top most associated SNPs were previously analyzed for correlation to mRNA expression in three different haplotypes: two synonymous SNPs in exon 3 and 18, one on intron 18 and one in the 3'UTR (Wilbe et al. 2013)*

2.8 PTPN3 and Tyrosine Phosphorylation in TCR Signal Transduction

Tyrosine phosphorylation classified by four different families is represented by 107 genes in the human genome (Alonso *et al.,* 2004). However, only 81 of these are known to be catalytically active and dephosphorylate tyrosine residues. The domain of PTP has around 280 amino acids and is characterized by the signature motif (I/V)HCxxGxxR(S/T), that constitutes an important part of the phosphate binding pocket. Besides this, there are still 9 more motifs spread along the PTP primary sequence domains involved in substrate recognition and catalysis that are highly conserved (Andersen *et al.,* 2005). The PTP families can be subdivided into receptor-like and non-receptor PTP (Bauler *et al.* 2008). The latter consists of 17 members, including PTPN3; most of them have a modular domain that affects localization and function. PTPs (and from which PTPN3 is part of) are signaling molecules that regulate a multitude of cellular processes like cell growth and differentiation, mitotic cell cycle, and oncogenic transformation. P97, a cell cycle regulator involved in a variety of membrane related functions, has been shown to be a substrate of this PTP (www.ncbi.nlm.nih.gov/refseq/, accessed April 2013).

PTPN3 consists of an NH2-terminal FERM domain, a central PDZ domain, and a COOH-terminal *PTP* domain. Both PDZ and FERM bind the cytosolic domains of transmembrane proteins as well as phospholipids PI(4,5)P2. In molecular biology, the FERM domain is a common protein domain involved in localizing proteins in the plasma membrane (Chrishti, 1998). The FERM domain, is located at the N terminus of PTPN3 and seem to be required for PTP membrane localization in T cells (Bauler *et al*., 2008). Studies suggest that the phosphatase expressed by the *PTPN3* gene also participates in multiple functions in signal transduction pathways depending on cell contexts. Human PTPH1 encoded by the *PTPN3* gene acts in TCR-signaling and transduction as a negative regulator dephosphorylating the immune tyrosine-based activation motifs (ITAM) in the TCRζ chain. Downstream in this pathway it will potentially inhibit activation of nuclear factor of activated T- cells (NFAT) (Sozio *et al.,* 2004). The previously detected reduction of the *PTPN3* mRNA levels in dogs carrying the risk haplotypes may cause a permanent activation of TCR signaling and lead to development of autoimmune disorders (Wilbe *et al*., 2013).

T lymphocytes originate from a common hematopoietic stem cell progenitor in bone marrow to all different blood cell types. Unlike B cells, which mature within the bone marrow, T cells migrate to the thymus gland where they achieve maturation. During this process within the thymus, the T cell will start to express a unique antigen-binding molecule, called the T-cell receptor (TCR), on its membrane (Kindt *et al.*, 2007).
The signaling pathways stimulated by these receptors use PTP to transduce such signals. Particular signals originating from TCR seem to be associated at least partially with T cell maturation and function (proliferation, differentiation growth, migration, B cell helping and target cell lysis) (Bauler *et al.,* 2008). T cells express around 14 non-receptor PTPNs and they are considered as negative regulators of TCR signaling (Musteling, 2005). In a study conducted by Gjorloff-Wingren, over-expression of *PTPN3* gene in T cell leukemia cell lines had deeply inhibited TCR-dependent signal transduction leading to constitutive activation of the T cell growth promoter cytokine IL-2. This interesting finding strongly suggests that PTPN3 may inhibit TCR-signaling in T cell leukemia, through

dephosphorylation of the plasma membrane-localized substrate, contributing to the idea that PTPN3 acts a negative-regulator of TCR signaling by dephosphorylating the TCR (Bauler, 2009).

4 MATERIALS AND METHODS

4.1 Primer Design

For assistance with designing and calculating the Tm of primers Amplify 3.1 software was used.
Primer design for Intron 18 for sequencing; 5'-ACCATCTGCATTCAGGTGTCGT-3´ (PTPN3-f-in18-1), 5'-AGTAACGCACATTATCTACATGCTAC-3' (rev-PTPN3-int18-1), (amplicon with 1311 bp). 5'- GAAGAGCAAACACCATAAAGCA-3' (PTPN3-f-int18-2) 5'-GCAGCGTGTCCATCAGAAATC-3' (Rev-PTPN3-int18-2), (amplicon with 1403 bp), and 5'-TATCCTTGGTCTGTAGCTGTTTCA-3' (Rev-PTPN3-in18-3), (amplicon with 715 bp)
Primers for intron 18 with restriction sites for cloning,;
5'GTACCgagctcAAGTGGGAACACAGGAGATTCG-3' (in18-f-clon1-SacI), 5'-GTACCagatctCCTTGGAGAGACACTACTCAAAACT-3' (Rev-in18-clon-BglII) (amplicon **1** with 1613 bp) and 5', covering 7 SNPs (Table 2)
GTACCgagctcTATACATTGTAAACACTGTAAAACTCA-3' (in18f-clon2-SacI) with same 5'-GTACCagatctCCTTGGAGAGACACTACTCAAAACT-3' (Rev-in18-clon-BglII) (for amplicon **2** with 838 bp).
Primer design for PTPN3, 3'UTR, for sequencing: 5'-AGCTGGACCTCAGGGCATATTTAT-3' (PTPN3-3'UTR-forw-1), 5'-GTTCCCATGGTTCCATGACTAGAT-3' (Rev1-PTPN3-3'UTR), 5'-CAACTAAATGTGCTAAGTCTGGAGATA-3' (Rev2-PTPN3-3'UTR): Amplicon 1 with Rev1 primer = 1174 bp and amplicon 2 with with Rev2 primer = 1785 bp.
Primers for cloning of 3'UTR: 5' GACAACTGTGAAAAACGTTCATTC-3' (PTPN3-3UTR-for-clon), amplicon with Rev2-1591 bp covering three SNPs (Table 2);
5'-GTCGCGGATCCCAACTAAATGTGCTAAGTCTGGAGATA-3' (Rev2-PTPN3-BamH1)

4.2 PCR amplification

For primer testing and DNA amplification for sequencing purposes, Platinum Taq DNA Polymerase (Invitrogen) was used. This is a widely used hot-start enzyme for robust and reliable amplifications. Platinum Taq is an antibody-mediated, hot-start enzyme that provides high PCR specificity ranging from cloning to genotyping to mutagenesis (Invitrogen, 2013). 11 DNA samples from risk and protective haplotypes were diluted in a 25 µl reaction mixture containing 0,1µl Platinum Taq, 1 µl dNTPs, 0,8µl $MgCl_2$, 2,5µl 10x PCR Rxn Buffer (Invitrogen) and 1 µl primers for 3'UTR (four sets) and intron18 (three sets). PCR conditions were as follows: 95°C for five min, followed by 45 cycles of 95°C for 15 sec, 60°C for 15 sec, then 72°C for 2 min and 72°C for 3 min final extension and 4°C

final hold. PCR product reactions were sent for sequencing at Sanger Sequencing Service (Uppsala Genome Center) after post PCR purification. PCR product purification was achieved using 0.1µl Exonuclease I (Fermentas) and 0.1µl Alkaline Phosphatase (Promega) added to 2µl Buffer SAP (Fermentas) per 25µl tube reaction, and incubated at 37°C for 1 h followed by 85°C for 15 min. Exonuclease I is a single strand exonuclease that allows degradation of primers still existent in the reaction tubes. Alkaline phosphatase on the other hand, allows removal of dNTPs still present in the PCR reactions. As for PCR amplification for cloning, incorporation fidelity for high fidelity amplification is an important asset. Therefore, we used a thermostable enzyme with proofreading to allow good fidelity in amplification reactions. Pfu proofreading enzyme has the potential to offer three- to six fold higher fidelity than standard DNA polymerase (Cline *et al*., 1996). PCR amplicons generated with proofreading Pfu polymerases are blunt ended giving an A-tail for T-vector subcloning (Knoche, 1999). DNA from risk and protective haplotypes was diluted in a 25 µl reaction mixture containing pfu Buffer, 1 µl dNTPs, 0,5 pfu ultra enzyme and 1 µl primers for 3'UTR and 2 amplicons of intron18 region. PCR conditions were as follows: 95°C for five min, followed by 45 cycles of 95°C for 15 sec, 60°C for 15 sec (57°C for intron 18) then 72°C for 2 min and 72°C for 4 min final extension and 4°C final hold. The oligonucleotide primers for PTPN3 3'UTR and intron18 were synthesized for us by IDT and were as previously described.

4.3 Analysis of PCR products by Agarose Gel Electrophoresis

PCR products were resolved using 1% agarose gel electrophoresis in TAE (Tris-Acetate-EDTA) Buffer. Gels were run at 120 volts for 20-25 min. To make the gels, 0,5g of Ultrapure agarose powder (Invitrogen) was measured in a 300ml conical flask and 50ml 1xTAE Buffer added. The flask was heated in a microwave oven until agarose was dissolved . 3 µl SYBR Safe DNA Gel Stain, a safe nucleic acid stain for detection DNA in agarose gels, was added to the flask and the gel poured into the gel tray with the comb. Gels were allowed to set and combs removed. 4 µl PCR products were loaded into the wells with ≈3µl of 6x Loading DNA Dye (Thermo Scientific). 2µl of GeneRuler 1Kb DNA ladder (Thermo Scientific) was also loaded for molecular weight reference. Gel was subsequently placed against UV Light using BioRad CCD camera image station and Quantity-One 1D analysis software for image capture.

4.4 Sequencing of *PTPN3* exon 18, intron 18 and 3'UTR regions

Results from sequencing were analyzed using Sequencher software, where all 11 samples including 4 protective haplotypes, 2 risk heterozygous and 5 risk homozygous haplotypes were genotyped and compared for 1 SNP in exon 18, 7 SNPs in the intron 18 and 3 SNPs for 3'UTR (Table 2).

Exon18

Dog ID		Dog1 (913)	Dog2 (848)	Dog3 (658)	Dog11 (634)	Dog4 (828)	Dog5 (900)	Dog6 (72)	Dog7 (73)	Dog8 (81)	Dog9 (100)	Dog10 (267)
SNP1	67538806	T/T	T/T	T/T	T/T	C/T	C/T	C/C	C/C	C/C	C/C	C/C

Intron18

Dog ID		Dog1 (913)	Dog2 (848)	Dog3 (658)	Dog11 (634)	Dog4 (828)	Dog5 (900)	Dog6 (72)	Dog7 (73)	Dog8 (81)	Dog9 (100)	Dog10 (267)
SNP1	67538032	G/G	G/G	G/G	G/G	G/G			T/T	T/T	T/T	T/T
SNP2	67537924	C/C	C/C	C/C	C/C	A/C		A/A	A/A	A/A	A/A	A/A
SNP3	67537493	C/C	C/C	C/C	C/C	C/T	C/T	T/T	T/T	T/T	T/T	T/T
SNP4	67537363	C/C	C/C	C/C	C/C	C/T	C/T	T/T	T/T	T/T	T/T	T/T
SNP5	67537177	A/A	A/A	A/A	A/A	A/G	A/G	G/G	G/G	G/G	G/G	G/G
SNP6	67536944	C/C	C/C			C/T			T/T	T/T	T/T	
SNP7	67536642	A/A	A/A	A/A		A/C			C/C	C/C	C/C	C/C

3'UTR

Dog ID		Dog1 (913)	Dog2 (848)	Dog3 (658)	Dog11 (634)	Dog4 (828)	Dog5 (900)	Dog6 (72)	Dog7 (73)	Dog8 (81)	Dog9 (100)	Dog10 (267)
SNP1	67516230	G/T	G/C	G/G	G/C	G/G	?	G/G	G/G	G/G	G/G	G/G
SNP2	67516211	T/T	T/T	T/T	T/T	T/T	?	T/T	T/T	T/T	T/T	T/T
SNP3	67516041	A/A	A/A	A/A	A/A	G/A	G/A	G/G	G/G	G/G	G/G	G/G

***PTPN3* Haplotypes**

Protective : **A/A-G/G-T/T**

Risk homozygous: **G/G-T/T-C/C**

Risk heterozygous: **G/A-G/T-C/T**

Table 2: *PTPN3* Sequencing analysis of different SNP genotypes of protective and risk haplotypes for exon 18, intron 18 and 3'UTR from NSDTRs. Analysis performed with Sequencher software. Highlighted indicates SNPs associated with defined haplotypes.

4.6 Molecular Cloning

DNA inserts from 3'UTR and intron 18, and pGL3 promoter plasmid were subject to different steps of purification and digestion with specific restriction enzymes prior to ligation.

For the 3'UTR region, DNA from protective and risk haplotypes were digested with *BamH*I only. pGL3 promoter was cleaved in two distinct restriction sites: *BamH*I and *Xba*I, leading to removal of 260 bp SV40 late poly(A) signal. *Xba*I end of plasmid was "polished" with pfu Ultra turning it into blunt ends on that side and kept as sticky ends on *BamH*I side (Fig.3).

For intron 18 region, DNA from both protective and risk haplotypes and pGL3 were digested with *Sac*I and *Bgl*II. A cleavage after synthetic poly(A) signal pause site was created for the insertion of the 2 intron 18 amplicons (Fig. 4).

To measure nucleic acids concentration (3'UTR, intron 18 inserts and pGL5 plasmid) 1µl of the samples was analysed on a NanoDrop ND-1000 Spectrophotometer according to the manufacturers instructions.

Ligation was performed according to manufacturer protocol with T4 DNA ligase (Thermo Scientific) and PEG in a 20 µl volume at a molar ratio of 3:1 of DNA to vector and incubated at 12°C for 15 min and then 22°C for 60 min.

Restriction sites for 3'UTR insert in pGL3 promoter vector:

*BamH*I
5′...G▼GATCC...3′
3′...CCTAG▲G...5′
*Xba*I
5′...T▼CTAGA...3′
3′...AGATC▲T...5′

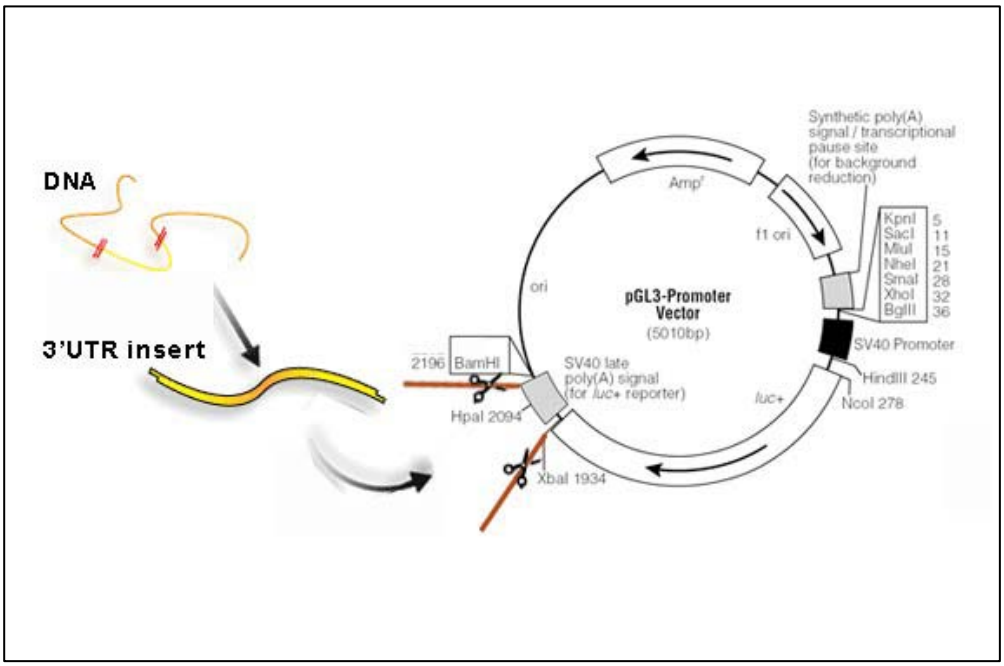

Figure 3*. Schematic representation of pGL3 promoter vector with restriction sites and 1.6 Kb insert location for 3'UTR.*

Restriction sites for intron 18 insert in pGL3 promoter vector:

*Sac*I
5′...GAGCT▼C...3′
3′...C▲TCGAG...5′
*Bgl*II
5′...A▼GATCT...3′
3′...TCTAG▲A...5′

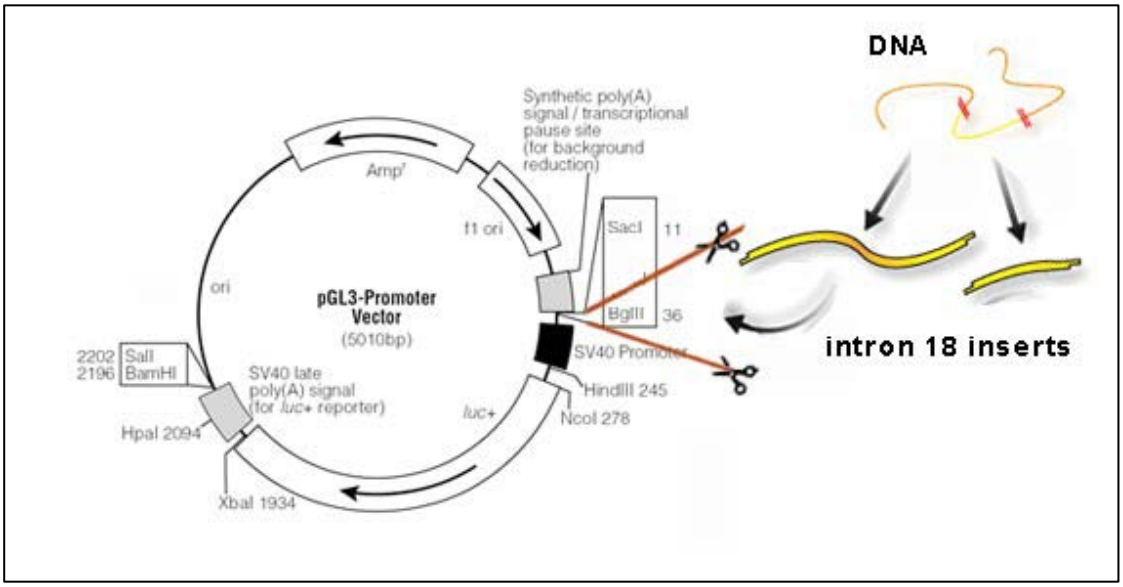

Figure 4*. Schematic representation of pGL3 promoter vector with restriction sites for two intron 18 inserts (1613bp and 838bp) ligation.*

50 µl chemically competent *E. coli* One Shot TOP10 (Invitrogen) cells were thawed on ice for each ligation/ transformation. 4 µl of each ligation reaction was pipetted directly into each vial of competent cells tube and tapped to mix. The tubes were then left on ice for 25 min, heat shocked at 42° C for exactly 30 sec in water bath and placed back on ice. 250 µl of pre-warmed S.O.C medium was added to each vial using sterile technique. Vials were then left in the shaker incubator at 225 rpm, 37°C for exactly one hour. Tubes were taken out of the shaker, and the contents streaked onto the selective LB (50 µg/ml Ampicillin) ~~separate and labeled~~ agar plates with 20 µl and 100 µl for each transformation and incubated overnight at 37°C.

In order to be sure about successful transformation of competent cells that have incorporated the plasmid vector pGL3 plus DNA inserts, 15 colonies from each 100 µl seeded petri dish plates were "picked up" with pipette tip and cultured in 1,5 ml properly labeled eppendorf tubes containing 150 µl of LB medium + 1µg/ml Ampicillin and incubated for 5-6 hours in shaker at 225 rpm, 37°C.

Plasmid DNA was initially extracted and purified for identification of potential positive colonies, using phenol/chloroform. 120 µl of cell suspension were aliquoted into new eppendorf tubes and centrifuged for 3 min at 13 000 rpm. After removal of supernatant by pipetting, 30 µl of TE buffer (10 mM Tris-Cl pH 8, 0.1 mM EDTA) were added to each tube and vortexed for 30 sec. 30 µl of phenol/chloroform were additionally added and vortexed again for another 30 sec. Centrifugation of the suspension allowed separation of purified DNA in the upper phase, which was afterwards loaded (20 µl) in a 1% agarose gel together with pGL3 control plasmid vector and 1Kb DNA ladder. Successful plasmid /insert incorporation was promptly identified by corresponding size bands after 1% agarose gel running for 25 min at 120 V. pGL3 vectors with 3'UTR revealed 6.3 Kb size bands (against 5Kb control vector and non transformed colonies), and intron 18 positive transformants revealed 6.6Kb (amplicon 1) and 5.8Kb (amplicon 2) size bands according to the two different size inserts. 1 µl of initial positive transformed cells previously aliquoted were

then recultured in 6 ml of ampicillin-enriched LB medium and incubated at 37°C at 200 rpm shaker incubator for approximately 14,5 hours, until culture medium became milky in appearance.

DNA extraction of transformed bacterial cultures for transfection was done using QIAprep Spin Miniprep Kit according to manufacturer's instructions.
The pelleted bacterial cells were resuspended in 250µl Buffer P1 and transferred to a microcentrifuge tube, 250µl Buffer P2 was added and mixed thoroughly by inverting the tube 4-6 times, 350µl Buffer N3 was added and mixed thoroughly by inverting the tube 4-6 times and then centrifuged 10 min at 13,000 rpm to obtain a compact white pellet. Supernatant was then decanted into the QIAprep spin column, centrifuged for 60 sec and the flow through was discarded. The column was washed by adding 750µl Buffer PE and spinning for 1 min and then again another minute after removal of any residual wash buffer. QIAprep spin column was finally placed in a 1.5ml microcentrifuge tube and 50µl Buffer EB (10mM Tris-Cl, pH 8.5) added to the center of the spin column and left to stand for 1 min, then centrifuged 1min to elute and collect DNA.

4.6.1 Sequencing of Plasmid DNA

Sequences of cloned DNA inserts were analyzed to confirm adequate integrity of alleles and to detect possible mutations. After PCR, all product reactions were sent for sequencing at Sanger Sequencing Service (Uppsala Genome Center).
Regarding 3'UTR two different alleles in protective haplotype were used for transfection to determine possible functional variations between these and the risk variant (Table 3).

Cloned 3'UTR-pGL3

Dog ID		sample3 (848)	sample10 (848)	sample11 (848)	sample13 (848)	sample19 100)	sample20 (100)	sample22 (100)
SNP1	67516230	G	G	G	C	G	G	G
SNP2	67516211	T	T	T	T	T	T	T
SNP3	67516041	A	A	A	A	G	G	G

Haplotypes for cloned PTPN3 3'UTR

2 Protective: G/A and C/A

1 Risk homozygous: G/G

Table 3: Alleles of 2 SNPS for 3 cloned variants of *PTPN3* 3'UTR-pGL3 construct.

4.7 Transfection

4.7.1 MDCK Cells

Madin Darby Canine Kidney cells (MDCK), were originally obtained by S. H. Madin and N. B. Darby from the kidney tissue of an adult female cocker spaniel dog, in September 1958. The MDCK line has been mainly used as a general model for epithelial cells such as epithelium (Hamamatsu Learning Center, April 2013), in this case they are useful because

of their culture adherence properties, resistance, suitable transfection host, and ability to express PTPN3. In our study we used MDKC NBL2 line from ATCC (American Type Culture Collection). Cell line was grown in DMEM (Gibco) supplemented with 10% FBS and antibiotics. Cells were cultured in T25 flasks at 37°C in a humidified air/CO_2 atmosphere and passaged after 70- 80% confluence which usually occurred in two days (Fig. 5).

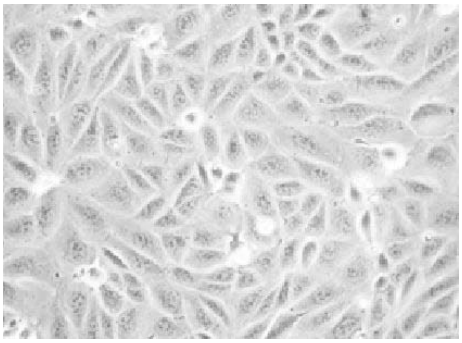

Figure 5*. MDCK cells in high confluence.*

Thawing of cryopreserved cells:
Stocks of MDCK cell cultures were stored in liquid nitrogen tanks. To revive cells, they were transferred to a 37°C water bath for about 2 min then rapidly thawed in 10ml of complete medium (DMEM) in a 15ml falcon tube. Cells were centrifuged at 125xg for 5 min, the supernatant discarded and the remaining pellet resuspended in 10ml of complete medium. Cell suspension was then afterwards transferred to a T-25 flask and incubated at the 5% CO_2 concentration and 37°C temperature. MDCK cells were maintained in Dulbecco's modified Eagle's medium, DMEM (Gibco), supplemented with 10% fetal calf serum (FCS) (Gibco) and penicillin/streptomycin solution (1%) (Gibco) and L-Glutamine solution (1%) (Gibco). Cells in culture T flasks were maintained by periodic passaging. When confluence was reached cells were trypsinised and subcultured.

4.7.2 Jurkat Cells

The Jurkat cell line is derived from a T-cell leukemia and was established from the peripheral blood of a 14 year old boy by Schneider et al., (year) and was originally designated JM. The line was cloned from cells obtained from Dr. Kendall Smith. This is a clone of the Jurkat-FHCRC cell line, a derivative of the original Jurkat cell line.
Jurkat cells are IL-2 producing T lymphocyte cells, commonly used to study T cell signaling. Jurkat cells are easy to culture, and grow rapidly (Fig. 6). The base medium for this cell line is RPMI Medium (Gibco), and to make the complete growth medium, FCS to a final concentration of 10% was added. Cultures were maintained in suspension by the addition of fresh medium or replacement of medium. New cultures were also established by centrifugation with subsequent resuspension at around 10^5 viable cells/mL. Incubation was maintained at 37°C and 5% CO_2 concentration.

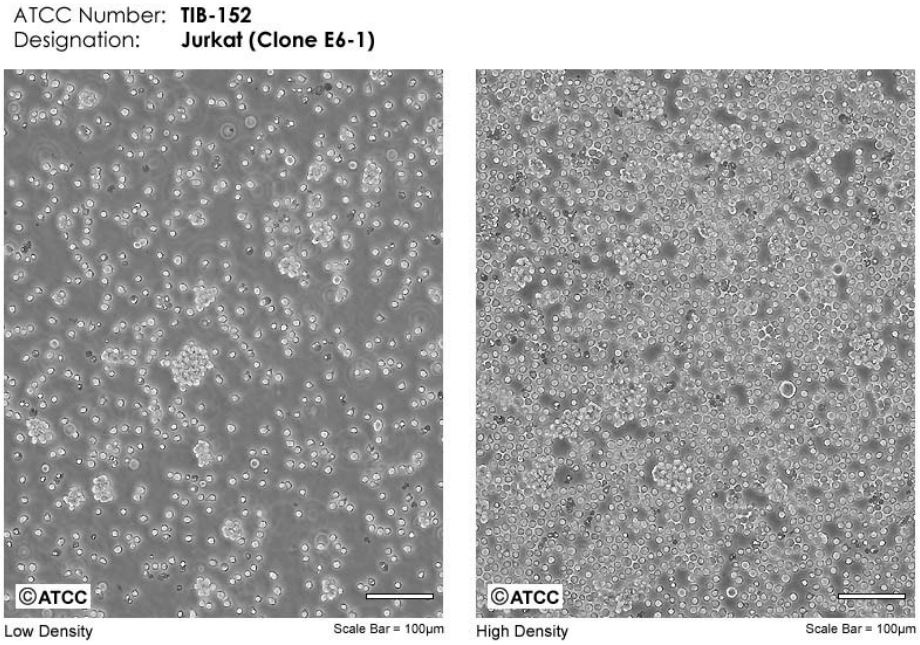

***Figure 6**: Jurkat T-Cells in suspension. Low and high confluence intensities are shown.*

4.7.3 Transfection with Plasmid DNA

Transfection is a method to introduce nucleic acids into eukaryotic cells using nonviral techniques. Although different approaches can be used for this method, we have successfully used lipid technology with Lipofectamine 2000 (Invitrogen).

Plasmid vectors containing luciferase reporter genes can be utilized to effectively quantify expression levels in the cells. Reporter gene assays may be performed 1 to 3 days after transfection, but usually 48 hours. MDCK cells were previously plated about 24h in advance prior to transfection. DMEM medium was aspirated and cells thoroughly washed with 10 ml PBS. PBS is carefully aspirated afterwards. Three ml of Trypsin (0,75%) in PBS was added to the culture flask and spread evenly and agitated to encourage detachment from the flask. Incubation at 37°C for 10-15 min in Trypsin allowed faster detachment of cells. Neutralization with 10 ml of complete Medium (DMEM + 10% FBS) without antibiotics was pipetted over the cells and pipetted up and down to avoid any cell clumps that may have formed. Cell suspension was transferred to a 50 ml falcon tube leaving 1 to 2 ml in the original flask to maintain cells stock (after adding 10 ml of complete DMEM fresh medium and transferred to incubator). Complete DMEM medium with antibiotics was added to falcon tube where remaining cells were until 25 ml were reached. Cells were centrifuged at 1200 rpm for 5 min and supernatant removed carefully without disturbing the pellet. Detachment of pellet rubbing the tube against a grid and addition of another 20 ml of medium for resuspension was preformed prior to cell counting. Cells were manually counted with haemocytometer (average of 4 grids), estimating the number of cells of round 99 $x10^4$ /ml. For transfection we needed 500 000 cells (0,5 ml) per well and 24 wells per plate. Cell culture plates were incubated at 37°C for 24 hours.

4.7.6 Reporter Gene Luciferase

Genetic reporter systems have an important role in the understanding of gene regulation and expression of eukaryotes mainly as indicators of transcription activity in cells. The reporter construct is inserted in a promoter sequence in a expression vector, in our case the pGL3 promoter, and transfected to *in vitro* cells followed by measurement of intensity of luciferase protein itself or its enzymatic activity. This type of assay is sensitive, quantitative, rapid, reproducible, and allow functional analysis of different elements in the genes since the expression of reporter protein is directly correlated with transcriptional activity of reporter gene. The Dual-Luciferase Reporter (DL) Assay System allows making efficiently two simultaneous assays. In the DL Assay, the activities of firefly and *Renilla* luciferases are quantitatively measured in a sequence from the same sample. By these means firefly luciferase is quantified by adding Luciferase Assay Reagent II to allow production of a chemiluminescent signal that is maintained by around one minute followed by the *Renilla* luciferase reaction adding a different reagent to the same sample. The luminometer may take as few as 4 sec per sample and is operated by reagent auto-injectors. This method is extremely sensitive providing a fast quantification method (www.promega.com, accessed May 2013).

4.8 RNA Degradation Assay in Jurkat T cells

The amount of mRNAs in the cell is the final result of mRNA synthesis and degradation. Reporter gene expression profiling experiments like the ones we performed in this project, indirectly take a snapshot of the total amount of mRNA levels and protein translation in the cell and do not capture all the dynamics of mRNA synthesis and breakdown during a certain amount of time ('t Hoen *et al*., 2010). The current assay determines and surveys mRNA degradation rates during a 12 hour period in specific time point intervals, for the Jurkat cells transfected with 3'UTR inserts. 3′-UTRs with different polymorphisms like those we studied here, may contribute to differential stability of transcripts and degradation rates of luciferase expression that might possibly change according to the different risk and protective variants, although final expression analysis seem to be similar between them. RNA measurement was performed by reverse transcribing RNA into cDNA, and quantify it with quantitative real time PCR (qPCR).

Jurkat T cells were grown in proliferation medium RPMI supplemented with 10% FBS without antibiotics and transfected with 1.4 μg of construct DNAs from the two protective and risk haplotypes as well as control transfections with pGL3 promoter vector without constructs, and 3.5 μg of Lipofectamine 2000 per well, each containing approximately 1,6 million cells. Cultures of transfected proliferating Jurkat cells were left incubating in 12-well dish plates for 24 hours containing 2-ml medium and activated with PMA+IO 5 hours after transfection according to same protocol used for pilot assay previously described. This period of time allows activation of all possible pathways for 3'UTR mRNA degradation including miRNA. After 10 hours of PMA+IO activation, 3 μl at 10 ηg/ml (per well) of actinomycin D was started to be added to arrest the transcription machinery and *de novo* RNA synthesis, at five different time points after addition of initial actinomycin D treatment

(0, 2, 4, 8 and 12 hours). RNA was extracted from around 800,000 cells/ well and DNA extracted from around 720,000 cells /well.

4.8.1 Two-step RT-qPCR

The quantification of mRNA using Reverse Transcription quantitative Real Time PCR (RT-qPCR) was performed using a two-step reaction. In this case it required that the RT reaction and qPCR amplification were performed in separate tubes. One major advantage of this approach is high reproducibility of the analysis. This method is also good when using DNA binding dyes like the one used, SYBR Green since the elimination of primer-dimers can be achieved by a simple change in the melting temperature (Wong *et al*., 2005).

RNA was isolated with TRIzol Reagent (Invitrogen) according to manufacturers instructions. After removal of growth media from culture dishes we added 1 ml TRIzol Reagent directly to the cells in the culture dish and cells were lysed directly by pipetting up and down several times. 0.2 mL of chloroform were added and used for homogenization shake tubes vigorously by hand for 15 sec. After 3-4 min incubation at room temperature samples were spun at 12,000 × g for 15 min at 4°C. Afterwards, we removed the aqueous phase of the sample by angling the tube at 45° and pipetting the solution out avoiding drawing any of the interphase or organic layer into the pipette when removing the aqueous phase. And transferred into new labeled eppendorf tubes. 0.6 mL of 100% isopropanol were added to the aqueous phase, and incubated at room temperature for 30 min and spun down at 12,000 × g for 15 min at 4°C. Finally, the supernatant was removed from the tubes and RNA pellet washed with 1 ml of ethanol 75%, vortexed briefly and centrifuged at 7500 x g for 5 min at 4°C. Wash was discarded and procedure was repeated once. We have made a final washing with 99% ethanol followed by a 3 min centrifugation at 7500 x g to allow for a better pellet drying. For RNA resuspension 25 µl of RNase-free water were added to the RNA pellet.

Since we convert the RNA back into DNA via reverse transcriptase, and then attempt to amplify and quantify target luciferase that is expressed in Jurkat T cells, any genomic DNA present could also be potentially amplified, leading us to question whether the final results of RT-qPCR are due to the cDNA that we want to generate, or potentially the presence of contaminating genomic DNA. To prevent this, RNA samples were treated with DNase, an enzyme that selectively degrades DNA according to the following protocol:

Reaction Components
RNA 6 µl
RQ1 DNase buffer (Promega) 2 µl
RQ1 DNase (Promega) 1.5 µl
Nuclease-free water to bring a total volume of 20 µl

Samples were incubated at 37°C for 55 min after which 2 µl of Stop Solution EGTA and incubated again at 65°C for additional 10 min to inactivate DNase. DNA was extracted to be used as “control” or for normalizing amounts of cDNA derived from RNA of

corresponding cell samples, in order to give us better reproducibility of results, calculating amplified cDNA/ DNA ratio. DNA extraction was performed using DNeasy Blood and Tissue (QIAGEN) according to manufacturer's recommendations. Around 100,000 cells per well were centrifuged for 5 min at 300 x g and the pellet resuspended in 200 µl PBS. 20 µl proteinase K and 4µl RNAse A at 30 mg/ml were added and left incubate for 2 min at room temperature. We added 200 µl Buffer AL (without added ethanol) and mixed thoroughly by vortexing. Samples were incubated at 56°C for 10 min, proceeded by addition 200 µl ethanol (99%) to the sample, and mixed thoroughly again by vortexing. The mixture was then pipetted into the DNeasy Mini spin column placed in a 2 ml collection tube and centrifuged at 8000 rpm for 1 min. The flow-through and collection tube were discarded and the DNeasy Mini spin column was placed in a new 2 ml collection tube and 500 µl of Buffer AW1 was added and centrifuged again for 1 min at 8000 rpm. Both flow-through and collection tube were discarded after this step. We placed the DNeasy Mini spin column in a new 2 ml collection tube and added 500 µl of Buffer AW2, and spin down for 3 min at 14,000 rpm in order to dry the DNeasy membrane. Again, the flow-through and collection tube were discarded in order to assure the membrane was left dry to assure no residual ethanol was left. Elution was carried by placin the DNeasy Mini spin column in a clean and properly labeled 1.5 ml o microcentrifuge tube, and by pipetting 200 µl of Buffer AE directly onto the DNeasy membrane and left Incubate at room temperature for 2 min. Centrifugation for 1 min at 13000 rpm allowed us to finally collect DNA from cell samples and estimate DNA concentration from all samples with nanodrop.

4.8.2 Reverse Transcription (RT)

RT reaction is also called first strand cDNA synthesis. Single-stranded RNA was reverse transcribed into cDNA by using 4 µl of the total cellular RNA or poly(A) RNA, 2 µl of RT Buffer II, 0.8 µl MULV-RT (Clontech) reverse transcriptase enzyme, 0.8 µl dNTPs (25 mM) and 0.4 µl RNase inhibitor. Two types of primers were used for RT reaction: oligo (dT) primers 0.5 µl and 1 µl of random (hexamer) primers in a final volume of 20 µl in PCR tubes. RT reaction was left incubate 10 min at room temperature and extended at 42° for 80 min and finally heated at 95° to inactivate the enzyme for 5 min. The resulting cDNA could then be used in qPCR reaction.

4.8.3 qPCR with SYBR green

Luciferase primer sets were used to target our samples for the real- time PCR reaction. SYBR green reagent was also used allowing detection of increases in DNA per cycle as binding double stranded DNA is being synthetized and light is emitted when excited. The light emitted during the process is quantified and used to determine the abundance of the gene product in the samples. Extracted DNA samples were diluted 1:500 (final concentration ± 50 pg/µl) and reverse transcribed RNA (cDNA) diluted 2 times in deionized water.
A MicroAmp optical 384-Well Reaction Plate (Applied Biosystems) were loaded in duplicates and for both cDNA and DNA with 19 µl of SYBR green master mix according to the following protocol (for each cDNA/ DNA sample):

cDNA/ DNA	1 µl
H_2O	14.2 µl
Primer F Luc (Thermo Scientific)	0.4 µl
Primer Rev Luc (Thermo Scientific)	0.4 µl
10x PCR Rxn Buffer (Invitrogen)	2 µl
$MgCl_2$ (1.5 mM final concentration)	0.7 µl
dNTPs	0.8 µl
Plat-Taq	0.1 µl
SYBR green (Invitrogen)	0.4 µl
Total	20 µl

Plate was loaded in a 7900 Fast Real-Time PCR System (Applied Biosystems) and PCR program set up according to following:

95° 5min
45 cycles
95° 15s
62° 15s
72° 20s
Melting Curve
95° 15s
62° 15s

5 RESULTS AND DISCUSSION OF BIOINFORMATICS AND FUNCTIONAL ANALYSIS

5.1 Bioinformatics and Molecular Analysis of PTPN3 intron 18

The first stretch of the sequence of intron 18 comprising around 2,000 bp of *PTPN3* gene was screened with different bioinformatics tools and databases in order to assess constraint regions and regulatory elements associated with 8 identified SNPs in this region to establish possible associations between the different SNP haplotypes and functional regulatory elements.

The sequence was initially screened on Santa Cruz database using Broad/canFam2 assembly and compared with Human hg/18 for constrained elements. Sequence was analyzed with different databases for SNP conservation, functional elements, consensus, conserved transcription factor binding sites and constrained elements and motifs between Dog, Human, Mouse and Rat and again between 29 mammals. A number of tracks were taken into consideration for this analysis.

The two identified constraint regions among 29 mammals were screened with TESS, CONCITE, TFSearch for Transcription Factor Binding Sites (TFBS). The regions were also screened with the software Sequence Manipulation Suite- DNA Finder

(http://bioinformatics.org/sms) for identifying DNA patterns and motifs. On the highest conserved region CONSITE detected the presence of 2 TFBS for AML 1 (cutoff 85 %). AML1 binds to the core site, 5'-PYGPYGGT-3', of a number of enhancers and promoters, including T-cell receptor enhancers. However, there were no consistent results between CONSITE and TESS. The first highly conserved region however does not reveal any binding sites or significant pattern or motif. This highly constrained region may be associated with additional TFBS, novel regulatory features or simply chance of stretches of sequence similarity.

The other conserved regions observed between Dog, Human, Mouse and Rat, were screened by TESS and CONSITE for putative TFBS and all revealed to have in common a GATA 3 binding motif within their regions for selected thresholds above 80% detected both in positive and negative strands. Similar screening was performed along the whole 2Kb sequence by CONCITE for GATA 3 with a transcription factor score cutoff of 90% and the results predict more putative sites along first 1Kb region (fig. 7).

Since GATA 3 is a transcription factor that is known to be functional for the TCR-beta – transcriptional enhancer and it binds to several T-cell receptor regulatory elements, this transcription factor was screened more cautiously around this whole region. TFBS GATA patterns were screened in the whole 2kb region by Sequence manipulation Suite software and 4 100% matches were found in the region around first 4 SNPs against 1 GATA match for second region around the other 4 SNPs, suggesting higher affinity for GATA family for first 1kb of intron 18.

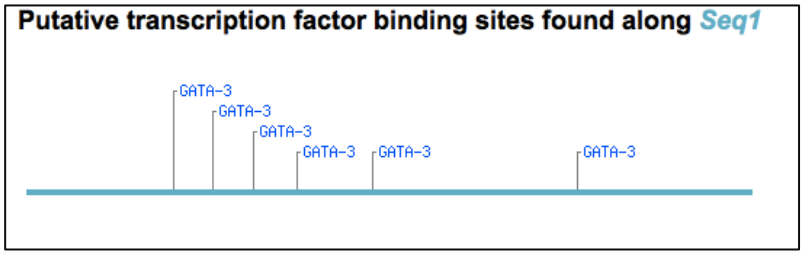

Figure 7*. CONSITE analysis of TFBS GATA 3 along 2Kb region of intron 18, cutoff 90%.*

In order to have a more precise functional insight around the identified SNPs, stretches of sequences of around 15 bp upstream and 15 bp downstream for each SNP and for both alleles were analyzed for TFBS prediction and the most striking results seem to be around SNP 67538032 (first) and SNP 67397924 (third) where presence of T alleles on these SNPs predict high scorings for TF GATA 3 against lower affinities for other alleles (Table 6). Besides, in all screened databases T alleles for both first and third SNPs seem to show affinities for higher number of many other different TFs than G or C alleles, even at higher cutoffs (> 85%).

SNP	allele	score	sequence

67538032	G (protective)	5.421	AAATGGAGTGTGTGTATCTTCTAAAT
	T (risk)	7.189	AAATGGAGTGTGTTTATCTTCTAAAT
67537924	C (protective)	0	TAGATCGTGATTAGCGACTACTTACAGAGTC
	A (risk)	0	TAGATCGTGATTAGAGACTACTTACAGAGTC
67537493	T (risk)	8.092	GGAGTACTATCATTACTTGG
	C (protective)	3.693	GGAGTACTACCATTACTTGG
67537177	A (protective)	0	GGGCTCATTCAGC
	G (risk)	0	GGGCTCGTTCAGC
67536944	C (protective)	0	GTCCTCCGGGGTGAGA
	T (risk)	0	GTCCTCTGGGGTGAGA
67536642	A (protective)	0	CCCTCCATGCGGCCAGC
	C (risk)	0	CCCTCCCTGCGGCCAGC

Table 6. CONCITE (http://asp.ii.uib.no:8090/cgi-bin/CONSITE, accessed Feb. 2013) analysis for TFBS GATA 3 around all SNPs. (red highlights the 2 top associated SNPs in this region)

According to TESS the position of T allele in SNP 67538032 determines 100% consensus for transcription factor FOXI 1 on that position, meaning that this TF is not predicted in other allele. Regarding top most associated SNP 67537177, G allele seems to have a 100% consensus for that nucleotide for both p53 and RAR. Since transcription factors rarely act alone in regulating gene expression and, many times, multiple factors may bind the DNA in close proximity to each other forming regulatory modules integrating multiple signals. The identification of GATA modules and higher-order regulatory structures around first and third SNPS may eventually be an important finding during this computational analysis of regulatory elements in intron 18 of *PTPN3*. Intronic Splicing Sites were also analyzed and for all SNP variants in intron 18 and results compared using different bioinformatics platforms. Human Splicing Finder software detected an introduction of 2 new silencer motifs in SNP 67538032 where there is a mutant sequence substitution at G>T. This site is however outside of the known exon/intron junction.

As for molecular analysis of Intron 18, and due to time restrictions the experiment couldn't be completed and we could only clone the two constructs from the risk haplotype. Protective haplotype had very low concentration yields of final purified and digested intron 18 amplicons (± 3 ηg/μl), not allowing to proceed with sufficient amounts for the pGL3 ligation step.

5.2 Bioinformatics Analysis of PTPN3 -3'UTR

The 3'-UTR is very important in gene expression by influencing the localization, stability, export, and translation efficiency of the mRNA. It has a number of sequences associated with gene expression, like microRNA response elements, AU-rich elements (ARE), IRON-responsive elements, cytoplasmic polyadenylation elements and the poly(A) tail. 3'UTR structural characteristics as well as its use of alternative polyadenylation strengthens its role in gene expression (Barrett *et al.*, 2012). Mutations in this region can have strong consequences because one change can be secondarily responsible for the altered expression of different genes. In other words, during transcription a mutation may affect not only the allele and genes that are physically linked but also binding proteins that are related to the processing and nuclear export of mRNA affecting other unrelated genes (Chatterjee *et al.*, 2009).

3’UTR in the canine *PTPN3* is 3452 bp length and 82% sequence similarity compared to human having many regions highly conserved between human/ mouse/ rat. Although none of the studied SNPs are included in these regions, they are flanking the polymorphic sites very closely (http://genome.ucsc.edu/index.html, accessed 22 May 2013). SNP rs22128547 (top most associated in 3’UTR region) is located in a region of constrained elements between 29 eutherian mammals predicting the existence of important regulatory elements around its location.

For the bioinformatics analysis different databases were used to determine different UTR motifs like potential predicted microRNA targets, ARE and poly(A) elements, to determine *in silico* if polymorphisms might have any causal effect. However, no evidence of direct influence of any of the studied SNPs with direct overlapping of any of the elements was found. For prediction of microRNA targets, TargetScanHuman database was used and retrieved information shows that targets are closely located both upstream and downstream of the SNPs but never overlapping them (Fig. 8). The target site with highest conservation score is the cfa-mir-145 and it happens to be simultaneously the closest (90bp downstream to the SNP rs22128547) (http://www.targetscan.org, accessed online 20 May 2013).

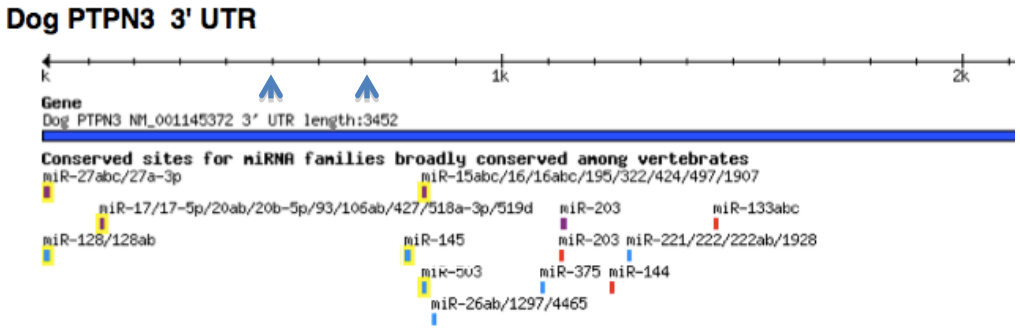

Figure 8*. Conserved sites for miRNA families broadly conserved among vertebrates relative to Dog PTPN3 3’UTR. Blue arrows indicate SNP positions (http://www.targetscan.org, accessed March 2013).*

Thermodynamic prediction of mRNA folding in 3’UTR region containing the two different alleles A and G for SNP rs22128547 was done using RNAfold webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi, accessed June 2013) in order to estimate possible secondary structure changes between the two variants. Results showed that there is an evident change in the stem loop where polymorphic site is located (Fig. 9).

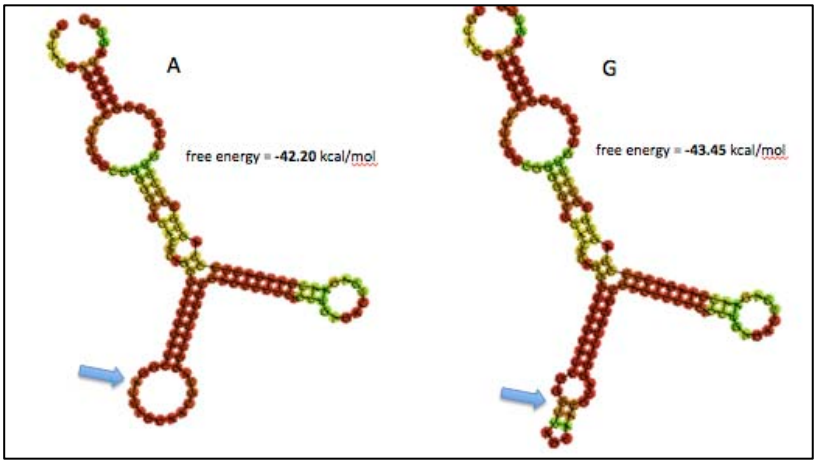

Figure 9*. Thermodynamic prediction of secondary folding of mRNA for the polymorphic site A/G in SNP rs22128547 from 3'UTR of PTPN3 gene (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi, accessed online 3 June, 2013).*

This change in mRNA secondary folding may potentially represent a dramatic change in de-regulation during translation or other important events during gene expression. RNA secondary structures have a preponderant influence in the function of different types of RNAs, like mRNAs and microRNAs in almost every step in gene expression (Seemann *et al*., 2008).

5.3 Results from Functional Molecular Analysis of *PTPN3* – 3'UTR

5.3.1 Optimization and Transfection for MDCK cells

Prior to transfection the plated cells were examined under the inverted microscope to ensure cells were 80% confluent. Optimization assay for MDCK cells with pGL3+ Renilla (pRL-TK) was preformed prior to final experiment in order to establish the optimal amount of DNA/ lipofectamine 2000 to be used in relation to two different concentrations Renilla pRL-TK, 50 ηg and 100 ηg A ratio of 1/2.5 of pGL3/Lipofectamine 2000 was kept all times. Protocol was developed according to Table 4.

Lipofectamine 2000	2.5 µl		2 µl		1.5 µl		0.75 µl	
pGL3 +pRL-TK (total)	1 µg		0.8 µg		0.6 µg		0.3 µg	
PGL3 + Renilla pRL-TK	0.9 µg /100ηg	0.95 µg /50ηg	0.7 µg /100ηg	0.75 µg /50ηg	0.5µg /100ηg	0.55 µg /50ηg	0.2 µg /100ηg	0.25 µg /50ηg

Table 4. Transfection optimization assay for Dual luciferase / Renilla expression with 1/25 ratio Lipofectamine 2000 in MDCK cells.

Fresh dilutions of plasmids were made as required before each transfection experiment (starting up with initial concentration of 671 ηg/ µl for pGL3 and 406 ηg/ µl for Renilla pRL-TK expressing vectors). Separate master mixes were made for each plasmid being transfected, and different amounts of Lipofectamine 2000/Opti-Mem mix reagent per reaction were made and incubated at room temperature for 5 min in the filtered hood.
Specified amounts of Lipofectamine 2000 /Opti-Mem mix was then added to each of the reporter/Opti-Mem mixes and incubated at room temperature in the filtered hood for another 20 min. Transfection mix was added in duplicates in one 24 well plate. Cells were

then incubated for 48h. For final transfection of MDCK cells we have used 400 ηg of total DNA plus 50 ηg of pRL-TK (1 μl). In order to achieve the same amounts of DNA, different volumes were used. Three independent DNA preparations in order to increase repeatability of results since many factors may interfere with DNA purity. 50 μl of Opti-Mem were added to 400 ηg of DNA and 50 ηg of pRL-TK in 7 separate and labeled eppendorf tubes and 50 μl of Opti-Mem plus 1,125 μl of Lipofectamine 2000 (ratio 1:2.5) were left incubated at room temperature for 5 min. After this time, both DNA+pRL-TK and Lipofectamine were combined and left again incubate for another 20 min at room temperature. DNA/pRL-TK/Lipofectamine 2000 combination was added to each of 24 well plates loaded in triplicates for each DNA sample plus controls with pGL3. Cells were left for transfection in incubator for 48h at 37°C and 5% CO_2. After this period cells were lysed and collected from the 24 well plates scratching the bottom with pipette tips and transferred to properly labeled eppendorf tubes. Cells were centrifuged for 2 min at 3000 rpm. Supernatant was discarded and cells were resuspended in PBS (700 μl) and centrifuged again for 2 min at 3000 rpm. Supernatant was discarded by suction pump and 90 μl of lysis buffer added to each tube followed by low speed vortexing and left 15 min at room temperature. Finally the lysed cells were spun for 30 sec and supernatant collected for Dual-Luciferase assay.

5.3.2 Pilot assay for Transfection of Jurkat cells

In order to assess transfection efficiency, two different batches “A” and “B” of Jurkat T-cells were cultured in RPMI + 10% FBS without antibiotics and divided in two subgroups, one stimulated with Phorbol 12-Myristate 13-Acetate (PMA) and Ionomycin (IO) and another one non-stimulated. PMA and IO are both immunogenic activators allowing Calcium flux leading to several downstream effects, such as up-regulation of CD7 in T cells (signal of activation) or the hydrolysis of phosphoinositides and activation of protein kinase C in T-cells. Activation of PMA+IO was performed 10 hours prior to collection and 24 hour after transfection. Stimulation was done with 2 μl of PMA at final concentration of 20 ηg/ml and IO with 0.5 μl at 0.5 mM. Transfection of pGL3 promoter and pRL-TK was done in two different amounts (300ηg/100ηg and 600ηg/200ηg) for both subgroups and 12 wells (600.000 cells/well) of a 24 well plate was used according to Table 5.

pGL3/pRL-TK	Non Stimulated (NS)	Stimulated (PMA+IO)		
300ηg/100ηg	1	2	9 (NS Cells)	Batch A
600ηg/200ηg	3	4	10 (Stimulated Cells)	
300ηg/100ηg	5	6	11 (NS Cells)	Batch B
600ηg/200ηg	7	8	12 (Stimulated Cells)	

Table 5. Distribution scheme of 2 different batches of Jurjat T-cells in 12 wells both stimulated and non-stimulated with PMA+IO and with different pGL3/pRL-TK amounts.

5.3.3 DUAL Luciferase Assay Analysis for 3’UTR in MDCK cells

Data obtained from reporter analysis from the two independent experiments was analyzed and compared between the 3 different 3’UTR variants of *PTPN3* in MDCK cells.

Unpaired *t* test results between protective G/A and C/A and risk G/G haplotypes:

P value and statistical significance G/A versus G/G:
The two-tailed P value between protective haplotype G/A *versus* Risk haplotype G/G equals 0.8121. By conventional criteria, this difference is considered not to be statistically significant.

P value and statistical significance C/A versus G/G:
The two-tailed P value between Protective haplotype C/T/A *versus* Risk haplotype G/T/G equals 0.5945. By conventional criteria, this difference is considered not to be statistically significant.
Luciferase analysis did not reveal a trend in Luciferase expression regarding different insert haplotypes. There are no consistent results that may give indication that protective and risk 3'UTR variants play a distinct role in PTPN3 gene regulation or expression (Fig. 10).

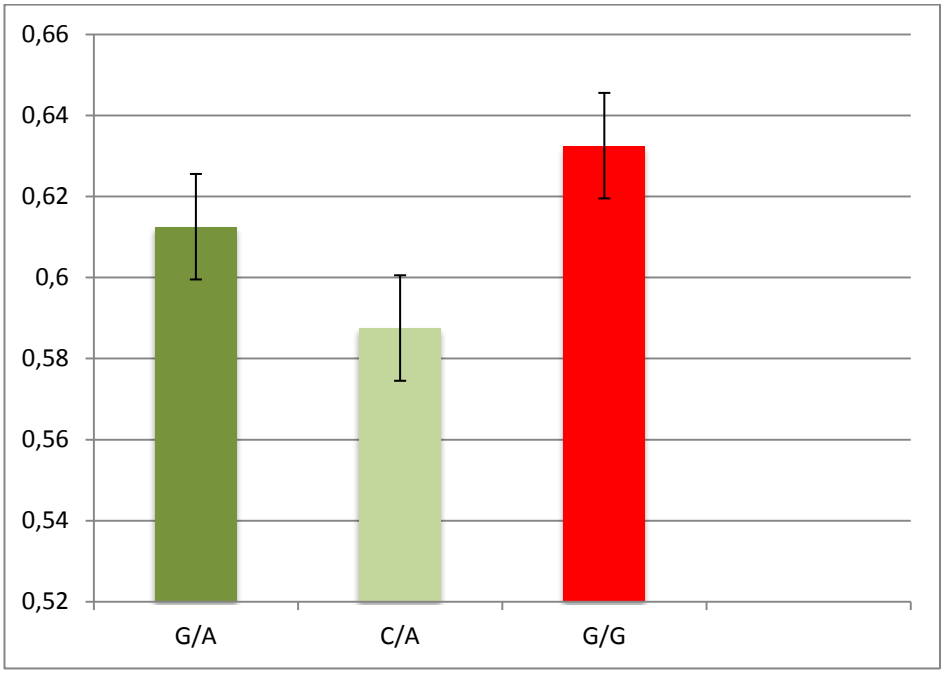

Figure 10. *Reporter analysis of the 3'- untranslated region (3'UTR) of PTPN3. The entire 1.6 kb region was cloned and inserted after the firefly luciferase in a pGL3 vector. Luciferase activity was examined. Bars show the mean results from 2 independent experiments. The firefly luciferase containing G/A construct has P=0.81 versus G/G, and C/A has P=0.59 versus G/G.*

5.3.4 DUAL Luciferase Assay Analysis for 3'UTR in Jurkat T-cells

Data obtained from reporter analysis from the two independent experiments was analyzed and compared between the 3 different 3'UTR variants of *PTPN3*.

Unpaired t test results between Non Stimulated Protective C/A and Risk G/G 3'UTR
P value and statistical significance:
The two-tailed P value equals 0.372. By conventional criteria, this difference is not considered to be statistically significant.
Unpaired t test results between Non Stimulated Protective G/A and Risk G/G 3'UTR

P value and statistical significance:

The two-tailed P value equals 0.5472. By conventional criteria, this difference is not considered to be statistically significant.

Unpaired t test results between Stimulated Protective G/A and Risk G/G 3'UTR

P value and statistical significance:

The two-tailed P value equals 0.7663. By conventional criteria, this difference is not considered to be statistically significant.

Unpaired t test results between Stimulated Protective C/A and Risk G/G 3'UTR

P value and statistical significance:

The two-tailed P value equals 0.8653. By conventional criteria, this difference is not considered to be statistically significant.

Once again, Luciferase analysis did not reveal a trend in Luciferase expression regarding different insert haplotypes in Jurkat cells. There are no consistent results that may give indication that protective and risk 3'UTR variants play a distinct role in PTPN3 gene regulation or expression (Fig. 11).

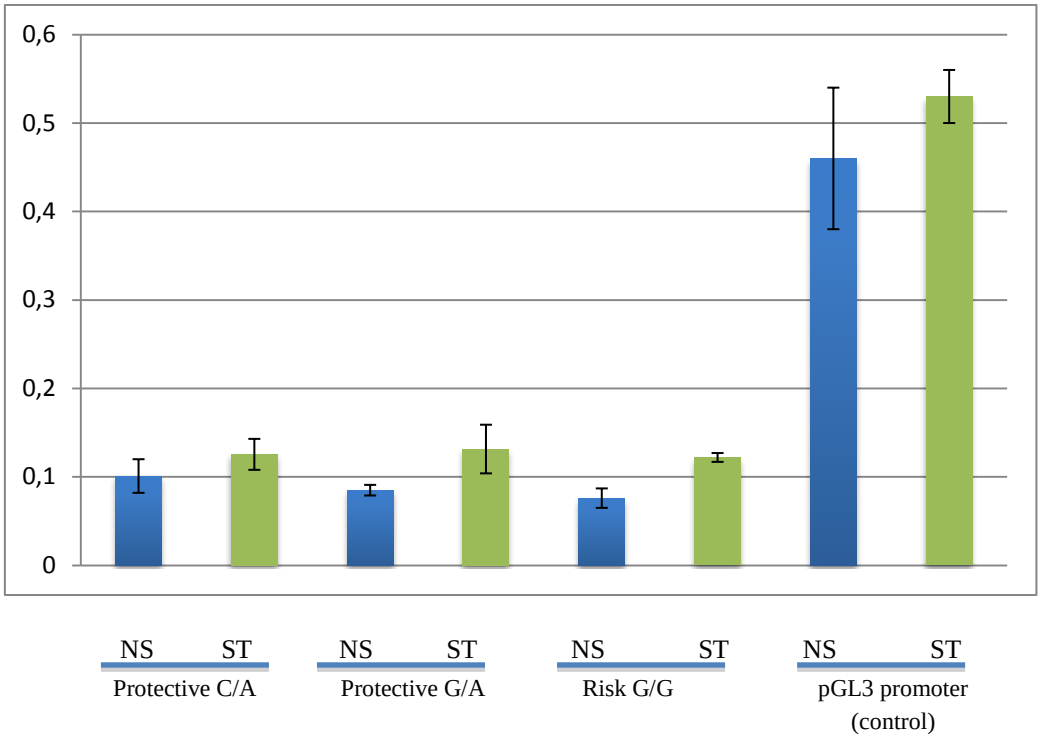

Figure 11. *Reporter analysis of the 3'UTR of PTPN3 for Jurkat T-cells stimulated with PMA+IO (ST) and non-stimulated (NS). The entire 1.6 kb region was cloned and inserted after the firefly luciferase in a pGL3 vector. Luciferase activity was examined. Bars show the mean results from 2 independent experiments in a total of 6 replicates. In experiments with the firefly luciferase in ST cells, G/A construct has P=0.372 versus G/G, and C/A has P=0.54 versus G/G. In ST cells we verified that protective G/A construct has P=0.86 versus risk G/G and protective C/A has P=0.76 versus risk G/G. pGL3 promoter was used without 3'UTR insert as a control where higher levels of protein expression with P=0.0086 versus protective G/A construct. SEM error bars are indicated.*

6 CONCLUSIONS AND FUTURE DIRECTIONS

The present investigation aimed to validate the association of PTPN3 -3'UTR and intron 18 regions containing some of the top highly associated SNPs to lower expression levels of this protein in IMRD in NSDTRs.

Luciferase expression levels in different haplotypes of 3'UTR did not reveal any trend and *p* values are not statistically significant between any of the protective and risk variants in the transfected MDCK and Jurkat T cells. As expected and as it is observed in most cells and tissues, PTPN3 protein is lowly expressed in transfected cells and that may be due to the presence of two ARE elements in 3'UTR that rapidly degrade mRNA.

To corroborate this hypothesis, RNA degradation analysis is still ongoing and might provide additional information about the dynamics of mRNA synthesis and degradation during several time-points during a 12 h time period. Since bioinformatics analysis revealed the presence of SNP rs22128547 within a highly constrained region it supports the fact that conformational change in thermodynamic secondary folding of mRNA for the polymorphic site (transition that exchanges a purine for purine (A ↔ G)) may potentially represent a dramatic change in de-regulation during important events in gene expression. The development of this research may provide insights into the complexity of PTPN3 expression. Additional experiments should be designed to clarify clearly if this region may play an important role in diseased haplotypes with for example further cloning of 3UTR inserts without one or both ARE elements that may potentially be overshadowing luciferase expression as we can clearly see in figure 11 by comparing protective and risk haplotypes with the control promoter pGL3.

To define the functional role of Intron 18 polymorphisms, ongoing experiments as outlined above must be proceeded and data analysed comparing both protective and risk haplotypes in luciferase expression. However, *in silico* analysis revealed some important clues about a specific highly conserved region comprising around 1 Kb, in which the identification of multiple GATA modules and higher-order regulatory structures around first and third SNPs may eventually be an important finding during computational analysis of regulatory elements in intron 18. Several diseases including autoimmune, have been already associated with intronic polymorphisms like autoimmune thyroid disease (Liu *et al*. 2012), SLE, and rheumatoid arthritis (Alfadhli, 2013)
Further development of research on intron 18 should be continued in order to uncover potential influence of this region in disease haplotypes over *PTPN3* expression.

BIBLIOGRAPHIC REFERENCES

Alonso A, Sasin J, Bottini N, Friedberg I, Osterman A, Godzik A, Hunter T, Dixon J, Mustelin T (2004) Protein tyrosine phosphatases in the human genome. Cell 117: 699-711

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

Andersen JN, Jansen PG, Echwald SM, Mortensen OH, Fukada T, Del Vecchio R, Tonks NK, Moller NP (2004) A genomic perspective on protein tyrosine phosphatases: gene structure, pseudogenes, and genetic disease linkage. FASEB J 18: 8-30

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

Atassi MZ, Casali P (2008) Molecular mechanisms of autoimmunity. Autoimmunity 41: 123-132

Barrett, Lucy W.; Fletcher, Sue; Wilton, Steve D. (27 April 2012). "Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements". *Cellular and Molecular Life Sciences* 69 (21): 3613–3634

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

Chabanne L, Fournel C, Monier JC, Rigal D: Canine systemic lupus erythematosus. Part II. Diagnosis and treatment. *Compen Contin Educ Pract Vet* 1999, 21:402-421

Chatterjee, Sangeeta; Pal, Jayanta K. (1 May 2009). "Role of 5'- and 3'-untranslated regions of mRNAs in human diseases". *Biology of the Cell* 101 (5): 251–262

Chishti AH, Kim AC, Marfatia SM, Lutchman M, Hanspal M, Jindal H, Liu SC, Low PS, Rouleau GA, Mohandas N, Chasis JA, Conboy JG, Gascard P, Takakuwa Y, Huang SC, Benz EJ, Bretscher A, Fehon RG, Gusella JF, Ramesh V, Solomon F, Marchesi VT, Tsukita S, Tsukita S, Hoover KB (August 1998). "The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane". *Trends Biochem. Sci.* 23 (8): 281–2

Cline, J., Braman, J.C. and Hogrefe, H.H. (1996) PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. Nucl. Acids Res. 24, 3546–51

Janeway C.A., Travers P., Walport M., Shlomchik M. ImmunobiologyGarland Science Publishing, 2005

Ermann, J., Fathman C. G. Autoimmune diseases: genes, bugs and failed regulation. Nature Immunology 2001; vol2:9, 759-761

Gjorloff-Wingren A, Saxena M, Han S, Wang X, Alonso A, Renedo M, Oh P, Williams S, Schnitzer J, Mustelin T (2000) Subcellular localization of intracellular protein tyrosine phosphatases in T cells. Eur J Immunol 30: 2412- 2421

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

Iwasaki H, Akashi K (2007) Myeloid Lineage Commitment from the Hematopoietic Stem Cell. Immunity 26: 726-740

James, William; Berger, Timothy, Elston & Dirk, Andrews' Diseases of the Skin: Clinical Dermatology. 10th Ed; 2005:1100-1107

Jokinen, P (2011) Identifying Genetic Risk Factors in Canine Autoimmune Disorders, Doctoral Thesis University of Helsinki

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

Kindt, T. J., Goldsby, R. A. and Osborne, B. A. (2007) Kuby Immunology (5th Ed), W. H. Freeman, CA, USA

Knoche, K. and Kephart, D. (1999) Cloning blunt-end Pfu DNA polymerase- generated PCR fragments into pGEM-T Vector Systems. Promega Notes 71, 10–13

Martens HA, Nolte IM, van der Steege G, *et al.* (March 2009). "An extensive screen of the HLA region reveals an independent association of HLA class I and class II with susceptibility for systemic lupus erythematosus". *Scand. J. Rheumatol.* 38 (4): 1–7. doi: 10.1080/03009740802552469

Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, Sloan CA, Rosenbloom KR, Roe G, Rhead B, Raney BJ, Pohl A, Malladi VS, Li CH, Lee BT, Learned K, Kirkup V, Hsu F, Heitner S, Harte RA, Haeussler M, Guruvadoo L, Goldman M, Giardine BM, Fujita PA, Dreszer TR, Diekhans M, Cline MS, Clawson H, Barber GP, Haussler D, and Kent WJ. The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res*. 2012 Nov 15.

Morel, L, Tian, Croker, B. P.& Weakland, E. K. Immunity 11, 131-139 (1999)

Mustelin T, Vang T, Bottini N (2005) Protein tyrosine phosphatases and the immune response. Nat Rev Immunol 5: 43-57

Reprogen, Faculty of Veterinary Science, University of Sydney. Online mendelian inheritance in animals, OMIA. 2013

Rahman Anisur and Isenberg A.David, Review Article: Systemic Lupus Erythematosus. North England Journal of Medicine 358 (9); 2008: 929–939

Seemann, S. E., Gorodkin, J. & Backofen, R. Unifying evolutionary and thermodynamic information for RNA folding of multiple alignments. Nucleic Acids Res. 36, 6355–6362 (2008

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

Strang A. and MacMillan G. The nova scotia duck tolling retriever. Alpine Publications, 1996

't Hoen PA,Hirsch M,de Meijer EJ,de Menezes RX,van Ommen GJ,den Dunnen JT.(2010). mRNA degradation controls differentiation state-dependent differences in transcript and splice variant abundanceNucl. Acids Res. doi:10.1093

Tan EM, Cohen AS, Fries AT & Mcshane DJ et.al , The 1982 revised criteria for the Classification of Systemic Lupus Erythematosus. Arthritis. Rheum.25; 1982: 1271-7

Tipold, A., A. Jaggy (1994): Steroid responsive meningitis-arteritis in dogs: long term study of 32 cases. J. Small Anim. Pract. 35, 311-316

Tipold, A., M. Vandevelde, A. Zurbriggen (1995): Neuroimmunological studies in steroid-responsive meningitis-arteritis in dogs. Res. Vet. Sci. 58, 103-108

Tipold A, and Schatzberg S. J. (2010), An update on steroid responsive meningitis-arteritis. Journal of Small Animal Practice, 51: 150–154. doi: 10.1111/j.1748-5827.2009.00848

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 com- plex. Immunogenetics 61, 557–564

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

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

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

Wong ML, Medrano JF (July 2005). "Real-time PCR for mRNA quantitation". *BioTechniques* 39 (1): 75–85

Wrzosev M, Konar M, Vandevelde M. & Oevermann, A. (2009) Cerebral extension of steroid-responsive meningitis arteritis in a boxer. *Journal Small Animal Practice* 50, 35–37

WEB resources:

http://bioinformatics.org/sms/

http://asp.ii.uib.no:8090/cgi-bin/CONSITE

http://learn.hamamatsu.com/

http://genome.ucsc.edu

http://itbtools.ba.itb.cnr.it/utrscan

http://www.targetscan.org/

http://products.invitrogen.com/

http://www.ncbi.nlm.nih.gov/refseq/

http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi

http://www.promega.com

http://omia.angis.org.au