UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1286

Epigenetic and Gene Expression Signatures in Systemic Inflammatory Autoimmune Diseases

JULIANA IMGENBERG-KREUZ

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017

ISSN 1651-6206
ISBN 978-91-554-9782-8
urn:nbn:se:uu:diva-310388

Dissertation presented at Uppsala University to be publicly examined in Enghoffsalen, entrance 50, ground floor, Uppsala University Hospital, Uppsala, Friday, 17 February 2017 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Prof Amr Sawalha (University of Michigan, Department of Internal Medicine).

Abstract
Imgenberg-Kreuz, J. 2017. Epigenetic and Gene Expression Signatures in Systemic Inflammatory Autoimmune Diseases. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1286. 76 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9782-8.

Autoimmune diseases are clinical manifestations of a loss-of-tolerance of the immune system against the body's own substances and healthy tissues. Primary Sjögren's syndrome (pSS) and systemic lupus erythematosus (SLE) are two chronic inflammatory autoimmune diseases characterized by autoantibody production and an activated type I interferon system. Although the precise mechanisms leading to autoimmune processes are not well defined, recent studies suggest that aberrant DNA methylation and gene expression patterns may play a central role in the pathogenesis of these disorders. The aim of this thesis was to investigate DNA methylation and gene expression in pSS and SLE on a genome-wide scale to advance our understanding of how these factors contribute to the diseases and to identify potential biomarkers and novel treatment targets.

In study I, differential DNA methylation was analyzed in multiple tissues from pSS patients and healthy controls. We identified thousands of CpG sites with perturbed methylation; the most prominent finding was a profound hypomethylation at regulatory regions of type I interferon induced genes in pSS. In study II, a cases-case study comparing DNA methylation in pSS patients with high fatigue to patients with low fatigue, we found methylation patterns associated to the degree of fatigue. In study III, RNA-sequencing was applied to investigate the transcriptome of B cells in pSS in comparison to controls. Increased expression of type I and type II interferon regulated genes in pSS was observed, indicating ongoing immune activation in B cells. In study IV, the impact of DNA methylation on disease susceptibility and phenotypic variability in SLE was investigated. We identified DNA methylation patterns associated to disease susceptibility, SLE manifestations and different treatments. In addition, we mapped methylation quantitative trait loci and observed evidence for genetic regulation of DNA methylation in SLE.

In conclusion, the results presented in this thesis provide new insights into the molecular mechanisms underlying autoimmunity in pSS and SLE. The studies confirm the central role of the interferon system in pSS and SLE and further suggest novel genes and mechanisms to be involved in the pathogenesis these autoimmune diseases.

Keywords: Autoimmunity, DNA methylation, Primary Sjögren's syndrome, Systemic lupus erythematosus, Gene expression, Type I interferon system, Epigenetics

Juliana Imgenberg-Kreuz, Department of Medical Sciences, Molecular Medicine, Akademiska sjukhuset, Uppsala University, SE-75185 Uppsala, Sweden.

ISSN 1651-6206
ISBN 978-91-554-9782-8
urn:nbn:se:uu:diva-310388 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-310388)

To Charlotta & Matilde

Supervisors

Assoc. Prof. Gunnel Nordmark
Prof. Ann-Christine Syvänen
Dr. Johanna Sandling

Department of Medical Sciences, Uppsala University, Uppsala, Sweden

Chair

Prof. Hans Törmä

Department of Medical Sciences, Uppsala University, Uppsala, Sweden

Faculty opponent

Prof. Amr Sawalha

Department of Internal Medicine, and Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

Review board

Prof. Tomas Ekström

Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Prof. Anca Catrina

Department of Medicine, Karolinska Institutet, Stockholm, Sweden

Prof. Birgitta Heyman

Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I. **Imgenberg-Kreuz J**[#], Sandling JK[#], Almlöf JC, Nordlund J, Signér L, Norheim KB, Omdal R, Rönnblom L, Eloranta ML, Syvänen AC, and Nordmark G. Genome-wide DNA methylation analysis in multiple tissues in primary Sjögren's syndrome reveals regulatory effects at interferon induced genes. *Annals of the Rheumatic Diseases*, 2016;75(11): 2029-2036.

II. Norheim KB, **Imgenberg-Kreuz J**, Jonsdottir K, Syvänen AC, Sandling JK, Janssen EAM, Nordmark G, and Omdal R. Epigenome-wide DNA methylation patterns associated with fatigue in primary Sjögren's syndrome. *Rheumatology (Oxford)*. 2016;55(6): 1074-82.

III. **Imgenberg-Kreuz J**, Sandling JK, Björk A, Nordlund J, Kvarnström M, Eloranta ML, Rönnblom L, Wahren-Herlenius M, Syvänen AC, and Nordmark G. Transcription profiling of CD19+ B cells in primary Sjögren's syndrome reveals an interferon-signature with up-regulated *BAFF* and *TLR7*. *Submitted manuscript*.

IV. **Imgenberg-Kreuz J**, Almlöf JC, Leonard D, Nordmark G, Eloranta ML, Padyukov L, Gunnarsson I, Svenungsson E, Sjöwall C, Rönnblom L, Syvänen AC, and Sandling JK. DNA methylation mapping identifies gene regulatory effects in patients with systemic lupus erythematosus (SLE). *Manuscript*.

[#] Equally contributing authors

Contents

Abbreviations

5-hmC	5-hydroxymethylcytosine
5-mC	5-methylcytosine
ACR	American College of Rheumatology
AECG	American European Consensus Group
ANA	Antinuclear antibody
β-value	Methylation beta-value
BAFF/ BLyS	B cell activating factor/ B lymphocyte stimulator
Breg	B regulatory cell
C	Complement
ChIP-seq	Chromatin immunoprecipitation sequencing
CNV	Copy number variation
CpG	Cytosine-phosphate-guanine dinucleotide
DHS	DNase hypersensitivity site
DMC	Differentially methylated CpG site
DNMT	DNA methyltransferase
ENCODE	Encyclopedia of DNA elements
ESSDAI	EULAR Sjögren's syndrome disease activity index
EULAR	European League Against Rheumatism
EWAS	Epigenome-wide association study
FC	Fold change
FDR	False discovery rate
FPKM	Fragments per kilobase of exon per million fragments mapped
f-VAS	Fatigue visual analogue scale
GWAS	Genome-wide association study
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HM450k	HumanMethylation450k BeadChip
IC	Immune complex
IFN	Interferon
IFNAR	Interferon-α/β receptor

Ig	Immunoglobulin
IL	Interleukin
IRF	Interferon regulatory factor
MAF	Minor allele frequency
meQTL	Methylation quantitative trait locus
MHC	Major histocompatibility complex
miRNA	Micro-RNA
MTX	Methotrexate
ncRNA	Non-coding RNA
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
PARP9/ BAL1	Poly(ADP-ribose) polymerase family member 9/ B-aggressive lymphoma-1
PBMC	Peripheral blood mononuclear cell
pDC	Plasmacytoid dendritic cell
pSS	Primary Sjögren's syndrome
QC	Quality control
RA	Rheumatoid arthritis
RNA-seq	RNA-sequencing
SLE	Systemic lupus erythematosus
SLICC	Systemic Lupus International Collaborating Clinics
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SOCS	Suppressor of cytokine signaling
STAT	Signal transducer and activator of transcription
SWAN	Subset-quantile within array normalization
TET	Ten-eleven translocation enzyme
TLR	Toll-like receptor
TNF	Tumor necrosis factor
WGBS	Whole-genome bisulfite sequencing
WGS	Whole-genome sequencing

Introduction

The primary function of the immune system is to protect the body against diseases and infections. In this tightly regulated process the body's ability to distinguish between self and non-self is crucial. Autoimmune diseases are manifestations of a loss-of-tolerance of the immune system against the body's own substances or healthy tissues. These diseases comprise a large and clinically heterogeneous group of more than 80 immune disorders, collectively affecting ~5% of the population [1]. The effects of autoimmune diseases can be organ-specific, such as in type I diabetes, which affects the insulin producing beta cells of the pancreas, or systemic, if the inappropriate immune response is directed against ubiquitous molecules, such as in systemic lupus erythematosus (SLE). Sjögren's syndrome can be considered both an organ-specific disease with autoimmunity against exocrine glands, as well as a systemic disease.

Sjögren's syndrome and SLE are two clinically and immunologically related chronic autoimmune inflammatory diseases with complex etiology. Sjögren's syndrome can occur either alone as primary Sjögren's syndrome (pSS) or as a secondary syndrome to SLE, rheumatoid arthritis (RA) or other autoimmune diseases. Some systemic clinical features are shared between pSS and SLE, such as non-erosive arthritis, leucopenia, myalgia, arthralgia, Raynaud's phenomenon and chronic fatigue, while e.g. severe renal involvement is more common in SLE. A course of disease flares and remissions is characteristic for SLE, whereas pSS is more often a stable disease. Autoreactive B cells play a key role in the pathophysiology of both pSS and SLE, and B cell hyperactivity leading to hypergammaglobulinemia and production of autoantibodies is a hallmark of the diseases [2, 3]. Furthermore, both diseases are associated with an increased risk for B cell lymphoma [4]. Activation of the type I interferon (IFN) system with increased plasma levels of IFN-α/β, and transcriptional upregulation of IFN induced genes, designated as *IFN signature*, has been demonstrated in many patients with pSS and SLE [5].

Although the precise etiology of pSS and SLE remains elusive at the molecular level, they are considered to be multifactorial diseases, where underlying genetic predisposition, environmental influences and epigenetic mechanisms are contributing factors to disease development. Epigenetic mechanisms may act as a dynamic link between genome, environment and phenotypic manifestation by their modulating effects on gene expression [6, 7].

Within the framework of this thesis the impact of epigenetic mechanisms and gene expression in pSS and SLE was studied.

Primary Sjögren's syndrome

Primary Sjögren's syndrome (pSS) has a population prevalence of approximately 0.01-0.06% and a distinct female predominance (female to male ratio of about 14:1) [8]. The disease is named after the Swedish ophthalmologist Henrik Sjögren. He presented his thesis "Zur kenntnis der Keratoconjunctivitis Sicca" in 1933 where he described the clinical and histological findings in 19 patients with dry eyes, dry mouth and arthritis [9].

PSS is a complex and chronic autoimmune rheumatic disease characterized by inflammation of the exocrine glands, mainly the salivary and lacrimal glands. Typical clinical consequences are dryness of the eyes (keratoconjuctivitis sicca) and of the mouth (xerostomia), referred to as sicca symptoms. Histopathological examination of the minor salivary glands typically identifies focal lymphocytic inflammation, which in a third of the patients shows germinal center-like formation. The focal inflammatory infiltrates consist mainly of T cells, with variable amounts of B cells and dendritic cells [10]. In addition to these glandular symptoms, pSS may affect multiple other organs resulting in systemic manifestations. Common extraglandular manifestations include arthritis, Raynaud's phenomenon, pulmonary or renal involvement, leucopenia and fatigue [11]. The most severe complication of pSS is the development of lymphoproliferative malignancy, which is seen in ~5% of the patients [12]. Due to the broad spectrum of symptoms, diagnosis and classification are challenging. The time from symptom onset to disease diagnosis may be up to ten years [11]. Typically, the disease affects adult females around 40-55 years of age, but juvenile forms of pSS have been described [13]. In 2002, the American European Consensus Group criteria (AECG) were proposed in an international effort to standardize classification criteria for pSS, and these criteria were used for patient inclusion in this thesis [14] (Table 1). Recently, the 2016 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) consensus classification criteria for Sjögren's syndrome were published [15]. The vast majority of patients fulfilling the AECG criteria will also fulfill these new criteria.

Table 1. *The American-European Consensus Group (AECG) criteria for classification of Sjögren's syndrome.*

Item	Parameter
I.	Ocular symptoms of dry eyes.
II.	Oral symptoms of dry mouth.
III.	Ocular signs defined as a positive result for at least one of the following two tests: Schirmer's I test (≤ 5 mm in 5 minutes) or Rose bengal score (≥ 4 according to van Bijsterveld's scoring system)
IV.	Histopathology: Minor salivary focus score ≥ 1, defined as a number of lymphocytic foci per 4 mm^2 of glandular tissue.
V.	Salivary gland involvement defined by a positive result for at least one of the following diagnostic tests: Unstimulated whole salivary flow (≤ 1.5 ml in 15 minutes) or Parotid sialography or Salivary scintigraphy
VI.	Presence in the serum of antibodies to Ro(SSA) or La(SSB) antigens, or both.

Modified from Vitali *et al*., 2002 [14]

The fulfillment of at least four out of the six items including either IV (a positive biopsy) or VI (positive autoantibodies), alternatively any three of the objective criteria (III, IV, V or VI) is indicative of pSS.

Etiology

The etiology of pSS is complex, where the interaction of genetic, environmental, hormonal and epigenetic factors is thought to contribute to disease development. The contribution of genetic variation on disease susceptibility has been studied in recent years by candidate gene and genome-wide association studies (GWAS) and to date, about ten risk loci for pSS have been identified at genome-wide significance ($p < 5x10^{-8}$) [16, 17]. However, these variants only account for a limited proportion of the heritability and susceptibility for pSS [18]. A role for viral infections in the pathogenesis of pSS has been suggested. Infections by several viruses such as Epstein-Barr virus (EBV), Hepatitis C virus (HCV) and retroviruses could lead to the development of autoimmunity by triggering a prolonged and excessive immune response, or by mechanisms of molecular mimicry between viral antigens and self-antigens [19]. Despite many efforts, no specific virus has been identified so far as being closely connected to the development of pSS. Given the strong overrepresentation of the disease in females and the observation that

clinical manifestation of pSS is usually seen at the age of 40 to 55 years, when both estrogen and androgen levels decrease, the role of hormonal influences on pSS has been discussed but remains inconclusive [20]. Having an extra X-chromosome as in Klinefelter's (47XXY) or triple X (47XXX) has shown to increase the risk for pSS [21, 22]. More recently, observations of alterations in the epigenetic landscape in several autoimmune diseases, including pSS, have suggested epigenetic mechanisms as important contributing factors in the pathogenic processes [23].

B cells and autoantibodies

B cells have multiple roles in the pathophysiology of pSS and are involved in pro-inflammatory as well as immunosuppressive functions of the immune system. Autoreactive B cells leading to hypergammaglobulinemia and autoantibody production are a hallmark of the disease [2]. Reported autoantibody frequencies vary with the methods used for detection and the cohort of patients investigated. Antinuclear antibodies (ANA) are found in sera from approximately 80% of pSS patients, and autoantibodies against the ribonucleoproteins SSA/Ro and SSB/La are present in about 75% and 45%, respectively [24]. SSA is composed of two subunits, Ro52 and Ro60, which refers to their molecular weight, and SSB is made up by only one protein termed La. These proteins are associated with Y RNA molecules, which are small non-coding RNAs (ncRNAs) implicated in DNA replication processes through their interaction with chromatin and replication initiation proteins [25]. During apoptotic processes SSA and SSB, which are usually distributed in the cytoplasm and the nucleus, can translocate to the cell membrane and become exposed to the immune system. Patients positive for SSB in their sera are virtually always positive for SSA as well, which may be an effect of B cell epitope spreading [26]. Antibody-positivity for SSA/-B in pSS is associated with distinct clinical phenotypes, such as earlier onset of disease, higher focus score in minor salivary gland biopsies, increased disease activity and an increased risk for lymphoproliferative disease [27].

Other types of systemic autoantibodies are detected in smaller subsets of patients; anti-centromere antibodies (ACA) in ~6 % [28], anti-ribonucleoprotein (anti-RNP) antibodies in ~5-10% [29], anti-Smith antibodies (anti-Sm) in < 1% [29] and antibodies to interferon-inducible protein-16 (anti-IFI16) in ~29 % of the patients [30]. Organ-specific autoantibodies against muscarinic receptor 3, the acetylcholine receptor responsible for saliva production, have been described. However, the prevalence and significance of these antibodies is controversial [31].

In addition to their role as autoantibody producing plasma cells, B cells are involved in infiltration of the salivary glands, where they, along with T cells and dendritic cells, participate actively in processes of local inflammation and tissue destruction. Total numbers of B cells are usually not altered,

but the distribution of B cell subsets has been described as aberrant in pSS patients, with depletion of CD27+/CD5+ memory B cells in peripheral blood. Whether memory B cells are accumulated in the minor salivary glands is controversial [32, 33]. One important example of a functional B cell subset, which is found to be decreased in many pSS patients, is regulatory B cells (Bregs) [34]. These cells have an important function in maintenance of immune tolerance by constricting excessive inflammatory responses. A central regulatory mechanism of Bregs is mediated by their release of cytokine interleukin (IL)-10, which counteracts pro-inflammatory responses of T effector cells, mainly Th1 and Th17, and instead supports differentiation of regulatory T cells (Tregs) leading to abatement of the immune response [35, 36].

The crucial role of B cells in pSS (and other systemic autoimmune diseases) has led to the development of targeted therapies against B cells. They may either target B cells directly, e.g. rituximab which is a monoclonal antibody against the B-lymphocyte antigen CD20, or indirectly by blocking B cell survival cytokines, e.g. IL-6 (tocilizumab) or B cell activating factor (BAFF)/B Lymphocyte stimulator (BLyS) (belimumab) [37]. However, B cell depletion therapies generally fail to discriminate between depletion of pro-inflammatory B cells and Breg populations. An exception may be given by the newly described programmed death-ligand 1 (PD-L1) expressing Breg subset, which appears to be largely resistant to anti-CD20 B cell depletion [38]. B cell depletion by rituximab has shown to increase serum levels of BAFF and a novel approach in clinical trials is the combination of rituximab and belimumab [39].

Lymphoma

While overall mortality is not elevated in pSS, the mortality due to lymphoproliferative disorders is almost 8-fold increased in pSS compared to the general population [40]. In pSS the risk of developing a non-Hodgkin lymphoma is approximately 16-fold increased, which makes pSS the autoimmune disease with the highest incidence of lymphoproliferative disease [12]. Predictive factors for lymphoma development are swelling of the parotid gland, palpable purpura, lymphadenopathy, germinal center-like formations in the minor salivary glands, increased disease activity as measured by EULAR Sjögren's syndrome disease activity index (ESSDAI) and low serum levels of complement C3 and C4 [12, 41-43]. Most of the non-Hodgkin lymphomas in pSS are of the B cell type and derive from mucosa-associated lymphoid tissues (MALT) of different organs, such as the major salivary glands or the tonsils. In the salivary glands lymphoma may arise at sites of local inflammation (myoepithelial sialoadenitis), where infiltrating B cells eventually undergo malignant transformation by transition from polyclonal activation to monoclonal proliferation [44].

Increased expression of BAFF by plasmacytoid dendritic cells (pDCs), monocytes and macrophages in the lymphocyte infiltrated salivary glands has been demonstrated [45]. BAFF is a key activator for B cell proliferation, differentiation, survival and immunoglobulin switch, and its expression is directly induced by type I IFNs and controlled by IFN regulatory factors (IRFs) [46]. BAFF overexpression has been associated with aberrant differentiation of B cells and may contribute to the increased lymphoma risk in pSS [43]. In addition, genetic variants in the *TNFSF13B* gene encoding BAFF have been linked to the risk of lymphoma [47].

Fatigue

Fatigue is common in chronic diseases and its etiology is poorly understood. From an evolutionary perspective, fatigue may be regarded as a deeply rooted phenomenon related to sickness-behavior that is generally seen in many species [48]. Sickness-behavior is observed, for instance, during viral infections and will usually disappear as the infection is cleared.

In systemic autoimmune diseases, fatigue may become a persisting condition. In pSS, about 65-70% of the patients are affected by fatigue and many of these patients report fatigue to be the most disabling symptom of their disease [49]. Fatigue is described as an overwhelming sense of tiredness, lack of energy, not to be overcome by sleep or rest. Studies in cancer-, post-stroke- and pSS-related fatigue indicate a genetic basis for fatigue [50-52]. Contributing roles of pro-inflammatory cytokines, in particular IL-1β [53], and heat shock proteins (HSPs), such as HSP90, have been postulated in pSS associated fatigue [54]. In addition, oxidative stress, down-regulatory mechanisms of the activated immune system as well as mitochondrial malfunction may be potential contributors to fatigue [55].

Measurement and quantification of fatigue are extremely challenging, which is reflected by the existence of more than 200 instruments for measuring fatigue. Fatigue may be understood as a number of different components, such as mental fatigue and physical fatigue, or as one entity [56], as it is described within this thesis. The subjective experience of fatigue can either be estimated by structured questionnaires, which typically also comprise questions related to possible causes and consequences of fatigue experience, or by a patient-reported numerical fatigue visual analogue scale (f-VAS) score [57, 58].

Treatment

Local and systemic treatment options for pSS do exist. Local symptomatic treatment of sicca symptoms includes artificial tears and salivary substitution or topical application of salivary flow stimulating agents [59]. Systemic symptomatic treatment is available in form of anticholinergic drugs to in-

duce increased salivary and lacrimal flow. The use of these drugs is limited by common side effects and risks. Hydroxychloroquine and chloroquine are used to alleviate arthritis, but do not elicit any effect on sicca symptoms [60]. For severe extraglandular manifestations such as interstitial nephritis, glucocorticoids and cytotoxic drugs like azathioprine or cyclophosphamide may be used [61]. During the past decade, several biologics have been investigated in clinical trials in pSS. Anti-TNF-α treatment has proven unsuccessful, whereas anti-B cell therapy has shown promising results in patients with pSS with systemic disease [37].

Systemic lupus erythematosus

Systemic lupus erythematosus (SLE) is widely regarded as the prototypic systemic inflammatory autoimmune disease as it can affect any system and organ of the body. Although the clinical appearance of SLE is very heterogeneous, the presence of autoantibodies, which are mainly directed against nuclear antigens, is seen in almost all patients with SLE. Formation and deposition of immune complexes (ICs) in tissues can cause inflammation and damage in multiple organs, including skin, muscles and joints, lungs, heart and kidneys, and the central and peripheral nervous system [62]. Further, apoptotic mechanisms are dysregulated in SLE. Increased apoptosis, together with an impaired clearance of the apoptotic material can lead to a prolonged exposure of the immune system to nuclear autoantigens and may trigger the development of autoreactive B cells [63].
Clinical manifestations of SLE range from rather mild skin or joint symptoms to severe and life-threating disease. Classification is based on the ACR criteria, where four out of eleven criteria need to be fulfilled to classify a patient as having SLE (Table 2), and these criteria are used in this thesis [64]. In 2012, new Systemic Lupus International Collaborating Clinics (SLICC) criteria were published showing a higher sensitivity but lower specificity for detecting SLE compared to the 1982 ACR criteria [65]. The 2012 SLICC criteria have not yet acquired general acceptance.

The prevalence of the disease varies considerably depending on geographic origin, ethnicity, age and sex. In the Northern European population the prevalence is estimated to be 20-69/100,000 individuals, with a strong sex bias towards women (female-male ratio of 9:1). The incidence rates range between 1 and 30 per 100,000 individuals per year [66]. Several studies have reported that prevalence and incidence of SLE are in general higher in individuals of non-Caucasian origin, together with an earlier onset of the disease [67]. The role of genetic and environmental factors behind the epidemiological heterogeneity is unclear.

Table 2. *The 1982 revised American College of Rheumatology (ACR) criteria for classification of SLE.*

Criterion	Definition
1	Malar rash: Fixed erythema over the malar eminence (Butterfly rash)
2	Discoid rash: Erythematosus raised patches
3	Photosensitivity: Skin rash caused by unusual reaction to sunlight
4	Oral ulcers: Oral or nasopharyngeal ulceration
5	Arthritis: Non-erosive arthritis involving two or more peripheral joints
6	Serositis: Pleuritis or pericarditis
7	Renal disorder: Persistent proteinuria or cellular casts
8	Neurologic disorder: Seizures or psychosis
9	Hematologic disorder: Hemolytic anemia, leukopenia, lymphopenia, or thrombocytopenia
10	Immunologic disorder: positive LE test, anti-dsDNA antibodies, anti-Sm antibodies or false positive syphilis test
11	Positive ANA: Abnormal titer of antinuclear antibodies (ANA)

Modified from Tan *et al*., 1982 [64]

SLE is a potentially life-threatening disease and survival rates were poor before treatment with corticosteroids became commonly available in the 1960s. Today, the reported 5-year survival rate is ~95%, and the 10-year survival rate is ~91% [68]. While life expectancy has continuously improved for patients with SLE, it still ranges clearly below the numbers of the general population. The increased mortality risk in SLE is mainly due to cardiovascular and renal disease [69, 70].

Similar to pSS, an activation of the type I IFN system with increased concentrations of IFN-α/β in sera and affected tissues and an IFN signature is observed in patients with SLE [5, 7, 71]. Genetic predisposition, epigenetic mechanisms and environmental factors, such as UV-light exposure, infectious diseases and hormonal influences, have been linked to SLE susceptibility [72-75]. The role of genetic and epigenetic mechanisms in the etiology of the disease is discussed further in the respective parts of this thesis.

The interferon system

Interferons (IFNs) are a group of cytokines that were first described close to 60 years ago by Isaacs and Lindenmann and named after their ability to *interfere* with viral replication [76]. Based on receptor usage the human IFN family is classified into three groups: type I (IFN-α, -β, -ω, -κ and -ε), type II (IFN-γ) and type III (IFN-λ1, -λ2, -λ3, -λ4). IFN-α is further divided into 13 subtypes, encoded by 13 highly homologous genes (*IFNA1*, *-2*, *-4*, *-5*, *-6*, *-7*, *-8*, *-10*, *-13*, *-14*, *-16*, *-17* and *-21*). All type I IFNs signal through the

same cell surface receptor, the IFN-α/β receptor (IFNAR). IFNAR is a ubiquitously expressed heterodimeric receptor, composed of IFNAR1 and IFNAR2 [77]. Type II IFN signals via IFN-γ receptor (IFNGR) and type III IFN via IFN-λ receptor (IFNLR).

During early responses to viral infections, several cell types are able to produce small amounts of IFN-α, but plasmacytoid dendritic cells (pDCs) are the major producer with the capability to secrete IFN-α at levels up to 100-fold above e.g. monocytes [78]. IFN-α production is induced by signaling via Toll-like receptor 7 (TLR7), TLR8 and TLR9, which are expressed in the endosomal compartment of pDCs and also B cells. TLRs are a class of innate immune receptors which detect specific molecular patterns of different pathogens in order to activate immune responses. TLR7 and TLR8 are activated by guanosine- and uridine-rich single stranded RNA (ssRNA), which is found e.g. in the influenza virus [79]. TLR9 recognizes single stranded DNA (ssDNA) molecules containing unmethylated CpGs, e.g. originating from DNA viruses or intracellular bacteria [80]. Binding of ssRNA and ssDNA to their respective TLR elicits myeloid differentiation factor 88 (MyD88)-dependent signaling cascades leading to phosphorylation and translocation of interferon regulatory factor 5 (IRF5) and IRF7 into the nucleus, where these transcription factors induce the expression of IFN-α/β.

The antiviral effects of IFNs include direct inhibition of viral replication and induction of apoptosis of virus-infected cells. In addition, IFNs have widespread indirect properties by mediating a general activation of the immune system. This activation is accomplished by signaling cascades through the Janus protein tyrosine kinases – signal transducer and activator of transcription (JAK-STAT) pathway leading to transcriptional induction of antiviral proteins and co-stimulatory molecules (Figure 1).

Type I IFNs can guide the recruitment of immune effector cells to the site of infection or inflammation. They stimulate the differentiation of monocytes into antigen-presenting cells, and enhance antigen presentation and expression of co-stimulatory molecules by dendritic cells, which in turn leads to the activation of T cells. Further immune-stimulatory effects of IFNs include induction of differentiation and antibody production in B cells, and in T cells suppression of differentiation into T regulatory (Tregs) cells and instead activation of cytotoxic T cells [81]. This broad activation of both innate and adaptive immunity is important in order to clear infections. However, a tight regulation and termination of the IFN response is crucial to avoid excessive stimulation of the immune system. Suppressor of cytokine signaling (SOCS) and protein inhibitor of activated STAT (PIAS) are examples of negative regulators of the JAK-STAT pathway [82].

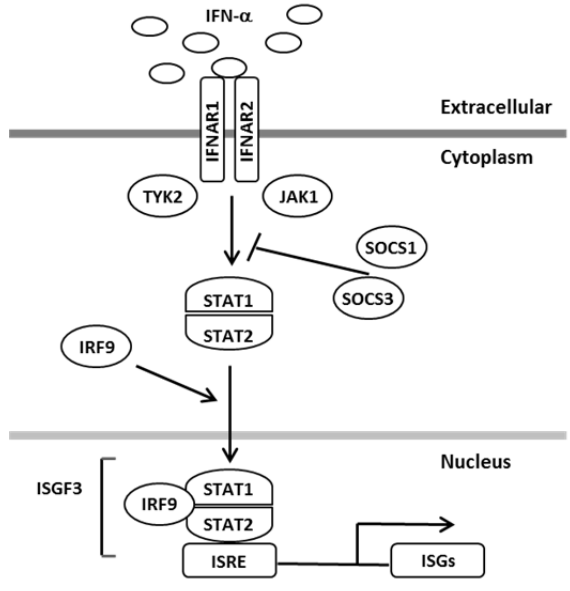

Figure 1. Schematic illustration of the activation of the antiviral response by type I IFN signaling via JAK-STAT. Upon binding of IFN-α to the interferon receptor (IFNAR, which is composed of the subunits IFNAR1 and IFNAR2) the intracellular parts of IFNAR associate with Janus kinase 1 (JAK1) and tyrosine kinase 2 (TYK2) leading to phosphorylation of the receptor, which in turn recruits signal transducer and activator of transcription protein 1 (STAT1) and STAT2. When the STAT1-STAT2 heterodimer becomes phosphorylated it translocates into the nucleus, where it builds a transcription factor complex (interferon stimulated gene factor 3 (ISGF3) complex) with interferon regulatory factor 9 (IRF9). ISGF3 binds to DNA sequences of IFN-stimulated response elements (ISRE) in order to induce transcription of antiviral genes such as *MX1* and *OAS1*, designated as IFN signature. The JAK-STAT pathway is negatively regulated by a feedback mechanism via suppressor of cytokine signaling 1 (SOCS1) and SOCS3.

Interferons in pSS and SLE

The type I IFN system is thought to play a major role in the pathogenesis of inflammatory autoimmune diseases, and a continuous activation has been demonstrated in patients with pSS and SLE. Early evidence for this hypothesis was given when increased levels of IFN in sera from patients with autoimmune diseases were detected more than 35 years ago [83]. A direct causative connection between IFN-α and autoimmunity became evident from the observation that IFN-α therapy, which is used in the clinic to treat certain viral infections and malignancies, gave rise to an SLE-like syndrome in some of the patients [84].

In pSS, elevated serum levels of IFN-α are detected in a proportion of patients and the presence of both IFN-α and pDCs has been demonstrated in minor salivary glands of pSS patients [5, 85]. IFN-α levels are also increased in SLE, and correlate with disease activity [86]. The majority of patients

with SLE have an upregulated expression of type I IFN induced genes in peripheral blood cells, denoted as IFN signature [71]. In patients with pSS, approximately 50% display an IFN signature, mainly seen in anti-SSA/-B antibody positive patients [87]. In addition to the upregulation of IFN induced genes, decreased expression of the negative type I IFN regulator *SOCS1* has been observed in peripheral blood mononuclear cells (PBMCs) from SLE patients and reduced expression of *SOCS3* has been demonstrated in minor salivary glands from patients with pSS [85, 88].

Endogenous IFN inducers appear to play a central role for the activation of the type I IFN system in pSS and SLE: Apart from motifs of infectious agents, also endogenous molecules can act as potent activators of TLR signaling [89, 90]. Immune complexes (IC) formed by nucleic acid containing material which is released from human cells during apoptotic or necrotic processes can appear as endogenous IFN inducers [91]. Upon binding to low-affinity Fc gamma receptor IIa (FcγRIIa) interferogenic ICs are endocytosed and their nucleic acid containing part activates TLR signaling leading to production of IFN-α [92]. Hydroxychloroquine and chloroquine, which were originally developed as antimalarial drugs, are frequently used to dampen the activation of endosomal TLRs in patients with autoimmune diseases [93]. The type I IFN system is an attractive therapeutic target in pSS and SLE and clinical trials are currently underway to evaluate treatment with monoclonal antibodies directly targeting IFN-α molecules (sifalimumab) and against IFNAR1 (anifrolumab) [94, 95]. Recent studies suggest that also the type II and III IFNs may contribute to a greater extent to autoimmunity than previously assumed [96-98].

Several genetic risk variants have been identified at genes connected to key functions in the IFN system (Table 3) and epigenetic studies in fractionated blood cells have revealed decreased DNA methylation levels at type I IFN induced genes in SLE and pSS [99, 100].

Genetics of autoimmune diseases

The human genome and genetic variation

The first drafts of the completed sequence of the ~3 billion base pairs that comprise the human genome were presented by the Human Genome Project and Celera Genomics in 2001 [101, 102]. According to recent estimations, a typical genome differs from the reference sequence at 4.1 to 5 million bases [103]. Any two unrelated individual human beings have the vast majority of their DNA sequence in common and it is the variation in the remaining part that makes us genetically unique and that contributes for example to differences in appearance and in predisposition to disease. Together with the rapid technological development of high-throughput methods, the reference se-

quence of the human genome has provided the basis for the increasing understanding of the impact of genetic variation in health and disease. Shortly after the presentation of the first human genome sequence, the International HapMap Project was initiated with the aim to catalog the common variation (single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) > 5% in a population) of the human genome in individuals with ancestry from four different populations (Northern and Western European, Yoruban, Japanese and Han-Chinese). In phase I and II of the HapMap Project SNP genotyping using microarray technology was performed and ~3.4 million SNPs in the 269 genotyped individuals of different ancestry were reported [104, 105]. Another large-scale effort was undertaken in 2008, including phase III of the HapMap Project and the 1000 Genomes Project; again with the goal to characterize the human genetic variation, but this time including a broader spectrum of ancestral populations and using next-generation sequencing methods. More than 2,000 genomes have been sequenced and have provided a rich source of information about allele frequencies and haplotype structure for ~15 million (SNPs), ~1 million short insertion/deletions and 20,000 structural variants [103].

By number the most common form of genetic variation are SNPs, which arise from germline mutations of a single base in the genetic code. A new variant is considered a SNP when it has reached a frequency of 1% or higher in the population, otherwise it is referred to as a single nucleotide variant (SNV). The vast majority of SNPs are located outside of protein-coding regions, which may be explained by strong negative selection on most of the new variants occurring in coding regions. SNPs located in intergenic regions may have a functional impact by influencing transcriptional and post-transcriptional gene regulation. The frequency of SNVs in the human genome is not yet well defined. While earlier studies estimated that roughly 1:1000 bases in an individual's genome differ from the human reference genome [106], more recent studies suggest the occurrence of a SNP at every 700 to 850 nucleotides, which would sum up to approximately 4 million SNVs in each individual human genome [103].

Genome variation and disease

Three major approaches are used for identifying genes underlying susceptibility to disease; linkage studies, candidate gene association studies and genome-wide association studies (GWAS), and different sequencing-based studies.

Linkage studies typically rely on genotyping of genetic markers across families with more than one affected member spanning several generations in order to identify markers that co-segregate with the disease. The theory behind this strategy is that a co-segregating marker is likely to be found in close proximity to the disease susceptibility locus. Linkage studies have been

particular successful in identifying causal genes for monogenetic disorders, while they have been less informative for complex diseases.

Association studies are based on the comparison of the allele frequencies of a polymorphism (usually a SNP) between two groups, typically cases and controls. Association studies can either be performed within a specific candidate gene or region, or in a more unbiased manner on a genome-wide scale. Since the first GWAS was published in 2005, which studied the genetic basis of macular degeneration [107], the research field has grown rapidly. The GWAS catalogue, a curated database for SNP-trait associations (http://ebi.ac.uk/gwas), contains as of today 2,670 studies and 30,321 unique SNP-trait associations [108]. These studies have changed our understanding of complex trait genetics. The hypothetical basis for GWAS is the *common disease – common variant* paradigm, which is founded on the assumption that for heritable diseases, that are common at the population level, the contributing genetic variants will be common in the population as well [109]. Studies comprising huge cohorts (beyond 100,000 samples for some common complex traits) have been carried out, often leading to the identification of numerous associated variants. However, the effect sizes are generally small and a large proportion of heritability still remains unexplained [110]. This *missing heritability* has led to considerations about the impact of rare variants, which are typically missed by GWAS, on susceptibility for complex common diseases [111]. Instruments for identification of rare variants with potentially large effect size or even *de novo* variants that may contribute to complex traits are various sequencing-based methods, either targeted to candidate genes or regions, or covering the whole exome or whole genome [112].

A major challenge in association studies is the functional biological interpretation of the identified associated variants, as they predominantly map to non-coding regions of the human genome. The estimated number of protein coding genes is around ~20,000, and they account for less than 2% of the sequence of our genome [113]. It has been suggested that intergenic variants are functional by modifying regulatory mechanism, and in accordance to their location and function they are divided into enhancers, promoters, silencers, locus control regions and insulators [114]. The ENCODE (ENCyclopedia Of DNA Elements) project is a long-term initiative involving a huge international consortium of research groups with the primary aim to identify all functional elements in the human genome sequence and to make these data publically available (https://www.encodeproject.org), along with the development and establishment of new high-throughput technology that is needed for this endeavor [115]. The ENCODE database comprises extensive information from experiments in different cell and tissue types about e.g. gene expression, transcription factor binding, histone mark distribution, chromatin states and variant annotation, and is continuously expanding [116].

Genetic variation in pSS and SLE

Evidence for the importance of the genetic component for susceptibility to pSS and SLE is revealed by genetic epidemiologic studies in affected families. The concordance rate of SLE in monozygotic twins is ~24-50% and in dizygotic twins ~2-5% [117]. Familial aggregation of SLE is well known, with a ~10% risk for first-degree relatives to be affected as well, and also in non-diseased relatives a positive ANA test is more commonly seen than in the general population [118]. Familial concordance rates are similar in other autoimmune diseases, including RA and pSS, although there are much fewer heritability studies in pSS [18]. Further, an increased prevalence for other autoimmune diseases in families with SLE and pSS has been reported [119, 120]. The observation of familial co-aggregation with other autoimmune diseases adds further evidence to a genetic contribution in SLE and pSS and suggests that the genetic background between different autoimmune diseases overlaps at least in parts. Indeed, a substantial proportion of susceptibility genes are shared among SLE and pSS suggesting common molecular pathways.

The majority of the identified genetic risk variants in SLE have only a modest individual effect size. Exceptions of rare variants with high prevalence for disease exist and they are mainly found within the complement (C) system, where ~90% of the individuals with homozygote C1q-deficiency develop SLE [121]. However, C1q-defiency is an extremely rare condition. The risk for SLE is about 75% for C4-deficiency and about 20% for C2-deficiency. Rare mutations in the three prime exonuclease I (*TREX1*) gene and in the deoxyribonuclease I (*DNASE1*) gene, both genes with implications in breakdown of nuclear components during apoptotic processes, have also been reported in SLE [122, 123].

With the advance of genotyping techniques it became possible to perform genome-wide linkage and large-scale association studies, and to identify more general risk genes for autoimmune diseases. The first GWAS in a small SLE cohort was published in 2007 [124]. This initial study was followed by a series of GWAS with improved power and study design analyzing genetic association with SLE in several different populations. Among common genetic variants, risk alleles within the major histocompatibility complex (MHC) region on chromosome 6 (6p21) display the strongest association with SLE and pSS. However, the extensive linkage disequilibrium (LD) in the MHC region has been a major obstacle for identifying the causal variants [125]. Most risk variants outside the MHC locus are linked to pathways of innate and adaptive immunity, such as B and T cell function, IC clearance, and type I IFN and NF-κB signaling. To date, more than 50 non-MHC risk genes for SLE and nearly10 for pSS have been identified with genome-wide significance (Table 3). The smaller number of risk variants in

pSS reflects the rather late start of genetic research in pSS compared to SLE and many other autoimmune diseases.

The strongest non-MHC association in both pSS and SLE is displayed at the interferon regulatory factor 5 (*IRF5*) gene [16, 126-128]. *IRF5* is a transcription factor involved in regulation of type I IFN genes and production of IFN-α [129]. Association between *IRF5* gene variants and IFN-α production by pDCs has been demonstrated [130].

Table 3. *Non-MHC risk loci associated with genome-wide significance or replicated for pSS and/or SLE.*

Pathway	Gene	Region	Disease
Innate Immunity			
TLR and IFN signaling	*IFIH1*	2q24	SLE [131-133]
	STAT4	2q32	pSS [16, 17], SLE [134-137]
	miR146a	5q34	SLE [131, 138, 139]
	PRDM1	6q21	pSS [16], SLE [140]
	IRF5-TNPO3	7q32	pSS [16, 128], SLE [115, 126, 134-137]
	IRF7-PHRF1	11p15.5	SLE [131, 134]
	CIITA-SOCS1	16p13.13	SLE [131]
	IRF8	16q24.1	SLE [132, 141]
	TYK2	19p13.2	SLE [126, 131, 132]
	TLR7	Xp22.3	SLE [142]
NF-κB signaling	*TNIP1*	5q33.1	pSS [16, 143], SLE [136, 140]
	TNFAIP3	6q23	pSS [16, 17, 143], SLE [137, 140]
	SLC15A4	12q24.32	SLE [136, 144]
	PRKCB	16p11.2	SLE [145]
	UBE2L3	22q11.21	SLE [134, 135]
	IRAK1-MECP2	Xq28	SLE [131, 136, 146, 147]
DNA degeneration and apoptosis	*ATG5*	6q21	SLE [134, 135]
	RAD51B	14q23-q24.2	SLE [131, 136]
	PRPS2	Xp22.3	SLE [136, 148]
Immune complex processing	*FCGR2A/B*	1q23	pSS [16], SLE [134, 135]
Neutrophil and monocyte function	*ITGAM*	16p11.2	SLE [134-136, 149]

		ICAM1,-4,-5	19p13.2	SLE [150]
Adaptive immunity				
	B cell function and signaling	*NCF2*	1q25	SLE [132, 134-136]
		IL10	1q31-q32	SLE [140]
		RASGRP3	2p25.1-p24.1	SLE [144]
		BANK1	4q24	SLE [137, 151, 152]
		MSH5	6p21.3	SLE [134, 153]
		BLK-FAM167A	8p23-p22	pSS [16, 154], SLE [135, 137, 155]
		LYN	8q12	SLE [134, 135]
		ETS1	11q23.3	SLE [137, 144, 146]
		ELF1	13q13	SLE [156]
		CXCR5	11q23.3	pSS [16]
	T cell function and signaling	*PTPN22*	1p13.2	SLE [134, 135, 140, 157]
		TNFSF4	1q25	SLE [136, 137, 152]
		CD80	3q13.3-q21	SLE [158]
		IL12A	3q25.33	pSS [16], SLE [131]
		TCF7	5q31.1	SLE [131]
		PDHX-CD44	11p13	SLE [159]
	B and T cell function and signaling	*IKZF2*	2q34	SLE [131]
		AFF1	4q21	SLE [160]
		IKZF1	7p12.2	SLE [132, 144, 146]
		CSK	15q24.1	SLE [131, 161]
		IKZF3	17q21	SLE [141]
Other genes				
		NMNAT2-SMG7	1q25.3	SLE [134, 144, 162]
		PTTG1	5q35.1	pSS [16]
		UHRF1BP1	6p21	SLE [140]
		GTF2I	7q11.23	pSS [17]
		JAZF1	7p15	SLE [136, 140]
		XKR6	8p23.1	SLE [134, 135]
		WDFY4	10q11.22	SLE [136, 144, 146]
		ARID5B	10q21	SLE [131, 158]

	DHCR7-NADSYN1	11q13.4	SLE [131]
	ATG16L2-FCHSD2-P2RY2	11q13.4	SLE [146]
	SH2B3	12q24	SLE [131]
	PLD2	17p13.1	SLE [131]

Modified from Teruel and Alarcón-Riquelme, 2016 [163], and Nezos and Mavragani, 2015 [164]

In light of the clinical heterogeneity of pSS and SLE, association studies aiming at identification of genetic variants in sub-phenotypes of the disease are performed. In pSS, genetic polymorphisms associated with increased risk for formation of ectopic germinal center-like structures have been reported [165] and multiple genes within the NF-κB signaling pathway, most notably in the TNFAIP3 interacting protein 1 (*TNIP1*) gene, are associated with anti-SSA/-B positive pSS [143]. In SLE, risk alleles of signal transducer and activator of transcription 4 (*STAT4)* have been implicated in anti-dsDNA and antiphospholipid autoantibody positivity, severe lupus nephritis and earlier onset of the disease [166-169], and variants at interferon regulatory factor 8 (*IRF8*), another transcription factor within the type I IFN system, are associated with cardiovascular disease in SLE [170].

Gene expression

Gene expression is the process by which the information in the DNA sequence is transcribed into a functional product. Apart from erythrocytes, which do not contain a nucleus, and gametes, which possess haploid genomes, the genetic content is the same in every cell of the human body. Cell and tissue type diversity as well as dynamic responses to changes in the environment are accomplished by transcriptional activation and silencing of certain parts of the genome in order to modify the expression levels of genes. Gene transcription and expression are highly regulated and coordinated processes, which are controlled in a spatial-temporal manner by genetic and epigenetic mechanisms. Transcription factors (TFs) are a diverse group of proteins with a common ability to recognize and bind to specific sites in the DNA sequence, thereby affecting the transcription machinery and promoting either increased or decreased transcription levels [171]. General TFs build the preinitiation complex (PIC) together with a variety of co-factors to recruit RNA polymerase II to the core promoter in order to initiate gene transcription.

RNA comprises both protein coding messenger RNA (mRNA) and various species of ncRNAs. Together these RNA molecules constitute the *transcriptome*. Ribosomal RNA (rRNA) and transfer RNA (tRNA) account for

the majority of ncRNAs in the cell and are fundamental for protein translation from mRNA. In addition, RNA molecules also act as important regulators with implications in tissue differentiation, development, proliferation and cell metabolism [172]. Long non-coding RNAs (lincRNAs), which are particularly important for genomic imprinting processes, and micro-RNAs (miRNAs), which can interfere with the conversion of mRNA into protein, are the most studied species of ncRNAs. MiRNAs are small RNA molecules with a length of around 19-25 nucleotides that can bind complementary to target mRNA for degradation or inhibition of translation by decapping and deadenylation [173]. Aberrant expression of miRNAs has been linked to a wide range of diseases, including cancer, infectious disease and autoimmunity. Notably, miRNA expression and DNA methylation appear to be tightly interconnected and deregulation of these processes has been described in SLE [174].

Epigenetics of autoimmune diseases

The human epigenome

The prefix *epi-* (Greek: επι- for *over*, *outer*, *above of*) indicates that *epigenetic mechanisms* act *on top* or *outside* of genetics. Epigenetic modifications comprise the heritable changes on the cellular level, which influence the phenotype, but without involving any direct alteration in the DNA sequence. Epigenetic regulation is primarily accomplished by DNA methylation, histone modifications and chromatin remodeling, and is central for development, differentiation and cell identity (Figure 2) [175]. DNA methylation and histone modifications involve covalent post-translational modifications, while chromatin remodeling enzymes utilize energy from ATP hydrolysis to facilitate nucleosome restructuring processes. In general, epigenetic marks are mitotically heritable on the cellular level. Whether they may be, in part, also trans-generationally heritable on the population level is a matter of ongoing debate and research [176]. Except in germ cells and pre-implantation embryos, epigenomic states are semi-stable, meaning that they are relatively stable over time and by that capable to maintain cell identity, but still dynamic enough to be able to change in response to external influences [177].

Complex autoimmune diseases cannot be explained by genetic predisposition and environmental factors alone, and reports of altered epigenetic landscapes in these diseases strongly indicate the contribution of epigenetic mechanisms in disease development and progression [6], and the reversible nature of epigenetic modifications suggest them as a potential target for therapeutic intervention.

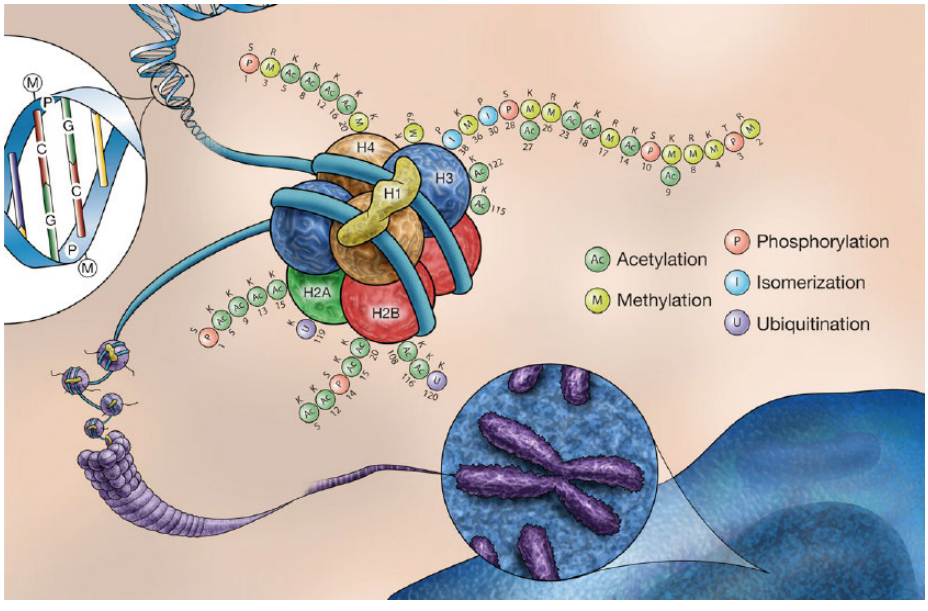

Figure 2. Schematic representation of chromatin organisation within the nucleus. The fundamental repeating units of chromatin are the nucleosomes, which consist of 147 bp double-stranded DNA wrapped around a histone octamer. Each histone octamer contains two H2A-H2B dimers and one H3-H4 tetramer, associated with one linker histone H1. Chromatin structure is influenced by post-translational modifications at core histones and DNA methylation at CpG dinucleotides. The figure was originally published by Barber *et al.* (2010) [178], and is reprinted with permission from the author.

DNA methylation

Methylation of DNA cytosine residues, mainly in the context of cytosine-guanine dinucleotides, denoted CpG sites (where *p* stands for phosphate separating the C and the G nucleotides), is the most widely studied epigenetic mark. The human genome comprises roughly 28 million CpG sites, which are unevenly distributed across the genome [179]. Due to an increased mutation rate by spontaneous deamination of 5-methylcytosine (5-mC) to thymidine, CpG sites are relatively underrepresented in the genome. The majority of CpG sites (around 70-80%) are sparsely scattered, enriched at repetitive elements and almost exclusively methylated [180]. In contrast, DNA sequences with high CG-density (> 50 %) are abundantly located in promoter regions of genes, are referred to as *CpG islands* and are predominantly non-methylated. The methylation status at CpG islands may interfere with the binding capabilities of transcription factors. Commonly, DNA methylation at CpG islands in promoter regions is regarded as repressive of gene transcription, whereas methylation of gene bodies has been associated with active transcription. However, this model is an oversimplification since the correlation (positive and negative) of DNA methylation levels and the expression of

nearby genes is often found to be limited [181]. In addition, a considerable number of CpG sites are located in intergenic regions, suggesting their involvement in long distance regulatory processes.

DNA methylation is established *de novo* and maintained during cell division by DNA methyltransferases (DNMTs) that transfer a methyl group from S-adenosyl-methionine to the carbon-5 position of cytosine, thereby generating 5-mC [182]. Removal of DNA methylation is achieved by ten-eleven translocation enzymes (TET), which actively remove the methyl group by iterative oxidation, followed by base excision repair. In this process 5-hydroxymethylcytosine (5-hmC) is passed as an intermediate state [183]. Functional roles for 5-hmC have been suggested particularly in embryonic development and in the nervous system [184]. Indirect removal of DNA methylation marks occurs when DNMTs are inhibited during the DNA replication process, which is for example observed when the demethylating drug 5-azacytidine (5-aza) is applied [185].

Methods for analysis of DNA methylation

Studying DNA methylation is an active area of research and several different methods have been developed within the field. Global DNA methylation can be measured by e.g. high performance liquid chromatography (HPLC) [186], while the methylation status of single genes or regions can be assessed by pyrosequencing [187]. Antibody- and restriction enzyme based methods, such as methylated DNA immunoprecipitation (MeDIP) and methyl-seq, give quantitative measurements on DNA methylation levels at genetic regions but lack single-base resolution. Whole-genome bisulfite sequencing (WGBS), which covers the complete genome, or reduced representation bisulfite sequencing (RRBS), where restriction enzymes are applied to limit the analysis to genomic regions of high CG abundance, are next-generation sequencing methods to quantify DNA methylation [188]. Although sequencing costs have dramatically decreased during the last couple of years, these methods are still cost intense, and involve complex data analyses. Methylation arrays targeting hundreds of thousands of predefined CpG sites are a widely used alternative. They combine comprehensive coverage of DNA methylation across the genome at base-pair resolution with high-throughput capabilities at prices that allow for investigation of larger cohorts. Sequencing- and array-based methods have in common that both rely on pretreatment of the DNA with sodium bisulfite, converting unmethylated cytosine into uracil and upon PCR amplification into thymidine, while leaving methylated cytosine unchanged (Figure 3) [189]. Bisulfite conversion is also the starting point for quantitative DNA methylation analysis by massARRAY technology (Sequenom). For this application bisulfite converted and PCR amplified DNA is subjected to base-specific enzymatic cleavage, where cleavage patterns depend on the methylation status of cytosine in the original genomic

DNA. The cleavage products are quantitatively analyzed by matrix-assisted laser desorption/ionization – time-of-flight (MALDI-TOF) mass spectrometry [190].

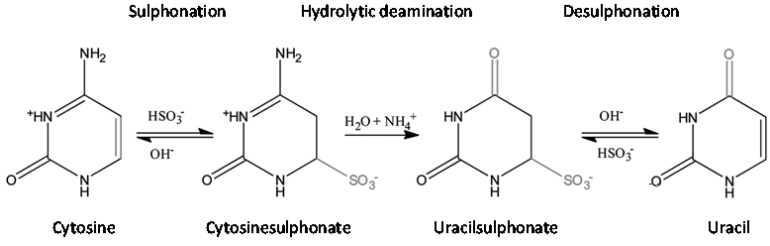

Figure 3. Sodium bisulfite conversion of genomic DNA. Bisulfite conversion involves the deamination of unmodified cytosine to uracil, whereas methylated cytosine remains as cytosine. The reaction comprises three steps: The first two steps are carried out under acidic conditions and high temperature and include sulphonation at the C-6 position of cytosine generating cytosinesulphonate, and subsequent irreversible hydrolytic deamination at the C-4 position of cytosinesulphonate to generate uracilsulphonate. The third step is carried out under alkaline conditions at room temperature and involves desulphonation of uracilsuphonate into uracil. A methyl group at the C-5 position of cytosine inhibits sulphonation at the C-6 position in the first reaction step and prevents 5-mC to from alteration by sodium bisulfite treatment.

Array-based analysis of DNA methylation

The Infinium HumanMethylation450 (HM450k) BeadChip (Illumina) interrogates quantitative DNA methylation levels at ~480,000 CpG sites with single-base resolution. The array covers ~99% of the Reference Sequence database (RefSeq) genes with multiple probes per gene (on average 17 probes per gene). These probes are distributed across promoter and non-promoter regions, CpG islands and regions outside of CpG islands, regions that are annotated to genes and intergenic regions [191].

The general design of the HM450k array is based on Infinium technology using base extension with biotin-labeled dinucleotides and subsequent multi-layer immunohistochemistry sandwich staining with fluorescent dyes [192]. The array comprises beads with attached probes that are complementary to specific target CpG sites. Methylation status-dependent base extension is performed and signal detection is achieved by optical scanning. A sample input of 250-500 ng genomic DNA is subjected to bisulfite conversion, amplification and fragmentation, and is thereafter applied to the HM450k array allowing the DNA fragments to hybridize to the attached probes. The HM450k array combines two different types of probe design in order to provide comprehensive coverage of the genome (Figure 4): About 30% of the probes on the array are Infinium type I probes, which utilize allele-specific hybridization of target alleles to two different bead types, that only differ in

the last base; one bead type for methylated alleles and the other for unmethylated alleles. The signal from the two different bead types is interrogated in one color channel. The remaining ~70% of probes have the Infinium type II design, which includes one bead type, and methylation status of either methylated or unmethylated cytosine is detected in two color channels. As read-out, DNA methylation at each probe is assessed as the quantitative measurement (calculated as beta-value, β-value) of bisulfite converted in relation to unconverted genomic DNA [191].

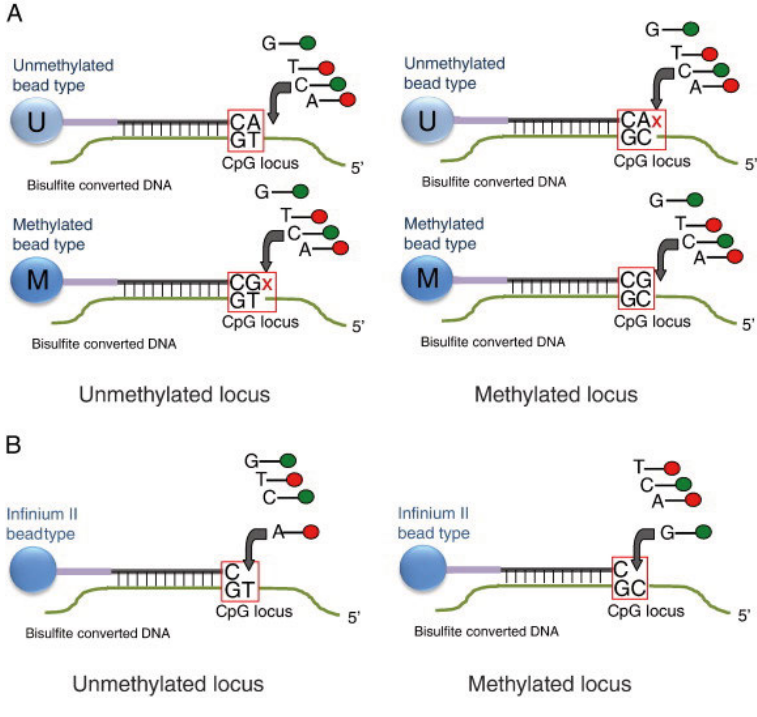

Figure 4. Probe design of the HM450k methylation assay. (A) Around 30% of the probes on the array have the type I probe design, which is based on allele-specific hybridization to target alleles. These probes come in pairs of two different bead types that only differ in the last nucleotide; one bead type for methylated cytosine, the other one for unmethylated cytosine. Signal detection from the two bead types is achieved by optical scanning using one color channel. (B) Type II probes only include one bead type and are based on single base extension depending on methylation status of the target CpG site. The signal from type II probes is interrogated in two color channels. The figure was originally published by Bibikova *et al.* (2011) [191], and is reprinted with permission from the author.

The HM450k array and its predecessor, the 27k array, have been widely used for *epigenome-wide association studies* (EWAS); a term that may be misleading, as the epigenome in not comprised by DNA methylation alone. Instead, *EWAS* refers to the concept of investigating an epigenetic mark (DNA methylation) in an association study on a genome-wide scale, analogously to the strategy of a GWAS.

Advances in genetics and epigenetics shed light on the importance of enhancers for regulation of gene function. Enhancers are regulatory genomic elements made up by short stretches of DNA (about 100-1500 bp) that serve as binding sequences for transcription factor complexes and interact with promoters to control gene expression. The majority of enhancers seem to act in *cis*, but *trans*-acting enhancers up to 1 Mbp away from the target gene exist as well [193]. GWAS variants are predominantly located outside of protein-coding regions and they are enriched for promoter and enhancer elements in many complex traits, including SLE [194]. Similarly, EWAS associations frequently overlap with genomic coordinates of enhancer marks. Since enhancers regions are in general underrepresented on the HM450k BeadChip, a novel array with extended enhancer coverage, the MethylationEPIC BeadChip (Illumina) containing over 850,000 CpG sites was released in 2015. The new EPIC array comprises > 90% of the probes from the HM450k array together with additional > 333,000 CpG sites located in enhancer regions as identified by the ENCODE and FANTOM5 projects [195]. The precise role of DNA methylation at enhancer elements is yet to be defined. However, it is suggested that the methylation status affects enhancer activity [116].

The role of DNA methylation in pSS and SLE

Until recently, DNA methylation in pSS has been mainly studied in candidate genes or on a global scale without single CpG site resolution. In CD4+ T cells inverse correlation of methylation and expression of *CD70* and *FOXP3* has been reported, with hypomethylation and overexpression of *CD70*, and hypermethylation and decreased expression of *FOXP3* in patients with pSS [196, 197]. In 2014, the first EWAS in pSS was published by Altorok *et al.* [100]. The authors investigated DNA methylation in naïve CD4+ T cells originating from eleven pSS patients and eleven matched controls and reported hypomethylation of several sites of lymphotoxin-alpha (*LTA*) and at IFN induced genes [100]. Miceli-Richard *et al.* studied methylation patterns in fractionated CD19+ B cells and CD4+ T cells in 26 female patients with pSS and 22 age-matched controls, and observed larger differences in DNA methylation between patients and controls in B cells compared to T cells [198].

Similar to pSS, during the last decade, studies of DNA methylation in SLE have moved from accessing candidate genes or global methylation lev-

els to the analysis of methylation patterns on the genome-wide scale. It has been known for a long time that DNA methylation inhibitors like procainamide and hydralazine can cause drug-induced lupus [199]. In 1990, Richardson *et al*. described global hypomethylation in CD4+ T cells in SLE [200]. About ten years later a number of follow-up studies investigating DNA methylation of candidate genes in SLE were published. In these studies of SLE CD4+ T cells it has been suggested that hypomethylation of genes encoding surface molecules like CD11a and CD70 may mediate the overexpression of genes that contribute to the formation of autoreactive T cells and to the increased production of immunoglobulins by B cells [201-203]. A role for miRNAs, which may target *DNMT1*, has been suggested as a contributing factor for alterations of T cell methylation in SLE [204, 205]. In addition, demethylation of CpG sites at *CD40LG* on the inactive X-chromosome has been proposed as a possible factor contributing to the predominance of SLE in women [206].

On the genome-wide scale differential methylation in SLE has been analyzed in fractionated blood cells in several studies of relatively limited sample size. The principle finding of these studies was hypomethylation of type I IFN regulated genes [99, 207, 208]. Differential methylation has also been suggested as one plausible mechanism behind monozygotic twin discordance in SLE [209]. In addition, associations between DNA methylation and different manifestations of the disease have been reported, including autoantibody production, renal disorder and different forms of skin rashes [210-213]. These findings are yet to be independently replicated.

Histone modifications

The nucleosome is the basic functional unit of chromatin. The core of the nucleosome is built up by histone octamers containing two copies each of histone H2A, H2B, H3 and H4. Each histone octamer is wrapped around by 147 bp of DNA and associated with one linker histone H1 (figure 2). Histone proteins are essential for nuclear architecture and genome stability, but they are also involved in regulatory processes of gene transcription by modifying the accessibility of the chromatin for the transcription machinery [214]. Much simplified, euchromatin is loosely packed and associated with active transcription, whereas heterochromatin has a very condensed conformation and is therefore associated with repression of transcription.

The N-terminal tails of histones protrude out of the nucleosome and are subject to a variety of covalent post-translational modifications, including methylation, acetylation, phosphorylation, ubiquitination, ADP-ribosylation, biotinylation, citrullination and SUMOylation, all of which thereby affecting the chromatin structure, and specific modifications have been associated with induction or inhibition of gene expression [215].

Acetylation and methylation of lysine (K) residues at histones H3 and H4 are the most studied marks. Histones are positively charged and form very tight interactions with the negatively charged DNA that is wrapped around them. When histone acetyltransferases (HATs) facilitate the acetylation of lysine residues at the histone tails, this process partly neutralizes the positive charge of the histone and consequently changes the chromatin into a more accessible conformation. Acetylation of lysine 27 at histone 3 is associated with active enhancer elements (H3K27ac). On the contrary, histone deacetylases (HDACs) remove acetyl groups from lysine residues, which leads to a more compact chromatin conformation and is therefore associated with transcriptional repression [216].

Lysine methylation does not impose an overall net charge on the nucleosome. Histone lysines can be mono-, di- or trimethylated by histone lysine methyltransferases (KMTs), while histone lysine demethylases (KDMs) remove methyl groups. Methylation of lysine 4 at histone 3 is enriched at active promoters (H3K4me3) and enhancers (H3K4me1), whereas methylation at lysine 9 and 27 (H3K9me3 and H3K27me3) is associated with transcriptional repression and heterochromatin state. Transcribed gene bodies are enriched for lysine 36 trimethylation (H3K36me3) [114].

Methods for analysis of histone modifications

A widely applied technology to assess histone modifications is chromatin immunoprecipitation (ChIP) followed by sequencing, or short ChIP-seq. For ChIP-seq the DNA of intact cells is cross-linked to the bound proteins by chemical or physical fixation. Upon cell lysis the chromatin is fragmented and immunoprecipitated by antibodies against the target protein (in this case certain histone proteins). Complex enrichment is achieved by the affinity between the antibody and IgG-binding protein-A or-G. This is followed by reverse-crosslinking, and purification and recovery of the DNA fragments. The fragments can then be analyzed by DNA sequencing. When aligned to the reference genome, sequencing reads originating from immunoprecipitated fragments pile up in form of *peaks*, marking the genomic regions to which the histone protein of interest had been bound [217].

The role of histone modifications in autoimmune diseases

The involvement of histone modification in the pathogenesis of autoimmune disease is suggested by altered patterns of histone modifications, including global H3 and H4 hypoacetylation and H3K9 hypomethylation in CD4+ T cells and CD14+ monocytes from SLE patients compared to controls [218, 219]. The inhibition of HDAC proteins resulting in global increase of histone acetylation was reported to improve symptoms in mice models for SLE and RA [220, 221]. HDAC inhibitors are exploited as cancer drugs, where

they have shown effectiveness in inhibition of tumor growth [222], while their potential role for treatment of autoimmunity is less clear.

It is important to keep in mind that neither DNA methylation nor histone modifications function on their own but instead are inherently interrelated and depending on each other to contribute to chromosome architecture and spatial-temporal regulation of genome function.

The present investigation

Aims

The general aim of this thesis was to investigate the role of DNA methylation in systemic autoimmune diseases, and to study gene expression patterns associated with autoimmunity. The specific objectives of the work in this thesis were:

- Paper I – To analyze the role of DNA methylation in the etiology of pSS.
- Paper II – To identify DNA methylation patterns associated with fatigue in pSS.
- Paper III – To decipher the role of B cell-specific gene expression in the pathophysiology of pSS.
- Paper IV – To study the impact of DNA methylation on SLE susceptibility and phenotype variability.

Patients and methods

Several methods have been applied in the different studies in this thesis, and they are specified in detail the in the individual papers. The most important methodology of the studies is described in brief in this section.

Patients and controls

All individuals included in the studies gave informed consent to participate and all studies were approved by the regional ethics boards.

Primary Sjögren's syndrome patients

All pSS patients (paper I-III) fulfilled the AECG criteria for pSS [14]. The first study (paper I) included 108 individual patients with pSS from the Rheumatology Clinic at Uppsala University Hospital, Sweden. Whole blood samples were collected from 100 patients, CD19+ B cells from 24 patients and minor salivary gland biopsies from 15 patients. As controls, whole blood samples from 400 and CD19+ B cells from 47 healthy blood donors from the Uppsala Bioresource [130], Uppsala University Hospital, were included. Minor salivary gland biopsies from 13 individuals undergoing evaluation for possible pSS, where clinical examination, biopsies and sera did not confirm a pSS diagnosis, were included as control biopsies. The second study (paper II) was conducted within a pSS cohort from Stavanger University Hospital, Norway, which has been described previously [223]. Briefly, the cohort comprises in total 124 individuals with pSS. For this study whole blood samples from the 24 patients presenting with the highest and the 24 patients presenting with the lowest fatigue visual analogue scores (f-VAS) were drawn from the cohort. The third study (paper III) was performed in a subset of pSS patients and controls included in study I, for which RNA samples from CD19+ B cells were available (patients n=12, controls n=20). The validation cohort consisted of 17 pSS patients from the Rheumatology Unit at the Karolinska University Hospital, Sweden, and 16 healthy controls from the same residential area. All patients participating in the third study were females, positive for anti-SSA autoantibodies in their sera and untreated regarding immunomodulatory drugs.

SLE patients

The fourth study (paper IV) comprised 347 SLE patients from the University Hospitals in Uppsala and Linköping, Sweden, and 400 blood donor controls from the Uppsala Bioresource [130] in the discovery phase. In the replication phase 201 SLE patients and 187 controls from the Karolinska University Hospital were included. Patients and controls were matched for age, sex, ethnicity and residential area. All SLE patients fulfilled at least four out of the eleven ACR 1982 revised criteria for classification of SLE [64].

Sample preparation

Genomic DNA was prepared from whole blood samples using the QIAamp DNA Blood Mini Kit (Qiagen). Primary CD19+ B cells were fractionated from buffy coats applying Ficoll density gradient centrifugation (GE Healthcare) for isolation of PBMCs followed by positive selection with CD19 specific microbeads (Miltenyi Biotec). The AllPrep DNA/RNA Mini and the AllPrep DNA/RNA/miRNA Universal kits (Qiagen) were used for preparation of DNA and RNA from CD19+ B cells. Preparation of genomic DNA from minor salivary gland biopsies was performed using the Qiagen DNeasy Blood and Tissue kit (Qiagen).

Methylation analysis

Methylation assay and data processing

Bisulfite treatment was performed on 500 ng of genomic DNA using the EZ DNA Methylation-Gold Kit (Zymo Research). Samples were randomized on each BeadChip to avoid batch effects. Amplification, enzymatic fragmentation and hybridization procedures were performed according the Illumina standard protocol. Signal intensities were obtained by standard Illumina imaging steps using the iScan instrument resulting in raw intensity data (idat files). All data processing was performed within the R Bioconductor environment [224]. The raw intensity data files were imported into the Minfi R software package, which was used for quality control (QC), Subset-quantile Within Array Normalization (SWAN) and calculation of methylation β-values as the fraction of signal intensity obtained from methylated beads over the total intensity (0-1, corresponding to 0-100% methylation) [225, 226]. Probes with a detection p-value > 0.01 were not considered to be significantly different from background noise and were filtered out. In addition, probes with incomplete data for any of the samples were removed. The post-QC data sets comprised only unambiguously mapping probes, that do not contain any SNP in the last 10 bp of the probe according to 1000 genomes CEU populations, release 3 [227] and that do not overlap any common copy number variation (CNV).

Estimation of cell type distribution

To avoid potential confounding by differential distribution of cell types between the investigated traits, an adjustment for cell type proportions was included in the statistical model. Cell type fractions for each whole blood sample were estimated applying a published algorithm that uses publicly available reference DNA methylation profiles derived from six flow sorted blood cell types; CD4+ and CD8+ T cells, NK cells, B cells, monocytes and granulocytes [228, 229].

Pathway analysis and characterization of IFN-regulated genes

Functional pathway analysis was conducted to identify molecular processes and biological functions and diseases the genes with differential methylation or differential expression were involved in. Functional gene-set enrichment analysis was performed using Ingenuity Pathway Analysis software (IPA, Qiagen). The Interferome v2.01 analysis tool was applied to characterize whether differentially methylated or differentially expressed genes were known to be to IFN-regulated [230].

Functional genomic annotation

Probe mapping and annotation in relation to CpG island context and in relation to functional gene regions were performed as previously described [231]. The functional regional distribution of differentially methylated CpG sites (DMCs) was compared with the distribution of probes on the HM450k array passing QC procedures. In order to investigate whether DMCs are enriched for overlap with specific chromatin modification marks, publicly available reference data from primary CD19+ B cells and CD3+ T cells derived from the NIH Roadmaps Epigenomics Project were used [232]. Histone modification peaks for H3K4me1, H3K4me3, H3K27ac, H3K36me3, H3K9me3 and H3K27me3, and DNase hypersensitivity sites (DHS) were analyzed for regional overlap with the genomic coordinates of DMCs.

Genotyping

SNP genotype data were generated on the Infinium ImmunoChip (Illumina) containing 196,524 probes [233]. Genotype QC at the sample level was performed as previously described [130] and SNP probes fulfilling ≥ 98% call rate, a Hardy-Weinberg equilibrium test p-value > $1x10^{-4}$ and a MAF of ≥ 1% were included, resulting in 133,115 genetic variants after filtering.

RNA-sequencing

Sequencing libraries were prepared from 1µg of total RNA applying the TruSeq stranded mRNA sample preparation kit including poly-A selection (Illumina), followed by paired-end sequencing (50bp read length) on a HiSeq2500 instrument (Illumina), resulting on an average of 20 M read-pairs per sample. QC of RNA-sequencing (RNA-seq) data was performed using RNA-SeQC [234] and sequencing reads were mapped to the human reference genome (GRCh37) with TopHat2 [235]. The analysis of differential expression was conducted within the Cufflinks pipeline, including correction for a false discovery rate (FDR) of 0.05. Results were reported using fold change (FC) of mean fragments per kilobase of exon per million fragments mapped (FPKM) between the two phenotypes compared [236].

Statistical testing

Association tests

To test the association of methylation status of a single CpG site with a trait, the methylation levels for each group (e.g. cases and controls) were calculated (as β-values) and compared in an additive logistic regression analysis using the `anova(lm(y ~ trait))` function in R predicting methylation β-values at each probe as a function of the phenotype. When appropriate, correction for potential confounding factors such as cell type distribution, age and sex was included in the model. P-values were adjusted for multiple testing by Bonferroni correction ($p < 1.3 \times 10^{-7}$ in studies I and IV) or by the Benjamini-Hochberg approach for controlling the FDR ($p < 0.05$ in study II).

Cis-meQTL-analysis

For analysis of methylation quantitative trait loci, DNA methylation levels were tested in PLINK [237] for genotype association separately in patients and controls assuming an additive model and including BeadChip, BeadChip position, age, sex and estimated cell type distribution as covariates.

In study I, meQTLs were tested in the controls at CpG sites within a ±100 kb window around genetic variants that had been previously associated to pSS risk at genome-wide significance [16]. A Bonferroni corrected $\alpha < 0.05$ ($p < 1.24 \times 10^{-7}$) was considered significant.

In study IV, meQTLs were tested in patients with SLE and controls for all CpG sites passing QC on the HM450k array and all polymorphisms passing QC on the ImmunoChip within a ±100 kb flanking region around each CpG site (in total 7,703,665 association tests, including 385,851 CpG sites and 130,115 SNPs). A Bonferroni corrected $\alpha < 0.05$ was considered significant ($p < 6.5 \times 10^{-9}$). A t-test (`t.test` function in R) testing the difference in slope from the meQTL regression analysis was used to determine which of the significant *cis*-meQTLs were differentially regulated between patients and controls. A Bonferroni corrected $\alpha < 0.05$ in combination with a delta-beta value from the model of at least |0.05| between cases and controls was applied to define meQTLs with differential regulation.

Enrichment tests

Enrichments for a given property i.e. overlap with a histone mark, in a specific set of CpG sites (paper I, II and IV) were tested by creating a two-by-two contingency table and performing a χ^2-test using the `quisq.test` function in R applying a Bonferroni corrected significance threshold. Enrichment testing in paper III was conducted as described above; however, a Fisher's exact test was applied using the R function `fisher.test`. This modification was due to the smaller number of some of the tested observations in paper III. Effect sizes of all enrichment tests were reported as FC.

Results and discussion

Paper I

Genome-wide DNA methylation analysis in multiple tissues in primary Sjögren's syndrome reveals regulatory effects at interferon-induced genes

In this study we analyzed differential DNA methylation in patients with pSS and control individuals in multiple tissues. We investigated whole blood (patients n=100, controls n=400), purified primary CD19+ B cells (patients n=24, controls n=47) and minor salivary gland biopsies (patients n=15, controls n=13) using the HM450k array.

Due to the cell type specificity of DNA methylation profiles we adjusted for estimated leukocyte distributions in the linear regression model. Estimated blood cell fractions in pSS patients were shifted towards increased granulocyte and monocyte abundance, while both the CD4+ and CD8+ T cell subsets were reduced. The fraction of CD19+ B cells was equal in patients and controls, which is in line with earlier reports [238].

The analysis of DNA methylation in whole blood identified disease associated alterations at 11,785 CpG sites across the genome; of these, 6,171 were hypo- and 5,614 were hypermethylated in patients. The DMCs were annotated to 5,623 unique genes with most significantly associated sites annotated to genes involved in immune responses, in particular to type I IFN induced genes such as *IFI44L*, *IFIT1, IFITM1*, *MX1* and *PLSCR1*. In addition, hypomethylation of multiple sites at the *PARP9* locus, also known as B-aggressive lymphoma-1 (*BAL1*), was observed in whole blood as well as in CD19+ B cells from pSS patients. We tested whether differential DNA methylation patterns were similar in anti-SSA/-B negative and positive patients with pSS and found that the hypomethylation of type I IFN regulated genes was mainly seen in antibody positive patients, while methylation levels in antibody negative patients were closer to those observed in controls (Table 4).

To avoid confounding factors due to gender-specific differences in methylation patterns, the sex chromosomes were analyzed separately in females and males. In females, we identified 85 DMCs located on the X-chromosome to be associated with pSS, with about half of them presenting as hypomethylated. These DMCs were annotated to 56 unique genes, including several genes with implications in immunity, such as *TLR8*, *CD40LG* and immunoglobulin (CD79A) binding protein 1 (*IGBP1*). Furthermore, multiple DMCs annotated to *miRNA-223* were identified, which has been implicated in a variety of immune processes in autoimmunity, viral infections and several malignancies [239]. We did not observe any significant DMCs on the sex chromosomes in men, probably due to the small number of men in the study.

Table 4. *Differential DNA methylation in patients with pSS stratified on the presence of anti-SSA and/or anti-SSB antibodies. Results are presented for the CpG sites with the largest difference in mean methylation (delta-beta) between all patients with pSS and control individuals in the whole blood association study.*

			Anti-SSA/SSB positive pSS patients versus controls			Anti-SSA/SSB negative pSS patients versus controls			Anti-SSA/SSB positive versus negative pSS patients	
CpG site	Gene Symbol	Meth β in controls, n=400	Meth β in anti-SSA/SSB pos pSS, n=75	Meth Δβ[a]	p-value	Meth β in anti-SSA/SSB neg pSS, n=25	Meth Δβ[a]	p-value	Meth Δβ[a]	p-value
cg21549285	*MX1*	0.83	0.49	-0.33	$4.1x10^{-82}$	0.79	-0.04	$8.3x10^{-03}$	-0.30	$7.3x10^{-09}$
cg05696877	*IFI44L*	0.67	0.42	-0.25	$6.1x10^{-70}$	0.64	-0.03	$5.6x10^{-02}$	-0.22	$3.8x10^{-10}$
cg22930808	*PARP9*	0.71	0.47	-0.24	$4.5x10^{-70}$	0.66	-0.05	$5.0x10^{-04}$	-0.19	$1.6x10^{-07}$
cg22862003	*MX1*	0.70	0.47	-0.23	$2.1x10^{-80}$	0.65	-0.05	$2.4x10^{-05}$	-0.18	$1.6x10^{-08}$
cg03607951	*IFI44L*	0.59	0.38	-0.21	$1.2x10^{-79}$	0.52	-0.07	$2.3x10^{-08}$	-0.13	$1.1x10^{-07}$
cg00959259	*PARP9*	0.58	0.36	-0.22	$1.7x10^{-67}$	0.56	-0.01	$3.4x10^{-01}$	-0.21	$8.6x10^{-10}$
cg06981309	*PLSCR1*	0.54	0.37	-0.17	$7.3x10^{-57}$	0.47	-0.07	$2.1x10^{-07}$	-0.10	$5.5x10^{-06}$
cg05552874	*IFIT1*	0.71	0.54	-0.17	$4.2x10^{-67}$	0.66	-0.05	$2.7x10^{-06}$	-0.12	$4.6x10^{-07}$
cg26312951	*MX1*	0.44	0.26	-0.18	$5.3x10^{-54}$	0.41	-0.03	$5.8x10^{-02}$	-0.15	$1.7x10^{-08}$
cg23570810	*IFITM1*	0.69	0.51	-0.18	$3.9x10^{-51}$	0.66	-0.02	$1.2x10^{-01}$	-0.16	$9.3x10^{-08}$
cg17608381	*HLA-A*	0.60	0.45	-0.15	$7.8x10^{-19}$	0.58	-0.02	$3.7x10^{-01}$	-0.13	$7.4x10^{-06}$
cg05825244	*EBF4*	0.47	0.57	0.10	$2.7x10^{-09}$	0.58	0.11	$7.3x10^{-05}$	-0.01	$9.5x10^{-01}$

[a] Methylation Δβ refers to the difference in mean methylation β (beta-values) between the two groups compared.

Pathway analysis of genes annotated to the most significantly associated DMCs was conducted to investigate the functional roles of the genes associated to the disease. Top gene-set enrichments were found for antigen-presentation and IFN signaling pathways, highlighting once again the central role of these signaling systems in pSS. A pronounced enrichment with respect to disease or functional annotation of DMCs was observed for lymphoproliferative malignant disorders ($p = 6x10^{-13}$), including B cell non-Hodgkin lymphoma. This is of interest, since the risk for lymphoma development is known to be elevated in pSS [12].

To gain further insights into possible functional implications of DMCs in pSS we analyzed the associated CpG sites in our data for overlap with chromatin marks in publicly available reference data sets from CD19+ B cells and CD3+ T cells [232]. Except for the repressive mark H3K27me3, for which both hypo- and hypermethylated DMCs were depleted, we observed opposed enrichment-depletion patterns for histone modification marks and DNase hypersensitivity sites (DHS) when comparing hypomethylated DMCs with hypermethylated DMCs (Figure 5). Hypomethylated DMCs were overrepresented in enhancer regions (H3K4me1 and H3K27ac) and DHS, while hypermethylated DMCs were underrepresented in these regions, and instead enriched for H3K36me3, which is a mark for actively transcribed gene bodies.

In the analysis of DNA methylation in purified primary CD19+ B cells we identified 453 DMCs, annotated to 303 unique genes. The prominent hypomethylation in whole blood at CpG sites in type I IFN induced genes and in *PARP9* was also observed in B cells, with an even more distinct difference in mean methylation values between patients and controls. This observation may indicate an advantage of investigating DNA methylation in a single cell type compared to studying whole blood samples adjusted for cell type distribution heterogeneity. However, all patients in the B cell analysis were positive for anti-SSA/-B autoantibodies and this created a more homogeneous study population, which may have contributed to the larger effect sizes seen in B cells compared to whole blood. For a subset of individuals (pSS n=16 and controls n=23) RNA samples from CD19+ B cells were available, and we investigated the RNA expression levels of the most differentially methylated genes. RNA-seq revealed a negative correlation between DNA methylation and expression levels for the genes with the ten most hypomethylated CpG sites in pSS.

		All associated DMCs	Hypomethylated DMCs	Hypermethylated DMCs
H3K4me1	CD19+	0.99	**1.45**	**0.53**
	CD3+	0.96	**1.31**	**0.62**
H3K4me3	CD19+	**0.63**	1.11	**0.14**
	CD3+	**0.68**	1.06	**0.30**
H3K27ac	CD19+	**0.85**	**1.34**	**0.36**
	CD3+	0.95	**1.27**	**0.63**
H3K36me3	CD19+	**1.45**	0.98	**1.91**
	CD3+	**1.31**	0.95	**1.66**
H3K27me3	CD19+	**0.40**	**0.60**	**0.20**
	CD3+	**0.26**	**0.37**	**0.15**
H3K9me3	CD19+	0.64	**0.29**	0.99
	CD3+	0.56	0.19	0.92
DHS	CD19+	**0.74**	**1.25**	**0.23**
	CD3+	**0.81**	**1.16**	**0.46**

Figure 5. Functional genomic distribution of differentially methylated CpG sites (DMCs) in whole blood from patients with pSS and controls. This heat-map shows the functional genomic distribution of all associated DMCs (first panel), of the hypomethylated DMCs (second panel) and of the hypermethylated DMCs (third panel) annotated in relation to chromatin states in CD19+ B cell and CD3+ T cell reference data sets [232]. The color scale indicates fold-enrichment (orange) or fold-depletion (blue) of the DMCs in each functionally annotated region. Bold numbers represent annotations to which DMCs significantly differ compared to the distribution of probes on the HM450k array (post-QC probe set) (Bonferroni corrected χ^2-test generated p-value < 0.005).

Since we were not able to account for differential cell type distribution in the association analysis in minor salivary gland biopsies, one must interpret these results with caution. Still, it is noteworthy that the most significant DMC (cg20870559) was annotated to *OAS2*, an IFN induced gene implicated in the innate immune response to viral infections [240]. Several hypomethylated *OAS2* sites were also found in the whole blood analysis, pointing to a potential role for this gene in the pathogenesis of pSS. In a recent study by Cole *et al.* the association at cg20870559 in pSS minor salivary gland tissue was replicated [241].

We investigated genetic loci of known pSS risk alleles for association with methylation levels at nearby CpG sites (±100kb from the associated SNP) and found evidence for genetic control of methylation levels at all pSS GWAS loci, in particular within the MHC region and at the *IRF5-TNPO3* locus. This indicates that several pSS risk alleles have the potential to affect DNA methylation levels of target genes.

In conclusion, evaluating DNA methylation profiles in multiple tissues from pSS patients and controls revealed distinct hypomethylation in enhancer and promoter regions of type I IFN regulated genes. In addition, differen-

tial methylation at genes with implication in lymphomagenesis was observed. We further report meQTLs of genetic variants at known pSS susceptibility loci with methylation levels at proximal CpG sites suggesting a potential functional mechanism for several of the pSS GWAS loci.

Paper II

Epigenome-wide DNA methylation patterns associated with fatigue in primary Sjögren's Syndrome

To investigate a possible association of differences in DNA methylation levels in whole blood with self-reported high versus low fatigue in pSS, we analyzed genome-wide DNA methylation in 24 pSS patients with high (f-VAS > 72 mm) and 24 pSS patients with low (f-VAS < 50 mm) score on the f-VAS using the HM450k array.

The relative cell type distribution in each sample was estimated using the same reference-based approach as applied in study I [228, 229], and we adjusted for cell type distribution along with age and sex in the linear regression model. Differential DNA methylation was identified at 251 CpG sites at an FDR of 0.05 and a cut-off of >|0.03| difference in mean methylation beta between patients with high and low fatigue (Figure 6). Of these DMCs, 166 were hypo- and 85 hypermethylated in the high fatigue group. The DMC with the most distinct difference in average methylation between the two groups was annotated to the ncRNA *SBF2-AS1*, displaying hypomethylation with a delta-beta of -0.09 in patients with high fatigue compared to patients with low fatigue. Functional implications of this antisense RNA have not been described previously and determining its function may be a target of future studies.

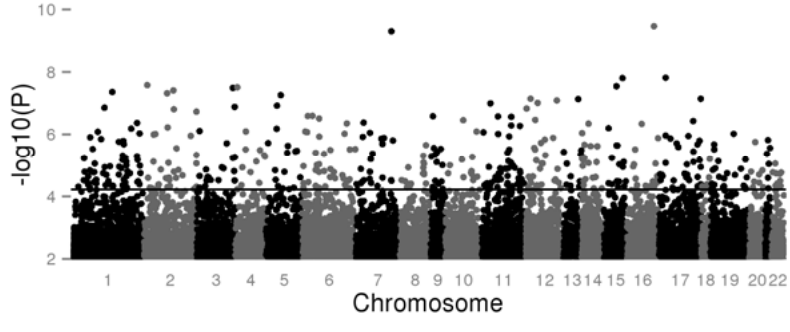

Figure 6. Epigenome-wide association results comparing DNA methylation levels in pSS patients with high versus pSS patients with low fatigue-VAS (f-VAS) score. The manhattan plot illustrates the $-\log_{10}$ p-value of each analyzed CpG site in relation to chromosomal position. A total of 251 CpG sites were identified as differentially methylated between patients with high and patients with low fatigue (FDR < 0.05 adjusted p-value < $6x10^{-5}$), as indicated by the black vertical line.

Functional pathway analysis of genes annotated to DMCs revealed enrichment for the role of nuclear factor of activated T cells (NFAT) in regulation of immune response as the most significantly associated pathway for our gene-set. Furthermore, phospholipidase C signaling and nerve growth factor IB (NGFIB, also denoted as Nur77) signaling in T lymphocytes were identified among the most enriched functional networks (Table 5). Nur77 is a key protein involved in inflammatory responses and apoptotic functions mediated by NF-κB signaling in macrophages and T cells [242]. Notably, IL-1 signaling was also among the overrepresented pathways in our study. IL-1, and particularly IL-1β, has been suggested as one of the key cytokines involved in fatigue generating and sustaining mechanisms [243].

Table 5. *Functional pathway analysis of differentially methylated CpG sites (DMCs) in pSS patients with high fatigue VAS (f-VAS) score (>72 mm) compared to patients with low f-VAS score (<50 mm).*

Pathway	Molecules	p-value
Role of NFAT in Regulation of the Immune Response	CD247, CD3G, GNAI2, GNG7, LCK, NFATC1,PLCB2, RCAN3, SYK	$2.6x10^{-5}$
Phospholipidase C Signaling	ADCY9, ARHGEF10, CD247, CD3G, GNG7, HDAC4, LCK, NFATC1, PLCB2, SYK	$6.3x10^{-5}$
Nur77 Signaling in T Lymphocytes	CD247, CD3G, MAP2K5, NFATC1, RXRA	$1x10^{-4}$
Cell Cycle G1/S Checkpoint Regulation	CUL1, E2F1, E2F3,, HDAC4, TFDP1	$2.5x10^{-4}$
Breast Cancer Regulation by Stathmin1	ADCY9, ARHGEF10, E2F1, E2F3, GNAI2, GNG7, LIMK2, PLCB2	$4x10^{-4}$
CCR5 Signaling in Macrophages	CD247, CD3G, FASLG, GNAI2, GNG7	$4.8x10^{-4}$
Cyclins and Cell Cycle Regulation	CUL1, E2F1, E2F3, HDAC4, TFDP1	$6.4x10^{-4}$
Hepatic Cholestasis	ADCY9, IRAK3, LTA, NR1H3, NR112, RXRA, TNF	$9.8x10^{-4}$
Estrogen-mediated S-phase Entry	E2F1, E2F3, TFDP1	$1.3x10^{-3}$
Role of NFAT in Cardiac Hypertrophy	ADCY9, GNAI2, GNG7, HDAC4, IGF1R, PLCB2, RCAN3	1.4x10-^{3}
FXR/RXR Activation	NR1H3, NR112, PLTP, RXRA, SLC51A, TNF	$1.4x10^{-3}$
IL-1 Signaling	ADCY9, GNAI2, GNG7, IRAK3, TOLLIP	$1.6x10^{-3}$
Acryl Hydrocarbon Receptor Signaling	AIP, E2F1, FASLG, RXRA, TFDP1, TNF	$1.7x10^{-3}$
Molecular Mechanisms of Cancer	ADCY9, ARHGEF10, E2F1, E2F3, FASLG, GNAI2, HIPK2, LRP5, PLCB2, TFDP1	$1.8x10^{-3}$
T Cell Receptor Signaling	CD247, CD3G, LCK, NFATC1, PTPN7	$2.2x10^{-3}$

The relative low number of samples combined with the investigation of a phenotype that is challenging to quantify are limitations to this study. Therefore it will be important to replicate the findings of this study in an independent cohort.

In summary, this study revealed differential DNA methylation profiles associated with fatigue levels in pSS. DNA methylation may be one mechanism contributing to development and regulation of fatigue, possibly via modification of NFAT and NF-κB regulation, and IL-1 mediated signaling.

Paper III

Transcription profiling of CD19+ B cells in primary Sjögren's syndrome reveals an interferon signature with upregulated BAFF and TLR7

Given the central role of B cells in the pathophysiology of pSS [2], the aim of this study was to investigate the B cell-specific contribution to the disease by gene expression profiling in patients and controls. Prior expression studies in PBMCs, pDCs, monocytes and salivary gland tissue from patients with pSS revealed overexpression of type I IFN inducible genes as well as of genes connected to other inflammatory and immune related pathways in pSS [87, 244-247], but differential gene expression in purified CD19+ B cells from patients with pSS has not been investigated on the whole-transcriptome level before. We performed RNA-seq of CD19+ B cells in a subgroup of individuals from study I (patients with pSS n=16, and blood donor controls n=20). All patients in the study were women, positive for anti-SSA and/or anti-SSB autoantibodies, untreated regarding immunosuppressive drugs and had no prior history of lymphoma.

Unsupervised analysis performed on the whole-transcriptome data of all samples included in the sequencing, resulted in clear separation between patients and control samples in the first two principal components (Figure 7), indicating wide-spread differences in gene expression between the two groups.

We detected in total 4,047 differentially expressed autosomal genes in B cells from patients with pSS compared to healthy controls. A profound IFN signature, including upregulation of *IFI44L*, *IFITM1*, *MX1*, *PARP9*, and *TNFSF13B*, encoding BAFF, was identified (Figure 8a). BAFF is important for B cell proliferation, survival of peripheral B cells and for induction of immunoglobulin switch [248]. Overexpression of *TNFSF13B/BAFF* has previously been reported from salivary glands and sera [45, 249], while this is the first time *TNFSF13B/BAFF* upregulation is observed directly in pSS B cells. In addition, the BAFF receptor *BCMA* was also found to be overexpressed in patients, while *TRAF3*, a negative regulator of BAFF function, was downregulated in pSS B cells. Taken together, the transcriptional

dysregulation of *BAFF* and genes associated with BAFF function may constitute a mechanism for ongoing B cell stimulation in pSS.

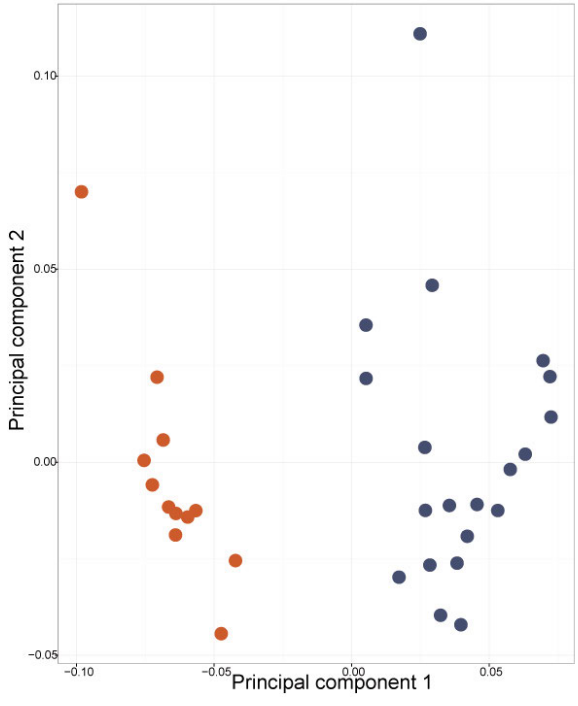

Figure 7. Principal component analysis (PCA) of CD19+ B cell gene expression data for all samples included in the RNA-sequencing analysis. The expression data from 12 patients with pSS and 20 control individuals are plotted using principal component one and two. Data from patient samples are shown in orange and data from controls are in blue.

Comparing gene expression in female patients (n=12) and controls (n=16), we identified 153 X-chromosomal differentially expressed genes, among these several genes with central functions for immune system signaling, such as *TLR7*, *TLR8* and complement factor properdin (*CFP*), which were upregulated in pSS. Notably, we also observed a number of genes implicated in epigenetic mechanisms such as *HDAC6*, *KDM6A*, *TET2* and *TET3* to be differentially expressed in B cells from patients, indicating their dysregulation in the disease. Intersection of expression data with DNA methylation data from study I revealed an overlap of differential expression and differential methylation in 77 genes, including typical IFN signature genes, for example *MX1*, *IFI44L* and *IFITM1*, as well as genes that constitute genetic risk loci for autoimmune diseases, such as *CXCR5*, which has been associated to pSS [16], and *TNFSF4*, associated to both pSS and SLE [136, 137, 152, 154]

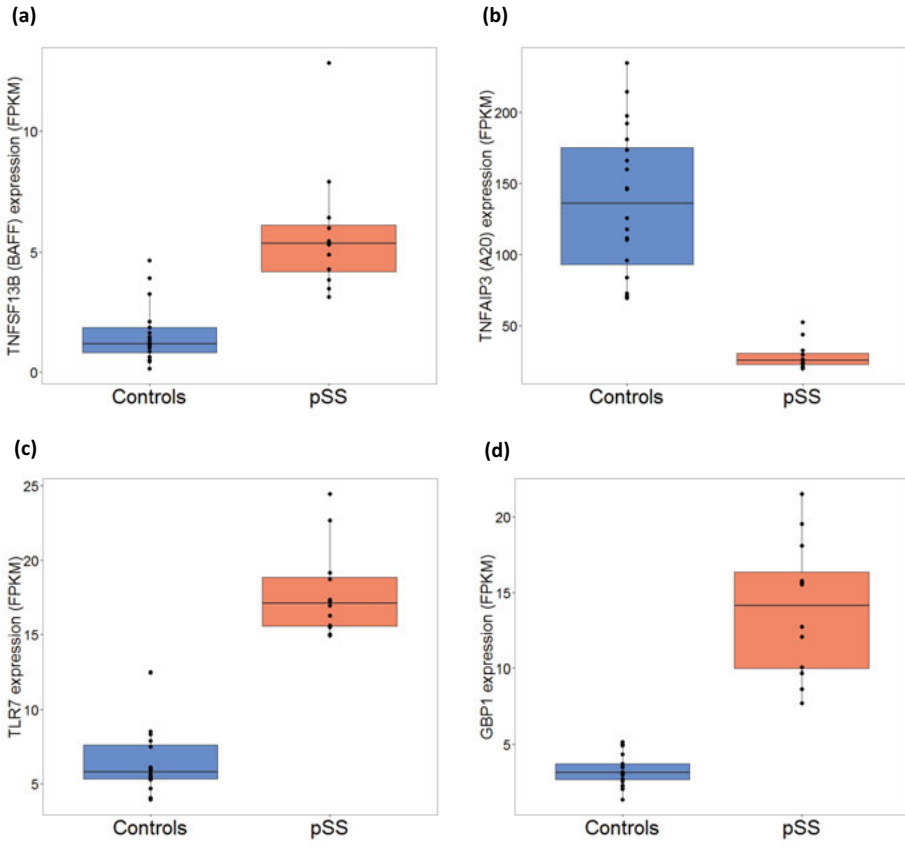

Figure 8. Differential gene expression in CD19+ B cells in pSS. Boxplots illustrating expression levels (in FPKM) of (a) *TNFSF13B* (*BAFF*), (b) *TNFAIP3* (*A20*), (c) *TLR7* and (d) *GBP1* in patients with pSS (n=12, indicated in orange) and controls (n=20, except for *TLR7* n=16, indicated in blue) analyzed by RNA-seq. FPKM, fragments per kilobase of exon per million fragments mapped.

TNFAIP3, encoding A20, showed downregulated expression in pSS B cells (Figure 8b), which is noteworthy given the importance of A20 in lymphomagenesis, where low protein levels of A20 in salivary glands [250] and genetic variants in *TNFAIP3* have been linked to increased risk for lymphoma [251, 252].

We further investigated differential gene expression in CD19+ B cells in a replication cohort (anti-SSA positive female patients with pSS n=17, and female controls n=16) using a NanoString panel (Human Immunology v2), which interrogates expression levels of 594 genes. Differential expression of 30 genes identified by RNA-seq was replicated, among these *TNFSF13B* (*BAFF*), *TNFSF4*, *TLR7* (Figure 8c) and *STAT1/-2*. Interestingly, we also replicated differential expression of genes, which are induced by type II IFN (IFN-γ), such as *GBP1* (Figure 8d) and *GBP5* [253].

Limitations to the study are the relative small sample size and that replication was performed using a different method only targeting a rather small set of predefined genes. Furthermore, our expression analysis was restricted to untreated female anti-SSA positive patients in order to limit the number of potential confounders. Follow-up studies interrogating gene expression in B cells from male patients as well as from autoantibody negative patients and patients undergoing treatment will be necessary in the future to provide a more comprehensive picture on the B cell transcriptome in pSS.

In conclusion, this is the first study investigating B cell-specific gene expression in pSS using RNA-seq. Overexpression of both type I and type II IFN induced genes was demonstrated and replicated in an independent cohort. Evidence for an ongoing immune activation in B cells strengthens the rational for efforts directed against B cells as therapeutic targets in pSS.

Paper IV

DNA methylation mapping identifies gene regulatory effects in patients with systemic lupus erythematosus (SLE)

In this study we investigated the impact of DNA methylation on SLE susceptibility and phenotype variability by analyzing DNA extracted from whole blood samples from patients with SLE and healthy controls using the HM450k array. The discovery cohort comprised 347 patients and 400 controls, while 201 patients and 188 controls were included in the replication cohort.

We accounted for cell type distribution heterogeneity along with age and sex in the case-control epigenome-wide association study (EWAS), where we identified and replicated 7,245 DMCs ($p < 1.3x10^{-7}$ and a delta-beta cut-off > |0.05| in the discovery cohort, and $p < 6.6x10^{-6}$ and same direction of effect in the replication cohort). Effect sizes of the majority of the individual DMCs were relatively moderate, but we observed prominent differential methylation (delta-beta >|0.1|) primarily at IFN regulated genes, which were typically hypomethylated in patients compared to controls. These results reflect observations of gene expression analyses, where a profound overexpression of type I IFN induced genes in SLE is well established [71, 254]. The most significant DMCs in our EWAS replicated results from prior studies, where DNA methylation profiles have been analyzed in fractionated CD4+ T cells and neutrophils from SLE patients and controls [99, 207, 208, 212]. In addition, we were able to identify and replicate many novel DMCs, probably due to the larger size of our cohorts compared to sample sizes included in previous studies.

SLE is a clinically highly heterogeneous disease, where novel biomarkers and more specific and more effective treatment options are demanded to meet requirements for improved patient care. In a case-case set up we ana-

lyzed DNA methylation patterns of SLE patients who fulfilled a specific disease manifestation as defined by the ACR 1982 classification criteria (Table 2) against all other patients without the specific manifestation. In the discovery cohort we observed 49 manifestation-specific DMCs ($p < 1.3x10^{-7}$ and delta-beta > |0.05|). The majority of these DMCs were associated with ANA positivity, a criterion fulfilled by almost all SLE patients. We only detected a single DMC located in the Fas associated via death domain (*FADD*) gene as specific for lupus nephritis. This low number was unexpected given previous reports about renal disorder associated DMCs in SLE [211, 255]. One DMC at the interferon induced transmembrane protein 1 (*IFITM1*) gene was associated with hematological manifestation in SLE. This CpG site was nominally significant in the replication cohort (p=0.001 and same direction and magnitude of effect). Applying Bonferroni correction, none of the 49 manifestation-specific DMCs from the discovery cohort reached significance in the replication cohort. Both the small number of DMCs observed in the discovery phase as well as the lack of replication may be explained by several factors: Disease criteria positivity is collected cumulatively over a patient's disease course. Hence, it does not represent active disease manifestations at the time of blood sampling. Another complicating circumstance is that any patient needs to fulfill at least four out of the eleven ACR criteria for SLE classification. Consequently, many different combinations of criteria in every single patient are possible, potentially confounding the analysis of manifestation-specific DNA methylation patterns. In addition, it is possible that some true positive association signals remained undetected due to the very stringent significance criteria used in our analyses.

Next, we investigated treatment-specific DNA methylation patterns in SLE, and we identified and replicated a total of 5,177 DMCs in the case-case analyses, of which the vast majority was associated with glucocorticoid treatment (n=5,046 DMCs). We further observed 130 DMCs associated with azathioprine treatment and one DMC associated with mycophenolate mofetil, while we were not able to replicate any DMCs associated with methotrexate (MTX) treatment. A previous study in RA patients reported reversion of hypomethylation in patients undergoing MTX treatment [256]. In our study we did not observe any indication for MTX-related reversion of hypomethylation in SLE. In general, associations of DNA methylation with treatment may be highly confounded by the underlying rational for prescribing a certain medication. Without prospective studies analyzing the same patients before and (at different time points) during medical treatment, we cannot conclude whether DNA methylation patterns in treated patients are indicative for their medication or for their underlying disease.

As DNA methylation may mediate genetic risk [257], we set out to investigate *cis*-meQTLs in patients with SLE (n=527) and controls (n=567) using DNA methylation data generated on the HM450k array and genotype data generated on the ImmunoChip. Analyzing SNPs within in ±100kb flanking

region at each CpG site, we identified 141,492 meQTLs in the SLE patients ($p < 6.5x10^{-9}$), of which 78% also were significant in the controls.

Next, we analyzed whether methylation levels at DMCs associated to SLE are under genetic control and observed that 466 CpG sites from the replicated 7,245 EWAS DMCs are meQTLs. This finding indicates that differences in DNA methylation between cases and controls at these sites are not reactive to the disease, but may instead contribute to SLE susceptibility via gene regulatory mechanism. We also investigated potential differential regulation of meQTLs between patients and controls by comparing the slope of the regression line from the meQTL analyses. A total of 5,961 SNP-CpG associations were identified as differential, and 34 of these sites were among our replicated SLE DMCs with multiple CpG sites within the MHC region as well as annotated to the type II IFN inducible genes *DAP* and *CASP*. Investigating specifically meQTLs at genetic risk loci for SLE as defined by GWAS (Table 3), we identified SNP-CpG associations overlapping with the replicated SLE DMCs at *IKZF3*, *IRF5* (Figure 9), *IRF7*, *MHC-class III*, *PTPRC*, *UBE2L3* and *UHRF1BP1*. Some of these meQTLs were specific to patients or controls while others were shared.

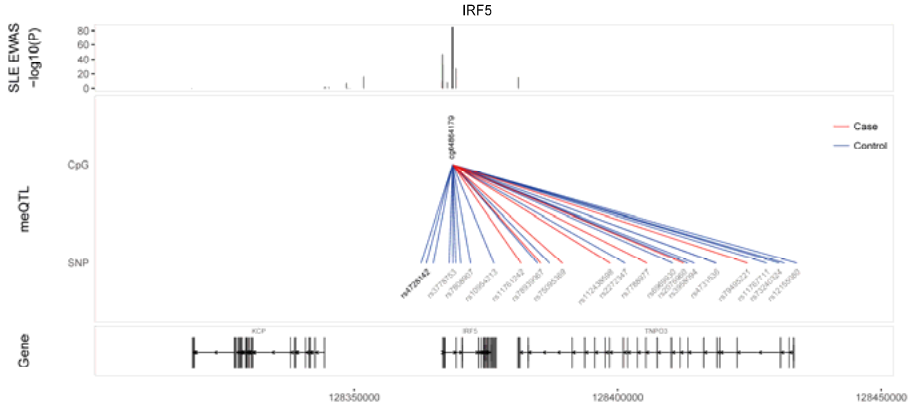

Figure 9. Genetic regulation of DNA methylation in SLE patients and controls at the *IRF5-TNPO3* locus. The top panel illustrates the results (as $-\log_{10}$ p-values) for the SLE case-control association in the discovery cohort (patients with SLE n=347, and control individuals n=400) of the EWAS, with subsequently replicated differentially methylated CpG sites (DMCs) in black. The panel in the middle represents significant methylation quantitative traits loci (meQTLs) in cases (indicated by red lines) and controls (indicated by blue lines). The meQTLs were analysed in both the discovery and the replication cohort, including a total of 524 patients with SLE and 567 controls. The SLE GWAS index SNP (rs4728142) is indicated in black. The bottom panel illustrates the relative genomic location of the RefSeq genes within the locus.

To conclude, our SLE EWAS results replicated previous findings and identified many novel DMCs associated with SLE. The results of the *cis*-meQTL analyses provide evidence for genetic control of DNA methylation in a subset of CpG sites and thus suggest a functional mechanism through DNA methylation for several of the genetic risk variants in SLE.

Concluding remarks

The studies presented in this thesis further support a contributing role for epigenetic mechanisms and dysregulated gene expression in the pathogenesis of systemic inflammatory autoimmune diseases. The most profound differences in DNA methylation and gene expression between cases and controls were observed at type I IFN regulated genes. These results underline the prominent role of the type I IFN system in pSS and SLE. In addition, aberrant patterns of DNA methylation and gene expression were also identified at type II IFN regulated genes. It will be important to further investigate the role of the type II IFN system in the pathophysiology of systemic autoimmune diseases.

While our EWAS in pSS and SLE confirmed association results from previous studies in fractionated blood cells, future studies need to replicate novel findings and to further address the functional implications of the identified variations. For this purpose the importance of clinically well-characterized cohorts comprising longitudinal data is immense, in particular for studies aiming at subgroups of patients with specific disease manifestation or treatment.

DNA methylation patterns are influenced by multiple factors such as cell type distribution, age, sex, treatment at time of sampling et cetera. It is therefore crucial to account for potential confounders in the analysis. In the standard regression models for analysis of EWAS data, cell type fractions can be included as covariates. For the analysis of whole blood or PBMC derived samples reference-based methods for cell type deconvolution are available [228, 229], while to date no reference DNA methylation profiles for other samples containing mixed cell types do exist. The recent development of reference-free methods, which instead are based on surrogate variable analysis (SVA) or on sparse PCA, may constitute an alternative approach when reference DNA methylation data sets are not available [258, 259]. It would be interesting to apply these methods on our salivary gland tissue data. Reference-independent adjustment for cell type heterogeneity may advance the knowledge gained from studies interrogating DNA methylation in specific target tissues of the diseases, such as salivary glands, skin or kidneys. Investigating epigenetic marks in these tissue types is suitable to shed light on the organ-specific disease mechanisms of pSS and SLE, while whole blood samples are probably best reflecting the systemic features of the

diseases. Whole blood, which is more easily available, would also be a preferred sample type for a clinical biomarker.

A comprehensive understanding of the dynamics of the epigenetic landscape in the interplay with genetic and environmental factors will be one of the key requirements for improvement of diagnosis and prediction, identification of novel and more specific therapeutic targets, and maybe even prevention of the development of complex inflammatory autoimmune diseases. The past two decades have been tremendously exciting from a genetics and epigenetics perspective, and exciting times are still ahead. We have only started to decipher the role of epigenetic mechanisms and cell type specific gene expression in autoimmune diseases. Current EWAS aim at inferring function through correlation [260]. The integration of additional types of data such as SNP genotyping, gene expression, histone modifications, multiple QTL analyses and proteomics can provide further indications for functional roles of epigenetic mechanisms. However, to deduce actual causative roles of epigenetic marks on for example gene regulation or pathogenesis remains challenging. Epigenome editing approaches, which apply the rapid developments in genome editing on the analysis of epigenetic marks, may provide definitive functional evidence [261, 262]. In particular the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system has created novel opportunities for high-throughput analyses of the mechanistic relationships between histone and DNA modifications and gene expression [263].

Profiling strategies can identify candidate genes or regions for functional testing and thus build the foundation for more targeted investigations aiming at clarifying the causative functional impact of genetic and epigenetic variation in health and disease. The studies presented in this thesis will hopefully serve as a contribution to the endeavor towards a comprehensive understanding of the functional molecular mechanisms of immunity and autoimmune disease.

Acknowledgements

This thesis was carried out at the group of Molecular Medicine at the Department of Medical Sciences at Uppsala University. I am thankful to the patients and blood donors who participated in the studies. Numerous of people have contributed to this thesis in many different ways, and I want to express my sincere gratitude to all of you for supporting me along the way. Especially I would like to thank:

My supervisors, you were a fantastic trio, each of you with characteristics that made this thesis both possible and enjoyable.
Ann-Christine, thank you for giving me the opportunity to perform my PhD studies in your group. I appreciated the possibility to work within the autoimmunity project immensely. You are an amazing and inspiring person, and your enthusiasm and experience provided a strong fundament for the studies. I definitely learned a lot from you.
Gunnel, you were a great supervisor. I valuate your organization, accuracy and scientific curiosity. You brought the clinical perspective into our studies, and I appreciated to visit the clinic together with you to meet the patients. Thank you also for being such a nice travel-company to conferences and meetings. You always encouraged me when I needed it and you were always available when your help was requested. I am looking forward to working together with you on the new exciting projects.
Johanna Sandling, you are the person one can wish for to have as a supervisor and colleague. It is a pleasure to work with you. Thank you for sharing your profound knowledge with me, for your kindness, your thoroughness and for your consideration and reliability.
Thanks to all the three of you for your support. It was a privilege having you as my supervisors during these years!

All members of the Swedish SLE network and the Scandinavian Sjögren's network, I am very grateful for our good collaboration. In particular, I would like to thank the co-authors on the different studies, Christopher, Elisabet, Iva, Leonid and Marika. Present and past members of the Rheumatology group, Lars, Maija-Leena, Lina, Niklas, Olle, Andrei, Dag, Rezvan, Karin B and Karin H and all the others, thank you for our close, friendly and fruitful collaboration. Katrine and Roald from Stavanger, thank you for working together on the fatigue paper and for the good time together at conferences.

Marie, Albin and Gudny Ella from KI, for the collaboration within the Sjögren's projects, and for your fantastic company at the meeting in Oklahoma.

Everybody at the SNP&SEQ technology platform, thank you for friendly and inspiring atmosphere during lab meetings and fika breaks, and of course, for the excellent data supplied by you. In particular, I like to thank Tomas and Anna for their work with the methylation arrays, and Johanna L, Jessica, Jon and Maria for the sequencing data. Katarina, thank you for all your help with the administration.

The MolMed research group, present and past members, including Amanda, Christofer, Chuan, Eva, Frida, Johan, Mathias, Michelle, Per and Tom, you truly turned BMC into a pleasant place to be! Especially all current and former office mates throughout the years, Behrooz, Ekatarina, Lina S, Niklas, Nour, Sara Ni, Sara Ny and Yanara, thank you for the good time, for all intellectual and all not so intellectual discussions, for your support and kindness. Special thanks to Yanara and Nour for the late night popcorn and pizza dinners at BMC when days were long. Yanara, now you are senior soon; I wish you the very best of luck for the final spurt! Jessica, thank you for fueling my thesis writing with cake, and for sharing your sound knowledge. Mårten, our former senior-PhD student, thanks for all your help, your kindness and calmness; I miss you in Uppsala. Jonas, I am grateful for the good collaboration on the DNA methylation studies, for your patience and for help with R. Anders, your support with the plotting of the meQTLs was invaluable. Special thanks to Sara Ny and Jessica for careful proof-reading.

I also like to thank the members of the former Genetic Medicine group, where I carried out my Master degree project on functional studies of genes associated to SLE, for introducing me into the fantastic world of genetics & immunity. I am also thankful to the present and past members of the Molecular Pathology group for sharing your knowledge with me about cancer biology, biobanking, sequencing of ffpe-samples, and much more.

Friends and family in Sweden, Germany and other corners around the world, my warmest thanks for your support. Especially Johan and your family, for your open-heartedness and generosity, for joined fjällsemester and trips to Öland, for providing me with good food and with even better champagne, and for all the other things that make life much more worth; your friendship means a lot to me. Janne, you are the very best travel-company, special thanks for Svalbard midsummer magic; I am lucky to have met you. Past and present book-circle companions, Johanna M, Cecilia, Angelika, Karolina and Viktoria, for sharing the passion for literature, delicious food and wine, and long discussions about books and everything beyond.

Holger H and Irene, without you this thesis probably would not have been accomplished. I cannot thank you enough for your support, encouragement and for trusting in me.

Holger, thank you for putting up with me and my ideas throughout the years, for your patience, consideration and your unconditional support.

Meine Eltern, Danke für eure Unterstützung. Auch wenn uns 1400 km trennen, so wissen wir doch, dass wir füreinander da sind, wenn wir einander brauchen. Oma Lilo, leider hast du die Fertigstellung dieser Arbeit nicht mehr erlebt. Du hast grossen Anteil an meinem Interesse für Medizin und Chemie; du hast für immer einen speziellen Platz in meinem Herzen.

Matilde and Charlotta, thank you for teaching me what matters in life. For being such wonderful people; I am happy to have you in my life. I have sometimes been a rare guest at our house during the past weeks and months, and I hope we can share more time together now as *the book* is written. Thank you for all the laughers you bring me, for raising questions I cannot not answer, for your curiosity, your warm-heartedness and your crazy ideas. You are the best inspiration.

I would like to leave the last citation to Karl Georg Büchner, poet, revolutionary, medical doctor, and natural scientist, one of my most favorite authors.

Die allgemeinen fixen Ideen, welche
man die gesunde Vernunft tauft,
sind unerträglich langweilig.
(Dantons Tod, 1835 [264])

References

1. Hayter, S.M. and M.C. Cook, *Updated assessment of the prevalence, spectrum and case definition of autoimmune disease.* Autoimmun Rev, 2012. **11**(10): p. 754-65.
2. Cornec, D., et al., *B cells in Sjogren's syndrome: from pathophysiology to diagnosis and treatment.* J Autoimmun, 2012. **39**(3): p. 161-7.
3. Anolik, J.H., *B cell biology: implications for treatment of systemic lupus erythematosus.* Lupus, 2013. **22**(4): p. 342-9.
4. Baecklund, E., et al., *Lymphoma development in patients with autoimmune and inflammatory disorders--what are the driving forces?* Semin Cancer Biol, 2014. **24**: p. 61-70.
5. Bave, U., et al., *Activation of the type I interferon system in primary Sjogren's syndrome: a possible etiopathogenic mechanism.* Arthritis Rheum, 2005. **52**(4): p. 1185-95.
6. Gupta, B. and R.D. Hawkins, *Epigenomics of autoimmune diseases.* Immunol Cell Biol, 2015. **93**(3): p. 271-6.
7. Ronnblom, L., G.V. Alm, and M.L. Eloranta, *The type I interferon system in the development of lupus.* Semin Immunol, 2011. **23**(2): p. 113-21.
8. Qin, B., et al., *Epidemiology of primary Sjogren's syndrome: a systematic review and meta-analysis.* Ann Rheum Dis, 2015. **74**(11): p. 1983-9.
9. Sjögren, H.S., *Zur kenntnis der Keratoconjunctivitis Sicca (Keratitis filiformis bei hypofunktion der Tränendrüsen)*, in *Acta Ophtal (Suppl) (Copenh)*1933. p. 1-151.
10. Jonsson, R., et al., *The complexity of Sjogren's syndrome: novel aspects on pathogenesis.* Immunol Lett, 2011. **141**(1): p. 1-9.
11. Brito-Zeron, P., et al., *Early diagnosis of primary Sjogren's syndrome: EULAR-SS task force clinical recommendations.* Expert Rev Clin Immunol, 2016. **12**(2): p. 137-56.
12. Theander, E., et al., *Lymphoma and other malignancies in primary Sjogren's syndrome: a cohort study on cancer incidence and lymphoma predictors.* Ann Rheum Dis, 2006. **65**(6): p. 796-803.
13. Houghton, K., et al., *Primary Sjogren's syndrome in children and adolescents: are proposed diagnostic criteria applicable?* J Rheumatol, 2005. **32**(11): p. 2225-32.
14. Vitali, C., et al., *Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group.* Ann Rheum Dis, 2002. **61**(6): p. 554-8.

15. Shiboski, C.H., et al., *2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjogren's syndrome A consensus and data-driven methodology involving three international patient cohorts.* Ann Rheum Dis, 2016.
16. Lessard, C.J., et al., *Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjogren's syndrome.* Nat Genet, 2013. **45**(11): p. 1284-92.
17. Li, Y., et al., *A genome-wide association study in Han Chinese identifies a susceptibility locus for primary Sjogren's syndrome at 7q11.23.* Nat Genet, 2013. **45**(11): p. 1361-5.
18. Ice, J.A., et al., *Genetics of Sjogren's syndrome in the genome-wide association era.* J Autoimmun, 2012. **39**(1-2): p. 57-63.
19. Igoe, A. and R.H. Scofield, *Autoimmunity and infection in Sjogren's syndrome.* Curr Opin Rheumatol, 2013. **25**(4): p. 480-7.
20. Brennan, M.T., et al., *Sex steroid hormones in primary Sjogren's syndrome.* J Rheumatol, 2003. **30**(6): p. 1267-71.
21. Harris, V.M., et al., *Klinefelter's syndrome (47,XXY) is in excess among men with Sjogren's syndrome.* Clin Immunol, 2016. **168**: p. 25-9.
22. Liu, K., et al., *X Chromosome Dose and Sex Bias in Autoimmune Diseases: Increased Prevalence of 47,XXX in Systemic Lupus Erythematosus and Sjogren's Syndrome.* Arthritis Rheumatol, 2016. **68**(5): p. 1290-300.
23. Konsta, O.D., et al., *The contribution of epigenetics in Sjogren's Syndrome.* Front Genet, 2014. **5**: p. 71.
24. Jonsson, R., et al., *Sjogren's syndrome--a plethora of clinical and immunological phenotypes with a complex genetic background.* Ann N Y Acad Sci, 2007. **1108**: p. 433-47.
25. Zhang, A.T., et al., *Dynamic interaction of Y RNAs with chromatin and initiation proteins during human DNA replication.* J Cell Sci, 2011. **124**(Pt 12): p. 2058-69.
26. Cornaby, C., et al., *B cell epitope spreading: mechanisms and contribution to autoimmune diseases.* Immunol Lett, 2015. **163**(1): p. 56-68.
27. Quartuccio, L., et al., *Anti-SSA/SSB-negative Sjogren's syndrome shows a lower prevalence of lymphoproliferative manifestations, and a lower risk of lymphoma evolution.* Autoimmun Rev, 2015. **14**(11): p. 1019-22.
28. Baer, A.N., et al., *Association of Anticentromere Antibodies With More Severe Exocrine Glandular Dysfunction in Sjogren's Syndrome: Analysis of the Sjogren's International Collaborative Clinical Alliance Cohort.* Arthritis Care Res (Hoboken), 2016. **68**(10): p. 1554-9.
29. Migliorini, P., et al., *Anti-Sm and anti-RNP antibodies.* Autoimmunity, 2005. **38**(1): p. 47-54.
30. Baer, A.N., et al., *Antibodies to interferon-inducible protein-16 in primary Sjogren's syndrome are associated with markers of more severe diseas.* Arthritis Care Res (Hoboken), 2016.

31. Deng, C., et al., *Meta-analysis of anti-muscarinic receptor type 3 antibodies for the diagnosis of Sjogren syndrome.* PLoS One, 2015. **10**(1): p. e0116744.
32. Hansen, A., et al., *Diminished peripheral blood memory B cells and accumulation of memory B cells in the salivary glands of patients with Sjogren's syndrome.* Arthritis Rheum, 2002. **46**(8): p. 2160-71.
33. Aqrawi, L.A., et al., *Low number of memory B cells in the salivary glands of patients with primary Sjogren's syndrome.* Autoimmunity, 2012. **45**(7): p. 547-55.
34. Furuzawa-Carballeda, J., et al., *Peripheral regulatory cells immunophenotyping in primary Sjogren's syndrome: a cross-sectional study.* Arthritis Res Ther, 2013. **15**(3): p. R68.
35. Mauri, C. and A. Bosma, *Immune regulatory function of B cells.* Annu Rev Immunol, 2012. **30**: p. 221-41.
36. Berthelot, J.M., et al., *Regulatory B cells play a key role in immune system balance.* Joint Bone Spine, 2013. **80**(1): p. 18-22.
37. Nocturne, G., et al., *Use of Biologics in Sjogren's Syndrome.* Rheum Dis Clin North Am, 2016. **42**(3): p. 407-17.
38. Khan, A.R., et al., *PD-L1hi B cells are critical regulators of humoral immunity.* Nat Commun, 2015. **6**: p. 5997.
39. De Vita, S., et al., *Sequential therapy with belimumab followed by rituximab in Sjogren's syndrome associated with B-cell lymphoproliferation and overexpression of BAFF: evidence for long-term efficacy.* Clin Exp Rheumatol, 2014. **32**(4): p. 490-4.
40. Theander, E., R. Manthorpe, and L.T. Jacobsson, *Mortality and causes of death in primary Sjogren's syndrome: a prospective cohort study.* Arthritis Rheum, 2004. **50**(4): p. 1262-9.
41. Ioannidis, J.P., V.A. Vassiliou, and H.M. Moutsopoulos, *Long-term risk of mortality and lymphoproliferative disease and predictive classification of primary Sjogren's syndrome.* Arthritis Rheum, 2002. **46**(3): p. 741-7.
42. Theander, E., et al., *Lymphoid organisation in labial salivary gland biopsies is a possible predictor for the development of malignant lymphoma in primary Sjogren's syndrome.* Ann Rheum Dis, 2011. **70**(8): p. 1363-8.
43. Quartuccio, L., et al., *BLyS upregulation in Sjogren's syndrome associated with lymphoproliferative disorders, higher ESSDAI score and B-cell clonal expansion in the salivary glands.* Rheumatology (Oxford), 2013. **52**(2): p. 276-81.
44. Nocturne, G. and X. Mariette, *Sjogren Syndrome-associated lymphomas: an update on pathogenesis and management.* Br J Haematol, 2015. **168**(3): p. 317-27.
45. Jonsson, M.V., et al., *Association between circulating levels of the novel TNF family members APRIL and BAFF and lymphoid organization in primary Sjogren's syndrome.* J Clin Immunol, 2005. **25**(3): p. 189-201.

46. Sjostrand, M., et al., *The Expression of BAFF Is Controlled by IRF Transcription Factors.* J Immunol, 2016. **196**(1): p. 91-6.
47. Nezos, A., et al., *B-cell activating factor genetic variants in lymphomagenesis associated with primary Sjogren's syndrome.* J Autoimmun, 2014. **51**: p. 89-98.
48. Harboe, E., et al., *Fatigue in primary Sjogren's syndrome--a link to sickness behaviour in animals?* Brain Behav Immun, 2009. **23**(8): p. 1104-8.
49. Meijer, J.M., et al., *Health-related quality of life, employment and disability in patients with Sjogren's syndrome.* Rheumatology (Oxford), 2009. **48**(9): p. 1077-82.
50. Bower, J.E., *Cancer-related fatigue--mechanisms, risk factors, and treatments.* Nat Rev Clin Oncol, 2014. **11**(10): p. 597-609.
51. Becker, K., et al., *Poststroke fatigue: hints to a biological mechanism.* J Stroke Cerebrovasc Dis, 2015. **24**(3): p. 618-21.
52. Norheim, K.B., et al., *A possible genetic association with chronic fatigue in primary Sjogren's syndrome: a candidate gene study.* Rheumatol Int, 2014. **34**(2): p. 191-7.
53. Hartkamp, A., et al., *Serum cytokine levels related to multiple dimensions of fatigue in patients with primary Sjogren's syndrome.* Ann Rheum Dis, 2004. **63**(10): p. 1335-7.
54. Bardsen, K., et al., *Heat shock proteins and chronic fatigue in primary Sjogren's syndrome.* Innate Immun, 2016. **22**(3): p. 162-7.
55. Norheim, K.B., G. Jonsson, and R. Omdal, *Biological mechanisms of chronic fatigue.* Rheumatology (Oxford), 2011. **50**(6): p. 1009-18.
56. Whitehead, L., *The measurement of fatigue in chronic illness: a systematic review of unidimensional and multidimensional fatigue measures.* J Pain Symptom Manage, 2009. **37**(1): p. 107-28.
57. Bowman, S.J., et al., *Patient-reported outcomes in primary Sjogren's syndrome: comparison of the long and short versions of the Profile of Fatigue and Discomfort--Sicca Symptoms Inventory.* Rheumatology (Oxford), 2009. **48**(2): p. 140-3.
58. Seror, R., et al., *Outcome measures for primary Sjogren's syndrome: a comprehensive review.* J Autoimmun, 2014. **51**: p. 51-6.
59. Ramos-Casals, M., et al., *Treatment of primary Sjogren syndrome: a systematic review.* JAMA, 2010. **304**(4): p. 452-60.
60. Gottenberg, J.E., et al., *Effects of hydroxychloroquine on symptomatic improvement in primary Sjogren syndrome: the JOQUER randomized clinical trial.* JAMA, 2014. **312**(3): p. 249-58.
61. Brito-Zeron, P., M. Ramos-Casals, and E.-S.t.f. group, *Advances in the understanding and treatment of systemic complications in Sjogren's syndrome.* Curr Opin Rheumatol, 2014. **26**(5): p. 520-7.
62. Bengtsson, A.A. and L. Ronnblom, *Systemic lupus erythematosus: still a challenge for physicians.* J Intern Med, 2016.
63. Munoz, L.E., et al., *Apoptosis in the pathogenesis of systemic lupus erythematosus.* Lupus, 2008. **17**(5): p. 371-5.

64. Tan, E.M., et al., *The 1982 revised criteria for the classification of systemic lupus erythematosus.* Arthritis Rheum, 1982. **25**(11): p. 1271-7.
65. Ighe, A., et al., *Application of the 2012 Systemic Lupus International Collaborating Clinics classification criteria to patients in a regional Swedish systemic lupus erythematosus register.* Arthritis Res Ther, 2015. **17**: p. 3.
66. Pons-Estel, G.J., et al., *Understanding the epidemiology and progression of systemic lupus erythematosus.* Semin Arthritis Rheum, 2010. **39**(4): p. 257-68.
67. Borchers, A.T., et al., *The geoepidemiology of systemic lupus erythematosus.* Autoimmun Rev, 2010. **9**(5): p. A277-87.
68. Mak, A., et al., *Global trend of survival and damage of systemic lupus erythematosus: meta-analysis and meta-regression of observational studies from the 1950s to 2000s.* Semin Arthritis Rheum, 2012. **41**(6): p. 830-9.
69. Bjornadal, L., et al., *Cardiovascular disease a hazard despite improved prognosis in patients with systemic lupus erythematosus: results from a Swedish population based study 1964-95.* J Rheumatol, 2004. **31**(4): p. 713-9.
70. Mok, C.C., R.C. Kwok, and P.S. Yip, *Effect of renal disease on the standardized mortality ratio and life expectancy of patients with systemic lupus erythematosus.* Arthritis Rheum, 2013. **65**(8): p. 2154-60.
71. Baechler, E.C., et al., *Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus.* Proc Natl Acad Sci U S A, 2003. **100**(5): p. 2610-5.
72. Rullo, O.J. and B.P. Tsao, *Recent insights into the genetic basis of systemic lupus erythematosus.* Ann Rheum Dis, 2013. **72 Suppl 2**: p. ii56-61.
73. Hedrich, C.M. and G.C. Tsokos, *Epigenetic mechanisms in systemic lupus erythematosus and other autoimmune diseases.* Trends Mol Med, 2011. **17**(12): p. 714-24.
74. Bijl, M. and C.G. Kallenberg, *Ultraviolet light and cutaneous lupus.* Lupus, 2006. **15**(11): p. 724-7.
75. McClain, M.T., et al., *Early events in lupus humoral autoimmunity suggest initiation through molecular mimicry.* Nat Med, 2005. **11**(1): p. 85-9.
76. Isaacs, A. and J. Lindenmann, *Virus interference. I. The interferon.* Proc R Soc Lond B Biol Sci, 1957. **147**(927): p. 258-67.
77. Hagberg, N. and L. Ronnblom, *Systemic Lupus Erythematosus--A Disease with A Dysregulated Type I Interferon System.* Scand J Immunol, 2015. **82**(3): p. 199-207.
78. Ronnblom, L. and G.V. Alm, *The natural interferon-alpha producing cells in systemic lupus erythematosus.* Hum Immunol, 2002. **63**(12): p. 1181-93.

79. Tanji, H., et al., *Toll-like receptor 8 senses degradation products of single-stranded RNA.* Nat Struct Mol Biol, 2015. **22**(2): p. 109-15.
80. Ohto, U., et al., *Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9.* Nature, 2015. **520**(7549): p. 702-5.
81. Eloranta, M.L. and L. Ronnblom, *Cause and consequences of the activated type I interferon system in SLE.* J Mol Med (Berl), 2016. **94**(10): p. 1103-1110.
82. Yoshimura, A., T. Naka, and M. Kubo, *SOCS proteins, cytokine signalling and immune regulation.* Nat Rev Immunol, 2007. **7**(6): p. 454-65.
83. Hooks, J.J., et al., *Immune interferon in the circulation of patients with autoimmune disease.* N Engl J Med, 1979. **301**(1): p. 5-8.
84. Ronnblom, L.E., G.V. Alm, and K.E. Oberg, *Possible induction of systemic lupus erythematosus by interferon-alpha treatment in a patient with a malignant carcinoid tumour.* J Intern Med, 1990. **227**(3): p. 207-10.
85. Gottenberg, J.E., et al., *Activation of IFN pathways and plasmacytoid dendritic cell recruitment in target organs of primary Sjogren's syndrome.* Proc Natl Acad Sci U S A, 2006. **103**(8): p. 2770-5.
86. Bengtsson, A.A., et al., *Activation of type I interferon system in systemic lupus erythematosus correlates with disease activity but not with antiretroviral antibodies.* Lupus, 2000. **9**(9): p. 664-71.
87. Brkic, Z., et al., *Prevalence of interferon type I signature in CD14 monocytes of patients with Sjogren's syndrome and association with disease activity and BAFF gene expression.* Ann Rheum Dis, 2013. **72**(5): p. 728-35.
88. Qiu, L.J., et al., *Decreased SOCS1 mRNA expression levels in peripheral blood mononuclear cells from patients with systemic lupus erythematosus in a Chinese population.* Clin Exp Med, 2015. **15**(3): p. 261-7.
89. Vallin, H., et al., *Anti-double-stranded DNA antibodies and immunostimulatory plasmid DNA in combination mimic the endogenous IFN-alpha inducer in systemic lupus erythematosus.* J Immunol, 1999. **163**(11): p. 6306-13.
90. Roers, A., B. Hiller, and V. Hornung, *Recognition of Endogenous Nucleic Acids by the Innate Immune System.* Immunity, 2016. **44**(4): p. 739-54.
91. Lovgren, T., et al., *Induction of interferon-alpha production in plasmacytoid dendritic cells by immune complexes containing nucleic acid released by necrotic or late apoptotic cells and lupus IgG.* Arthritis Rheum, 2004. **50**(6): p. 1861-72.
92. Nimmerjahn, F. and J.V. Ravetch, *Fcgamma receptors as regulators of immune responses.* Nat Rev Immunol, 2008. **8**(1): p. 34-47.
93. Kuznik, A., et al., *Mechanism of endosomal TLR inhibition by antimalarial drugs and imidazoquinolines.* J Immunol, 2011. **186**(8): p. 4794-804.

94. Khamashta, M., et al., *Sifalimumab, an anti-interferon-alpha monoclonal antibody, in moderate to severe systemic lupus erythematosus: a randomised, double-blind, placebo-controlled study.* Ann Rheum Dis, 2016. **75**(11): p. 1909-1916.
95. Furie, R., et al., *Anifrolumab, an Anti-Interferon-Alpha Receptor Monoclonal Antibody, in Moderate to Severe Systemic Lupus Erythematosus.* Arthritis Rheum, 2016.
96. Nezos, A., et al., *Type I and II interferon signatures in Sjogren's syndrome pathogenesis: Contributions in distinct clinical phenotypes and Sjogren's related lymphomagenesis.* J Autoimmun, 2015. **63**: p. 47-58.
97. Munroe, M.E., et al., *Altered type II interferon precedes autoantibody accrual and elevated type I interferon activity prior to systemic lupus erythematosus classification.* Ann Rheum Dis, 2016. **75**(11): p. 2014-2021.
98. Apostolou, E., et al., *Expression of type III interferons (IFNlambdas) and their receptor in Sjogren's syndrome.* Clin Exp Immunol, 2016. **186**(3): p. 304-312.
99. Absher, D.M., et al., *Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations.* PLoS Genet, 2013. **9**(8): p. e1003678.
100. Altorok, N., et al., *Genome-wide DNA methylation patterns in naive CD4+ T cells from patients with primary Sjogren's syndrome.* Arthritis Rheumatol, 2014. **66**(3): p. 731-9.
101. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
102. Venter, J.C., et al., *The sequence of the human genome.* Science, 2001. **291**(5507): p. 1304-51.
103. 1000 Genomes Project Consortium, et al., *A global reference for human genetic variation.* Nature, 2015. **526**(7571): p. 68-74.
104. International HapMap Consortium, *A haplotype map of the human genome.* Nature, 2005. **437**(7063): p. 1299-320.
105. International HapMap Consortium, et al., *A second generation human haplotype map of over 3.1 million SNPs.* Nature, 2007. **449**(7164): p. 851-61.
106. Ng, P.C., et al., *Genetic variation in an individual human exome.* PLoS Genet, 2008. **4**(8): p. e1000160.
107. Klein, R.J., et al., *Complement factor H polymorphism in age-related macular degeneration.* Science, 2005. **308**(5720): p. 385-9.
108. *GWAS catalogue. The NHGRI-EBI Catalog of published genome-wide association studies.* Access date: 2016-12-22]; Available from: http://ebi.ac.uk/gwas.
109. Pritchard, J.K. and N.J. Cox, *The allelic architecture of human disease genes: common disease-common variant...or not?* Hum Mol Genet, 2002. **11**(20): p. 2417-23.

110. Manolio, T.A., et al., *Finding the missing heritability of complex diseases.* Nature, 2009. **461**(7265): p. 747-53.
111. Zuk, O., et al., *Searching for missing heritability: designing rare variant association studies.* Proc Natl Acad Sci U S A, 2014. **111**(4): p. E455-64.
112. Cirulli, E.T. and D.B. Goldstein, *Uncovering the roles of rare variants in common disease through whole-genome sequencing.* Nat Rev Genet, 2010. **11**(6): p. 415-25.
113. International Human Genome Sequencing Consortium, *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-45.
114. Roadmap Epigenomics Consortium, et al., *Integrative analysis of 111 reference human epigenomes.* Nature, 2015. **518**(7539): p. 317-30.
115. Encode Project Consortium, et al., *Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.* Nature, 2007. **447**(7146): p. 799-816.
116. Colwell, J., *Expanding the Scope of ENCODE.* Cancer Discov, 2016. **6**(4): p. OF4.
117. Deapen, D., et al., *A revised estimate of twin concordance in systemic lupus erythematosus.* Arthritis Rheum, 1992. **35**(3): p. 311-8.
118. Giles, I. and D. Isenberg, *Lupus in the family--analysis of a cohort followed from 1978 to 1999.* Lupus, 2001. **10**(1): p. 38-44.
119. Alarcon-Segovia, D., et al., *Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort.* Arthritis Rheum, 2005. **52**(4): p. 1138-47.
120. Kuo, C.F., et al., *Familial Risk of Sjogren's Syndrome and Co-aggregation of Autoimmune Diseases in Affected Families: A Nationwide Population Study.* Arthritis Rheumatol, 2015. **67**(7): p. 1904-12.
121. Truedsson, L., A.A. Bengtsson, and G. Sturfelt, *Complement deficiencies and systemic lupus erythematosus.* Autoimmunity, 2007. **40**(8): p. 560-6.
122. Lee-Kirsch, M.A., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus.* Nat Genet, 2007. **39**(9): p. 1065-7.
123. Yasutomo, K., et al., *Mutation of DNASE1 in people with systemic lupus erythematosus.* Nat Genet, 2001. **28**(4): p. 313-4.
124. Cervino, A.C., N.F. Tsinoremas, and R.W. Hoffman, *A genome-wide study of lupus: preliminary analysis and data release.* Ann N Y Acad Sci, 2007. **1110**: p. 131-9.
125. Morris, D.L., et al., *Unraveling multiple MHC gene associations with systemic lupus erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans.* Am J Hum Genet, 2012. **91**(5): p. 778-93.

126. Sigurdsson, S., et al., *Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus.* Am J Hum Genet, 2005. **76**(3): p. 528-37.
127. Sigurdsson, S., et al., *Comprehensive evaluation of the genetic variants of interferon regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus.* Hum Mol Genet, 2008. **17**(6): p. 872-81.
128. Nordmark, G., et al., *Additive effects of the major risk alleles of IRF5 and STAT4 in primary Sjogren's syndrome.* Genes Immun, 2009. **10**(1): p. 68-76.
129. Feng, D., et al., *Genetic variants and disease-associated factors contribute to enhanced interferon regulatory factor 5 expression in blood cells of patients with systemic lupus erythematosus.* Arthritis Rheum, 2010. **62**(2): p. 562-73.
130. Berggren, O., et al., *IFN-alpha production by plasmacytoid dendritic cell associations with polymorphisms in gene loci related to autoimmune and inflammatory diseases.* Hum Mol Genet, 2015. **24**(12): p. 3571-81.
131. Bentham, J., et al., *Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus.* Nat Genet, 2015. **47**(12): p. 1457-64.
132. Cunninghame Graham, D.S., et al., *Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus.* PLoS Genet, 2011. **7**(10): p. e1002341.
133. Molineros, J.E., et al., *Admixture mapping in lupus identifies multiple functional variants within IFIH1 associated with apoptosis, inflammation, and autoantibody production.* PLoS Genet, 2013. **9**(2): p. e1003222.
134. International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN), et al., *Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci.* Nat Genet, 2008. **40**(2): p. 204-10.
135. Hom, G., et al., *Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX.* N Engl J Med, 2008. **358**(9): p. 900-9.
136. Alarcon-Riquelme, M.E., et al., *Genome-Wide Association Study in an Amerindian Ancestry Population Reveals Novel Systemic Lupus Erythematosus Risk Loci and the Role of European Admixture.* Arthritis Rheumatol, 2016. **68**(4): p. 932-43.
137. Yang, W., et al., *Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus.* PLoS Genet, 2010. **6**(2): p. e1000841.
138. Lofgren, S.E., et al., *Genetic association of miRNA-146a with systemic lupus erythematosus in Europeans through decreased expression of the gene.* Genes Immun, 2012. **13**(3): p. 268-74.

139. Luo, X., et al., *A functional variant in microRNA-146a promoter modulates its expression and confers disease risk for systemic lupus erythematosus.* PLoS Genet, 2011. **7**(6): p. e1002128.
140. Gateva, V., et al., *A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus.* Nat Genet, 2009. **41**(11): p. 1228-33.
141. Lessard, C.J., et al., *Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study.* Am J Hum Genet, 2012. **90**(4): p. 648-60.
142. Deng, Y., et al., *MicroRNA-3148 modulates allelic expression of toll-like receptor 7 variant associated with systemic lupus erythematosus.* PLoS Genet, 2013. **9**(2): p. e1003336.
143. Nordmark, G., et al., *Association of genes in the NF-kappaB pathway with antibody-positive primary Sjogren's syndrome.* Scand J Immunol, 2013. **78**(5): p. 447-54.
144. Han, J.W., et al., *Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus.* Nat Genet, 2009. **41**(11): p. 1234-7.
145. Sheng, Y.J., et al., *Follow-up study identifies two novel susceptibility loci PRKCB and 8p11.21 for systemic lupus erythematosus.* Rheumatology (Oxford), 2011. **50**(4): p. 682-8.
146. Lessard, C.J., et al., *Identification of a Systemic Lupus Erythematosus Risk Locus Spanning ATG16L2, FCHSD2, and P2RY2 in Koreans.* Arthritis Rheumatol, 2016. **68**(5): p. 1197-209.
147. Sawalha, A.H., et al., *Common variants within MECP2 confer risk of systemic lupus erythematosus.* PLoS One, 2008. **3**(3): p. e1727.
148. Zhang, Y., et al., *Meta-analysis of GWAS on two Chinese populations followed by replication identifies novel genetic variants on the X chromosome associated with systemic lupus erythematosus.* Hum Mol Genet, 2015. **24**(1): p. 274-84.
149. Nath, S.K., et al., *A nonsynonymous functional variant in integrin-alpha(M) (encoded by ITGAM) is associated with systemic lupus erythematosus.* Nat Genet, 2008. **40**(2): p. 152-4.
150. Kim, K., et al., *Variation in the ICAM1-ICAM4-ICAM5 locus is associated with systemic lupus erythematosus susceptibility in multiple ancestries.* Ann Rheum Dis, 2012. **71**(11): p. 1809-14.
151. Kozyrev, S.V., et al., *Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus.* Nat Genet, 2008. **40**(2): p. 211-6.
152. Chang, Y.K., et al., *Association of BANK1 and TNFSF4 with systemic lupus erythematosus in Hong Kong Chinese.* Genes Immun, 2009. **10**(5): p. 414-20.
153. Fernando, M.M., et al., *Transancestral mapping of the MHC region in systemic lupus erythematosus identifies new independent and interacting loci at MSH5, HLA-DPB1 and HLA-G.* Ann Rheum Dis, 2012. **71**(5): p. 777-84.

154. Nordmark, G., et al., *Association of EBF1, FAM167A(C8orf13)-BLK and TNFSF4 gene variants with primary Sjogren's syndrome.* Genes Immun, 2011. **12**(2): p. 100-9.
155. Delgado-Vega, A.M., et al., *Fine mapping and conditional analysis identify a new mutation in the autoimmunity susceptibility gene BLK that leads to reduced half-life of the BLK protein.* Ann Rheum Dis, 2012. **71**(7): p. 1219-26.
156. Yang, J., et al., *ELF1 is associated with systemic lupus erythematosus in Asian populations.* Hum Mol Genet, 2011. **20**(3): p. 601-7.
157. Kyogoku, C., et al., *Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE.* Am J Hum Genet, 2004. **75**(3): p. 504-7.
158. Yang, W., et al., *Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians.* Am J Hum Genet, 2013. **92**(1): p. 41-51.
159. Lessard, C.J., et al., *Identification of a systemic lupus erythematosus susceptibility locus at 11p13 between PDHX and CD44 in a multiethnic study.* Am J Hum Genet, 2011. **88**(1): p. 83-91.
160. Okada, Y., et al., *A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus eyrthematosus in Japanese.* PLoS Genet, 2012. **8**(1): p. e1002455.
161. Manjarrez-Orduno, N., et al., *CSK regulatory polymorphism is associated with systemic lupus erythematosus and influences B-cell signaling and activation.* Nat Genet, 2012. **44**(11): p. 1227-30.
162. Deng, Y., et al., *Decreased SMG7 expression associates with lupus-risk variants and elevated antinuclear antibody production.* Ann Rheum Dis, 2016. **75**(11): p. 2007-2013.
163. Teruel, M. and M.E. Alarcon-Riquelme, *Genetics of systemic lupus erythematosus and Sjogren's syndrome: an update.* Curr Opin Rheumatol, 2016.
164. Nezos, A. and C.P. Mavragani, *Contribution of Genetic Factors to Sjogren's Syndrome and Sjogren's Syndrome Related Lymphomagenesis.* J Immunol Res, 2015. **2015**: p. 754825.
165. Reksten, T.R., et al., *Genetic associations to germinal centre formation in primary Sjogren's syndrome.* Ann Rheum Dis, 2014. **73**(6): p. 1253-8.
166. Sigurdsson, S., et al., *A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5.* Hum Mol Genet, 2008. **17**(18): p. 2868-76.
167. Bolin, K., et al., *Association of STAT4 polymorphism with severe renal insufficiency in lupus nephritis.* PLoS One, 2013. **8**(12): p. e84450.
168. Taylor, K.E., et al., *Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus.* PLoS Genet, 2008. **4**(5): p. e1000084.

169. Svenungsson, E., et al., *A STAT4 risk allele is associated with ischaemic cerebrovascular events and anti-phospholipid antibodies in systemic lupus erythematosus.* Ann Rheum Dis, 2010. **69**(5): p. 834-40.
170. Leonard, D., et al., *Coronary heart disease in systemic lupus erythematosus is associated with interferon regulatory factor-8 gene variants.* Circ Cardiovasc Genet, 2013. **6**(3): p. 255-63.
171. Lee, T.I. and R.A. Young, *Transcriptional regulation and its misregulation in disease.* Cell, 2013. **152**(6): p. 1237-51.
172. Holoch, D. and D. Moazed, *RNA-mediated epigenetic regulation of gene expression.* Nat Rev Genet, 2015. **16**(2): p. 71-84.
173. Bartel, D.P., *MicroRNAs: target recognition and regulatory functions.* Cell, 2009. **136**(2): p. 215-33.
174. Qin, H., et al., *MicroRNA-29b contributes to DNA hypomethylation of CD4+ T cells in systemic lupus erythematosus by indirectly targeting DNA methyltransferase 1.* J Dermatol Sci, 2013. **69**(1): p. 61-7.
175. Berger, S.L., et al., *An operational definition of epigenetics.* Genes Dev, 2009. **23**(7): p. 781-3.
176. Heard, E. and R.A. Martienssen, *Transgenerational epigenetic inheritance: myths and mechanisms.* Cell, 2014. **157**(1): p. 95-109.
177. Reik, W., *Stability and flexibility of epigenetic gene regulation in mammalian development.* Nature, 2007. **447**(7143): p. 425-32.
178. Barber, B.A. and M. Rastegar, *Epigenetic control of Hox genes during neurogenesis, development, and disease.* Ann Anat, 2010. **192**(5): p. 261-74.
179. Ziller, M.J., et al., *Charting a dynamic DNA methylation landscape of the human genome.* Nature, 2013. **500**(7463): p. 477-81.
180. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.
181. Schubeler, D., *Function and information content of DNA methylation.* Nature, 2015. **517**(7534): p. 321-6.
182. Denis, H., M.N. Ndlovu, and F. Fuks, *Regulation of mammalian DNA methyltransferases: a route to new mechanisms.* EMBO Rep, 2011. **12**(7): p. 647-56.
183. Kohli, R.M. and Y. Zhang, *TET enzymes, TDG and the dynamics of DNA demethylation.* Nature, 2013. **502**(7472): p. 472-9.
184. Globisch, D., et al., *Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates.* PLoS One, 2010. **5**(12): p. e15367.
185. Christman, J.K., *5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy.* Oncogene, 2002. **21**(35): p. 5483-95.
186. Lisanti, S., et al., *Comparison of methods for quantification of global DNA methylation in human cells and tissues.* PLoS One, 2013. **8**(11): p. e79044.
187. Tost, J. and I.G. Gut, *DNA methylation analysis by pyrosequencing.* Nat Protoc, 2007. **2**(9): p. 2265-75.

188. Bock, C., *Analysing and interpreting DNA methylation data.* Nat Rev Genet, 2012. **13**(10): p. 705-19.
189. Frommer, M., et al., *A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.* Proc Natl Acad Sci U S A, 1992. **89**(5): p. 1827-31.
190. Ehrich, M., et al., *Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry.* Proc Natl Acad Sci U S A, 2005. **102**(44): p. 15785-90.
191. Bibikova, M., et al., *High density DNA methylation array with single CpG site resolution.* Genomics, 2011. **98**(4): p. 288-95.
192. Gunderson, K.L., et al., *A genome-wide scalable SNP genotyping assay using microarray technology.* Nat Genet, 2005. **37**(5): p. 549-54.
193. Andersson, R., et al., *An atlas of active enhancers across human cell types and tissues.* Nature, 2014. **507**(7493): p. 455-61.
194. Dozmorov, M.G., J.D. Wren, and M.E. Alarcon-Riquelme, *Epigenomic elements enriched in the promoters of autoimmunity susceptibility genes.* Epigenetics, 2014. **9**(2): p. 276-85.
195. Moran, S., C. Arribas, and M. Esteller, *Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences.* Epigenomics, 2016. **8**(3): p. 389-99.
196. Yu, X., et al., *DNA hypermethylation leads to lower FOXP3 expression in CD4+ T cells of patients with primary Sjogren's syndrome.* Clin Immunol, 2013. **148**(2): p. 254-7.
197. Yin, H., et al., *Hypomethylation and overexpression of CD70 (TNFSF7) in CD4+ T cells of patients with primary Sjogren's syndrome.* J Dermatol Sci, 2010. **59**(3): p. 198-203.
198. Miceli-Richard, C., et al., *Overlap between differentially methylated DNA regions in blood B lymphocytes and genetic at-risk loci in primary Sjogren's syndrome.* Ann Rheum Dis, 2016. **75**(5): p. 933-40.
199. Rubin, R.L., *Drug-induced lupus.* Expert Opin Drug Saf, 2015. **14**(3): p. 361-78.
200. Richardson, B., et al., *Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis.* Arthritis Rheum, 1990. **33**(11): p. 1665-73.
201. Lu, Q., et al., *Demethylation of ITGAL (CD11a) regulatory sequences in systemic lupus erythematosus.* Arthritis Rheum, 2002. **46**(5): p. 1282-91.
202. Kaplan, M.J., et al., *Demethylation of promoter regulatory elements contributes to perforin overexpression in CD4+ lupus T cells.* J Immunol, 2004. **172**(6): p. 3652-61.
203. Oelke, K., et al., *Overexpression of CD70 and overstimulation of IgG synthesis by lupus T cells and T cells treated with DNA methylation inhibitors.* Arthritis Rheum, 2004. **50**(6): p. 1850-60.
204. Pan, W., et al., *MicroRNA-21 and microRNA-148a contribute to DNA hypomethylation in lupus CD4+ T cells by directly and indirectly targeting DNA methyltransferase 1.* J Immunol, 2010. **184**(12): p. 6773-81.

205. Zhao, S., et al., *MicroRNA-126 regulates DNA methylation in CD4+ T cells and contributes to systemic lupus erythematosus by targeting DNA methyltransferase 1.* Arthritis Rheum, 2011. **63**(5): p. 1376-86.
206. Lu, Q., et al., *Demethylation of CD40LG on the inactive X in T cells from women with lupus.* J Immunol, 2007. **179**(9): p. 6352-8.
207. Coit, P., et al., *Genome-wide DNA methylation study suggests epigenetic accessibility and transcriptional poising of interferon-regulated genes in naive CD4+ T cells from lupus patients.* J Autoimmun, 2013. **43**: p. 78-84.
208. Coit, P., et al., *Epigenome profiling reveals significant DNA demethylation of interferon signature genes in lupus neutrophils.* J Autoimmun, 2015. **58**: p. 59-66.
209. Javierre, B.M., et al., *Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus.* Genome Res, 2010. **20**(2): p. 170-9.
210. Chung, S.A., et al., *Genome-Wide Assessment of Differential DNA Methylation Associated with Autoantibody Production in Systemic Lupus Erythematosus.* PLoS One, 2015. **10**(7): p. e0129813.
211. Coit, P., et al., *Renal involvement in lupus is characterized by unique DNA methylation changes in naive CD4+ T cells.* J Autoimmun, 2015. **61**: p. 29-35.
212. Zhao, M., et al., *DNA methylation and mRNA and microRNA expression of SLE CD4+ T cells correlate with disease phenotype.* J Autoimmun, 2014. **54**: p. 127-36.
213. Renauer, P., et al., *DNA methylation patterns in naive CD4+ T cells identify epigenetic susceptibility loci for malar rash and discoid rash in systemic lupus erythematosus.* Lupus Sci Med, 2015. **2**(1): p. e000101.
214. Tessarz, P. and T. Kouzarides, *Histone core modifications regulating nucleosome structure and dynamics.* Nat Rev Mol Cell Biol, 2014. **15**(11): p. 703-8.
215. Smolle, M. and J.L. Workman, *Transcription-associated histone modifications and cryptic transcription.* Biochim Biophys Acta, 2013. **1829**(1): p. 84-97.
216. Choudhary, C., et al., *The growing landscape of lysine acetylation links metabolism and cell signalling.* Nat Rev Mol Cell Biol, 2014. **15**(8): p. 536-50.
217. Landt, S.G., et al., *ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia.* Genome Res, 2012. **22**(9): p. 1813-31.
218. Zhang, Z., et al., *Global H4 acetylation analysis by ChIP-chip in systemic lupus erythematosus monocytes.* Genes Immun, 2010. **11**(2): p. 124-33.
219. Hu, N., et al., *Abnormal histone modification patterns in lupus CD4+ T cells.* J Rheumatol, 2008. **35**(5): p. 804-10.
220. Reilly, C.M., N. Regna, and N. Mishra, *HDAC inhibition in lupus models.* Mol Med, 2011. **17**(5-6): p. 417-25.

221. Nishida, K., et al., *Histone deacetylase inhibitor suppression of autoantibody-mediated arthritis in mice via regulation of p16INK4a and p21(WAF1/Cip1) expression.* Arthritis Rheum, 2004. **50**(10): p. 3365-76.
222. West, A.C. and R.W. Johnstone, *New and emerging HDAC inhibitors for cancer treatment.* J Clin Invest, 2014. **124**(1): p. 30-9.
223. Norheim, K.B., et al., *Oxidative stress, as measured by protein oxidation, is increased in primary Sjogren's syndrome.* Free Radic Res, 2012. **46**(2): p. 141-6.
224. R Core Development Team, *R: A Language and Environment for Statistical Computing*, in *Vienna, Austria: R Foundation for Statistical Computing*2011.
225. Aryee, M.J., et al., *Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays.* Bioinformatics, 2014. **30**(10): p. 1363-9.
226. Maksimovic, J., L. Gordon, and A. Oshlack, *SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips.* Genome Biol, 2012. **13**(6): p. R44.
227. 1000 Genomes Project Consortium, et al., *An integrated map of genetic variation from 1,092 human genomes.* Nature, 2012. **491**(7422): p. 56-65.
228. Houseman, E.A., et al., *DNA methylation arrays as surrogate measures of cell mixture distribution.* BMC Bioinformatics, 2012. **13**: p. 86.
229. Reinius, L.E., et al., *Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility.* PLoS One, 2012. **7**(7): p. e41361.
230. Rusinova, I., et al., *Interferome v2.0: an updated database of annotated interferon-regulated genes.* Nucleic Acids Res, 2013. **41**(Database issue): p. D1040-6.
231. Grundberg, E., et al., *Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements.* Am J Hum Genet, 2013. **93**(5): p. 876-90.
232. Bernstein, B.E., et al., *The NIH Roadmap Epigenomics Mapping Consortium.* Nat Biotechnol, 2010. **28**(10): p. 1045-8.
233. Cortes, A. and M.A. Brown, *Promise and pitfalls of the Immunochip.* Arthritis Res Ther, 2011. **13**(1): p. 101.
234. DeLuca, D.S., et al., *RNA-SeQC: RNA-seq metrics for quality control and process optimization.* Bioinformatics, 2012. **28**(11): p. 1530-2.
235. Kim, D., et al., *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.* Genome Biol, 2013. **14**(4): p. R36.
236. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.

237. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* Am J Hum Genet, 2007. **81**(3): p. 559-75.
238. Binard, A., et al., *Is the blood B-cell subset profile diagnostic for Sjogren syndrome?* Ann Rheum Dis, 2009. **68**(9): p. 1447-52.
239. Haneklaus, M., et al., *miR-223: infection, inflammation and cancer.* J Intern Med, 2013. **274**(3): p. 215-26.
240. Hovanessian, A.G. and J. Justesen, *The human 2'-5'oligoadenylate synthetase family: unique interferon-inducible enzymes catalyzing 2'-5' instead of 3'-5' phosphodiester bond formation.* Biochimie, 2007. **89**(6-7): p. 779-88.
241. Cole, M.B., et al., *Epigenetic signatures of salivary gland inflammation in sjogren's syndrome.* Arthritis Rheumatol, 2016.
242. Pei, L., A. Castrillo, and P. Tontonoz, *Regulation of macrophage inflammatory gene expression by the orphan nuclear receptor Nur77.* Mol Endocrinol, 2006. **20**(4): p. 786-94.
243. Norheim, K.B., et al., *Interleukin-1 inhibition and fatigue in primary Sjogren's syndrome--a double blind, randomised clinical trial.* PLoS One, 2012. **7**(1): p. e30123.
244. Emamian, E.S., et al., *Peripheral blood gene expression profiling in Sjogren's syndrome.* Genes Immun, 2009. **10**(4): p. 285-96.
245. Hjelmervik, T.O., et al., *Gene expression profiling of minor salivary glands clearly distinguishes primary Sjogren's syndrome patients from healthy control subjects.* Arthritis Rheum, 2005. **52**(5): p. 1534-44.
246. Wildenberg, M.E., et al., *Systemic increase in type I interferon activity in Sjogren's syndrome: a putative role for plasmacytoid dendritic cells.* Eur J Immunol, 2008. **38**(7): p. 2024-33.
247. Devauchelle-Pensec, V., et al., *Gene expression profile in the salivary glands of primary Sjogren's syndrome patients before and after treatment with rituximab.* Arthritis Rheum, 2010. **62**(8): p. 2262-71.
248. Thompson, N., et al., *Exploring BAFF: its expression, receptors and contribution to the immunopathogenesis of Sjogren's syndrome.* Rheumatology (Oxford), 2016. **55**(9): p. 1548-55.
249. Lavie, F., et al., *Increase of B cell-activating factor of the TNF family (BAFF) after rituximab treatment: insights into a new regulating system of BAFF production.* Ann Rheum Dis, 2007. **66**(5): p. 700-3.
250. Johnsen, S.J., et al., *Low Protein A20 in Minor Salivary Glands is Associated with Lymphoma in Primary Sjogren's Syndrome.* Scand J Immunol, 2016. **83**(3): p. 181-7.
251. Nocturne, G., et al., *Germline and somatic genetic variations of TNFAIP3 in lymphoma complicating primary Sjogren's syndrome.* Blood, 2013. **122**(25): p. 4068-76.
252. Nocturne, G., et al., *Germline variation of TNFAIP3 in primary Sjogren's syndrome-associated lymphoma.* Ann Rheum Dis, 2016. **75**(4): p. 780-3.

253. Kim, B.H., et al., *A family of IFN-gamma-inducible 65-kD GTPases protects against bacterial infection.* Science, 2011. **332**(6030): p. 717-21.
254. Ronnblom, L. and M.L. Eloranta, *The interferon signature in autoimmune diseases.* Curr Opin Rheumatol, 2013. **25**(2): p. 248-53.
255. Mok, A., et al., *Genome-wide profiling identifies associations between lupus nephritis and differential methylation of genes regulating tissue hypoxia and type 1 interferon responses* Lupus Sci Med, 2016(3:e000183).
256. de Andres, M.C., et al., *Assessment of global DNA methylation in peripheral blood cell subpopulations of early rheumatoid arthritis before and after methotrexate.* Arthritis Res Ther, 2015. **17**: p. 233.
257. Liu, Y., et al., *Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis.* Nat Biotechnol, 2013. **31**(2): p. 142-7.
258. Houseman, E.A., J. Molitor, and C.J. Marsit, *Reference-free cell mixture adjustments in analysis of DNA methylation data.* Bioinformatics, 2014. **30**(10): p. 1431-9.
259. Rahmani, E., et al., *Sparse PCA corrects for cell type heterogeneity in epigenome-wide association studies.* Nat Methods, 2016. **13**(5): p. 443-5.
260. Stricker, S.H., A. Koferle, and S. Beck, *From profiles to function in epigenomics.* Nat Rev Genet, 2017. **18**(1): p. 51-66.
261. Thakore, P.I., et al., *Editing the epigenome: technologies for programmable transcription and epigenetic modulation.* Nat Methods, 2016. **13**(2): p. 127-37.
262. Liu, X.S., et al., *Editing DNA Methylation in the Mammalian Genome.* Cell, 2016. **167**(1): p. 233-247 e17.
263. Hilton, I.B., et al., *Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers.* Nat Biotechnol, 2015. **33**(5): p. 510-7.
264. Büchner, K.G., *Dantons Tod*. 1835. 7 ed. 2007: Suhrkamp.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1286

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-310388

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017