Exploring the world of non-coding genes in stem cells and autoimmunity.

Messemaker, T.C.

Chapter 1

General introduction

Autoimmune diseases (AIDs) are common and can affect a wide variety of organs. Understanding why the immune system attacks the body's own cells is crucial in order to treat patients and prevent the onset of autoimmunity. In the past 100 years great efforts have been undertaken to gather insight into AIDs. Despite the advances, these studies have revealed that the complexity of AIDs is enormous. A wide range of AIDs exists. A few of the most common AIDs are rheumatoid arthritis, systemic lupus erythematosus, type I diabetes, thyroiditis, multiple sclerosis and psoriasis[1–3]. However, not all AIDs affect a large proportion of the population. For example the prevalence of systemic sclerosis is ~100 times lower than rheumatoid arthritis[4]. AIDs display a typical preference for females albeit the reason for this is still unknown[5–8]. Both genetic and environmental factors contribute to dysregulation of the immune system and disease pathogenesis. Some of these genetic and environmental factors overlap among different AIDs, but also disease-specific factors have been identified[9,10]. Genes identified through genetic studies, have pinpointed to the involvement of multiple pathways, which also act in cell-type specific manners[10,11]. Multiple environmental factors have been identified and are thought to play a role in the onset and development of AIDs. These include smoking, exposure to UV, microbes, nutrients and exposure to organic substances[12–15]. This variety of contributing factors illustrates the complexity of AID and indicates why causal factors are notoriously hard to be identified. In this thesis, rheumatoid arthritis and systemic sclerosis were more closely investigated. Which genes play a role and how these genes are deregulated were the main objectives of these studies. Moreover, the role of non-coding RNAs was studied in the context of rheumatoid arthritis, systemic sclerosis, and more basic transcriptional regulation.

Rheumatoid arthritis

Rheumatoid arthritis (RA) is the most common autoimmune disease with a prevalence of 0.5 to 1% in the adult population worldwide[16]. Prime characteristics of RA are inflammation of the joints leading to cartilage damage and bone damage. Many cell types are involved in this process including T-cells, B-cells, monocytes, macrophages, dendritic cells, synovial fibroblasts, synoviocytes, neutrophils, osteoclasts and mast cells[17]. These cell types are involved in i.) recognizing the self-proteins as foreign proteins, ii.) enhancing inflammation by cytokine production and the recruitment of other immune cells and iii.) secreting

enzymes involved in bone erosion and destruction[18]. The interplay between these cells and processes likely contributes to a self-stimulating process that results in the chronic nature of RA. The disease is more prevalent in women with a 2-3x higher incidence[19]. RA is a heterogeneous disease indicated by both seropositive and seronegative patients. Seropositivity is indicated by various autoantibodies, and associates with more severe symptoms, joint damage and higher mortality[20–24]. The most prevalent autoantibody known is Rheumatoid Factor (RF), which recognizes the Fc part of an IgG molecule. RF is found in approximately 75% of patients, however this autoantibody is also found in other diseases and in healthy individuals upon ageing[25]. A more RA-specific autoantibody is the anti-citrullinated antibody (ACPA), which is directed against citrullinated proteins. ACPAs are found in approximately 70% of patients and are highly specific for RA[26]. A more recently discovered autoantibody in RA patients is the anti-carbamylated protein antibody (anti-CarP), which recognizes carbamylated proteins. These anti-CarP antibodies are present in ~40% of the patients and associate with disease activity and bone damage[23,27]. The positivity for some of these autoantibodies is linked to environmental factors. ACPA-positivity and RF-positivity are both higher in patients who have been smoking compared to non-smokers, while no specific relationship with anti-Carp positivity exists[27,28]. Interestingly these autoantibodies are present years before disease onset and can therefore be used for diagnosing RA[29,30]. Although these autoantibodies are key in diagnoses, classification, and prediction of disease severity, they are currently not exploited as targets for treatment.

Currently, the best treatment is provided by inhibition of pro-inflammatory cytokines. In RA patients, cytokines like TNFα and IL6 are elevated in the serum and synovium compared with healthy individuals[31–33]. Both anti-TNFα and anti-IL6 treatment are currently successful in alleviating rheumatoid arthritis[34,35]. TNFα is a potent inducer of inflammatory genes, resulting in increased local inflammation and bone degradation[33,34]. Prominent TNF inhibitors include infliximab, etanercept, adalimumab, golimumab and certolizumab pegol[36]. IL6 is a cytokine that activates the immune response of several cell types and is also involved in the maturation of B-cells[37]. Tocilizumab is an IL6-inhibitor which can bind soluble and membrane-bound IL6-receptors and is used as a therapeutic strategy to treat RA[38]. However, the mechanism by which the immune system is activated and the mechanism by which TNFα and IL6 are enhanced remain elucidative.

Both environmental and genetic components have been identified in RA. On basis of twin studies, the genetic component is estimated to account for 60% of the susceptibility to RA and is even more contributing in seropositive RA[39,40]. Besides familial or twin studies, a genetic contribution can also be investigated by analysing frequencies of genetic variants in large populations referred to as genome wide association studies (GWAS). In GWAS, the prevalence of genetic variants (such as single nucleotide polymorphism (SNPs)) in a disease population is compared with the prevalence of the same genetic variants in a healthy population thereby investigating whether certain variants are more frequent in diseased individuals. GWAS have identified over 100 associated loci that contribute to RA[41]. The strongest associating variants are located in the *HLA*-region on chromosome 6, which explains approximately 80% of the genetic contribution[41]. Specifically, variants in the *HLA-DRB1* gene, a gene involved in peptide presentation, associates strongly with RA susceptibility. The RA associated variants encode amino acid sequences in the peptide-binding groove, which is known as the shared epitope (SE)[42]. The SE epitope includes QKRAA, QQRAA and KKRAA on position 70-74 of the *HLA-DRB1* chain and points to a crucial role for peptide (and self-peptide) binding in RA pathogenesis[43].

Next to the association of the HLA locus, many non-HLA regions have been associated to RA. These non-HLA regions have lower odds ratios and are probably involved with a smaller functional contribution[41]. Several studies have investigated how the variants in non-HLA genes may translate to the increased onset and development of RA and other AIDs[44]. For example, *PTPN22* is a gene which acts as a negative regulator of the T-cell receptor of which several variants have been associated with multiple autoimmune diseases[45]. These variants are thought to interfere with PTPN22 functioning resulting in a diminished inhibitory effect and therefore increased T-cell activation[46,47]. Overall, genes located in these non-HLA regions are significantly enriched for immune-related pathways like NF-kb signalling pathway, T-cell receptor signalling pathway and the JAK-STAT signalling pathway[44]. Likely, genetic variants in the non-HLA regions disrupt genes within these pathways making an individual more susceptible to inflammatory diseases like RA. Together, RA is a multifactorial disease influenced by genetic and environmental factors, which together likely result in disturbed immune homeostasis in synovial areas eventually leading to disease pathogenesis, figure 1.

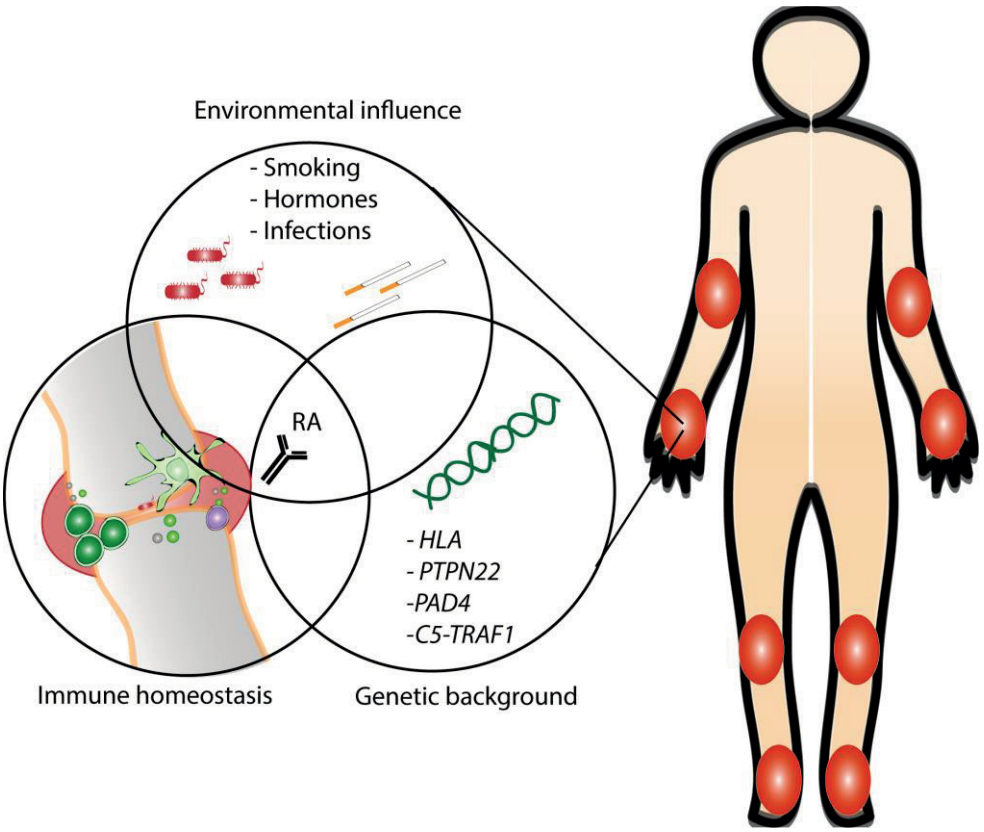

Figure 1. Schematic overview of factors influencing rheumatoid arthritis.

Systemic sclerosis

Systemic sclerosis (SSc) is a heterogeneous autoimmune disease with a prevalence of ~0.02% in the western population[4]. Similar to rheumatoid arthritis, females are more prone to develop this autoimmune disease with an observed female-to-male ratio of up to 1:5[48,49]. The typical diagnosis of SSc patients is based on fibrosis of the skin and complications of other internal organs. Over 90% of patients show skin fibrosis, ~90% gastrointestinal complications, ~65% musculoskeletal problems, ~40% interstitial lung disease, and ~15% of patients suffer from pulmonary arterial hypertension (PAH)[50]. Moreover typical characteristics of SSc are vascular manifestations reminiscent of Raynaud phenomenon, which are often observed prior to diagnosis of SSc[51]. SSc patients are grouped into limited cutaneous systemic sclerosis (lcSSc) and diffuse

cutaneous systemic sclerosis (dcSSc) on basis of their skin involvement. In lcSSc, skin involvement is restricted to the region between fingers and elbow, and face, while in dcSSc proximal regions are also affected[52]. Overall, dcSSc patients display a rapid disease progression with extensive skin fibrosis and development of complications of the internal organs but severity of the disease may differ between patients and between disease-subtypes[52].

A number of autoantibodies have been detected in SSc. The presence of autoantibodies is used for SSc diagnosis and classification of patients. Autoantibodies are mainly directed against nuclear components and are described as anti-nuclear antibodies (ANAs). ANAs include anti-centromeric antibodies, anti-topoisomerase I, anti-RNA polymerase III, anti-U1-RNP, anti-U3-RNP, anti-Th/To, anti-Pm/Scl and anti-nucleolar antibodies[53]. lcSSc patients display a stronger association with anti-centromeric antibodies, while dcSSc patients often have anti-topoisomerase and anti-RNA polymerase antibodies[53]. Finally, the complexity and heterogeneity of SSc is illustrated by the fact that some individuals are positive for SSc serology but lack the presence of detectable skin involvement[54].

Similar to RA and other AIDs, environmental and genetic components have been identified in SSc, figure 2. Known environmental factors are pollutants and chemicals including silica dust, vinyl chloride and organic substances[55,56]. Moreover, infectious agents, like viruses, have been reported to be associated with risk of developing SSc[55]. Despite numerous studies, the molecular mechanisms underlying SSc remain elusive. Approximately 40 genes have been linked to SSC by multiple genetic studies[57–59]. The *HLA*-locus (*HLA-DR* and *HLA-DP*) shows the strongest genetic association with SSc, indicating that (self) antigen presentation plays a role in SSc. Clinical implications of these genetic associations have been shown by correlation analysis with the presence of autoantibodies. Anti-topoisomerase antibodies correlated strongly with DPB1*1301 and DRB1*1101–21, while anti-centromeric antibodies were positively correlated with the presence of DRB1*0401–22 and DRB1*0801–11[60]. Besides the *HLA*-genes, other immunological genes are enriched in SSc associated regions, for example genes belonging to the interferon pathway[57].

Besides genetic evidence expression studies have shown that interferon genes are deregulated in patients. Expression studies have been performed in SSc tissues to investigate deregulated genes and altered pathways, and to discover

new drug targets. Affected tissues that have been investigated included skin, cultured fibroblasts, keratinocytes and various cells of the immune system. Besides, gene expression profiles are under investigation for the classification of SSc patients[61–64]. Milano *et al*. showed that patients can be subdivided into four groups on basis of gene expression profiles in the skin: i) deregulated expression of proliferative genes, ii) altered expression of inflammatory genes, iii) aberrant expression of fibrotic genes or iv) a normal-like gene expression profile[64].

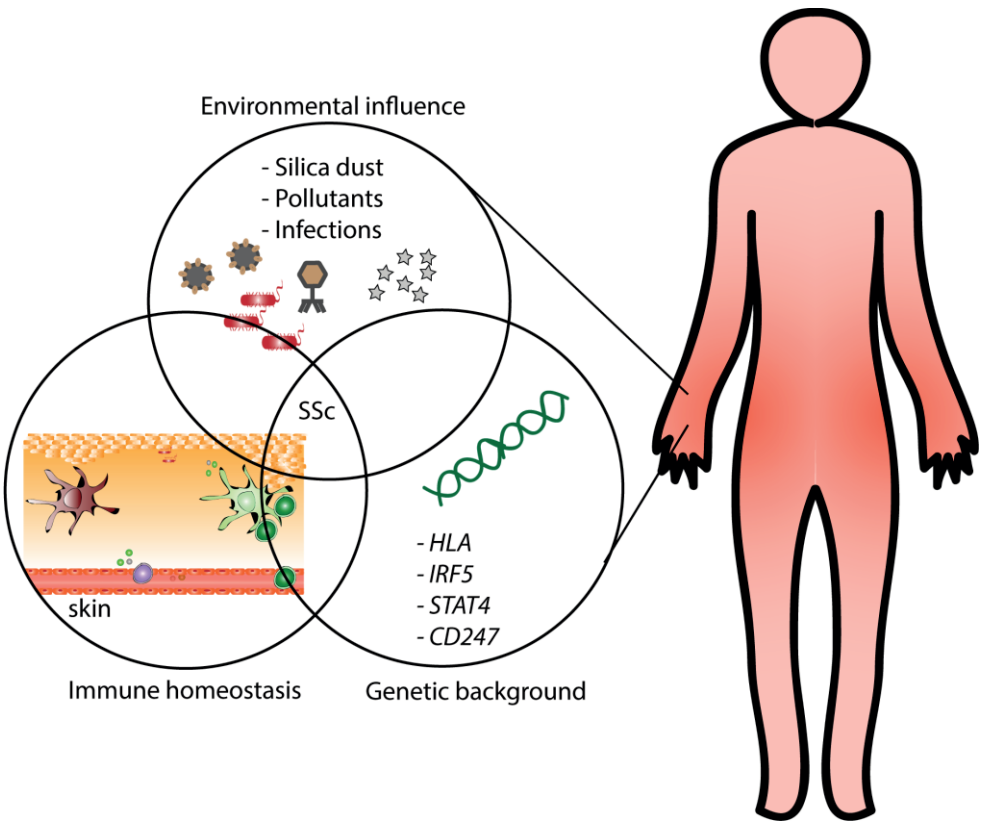

Figure 2. Schematic overview of factors influencing systemic sclerosis.

Although other studies did not classify SSc patients they observed deregulation of genes of immunological and fibro-proliferative nature[62,63]. Assassi *et al*. showed that many keratin-related genes are altered in SSc patients and that this keratin signature associate with early patients. On the other hand, a fibroinflammatory

signature correlates with dcSSc and higher skin scores[62]. Studies investigating gene expression in immune cells of SSc patients have described increased expression of macrophage and interferon genes including *CCR1*, *IL1B*, *IL13* and *JAK2*[65,66]. Together, these studies point to a disease mechanism in which a genetically primed individual is exposed to environmental factors triggering a chronic inflammatory process. This inflammatory process is characterized by vascular alterations and fibrosis. Early responses are likely mediated through macrophages, which induce the expression of interferon-related cytokines along with fibrotic mediators such as TGFβ. A continuous interplay between these pathways further exacerbates the fibrotic phenotype of the skin of patients which might be propelled further into other internal organs[67].
Current treatment of SSc is limited and consists mainly of immunosuppressive medication and haematopoietic stem cell transplantation. However, anti-fibrotic compounds are under investigation[68,69]. While current drug targets are based on coding genes only, a large proportion of the transcriptome is annotated as non-coding[70]. Thus investigating deregulated non-coding genes in SSc patients may lead to the identification of novel SSc biomarker genes and potentially new druggable targets.

Gene transcription and regulation and its relevance to rheumatic diseases.
Gene transcription and regulation is fundamental for all biological processes. Disturbed regulation of gene expression may lead to disease, including autoimmunity. Thus, insight into the complexity of gene regulation has broad implications for disease understanding and disease treatment.
In the early 1960's Francois Jacob and Jacques Monod pioneered the first model for gene regulation and introduced the phenomenon of gene activation and repression[71]. In the years after, many additional factors have been described to influence this process. Recruitment of RNA polymerases is essential for a gene to be transcribed into RNA. Recruitment and activation of RNA polymerases is mediated by transcription factors (TFs), which can bind specific DNA sequences known as promotors and enhancers. The interaction between DNA and transcription factors is defined by both the DNA sequence and the accessibility of the DNA. Importantly, this DNA accessibility is determined by the 3-dimensional structure and how the DNA is packed in the nucleus[72]. Alterations in the DNA structure, but not sequences are known as epigenetic changes and describes that these structural changes can be heritable to daughter cells and offspring[73]. In

humans, DNA is packed into chromatin in which the DNA is wrapped around histones. DNA with an open chromatin structure is called euchromatin and is associated with active gene transcription[74]. DNA with a dense chromatin structure is known as heterochromatin and is associated with gene silencing or gene repression[74]. An important component of the epigenetic landscape is formed by the histones which are modulators of the chromatin structure. Histones consist of 5 families, the linker histone H1 and 4 core histones: H2A, H2B, H3 and H4. Eight histones (2 of each 4 core histones) form together a nucleosome wrapping approximately 150bp of DNA. The N-terminal histone tails stick out making them accessible for modifications. These modification can steer both transcriptional repression and activation dependent on the type of modification, see table I[75–77].

Table I. Histone modifications and their role on gene transcription

Functional association	Modification	Modification site
Gene activation	Acetylation	H3K9, H3K14, H3K18, H3K27, H3K56, H4K5, H4K8, H4K12, H4K16, H2BK6, H2BK7, H2BK16, H2BK17.
	Methylation	H3K4me1, H3K4me2, H3K4me3, H3K36me3, H3K79me2.
	Phosphorylation	H3S10ph.
Gene repression	Methylation	H3K9me2, H3K9me3, H3K27me2, H3K27me3, H4K20me3.

K = Lysine
S = Serine
me1 = monomethylation
me2 = demethylation
me3 = trimethylation
ph = phosphorylation.

One of the best-studied modifications is the trimethyl-modification on the position 4 (lysine) of histone 3 (H3K4me3), which is associated with active promotors[77]. Many of these histone modifications regulate gene expression by interacting with other proteins. For example, H3K4me3 can interact with TFIID, a

protein involved in the initiation of gene transcription[78]. In contrast, histone modification H3K9me3 interacts with Heterochromatin Protein 1 (HP1) thereby silencing gene transcription[79]. Histone modifications are generated by histone methylation transferases (HMTs) and histone acetylation transferases (HATs) and can be removed by histone deacetylases (HDACs) and histone demethylases (KDMs)[80,81].

Several studies have studied the role of epigenetic changes in autoimmune diseases including RA and SSc[82]. In RA, especially synovial fibroblasts (SFs) seem to display epigenetic alterations in patients[83–85]. In RA-specific SFs global hypomethylation correlates with increased levels of multiple receptors, adhesion molecules, and matrix-degrading enzymes and correlated with an activated phenotype[84]. Moreover, changes in DNA methylation have been observed in various immune cells of RA patients. The promotor of *CD40LG* in CD4+ T-cells and the promotor of *IL6* in B-cells are hypomethylated in RA patients[86,87].

Furthermore, behaviour and levels of the enzymes modifying the histones have been investigated in RA patients. Gillespie *et al*. showed that PBMCs isolated from RA patients exhibit enhanced histone deacetylase (HDAC) activity and inhibition of this class of enzymes showed the potential to reduce IL6 and TNFα proteins in a cell type and compound-dependent manner[88]. Moreover, levels of EZH2 (Enhancer of Zeste Homolog 2), a histone methyl transferase creating H3K27Me3 are increased in synovial fibroblasts isolated from RA patients[89]. Finally, in mouse models beneficial effects such as, reduced joint swelling, inflammation and cartilage destruction, have been obtained using inhibitors of HDACs[90,91].

In SSc patients, altered DNA methylation influenced the expression of both immune and fibrotic genes including *CD40L*, *TNFSF7*, *CD11a* and *FLI-1*[92–96]. Moreover, a genome-wide study identified several collagen genes both methylation and differential expressed in dermal fibroblasts[97].

Finally, histone modifying enzymes have been investigated as potential therapeutic targets in SSc. Inhibition of HDAC7 showed reduced cytokine-induced production of type I and type III collagen[98], whereas interfering with other histone modifying enzymes was found to reduce the accumulation of extracellular matrix in SSc mouse models[99].

Histone modifications play an important role in the regulation of immune-responsive genes as many immune-related genes are under control of epigenetic mechanisms. Various studies have been performed investigating which epigenetic marks play a role in the regulation of immune cells[100,101]. These changes seem specific for pathogens and environmental stimuli. For example stimulation of monocytes by either LPS (bacterial origin) or B-glucan (yeast origin) induces a different response leaving different epigenetic traces[102,103]. Interestingly, B-glucan exposure of monocytes induces long-lasting epigenetic changes[104]. Upon restimulation of these cells, induced cytokine production is observed. This phenomenon is called 'trained immunity' and the identified epigenetic changes include increased H3K4me3 levels and reduced H3K27me3 levels on genes that are enriched for immunological pathways[104]. These findings indicate that these genes are more easily accessible and ready for transcription upon new activation.

Besides regulation at the transcriptional level, gene expression is also regulated at the RNA level. Specifically, a subset of small non-coding RNAs have been discovered as regulator of RNA transcripts known as miRNAs[105,106]. miRNAs are small RNA molecules that can bind mRNA (often in the 3' UTR regions of mRNAs) thereby blocking translation and accelerating degradation of the mRNA in a process called RNA interference[107]. In short, miRNA are processed transcripts of ~20 nucleotides that bind complementary RNA often including several mismatches. Upon binding, the RISC complex can recognize the double stranded RNA and the targeted mRNA gets degraded by Dicer, an enzyme capable of cleaving double stranded RNA[107]. The opportunity that mRNA levels can be regulated qualifies miRNA as possible therapeutic targets in autoimmunity. For example, miRNA29 a key regulator of collagen expression was found decreased in patients with systemic sclerosis[108]. More examples of deregulated miRNAs are reviewed in[82,109]. Besides these small non-coding RNAs, other non-coding RNAs have surfaced as new players in disease and development.

Long non-coding RNAs and their relevance to rheumatic diseases

The human genome encompasses roughly 60,000 genes. As approximately 20,000 genes encode proteins, a larger proportion is of non-coding nature. Non-coding genes can be divided in groups of small and long non-coding RNAs, figure

3A. Besides a role of small non-coding RNAs (eg. miRNAs) in autoimmune disease, long non-coding RNAs have been linked to functions in immunity and autoimmune diseases[110]. lncRNAs are RNA transcripts with low or no coding potential with a length of over 200 nucleotides[111]. lncRNAs are often polyadenylated, and lack open reading frames (ORFs)[111,112].

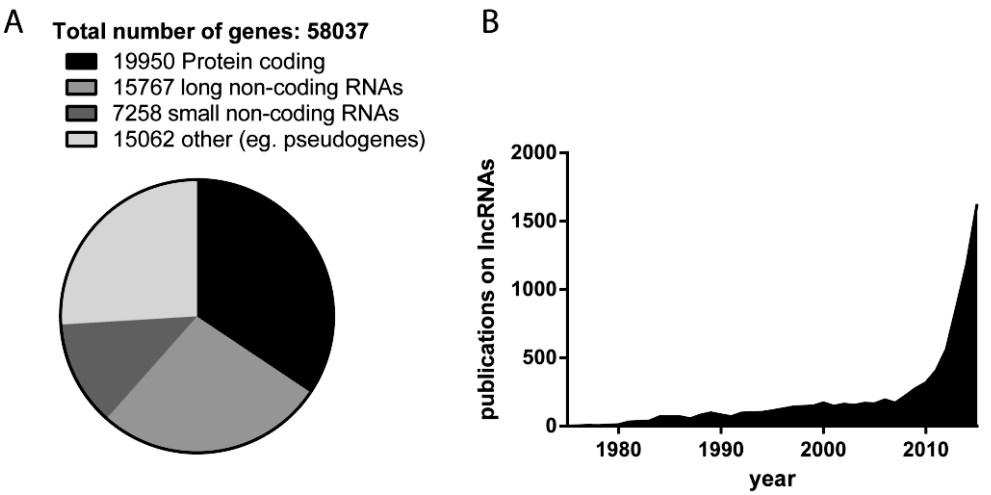

Figure 3. (**A**) The total number of genes divided into protein-coding genes, long non-coding RNAs (length of over 200 nucleotides), small non-coding RNAs and others (for example pseudogenes and immunoglobulin/T-cell receptor segments). Data was obtained from Gencode Version 25 (March 2016 freeze, GRCh38) - Ensembl 87. (**B**). Number of publications per year using search-term "Long non-coding RNAs", data subtracted from PubMed.

Since 2010, research and the number of publications studying the role of lncRNAs have exponentially increased (figure 3B). From these studies it has emerged that lncRNAs are important regulators of tissue physiology and disease processes[113]. Overall, lncRNAs are expressed at lower levels compared to coding genes and one of the most striking differences with coding RNAs is a more tissue-specific expression of lncRNAs[114,115]. Because of this tissue-specific expression, lncRNAs are considered important regulators of tissue specific physiology during development and during life[113,116,117]. It is key that during development gene expression is tightly regulated to prevent malformation of tissues and organs. For example, Sox2 is a transcription factor important for pluripotency and neuronal

development. Deregulation of Sox2 levels leads to malformation of neural tissues such as aberrant eye and brain development[114,118–121]. A lncRNA known as *Sox2* overlapping transcript (*Sox2ot*) is thought to play a role by safeguarding levels of Sox2 during neural development[122–124]. Humans have the highest count of lncRNAs and the number of lncRNAs has been shown to correlate with organism complexity[125,126]. Besides mammalian cells, lncRNAs are found in plants, yeast and even bacteria[127–129]. Although lncRNAs are found in many organisms, they are poorly conserved across species[115,130]. For example, only 14% of the mouse lncRNAs have a human orthologue[131], while another study shows that 12% of human lincRNAs have orthologous transcripts in other species[115]. Various types of lncRNAs have been described and can be divided into long intergenic non-coding RNAs (lincRNAs), intronic long non-coding RNAs, sense lncRNAs and antisense lncRNAs, figure 4.

Figure 4. Types of lncRNAs divided into intronic lncRNAs, intergenic lncRNAs, sense lncRNAs and antisense lncRNAs.

In comparison with small non-coding RNAs, the longer length of lncRNAs allows additional folding properties as stability and interaction abilities with DNA, RNA and protein[132]. lncRNAs can interact with single stranded DNA by direct base pairing or with double stranded DNA via triplex RNA-DNA structures[133]. lncRNAs are typically coexpressed with their neighbouring genes and are thought be involved with various gene regulatory processes[115]. Examples of lncRNAs that interfere with transcriptional and translational processes are summarized in figure 5[134,135]. Transcriptionally, lncRNAs are potent guiding molecules because of their ability to bind both DNA and proteins. By doing so, lncRNAs can bind co-activating or repressing proteins to specific genes and loci. Moreover, lncRNA have shown to play an important role in the modulation of epigenetic marks by recruiting histone modifying enzymes[136,137].

The other way around, lncRNAs can decoy proteins thereby preventing the interaction with specific regions[138–141]. Finally, expression of lncRNAs alone can result in interference with transcription of other genes. Transcriptional overlap of lncRNA *Airn*, but not its RNA transcripts were found crucial for silencing *Igf2r*[142]. Similarly, some lncRNAs are thought to function by disrupting DNA loops[143], while other lncRNAs (like Dum and HOTTIP) have been shown to establish DNA loops and coordinate gene expression[144,145].

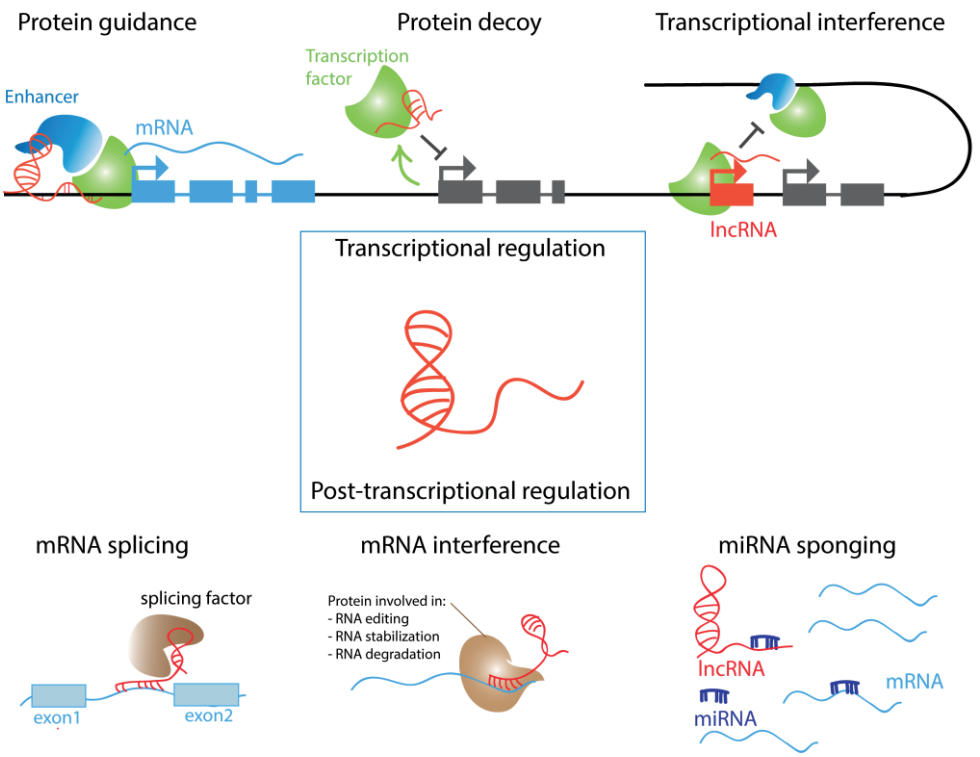

Figure 5. Mechanisms and function by lncRNAs. lncRNAs have potential roles in the regulation of transcription by protein guidance, protein decoy or via transcriptional interference (also mediated via DNA looping). Regulation of post-transcriptional processes by lncRNAs include the interference with mRNA splicing, stability, degradation, RNA editing and its interaction with miRNAs.

In addition, lncRNAs can regulate processes at the post-transcriptional level including, mRNA splicing, and interaction with mRNAs and miRNAs[146–148]. For example, ZEB2-AS1 and BACE1-AS1 are two antisense genes that interact with

mRNA by either influencing splicing or mRNA stability[149,150]. Antisense RNA genes are one of the largest subtypes of lncRNAs. These antisense RNA transcripts are transcribed from the opposite DNA strand of a sense gene in either close proximity or by partly overlapping, figure 4. The most prominent form of antisense transcription in the mammalian genome is of non-protein-coding nature[151]. Particularly, antisense ncRNAs represent an interesting subset as they are often involved in the regulation of its sense counterpart forming sense-antisense (SAS) gene pairs[152]. Besides ZEB2-AS1 and BACE1-AS, other SAS gene pairs have been shown to have implications in disease and development[153–156].

Studies with immune cells have shown that lncRNAs are important during the differentiation of immune cells[157,158]. Moreover, lncRNAs also play a role in differentiated immune cells and influence processes involved in innate and adaptive immunity[159–163]. Particularly in innate immunity lncRNA have been identified as important regulators. Many lncRNAs are under immunoregulatory control as shown by their responsiveness to external immune stimuli[164,165]. In both mouse and human derived monocytes/macrophages lncRNAs are responsive to LPS stimulation[156,165–167]. One of these studies shows that many of the lncRNA are coregulated or coexpressed with neighbouring protein-coding genes including *Nfkb2* and *Rel* two genes involved in NFKB-signalling[166]. Moreover, in human monocytes, 182 lncRNAs were induced by LPS, amongst others two lncRNAs in the *IL1B* locus that were identified as regulators of *IL1B* transcription and protein release[167]. Similarly, in monocyte-like cell lines, lncRNAs expression can be induced by various immune-stimuli[168,169]. For example, stimulation of THP1 cells with LPS showed 1161 lncRNA with differential expression[168]. Further knockdown experiments for some of these immune responsive-lncRNAs revealed their involvement with TNFα and IL6 levels[169]. Together, these studies show that lncRNAs have functional roles in our immune system and that they can influence the release of a variety of cytokines.

Besides these functional studies, genetic studies have linked lncRNAs to immunity and autoimmune diseases. Multiple GWAS have been performed to identify functional genetic regions that contribute to autoimmunity. Many of these associated regions are located in loci without coding genes, however can contain unannotated non-coding genes[164,170,171]. Nonetheless, SNPs occurring in these non-coding genes could have functional effects and subsequently result in

phenotypic changes contributing to complex diseases, including cancers and AID[44,172,173]. SNPs can strongly affect gene expression and are also known as expression quantitative trait loci (*cis*-eQTL)[174,175]. Many of these variants interfere with the binding of transcription factors or enhancers, however also lncRNAs could be functionally affected[176,177]. Several studies have linked genetic variants to changes in lncRNA levels through eQTL studies. Almhof *et al*. investigated cis-acting genetic variation that regulate expression of lncRNAs in monocytes isolated from 188 healthy donors[178]. 258 lncRNAs were detected with at least one associated cis-regulatory SNP of which 20% co-regulated with the closest protein coding gene[178]. Not only expression levels could be affected by these SNPs but studies have shown that genetic variants influence the predicted structure of lncRNAs related to immune diseases as T1D and inflammatory bowel disease[179]. Together, these studies show that associated variants can affect non-coding genes by altering the RNA stability, RNA expression levels and RNA structure and are therefore important candidates for further investigation in disease pathogenesis.

Outline of this thesis

Multiple factors have been identified that contribute to autoimmunity. Studies like association studies, gene expression studies and familial studies have revealed that both environmental and genetic factors contribute to disease pathogenesis. Although a dominant role for HLA genes in autoimmunity is very likely, non-HLA protein encoding genes have been identified as important clinical drug targets as well. In RA, anti-TNFα and anti-IL6 therapies are two successful strategies to treat patients. However, the mechanism by which TNFα and IL6 are upregulated in RA patients is yet unknown. In **Chapter 2** of this thesis, transcriptional regulation and the protein levels of IL6 and TNFα were investigated in RA patients. In particular, this study addressed whether epigenetic changes mediate enhanced IL6 and TNFα levels in early untreated RA patients and if these epigenetic changes can be detected in circulating monocytes isolated from these patients.

More recently, it was postulated that besides coding genes, non-coding genes also possess functional roles in development, immunity and various diseases. However, the involvement of these non-coding genes in autoimmunity and their

mode of action is poorly understood. **Chapter 3** of this thesis summarizes how both coding and non-coding variants can affect genes and how these variants may contribute to RA and other autoimmune diseases[44]. Interestingly, the majority of risk variants locate to non-coding regions of which the underlying disease-contributing mechanism is often unknown[180]. A RA-associated locus with many non-coding variants is the *TRAF1-C5* risk locus located on chromosome 9[181,182]. As described in **Chapter 4 and 5**, none of the identified variants are located in known coding regions but instead are synonymous, intronic or located in intergenic and UTR regions[183]. Such variants might also cover non-coding genes which are now hypothesized as a novel group of candidate genes that might explain part of the genetic variants associated with disease pathogenesis[164,172]. In addition, **Chapter 5** describes the identification of such a novel disease candidate gene, of non-coding nature, in the *TRAF1-C5* locus[164].

Besides a role for lncRNAs in RA, also lncRNAs are proposed to play pathogenic roles in SSc[184]. Investigating deregulated non-coding genes in SSc patients may lead to the identification of novel SSc biomarker genes and potentially new druggable targets. In **Chapter 6** of this thesis, expression of both coding and non-coding genes of SSc patients was compared to healthy controls to identify such targets. Moreover, the molecular mechanism by which deregulated antisense RNA genes play a role in SSc was investigated particularly. In **Chapter 7** of this thesis, the molecular mechanism of another lncRNA known as Sox2ot was further investigated. Sox2ot is a non-coding RNA that is located near Sox2, an important transcription factor for pluripotency and neural development. Together, these studies aid in unravelling the function of lncRNAs and to understand their role and involvement in development and disease pathogenesis. Finally, general remarks, implication and future directions are discussed in **Chapter 8**. To summarize, in this thesis we aimed at identifying and unravelling of enigmatic transcriptional mechanisms that contribute to rheumatoid arthritis (RA) but also to systemic sclerosis (SSc) with a particular focus on non-coding genes.

Answering the following questions was of main interest in this thesis:

i.) Are epigenetic changes underlying the development of RA and can they be detected in early untreated RA patients?

ii.) What are the mechanisms by which genetic variation contributes to autoimmunity and how do non-coding genes play a role in this process?

iii.) Are non-coding genes deregulated in AID and how can these deregulated genes contribute to autoimmunity?

iv.) What are the molecular mechanisms by which non-coding genes can function and play a role in disease and development?

References

1 R0ichard-Miceli C, Criswell LA. Emerging patterns of genetic overlap across autoimmune disorders. *Genome Med* 2012; **4**: 1–9.

2 Somers EC, Thomas SL, Smeeth L, Hall AJ. Autoimmune diseases co-occurring within individuals and within families: a systematic review. *Epidemiology* 2006; **17**: 202–17.

3 Cooper GS, Stroehla BC. The epidemiology of autoimmune diseases. Autoimmun. Rev. 2003; **2**: 119–125.

4 Barnes J, Mayes MD. Epidemiology of systemic sclerosis: incidence, prevalence, survival, risk factors, malignancy, and environmental triggers. *Curr Opin Rheumatol* 2012; **24**: 165–70.

5 van Vollenhoven RF. Sex differences in rheumatoid arthritis: more than meets the eye... *BMC Med* 2009; **7**: 12.

6 Kivity S, Ehrenfeld M. Can we explain the higher prevalence of autoimmune disease in women? *Expert Rev Clin Immunol* 2010; **6**: 691–4.

7 Amur S, Parekh A, Mummaneni P. Sex differences and genomics in autoimmune diseases. *J Autoimmun* 2012; **38**: J254-65.

8 Tiniakou E, Costenbader KH, Kriegel M a. Sex-specific environmental influences on the development of autoimmune diseases. *Clin Immunol* 2013; **149**: 182–91.

9 Zhernakova A, Withoff S, Wijmenga C. Clinical implications of shared genetics and pathogenesis in autoimmune diseases. *Nat Rev Endocrinol* 2013; **9**: 646–59.

10 Cotsapas C, Voight BF, Rossin E, Lage K, Neale BM, Wallace C *et al.* Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet* 2011; **7**: e1002254.

11 Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015; **518**: 337–43.

12 Danzer C, Mattner J. Impact of microbes on autoimmune diseases. *Arch Immunol Ther Exp (Warsz)* 2013; **61**: 175–86.

13 Barragán-Martínez C, Speck-Hernández C a, Montoya-Ortiz G, Mantilla RD, Anaya J-M, Rojas-Villarraga A. Organic solvents as risk factor for autoimmune diseases: a systematic review and meta-analysis. *PLoS One* 2012; **7**: e51506.

14 Draborg AH, Duus K, Houen G. Epstein-Barr virus in systemic autoimmune diseases. *Clin Dev Immunol* 2013; **2013**: 535738.

15 Wang L, Wang F-S, Gershwin ME. Human autoimmune diseases: a comprehensive update. *J Intern Med* 2015; **278**: 369–95.

16 Alamanos Y, Voulgari P V, Drosos A a. Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. *Semin Arthritis Rheum* 2006; **36**: 182–8.

17 Smolen JS, Steiner G. Therapeutic strategies for rheumatoid arthritis. *NatRev Drug Discov* 2003; **2**: 473–488.

18 Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. Lancet. 2016; **388**: 2023–2038.

19 Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F *et al.* Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. *Arthritis Res Ther* 2009; **11**: R7.

20 Aletaha D, Alasti F, Smolen JS. Rheumatoid factor, not antibodies against citrullinated proteins, is associated with baseline disease activity in rheumatoid arthritis clinical trials. *Arthritis Res Ther* 2015; **17**: 229.

21 van Gaalen F a, van Aken J, Huizinga TWJ, Schreuder GMT, Breedveld FC, Zanelli E *et al.* Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. *Arthritis Rheum* 2004; **50**: 2113–21.

22 Brink M, Verheul MK, Rönnelid J, Berglin E, Holmdahl R, Toes REM *et al.* Anti-carbamylated protein antibodies in the pre-symptomatic phase of rheumatoid arthritis, their relationship with multiple anti-citrulline peptide antibodies and association with radiological damage. *Arthritis Res Ther* 2015; **17**: 25.

23 Shi J, Knevel R, Suwannalai P, van der Linden MP, Janssen GMC, van Veelen P a *et al.* Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proc Natl Acad Sci U S A* 2011; **108**: 17372–7.

24 Gonzalez A, Icen M, Kremers HM, Crowson CS, Davis JM, Therneau TM *et al.* Mortality trends in rheumatoid arthritis: The role of rheumatoid factor. *J Rheumatol* 2008; **35**: 1009–1014.

25 Nell VPK, Machold KP, Stamm T a, Eberl G, Heinzl H, Uffmann M *et al.* Autoantibody profiling as early diagnostic and prognostic tool for rheumatoid arthritis. *Ann Rheum Dis* 2005; **64**: 1731–6.

26 Suzuki K, Sawada T, Murakami a, Matsui T, Tohma S, Nakazono K *et al.* High diagnostic performance of ELISA detection of antibodies to citrullinated antigens in rheumatoid arthritis. *Scand J Rheumatol* 2009; **32**: 197–204.

27 Jiang X, Trouw L a, van Wesemael TJ, Shi J, Bengtsson C, Källberg H *et al.* Anti-CarP antibodies in two large cohorts of patients with rheumatoid arthritis and their relationship to genetic risk factors, cigarette smoking and other autoantibodies. *Ann Rheum Dis* 2014; **73**: 1761–8.

28 Krol a, Garred P, Heegaard NHH, Christensen a F, Hetland ML, Stengaard-Pedersen K *et al.* Interactions between smoking, increased serum levels of anti-CCP antibodies, rheumatoid factors, and erosive joint disease in patients with early, untreated rheumatoid arthritis. *Scand J Rheumatol* 2015; **44**: 8–12.

29 Shi J, van de Stadt L a, Levarht EWN, Huizinga TWJ, Hamann D, van Schaardenburg D *et al.* Anti-carbamylated protein (anti-CarP) antibodies precede the onset of rheumatoid arthritis. *Ann Rheum Dis* 2014; **73**: 780–3.

30 Shi J, van Steenbergen HW, van Nies J a. B, Levarht EWN, Huizinga TWJ, van der Helm-van Mil AHM *et al.* The specificity of anti-carbamylated protein antibodies for rheumatoid arthritis in a setting of early arthritis. *Arthritis Res Ther* 2015; **17**: 339.

31 Wei S-T, Sun Y-H, Zong S-H, Xiang Y-B. Serum Levels of IL-6 and TNF-α May Correlate with Activity and Severity of Rheumatoid Arthritis. *Med Sci Monit* 2015; **21**: 4030–4038.

32 Houssiau FA, Devogelaer J, Damme JOVAN, Deuxchaisnes CNDE, Snick JVAN. Interleukin-6 in synovial fluid and serum of patients with rheumatoid arthritis and other inflammatory arthritides. *Arthritis Rheum* 1988; **31**: 784–788.

33 Tetta C, Camussi G, Modena V, Vittorio C Di, Baglioni C. Tumour necrosis factor in serum and synovial fluid of patients with active and severe rheumatoid arthritis. *Ann Rheum Dis* 1990; **49**: 665–667.

34 Xu S. TNF inhibitor therapy for rheumatoid arthritis. *Biomed Reports* 2012; **1**: 177–184.

35 Woodrick R, Ruderman EM. Anti-Interleukin-6 Therapy in Rheumatoid Arthritis. 2010; **68**.

36 Ma X, Xu S. TNF inhibitor therapy for rheumatoid arthritis. *Biomed reports* 2013; **1**: 177–184.

37 Srirangan S, Choy EH. The role of interleukin 6 in the pathophysiology of rheumatoid arthritis. *Ther Adv Musculoskelet Dis* 2010; **2**: 247–56.

38 Hennigan S, Kavanaugh A. Interleukin-6 inhibitors in the treatment of rheumatoid arthritis. *Ther Clin Risk Manag* 2008; **4**: 767–775.

39 Macgregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K *et al.* Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum* 2000; **43**: 30–37.

40 Frisell T, Holmqvist M, Källberg H, Klareskog L, Alfredsson L, Askling J. Familial risks and heritability of rheumatoid arthritis: role of rheumatoid factor/anti-citrullinated protein antibody status, number and type of affected relatives, sex, and age. *Arthritis Rheum* 2013; **65**: 2773–82.

41 Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2013. doi:10.1038/nature12873.

42 Gregersen PK, Silver J, Winchester RJ. The Shared Epitope Hypothesis. *Arthritis Rheum* 1987; **30**: 1205–1212.

43 Scally SW, Petersen J, Law SC, Dudek NL, Nel HJ, Loh KL *et al.* A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med* 2013; **210**: 2569–82.

44 Messemaker TC, Huizinga TW, Kurreeman F. Immunogenetics of rheumatoid arthritis: Understanding functional implications. *J Autoimmun* 2015; **64**: 7–14.

45 Rawlings DJ, Dai X, Buckner JH, Alerts E. The Role of PTPN22 Risk Variant in the Development of Autoimmunity: Finding Common Ground between Mouse and Human. *J Immunol* 2015; **194**: 2977–84.

46 Zhang J, Zahir N, Jiang Q, Miliotis H, Heyraud S, Meng X *et al.* The autoimmune disease – associated PTPN22 variant promotes calpain-mediated Lyp / Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. *Nat Genet* 2011; **43**: 902–907.

47 Rieck M, Arechiga A, Onengut-gumuscu S, Greenbaum C, Concannon P, Buckner JH *et al.* Genetic variation in PTPN22 corresponds to altered function of T and B lymphocytes. *J Immunol* 2007; **179**: 4704–4710.

48 Al MASET. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum* 1980; **23**: 581–90.

49 Steen D V. Epidemiology of systemic sclerosis. *Rheumatology* 2003; **12**: 1455–1462.

50 Allanore Y, Simms R, Distler O, Trojanowska M, Pope J, Denton CP *et al.* Systemic sclerosis. *Nat Rev Dis Prim* 2015; **1**: 15002.

51 Sunderkötter C, Riemekasten G. Pathophysiology and clinical consequences of Raynaud's phenomenon related to systemic sclerosis. *Rheumatology (Oxford)* 2006; **45 Suppl 3**: iii33-5.

52 Hachulla E, Launay D. Diagnosis and classification of systemic sclerosis. *Clin Rev Allergy Immunol* 2011; **40**: 78–83.

53 Steen VD. Autoantibodies in systemic sclerosis. *Semin Arthritis Rheum* 2005; **35**: 35–42.

54 Lomeo RM. Progressive Systemic Sclerosis Sine Scleroderma Presenting as Pulmonary Interstitial Fibrosis. *Am J Med* 1989; **87**: 525–527.

55 De Martinis M, Ciccarelli F, Sirufo MM, Ginaldi L. An overview of environmental risk factors in systemic sclerosis. *Expert Rev Clin Immunol* 2016; **12**: 465–78.

56 Dospinescu P, Jones GT, Basu N. Environmental risk factors in systemic sclerosis. *Curr Opin Rheumatol* 2013; **25**: 179–83.

57 Korman BD, Criswell L a. Recent advances in the genetics of systemic sclerosis: toward biological and clinical significance. *Curr Rheumatol Rep* 2015; **17**: 21.

58 Mayes MD. The genetics of scleroderma: looking into the postgenomic era. *Curr Opin Rheumatol* 2012; **24**: 677–84.

59 Zochling J, Newell F, Charlesworth JC, Leo P, Stankovich J, Cortes A *et al.* An Immunochip-based interrogation of scleroderma susceptibility variants identifies a novel association at DNASE1L3. *Arthritis Res Ther* 2014; **16**: 438.

60 Gilchrist FC, Bunn C, Foley PJ, Lympany PA, Black CM, Welsh KI *et al.* Class II HLA associations with autoantibodies in scleroderma: a highly significant role for HLA-DP. *Genes Immun* 2001; **2**: 76–81.

61 Whitfield ML, Finlay DR, Murray JI, Troyanskaya OG, Chi J, Pergamenschikov A *et al.* Systemic and cell type-specific gene expression patterns in scleroderma skin. *Proc Natl Acad Sci U S A* 2003; **100**: 12319–12324.

62 Assassi S, Swindell WR, Wu M, Tan FD, Khanna D, Furst DE *et al.* Dissecting the heterogeneity of skin gene expression patterns in systemic sclerosis. *Arthritis Rheumatol (Hoboken, NJ)* 2015; **67**: 3016–26.

63 Gardner H, Shearstone JR, Bandaru R, Crowell T, Lynes M, Trojanowska M *et al.* Gene profiling of scleroderma skin reveals robust signatures of disease that are imperfectly reflected in the transcript profiles of explanted fibroblasts. *Arthritis Rheum* 2006; **54**: 1961–73.

64 Milano A, Pendergrass S a, Sargent JL, George LK, McCalmont TH, Connolly MK *et al.* Molecular subsets in the gene expression signatures of scleroderma skin. *PLoS One* 2008; **3**: e2696.

65 Pendergrass S a, Hayes E, Farina G, Lemaire R, Farber HW, Whitfield ML *et al.* Limited systemic sclerosis patients with pulmonary arterial hypertension show biomarkers of inflammation and vascular injury. *PLoS One* 2010; **5**: e12106.

66 Christmann RB, Hayes E, Pendergrass S, Padilla C, Farina G, Affandi AJ *et al.* Interferon and alternative activation of monocyte/macrophages in systemic sclerosis-associated pulmonary arterial hypertension. *Arthritis Rheum* 2011; **63**: 1718–28.

67 Johnson ME, Pioli P a, Whitfield ML. Gene expression profiling offers insights into the role of innate immune signaling in SSc. *Semin Immunopathol* 2015; **37**: 501–9.

68 Bournia VKK, Vlachoyiannopoulos PG, Selmi C, Moutsopoulos HM, Gershwin ME. Recent advances in the treatment of systemic sclerosis. *Clin Rev Allergy Immunol* 2009; **36**: 176–200.

69 Taroni JN, Martyanov V, Mahoney JM, Whitfield ML. A functional genomic meta-analysis of clinical trials in systemic sclerosis: towards precision medicine and combination therapy. *J Invest Dermatol* 2016. doi:10.1016/j.jid.2016.12.007.

70 Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F *et al.* GENCODE: The reference human genome annotation for the ENCODE project. *Genome Res* 2012; **22**: 1760–1774.

71 Jacob F, Monod J. Genetic Regulatory Mechanisms in the Synthesis. *J Mol Miol* 1960; **3**: 318–356.

72 Song L, Crawford GE. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb Protoc* 2010; **2010**: 1–13.

73 Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 1999; **23**: 314–318.

74 Hood L, Galas D. The digital code of DNA. *Nature* 2003; **421**: 444–8.

75 Tessarz P, Kouzarides T. Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol* 2014; **15**: 703–8.

76 Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res* 2011; **21**: 381–95.

77 Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* 2007; **129**: 823–37.

78 Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH *et al.* H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* 2013; **152**: 1021–36.

79 Boros J, Arnoult N, Stroobant V, Collet J-F, Decottignies A. Polycomb repressive complex 2 and H3K27me3 cooperate with H3K9 methylation to maintain heterochromatin protein 1α at chromatin. *Mol Cell Biol* 2014; **34**: 3662–74.

80 Wendler DS. Problems with the consensus definition of the therapeutic misconception. *J Clin Ethics* 2013; **24**: 387–394.

81 Kelley RK, Wang G, Venook AP. Biomarker use in colorectal cancer therapy. JNCCN J. Natl. Compr. Cancer Netw. 2011; **9**: 1293–1303.

82 Aslani S, Mahmoudi M, Karami J, Jamshidi AR, Malekshahi Z, Nicknam MH. Epigenetic alterations underlying autoimmune diseases. *Autoimmunity* 2016; **49**: 69–83.

83 Takami N, Osawa K, Miura Y, Komai K, Taniguchi M, Shiraishi M *et al.* Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. *Arthritis Rheum* 2006; **54**: 779–87.

84 Karouzakis E, Gay RE, Michel B a, Gay S, Neidhart M. DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. *Arthritis Rheum* 2009; **60**: 3613–22.

85 Neidhart M, Rethage J, Kuchen S, Ku P, Billingham ME, Gay RE *et al.* Retrotransposable l1 elements expressed in rheumatoid arthritis synovial tissue. *Arthritis Rheum* 2000; **43**: 2634–2647.

86 Nile CJ, Read RC, Akil M, Duff GW, Wilson AG. Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. *Arthritis Rheum* 2008; **58**: 2686–93.

87 Liao J, Liang G, Xie S, Zhao H, Zuo X, Li F *et al.* CD40L demethylation in CD4(+) T cells from women with rheumatoid arthritis. *Clin Immunol* 2012; **145**: 13–8.

88 Gillespie J, Savic S, Wong C, Hempshall A, Inman M, Emery P *et al.* Histone deacetylases are dysregulated in rheumatoid arthritis and a novel histone deacetylase 3-selective inhibitor reduces interleukin-6 production by peripheral blood mononuclear cells from rheumatoid arthritis patients. *Arthritis Rheum* 2012; **64**: 418–22.

89 Trenkmann M, Brock M, Gay RE, Kolling C, Speich R, Michel B a *et al.* Expression and function of EZH2 in synovial fibroblasts: epigenetic repression of the Wnt inhibitor SFRP1 in rheumatoid arthritis. *Ann Rheum Dis* 2011; **70**: 1482–8.

90 Nishida K, Komiyama T, Miyazawa S-I, Shen Z-N, Furumatsu T, Doi H *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**: 3365–76.

91 Manabe H, Nasu Y, Komiyama T, Furumatsu T, Kitamura A, Miyazawa S *et al.* Inhibition of histone deacetylase down-regulates the expression of hypoxia-induced vascular endothelial growth factor by rheumatoid synovial fi broblasts. *Inflamm Res* 2008; **57**: 4–10.

92 Czuwara-Ladykowska J, Shirasaki F, Jackers P, Watson DK, Trojanowska M. Fli-1 inhibits collagen type I production in dermal fibroblasts via an Sp1-dependent pathway. *J Biol Chem* 2001; **276**: 20839–48.

93 Jiang H, Xiao R, Lian X, Kanekura T, Luo Y, Yin Y *et al.* Demethylation of TNFSF7 contributes to CD70 overexpression in CD4+ T cells from patients with systemic sclerosis. *Clin Immunol* 2012; **143**: 39–44.

94 Lian X, Xiao R, Hu X, Kanekura T, Jiang H, Li Y *et al.* DNA demethylation of CD40l in CD4+ T cells from women with ystemic sclerosis: a possible explanation for female susceptibility. *Arthritis Rheum* 2012; **64**: 2338–45.

95 Wang Y, Shu Y, Xiao Y, Wang Q, Kanekura T, Li Y *et al.* Hypomethylation and overexpression of ITGAL (CD11a) in CD4(+) T cells in systemic sclerosis. *Clin Epigenetics* 2014; **6**: 25.

96 Wang Y, Fan P-S, Kahaleh B. Association between enhanced type I collagen expression and epigenetic repression of the FLI1 gene in scleroderma fibroblasts. *Arthritis Rheum* 2006; **54**: 2271–9.

97 Altorok N, Tsou P-S, Coit P, Khanna D, Sawalha AH. Genome-wide DNA methylation analysis in dermal fibroblasts from patients with diffuse and limited systemic sclerosis reveals common and subset-specific DNA methylation aberrancies. *Ann Rheum Dis* 2015; **74**: 1612–20.

98 Hemmatazad H, Rodrigues HM, Maurer B, Brentano F, Pileckyte M, Distler JHW *et al.* Histone deacetylase 7, a potential target for the antifibrotic treatment of systemic sclerosis. *Arthritis Rheum* 2009; **60**: 1519–29.

99 Huber LC, Distler JHW, Moritz F, Hemmatazad H, Hauser T, Michel B a *et al.* Trichostatin A prevents the accumulation of extracellular matrix in a mouse model of bleomycin-induced skin fibrosis. *Arthritis Rheum* 2007; **56**: 2755–64.

100 Reiner SL. Epigenetic control in the immune response. *Hum Mol Genet* 2005; **14 Spec No**: R41-6.

101 Lim PS, Li J, Holloway AF, Rao S. Epigenetic regulation of inducible gene expression in the immune system. *Immunology* 2013; **139**: 285–93.

102 Ifrim DC, Quintin J, Joosten L a B, Jacobs C, Jansen T, Jacobs L *et al.* Trained immunity or tolerance: opposing functional programs induced in human monocytes after engagement of various pattern recognition receptors. *Clin Vaccine Immunol* 2014; **21**: 534–45.

103 Saeed S, Quintin J, Kerstens HHD, Rao N a, Aghajanirefah A, Matarese F *et al.* Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. *Science (80-)* 2014; **345**: 1251086.

104 Quintin J, Saeed S, Martens JH a, Giamarellos-Bourboulis EJ, Ifrim DC, Logie C *et al.* Candida albicans infection affords protection against reinfection via functional reprogramming of monocytes. *Cell Host Microbe* 2012; **12**: 223–32.

105 Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* 2010; **79**: 351–79.

106 Afonso-Grunz F, Müller S. Principles of miRNA-mRNA interactions: beyond sequence complementarity. *Cell Mol Life Sci* 2015; **72**: 3127–41.

107 Jonas S, Izaurralde E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* 2015; **16**: 421–33.

108 Maurer B, Stanczyk J, Jüngel A, Akhmetshina A, Trenkmann M, Brock M *et al.* MicroRNA-29, a key regulator of collagen expression in systemic sclerosis. *Arthritis Rheum* 2010; **62**: 1733–43.

109 Chen X-M, Huang Q-C, Yang S-L, Chu Y-L, Yan Y-H, Han L *et al.* Role of Micro RNAs in the Pathogenesis of Rheumatoid Arthritis: Novel Perspectives Based on Review of the Literature. *Medicine (Baltimore)* 2015; **94**: e1326.

110 Esteller M. Non-coding RNAs in human disease. *Nat Rev Genet* 2011; **12**: 861–74.

111 Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 2012; **22**: 1775–89.

112 Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT *et al.* RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science (80-)* 2007; **316**: 1484–1488.

113 Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. *Cell* 2013; **152**: 1298–307.

114 Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y *et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet* 2015; **47**: 199–208.

115 Pj B, Cao H, Np C, Gk C, Davis S, Day N *et al.* Mohammad Heydarian thesis proposal - 1.19.2012 Page 27. *Genes …* 2012; **447**: 27–32.

116 Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet* 2014; **15**: 7–21.

117 van Leeuwen S, Mikkers H. Long non-coding RNAs: Guardians of development. *Differentiation* 2010; **80**: 175–83.

118 Taranova O V, Magness ST, Fagan BM, Wu Y, Surzenko N, Hutton SR *et al.* SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Genes Dev* 2006; **20**: 1187–1202.

119 Niu W, Zang T, Zou Y, Fang S, Smith DK, Bachoo R *et al.* In vivo reprogramming of astrocytes to neuroblasts in the adult brain. *Nat Cell Biol* 2013; **15**: 1164–75.

120 Bylund M, Andersson E, Novitch BG, Muhr J. Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nat Neurosci* 2003; **6**: 1162–8.

121 Graham V, Khudyakov J, Ellis P, Pevny L. SOX2 functions to maintain neural progenitor identity. *Neuron* 2003; **39**: 749–765.

122 Askarian-Amiri ME, Seyfoddin V, Smart CE, Wang J, Kim JE, Hansji H *et al.* Emerging role of long non-coding RNA SOX2OT in SOX2 regulation in breast cancer. *PLoS One* 2014; **9**: e102140.

123 Hou Z, Zhao W, Zhou J, Shen L, Zhan P, Xu C *et al.* A long noncoding RNA Sox2ot regulates lung cancer cell proliferation and is a prognostic indicator of poor survival. *Int J Biochem Cell Biol* 2014; **53**: 380–8.

124 Shahryari A, Rafiee MR, Fouani Y, Oliae NA, Samaei NM, Shafiee M *et al.* Two novel splice variants of SOX2OT, SOX2OT-S1, and SOX2OT-S2 are coupregulated with SOX2 and OCT4 in esophageal squamous cell carcinoma. *Stem Cells* 2014; **32**: 126–134.

125 Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U *et al.* The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* 2014; **505**: 635–40.

126 Taft RJ, Pheasant M, Mattick JS. The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* 2007; **29**: 288–99.

127 Chekanova J a. Long non-coding RNAs and their functions in plants. *Curr Opin Plant Biol* 2015; **27**: 207–16.

128 Yamashita A, Shichino Y, Yamamoto M. The long non-coding RNA world in yeasts. *Biochim Biophys Acta* 2016; **1859**: 147–54.

129 Weinberg Z, Perreault J, Meyer MM, Breaker RR. Exceptional structured noncoding RNAs revealed by bacterial metagenome analysis. *Nature* 2009; **462**: 656–9.

130 Chodroff RA, Goodstadt L, Sirey TM, Oliver PL, Davies KE, Green ED *et al.* Long noncoding RNA genes: conservation of sequence and brain expression among diverse amniotes. *Genome Biol* 2010; **11**: R72.

131 Church DM, Goodstadt L, Hillier LW, Zody MC, Goldstein S, She X *et al.* Lineage-specific biology revealed by a finished genome assembly of the mouse. *PLoS Biol* 2009; **7**: e1000112.

132 Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. *Nature* 2012; **482**: 339–46.

133 Li Y, Syed J, Sugiyama H. RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chem Biol* 2016; **23**: 1325–1333.

134 Guil S, Esteller M. Cis-acting noncoding RNAs: friends and foes. *Nat Struct Mol Biol* 2012; **19**: 1068–75.

135 Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem* 2012; **81**: 145–66.

136 Richard I. Joh, Christina M. Palmieri, Ian T Hill and MM. Regulation of histone methylation by noncoding RNAs. *Biochim Biophys Acta* 2014; **1839**: 1385–1394.

137 Holdt LM, Hoffmann S, Sass K, Langenberger D, Scholz M, Krohn K *et al.* Alu elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. *PLoS Genet* 2013; **9**: e1003588.

138 Xue Z, Hennelly S, Doyle B, Gulati A a, Novikova I V, Sanbonmatsu KY *et al.* A G-Rich Motif in the lncRNA Braveheart Interacts with a Zinc-Finger Transcription Factor to Specify the Cardiovascular Lineage. *Mol Cell* 2016; **64**: 37–50.

139 Klattenhoff C a, Scheuermann JC, Surface LE, Bradley RK, Fields P a, Steinhauser ML *et al.* Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* 2013; **152**: 570–83.

140 Grote P, Herrmann BG. The long non-coding RNA Fendrr links epigenetic control mechanisms to gene regulatory networks in mammalian embryogenesis. *RNA Biol* 2013; **10**: 1579–1585.

141 Bläss G, Kellis M, Werber M, Herrmann BG. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell* 2013; **24**: 206–214.

142 Latos PA, Pauler FM, Koerner M V, Ba H, Hudson QJ, Stocsits RR *et al.* Airn Transcriptional Overlap, But Not Its lncRNA Products, Induces Imprinted Igf2r Silencing. *Science (80-)* 2012; **338**.

143 Quinodoz S, Guttman M. Long noncoding RNAs: an emerging link between gene regulation and nuclear organization. *Trends Cell Biol* 2014; **24**: 651–63.

144 Wang L, Zhao Y, Bao X, Zhu X, Kwok YK-Y, Sun K *et al.* LncRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. *Cell Res* 2015; **25**: 335–50.

145 Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 2011; **472**: 120–4.

146 Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet* 2016; **17**: 272–83.

147 Bayoumi AS, Sayed A, Broskova Z, Teoh J-P, Wilson J, Su H *et al.* Crosstalk between Long Noncoding RNAs and MicroRNAs in Health and Disease. *Int J Mol Sci* 2016; **17**: 356.

148 Chen M-T, Lin H-S, Shen C, Ma Y-N, Wang F, Zhao H-L *et al.* PU.1-Regulated Long Noncoding RNA lnc-MC Controls Human Monocyte/Macrophage Differentiation through Interaction with MicroRNA 199a-5p. *Mol Cell Biol* 2015; **35**: 3212–24.

149 Lan T, Chang L, Wu L, Yuan Y. Downregulation of ZEB2-AS1 decreased tumor growth and metastasis in hepatocellular carcinoma. *Mol Med Rep* 2016; **14**: 4606–4612.

150 Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE *et al.* Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β-secretase. *Nat Med* 2008; **14**: 723–730.

151 Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M *et al.* Antisense transcription in the mammalian transcriptome. *Science (80-)* 2005; **309**: 1564–1566.

152 Chen J, Sun M, Hurst LD, Carmichael GG, Rowley JD. Genome-wide analysis of coordinate expression and evolution of human cis- encoded sense-antisense transcripts. *Trends Genet* 2005; **21**: 326–329.

153 Peng H, Liu Y, Tian J, Ma J, Tang X, Rui K *et al.* The Long Noncoding RNA IFNG-AS1 Promotes T Helper Type 1 Cells Response in Patients with Hashimoto's Thyroiditis. *Sci Rep* 2015; **5**: 17702.

154 Chen N, Guo D, Xu Q, Yang M, Wang D, Peng M. Long non-coding RNA FEZF1-AS1 facilitates cell proliferation and migration in colorectal carcinoma. *Oncotarget* 2016; **7**.

155 Liu C, Lin J. Long noncoding RNA ZEB1-AS1 acts as an oncogene in osteosarcoma by epigenetically activating ZEB1. *Am J Transl Res* 2016; **8**: 4095–4105.

156 Chan J, Atianand M, Jiang Z, Carpenter S, Aiello D, Elling R *et al.* Cutting Edge: A Natural Antisense Transcript, AS-IL1α, Controls Inducible Transcription of the Proinflammatory Cytokine IL-1α. *J Immunol* 2015; **195**: 1359–63.

157 Ranzani V, Rossetti G, Panzeri I, Arrigoni A, Bonnal RJP, Curti S *et al.* The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat Immunol* 2015; **16**: 318–25.

158 Muljo SA, Zhu J, Zhao K. Expression and regulation of lincRNAs during T cell development and differentiation. *Nat Immunol* 2013; **14**: 1190–1198.

159 Zhang Y, Cao X. Long noncoding RNAs in innate immunity. *Cell Mol Immunol* 2016; **13**: 138–47.

160 Sigdel KR, Cheng A, Wang Y, Duan L, Zhang Y. The Emerging Functions of Long Noncoding RNA in Immune Cells: Autoimmune Diseases. *J Immunol Res* 2015; **2015**: 848790.

161 Imamura K, Akimitsu N. Long Non-Coding RNAs Involved in Immune Responses. *Front Immunol* 2014; **5**: 573.

162 Rapicavoli N a, Qu K, Zhang J, Mikhail M, Laberge R-M, Chang HY. A mammalian pseudogene lncRNA at the interface of inflammation and anti-inflammatory therapeutics. *Elife* 2013; **2**: e00762.

163 Liu B, Sun L, Liu Q, Gong C, Yao Y, Lv X *et al.* A cytoplasmic NF-κB interacting long noncoding RNA blocks IκB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* 2015; **27**: 370–81.

164 Messemaker TC, Frank-Bertoncelj M, Marques RB, Adriaans a, Bakker a M, Daha N *et al.* A novel long non-coding RNA in the rheumatoid arthritis risk locus TRAF1-C5 influences C5 mRNA levels. *Genes Immun* 2016; **17**: 85–92.

165 Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL *et al.* A long noncoding RNA induced by TLRs mediates both activation and repression of immune response genes. *Science (80-)* 2014; **341**.

166 Mao A-P, Shen J, Zuo Z. Expression and regulation of long noncoding RNAs in TLR4 signaling in mouse macrophages. *BMC Genomics* 2015; **16**: 45.

167 IIott NE, Heward J a, Roux B, Tsitsiou E, Fenwick PS, Lenzi L *et al.* Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes. *Nat Commun* 2014; **5**: 3979.

168 Cui H, Xie N, Tan Z, Banerjee S, Thannickal VJ, Abraham E *et al.* The human long noncoding RNA, lnc-IL7R, regulates inflammatory response. *Eur J Immunol* 2014; : 1–36.

169 Li Z, Chao T-C, Chang K-Y, Lin N, Patil VS, Shimizu C *et al.* The long noncoding RNA THRIL regulates TNFα expression through its interaction with hnRNPL. Proc. Natl. Acad. Sci. U. S. A. 2014; **111**: 1002–7.

170 Ward LD, Kellis M. Interpreting noncoding genetic variation in complex traits and human disease. *Nat Biotechnol* 2012; **30**: 1095–106.

171 Chen G, Qiu C, Zhang Q, Liu B, Cui Q. Genome-wide analysis of human SNPs at long intergenic noncoding RNAs. *Hum Mutat* 2013; **34**: 338–44.

172 Ding J, Eyre S, Worthington J. Genetics of RA susceptibility, what comes next? *RMD Open* 2015; **1**: e000028–e000028.

173 Pasmant E, Sabbagh A, Vidaud M, Bièche I. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J* 2011; **25**: 444–8.

174 Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C *et al.* Population genomics of human gene expression. *Nat Genet* 2007; **39**: 1217–1224.

175 Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. *Nature* 2005; **437**: 1365–9.

176 Hrdlickova B, Almeida RC De, Borek Z, Withoff S. Biochimica et Biophysica Acta Genetic variation in the non-coding genome : Involvement of micro-RNAs and long non-coding RNAs in disease ☆. *BBA - Mol Basis Dis* 2014; **1842**: 1910–1922.

177 Popadin K, Gutierrez-Arcelus M, Dermitzakis ET, Antonarakis SE. Genetic and epigenetic regulation of human lincRNA gene expression. *Am J Hum Genet* 2013; **93**: 1015–26.

178 Almlöf JC, Lundmark P, Lundmark A, Ge B, Pastinen T, Goodall AH *et al.* Single nucleotide polymorphisms with cis-regulatory effects on long non-coding transcripts in human primary monocytes. *PLoS One* 2014; **9**: e102612.

179 Mirza AH, Kaur S, Brorsson C a, Pociot F. Effects of GWAS-associated genetic variants on lncRNAs within IBD and T1D candidate loci. *PLoS One* 2014; **9**: e105723.

180 Spielmann M, Mundlos S. Looking beyond the genes: the role of non-coding variants in human disease. *Hum Mol Genet* 2016; **25**: R157–R165.

181 Plenge RM, Ph D, Seielstad M, Padyukov L, Lee T, Remmers EF *et al.* TRAF1–C5 as a Risk Locus for Rheumatoid Arthritis — A Genomewide Study. *NEJM* 2007; **357**: 1199–1209.

182 Kurreeman F a S, Goulielmos GN, Alizadeh BZ, Rueda B, Houwing-Duistermaat J, Sanchez E *et al.* The TRAF1-C5 region on chromosome 9q33 is associated with multiple autoimmune diseases. *Ann Rheum Dis* 2010; **69**: 696–9.

183 Messemaker T, Toes REM, Kurreeman F, Mikkers HMM. Comment on 'Functional Analysis of a Complement Polymorphism (rs17611) Associated with Rheumatoid Arthritis'. *J Immunol*; : 2.

184 Wang Z, Jinnin M, Nakamura K, Harada M, Kudo H, Nakayama W *et al.* Long non-coding RNA TSIX is upregulated in scleroderma dermal fibroblasts and controls collagen mRNA stabilization. *Exp Dermatol* 2016; **25**: 131–136.