UPPSALA UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1700

Animal genomics – gene discovery and gene characterization

RAKAN NABOULSI

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2020

ISSN 1651-6206
ISBN 978-91-513-1060-2
urn:nbn:se:uu:diva-423787

Dissertation presented at Uppsala University to be publicly examined in B41, BMC, Husargatan 3, 751 23 Uppsala, Thursday, 17 December 2020 at 08:00 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English. Faculty examiner: Professor Tosso Leeb (Institute of Genetics, University of Bern).

Abstract
Naboulsi, R. 2020. Animal genomics – gene discovery and gene characterization. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1700. 53 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1060-2.

This thesis involves two projects. The aim in the first project was to identify genomic regions associated with spontaneous autoimmune thyroiditis (SAT), which is a hereditary autoimmune disease that affects the obese strain (OS) of chicken, an animal model for human Hashimoto's thyroiditis (HT). In the second project, we study ZBED6, a highly conserved protein unique to placental mammals. Here we explore the functional significance of ZBED6 in general, and its effect on the regulation of *Igf2* and *miR483* in specific.

To identify genomic regions predisposing to SAT, a nine-generation intercross between OS and their wild ancestor, the red junglefowl (RJF), was previously generated. In paper I, we developed a cell-based assay to phenotype the F_9 chickens by measuring the TSH levels in their serum. We found that 1) SAT is similar to HT in the sense that the serum-TSH levels increase in affected individuals, and 2) that TSH levels in SAT-affected chickens starts to increase after 20 weeks of age. In paper II, a whole genome sequencing experiment was performed to compare a healthy and a severely SAT-affected groups of chicken. This analysis revealed 12 genomic loci to be significantly different between the two groups.

In the second project, we utilized a mouse myoblast cell line, C2C12, to characterize the function of ZBED6. In paper III, we affect ZBED6 function, by either mutating its binding site in *Igf2* (*Igf2*^{dGGCT}), or by completely knocking it out (*Zbed6*$^{-/-}$). Functional analysis of the mutant cells revealed that ZBED6 overexpression induces cell cycle arrest and apoptosis, that ZBED6 directly affects mitochondrial activity, and that ZBED6 in myoblast cells mainly exerts its effect through regulating *Igf2*. In paper IV, we use ZBED6 knock-out and knock-in mice to investigate the effect of ZBED6 on the regulation of miRNA expression. We found that ZBED6 is not a general regulator for miRNA, with the exception of *miR483*, which exists in an intron of *Igf2*. Thereafter, we generated *miR483*$^{-/-}$ cells, using the *Igf2*^{dGGCT} cell line. In this analysis we found that the main function of *miR483* in myoblast cells is to regulate the expression of *Igf2*, and that ZBED6 partially regulates *Igf2* through regulating *miR483*.

Keywords: Spontaneous autoimmune thyroiditis, SAT, Hashimoto's thyroiditis, HT, Obese strain, transcriptional regulation, ZBED6, Igf2, miRNA, miR483, Whole-genome sequencing, RNA-seq, miRNA-seq

Rakan Naboulsi, Department of Medical Biochemistry and Microbiology, Box 582, Uppsala University, SE-75123 Uppsala, Sweden.

ISSN 1651-6206
ISBN 978-91-513-1060-2
urn:nbn:se:uu:diva-423787 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-423787)

لوجهه تعالى لا شريك له

The image of the front cover was created using Biorender.com.

List of Papers

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

I. **Naboulsi, R.**, Kerje, S., Larsson, M., Andersson, L. (2020) Development of a cell-based luciferase assay for estimating TSH hormone levels in chicken serum. (Manuscript).

II. **Naboulsi, R.**, Kerje, S., Pettersson, M., Rafati, N., Larsson, M., Gunnarsson, U, Andersson, L. (2020) A genomic screen for loci causing spontaneous autoimmune thyroiditis in the Obese strain of chickens. (Manuscript).

III. Younis, S., **Naboulsi, R.**, Wang, X., Cao, X., Larsson, M., Sargsyan, E., Bergsten, P., Welsh, N., Andersson, L. (2019) The importance of the ZBED6-IGF2 axis for metabolic regulation in mouse myoblast cells. *FASEB J. 34(8):10250-66*.

IV. **Naboulsi, R.**, Larsson, M., Andersson, L., Younis, S. (2020) ZBED6 affects the expression of the oncogenic miR483 through its interaction with the *Igf2* locus. (Submitted manuscript).

Reprint of paper III was made with permission from the publisher.

Contents

Abbreviations

AID	Autoimmune disease
Bp	Base-pair
DNA	Deoxyribonucleic acid
EdU	5-ethynyl-2'-deoxyuridine
PCR	Polymerase chain reaction
MHC	Major histocompatibility complex
OS	Obese strain
PI	Propidium iodide
PS	Phosphatidylserine
QTL	Quantitative trait locus
QTN	Quantitative trait nucleotide
RJF	Red junglefowl
SAT	Spontaneous autoimmune thyroiditis
SNP	Single nucleotide polymorphism
TSH	Thyroid stimulating hormone
TSHR	Thyroid stimulating hormone receptor

Introduction

Life on earth exists in diverse forms. This astonishing variation scratched the curiosity of humans and led them to observe how life began, study how phenotypic characteristics are inherited, and explore the underlying causes by studying genetics in different life systems. Scientists have been studying how phenotypic characteristics are inherited, and explore how hereditary diseases arise. These are questions that can be answered by genetics, which is the field that studies the inheritance of hereditary traits. Humans shaped desired characteristics in animals through artificial selection without any knowledge about the underlying heritable elements that we now know as genetic factors. We have been selecting animals and plants for different traits long before 1866 when Gregor Mendel published his work on hybrid pea plants establishing the basic principles of genetics [1, 2].

The first isolation of deoxyribonucleic acids (DNA), by Friedrich Miescher in 1869 [3-5] and the description of chromatin by Walther Flemming in 1882 [6, 7], paved the way to Walter Sutton in 1903 to articulate the chromosomal theory of inheritance, which described the physical basis of the Mendelian law of heredity [8]. A theory that was first demonstrated by Tomas Hunt Morgan in 1910 who found out that the white eye color trait in *Drosophila melanogaster* is caused by a mutation that co-segregates with the sex chromosome X. Morgan's work provided experimental evidence supporting the chromosomal theory of inheritance.

In 1916, Morgan's doctoral student, Calvin Bridges, further asserted the chromosomal theory of inheritance and stated that genes are real entities that exist on chromosomes [9]. The findings of Morgan and his colleagues about the chromosome's involvement in heredity inaugurated research in the field of genetics and revolutionized biology.

The double-helix structural model of DNA was described by Watson and Crick in 1953 [10, 11]. This finding was a prerequisite for the development of primer-based DNA sequencing, developed by Frederick Sanger in 1977 [12] and known as Sanger sequencing. After this advent, DNA sequencing became a widespread and effective tool used by geneticists in their investigation of heredity. Later on, the polymerase chain reaction (PCR) was developed by Kary Mullis in 1986 [13]. This was subsequently followed by multiple technological advances that grew from it, such as the availability of the DNA-sequencers, which further aided the genetics revolution.

With the availability of these technological advances, geneticists determine the DNA sequences, the genotype, and try to associate them with a certain disease or trait, the phenotype. For example, one of the first genetic disorders ascribed to heredity that follows a mendelian recessive inheritance pattern [14], alkaptonuria, a monogenic disease, was found to be caused by variations in a single genetic locus, the homogentisate 1,2 dioxygenase gene [15]. However, the identification of the genetic determinants of polygenic conditions is not trivial as they are caused by variations in several genomic loci and deviate from mendelian inheritance patterns. Therefore, this process poses a considerable challenge and a significant motivation to researchers, especially when the studied trait is a polygenic human hereditary disease, such as Hashimoto's thyroiditis.

Genetic mapping

To identify genes important for specific phenotypes, a genetic mapping approach can be used. This approach requires the availability of a pedigree, genotyping of genetic markers to build a linkage map, and adequate phenotype data.

A pedigree can be created by choosing two lines that are phenotypically different in the trait of interest, given that the phenotypic difference corresponds to a clear genetic difference. A cross between two such lines (parental F_0) generates an F_1 generation that is expected to show a high heterozygosity at the phenotype-causing loci. The F_1 generation is backcrossed to either line of the parents, or intercrossed to expand the pedigree into further generations (F_2 or more).

Genetic markers are DNA sequence variants that are easy to genotype and show a simple Mendelian inheritance. One type of genetic marker is microsatellites that are short simple repetitive stretches of DNA sequences (1-5 bp) that show hypervariability in their size among individuals or lines, and show a random distribution throughout the genome [16]. Another type of genetic markers, that is the most abundant in genomes, is single nucleotide polymorphisms (SNPs) that are variations in the DNA sequence that occur when a single nucleotide varies among individuals.

After determining the genotypes of genetic markers, the most informative markers, i.e. that show strong differentiation between parental lines, are chosen for further analysis. These informative markers can be used to build a linkage map that shows the order and relative distance between markers, and that can be used to locate loci controlling phenotypic traits.

Additionally, phenotype data is an important factor needed to initiate the analysis. The type of available phenotype data such as being binary/continuous or complete/incomplete penetrance might have an impact on the choice of the analysis method. Finally, the statistical analysis of the co-segregation of genetic markers and the studied phenotype might identify the causative locus.

Examples of linkage studies for mapping simple monogenic, complete penetrance traits is the mapping of the chicken comb morphology [17] or plumage variants [18]. However, for traits that show an incomplete penetrance pattern, influenced by many genes, affected by environmental factors, and where phenotype data for individuals in a pedigree are available, a quantitative trait loci (QTL) analysis can be performed as it does not require any previous knowledge about the number of responsible loci or the mode of inheritance [19]. Although QTL mapping is a powerful method, it has its limitations, such as the number of crossing-over events that can occur in the generated pedigree [20]. An example of a method that overcomes the limitations of QTL studies, but still has its own limitations, is genome wide association studies (GWAS) which measures the association between phenotype and genotype using a large number of genetic markers in a large number of phenotyped unrelated individuals.

A more fine-mapping method as compared to GWAS is whole genome sequencing of individuals of known phenotypes. In this method, almost all SNPs in a given individual are sequenced to determine the genotype of the SNPs. Groups of individuals with different phenotypes are contrasted to identify the phenotype-causative genomic locus or loci. A variation of whole genome sequencing performed on individuals is the pooled whole genome sequencing, or simply pool-seq. Here, DNA samples of a group of individuals are mixed into a single pool that is sequenced and contrasted with a second pool derived from another group of individuals that have a different phenotype.

The ability to determine DNA sequences not only allows researchers to unravel the genetics predisposing to different hereditary diseases and phenotypic traits, but also expands beyond inheritance and renders the transcriptional regulation and functional characterization of genes feasible. These two features of genetic research are the basis of the current thesis, in which the hereditary spontaneous autoimmune thyroiditis (SAT), that occurs in the Obese strain (OS) of chickens, is used as an animal disease model to study human Hashimoto's thyroiditis. Additionally, the transcriptional regulation of *Igf2* by ZBED6 and the functional characterization of *miR483* are investigated.

Spontaneous autoimmune thyroiditis in the Obese strain of chickens as an animal model for human Hashimoto's thyroiditis

Autoimmune diseases

A properly functioning immune system defends the body against cancer cells and pathogens like viruses, bacteria. However, any abnormality in the function of the immune system might lead to an immunodeficiency, allergy or autoimmunity. Autoimmune diseases are a set of complex diseases characterized by an immune response mounted against self-antigens. This response starts when an antigen-presenting cell presents an antigen on its surface, in a major histocompatibility complex (MHC) class II dependent manner, to helper T lymphocytes that activate B lymphocytes or cytotoxic T lymphocytes to mount a humoral or cellular immune response, respectively. These responses aim to neutralize the targeted self-antigen, falsely recognized as foreign, resulting in the destruction of the tissues or organs that contain and express the antigen.

Autoimmunity is usually prevented by central and peripheral self-tolerance mechanisms that are mediated in central lymphoid organs (thymus and bone marrow), and in peripheral lymphoid tissues (spleen, lymph nodes and tonsils), respectively. Central tolerance induces apoptosis or anergy, a state of inactivity or dormancy, in autoreactive B and T lymphocytes that react to self-antigens. Conversely, peripheral tolerance neutralizes the lymphocytes that escaped the central negative selection process because of being weakly activated by self-antigens [21]. Any failure in the central or peripheral mechanisms might result in a loss of tolerance and trigger an autoimmune disease.

Hashimoto's thyroiditis

Hashimoto's thyroiditis in humans is an autoimmune disease with a complex genetic background [22]. It is characterized by the presence of autoantibodies against thyroid peroxidase (TPO), a thyroid specific enzyme [23], lymphocytic infiltration of the thyroid gland causing its impairment and leading to hypothyroidism, and elevated levels of thyroid stimulating hormone (TSH) that is secreted by the anterior pituitary gland.

Therefore, Hashimoto's thyroiditis is diagnosed by a blood test to check the levels of the thyroid's secreted hormones, triiodothyronine (T3) and thyroxine (T4), TSH hormone levels, and for the presence of anti-TPO antibodies. The reduction in T3 and T4 hormones below normal levels, hence the name hypothyroidism, causes many symptoms which include sensitivity to cold, pale and dry skin, muscle weakness, depression and weight gain [24].

Hashimoto's thyroiditis has a complex genetic background. Many different genetic risk factors were found to be associated with the susceptibility to the

disease, such as specific genes in the major histocompatibility complex (MHC) region, also known in humans as human leukocyte antigens (HLA). This locus includes a set of highly polymorphic genes that play a major role in the regulation of the immune system. These genes have several alleles that were found to present auto-antigenic peptides with higher affinity to T lymphocytes, thus increasing the susceptibility to Autoimmune diseases [22, 25, 26].

Several non-MHC immune-related genes have also been associated with the susceptibility to Hashimoto's thyroiditis. The cytotoxic T lymphocyte antigen-4 (CTLA-4) encodes a protein that has an effect on T lymphocyte activation. Certain polymorphisms of this gene reduce the inhibition of T lymphocyte proliferation thus increasing the risk of developing Hashimoto's thyroiditis [27]. Protein tyrosine phosphate (PTPN22) is a negative regulator of T-cell activation. It may play a role in the inhibition of T regulatory cells, which protect against autoimmunity [28]. SH2 adaptor protein 3 (SH2B3) is involved in cytokine receptor signaling and cellular transformation and has been associated with the development of several autoimmune diseases [22, 29, 30]. *PDE8B* gene is associated with TSH levels and modulation of thyroid physiology [30]. A splice variant of thyroglobulin was found to be present in almost all affected individuals in a family in which the disease showed an autosomal dominant inheritance with an early disease onset [31]. *TPO*, *ATXN2*, *RASGRP1* have been associated with TPO antibody levels and thyroid gland volume, thus contributing to disease susceptibility [32].

Spontaneous autoimmune thyroiditis (SAT) in chicken

The Obese strain (OS) of chicken, studied in this thesis, is a strain that develops spontaneous autoimmune thyroiditis (SAT). OS chickens emerged from a White Leghorn line, a layer breed of chicken, when some individuals were observed to deviate from the normal phenotype of the flock [33]. SAT in chicken constitutes an animal disease model for human Hashimoto's thyroiditis. Therefore, SAT in chickens is studied in an attempt to identify the causative gene(s), a step required for the translational research from the OS chicken as an animal model to Hashimoto's thyroiditis. The thyroid glands of SAT-affected chickens are destroyed leading to hypothyroidism. This thyroid destruction results from the infiltration of CD4+ T lymphocytes expressing T-cell receptor α/β (TCRα/β), Interleukin-2 receptor (IL2-R) and MHC-II molecules into the thyroid glands [34, 35]. OS chickens affected with SAT show specific attributes such as long silky feathers, small comb size, small skeletal size, high body weight, lipemia, and obesity (Figure 1) [33]. The genetic background of SAT is unknown and was explored in this thesis in an attempt to identify genomic regions predisposing SAT in chickens.

Figure 1. A male chicken from the OS line showing long silky feathers. Photo taken by Susanne Kerje.

Thyroid stimulating hormone

Thyroid stimulating hormone (TSH), also known as thyrotropin, was discovered in the early 20th century [36]. It consists of two subunits, TSHα and TSHβ that dimerize to form a functional hormone that is secreted by the anterior pituitary gland into the blood stream [37]. TSH binds to its receptor (TSHR) and this interaction is important for many biological functions such as reproduction, brain development and maintaining the body's metabolic rate in different tissues. TSHR is a seven trans-membrane G protein-coupled glycoprotein that is expressed in different tissues and organs, with a particularly high expression in the thyroid gland [38]. In a normal thyroid gland, the binding of TSH to TSHR stimulates and induces the secretion of T3 and T4 hormones that inhibit the anterior pituitary from secreting more TSH, via a negative feedback loop. However, when the thyroid gland is destroyed by thyroiditis, the resulting low levels of T3 and T4 hormones fail to establish the negative feedback loop, thus allowing the anterior pituitary to increase its TSH secretion in an attempt to induce the thyroid to restore normal T3 and T4 levels (Figure 2).

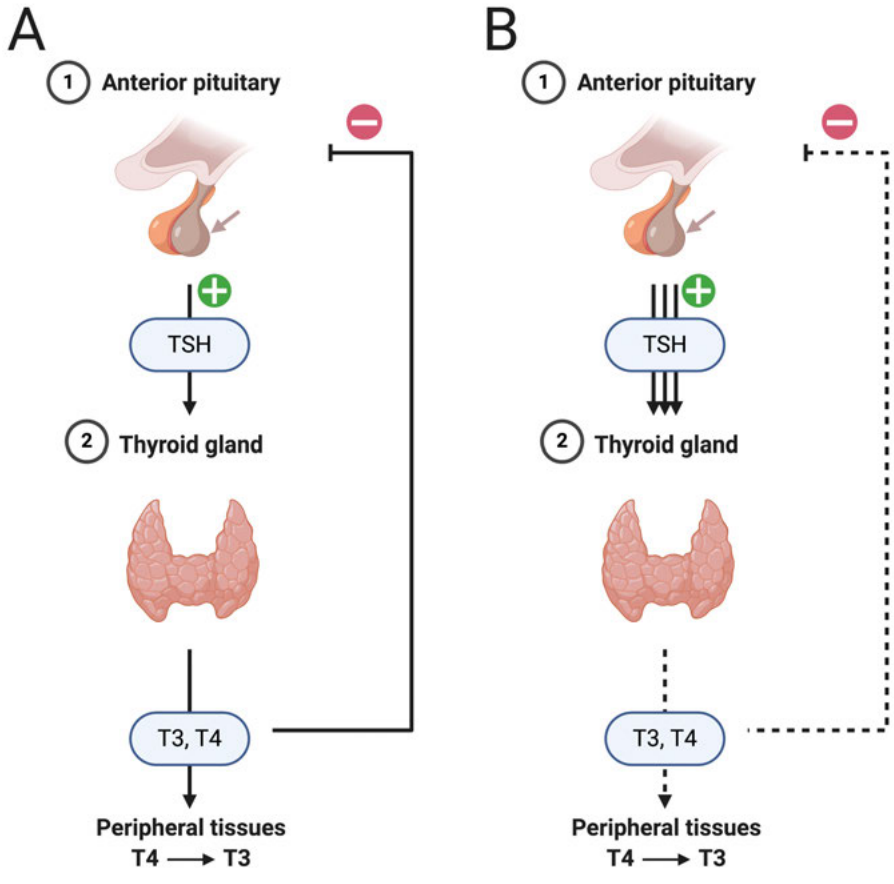

Figure 2. An illustration showing the thyroid gland secreting T3 and T4 hormones in normal levels, which establishes a feedback loop on the anterior pituitary to secrete TSH in normal levels (A). The failure of the thyroid gland to secret T3 and T4 in normal levels, in the case of hypothyroidism, fails to establish the feedback loop, resulting in an elevation in the secretion of TSH by the anterior pituitary (B). The illustration was created with BioRender.com.

Characterization of the *Zbed6-Igf2* genes underlying a major QTL affecting pig muscle growth

The second project in this thesis explores how the insulin-like growth factor 2 (*IGF2*) gene is regulated on the transcription level by the zinc finger BED-type containing 6 (ZBED6) protein and on the translation level by miR483. Moreover, the functional characterization of miR483 is carried out.

IGF2 was first found to have an effect on muscle growth, size of the heart and fat deposition in pigs when a major QTL was identified using an intercross between wild boar and domestic pigs [39]. The QTL was found to be located on chromosome 2 in a genomic region harboring the IGF2 locus, to be paternally expressed, and that the allele coming from the Large White domestic pigs is associated with reduced back-fat thickness and larger muscle mass [40]. A cross between Large White and Piétrain pigs also mapped an imprinted QTL to the *IGF2* locus with effects on fat deposition and muscle mass [41].

An *IGF2* mutation abrogating the binding of a repressor

The characterization of the QTL on pig chromosome 2 revealed a guanine-to-adenine single nucleotide substitution at position 3,072, which is located in an evolutionary conserved CpG island in the third intron of *IGF2* [39]. When this single-base quantitative trait nucleotide (QTN) occurs on the paternally inherited allele, it causes a three-fold increase in the IGF2 mRNA [39] due to the disruption of the recognition and binding of a transcriptional repressor to the mutated site.

Identifying the QTL-related repressor of IGF2 transcription

The G3072A mutation in the third intron of *IGF2* abrogates the binding of a transcription factor that represses the transcription of *IGF2* [39]. To identify this unknown transcription factor, biotinylated wild type (WT) and mutated oligonucleotides were used to capture the protein from C2C12 mouse myoblast cell lysates, which resulted in the identification of the ZBED6 protein by mass spectrometry [42]. This finding was supported by an electrophoretic mobility shift assay (EMSA) experiment that was done to capture regulatory proteins that bind to this location in the *IGF2* gene; and anti-ZBED6 antibodies produced a supershift when added to the oligonucleotide-protein complex [42].

Characterization of ZBED6

ZBED6 is an intron-less domesticated DNA transposon that was integrated within the first intron of the zinc finger CCCH-type containing 11A gene (*ZC3H11A*) in an ancestor of placental mammals [42]. ZBED6 is exclusive to

placental mammals and is a highly conserved protein that shows a nuclear localization and broad expression in different tissues [42]. These data suggested the importance of ZBED6 as a transcriptional regulator. A ChIP-sequencing experiment that revealed around 1,200 annotated genes with putative ZBED6 binding sites (total number of peaks is 2,499); and a consensus sequence of the ZBED6 binding site to be 5'-GCTCGC-3' [42]. A Gene-Ontology (GO) analysis was done for these ~1,200 genes and some of the highly enriched processes that were observed are cellular differentiation, transcriptional regulation, and muscle development [42]. RNA-seq analysis of C2C12 myoblasts overexpressing *ZBED6* was also performed and revealed a down-regulation in genes involved in cell cycle and cell division processes.

Cell cycle analysis

Generally, after the completion of mitosis, cells start a new cell cycle by entering the G0/G1 growth phase where they increase their size and internal cellular organelles and contents, during which a cell has only one pair of each chromosome (2n DNA content). After reaching an optimal size and successfully crossing G1-S checkpoint, cells enter the DNA synthesis phase (S-phase), which is characterized by a variable DNA content in different cells due to the active synthesis and duplication of chromosomes. When DNA synthesis finishes and the cells contain two sets of paired chromosomes (4n DNA content), cells enter the G2/M-phase during which a cell division occurs. Then cells either undergo cell cycle arrest and stay in the G0-phase, similar to what happens to memory T-lymphocytes or differentiated cells, or start a new cell cycle process.

Several methods are available for performing a cell cycle analysis using flow cytometry. Propidium iodide (PI) is a commonly used DNA staining dye, but its broad emission spectrum and its RNA-staining ability make it less favorable, even when RNases are used in the experiment. Therefore, FxCycle Violet stain was preferred over PI because it preferentially stains dsDNA by binding to the minor groove of AT clusters [43]. Moreover, the FxCycle Violet stain is excited by the less commonly used UV laser, keeping the more commonly used lasers, such as the 488 nm laser which is used by many dyes including GFP, available to other dyes. The signal obtained from the FxCycle Violet stain is proportional to the amount of DNA content within a cell and thus accurately places the cells on a linear scale sorting them according to their DNA content.

A second reagent that has been used is 5-ethynyl-2'-deoxyuridine (EdU). EdU is a thymidine analogue that, when added to the culture medium, will be incorporated in the new DNA of the active DNA-synthesizing cells. EdU is then coupled to a photo-stable Alexa Fluor dye in a simple and highly specific click reaction. The incorporation of EdU and its subsequent labeling allows

the detection of DNA-synthesizing cells from cells in either G0/G1 or G2/M phases.

This dual usage of FxCycle Violet stain and Click-iT EdU to perform a cell cycle analysis using flow cytometry allows obtaining a two-dimensional plot which makes it easy to gate/sort different cell populations depending on the cell cycle phase in which they are currently in.

Apoptosis analysis

Apoptosis is composed of three distinct phases common to all cell types [44]. These phases are: the initiation phase initiated by death signals; decision phase where the cell can still decide to abolish the apoptosis process and continue to survive; and execution or "no-return" phase, where effector caspases are hierarchically activated [44]. The execution phase of apoptosis occurs in an orderly manner as follows: i) shrinkage of cells, ii) cell-surface exposure of phosphatidylserine (PS), iii) condensation, margination and degradation of chromatin, iv) nuclear fragmentation, and v) cellular disintegration into apoptotic bodies [44].

As the cells are able to revert undergoing apoptosis before reaching the point of no return, the earliest time point during which apoptosis can be studied is the transition between the decision phase and the execution phase. During this period that precedes nuclear condensation and disintegration, and while cell membrane integrity is still intact [45], cells start losing their plasma membrane asymmetry and externalize phospholipids, such as PS, that are predominantly present in the internal leaflet of the plasma membrane [44, 46]. PS externalization triggers receptor-mediated recognition and elimination of apoptotic PS-exposing cells by macrophages [46], and lasts late until cellular disintegration [47].

PS externalization can be detected using annexin V, which is a calcium-dependent phospholipid binding protein [48] that binds with a high affinity to PS [49] and does not penetrate the plasma membrane [44, 49]. Therefore, it selectively binds to apoptotic and necrotic cells due to their externalization of PS [44, 50]. However, as apoptotic and necrotic cells both externalize PS, a method utilizing annexing V and a DNA-staining dye, such as PI, has been developed and successfully used to differentiate between viable Annexin-V-negative and PI-negative (AnnV-, PI-) cells, cells in early apoptosis (AnnV+, PI-), and cells in late-apoptosis/necrosis (AnnV +, PI+) [45, 50-52]. Moreover, PI seems to be a good DNA stain due to its ability to exclude dye in living cells [53]; and as it cannot penetrate an intact plasma membrane [52].

Although the annexin-V/PI method for studying apoptosis is commonly used, yet it was shown that it leads up to 40% false positive events as a result of PI binding to cytoplasmic RNA [54]. To overcome the drawbacks of this method, a more accurate modified protocol has been proposed [54], that shows < 5% events with false-positive PI staining of the cytoplasm [54, 55]. This

method was modified to be applicable to adherent cells [56]; however, it requires using a flow cytometry imaging technique to discriminate between the cytoplasmic false-positive PI-signals and the true nuclear signals [55]. Another option is using DRAQ7 instead of PI as a DNA dye as it does not bind RNA and selectively binds DNA, is not able to penetrate the cytoplasmic membrane, and has a much narrower far red emission signal compared to PI.

Apoptosis can be evaluated by other methods such as measuring caspase activity by western blot or enzymatic assays; measuring translocation of mitochondrial proteins into the cytoplasm; or by estimating DNA fragmentation [44]. However, the common drawback of these methods is that they rely on not-so-early events in apoptosis. Therefore, flow cytometry with annexin V and DRAQ7 stains was used to analyze apoptosis in C2C12 myoblasts transiently transfected to overexpress GFP-tagged *ZBED6* and its GFP control.

Investigating the effect of miR483 on the expression of *IGF2*

The *miR483* gene resides in an intron of *IGF2* gene itself, enhances its own transcription and the transcription of *IGF2*, and is linked to P53 expression and different cancers [73-77]. *MiR483* is transcribed and processed to form the premature pre-miR483 hairpin. Maturation of this double-stranded pre-miR483 hairpin occurs when the endoribonuclease dicer enzyme cleaves it into two mature miRNAs, namely miR483-3p and miR483-5p derived from the 3'- and 5'-ends, respectively.

MiR483 is co-expressed with its host gene [57] and is activated by the transcriptional activator USF1 [58, 59] that acts as a mediator between the oncoprotein β-catenin (CTNNβ1) and the miR483 locus by forming a complex that binds to the E-box element with motif sequence being CACGTG upstream of the miR483 gene [60]. The binding of the CTNNB1/USF1 complex to the E-Box element is regulated by N-acetyl-glucosamine transferase (OGT) in a glucose dependent manner; as the inhibition of OGT by the addition of azaserine, its knockdown by siRNA, or inducing glucose deprivation by 2-deoxy-D-glucose (2-DG) decreases the binding affinity of the CTNNB1/USF1 complex to the E-box element [59]. Moreover, miR483 can self-regulate its expression by indirectly enhancing the expression of USF1, which in turn stabilizes the transcriptional complex on the miR483 locus [60].

MiR483-3p is overexpressed in many cancers, such as human colorectal cancer (HCC) [57]. Deleted in liver cancer 1 (DLC-1), a candidate tumor repressor, was shown to be a direct target of miR483-3p and that its protein level was significantly reduced in HCC tissues compared to matched adjacent normal controls [57]. Similar to DLC-1, breast cancer 1 (BRCA-1) is another tumor repressor that was also found to have a target site for miR483-3p in its 3'-UTR [61]. Moreover, a meta-analysis study comparing human colorectal cancers that have WT or mutated tumor protein 53 (P53) found that

miR483-3p is significantly upregulated in WT-P53 HCCs [62]. These oncogenic effects of miR483-3p were further supported by finding that it inhibits P53-dependent apoptosis by targeting the P53-upregulated-modulator-of-apoptosis (PUMA) gene [63].

MiR483-5p has been found to enhance the association of the DHX9 RNA helicase with the 5'-UTR of the *IGF2* mRNA transcripts, thus promoting their transcription [64]. This is supported by the increase in *IGF2* mRNA levels upon the transient transfection of a miR483-5p mimic in different cell lines (human-MMH-ES-1 and RD, and mouse-F4328 and F4864 cells) [64]. Moreover, miR483-5p increases tumorigenesis upon injecting MMH-ES-1 cells that stably express miR483-5p in nude mice [64]. And although the expression levels of IGF2, miR483-3p and miR483-5p increase in colorectal cancer (CRC) tissues relative to matched adjacent normal tissues, the serum levels of miR483-5p significantly increase in CRC patients compared to normal controls; to the extent that it is suggested to be used as a marker for the early detection of CRC [57]. Furthermore, miR483-5p was found to be an accurate marker to categorize tumors as benign or malignant [65].

These oncogenic characteristics of miR483-3p and miR483-5p, and their relation with *IGF2* expression triggered our interest in studying their relation with ZBED6 especially in *IGF2*-knock-in C2C12 myoblasts that show a high expression of *IGF2*.

Aims of the thesis

I The aim of the project on autoimmunity was to develop a cell-based assay as an efficient tool for measuring TSH levels in chicken serum (Paper I), and to perform a whole genome sequencing experiment to identify genomic regions co-segregating with SAT in chickens which might aid in the identification of genomic regions associated with Hashimoto's thyroiditis in humans (Paper II).

II The aim of the second project is to explore how the transcription factor ZBED6 regulates the expression of *Igf2* (Paper III). Furthermore, we wanted to study the effect of how ZBED6 regulates the expression of miRNAs in general and miR483 in particular (Paper IV).

Present investigations

Spontaneous autoimmune thyroiditis in the Obese strain of chickens as an animal model for human Hashimoto's thyroiditis: Development of a phenotyping assay and dissection of the genetic background (Papers I and II)

Spontaneous autoimmune thyroiditis (SAT) is an organ-specific autoimmune disease that spontaneously develops in the Obese strain (OS) of chickens, resulting in the destruction of the thyroid gland. This destruction causes hypothyroidism, which is expected to cause an increase in the secretion of the TSH hormone from the anterior pituitary gland. The genetic and pathophysiological background of SAT in chicken, which largely resembles human Hashimoto's thyroiditis, are unknown. However, the spontaneous nature of the chicken and human diseases makes SAT in the OS chickens a good animal model for studying human Hashimoto's thyroiditis.

To study SAT, the OS chickens were crossed to their wild ancestor, the red jungle fowl (RJF) to generate a nine-generation pedigree by intercrossing generations F_1 through F_8. A cell-based luciferase assay that estimates the TSH levels in chicken serum was developed to phenotype the chickens at different ages without the need to sacrifice the animals. Moreover, individuals of the F_9 generation were also phenotyped by measuring the degree of infiltration of lymphocytes into their thyroid glands. Four groups of chicken with different degrees of thyroid infiltration were chosen and used in a whole genome sequencing experiment to identify genomic regions associated with the susceptibility to the disease.

Paper I: Development of a cell-based luciferase assay for estimating TSH hormone levels in chicken serum

Background

During thyroiditis, the immune system mounts a reaction against the thyroid gland which results in the infiltration of lymphocytes into the gland leading to its destruction. As the disease progresses and the infiltration of the thyroid increases, its ability to secret T3 and T4 hormones decreases. This decrease in the levels of T3 and T4 hormones induces the pituitary gland to increase its production of TSH which in turn induces the undamaged thyroid cells to increase their production of thyroidal hormones. The increasing levels of TSH succeed in maintaining the T3 and T4 at normal levels until a late stage of thyroid infiltration, when a large proportion of the gland is damaged. Therefore, measuring T3 and T4 might not be a good indication for the severity of the disease, whereas the level of TSH is a good and accurate indicator.

Aim

The plan was to develop a TSHR-Luciferase cell-based assay to be able to phenotype chickens as healthy or affected by SAT and to study the progression of SAT in the affected chickens by measuring serum TSH levels at different time points of age.

Results and discussion

A TSHR-Luciferase cell-based assay was developed by generating TSHR-expressing cAMP-responsive luciferase-reporter cells (Figure 3). The cells respond to the amount of TSH in the serum added to the cell culture medium by activating the TSHR. The activated receptor induces the synthesis of cAMP which accumulates in the cells. The cAMP accumulation induces the transcription of luciferase and finally the luciferase activity, which is proportional to the amount of TSH in the serum, can be measured on a plate reader.

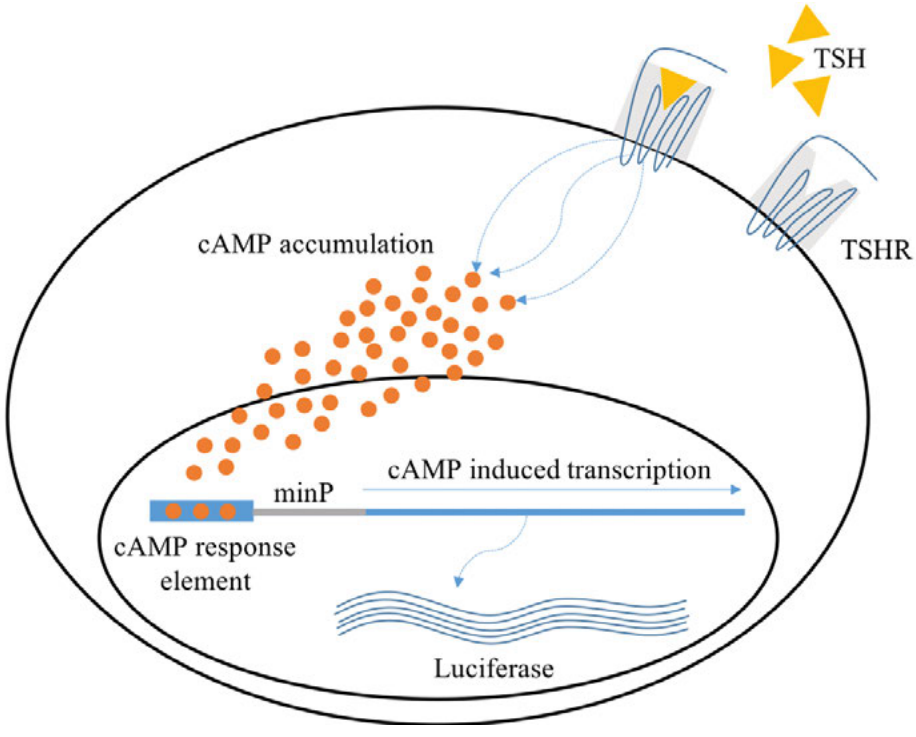

Figure 3. A cell-based luciferase assay was developed to estimate the levels of TSH in chicken serum. Upon the addition of chicken serum to the culture medium of the cells, the TSH present in the serum binds to the TSHR that is expressed on the cell surface. This receptor-ligand binding activates the receptor that converts ATP to cAMP which accumulates in the cells. A cAMP-response-element (CRE) is then activated by the increasing amounts of cAMP, resulting in the induction of luciferase transcription. The level of luciferase is measured by adding the luciferin substrate to start a reaction that emits luminescence. Finally, luminescence is measured on a luminometer to estimate the luciferase activity level, which is proportional to the amount of added serum TSH.

The inducibility of the transcription of luciferase was confirmed by the addition of forskolin, a plant extract that is a potent activator of adenylyl cyclase, to the cell culture medium (Figure 4A). Thereafter, to test the ability of activating the assay through the TSH-TSHR binding, the cells were incubated in cell culture medium with 0-20% chicken serum (Figure 4B). This showed that the two parts of the assay, the induction of cAMP synthesis through activating the TSHR and the subsequent induction of luciferase transcription, were functional. Moreover, it was observed that the luciferase production reached a plateau upon adding 10% chicken serum to the culture medium, indicating that 10% chicken serum is the optimal amount to be used.

Furthermore, although the addition of either OS or RJF serum significantly induced the luciferase expression in TSHR-expressing cells, the OS serum resulted in higher induction as compared to the RJF serum, indicating that the TSH levels in OS chickens is higher than in RJF chickens (Figure 4C).

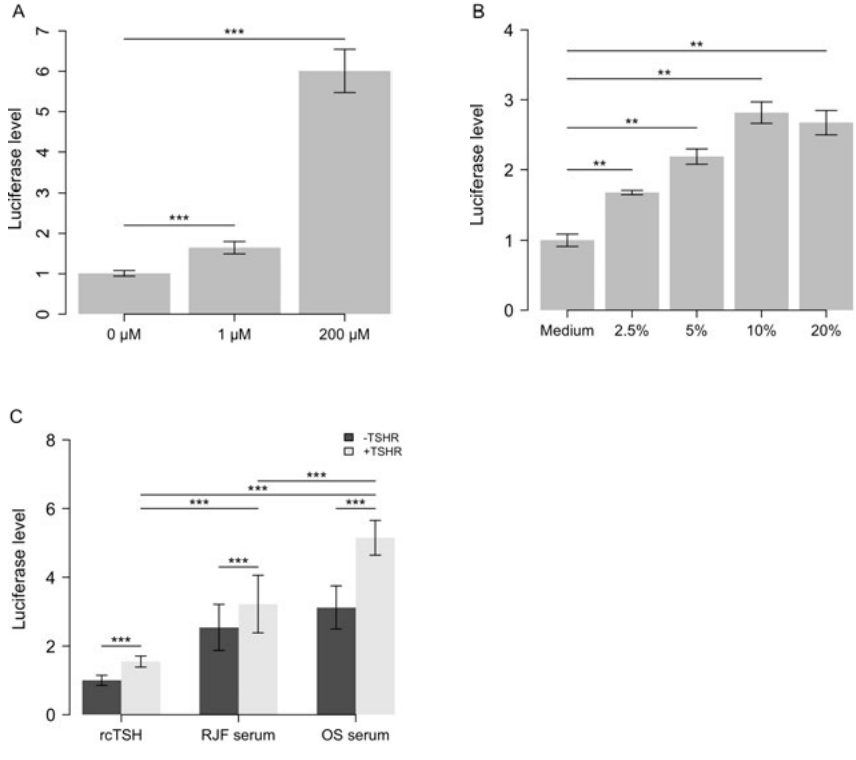

Figure 4. Induction of luciferase production in the cAMP-responsive luciferase-reporter cells by different amounts of forskolin (A). Induction of TSHR-expressing cAMP-responsive luciferase-reporter cells by increasing amounts of chicken serum (B). Induction of cAMP-responsive luciferase-reporter cells either expressing TSHR (+TSHR) or not expressing it (-TSHR) by recombinant chicken TSH (rcTSH), RJF serum and OS chicken serum (C).

Finally, the amount of TSH in the F_9 chickens at 10, 20 and 28 weeks of age, was estimated by adding the serum from 104 chickens to the cell culture medium (Figure 5). The correlations between the degree of infiltration of the thyroid glands and the TSH levels at the three ages were found to be 0.03 (P=0.79), 0.05 (P=0.6) and 0.72 (P <0.001) for the age of 10, 20 and 28 weeks, respectively. This shows that the correlation was highest at 28 weeks of age whereas there was no correlation between thyroid infiltration and the increase in TSH at the ages of 10 or 20 weeks, thus indicating that the increase in TSH levels starts between weeks 20 and 28.

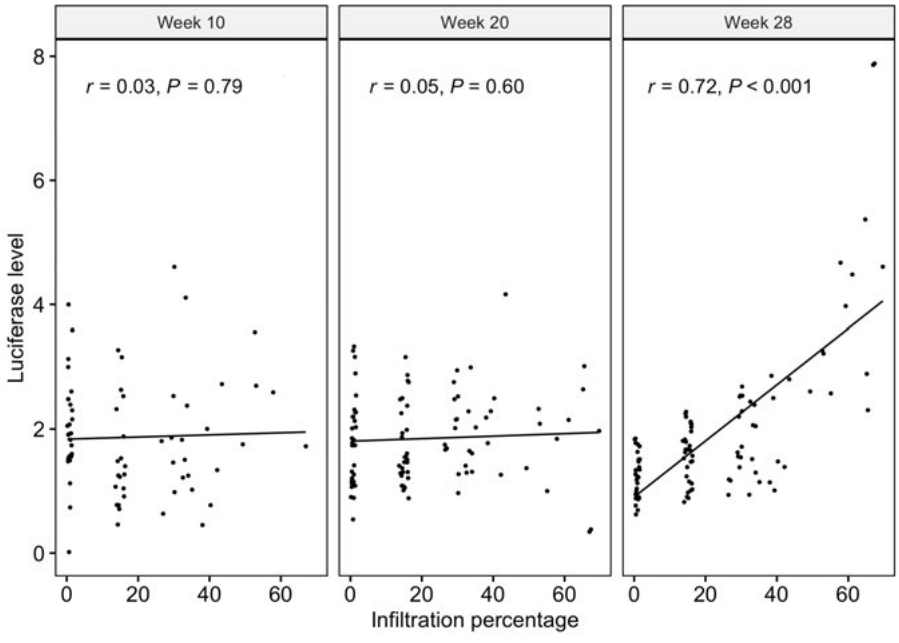

Figure 5. Correlation of the serum TSH levels and the degree of thyroid gland infiltration of F_9 individuals from the OS x RJF cross at different ages. Luciferase levels were normalized against the FBS-free medium control sample. Pearson's correlation and the respective P-values are presented.

Conclusions

The TSHR-expressing cAMP-responsive luciferase-reporter cell-based system developed in this study was used to estimate the levels of TSH in chicken serum. Moreover, the levels of TSH were shown to increase after week 20 in chickens affected by SAT in comparison with healthy chickens. These results indicate that in SAT-affected chickens, the levels of serum-TSH increase in a manner similar to that of human Hashimoto's thyroiditis.

Paper II: A genomic screen for loci causing spontaneous autoimmune thyroiditis in the Obese strain of chickens

Background

Hashimoto's thyroiditis is a polygenic autoimmune disease that affects humans with a women:men ratio of 10:1, that are usually diagnosed at 30-50 years of age [66, 67]. Hashimoto's thyroiditis patients are positive for anti-thyroid peroxidase (anti-TPO) and anti-thyroglobulin (anti-Tg) auto-antibodies [23]. The presence of these antibodies indicates an ongoing autoimmune reaction against the thyroid gland, the tissue that expresses TPO and Tg. This immune reaction causes a lymphocytic infiltration into the thyroid gland which leads to its destruction and eventually leads to hypothyroidism.

Spontaneous autoimmune thyroiditis (SAT) occurs in the Obese strain (OS) of chickens [33], which is a model for human Hashimoto's thyroiditis making OS chickens a good animal model for studying Hashimoto's thyroiditis. The destruction of the thyroid glands in the OS chickens starts after a few weeks of age [68] which leads to hypothyroidism and causes the OS chickens to have long silky feathers, be cold-sensitive and have a slightly smaller skeletal size [33].

Aim

To investigate the genetic basis underlying SAT using a nine-generation intercross between OS and RJF chickens [69] and a whole genome sequencing approach.

Results and discussion

In an attempt to identify genomic loci co-segregating with SAT, OS and RJF chickens were crossed to generate a nine-generation pedigree [69]. To phenotype the F_9 chickens as healthy or affected by SAT, their thyroid glands were dissected, sectioned and stained with hematoxylin and eosin, to score the degree of lymphocytic infiltration into the glands. An intact thyroid gland is characterized by regular and circular shaped follicles, known as colloids, and a pink color due to the eosin staining of prohormones contained within these colloids. An infiltrated thyroid gland is characterized by an irregular shape of the colloids due to the lymphocytic infiltration, and a dark purple color due to the hematoxylin staining of the nuclei of the infiltrating lymphocytes. These images were used to phenotype chickens as healthy (Figure 6A), suffering from mild thyroiditis (Figure 6B) or severe thyroiditis (Figure 6C).

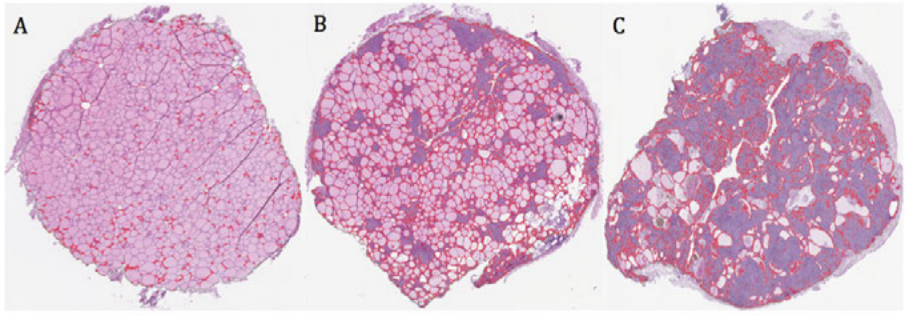

Figure 6. Images of hematoxylin and eosin stained thyroid glands of 28-week old chickens from the OS x RJF cross. Normal healthy thyroid gland (A). Medium infiltration of T-lymphocytes into the thyroid gland causing mild damage (B). Severe infiltration of T-lymphocytes into the thyroid gland leading to almost total destruction (C).

Furthermore, there was no difference in the degree of lymphocytic infiltration into the thyroid glands between males and females (Figure 7). This is in contrast to human Hashimoto's thyroiditis that affects more women than men with a ratio of 10:1.

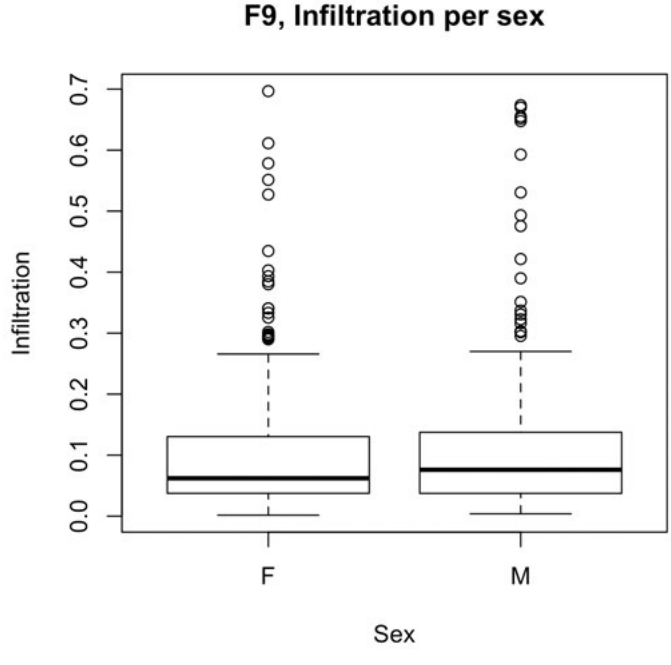

Figure 7. A boxplot showing the difference between males and females Degree of lymphocytic infiltration into the thyroid glands of F9 individuals from the OS x RJF cross in the

Thereafter, from 852 individuals in the F_9 generation of the OS x RJF cross, four phenotypic chicken groups were chosen and used in a whole genome sequencing experiment. The degree of lymphocytic infiltration in the thyroid glands of the four groups were <1.6% (30 individuals), 13-17% (30 individuals), 26-43% (30 individuals) and 47-70% (14 individuals). The analysis of the whole genome sequencing data comparing the most extreme groups, in which Fst values were calculated revealed a total of 16 regions to be significantly different between the healthies and the most affected groups (Figure 8). However, four regions were manually identified as artifacts due to a poor indel realignment process in highly repetitive regions and regions with a very low complexity. The remaining 12 regions were spread over 9 autosomes and harbor or exist in the vicinity of many genes, some of which are related to the immune system. One example is region 12 that exists on chromosome 31 upstream of the immunoglobulin heavy chain locus. However, most of the regions are wide and thus further investigation is needed to identify SAT-associated genes.

This finding was in contrast to our anticipation, as we expected SAT to be caused by more than one gene [33] and to involve a minimum of three genetic loci [70], but not as many as 12 regions. These deductions were based on selective intercross and backcross experiments that were previously performed [68] and where a few generations were sufficient for many affected individuals to appear [71].

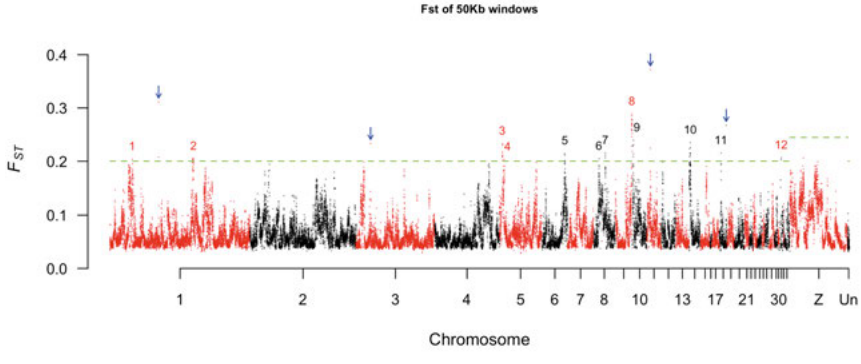

Figure 8. Manhattan plot based on SNP data from whole genome pooled sequencing. Average F_{ST} values between the healthiest (pool 1) and most sick (pool 4) pools based on 50 kb windows with steps of 25 kb for the autosomes, chromosome Z and unassigned scaffolds (Un) are plotted. The green horizontal lines represent the significance threshold levels for the autosomes and chromosome Z. The 12 regions showing genome-wide significance are numbered 1-12. Four other regions, indicated by blue arrows, passed the significance threshold but were excluded because they were identified as artifacts due to difficult alignment of highly repetitive or low complexity regions.

Conclusion

Our results support the idea that SAT is similar to Hashimoto's thyroiditis in its polygenic nature as we identified a total of 12 genomic loci to contribute to SAT in the studied pedigree. Moreover, despite of the fact that we have used F_9 generation to increase recombination, which is needed for the fine-mapping of SAT-associated regions, we still have broad regions with multiple genes, indicating that this experimental design has not allowed us to identify SAT-associated genes, simply because SAT seems to be more polygenic than we initially thought.

Transcriptional regulation of IGF2 by ZBED6 and miR483 (Papers III and IV)

Insulin-like growth factor 2 (IGF2) is an essential growth factor that plays an important role in fetal growth, initiation of myoblast differentiation, the regulation of muscle mass, fat deposition and the size of internal organs [41, 72, 73]. Its expression has a complex regulation where, at the genomic level, only the paternal allele gets expressed [74, 75].

The expression of *Igf2* was also found to be repressed by ZBED6 which is a transcriptional repressor that binds to a specific motif (5'-GCTCGC-3') in an intron of *Igf2* [42]. This mechanism for regulating *Igf2* transcription was discovered following the identification of a single base substitution in an intron of *Igf2,* which disrupts the binding of ZBED6 to the motif and upregulates *Igf2* expression in domestic pigs [39, 42].

Igf2 knock-in mice (IGF2-KI) carrying the same mutation, and *Zbed6* knock-out mice (ZBED6-KO) were generated and showed an upregulation of IGF2 expression up to >10-folds and 8-folds at the mRNA and protein levels, respectively [42, 72]. The generated mice also exhibited increased skeletal muscle growth and body weight [72].

Moreover, *miR483*, a gene hosted in an intron of *Igf2,* was found to regulate its expression by the binding of miR483-5p to the 5'-UTR of the *Igf2* transcripts to stabilize them and promote their transcription [64]. Additionally, the *miR483* gene was shown to self-regulate its own expression by indirectly enhancing the expression of USF1, which in turn stabilizes the transcriptional complex on the *miR483* locus [60], and that the *miR483* and *Igf2* genes were found to be co-expressed [57].

Paper III: The importance of the ZBED6-*Igf2* axis for metabolic regulation in mouse myoblast cells

Background

ZBED6 is a transcription factor that was first identified as a repressor of *Igf2* [39, 42]. It was later found to have thousands of binding sites near transcription start sites and to modulate the expression of myogenic genes, regulate cell aggregation, neuronal development and the production of insulin [76, 77]. Moreover, *Igf2*, one of the main targets of ZBED6, was found to play an important role in myoblast differentiation and muscle development [73]. However, how ZBED6 impacts the phenotype by regulating *Igf2* expression during myogenesis is not yet known.

Aim

The aim is to investigate the functional significance of ZBED6 and its effect on transcriptional regulation in general, and to explore the interaction between ZBED6 and *Igf2* during myogenesis in specific. Therefore, we generated a *Zbed6* knockout (*Zbed6*$^{-/-}$) and a deletion in the ZBED6 binding site in an *Igf2* intron (*Igf2*^{dGGCT}) in C2C12 myoblast cells. To establish the mentioned cell types, the C2C12 mouse myoblast cell line was used due to its ability to differentiate into mature myotubes [78]. The generated *Zbed6*$^{-/-}$ and *Igf2*^{dGGCT} cell types were induced to differentiate and subjected to transcriptome and proteome analysis.

Results and discussion

To explore how ZBED6 affects *Igf2*, our initial plan was to establish C2C12 cells that stably overexpress ZBED6 (ZBED6-OE). However, the multiple unsuccessful attempts to generate the stable cells indicated the lethality of ZBED6-OE. Therefore, the inducible "Tet-on" system was used. However, even though this system is thought to be one of the tightest available systems, the generation of ZBED6-OE cells was still unsuccessful, probably due a small leakage in the system and the high lethality of ZBED6-OE.

Thereafter, a 2.5 kb region was deleted from *Zbed6* to generate *Zbed6*$^{-/-}$ C2C12 mouse myoblasts in which *Zbed6* is completely inactivated. Moreover, the binding motif of ZBED6 in an intron of *Igf2* was mutated by deleting four nucleotides (GGCT) from the 5'-GGCTCG-3' ZBED6 binding motif, generating *Igf2*^{dGGCT} cells. The two mutations generated in C2C12 myoblast cells were found to cause a similar phenotype, as they caused the cells to have >30-fold upregulation of *Igf2* mRNA, exhibit a faster growth rate and develop myotube hypertrophy compared to the control WT C2C12 cells. These observations were opposite to the lethality of ZBED6-OE, and indicate that ZBED6 exerts its main effect though regulating *Igf2*.

Furthermore, proteome and transcriptome analyses of *Zbed6*$^{-/-}$, *Igf2*^{dGGCT} and WT C2C12 cells were performed. The proteome analysis of the mutant

myoblasts identified 1167 and 570 DE proteins in *Zbed6*$^{-/-}$ and *Igf2*^{dGGCT}, respectively. Whereas the transcriptome analysis identified around ~3000 and ~2500 DE genes in *Zbed6*$^{-/-}$ and *Igf2*^{dGGCT} myoblasts. Integrating the proteome and transcriptome data revealed 381 and 196 genes to be DE at both of the protein and mRNA levels in *Zbed6*$^{-/-}$ and *Igf2*^{dGGCT}, respectively (Figure 9).

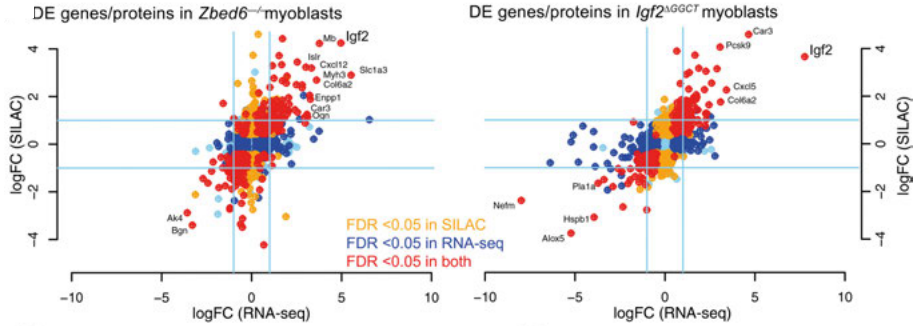

Figure 9. Proteome (SILAC) and transcriptome (RNA-seq) analyses of *Zbed6*$^{-/-}$ (left) and *Igf2*^{dGGCT} (right) myoblasts presented as log fold change (logFC) to WT cells.

KEGG pathway analysis of DE proteins unique for the *ZBED6*$^{-/-}$ cells revealed an enrichment of mitochondrial membrane proteins. This was also shown in the analysis of live mitochondria that revealed an increase in mitochondrial mass and oxidation rate, and a decrease in extracellular acidification rate in *Zbed6*$^{-/-}$ cells. In contrast, a significant reduction of mitochondrial mass and mitochondrial membrane potential was observed in transient ZBED6 overexpressing (ZBED6-OE) cells.

The results of the transcriptome analyses revealed an opposite trend in the expression of genes involved in cell division and regulation between *Zbed6*$^{-/-}$ and *Igf2*^{dGGCT} cells on one hand, and ZBED6-OE cells on the other hand. These gene were upregulated in *Zbed6*$^{-/-}$ and *Igf2*^{dGGCT}, and downregulated in ZBED6-OE cells, explaining the opposite phenotypes observed in the mutant cells.

Therefore, we analyzed the cell cycle and apoptosis in ZBED6-OE C2C12 myoblasts. Cell cycle analysis confirmed the transcriptome analysis and revealed that 82% of ZBED6-OE cells appeared to be in the G0/G1 phase and 10% in the S-phase, compared to 58% and 35% in control cells, respectively (Figure 10). Whereas apoptosis analysis revealed a significant reduction in the number of live ZBED6-OE cells.

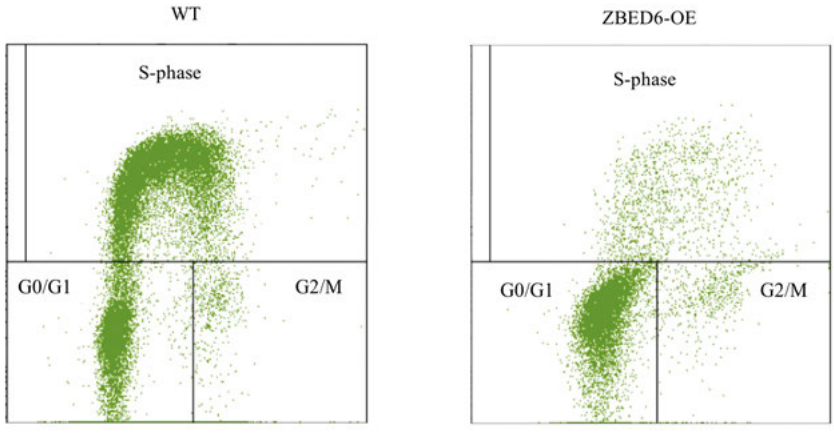

Figure 10. Distribution of cells in the different cell cycle phases analyzed using flow cytometry of ZBED6-OE cells compared to WT cells.

Conclusion

Our results confirm that ZBED6 is an important transcription factor with a large impact on cellular processes. Moreover, ZBED6 was found have a direct effect on mitochondrial activity and to affect different cellular functions mainly through regulating the expression of *Igf2*. Furthermore, ZBED6-OE was found to be lethal and to induce cell cycle arrest and apoptosis. Whereas ZBED6 inactivation inn ZBED$^{-/-}$ and *Igf2*^{dGGCT} induced faster growth rate and hypertrophy. These observations indicate that ZBED6-OE seems to be more deleterious than knocking it out.

Paper IV: ZBED6 affects the expression of the oncogenic miR483 through its interaction with the *Igf2* locus

Background

Insulin-like growth factor 2 (*Igf2*) is a paternally expressed gene that shows a loss of imprinting and biallelic expression leading to its upregulation in several tumors [59]. *miR483*, a gene hosted in an intron of *Igf2* and co-expressed with it, can self-regulate its own expression [60], and plays a prominent role for enhanced *Igf2* expression in tumors [57]. Moreover, miR483-3p and miR483-5p, the two miRNA products of *miR483*, are overexpressed in many cancers [57, 62]. miR483-3p was found to target candidate tumor repressors [57, 61, 63]. Whereas miR483-5p was found to increase tumorigenesis [64]. Interestingly, it has been reported that miR483-5p binds directly to the 5'UTR region of *Igf2* and thereby enhances *Igf2* transcription [64].

Aim

We aim to investigate whether ZBED6 has an effect on miRNA expression in general and *miR483* in particular using *Zbed6*-KO and *Igf2*-KI mice. Additionally, we explore how *miR483* affects the expression of *Igf2* by knocking it out from $Igf2^{+/+}$ and $Igf2^{dGGCT}$ C2C12 mouse myoblasts.

Results and discussion

A miRNA-seq screen using kidney, liver and skeletal muscle tissues from *Zbed6*-KO, *Igf2*-KI and WT mice was performed. Although ZBED6 has a well-established role as a transcriptional regulator with an effect on several cellular processes [79], the results show that it is not a general regulator of miRNA expression as it has a minimal effect on the expression of miRNAs with the exception of miR483-3p, miR483-5p and miR3105-3p in muscle and kidney tissues. Moreover, the co-overexpression of *Igf2* and *miR483* in *Zbed6*-KO and *Igf2*-KI mice indicated that ZBED6 acts as a repressor for both genes. The expression of miR483-5p and miR483-3p in liver was unaltered in both mouse models. This is in line with a previous study in pigs that shows that ZBED6 inactivation did not affect the expression of *Igf2* in liver.

The previously established $Igf2^{dGGCT}$ C2C12 myoblast cell line [79] that carries a four base pairs (GGCT) deletion of the ZBED6 binding site in an intron of *Igf2* was shown to have a 12-fold increase of miR483-3p expression and more than two-fold increase in miR483-5p expression compared to $Igf2^{+/+}$ C2C12 cells. Transient transfection of $Igf2^{+/+}$ and $Igf2^{dGGCT}$ with miR483-5p inhibitors caused a significant downregulation of *Igf2* expression.

Thereafter, $Igf2^{+/+}$ and $Igf2^{dGGCT}$ cells were used to knock-out the *miR483* gene generating two $miR483^{-/-}$ clones in which *Igf2* was shown to be significantly downregulated in comparison with the $miR483^{+/+}$ controls both on the $Igf2^{+/+}$ and the $Igf2^{dGGCT}$ backgrounds. This confirms the dependence of *Igf2* expression on the expression of *miR483* [57] probably due to the stabilization

of the *Igf2* transcripts by miR483-5p [64]. Transient transfection of the *miR483*$^{-/-}$ clones with miR483-3p and miR483-5p mimics induced an upregulation of *Igf2* expression. Moreover, the dramatic increase in the expression of *Igf2* in the *Igf2*^{dGGCT}*_miR483*$^{+/+}$ in comparison to the *Igf2*$^{+/+}$*_miR483*$^{+/+}$, and the intermediate expression of *Igf2* in the *Igf2*^{dGGCT}*_miR483*$^{-/-}$ indicates that ZBED6 regulates *Igf2* partially through regulating *miR483*.

Furthermore, the generated *Igf2*$^{+/+}$*_miR483*$^{-/-}$ and *Igf2*^{dGGCT}*_miR483*$^{-/-}$ clones were used in an RNA-seq experiment. The *Igf2*^{dGGCT}/*Igf2*$^{+/+}$ contrast performed on a *miR483*$^{+/+}$ background showed >1000 statistically significant, differentially expressed (DE) genes with *Igf2* showing a high fold-change (Figure 11).

Additionally, the *miR483*$^{-/-}$/*miR483*$^{+/+}$ contrast, performed on an *Igf2*$^{+/+}$ background, showed 182 DE genes with only 96 genes being unique to the mentioned contrast. Similarly, the *miR483*$^{-/-}$/*miR483*$^{+/+}$ contrast, performed on an *Igf2*^{dGGCT} background, showed 164 DE genes with only 40 genes being unique to the mentioned contrast (Figure 11). These 96 and 40 genes unique for the *miR483*$^{-/-}$/*miR483*$^{+/+}$ contrasts included 43 and 24 genes with a $\log_2$ fold-change greater than one in the *Igf2*^{dGGCT}/*Igf2*$^{+/+}$ contrast but did not reach statistical significance probably due to a limited statistical power. This observation indicates that the major function of *miR483* in C2C12 myoblasts in to regulate the expression of *Igf2*.

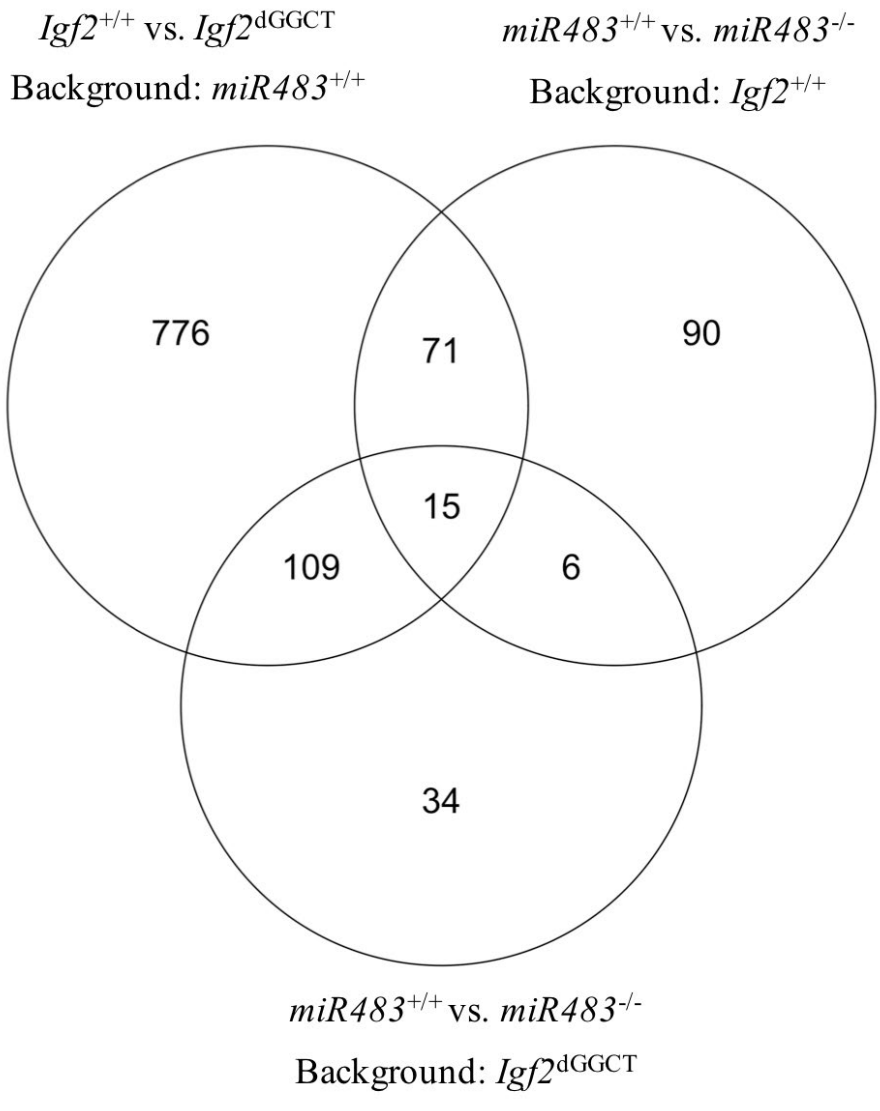

Figure 11. Venn diagram showing the distribution of statistically significant differentially expressed genes with at least one-fold change in one or more of the three contrasts.

Conclusion

Our results indicate that *miR483* is an important component of the ZBED6-*Igf2* axis, where miR483-5p enhances *Igf2* expression in mouse C2C12 cells and that ZBED6 suppresses *Igf2* expression partially by suppressing *miR483* expression. This, combined with the well-established association between up-regulated expression of miR483 and tumorigenesis [57, 62-65] implies that disruption of the ZBED6-*Igf2* axis can promote tumor growth in cell types where ZBED6 suppresses *miR483* expression.

Concluding remarks and future perspectives

This thesis presents a new insight into the genetic basis and the onset time of SAT in chickens. Furthermore, it provides an understanding of the functional significance of ZBED6 and its interaction with the *Igf2* and *miR483* genes.

In paper I, we developed for the first time an assay to measure the levels of TSH in chicken serum as means to phenotype the chickens as healthy or affected by SAT. The developed assay was successfully used to measure the serum TSH levels in F_9 individuals of an OS x RJF cross. Furthermore, the levels of TSH in chicken serum were found to be correlated with the degree of infiltration of the thyroid glands tested at 28 weeks of age. This assay provides an efficient tool for the diagnosis of SAT in chicken serum in future experiments without the need to sacrifice the animals.

In paper II, we aimed to identify the genomic mechanism underlying SAT in chickens. Therefore, we used a nine-generation intercross between the Obese strain of chicken, where almost all individuals develop SAT, and their wild ancestor, the red junglefowl, which is essentially disease-free. The F_9 individuals from this cross were phenotyped using histological sections of their thyroid glands. The phenotype data was used to choose groups of chicken, suffering from different degrees of SAT, to be used in a whole genome sequencing experiment. The results indicate that SAT is a highly polygenic disease, as we identified 12 genomic loci to be significantly associated with the disease, which is in agreement with the genetic background of human Hashimoto's thyroiditis.

In the future, we plan to perform further investigations of the 12 regions in general, and region 12 which exists on chromosome 31 upstream of the immunoglobulin heavy chain locus in specific. Furthermore, one option to fine map the regions associated with SAT is to increase the population size. This is because out of >800 F_9 individuals in the current population, only 14 individuals were found to have >47% lymphocytic infiltration of the thyroid gland. Thus, increasing the population size will increase the number of individuals severely affected with SAT, which increases the statistical power of the analysis and aids in the mapping process.

Alternatively, a genome-wide association study (GWAS) represents a good approach, where we would collect a large number of samples from unrelated SAT-affected chickens in a certain breed and compare them with healthy controls.

Furthermore, the availability of enormous databases of genetic variants in humans, that were built during the last few years, caused a decline in the importance of animal models for studying human diseases. This is because such databases of genetic variants directly associated with disease, can now be used to directly perform GWAS studies on human samples derived from thousands of patients and controls

The second project in this thesis revolved around the functional characterization of ZBED6, its effect on the general regulation of transcription, and more specifically on regulating the transcription of the *Igf2* gene and its intronic miRNA, miR483.

Paper III of this thesis, dealt with the functional characterization of ZBED6, which is a transcription factor that regulates muscle growth and the size of internal organs in placental mammals. We investigated the mechanisms of how ZBED6 affects the metabolic regulation in C2C12 mouse myoblasts. Therefore, we generated *Zbed6*-KO and *Igf2*^{dGGCT} cells, the latter carrying a four-base deletion in the ZBED6-binding site in *Igf2*. The mutant clones were used to perform transcriptome and proteome experiments. These analyses revealed that ZBED6 mainly functions through modulating the expression of *Igf2*, but also has a direct effect on mitochondrial activity. Moreover, the study showed that the overexpression of ZBED6 has a more dramatic effect than ZBED6 inactivation as it induces cell cycle arrest, reduces mitochondrial activity, halts myoblast differentiation and induces apoptosis in C2C12 mouse myoblasts.

Paper IV in this thesis expanded our functional characterization of ZBED6 into investigating its effect on the expression of miRNAs. We used *Igf2*-KI mice, where the ZBED6 binding site in the *Igf2* gene is mutated, and *Zbed6*-KO mice. A miRNA-seq experiment using muscle, kidney and liver tissues from the mutant transgenic mice was performed and revealed that ZBED6 is not a regulator of miRNA expression. However, one notable exception is the regulation of miR483, which exists in an intron of *Igf2* and is associated with the development of several tumors.

Thereafter, we used the *Igf2*^{dGGCT} cells, developed in paper III, and WT C2C12 cells to generate miR483$^{-/-}$ cells. These cells were used to explore the ZBED6-IGF2-miR483 interaction. Our investigation revealed a positive feedback between miR483 and *Igf2* in C2C12 myoblasts. And that ZBED6 regulates the expression of *Igf2* partially through regulating the expression of *miR483*. This finding is supported by the fact that ZBED6 plays an important role in the regulation of *Igf2* in muscle cells, where *miR483* is expressed, but not an important role in liver cells, where *miR483* is not expressed.

In regards to the ZBED6 project, further future investigations of ZBED6 are planned. We are mainly interested in the identification of the partners of ZBED6, and what is the mechanism by which ZBED6 represses transcription. Furthermore, the fact that *Igf2*, *miR483*-3p and *miR483*-5p are over-expressed in several cancer biology, and the established effect of ZBED6 on the regulation of *Igf2* and *miR483*, led to an increased interest in exploring the role of ZBED6 in tumors.

Summary in Swedish

All variation i de olika livsformer som ses på jorden beror på olika kombinationer av endast fyra nukleotider som kallas adenin (A), cytosin (C), guanin (G) och tymin (T). Dessa nukleotider är byggstenarna för det vi kallar deoxiribonukleinsyra (DNA). Det mänskliga genomet består av 3,2 miljarder nukleotider fördelade på två uppsättningar av vardera 23 enheter som kallas kromosomer. Varje avkomma får en uppsättning av 23 kromosomer från varje förälder, och dessa två uppsättningar bildar tillsammans hela genomet. Kromosomerna rekombineras innan de överförs till en avkomma, och denna rekombinationsprocess ger olika DNA-kombinationer för alla avkommor i en familj, vilket illustreras i de olika observerbara egenskaperna (fenotypen) för varje individ inom familjen. Endast cirka 1-2% av det mänskliga genomet kodar för gener. Variationen i en enda nukleotid vid en given position i genomet räcker ibland för att ha en dramatisk effekt på fenotypen hos en organism.

I det första av de två projekten i denna avhandling studerade vi en autoimmun sjukdom kallad spontan autoimmun tyreoidit (SAT). SAT förekommer hos en stam av kycklingar som kallas Obese Strain (OS), och dessa utgör en djurmodell för den humana sjukdomen Hashimotos tyreoidit. I denna studie försökte vi identifiera regionerna i kycklinggenomet som associeras med SAT. Vi korsade därför kycklingar från OS-linjen, där nästan alla fåglar utvecklar SAT, med kycklingar från den röda djungelhönan (RJF), där alla fåglar är friska. Individer från första generationen (F_1) korsades för att generera F_2. På samma sätt korsades alla efterföljande generationer (F_2-F_8) för att nå generation F_9. Denna avancerade korsningsstrategi ökar antalet rekombinationshändelser, vilket underlättar processen att identifiera genomregioner associerade med SAT.

Därefter utvecklade vi ett verktyg för att klassificera F_9-kycklingarna som friska eller SAT-påverkade genom att mäta nivåerna av sköldkörtelstimulerande hormon (TSH) i serumprover extraherade vid 10, 20 och 28 veckors ålder. Sköldkörtlarna hos F_9-kycklingarna dissekerades, sektionerades och färgades sedan också för att dessutom klassificera kycklingarna med hjälp av en andra mikroskopbaserad metod. Dessa fenotypdata användes för att selektera en frisk och en SAT-påverkad grupp av kycklingar. DNA från de två grupperna extraherades, sekvenserades och jämfördes mot varandra för att identifiera SAT-associerade regioner. Denna analys av DNA-sekvenserna identifierade totalt 12 regioner som var signifikant associerade med SAT. Ytterligare analys krävs emellertid fortfarande för att kartlägga loci som orsakar SAT. Vi drog dessutom slutsatsen att SAT i kyckling liknar den humana sjukdomen Hashimotos tyreoidit i förhöjningen av nivåerna av serum-TSH och den polygena (orsakad av mer än en gen) karaktär de två sjukdomarna har.

I det andra projektet i denna avhandling ville vi undersöka hur det DNA-bindande proteinet ZBED6, som fungerar som en transkriptionsfaktor, reglerar uttrycket av två gener, *Igf2* och *miR483*. För att undersöka detta använde vi en myoblastcellinje från mus (C2C12) där vi inaktiverade ZBED6 genom att antingen mutera dess bindningsställe i *Igf2*-genen, vilket genererade *Igf2*^{dGGCT}-celler, eller genom att radera *Zbed6*-genen och därmed generera *Zbed6*$^{-/-}$-celler. Dessa två mutantceller jämfördes med vildtypscellerna på RNA- och proteinnivå. Vidare undersöktes effekten av dessa mutationer på cellulära processer. Vi fann att ZBED6 har en direkt effekt på regleringen av mitokondriell aktivitet, men också på uttrycket av RNA främst genom reglering av *Igf2*. Vi visade också att överuttryck av ZBED6 (ZBED6-OE) var mer kritisk än inaktivering av densamma. Detta baserades på observationen att ZBED6-OE inducerar cellcykelarrest och programmerad celldöd (apoptos), medan inaktivering av ZBED6 inducerar en snabbare tillväxthastighet och ökad hypertrofi jämfört med vildtypscellerna.

Vår undersökning utvidgades till effekten av ZBED6 på regleringen av en speciell typ av RNA, så kallade mikro-RNA eller miRNA, i muskel-, njur- och levervävnad från mus. Vår analys visade att ZBED6 inte är en allmän regulator av miRNA, med undantag av *miR483*, som återfinns i ett intron av *Igf2*-genen. Vi fann att ZBED6 hämmar uttrycket av Igf2 delvis genom att dämpa uttrycket av miR483. Dessutom fann vi att den huvudsakliga funktionen för miR483 i C2C12-myoblaster är att reglera uttrycket av *Igf2* genom att binda till dess 5'-UTR och stabilisera dess transkript.

Sammanfattningsvis har vi framgångsrikt utforskat ZBED6-*Igf2*-*miR483*-interaktionen på RNA- och proteinnivå i C2C12-myoblaster samt i olika musvävnader.

كل التنوع الظاهر بالأشكال المختلفة للحياة على سطح الأرض سببه اختلاف تسلسل أربعة أحماض نووية هي: الأدينين (A)، والسايتوسين (C)، والغاوانين (G)، وأخيراً الثايمين (T) أو اليوراسيل (U). تشكّل هذه الأحماض النوويّة أحجار الأساس لبناء ما يسمى بالحمض النووي الريبوزي منزوع الأوكسيجين، أو ما يعرف اصطلاحاً بال .(DNA) تتشكل المادة الوراثيّة عند البشر من حوالي ٦،٤ مليار حمض نووي، تتوزع على شكل مجموعتين من الأصباغ الوراثيّة. إذ يرث الانسان مجموعةً واحدةً مؤلفةً من ٢٣ صبغاً من كلّ من الوالدين ليصبح لديه ٤٦ صبغاً وراثياً يشكّلون مادته الوراثية، أي الجينوم. تختلط الأصباغ الوراثية من المجموعة الأولى مع الأصباغ الموازية لها في المجموعة الثانية، فينتج عن ذلك مجموعتين جديدتين تحتويان على مزيج من المجموعتين الأساسيتين .

يرث الأبناء إحدى هاتين المجموعتين من الأهل، وبذلك يحصل الابن على مزيج من المواد الوراثية من أجداده الأربعة. ويسمح هذا الامتزاج بتأمين سلسلة وراثيّة فريدة لكل ابن، تتجلى عبر اكتساب مواصفات فريدة وأنماط ظاهريّة خاصة بكل ابن من أبناء العائلة الواحدة. تتراوح النسبة المؤثرة في المادة الوراثية، أي التي ترمز لبروتين ما على سبيل المثال، بين١٪ و٢٪. إلا أنّ أي اختلاف في حمض نووي واحد قد يكون له تأثير كبير على النمط الظاهريّ للشخص.

تتضمن هذه الأطروحة مشروعين؛ قام الأول بدراسة مرضٍ يخصّ المناعة الذاتية يسمى "التهاب الغدة الدرقية المناعي الذاتي التلقائي". يصيب هذا المرضُ بشكل تلقائي سلالة معينة من الدجاج والتي تسمى بالسلالة السمينة نظراً لإصابتها بهذا المرض المسبب للسمنة، والذي يُعدُّ نموذجًا حيوانيًّا لدراسة داء "هاشيموتو" الذي يصيب البشر، المعروف أيضًا باسم "التهاب الغدّة الدرقيّة اللمفيّ المزمن ."

بحثنا خلال هذه الدراسة في الجينوم عن أماكن مرتبطة بظهور التهاب الغدة الدرقية عند سلالة الدجاج السمين، وهو ما قد يساهم بتحديد أماكن وراثية ترتبط بداء "هاشيموتو". ومن أجل الوصول إلى تلك الغاية، قمنا بتزويج دجاج من السلالة السمينة، (حيث تقارب نسبة الإصابة بالتهاب الغدة الدرقية ١٠٠٪) مع دجاج الادغال الاحمر (وهي سلالة بريّة خالية تقريبًا من هذا المرض). ثم قمنا كخطوة ثانية بتزويج إناث الجيل الأول مع ذكور الجيل نفسه للحصول على الجيل الثاني. ثم كرّرنا الامر نفسه مراراً إلى أن حصلنا على ما يقارب ال ٨٥٠ دجاجة من الجيل التاسع. تعتبر هذه الاستراتيجية المتمثلة بعملية التزويج المتكررة طريقة فعالة لزيادة حدوث امتزاج الاصباغ الوراثية، مما يساهم بزيادة دقّة تحديد الأماكن في الجينوم المتوقع اكتشاف ارتباطها بمرض التهاب الغدة الدرقية.

قمنا لاحقًا بتطوير وإنتاج أداة تُستعمل لتحديد النمط الظاهري للدجاج من الجيل التاسع. مما يمكننا من تصنيف الدجاج بين معافى مصاب بمرض التهاب الغدة الدرقية. ويتمّ هذا الامر عبر قياس كمية هرمون تحفيز الغدة الدرقية (TSH) في دم الدجاج: إذ ترتفع معدلات الـ TSH بالدم كنتيجة لإصابة الغدة الدرقية بالاتهاب المزمن الذي يسبب تدميرها التدريجي مما يؤدي لانخفاض قدرتها على العمل بشكل سليم والاستمرار بإنتاج هرموناتها. عند الانتهاء من تطوير هذه الأداة، تم استخدامها لفحص معدلات هرمون الـ TSH عبر سحب دماء من الجيل التاسع من الدجاج في عمر الـ١٠ و٢٠ و٢٨ أسبوعا.

ثمّ اتّبعنا أسلوبًا اخر لتحديد النمط الظاهري للدجاج لنؤكد فعالية ودقّة الأداة التي قمنا بتطويرها. ولذلك قمنا بالتضحية بدجاج الجيل التاسع واستخرجنا منها الغدد الدرقيّة، ثمّ قسمّناها إلى شرائح رقيقة تتم معالجتها بمواد كيماويّة تسمح بتحديد درجة الالتهاب الحاصل للغدة والمتمثّل بازدياد في اعداد كريات الدم البيضاء فيها.

تم استخدام نتائج تحليل الأنماط الظاهرية للدجاج لتصنيفها واختيار مجموعة مصابة وأخرى غير مصابة بالتهاب الغدّة الدرقيّة. وتم استخراج الحمض النووي من هاتين المجموعتين وتحديد تسلسلها ومن ثمّ مقارنتها ببعضها البعض لتحديد الأماكن في الجينوم المرتبطة بهذا المرض. أدى تحليل مقارنة الجينوم بين المجموعتين، الصحيحة والمصابة، الى عدّة نتائج، هي:

-تحديد ١٢ مكانا في الجينوم مرتبط بمرض التهاب الغدة الدرقية ارتباطا ذا دلالة إحصائيّة. الا اننا ما زلنا بحاجة لمتابعة التحليل للوصول الى نتائج أكثر دقّة.
-إستنتاج أنّ مرض التهاب الغدة الدرقيّة المناعي الذاتي التلقائي الذي يصيب السلالة السمينة من الدجاج يشبه داء "هاشيموتو" الذي يصيب البشر؛ إذ أنّ كلاهما يتسببان بازدياد معدلات هرمون الـ TSHبالدم، ويرتبطان بأماكن متعددة في الجينوم.

امّا المشروع الثاني، فقد أردنا استكشاف كيفية تنظيم بروتين الـ " ZBED6 " لعملية نسخ الـ DNA لمختلف الجينات، وعلى وجه الخصوص للجينان المسميان "*Igf2* " و"*miR483*". ومن أجل الوصول إلى غايتنا، استخدمنا خلايا مخبرية "C2C12" أُستخرجَت من خلايا عضلة فخذ فأر مخبريّ. وكانت هذه الخلايا قد تعرّضت لاستحداث طفرة اولى داخل جين الـ "*Igf2* " تمنع بروتين الـ "ZBED6 " من التعرف والارتباط بالحمض النووي الخاص بهذا الجين، وسميّت هذه الخلايا بـ "$Igf2^{dGGCT}$ ". اما الطفرة الثانية فتتمثل بقصّ وإزالة جين الـ "*Zbed6*" من هذه الخلايا لتتخذ اسمًا جديدًا معدّلاً هو "$Zbed6^{-/-}$".
بعدها قمنا بدراسة هذه الخلايا التي تحتوي على الطفرات، وقارنّاها بخلايا خالية من الطفرات فوجدنا أنّ بروتين الـ "ZBED6 " لديه تأثير مباشر على نشاط "المتقدّرة"، أي مصنع الطاقة، بالخلايا. كما وجدنا أيضا أنّ له تأثير عام على عملية نسخ الـ " DNA"، إلاّ أنّ هذا التأثير يحدث بشكل ثانوي عبر الـ "*Igf2*". بعد ذلك لاحظنا أنّ زيادة انتاج بروتين الـ " ZBED6 " داخل الخلايا لديه تأثير أكبر وأهم من خفض انتاجه؛ إذ أنّ زيادة انتاجه تسبب إحباط العملية الطبيعية لانقسام الخلايا وبدخول الخلايا في عملية موت الخلايا المبرمج. بينما تسبب خفض انتاجه بازدياد سرعة انقسام الخلايا وتضخم حجمها.
توسّعت أبحاثنا بعد ذلك لدراسة تأثير الـ "ZBED6 " في العضلات والكبد والكلى لدى الفئران على نوع خاص من الاحماض النووية يسمى بـ "microRNA" او "miRNA". وتوصلنا لاستنتاج أنّ تأثير الـ "ZBED6 "على الـ "miRNA" ضعيف جدًا، مع استثناء تأثيره الكبير على جين الـ "*miR483*" الذي يقع داخل جين الـ "*Igf2*". كما وجدنا أيضا ان بروتين الـ "ZBED6 " يُحبط نسخ جين الـ "*Igf2*" جزئيًا من خلال إحباط نسخ جين الـ "*miR483*". فضلا عن ذلك وجدنا ان الـ "*miR483*" يساهم بتنظيم انتاج الـ "*Igf2*" عبر الارتباط به وزيادة استقراره. بإختصار كشف البحث عن كيفية تنظيم العلاقة بين الـ " ZBED6 " و الـ "*Igf2*" و الـ "*miR483*".

Additional work performed by the author

Lindgren, G., **Naboulsi, R.**, Frey, R., Solé, M. Genetics of skin disease in horses. Vet Clin North Am Equine Pract, 2020. 36(2): p. 323-339.

Acknowledgments

This thesis, among other achievements, would not have been possible without the combined help and support of many people in my life. In spite of the fact that what is written below only scratches the tip of the iceberg of my feelings and appreciation that cannot be expressed in a few sentences, I would like to express my sincere gratitude to all of them, especially to:

Leif Andersson, my main supervisor. I met you for the first time in the gel room when I visited for a trial of the ddPCR instrument. I mentioned this meeting to different people and one of them described you as one of the best scientists in the world of genetics. During the years I spent in your group, I was privileged to see why you are actually the best geneticist in the world. Your knowledge is a treasure, your guidance is invaluable, and your patience and care for your co-workers is unique. Simply, you are a great scientist and a personal mentor. I developed my scientific knowledge under your supervision and I learned too many things on the personal level as well. For all of that, I will always be very grateful to you.

Susanne Kerje, my co-supervisor. It has been a great pleasure to work with you. You have always supported me during my PhD journey, in times of success and failure. I feel honored for having the chance to work with a kind and understanding person like you.

Mårten Larsson, my co-supervisor. The support I got from you during the years was outstanding. You always knew the right thing to be done. Your optimism was contagious. I am very grateful for your belief in me. I thank you from my heart.

Mats Pettersson, I admire your never-ending wisdom especially in speeding up my slow R commands.

Nima Rafati, my colleague since I joined this group and my friend even after you moved to a new job. Your kindness and helpfulness are rare and your positive energy is admirable.

Ulrika Gunnarsson, for the help with the chicken thyroid images. **Shady** Younis, for the joint experiments we performed, **Calle** Rubin, for all of the valuable inputs during our scientific discussions.

The very nice present and past colleagues in the corridor, **Ulla** G., **Åsa** K., **Eva** M, **Sergey** K., **Sergey** A., **Jennifer** M., **Matteo** B., **Angela** F., **Débora** P., **Khurram** M., **Akhtar**, A. **Anna** O., **Alvaro** M., **Jessica** P., **Elisabeth** S., **Ola** W.

The heads of the IMBIM department and all of the administrative staff **Göran** Ak., **Catharina** S., **Rehné** Å., **Veronica**, H., **Alexis** F., **Malin** R. and **Malin** S.

Gabriella L. and **Göran** An., I met you at the beginning of my PhD career in the "Genome analysis" course. I learned too many things from you and yet hope for much more to learn in the coming period.

Knut and Alice Wallenberg and **Agnes and Mac Rudberg** foundations, for making my PhD studies possible.

Uppsala University and the **Swedish government** that gave me a chance to pursue my studies at the Master and PhD degrees.

Vanco N., you treated me as a brother since our first meeting back in 2010. This cannot be overlooked or forgotten. I will always appreciate your help and will miss our talks and fun times.

جدتي، كم تمنيت ان تشاركيني افراحي. الى روحك الطاهرة التي لم تفارقني ابدا، اساله تعالى ان يتغمّدك برحمته وان يجزيك خيرا عني وان يجعل اجر هذا العمل بميزان حسناتك.

ابي الغالي يا ماء زمزم، اشكرك على تعبك وتضحيتك اللامتناهية لنا. لن يوفيك حقك كل الكلام مهما كثر. احبك من أعماق قلبي وأكثر من نفسي.

امي، يا جنتي على الأرض ودار الحب والحنان، لم أصل لهذه اللحظة الا بإيمانك بي وتشجيعك لي وسهرك وتعبك. لا تسعفني قدرتي على صياغة الكلام لقول ما يليق بك. اشكرك جزيل الشكر. أساله تعالى ان يتكفل بإرضائكم عني.

اخي، قرة عيني وصديقي الدائم، واخواتي حبيباتي الحنونات. وجودكم بحياتي يضفي بهجة وسرور ويهون الصعاب.

زوجتي، حبيبتي وصغيرتي ورفيقة دربي، واجهنا اوقاتا صعبة ارهقتنا. الا ان وجودك الدائم بجانبي كان سببا ينمح القوة ويحفز على تخطي الصعاب. ابارك لك ولنفسي تخرجك وتخرجي. اشكرك واحبك جدا.

ابني، قلبي وابن روحي ومدللي، ستتمكن من قراءة هذه السطور يوما ما. فلتعلم بأني كنت الجأ لاحتضانك لارتاح من ارهاق العمل وأن النظر في عيونك البريئة كفيل بإسعادي وبأن ينسيني تعبي وبمنحني القوة على النهوض للعمل في اليوم التالي. اشكرك يا ملاكي وصنديدي.

الجنين الذي لا نعرف اسمه بعد، ننتظر قدومك بشوق. ادعوه تعالى ان يكتب لك ولوالدتك ولادة يسيرة.

الرسول الكريم، عليك أفضل الصلاة والتسليم، وعلى آلك وصحبك اجمعين. يأبى القلب والعقل قبول محاولات الظالمين النيل من مقامك الكريم ونسبك الرفيع. احمده تعالى على ارسالك رحمة لنا واشكرك على اتمامك الرسالة والا لكنا في ضلال مبين.

في النهاية اشهده تعالى ان هذا العمل خالص لوجهه الكريم لا أشرك به أحدا ولا شيئا.

References

1. Mendel, G., *Versuche über Plflanzenhybriden.* Verhand- lungen des natur forschenden Vereines in Brünn, 1866. **Bd. IV für das Jahr 1865**: p. 3-47.
2. Bateson, W. and G. Mendel, *Mendel's principles of heredity; a defence*. 1902, Cambridge,: University press. xiv p., 1 l., 212 p.
3. Miescher, F., *Die Histochemischen und Physiologischen Arbeiten von Friedrich Miescher—Aus dem wissenschaftlichen Briefwechsel von F. Miescher.* F. C. W. Vogel, 1869. **vol. 1**: p. 33-38.
4. Miescher, F., *Ueber die chemische Zusammensetzung der Eiterzellen.* Medicinisch-chemische Untersuchungen, 1871. **4**: p. 441-460.
5. Dahm, R., *Discovering DNA: Friedrich Miescher and the early years of nucleic acid research.* Hum Genet, 2008. **122**(6): p. 565-81.
6. Flemming, W., *Zellsubstanz, Kern und Zelltheilung.* F.C.W. Vogel 1882.
7. Flemming, W., *Historical Paper. Contributions to the Knowledge of the Cell and Its Vital Processes.* J Cell Biol, 1965. **25**: p. SUPPL:1-69.
8. Sutton, W.S., *The chromosomes in heredity.* Biological Bulletin, 1903. **4**(5): p. 231-251.
9. Bridges, C.B., *Non-Disjunction as Proof of the Chromosome Theory of Heredity.* Genetics, 1916. **1**(1): p. 1-52.
10. Watson, J.D. and F.H. Crick, *The structure of DNA.* Cold Spring Harb Symp Quant Biol, 1953. **18**: p. 123-31.
11. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-8.
12. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
13. Mullis, K., et al., *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Cold Spring Harb Symp Quant Biol, 1986. **51 Pt 1**: p. 263-73.
14. Garrod, A.E., *About Alkaptonuria.* Med Chir Trans, 1902. **85**: p. 69-78.
15. Fernandez-Canon, J.M., et al., *The molecular basis of alkaptonuria.* Nat Genet, 1996. **14**(1): p. 19-24.
16. Tautz, D., *Hypervariability of simple sequences as a general source for polymorphic DNA markers.* Nucleic Acids Res, 1989. **17**(16): p. 6463-71.
17. Imsland, F., et al., *The Rose-comb mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective sperm motility.* PLoS Genet, 2012. **8**(6): p. e1002775.
18. Kerje, S., et al., *Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken.* Anim Genet, 2003. **34**(4): p. 241-8.
19. Andersson-Eklund, L., et al., *Mapping quantitative trait loci for carcass and meat quality traits in a wild boar x Large White intercross.* J Anim Sci, 1998. **76**(3): p. 694-700.
20. Darvasi, A. and M. Soller, *Advanced intercross lines, an experimental population for fine genetic mapping.* Genetics, 1995. **141**(3): p. 1199-207.

21. Goodnow, C.C., et al., *Cellular and genetic mechanisms of self tolerance and autoimmunity.* Nature, 2005. **435**(7042): p. 590-7.
22. Jacobson, E.M. and Y. Tomer, *The genetic basis of thyroid autoimmunity.* Thyroid, 2007. **17**(10): p. 949-61.
23. Chardes, T., et al., *The human anti-thyroid peroxidase autoantibody repertoire in Graves' and Hashimoto's autoimmune thyroid diseases.* Immunogenetics, 2002. **54**(3): p. 141-57.
24. Lorini, R., et al., *Hashimoto's Thyroiditis.* Pediatr Endocrinol Rev, 2003. **1 Suppl 2**: p. 205-11; discussion 211.
25. Bodis, G., V. Toth, and A. Schwarting, *Role of Human Leukocyte Antigens (HLA) in Autoimmune Diseases.* Rheumatol Ther, 2018. **5**(1): p. 5-20.
26. Ceccarelli, F., N. Agmon-Levin, and C. Perricone, *Genetic Factors of Autoimmune Diseases 2017.* J Immunol Res, 2017. **2017**: p. 2789242.
27. Zaletel, K. and S. Gaberscek, *Hashimoto's Thyroiditis: From Genes to the Disease.* Curr Genomics, 2011. **12**(8): p. 576-88.
28. Burn, G.L., et al., *Why is PTPN22 a good candidate susceptibility gene for autoimmune disease?* FEBS Lett, 2011. **585**(23): p. 3689-98.
29. Serrano, N.C., P. Millan, and M.C. Paez, *Non-HLA associations with autoimmune diseases.* Autoimmun Rev, 2006. **5**(3): p. 209-14.
30. Baric, A., et al., *Association of established hypothyroidism-associated genetic variants with Hashimoto's thyroiditis.* J Endocrinol Invest, 2017. **40**(10): p. 1061-1067.
31. Lo, M.S., et al., *Monogenic Hashimoto thyroiditis associated with a variant in the thyroglobulin (TG) gene.* J Autoimmun, 2018. **86**: p. 116-119.
32. Brcic, L., et al., *Association of established thyroid peroxidase autoantibody (TPOAb) genetic variants with Hashimoto's thyroiditis.* Autoimmunity, 2016. **49**(7): p. 480-485.
33. Cole, R.K., *Hereditary hypothyroidism in the domestic fowl.* Genetics, 1966. **53**(6): p. 1021-33.
34. Cihak, J., et al., *Preferential TCR V beta 1 gene usage by autoreactive T cells in spontaneous autoimmune thyroiditis of the obese strain of chickens.* J Autoimmun, 1995. **8**(4): p. 507-20.
35. Wick, G., et al., *Distribution and functional analysis of B-L/Ia-positive cells in the chicken: expression of B-L/Ia antigens on thyroid epithelial cells in spontaneous autoimmune thyroiditis.* Mol Immunol, 1984. **21**(12): p. 1259-65.
36. Magner, J., *Historical note: many steps led to the 'discovery' of thyroid-stimulating hormone.* Eur Thyroid J, 2014. **3**(2): p. 95-100.
37. Edmonds, M., et al., *Secretion of alpha and beta subunits of TSH by the anterior pituitary.* Clin Endocrinol (Oxf), 1975. **4**(5): p. 525-30.
38. Grommen, S.V., et al., *Identification of unique thyrotropin receptor (TSHR) splice variants in the chicken: the chicken TSHR gene revisited.* Gen Comp Endocrinol, 2008. **156**(3): p. 460-3.
39. Van Laere, A.S., et al., *A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig.* Nature, 2003. **425**(6960): p. 832-6.
40. Jeon, J.T., et al., *A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus.* Nat Genet, 1999. **21**(2): p. 157-8.
41. Nezer, C., et al., *An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs.* Nat Genet, 1999. **21**(2): p. 155-6.
42. Markljung, E., et al., *ZBED6, a novel transcription factor derived from a domesticated DNA transposon regulates IGF2 expression and muscle growth.* PLoS Biol, 2009. **7**(12): p. e1000256.

43. Kubista, M., B. Akerman, and B. Norden, *Characterization of interaction between DNA and 4',6-diamidino-2-phenylindole by optical spectroscopy.* Biochemistry, 1987. **26**(14): p. 4545-53.
44. van Heerde, W.L., et al., *Markers of apoptosis in cardiovascular tissues: focus on Annexin V.* Cardiovasc Res, 2000. **45**(3): p. 549-59.
45. Cornelissen, M., et al., *Annexin V expression in apoptotic peripheral blood lymphocytes: an electron microscopic evaluation.* Apoptosis, 2002. **7**(1): p. 41-7.
46. Fadok, V.A., et al., *Exposure of phosphatidylserine on the surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages.* J Immunol, 1992. **148**(7): p. 2207-16.
47. van Engeland, M., et al., *Annexin V-affinity assay: a review on an apoptosis detection system based on phosphatidylserine exposure.* Cytometry, 1998. **31**(1): p. 1-9.
48. Raynal, P. and H.B. Pollard, *Annexins: the problem of assessing the biological role for a gene family of multifunctional calcium- and phospholipid-binding proteins.* Biochim Biophys Acta, 1994. **1197**(1): p. 63-93.
49. Andree, H.A., et al., *Binding of vascular anticoagulant alpha (VAC alpha) to planar phospholipid bilayers.* J Biol Chem, 1990. **265**(9): p. 4923-8.
50. Koopman, G., et al., *Annexin V for flow cytometric detection of phosphatidylserine expression on B cells undergoing apoptosis.* Blood, 1994. **84**(5): p. 1415-20.
51. Hertveldt, K., et al., *Flow cytometry as a quantitative and sensitive method to evaluate low dose radiation induced apoptosis in vitro in human peripheral blood lymphocytes.* Int J Radiat Biol, 1997. **71**(4): p. 429-33.
52. Vermes, I., et al., *A novel assay for apoptosis. Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V.* J Immunol Methods, 1995. **184**(1): p. 39-51.
53. Bacso, Z., R.B. Everson, and J.F. Eliason, *The DNA of annexin V-binding apoptotic cells is highly fragmented.* Cancer Res, 2000. **60**(16): p. 4623-8.
54. Rieger, A.M., et al., *Conventional apoptosis assays using propidium iodide generate a significant number of false positives that prevent accurate assessment of cell death.* J Immunol Methods, 2010. **358**(1-2): p. 81-92.
55. Rieger, A.M., et al., *Modified annexin V/propidium iodide apoptosis assay for accurate assessment of cell death.* J Vis Exp, 2011(50).
56. van Engeland, M., et al., *A novel assay to measure loss of plasma membrane asymmetry during apoptosis of adherent cells in culture.* Cytometry, 1996. **24**(2): p. 131-9.
57. Cui, H., et al., *IGF2-derived miR-483 mediated oncofunction by suppressing DLC-1 and associated with colorectal cancer.* Oncotarget, 2016. **7**(30): p. 48456-48466.
58. Emmerling, V.V., et al., *miR-483 is a self-regulating microRNA and can activate its own expression via USF1 in HeLa cells.* Int J Biochem Cell Biol, 2016. **80**: p. 81-86.
59. Pepe, F., et al., *Regulation of miR-483-3p by the O-linked N-acetylglucosamine transferase links chemosensitivity to glucose metabolism in liver cancer cells.* Oncogenesis, 2017. **6**(5): p. e328.
60. Veronese, A., et al., *Mutated beta-catenin evades a microRNA-dependent regulatory loop.* Proc Natl Acad Sci U S A, 2011. **108**(12): p. 4840-5.
61. Liu, H., et al., *Altered regulation of miR-34a and miR-483-3p in alcoholic hepatitis and DDC fed mice.* Exp Mol Pathol, 2015. **99**(3): p. 552-7.

62. Lupini, L., et al., *Over-expression of the miR-483-3p overcomes the miR-145/TP53 pro-apoptotic loop in hepatocellular carcinoma.* Oncotarget, 2016. **7**(21): p. 31361-71.
63. Veronese, A., et al., *Oncogenic role of miR-483-3p at the IGF2/483 locus.* Cancer Res, 2010. **70**(8): p. 3140-9.
64. Liu, M., et al., *The IGF2 intronic miR-483 selectively enhances transcription from IGF2 fetal promoters and enhances tumorigenesis.* Genes Dev, 2013. **27**(23): p. 2543-8.
65. Patterson, E.E., et al., *MicroRNA profiling of adrenocortical tumors reveals miR-483 as a marker of malignancy.* Cancer, 2011. **117**(8): p. 1630-9.
66. Mincer, D.L. and I. Jialal, *Hashimoto Thyroiditis*, in *StatPearls*. 2019: Treasure Island (FL).
67. Hutfless, S., et al., *Significance of prediagnostic thyroid antibodies in women with autoimmune thyroid disease.* J Clin Endocrinol Metab, 2011. **96**(9): p. E1466-71.
68. Wick, G., et al., *The obese strain (OS) of chickens with spontaneous autoimmune thyroiditis: review of recent data.* Curr Top Microbiol Immunol, 1981. **91**: p. 109-28.
69. Sahlqvist, A.-S., *Genetic Characterization of Chicken Models for Autoimmune Disease.* Digital comprehensive summaries of Uppsala dissertations from the facuty of medicine 828, 2012.
70. Rose, N.R., L.D. Bacon, and R.S. Sundick, *Genetic determinants of thyroiditis in the OS chicken.* Transplant Rev, 1976. **31**: p. 264-85.
71. Wick, G., et al., *Avian models with spontaneous autoimmune diseases.* Adv Immunol, 2006. **92**: p. 71-117.
72. Younis, S., et al., *The ZBED6-IGF2 axis has a major effect on growth of skeletal muscle and internal organs in placental mammals.* Proc Natl Acad Sci U S A, 2018. **115**(9): p. E2048-E2057.
73. Florini, J.R., et al., *"Spontaneous" differentiation of skeletal myoblasts is dependent upon autocrine secretion of insulin-like growth factor-II.* J Biol Chem, 1991. **266**(24): p. 15917-23.
74. DeChiara, T.M., E.J. Robertson, and A. Efstratiadis, *Parental imprinting of the mouse insulin-like growth factor II gene.* Cell, 1991. **64**(4): p. 849-59.
75. Sasaki, H., K. Ishihara, and R. Kato, *Mechanisms of Igf2/H19 imprinting: DNA methylation, chromatin and long-distance gene regulation.* J Biochem, 2000. **127**(5): p. 711-5.
76. Jiang, L., et al., *ZBED6 modulates the transcription of myogenic genes in mouse myoblast cells.* PLoS One, 2014. **9**(4): p. e94187.
77. Wang, X., et al., *ZBED6 negatively regulates insulin production, neuronal differentiation, and cell aggregation in MIN6 cells.* FASEB J, 2019. **33**(1): p. 88-100.
78. Yaffe, D. and O. Saxel, *Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle.* Nature, 1977. **270**(5639): p. 725-7.
79. Younis, S., et al., *The importance of the ZBED6-IGF2 axis for metabolic regulation in mouse myoblast cells.* The FASEB Journal, 2020.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1700

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-423787

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2020