UNIVERSITÄT ZU LÜBECK

**From the Medical Department I**
**of the University of Lübeck**
**Director: Prof. Dr. med. Marquardt**

# Gene expression in Liver is Altered by DNA methylation and miRNAs in Obese Subjects

Dissertation
for Fulfillment of
Requirements
for the Doctoral Degree
of the University of Lübeck

from the Department of Natural Sciences

Submitted by

Christin Krause
from Geesthacht

Lübeck 2019

First referee: Dr. rer. nat. Henriette Kirchner
Second referee: Prof. Dr. rer. nat. Henrik Oster
Date of oral examination: 28.01.2020
Approved for printing. Lübeck, 29.01.2020

# Contents

# 1 Abstract

Though Type 2 diabetes (T2D) is a global health threat, the underlying molecular mechanism of disease etiology remains incompletely known. Manifestation of systemic insulin resistance is often noticed in a late state and the lack of effective treatments results into missing reversion of disease progress.
The liver is the body's most important organ for the maintenance of blood glucose, therefore a disruption of hepatic metabolic pathways is likely causal for the hyperglycemic state in insulin resistance and T2D. We suppose that this observed hepatic insulin resistance is mediated by epigenetic alterations, which mirror influences of modern sedentary lifestyle and high fat and high sugar nutrition. The aim of this thesis is to identify novel target genes and the underlying epigenetic regulatory mechanisms which contribute to metabolic liver diseases in humans.
Dysregulated hepatic gene expression of pathways involved in hepatic lipid and glucose metabolism was identified by TaqMan and SYBR green qPCR. Site specific DNA methylation was analyzed by bisulfite pyrosequencing in liver biopsies of 101 human obese T2D and non-diabetic (ND) subjects with different stages of Non-alcoholic fatty liver disease (NAFLD) and carefully evaluated by respective regression models. Functionality of altered methylation sites was estimated by Luciferase reporter gene assay and electrophoretic mobility shift assay. Moreover, a systematic approach by microarray analysis was used for the characterization of aberrant miRNA expression in diverse T2D related metabolic traits. Therefore a smaller subset cohort of 40 liver biopsies, matched for age, gender, and BMI, was chosen as discovery cohort and candidate miRNAs were validated by qPCR in the whole cohort. Potentially repressed target genes were identified by *in silico* binding site analysis and validated by mimic transfection in cell culture and subsequent gene expression measurement by qPCR. For a comprehensive evaluation, also the induction of miRNA expression by insulin stimulation and the concentration in human serum was examined.
We described for the first time a decrease of human hepatic TH signaling in T2D and NAFLD which is mediated by reduced *THRB* expression. Moreover, we identified *GALNT18* as novel gene involved in age and obesity driven T2D manifestation, which is likely regulated by intragenic DNA methylation. Methylation markers within *ABCG1* and *SREBF1* emerged from blood-based epigenome-wide methylation studies and hence generated methylation scores could not be reproduced as significantly altered in human liver. These results prove a missing functionality of these methylation sites in T2D manifestation. Increased hepatic hsa-miR-182-5p expression was identified as an regulatory mechanism for repression of *FOXO1* and *LRP6* leading to enhanced *de novo* lipogenesis in T2D and NAFLD. We were also able to reproduce an association of altered liver-specific hsa-miR-122-5p expression to aberrant glycemic traits (blood glucose and HbA1c level) as well as for liver fibrosis and steatosis state in T2D. Moreover, hepatic hsa-miR-223-3p expression correlated with lipid traits in an age dependent manner. Repression of possible target genes involved in hepatic lipid metabolism could not be validated by correlation analysis for both lastly mentioned miRNAs.
Finally, a complex regulatory network of *IRS2* repression mediated by SREBF1 and Sp1 motif hypermethylation and hepatic hsa-let-7e-5p overexpression emerged as proposed mechanism for the phenotype of selective insulin resistance. Thereby CpG methylation is influenced by a genome-wide associated SNP rs4547213 for T2D. Moreover, hepatic hsa-let-7e-5p expression is induced by acute hyperinsulinemia in cell culture. This network was used to generate and evaluate different hepatic and blood-based models by robust machine learning approaches as translation into clinical application.

# 2 Zusammenfassung

Die weltweite Prävalenz von Typ 2 Diabetes (T2D) hat in den vergangenen Jahren drastisch zugenommen und betrifft nicht nur Industrieländer, sondern zunehmend auch Entwicklungsländer. Eine Vorstufe von T2D stellt die Insulinresistenz dar, die durch einen permanenten hyperglykämischen und hyperinsulinämischen Zustand des Organismus gekennzeichnet ist. Im Verlauf der Erkrankung kann die Insulinsekretion des Pankreas den steigenden Bedarf an Insulin nicht weiter kompensieren, weshalb es zur schlussendlichen Manifestierung des T2D kommt. Die genaue molekulare Ursache für den Beginn der Insulinresistenz ist weitgehend unbekannt. Genetische Faktoren erklären nur einen marginalen Teil der zunehmenden Häufigkeit von T2D, wohingegen Umwelteinflüsse, wie durch ungesunde Ernährung verursachtes Übergewicht oder eine moderne, bewegungsarme Lebensweise mit wenig Sport, stark mit dem vermehrten Auftreten von T2D verbunden sind. Diese Einflüsse spiegeln sich in der fehlerhaften Regulation von Genen in Zucker- und Fettstoffwechselwegen wider, welche durch epigenetische Mechanismen, wie DNA-Methylierung und Repression durch micro RNAs (miRNAs), verursacht werden.

Da besonders die Leber für die allgemeine Energiehomeostase verantwortlich ist, wird vermutet, dass eine hepatische Insulinresistenz mitunter ursächlich ist für die Entwicklung von T2D. Diese Arbeit beschäftigt sich daher mit der Identifikation von Schlüsselgenen und dessen epigenetische Regulation, die in T2D verändert ist.

Hierzu wurde die DNA-Methylierung an CpG-Oligonucleotiden und Genexpression spezifischer Gene von 101 Leberbiopsien adipöser Individuen mit verschiedenen Stadien der Leberverfettung, Insulinresistenz oder Erkrankung an T2D, gemessen. Da durch Leberverfettung und -fibrose assoziierte Erkrankungen, auch als *Non-alcoholic fatty liver disease* (NAFLD) bezeichnet, stark mit der hepatischen Insulinresistenz verbunden sind, wurden ebenso die, durch Schulddrüsenhormonen aktivierten, hepatische Signalwege des Lipidmetabolismus untersucht. Dabei wurde zum ersten Mal im humanen System eine Repression der *THRB*-Expression als mögliche Ursache der hepatischen Resistenz gegenüber Schilddrüsenhormonen identifiziert.

Signifikante Unterschiede in der Methylierung wurden hinsichtlich eines regulatorischen Effekts durch verschiedene Reporter- und Bindungsassays analysiert. *GALNT18* wurde hierbei als neues Gen in der T2D Manifestation identifiziert. Blutstudien-basierte Methylierungsmarker für T2D und Übergewicht sind hingegen in der Leber nicht verändert.

In einer kleineren Subkohorte aus 20 Typ 2-diabetischen und 20 nicht-diabetischen Individuen wurde das hepatische miRNA-Expressionsprofil untersucht und mit verschiedenen T2D assoziierten Blutparametern korreliert. Die miRNA hsa-miR-182-5p zeigte dabei eine starke hepatische Überexpression in der gesamten Kohorte, was die Repression zweier Gene, *FOXO1* und *LRP6*, verursacht und so zu einer Verstärkten *de novo* Lipogenese führt. Die miRNAs hsa-miR-122-5p und hsa-miR-223-3p zeigten Assoziationen mit Blutparametern, jedoch konnte kein Zielgen hierfür identifiziert werden.

Selektive Insulinresistenz der Leber, die vor allem durch anhaltende Gluconeogenese und gleichzeitiger Lipogenese gekennzeichnet ist, wird unter anderem durch eine fehlerhafte Regulation der *IRS2*-Expression ausgelöst. Als regulatorische Ursache dieser Repression wurde ein humanes Modell generiert, welches eine Kombination aus intergenischer Hypermethylierung zweier Bindungsmotive für die Transkriptionsfaktoren Sp1 und SREBF1, Beeinflussung dieser Methylierung durch einen benachbarten T2D assoziierten Polymorphismus, und Hyperinsulinämie induzierter hepatischer Überexpression der miRNA hsa-let-7e-5p beschreibt.

Schlussendlich wurden diese hepatischen regulatorischen Mechanismen der selektiven Insulinresistenz und entprechend validierte, repräsentierende Parameter im Blut genutzt, um Klassifikations- und Vorhersagesysteme für T2D zu ergänzen und zu beurteilen.

# 3 Introduction

## 3.1 Type 2 diabetes - A modern pandemic

The number of patients diagnosed with diabetes rises continually and world-wide. This tendency is shown by a drastic increase of patients with diabetes from 1980 (108 million) to 2014 (422 million) (Roglic and Organization, 2016). This rise is not limited to the industrialized world, but affects also low- and middle income countries (Whiting et al., 2011). The majority of all cases represents as Type 2 diabetes (T2D) (Tamayo et al., 2016). In contrast to Type 1 diabetes (T1D), which has mostly a childhood or juvenile onset and is characterized by deficient insulin production, T2D manifests mostly during adulthood due to the inability of the body to response to insulin action and the inadequate response of pancreatic beta cells to compensate for the higher demands on insulin.

Health cost estimates from Germany show that about 10 % of healthcare expenses are used for T2D treatment costs with an upward tendency (Jacobs et al., 2017). Besides the cost factor, T2D as a disease itself is a strong psychological burden for patients and family members. T2D is incurable and effectiveness of common anti-diabetic treatments to control body glucose vary for each individual. Patients have to monitor their blood glucose and insulin levels closely to avoid hyper- or hypoglycemia and -insulinaemia which limits life quality. Carelessness could be fatal leading to coma and death. In 2016, the Word Health Organization (WHO) states diabetes as the seventh leading cause of death with about 3.7 million preventable incidences.

T2D is a chronic disease and characterized by progressive hepatic and systemic insulin resistance (IR), decreased ability to produce endogenous insulin and sustained hyperglycemia and hypertriglyceridemia. Comorbidities like hepatic fat accumulation, increased blood pressure and cardiovascular diseases might emerge as interplay from additional risk factors like aging and obesity. Especially obesity and excess visceral fat accumulation is a key driver of metabolic dysfunction and tightly linked to T2D. Furthermore, T2D is a leading cause for blindness due to diabetic retinopathy, neuropathy, lower limb amputation and kidney failure (Roglic and Organization, 2016).

A high fat and high sugar diet together with low exercise levels is leading to obesity. This excess in energy uptake is stored as lipids in adipose tissue or fat depots of other organs, like liver or pancreas, damaging and hindering other organs in their functionality (Liu et al., 2010; Singh et al., 2017). Besides obesity, age is a major risk factor for T2D based on the senescence of cells within metabolic organs (especially pancreatic beta cells) (Gong and Muzumdar, 2012; Palmer et al., 2015; Helman et al., 2016).

A genetic component, indicated by a history of family members suffering from T2D, is a strong risk factor to develop IR and T2D. Therefore these factors besides lifestyle habits are mostly used to "predict" one individual's risk to manifest T2D. Nevertheless, this rise in prevalence cannot be explained solely by genetic inheritability and predisposition. It is rather assumed that environmental influences regarding a sedentary and unhealthy lifestyle with less exercise, exposure to toxins (smoking) and high fat/high sugar nourishment are the driving factors initiating the vicious cycle towards IR. To date, the underlying molecular mechanisms of IR and T2D are not completely understood but hepatic dysregulated glucose and lipid metabolic pathways contribute as key driver (Loria et al., 2013). Lifestyle interventions including a change in diet, increased physical activity and weight loss result into more effective T2D prevention strategies than medication.

It is mandatory to understand how modern lifestyle influences the manifestation of IR, disease progression towards T2D and the inheritability between generations to improve life quality of patients and to counteract this modern pandemic.

## 3.2 The hepatic paradox of selective insulin resistance

The liver is the body's most important organ to maintain normal glucose homeostasis during fasting and feeding. During hunger periods, blood glucose levels are maintained for whole body energy homeostasis by secretion of hepatic glucose due to gluconeogenesis or degradation of glycogen. Other lipid metabolizing organs are also able to maintain tissue energy homeostasis by hepatic ketone bodies. Moreover, the liver is able to store excess glucose as glycogen. Also free fatty acids can be metabolized by hepatic synthesis of triacylglycerols, phospholipids or cholesterol, which are stored as lipid droplets or secreted as very low density lipoprotein (VLDL) and high density lipoprotein (HDL).
Since a dysregulation of hepatic glucose or lipid metabolism leads towards an imbalance in energy homeostasis and the hallmark markers of T2D, hyperglycemia and hypertriglyceridemia, these key pathways of the physiological and (known) pathological state should be mentioned.

### 3.2.1 The physiological state

In normal metabolism, an increase in blood glucose results into enhanced secretion of insulin by pancreatic beta cells. Insulin responsive tissues (liver, adipose tissue, skeletal muscle) increase glucose uptake by specific transporters. In liver, insulin activates upon binding to the insulin receptor (INSR) a signal cascade via Insulin Receptor Substrate 1 (IRS1), IRS2 and Protein Kinase B (known as PKB or AKT). This cascades results into a decrease of hepatic glucose production by phosphorylation and cytoplasmatic retention of Forkhead Box proteins (like FOXO1). Without phosphorylation, these FOXO transcription factors translocate to the nucleus and enhance for example the expression of genes involved in gluconeogenesis. As an overall result, the liver maintains low blood glucose levels by storage of excess glucose as glycogen by insulin induced inhibition of Glycogen Synthase Kinase 3 Beta (GSK3$\beta$) and subsequent activation of the Glycogen Synthase (GYS).
Also the hepatic conversion of glucose into lipids is tightly regulated by insulin. Insulin stimulates the activation of Sterol Regulatory Element Binding Transcription Factor (SREBP-1c) and transcription of its gene *SREBF*, which is a key transcription factor for enzymes involved in fatty acid and triglyceride biosynthesis. As a result of this activated *de novo* lipogenesis, VLDL lipids are synthesized and either secreted into the blood stream to be stored in adipose tissue or in hepatic lipid droplets. Hepatic cholesterol storage is degraded either for fatty acid $\beta$-oxidation for extra hepatic energy gain during fasting periods, synthesis of ketone bodies or used for the synthesis of bile acids.

### 3.2.2 Pathological state in Type 2 diabetes and liver steatosis manifestation

IR is a pathogenic state where the liver is not physiologically responding to elevated serum insulin concentrations during feeding periods. A total hepatic insulin resistance caused by a complete block of downstream signaling would rather lead to increased gluconeogenesis and simultaneously decreased lipogenesis. Instead, the phenotype of selective insulin resistance is observed which is characterized by persistent hepatic lipogenesis besides sustained gluconeogenesis (Brown and Goldstein, 2008; Samuel and Shulman, 2012).
Careful analysis of two directly linked downstream molecular switches, Irs1 and Irs2, in respective knockout mice propose a potential mechanism behind this contradictory metabolic dysregulation, which depend on the state (fasting or fed) and the zonation of the liver (Figure 3.1). This mechanism proposed sustained gluconeogenesis in the hepatic periportal (PP) zone by impaired Irs2 signaling causing activation of respective enzymes and pathways. In the perivenous (PV) zone, prolonged Irs1 expression compensated for decreased Irs2 and thereby maintains hepatic lipogenesis (Kubota et al., 2016). This is in accordance to the observation of decreased *IRS2* and unchanged *IRS1* expression in liver samples of human overweight subjects with NAFLD (Honma et al., 2018). In contrast, also enhanced *de novo* lipogenesis induced by increased *IRS2* expression besides unchanged *IRS1* levels was observed in liver samples of obese subjects

with progressive hepatic steatosis and inflammation (Rametta et al., 2013). These inconsistent results indicate a high dynamic in overall disease progression and manifestation. Further studies in mice propose an additional small scale time dependency on which protein is needed to transmit insulin signaling. Irs2 mediated signaling is active in both zones during fasting periods with low plasma glucose levels and Irs1 is mostly active after feeding in the PV zone (Kubota et al., 2008). Hence, the mechanisms which lead to selective insulin resistance remain unclear.

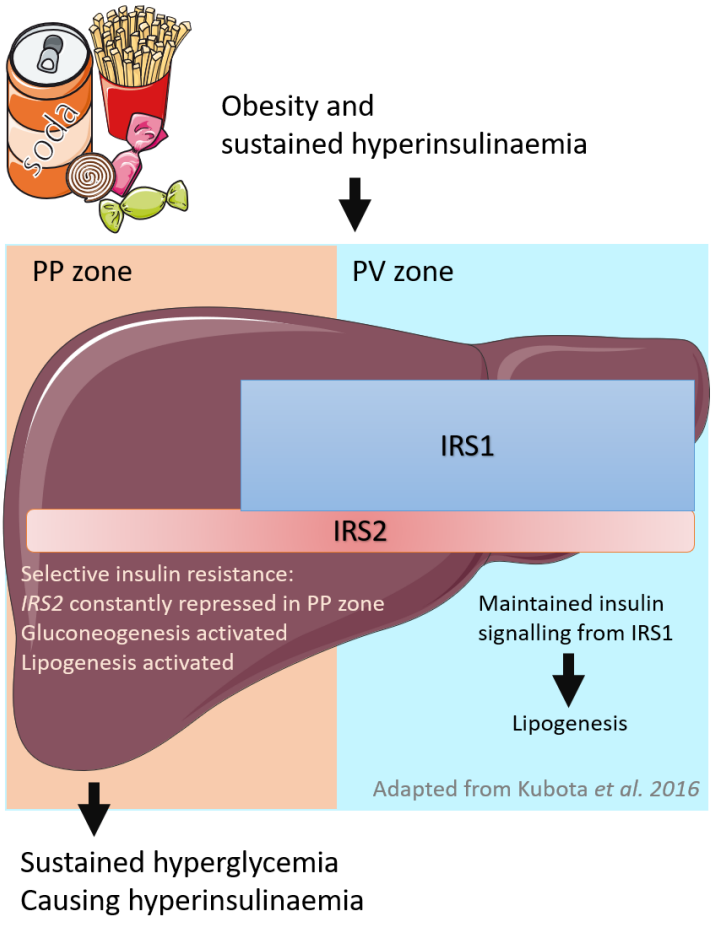


Figure 3.1: **Mode of selective insulin resistance based on the expression of *IRS2*.** Lasting *IRS1* expression and repression of *IRS2* in the periportal (PP) and perivenous (PV) zone causing sustained gluconeogenesis and lipogenesis, leading to enduring hyperglycemia and hyperinsulinaemia. Figure adapted from (Kubota et al., 2016).

As already mentioned, sustained hepatic (*de novo*) lipogenesis results into obesity and fatty liver. Increased adipose tissue lipolysis and enrichment in serum free fatty acids have an reinforcing function on hepatic fat accumulation. Fatty acid overdose lead to lipotoxicity and inflammation by sustained $\beta$ oxidation for fatty acid disposal or formation of triglycerides to store fat as lipid droplets. Associated diseases are called non-alcoholic fatty liver diseases (NAFLD) and contain different stages of progression and severity. Simple steatosis and reinforcing hepatic inflammation lead to non-alcoholic seatohepatitis (NASH), which is a precursor to cirrhosis due to increased tissue fibrosis. Persistent inflammation and fibrosis is leading to hepatocellular cancer (HCC). NAFLD also worsens insulin resistance by inflammatory responses to lipid accumulation and fibrosis which compromises hepatic metabolism. It is unclear which mechanism initiated this cycle of lipid accumulation and insulin resistance though both hepatic diseases are closely linked to each other. A recent meta analysis for NAFLD of approximately 8.5 million subjects in scientific literature revealed a global prevalence of 25 %, whereby at least 50 % was further associated to obesity and 25 % to T2D (Younossi et al., 2016).

A lack of hepatic thyroid hormone (TH) signaling was previously identified as contributor to the pathogenesis of hepatic lipid accumulation and NAFLD (Pagadala et al., 2012). Hypothyroidism is also associated with increased serum triglycerides and cholesterol levels and might be an important risk factor in metabolic syndrome (Sinha et al., 2018).
In the physiological state, TH stimulate adipose lipolysis and the release of free fatty acid, which are used by the liver for the synthesis of cholesterol or metabolized by $\beta$-oxidation. Hepatic *de novo* lipogenesis is stimulated by TH and subsequent transcriptional activation of respective key enzymes (for example *FASN* encoding Fatty Acid Synthase or *ACACA* encoding Acetyl-CoA Carboxylase Beta). Furthermore, TH signaling induces *Low Density Lipoprotein Receptor* (*LDLR*) and *Cytochrome P450 Family 7 Subfamily A Member 1* (*CYP7A1*) expression, which results into cholesterol clearance by conversion of LDL into bile acids (Ness, 1991).
The underlying mechanism of an apparent hepatic TH resistance which results into enhanced hepatic lipid accumulation is not well understood. The coherence between TH, obesity and fat accumulation emerged from human observational studies where active TH analoga resulted into improvement of obesity and hypercholesterolemia (Krotkiewski, 2000). Additionally, the functionality of TH receptor $\beta$ in respective knock out mice is evaluated regarding an increased hepatic lipid accumulation upon disruption of *THRB* gene (Araki et al., 2009). Thus a contribution of hepatic TH signaling to selective hepatic insulin resistance besides NAFLD is feasible.

## 3.3 The deadlock of genome-wide association studies in deciphering a cause of Type 2 diabetes

High throughput sequencing tools enabled the era of genome-wide association studies (GWASs) to find possible genetic predispositions and causes for a number of complex (non-Mendelian inherited) diseases. In 2007, the first T2D related GWAS was performed identifying the single nucleotide polymorphism (SNP) rs7903146 within *Transcription Factor 7 Like 2* (*TCF7L2*), which is one of the most prominent risk loci (Sladek et al., 2007). Since then, several GWAS identified additional 152 other loci in about 120 genes (Prasad and Groop, 2015). Altogether, these variants explain only a small proportion (less than 20 %) of the disease inheritability.
Association results from GWAS depend strongly on sample size. Therefore a recent large meta-study including almost 900.000 individuals from 32 GWASs with European descent revealed about 243 loci which were associated with a predisposition to T2D (Mahajan et al., 2018). Statistical analysis identified 51 signals with a causal relationship to T2D and within those markers, 43 SNPs which might contribute to actual gene regulation. A validation of these statistical results by functional experiments or how these alterations affect which tissue dependent metabolic pathway is still missing.
Besides studies for the association to the incidence of T2D, also GWAS regarding common risk factors (obesity and age) or metabolic traits (fasting plasma glucose or insulin secretion) were performed (Speliotes et al., 2010; Andersson et al., 2013; Hong et al., 2014). These results are mostly descriptive and rely solely on statistical significance. Further evaluation whether these markers are associated to T2D manifestation is performed by calculation of GWAS based genetic risk scores and cohort follow-up studies (Lango et al., 2008; Meigs et al., 2008; van Hoek et al., 2008; Lin et al., 2009; Andersson et al., 2013). Hence, addition of genetic scores to common risk factors only slightly improves risk group stratification (Lango et al., 2008; van Hoek et al., 2008).
GWAS represent a suitable tool for large-scale data mining regarding disease associated biomarkers. Unfortunately, in the field of causal T2D research, they lack for meaningful statements due to the missing proof of function. Another limiting factor of GWAS is the diversity within populations, so these results may apply only for one ethnic cohort. Also most T2D GWAS neglect the interplay between environment and genetics or between several alterations.

## 3.4 Epigenetic - A fresh breeze in explaining the manifestation of complex diseases

In general, the field of epigenetic is defined as the totality of all non-genetic modifications which influence gene expression. The occurrence of epigenetic regulation on the genome is intuitive since it is mandatory to allow the transcription of only a cell type specific set of genes for development of different cell and tissue types within one organism.

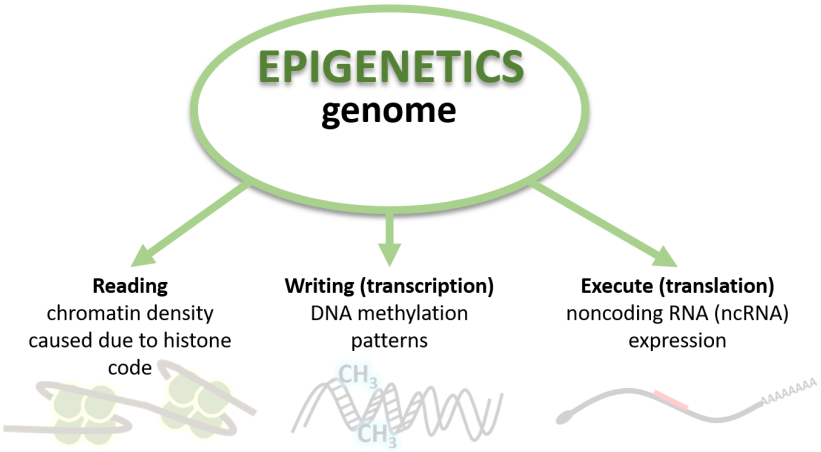


Figure 3.2: **Layers of epigenetic regulation.** Epigenetic mechanisms act on different layers including chromatin remodelling which inhibits reading of genes, DNA methylation which inhibits binding of transcription factors to the DNA and noncoding RNAs (especially micro RNAs) which mostly inhibit translation of genes to proteins.

Three distinct methods of epigenetic regulation are known, which act on different transcriptional layers (see Figure 3.2): The first layer is about changes in chromatin density which allows or represses transcription (reading). Diverse modifications of specific histone protein residues, most commonly methylation, acetylation, phosphorylation and ubiquitinylation, are responsible for either a dense or a loose wrapping of DNA around histones. These modifications are often referred as histone code to state the DNA density. The second layer contains covalent attachment of methyl groups to DNA bases, most commonly on Cytosine of CpG oligonucleotides. The underlying mechanism are explained later in detail. Since DNA methylation allows or represses transcription factor binding among others, this mechanism regulates the writing process of genes into messenger (m)RNA. The last layer of epigenetic regulation interferes with the translation of mRNA into proteins by binding of noncoding (nc)RNAs. These ncRNAs are most commonly short-interfering (si)RNAs (viral origin) or micro (mi)RNAs (endogenously transcribed). The processing of miRNAs and their function in translational interference will also be explained later.
The methods of this thesis will include only the analysis of transcriptional regulation by DNA methylation and translational interference by miRNA. Both mechanisms are highly side- and gene-specific. Changes can easily be quantified in small amounts of tissue. The scarce accessibility of human liver tissue is a limiting factor which makes it impractical to perform studies for the discovery of chromatin structure changes in T2D.

### 3.4.1 DNA methylation in physiology

DNA methylation is the covalent attachment of a methyl group to a nucleotide base which occurs almost exclusively on Cytosines of CpG oligonucleotides. Also methylation in the context of CpT, CpGA or CpC is described but in a way less pronounced manner. This non-CpG methylation is mostly observed in DNA of pluripotent stem cells, neuronal and glia cells or in oocytes and therefore might have a neglectable

impact on metabolic diseases (Jang et al., 2017).
The human genome contains CpG island (CGI) within or distal to gene promoters, which are characterized by a high abundance of CpG oligonucleotides (> 50 %) across a stretch of 1000 bp. These CGI are mostly unmethylated in the physiological state. Only a small tissue specific proportion is hypermethlated leading to silenced gene expression. CGI hypermethylation of especially tumor suppressor genes is commonly associated to cancer (Esteller, 2002). In contrast to promoter hypomethylation, gene bodies are hyper- or fully methylated which correlates in a non-linear manner with enhanced gene expression (Jjingo et al., 2012). This intragenic methylation occurs on single CpG sites or within intragenic CGIs, which regulate promoters for alternative transcripts (Maunakea et al., 2010). An influence of gene-body methylation on splicing or transcriptional elongation is also reported (Shayevitch et al., 2018; Veloso et al., 2014). Lastly, intergenic DNA methylation occurs on long distance regulatory elements which allows or prevents the binding of transcriptional enhancer and repressor proteins. Binding of transcription factors is not neccesscarily abolished upon methylation of CpG sites within the recognition motif. Methylation sensitivity or even a higher affinity to methylated DNA is strongly dependent on the transcription factor and presence of respective methyl-CpG-binding domains (MBDs) (Schübeler, 2015).
Though the attachment of a methyl group is a covalent bond, methylation and demethylation is a highly dynamic progress and changes during the differentiation of tissue types and development of organisms. Hypermethylation occurs to silence genes in the context of genomic imprinting, as well to form Barr bodies (X chromosomal inactivation in females). Also repetitive sequences which mostly originate from viral genomes are highly methylated to silence possible damage. During embryonic development, the parental imprint is completely deleted to achieve toti- and pluripotency and restored again by an unknown mechanism. Also parental alterations of methylation are inherited throughout generations.
Only a small proportion of the genome is differentially methylated since most CpG sites are within unmethylated CGI. These deferentially methylated regions (DMR) are highly tissue, cell and disease specific and occur on transcriptional start sited (TSS), enhancer elements or transcription factor binding motifs (Kitamura et al., 2007).
It is not clear, how timing and estabishment of methylation patterns on DNA is regulated. Nevertheless, methylation and demethylation is an active progress performed by DNA-methyltransferases (DMNTs) or ten-eleven translocation (TET) enzymes (Chen and Riggs, 2011). There are two types of DMNTs. DMNT1 has a high affinity to hemi-methylated DNA and copies the parental methylation pattern during DNA replication. DMNT3 (DMNT3a and b) appears to establish *de novo* DNA methylation by using a third cofactor (DMNT3L) without catalytic function. It is hypothesized that either non-coding RNAs or DNA binding transcription factors recruit DNA-methyltransferases. It is also controversially discussed whether the establishment DNA methylation patterns is a cause of protein binding or a mechanism to protect against gene activation by transcription factors (Schübeler, 2015). The more consensual mechanism how DNA methylation inhibits transcription factor binding is visualized in Fig 3.3. Demethylation can occur by modification through TET enzymes or by basepair excision repair of the DNA by unknown mechanisms (Chen and Riggs, 2011; He et al., 2011).
DNA methylation is known to interact with other layers of epigenetic regulation. DMNTs can recruit enzymes needed for histone modification to completely silence genes by tightening the chromatin structure (Cedar and Bergman, 2009). Furthermore, methylation of promoters of miRNA genes inhibit miRNA expression (Pheiffer et al., 2016). On the other hand, some miRNAs can bind to DMNT or TET mRNA and inhibit translation of those enzymes (Wang et al., 2012).
Measurement of DNA methylation represents only a snap shot of the actual situation. Several methods exist which work on a global or site specific level and which achieve different resolutions and valid differences in DNA methylation. The following explanation will give a short description of three commonly used techniques. All these methods have in common that they work with bisulfite treated DNA, where unmethylated cytosines are converted into uracil and methylated cytosines remain.

Whole genome bisulfite sequencing (WGBS) based on Flow cells sequences the whole genome on a low resolution level by a low average coverage of 5 to 10 reads per CpG site (Ziller et al., 2015). This enabled the detection of DMRs instead of single CpGs sites. Besides a low resolution, DNA library preparation and subsequent sequencing is expensive. Furthermore, data collection and analysis is time consuming and has enormous memory requirements. The advantage of independence of assay design has to be evaluated against sequencing costs and low coverage.

High throughput methods based on microarrays, as seen for GWAS, are also commonly used in larger Epigenome-wide association studies (EWAS). Probe selection relies solely on assay design by manufacturer and is rather optimized for analysis of aberrant promoter methylation in cancer. A revised version of these arrays include more probes (approximately 850.000 instead of 450.000 probes) with a higher coverage of predicted regulatory elements (Moran et al., 2015). Nevertheless, only single CpG sites are analyzed which might only point towards a neighboring effect causing site. This method needs also a complex measurement preparation, technical equipment and data analysis which causes high costs. Measurements for data collection relies on probe hybridization which might be biased due to polymorphisms within a CpG site itself or within the probe.

Due to the low coverage and concomitant low resolution, aberrant DNA methylation of candidate CpG sites have to be validated by more sensitive methods. For this purpose, bisulfite pyrosequencing represents the gold standard for the site-specific quantification of DNA methylation. This method is exclusively used within this thesis and therefore will be explained in more detail in the methods section. It offers a high resolution which is limited by the machine precision. Careful assay validation is further used to estimate probable measurement bias which could have been introduced by upstream applications (DNA extraction, assay design, amplification bias by polymerase efficiency). Since bisulfite pyrosequencing is CpG site specific and limited to only few bases (mostly up to 100 bp), it is not applicable for whole genome data mining. Nonetheless, a global comparison between 18 different laboratories revealed bisulfite pyrosequencing as method with most consistent results (The BLUEPRINT Consortium et al., 2016).

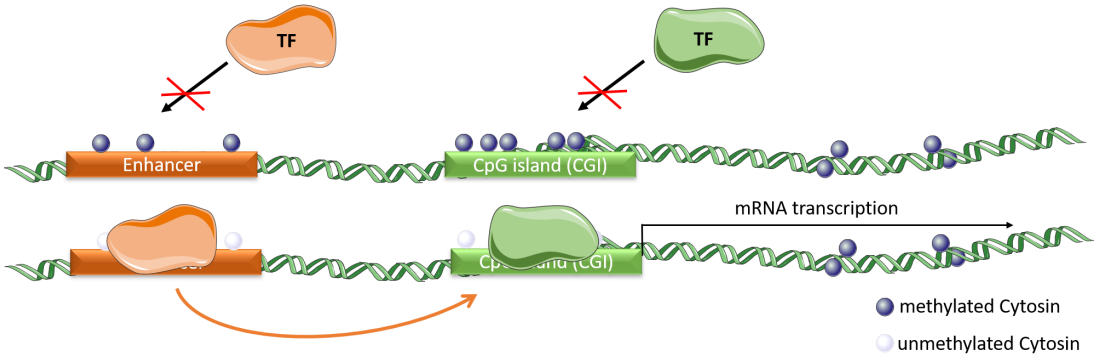


Figure 3.3: **DNA methylation abolishes binding of transcription factors (TF) to DNA motifs.** Physiological DNA methylation occurs on regions within gene promoters with high density of CpG oligonculeotides, called CpG island (CGI), or transcriptional regulatory sequences (for example enhancer elements). These covalent modifications are highly dynamic and can abolish or facilitate binding of transciption factors. CpG oligonucleotides within gene bodies are mostly methylated.

### 3.4.2 Epigenome-wide association studies to elucidate the mechanisms of Type 2 diabetes manifestation

Advanced technology enabled the era of EWAS in the context of complex diseases. Since non-genetic risk factors like aging, weight gain, unhealthy food, a sedentary lifestyle and intrauterine inheritance of parental DNA methylation patterns contribute more to the increasing prevalence of T2D, EWAS gained particular interest for diabetes associated glycemic traits, obesity and lipid traits. Different human studies on genome-wide changes in methylation were performed in DNA of blood leukocytes, skeletal

muscle, pancreas, adipose tissue and liver (see Table 1 for a detailed summary). Thereby two recurring methylation markers, *ABCG1* cg06500161 and *SREBF1* cg11024682, emerge from blood-based EWAS. Human liver, adipose tissue, pancreas or skeletal muscle specific studies are mostly in a small scale case-control design and mirror only a snap shot of the actual methylation pattern. Also methylation studies on monocygotic twin discordant for T2D exist in a tissue-specific manner. However, blood-based human studies offer the possibility of a follow-up analysis and offer a longitudinal study design. Blood drawing is also a safe method which enables the collection of a large sample size and to increased statistical power. As a result, most of these studies rely on small impact changes which become only apparent due to the number of cases. These high throughput studies also lack any mechanistic information or reliable high resolution validation. Furthermore, tissue specificity of DNA methylation patterns lead to the urgent need for careful evaluation of blood-based candidate CpG site in the respective relevant tissue. Lastly, DNA methylation is also dependent of the ethnicity, so further cross validation between populations are needed.

The identification of biomarkers or the calculation of DNA methylation-based risk scores for T2D prediction are two common approaches to translate EWAS results into a clinical application (Willmer et al., 2018). But to the present day, these approaches preform rarely better than disease process prediction by blood parameters.

Table 1: **Summary of recent metabolic Epigenome-wide studies.** Two most mentioned genes ABCG1 and SREBF1 are indicated in bold.

| Autor | Top Candidates/Marker | Significance | Data | Comments |
|---|---|---|---|---|
| Ahrens, 2016 | PC, ACLY, PLCG1, IGF1, IGFBP2, PRKCE | FDR < 0.05 | n = 63 liver biopsies, n = 23 postbariatric liver biopsies | NAFLD, effect of bariatric surgery |
| Akinyemiju, 2018 | IGF2BP1, **ABCG1** | p < E-07 | HyperGen n = 614, buffy coat | SNP information |
| Braun, 2017 | DHCR24, CPT1A, **ABCG1**, **SREBF1** | p < 1.08E-07 | n1 = 725 discovery, n2 = 760 replication (Rotterdam), blood | |
| Campanella, 2018 | **ABCG1** | p < 9.07E-08 - 3.27E-18 | n1 = 1941, n2 = 358, blood | Obesity and colorectal cancer |
| Chambers, 2015 | **ABCG1**, PHOSPHO1, SOCS3, **SREBF1**, TXNIP | p < E-17 | n = 13535 Indian Asians, n = 7066 Europeans (LOLIPOP), blood | Score based on 5 markers, longitudinal with 8 year follow-up |
| Dayeh, 2014 | 1649 CpG sites | FDR | n = 34 (ND) and n = 15 (T2D), pancreatic islet cells | T2D, further validation |
| Dayeh, 2016 | **ABCG1**, PHOSPHO1, SOCS3, **SREBF1**, TXNIP (Only these with T2D) | FDR < 0.018 | N = 258 (Botnia), blood | Replication of markers from other study |
| Dekkers, 2016 | CPT1A, **SREBF1**, **ABCG1** | | n = 3296 (Netherlands biobank), blood profile, microarray | Mendelian randomization for cause-consequence |
| Demerath, 2015 | HIF3A, CPT1A, **ABCG1**, LGALS3BP, KDM2B, PBX1, BBS2 | p < E-07 | n1 = 2097 African-American (ARIC), n2 = 2377 (Framingham Heart Study), n3 = 991 (medication), blood and n = 648 women (MuTHER cohort), blood | Obesity |
| Dhana, 2018 | BRDT, PSMD1, IFI44L, MAP1A, MAP3K5; **ABCG1** and **SREBF1** replicated | p < 1.08E-07, replication p < E-08 | Rotterdam n = 1450 and ARIC n = 2097 replication (African Americans), blood | BMI and waist circumference |
| Dick, 2014 | HIF3A | FDR = 0.05 | European origin: Cardiogenics consortiom n1 = 479 discovery, n2 = 339 first replication (MARTHA study), n3 = 1789 second replication (KORA study), blood; n = 635 (MuTHER study), adipose tissue | Influence of a SNP on methylation, significant in blood and adipose tissue |
| Hedman, 2017 | **ABCG1**, APOB | p < 1.08E-07 | n1 = 2306 discovery, n2 = 2025 replication, blood | Cardiovascular, lipid traits |
| Hidalgo, 2014 | **ABCG1** | p < 1.1E-7 | Discovery n1 = 544, replication n2 = 293, blood | HOMA-IR, SNP information |
| Kirchner, 2016 | binding motifs of ATF | FDR < 0.10 | n = 34 liver biopsies | T2D |
| Kriebel, 2016 | **ABCG1** | p < 9.2E-5 - 0.047 | n = 533 (KORA F4 cohort), blood | Glucose metabolism and obesity |
| Kulkarni, 2015 | TXNIP, **ABCG1** | FDR < 0.05 and P < E-22 | n = 850 Mexican-American, blood | T2D, HOMA-IR and fasting blood glucose; SNP information |
| Lindholm, 2014 | binding motifs of MRF, MEF2 and ETS | FDR < 0.05 | n = 23, skeletal muscle | before and after training |
| Nilsson, 2014 | PPARG, KCNQ1, TCF7L2, IRS1 | FDR | n1 = 120, n2 = 56, n3 = 40, n4 = 28, adipose tissue | T2D, Twin-study (n4, n3) and case-control (n1, n2) |
| Nilsson, 2015 | GRB10, ABCC3, MOGAT1, PRDM16, H19 | FDR < 0.05 | n = 95 liver biopsies | T2D |

| | | | | |
|---|---|---|---|---|
| Nitert, 2012 | MAPK1, MYO18B, HOXC6, PRKAB1 | ? | n = 28 with and without family histroy, n = 18 twins, skeletal muscle | T2D, heritariy and Twin-study |
| Orozco, 2018 | 24 genes, also FASN, SLC1A4, CPEB4 | p < E-08 | n = 201 (METSIM cohort), adipose tissue | Different traits, Biomarker |
| Pfeiffer, 2015 | **ABCG1**, MIR133B/**SREBF1**, TNIP1 | p < 1.1E-07 | n1 = 1776 (KORA F4 study) discovery, n2 = 499 (KORA F3 study) and n3 = 472 (inCHIANTI study), blood; n = 634 (MuTHER study), adipose tissue | Blood lipids |
| Ribel-Madsen, 2012 | PPARGC1A (muscle), HNF4A (adipose) | p < 0.05 (Westfall-Young resampling adjustment) | n = 22, skeletal muscle; n = 10, adipose tissue | T2D, Twin-study |
| Sayols-Baixeras, 2016 | **SREBF1**, SREBF2, PHOSPHO1, SYNGAP1, **ABCG1**, CPT1A, MYLIP, TXNIP, SLC7A11 | ? | n1 = 645 (REGICOR), n2 = 2542 (Framingham Offspring ), n3 = 98 (GOLDN) for gene expression in blood | Cardiovascular, lipid traits |
| Truong, 2017 | **ABCG1**, PHGDH | p < 8.4E-6 discovery, p < 0.0091 replication | n1 = 199 (5 french Canadian families), n2 = 324 French (MARTHA), blood | Also SNP – Methylation context |
| Volkmar, 2012 | 276 CpG sites | ? | n = 16 pancreatic islet cells; replication in n = 24, blood | T2D, methylation from pancreas not reproducable in blood DNA |
| Wahl, 2016 | 187 CpG sites (including **SREBF1** and **ABCG1**) | p < E-07 | n = 5387 discovery and n = 4874 replication, blood; different settings for tissue specific; n3 = 6 data base entries of blood, liver, muscle, pancreas, subcutaneous fat, omentum and spleen for cross-tissue | Meta Study for obesity, generation of risk score for T2D, Mendelian randomization for cause-consequence, SNP information |
| Walaszczyk, 2017 | **ABCG1**, LOXL2, TXNIP, SLC1A5, **SREBF1** | p < 0.00009 | 15 EWAS from blood, adipose and liver, replication in 100/100 T2D/ND | Only blood was replicated, other tissue specific not |

LOLIPOP .. The London Life Sciences Prospective Population; ARIC .. The Atherosclerosis Risk in Communities; MuTHER .. Multiple Tissue Human Expression Resource; MARTHA .. MARseille THrombosis Association; KORA - Kooperative Gesundheitsforschung in der Region Augsburg; METSIM .. The Metabolic Syndrome in Men; REGICOR .. Registre Gironí del COR; GOLDN .. Genomics of post-prandial lipidomic phenotypes in the Genetics of Lipid lowering Drugs and Diet Network

### 3.4.3 Small molecules, huge potential - MicroRNAs

MicroRNAs (miRNAs) are 21-22 nt oligonucleotides which are transcribed as Pri-miRNA in a single or polycistronic transcript (Bartel, 2004). Genes for miRNA transcription are located both intragenic within introns of mRNA transcribing genes or intergenic. Resulting pri-miRNAs are small double stranded oligonucleotides which form a step loop structure. Further processing by Drosha and Pasha in the nucleus generates a pre-miRNA which is subsequently exported to the cytosol and cleaved by Dicer into to a miRNA-duplex. This mature double stranded miRNA complex contains a guide and a passenger strand (mostly referred as miR*). The guide strand is incorporated into different Argonaute (AGO) proteins (dependent of the downstream function) to form the RNA-induced silencing complex (RISC). Guide strand selection is performed by a largely unknown mechanism including a contribution of several proteins (AGO, Dicer, TRBP (trans-activation response RNA-binding protein), PACT (protein activator of dsRNA-dependent protein kinase) and Xrn-1/2) (Meijer et al., 2014). Candiate miRNA selection was used to be restricted to the guide strand reasoned by complex stability and binding energy calculation. However, for some passenger strands also a mechanistic function is reported (Mah et al., 2010).
The formed RISC is guided by miRNA strand complementary to the target mRNA-3' untranslated region (3'UTR). In dependency of the match between miRNA and mRNA and the bound AGO, either site specific cleavage, enhanced degradation of mRNA or inhibition of translation is achieved (Gu and Kay, 2010).
Target prediction is a highly artificial method based on bioinformatic algorithms. Recent development figured out, that solely a 6-8 nucleotides long seed sequence contributes most to efficient target mRNA identification (Bartel, 2004, 2009). Since the length of the 3'UTR is highly tissue specific and varies because of mRNA isoforms and different poly-A signals, most algorithms predict unequal sets of target genes in dependency of the used transcriptomic data. Some databases integrate additional information regarding laboratory results and target site mutation analysis, as well sequence conservation across species to filter an overflow of feasible target sites. This is only possible due to the rising interest in miRNA and the increasing collection of *in vitro* validated targets. Moreover, the diversity of miRNA-mRNA interaction mode is complicating in *in silico* target prediction.

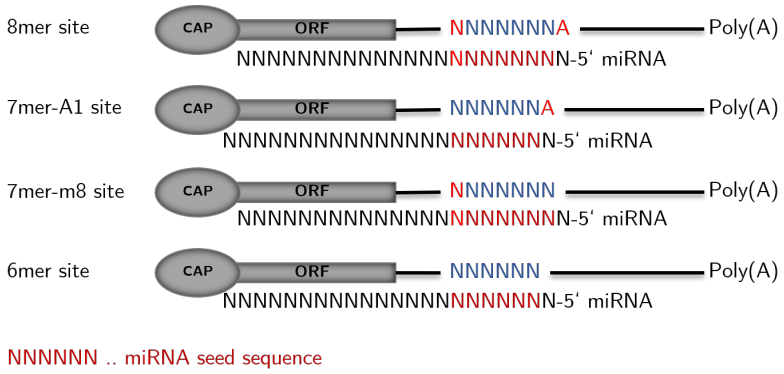


Figure 3.4: **Four canonical modes of miRNA-mRNA matches.** Each specific side change is indicated as red nucleotides in the target mRNA sequence. The miRNA seed is indicated as dark red sequence within the 5' end of the guide strand or as blue nucleotides within the 3'UTR of the target mRNA. Furthermore, a mRNA cap (CAP), the open reading frame (ORF) and the poly-A-tail (Poly(A)) is indicated for orientation on the mRNA.

Four types of these miRNA-mRNA matching modes are known (see Figure 3.4) (Peterson et al., 2014): A perfect base pair match between nucleotides 2-8 of the miRNA seed sequence and an A across the first nucleotide of the miRNA is called an 8mer. The 7mer binding mode can be distinguished by either complementary binding of the nucleotides 2-7 of the miRNA seed sequence and an A across the first

nucleotide of the miRNA (7mer-A1) or solely a perfect match between nucleotides 2-8 of the miRNA seed (7mer-m8). Lastly, a complementary match of 6 nucleotides of the seed sequence to its target is referred as a 6mer. A computational approach in comparing miRNA seed sequences revealed that most seeds are located from nucleotide 2 within the 5' end of the mature miRNA strand. This is not neccesscarily the rule for every miRNA related repression.

Since the first functional description, various implications of miRNAs especially in cancer were discovered (more than 30,000 entries for miRNAs in human cancer at PubMed). In a human T2D context, aberrant miRNA expression is best described in serum, plasma or leukocyte samples (He et al., 2017). Used high throughput methods to measure aberrant expression are based in microarrays or specific deep sequencing techniques. Interestingly, all candidate miRNA have a small or no overlap between those studies which mirrors again the complexity of T2D and the ethnicity specific gene regulation. The Subsequent detection of target genes is more elaborate and include either overexpression experiments in cell culture, co-transfections with reporter plasmids containing the 3'UTR of interest, mutation analysis or cross linking of RISC and DNA to extract bound mRNA species by antibody pull down assays. This is often missing for large-scale identified miRNA. Also the opposite direction is observed by identification of miRNAs emerged from mouse models or cell culture experiments whose expression is not evaluated in human tissue.

Only few human hepatic high throughput studies regarding miRNA expression in metabolic context were performed. Most information about miRNA expression changes in T2D or NAFLD was achieved by collection of liver from rodent disease models (Szabo and Bala, 2013). Nevertheless, some studies in human liver cover HCC development or disease progression in NAFLD. How miRNA contribute to the parthenogenesis of IR in humans is rather unknown due to a lack of systematic characterization.

Of all reported metabolically involved miRNAs, miR-122 is the most abundant hepatic miRNA (accounts for 70 % of total hepatic miRNAs) and urgently needed for maintenance of healthy liver homeostasis (Filipowicz and Grosshans, 2011; Esau et al., 2006; Wen and Friedman, 2012). However, this miRNA is also the most characterized hepatic miRNA in cancer, hepatitis C virus infection, fibrosis, steatosis and inflammation. Circulating miR-122 levels are associated with obesity or NAFLD, therefore possible metabolic target genes were particularly analyzed concerning hepatic lipid metabolism (Cermelli et al., 2011; Miyaaki et al., 2014). To date, silencing this miRNA by complementary binding of small oligonulceotides represents a potential therapeutic target (Krützfeldt et al., 2005). One antagomir, known as Miravirsen, is already successfully used for experimental treatment of Hepatitis C Virus infection (Gebert et al., 2014). Thus it is mandatory to identify and carefully characterize new miRNA which contribute to metabolic disease manifestation and might not act as a compensatory or protective mechanism. Furthermore, secreted miRNA might also act as potential circulating biomarker which can be assessed an monitored to counteract disease progression.

## 3.5 Motivation and aims

**Premise**

Every seventh death worldwide is due to T2D and respective comorbidities. Moreover, the increasing global prevalence of obesity and T2D will be one of the mayor challenges for each country's health care system. T2D is characterized by persistent hyperglycaemia, hyperinsulinemia and finally the inability of the pancreas to produce sufficient insulin. Since the liver is responsive for the maintenance of overall glucose homeostasis, pathological changes in hepatic glucose and lipid metabolism induced by insufficient insulin signaling are key drivers for the development of systemic insulin resistance and subsequent T2D manifestation.

Studies in liver samples of mice and human subjects with NAFLD already describe the persistent activation of genes involved in gluconeogenesis besides sustained lipogenesis in metabolic liver diseases,

which is referred as selective insulin resistance. However, the underlying molecular mechanism of disease etiology remains largely unclear which causes a lack of effective long-term therapies and missing reversion of insulin resistance. Large blood-based (epi-)genome wide studies provide approaches of disease onset and progress prediction to prevent T2D manifestation but lack in mechanistic analysis or validation in metabolically active tissue.

**Hypothesis**

T2D inheritability is only marginally explained by genomic alterations which indicates a significant contribution of lifestyle habits, obesity and malnutrition to disease onset. These environmental factors are mirrored by epigenetic mechanisms like DNA methylation and changes in miRNA expression, which interfere with physiological hepatic gene expression.

**Objective**

We aim to identify novel T2D target genes and respective epigenetic regulatory mechanisms, which includes alterations in DNA methylation and miRNA expression. Hepatic insulin resistance and NAFLD are closely connected to each other, therefore also fatty liver associated pathways including hepatic TH signaling are carefully elucidated. Human liver biopsies, collected during bariatric surgery, of 101 obese diabetic or non-diabetic subjects are analyzed regarding obesity-independent site specific DNA methylation, hepatic gene- and miRNA expression. Significant changes in DNA methylation are comprehensively analyzed regarding a regulatory function by reporter gene or electrophoretic mobility shift assay. A systematic approach for the characterization of hepatic miRNA expression in T2D and NAFLD is performed in a smaller subcohort of 20 diabetic and 20 non-diabetic subjects matched for age, gender and BMI, by microarray measurement and respective regression models for metabolic traits. An emerging candidate miRNA is evaluated in the complete cohort and its function is characterized regarding potentially repressed candidate genes. Lastly, the potential use of alterations as biomarker for disease stratification or T2D prediction is evaluated in human liver and in serum by usage of validated surrogate markers. Especially two recurring blood-based methylation markers for T2D and obesity within *ABCG1* and *SREBF1* are evaluated in human liver.

# 4 Subjects, Materials and Methods

## 4.1 Subjects

All liver wedge biopsies were obtained in a standardized fashion from segment III during bariatric surgery of obese subjects at University Hospital Eppendorf (UKE, Hamburg) in order to establish a tissue bank for metabolic disorders. Blood was drawn prior surgery and prepared serum was frozen at -80 $^{o}$C. Biopsies were taken in fasted state of the subjects. All participants signed informed consent. The study was approved by the local ethics committee "Ethik Komission der Ärztekammer Hamburg" (PV4889, 2015).

### 4.1.1 Subject stratification for type 2 diabetes and fatty liver

Subjects of this cohort were stratified into subjects with type 2 diabetes (T2D) or non-diabetic (ND) subjects by clinical examination. Therefore an acute treatment of subjects with anti diabetic medication (e.g. Metformin, insulin) was used as indicator for T2D. Since a glycemic index of hemoglobin (HbA1c) $\geq$ 6.5 % or a fasting blood glucose value of $\geq$ 126 mg/dL are official criteria stated in the American Diabetes Association (ADA), these factors were additionally considered by respective regression analysis (American Diabetes Association, 2018).
As this was no direct recruitment for a study, the cohort is only matched for BMI (Figure 4.1A) and not for age and gender (Figure 4.1B and C). In mean, T2D subjects are 14 years older. Gender proportion also shows a slightly more balanced cohort for T2D subjects (T2D: 61 % femals, 39 % males; ND: 81 % females, 19 % males). Further analysis will use age, gender and BMI as confounding factors.

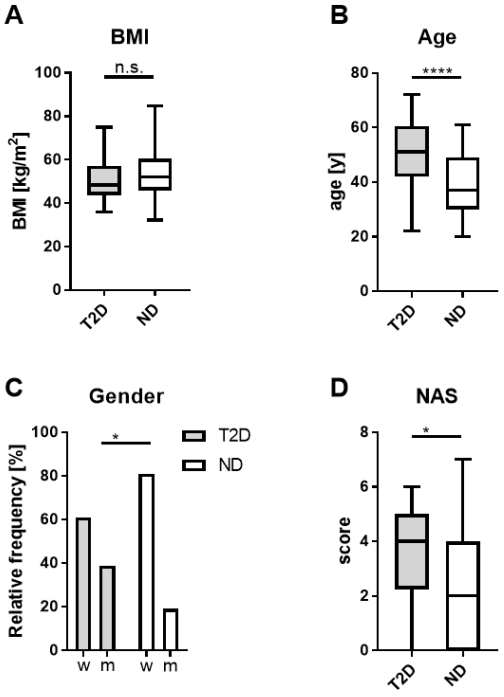


Figure 4.1: **Distribution of confounding factors whithin type 2 diabetic (T2D) and non-diabetic (ND) subjects.** This cohort is matched for BMI (A), nevertheless subjects with T2D are significantly older (B). The cohort contains more female ND subjects (C). The median NAS score (score for strength of liver steatosis and fibrosis) is higher in T2D subjects (D).

Besides official stratification criteria from the ADA and medical examination (diagnosis by UKE), the HOMA-IR (Homeostatic Model Assessment for Insulin Resistance) was calculated as estimate for insulin resistance (IR)(Matthews et al., 1985). The HOMA-IR was calculated by (fasting glucose mg/dL $*$

serum insulin mU/L)/405. A cut-off of 2.5 was applied for IR. As seen in Figure 4.2A some subjects are currently not diagnosed for T2D manifestation though they show an increased insulin resistance. A stratification for HOMA-IR was additionally applied for further analysis.
NAFLD activity score (NAS) was determined according to the current recommendations by two expert pathologists (Brunt et al., 2011). As the cohort differed also significantly for mean NAS (Figure 4.1D), the distribution of the score among T2D and ND subjects revealed more ND subjects with no hepatic steatosis or fibrosis (Figure 4.2B). Additional calculation of a respective score (Sterling et al., 2006) reveals a higher Fib-4 in subjects with T2D (Figure 4.2C). This score is based on age, blood levels of AST (aspartate aminotransferase), blood levels of ALT (alanine transaminase) and platelet count. No subject reaches the threshold of 3.25, which would result into a specificity of 97 % for hepatic fibrosis, and only few subjects reach a threshold of 1.45 which shows 90 % sensitivity. Nevertheless, NAS and Fib-4 score correlate positively ($p = 0.0059$, Figure 4.2D). Expression of *Apolipoprotein F* (*APOF*) which is reported to be a suitable inflammation marker is not altered between both subgroups (Figure 4.2E). As already indicated by NAS distribution and difference in mean, expression of *Cytochrome P450 Family 2 Subfamily C Member 19* (*CYP2C19*) which is reported to be a marker for NAS is decreased in subjects with T2D (Figure 4.2F).

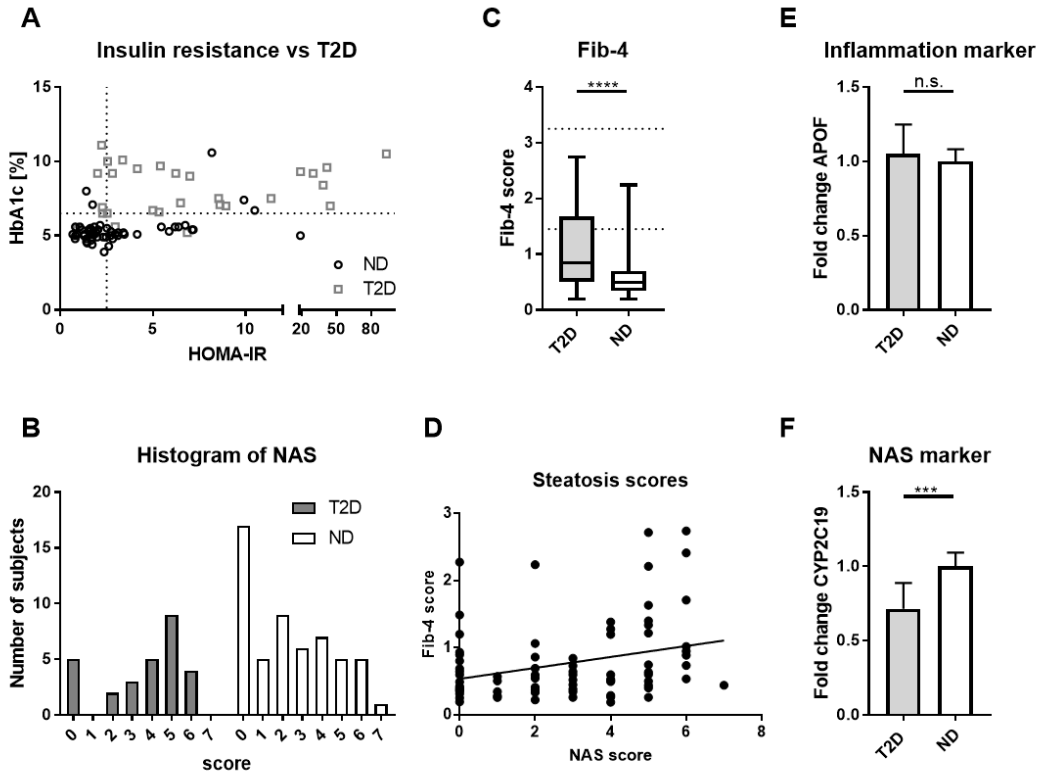


Figure 4.2: **Distribution of different scores and markers after stratification for type 2 diabetic (T2D) and non-diabetic (ND) subjects.** Most subjects with high HbA1c show also insulin resistance (indicated by HOMA-IR > 2.5) and are also diagnosed by medical examination as type 2 diabetic (T2D, A). Distribution of NAS score among both groups (B). Fib-4 score is increased in subjects with T2D, but no subject reaches indicated limits (C)(Sterling et al., 2006). NAS score correlated with Fib-4 score (D). Hepatic expression of *APOF* is not altered, indicating the same level of inflammation among both groups (E) (Ryaboshapkina and Hammar, 2017). Hepatic expression of *CYP2C19* is decreased in subjects with T2D, confirming distribution of NAS score (F) (Ryaboshapkina and Hammar, 2017).

### 4.1.2 Clinical parameters

All clinical parameters except serum insulin and serum C-peptide were measured prior surgery by the Institut für Klinische Chemie und Laboratoriumsmedizin, Zentrum für Diagnostik, Universitätsklinikum Eppendorf, Hamburg, Germany according to the DIN EN ISO 15189:2014 certification. Glucose, cholesterol, HDL cholesterol, triglycerides, AST and ALT were determined using photometric assays, HbA1c was quantified using capillary electrophoresis or turbidimetric inhibition assays, folic acid, vitamin B12 and TSH were measured using luminescent oxygen channeling immunoassays (LOCI), cortisol concentrations were quantified by electrochemiluminescence. LDL cholesterol was calculated by total cholesterol – (triglycerides/5) – HDL cholesterol. Serum insulin and C-peptide concentrations were measured by Enzyme-linked Immunosorbent Assays (ELISAs, see section 4.2.8) in duplicates according to manufacturer's protocol.
Fasting glucose, insulin and HbA1c are significantly increased in subjects with T2D (Figure 4.3A - C). The concentration of serum C-peptide, which is the connecting chain of pro insulin and cleaved off during insulin processing, does not differ between both subgroups. C-peptide concentrations are used to estimate the secretion efficiency of pancreatic beta cells whereby a decreased value indicates progressive damage of beta cells. Nevertheless, all subjects present values within normal range (230 - 1000 pmol/L, Figure 4.3D).

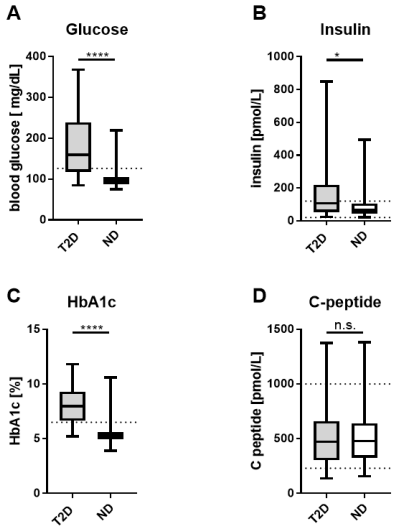


Figure 4.3: **Glycemic traits whithin type 2 diabetic (T2D) and non-diabetic (ND) subjects.** Blood glucose (A), serum insulin (B) and the glycemic index of hemoglobin (HbA1c, C) are elevated in subjects with T2D. Insulin secretion and processing indicated by C-peptide (D) is not significantly altered between both groups.

Whole cholesterol, LDL and HDL cholesterol are not significantly altered between both subgroups (Figure 4.4A - C). Serum triglycerides are significantly elevated in subjects with T2D (higher than normal range) which is a common feature of metabolic diseases. No subject reaches a value beyond 1000 mg/dL, therefore no pancreatitis induced T2D is observed (Sjoberg and Kidd, 1989; Kota et al., 2012).

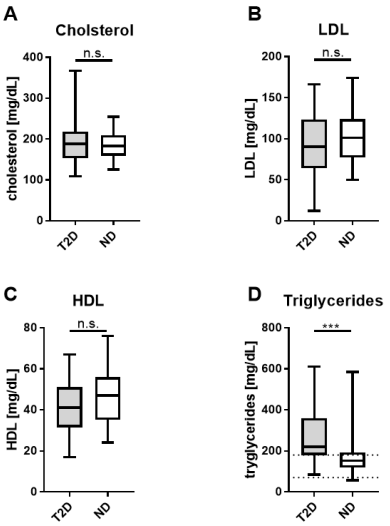


Figure 4.4: **Distribution of blood lipids whithin type 2 diabetic (T2D) and non-diabetic (ND) subjects.** The cohort is matched for cholesterol (A), LDL (B) and HDL (C) concentrations. Subjects with T2D have a higher amount of serum triglycerides (D).

Blood levels of liver enzymes AST and ALT are significantly elevated in subjects with T2D (Figure 4.5A - B) indicating a stronger hepatocellular damage in T2D. C reactive protein (CRP) is elevated (mean $>$ 5 mg/L) but not significantly altered between both groups which further proves the same level of inflammation in T2D and ND liver as already measured by *APOF* expression (Figure 4.5C). The thyroid-stimulating hormone (TSH) is slightly elevated in ND subjects (Figure 4.5D). Cortisol is a hormone which is secreted in a diurnal rhythm by the adrenal cortex and responsible for an increase in blood glucose. Cortisol levels are significantly elevated in T2D leading further to hyperglycemia and increase in insulin secretion (Figure 4.5E). Though a deficiency of vitamin B12 is reported as side effect of T2D medication, concentrations between both subgroups are not altered (Figure 4.5E) (Liu et al., 2006; Akinlade et al., 2015). Folic acid is needed for DNA methylation since the folic acid to folat metabolic pathway produces its substrate SAM (S-adenosyl methionine) using vitamin B12 as co-factor (Crider et al., 2012). A higher folic acid level is reported to have a favorable disease course (Zhao et al., 2018). Folic acid concentrations are also significantly elevated in T2D which might be induced by supplements (Figure 4.5G).

All in all, alterations in TSH, cortisol, vitamin B12 or folic acid might also be explained by additional medication and supplements. In conlucion, further regression analysis focuses on glycaemic traits (glucose, insulin and HbA1c) and triglycerides as major lipid trait. AST and ALT serve as a marker for hepatocellular damage.

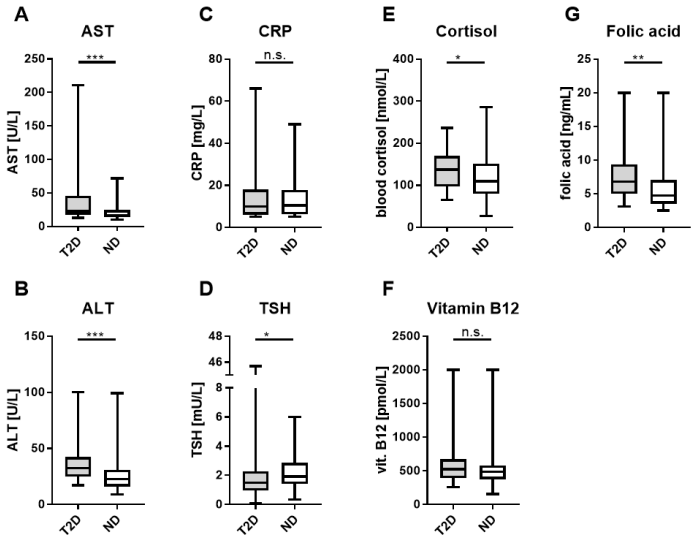


Figure 4.5: **Distribution of different enzymes, hormones and vitamins whithin type 2 diabetic (T2D) and non-diabetic (ND) subjects.** Blood levels of liver enzymes AST (aspartate aminotransferase, A) and ALT (alanine aminotransferase, B) are significantly altered between both groups. Inflammation marker CRP is not changed (C). The thyroid stimulating hormone (TSH) is altered between both groups (D). Cortisol is significantly altered between both groups (E). Vitamin B12 is not altered (F). Folic acid is significantly increased in subjects with T2D (G).

## 4.2 Materials

### 4.2.1 Equipment

Table 2: List of equipment

| Name | Manufacturer |
|---|---|
| Bead Ruptor 24 | Omni International, Kennesaw, US |
| Blotting chamber | Bio-Rad, Hercules, US |
| Cell culture clean bench Misc advanced | Thermo Fisher Scientific, Waltham, US |
| Centrifuge 5430R | Eppendorf, Hamburg, DE |
| Centrifuge MC 6 | Sarstedt, Nümbrecht, DE |
| ChemiDoc Touch Imager | Bio-Rad, Hercules, US |
| Fixed-angle rotor | Eppendorf, Hamburg, DE |
| Gene Amp PCR system 9700 | Applied Biosystems, Foster City, US |
| GeneChip Fluidics Station 450 | Applied Biosystems, Foster City, US |
| GeneChip Hybridization Oven 645 | Applied Biosystems, Foster City, US |
| GeneChip Scanner 3000 | Applied Biosystems, Foster City, US |
| Incubator with $CO_2$ Hera cell 150 | Thermo Fisher Scientific, Waltham, US |
| Light microscope Axiovert 40 CFL | Zeiss, Jena, DE |
| Microplate reader CLARIOstar Plus | BMG Labtech, Ortenberg, DE |
| Magnetic stirrer MOD 205 | VWR, Radnor, US |
| Mastercycler Nexus Gradient | Eppendorf, Hamburg, DE |
| Mastercycler Nexus X2 eco | Eppendorf, Hamburg, DE |
| Multifuge 35A Heraeus | Fisher Scientific, Schwerte, DE |
| Multipette E3 | Eppendorf, Hamburg, DE |
| Perfectblue gelsystem mini L,S | VWR, Radnor, US |
| Pipettes | Eppendorf, Hamburg, DE |
| Pipettus Akku | Hirschmann, Neckartenzlingen, DE |
| Powerpack basic power supply | Bio-Rad, Hercules, US |
| PyroMark Q48 Autoprep pyrosequencer "Quenzy" and cartridges | Qiagen, Venlo, NL |
| QuantStudio 5 Real-Time PCR System | Applied Biosystems, Foster City, US |
| Quantus Fluorometer | Promega, Madison, US |
| Rotor A-2-MTP | Eppendorf, Hamburg, DE |
| Scales Kern PCB 1000-1 | KERN & SOHN GmbH, Balingen, DE |
| ThermoMixer C | Eppendorf, Hamburg, DE |
| Vortex mixer orbit | Fisher Scientific, Schwerte, DE |
| Vortex mixer 7-2020 | neoLab Migge, Heidelberg, DE |
| Upright gel chamber Minigel-Twin | Biometra, Analytik Jena AG, Jena, DE |
| Water bath Grant OLS 200 | Grant Instruments, Cambridge, UK |

### 4.2.2 Consumable

Table 3: List of consumables

| Name | Manufacturer |
|---|---|
| Biodyne B nylon membranes | Life Technologies, Carlsbad, US |
| Biosphere Filter Tip 10 neutral | Sarstedt, Nümbrecht, DE |
| Cell culture plate, 6 well, 12 well, 96 well | Sarstedt, Nümbrecht, DE |
| Cell lifter | Sarstedt, Nümbrecht, DE |
| Combitips advanced PCR clean | Eppendorf, Hamburg, DE |
| Combitips Plus Biopur | Eppendorf, Hamburg, DE |
| Cryo tubes 2 mL | Sarstedt, Nümbrecht, DE |
| Falcon tubes, 15 mL, 50 mL | Sarstedt, Nümbrecht, DE |
| Felt-tip pen | Sarstedt, Nümbrecht, DE |
| Filter paper, extra thick | Bio-Rad, Hercules, US |
| Filtertips, 0-1000 $\mu$L | Nerbe plus, Winsen, DE |
| Filtropur V25 250 mL, 0.2 $\mu$m | Sarstedt, Nümbrecht, DE |
| Foam pads | Bio-Rad, Hercules, US |
| Gelloader pipette tips, 200 $\mu$L | Sarstedt, Nümbrecht, DE |
| GeneChip miRNA Array 4.0 | (Affymetrix) Applied Biosystems, Foster City, US |
| MicroAmp clear adhesive film | Applied Biosystems, Foster City, US |
| MicroAmp Fast 96-Well Reaction Plate | Applied Biosystems, Foster City, US |
| Membrane filter, 0.1 $\mu$L | Merck Millipore, Darmstadt, DE |
| Multiply Strip 0.2 mL | Sarstedt, NümbrechDE |
| Novex 6 % DNA Retardation Gels, 12 wells | Life Technologies, Carlsbad, US |
| Optical 96-Well Clear Reaction Plates | Applied Biosystems, Foster City, US |
| PCR cover chain | Sarstedt, Nümbrecht, DE |
| PCR tubes, 0.5 mL | Promega, Madison, US |
| Petri dish w/ cams | Sarstedt, Nümbrecht, DE |
| Pipettes, serological, 5 mL, 10 mL, 25 mL | Sarstedt, Nümbrecht, DE |
| Pipette tips, 20 $\mu$L | Sarstedt, Nümbrecht, DE |
| Plating spatula | Sarstedt, Nümbrecht, DE |
| Pleated filter | Whatman, GE Healthcare |
| Precellys ceramic beads | peqlab/VWR Life Science Competence Center, Erlangen, DE |
| PyroMark Q48 absorber stripes | Qiagen, Venlo, NL |
| PyroMark Q48 discs | Qiagen, Venlo, NL |
| PyroMark Q48 magnetic beads | Qiagen, Venlo, NL |
| Reaction tube, 1.5 mL, 2 mL, 5 mL | Sarstedt, Nümbrecht, DE |
| SafeSeal reaction tube, 1.5 mL | Sarstedt, Nümbrecht, DE |
| Screw cap micro tubes, 2 mL | Sarstedt, Nümbrecht, DE |
| Tissue culture flask T-75, polystyrene for adherent cells | Sarstedt, Nümbrecht, DE |

### 4.2.3 Chemicals

Table 4: List of chemicals

| Name | Manufacturer |
|---|---|
| Agarose Broad Range | Carl Roth, Karlsruhe, DE |
| Boric acid | Sigma Aldrich, St. Louis, US |
| Bovine Serum Albumin (BSA), 7.5 % | Life Technologies, Carlsbad, US |
| Chloroform | Carl Roth, Karlsruhe, DE |
| D (+) Glucose, 99.5 % | Carl Roth, Karlsruhe, DE |
| Diethyl pyrocarbonate (DEPC) | Carl Roth, Karlsruhe, DE |
| DEPC treated water | Life Technologies, Carlsbad, US |
| Dimethyl sulfoxide (DMSO) | AppliChem GmbH, Darmstadt, DE |
| Ethylenediaminetetraacetic acid (EDTA) solution 0.5 M, pH 8.0 | AppliChem GmbH, Darmstadt, DE |
| Ethanol 70 %, denatured | Carl Roth, Karlsruhe, DE |
| Ethanol 99.8 %, denatured | Carl Roth, Karlsruhe, DE |
| Ethanol 99.8 %, pure | Carl Roth, Karlsruhe, DE |
| Fetal Bovine Serum (FBS), heat-inactivated One Shot | Life Technologies, Carlsbad, US |
| GeneRuler 100 bp Ladder | Life Technologies, Carlsbad, US |
| High-performance liquid chromatography (HPLC) grade clean water, sterile | Biochrom, Berlin, DE |
| Insulin aspart NovoRapid | Novo Nordisk, Bagsvaerd, DK |
| Isopropanol, pure | Fisher BioReagents, Thermo Fisher Scientific, Waltham, US |
| LightShift Poly dI-dC | Life Technologies, Carlsbad, US |
| Lipofectamin 2000 | Life Technologies, Carlsbad, US |
| Lipofectamin 3000 | Life Technologies, Carlsbad, US |
| Loading buffer, 6x | Life Technologies, Carlsbad, US |
| Penicillin/Streptomycin | Life Technologies, Carlsbad, US |
| Proteinase inhibitor cocktail | Merck Millipore, Darmstadt,DE |
| Pyrophosphatase, inorganic (0.1 U/$\mu$L) | Life Technologies, Carlsbad, US |
| QIAzol Lysis Reagent | Quiagen, Venlo, NL |
| RNase Inhibitor, 2000 U | Life Technologies, Carlsbad, US |
| Sodium pyruvate, 100 mM | biowest, Nuaillé, FR |
| SYBR Safe DNA gel stain | Invitrogen, Carlsbad, US |
| Trizma Base, 2-Amino-2-(hydroxymethyl)-1,3-propanediol | Sigma Aldrich, St. Louis, US |
| Trypan Blue Solution, 0.4 % | Sigma Aldrich, St. Louis, US |

### 4.2.4 Solutions, media and buffers

Table 5: List of solutions, media and buffer

| Name | Ingredients | Manufacturer and usage |
|---|---|---|
| Dulbecco's Modified Eagle Medium (DMEM) | w/o glucose, sodium, HEPES | Life Technologies, Carlsbad, US. Basic cell culture medium |

| | | |
|---|---|---|
| Dulbecco's Modified Eagle Medium (DMEM) high glucose (4.5 g/L) | GlutaMAX Supplement | Life Technologies, Carlsbad, US. Basic cell culture medium |
| Dulbecco's Modified Eagle Medium (DMEM) low glucose (1.5 g/L) for HepG2 cultivation | Mixture of DMEM high glucose and w/o glucose, 10 % (v/v) FBS, 200 mM glutamine, 1 mM sodium pyruvate, 1 % penicillin/streptomycin | Self-diluted. For maintenance of HepG2 cells |
| Dulbecco's Phosphate-Buffered Saline (DPBS) | w/o calcium, magnesium | Life Technologies, Carlsbad, US. Basic cell culture buffer |
| Fast Media Amp Agar solid medium | 1 pouch in 200 mL water, w/ ampicillin | InvivoGen, San Diego, US. Selection of bacteria containing plasmids with ampicillin resistance |
| Fast Media Zeo Agar solid medium | 1 pouch in 200 mL water, w/ Zeocin | InvivoGen, San Diego, US. Selection of bacteria containing plasmids with Zeocin resistance |
| Fast Media Amp (Lysogency Broth) LB liquid medium | 1 pouch in 200 mL water, w/ ampicillin | InvivoGen, San Diego, US. Growth of bacteria containing plasmids with ampicillin resistance |
| Fast Media Zeo (Terrific Broth) TB liquid medium | 1 pouch in 200 mL water, w/ Zeocin | InvivoGen, San Diego, US. Growth of bacteria containing plasmids with Zeocin resistance |
| Freezing medium | DMEM low glucose medium + DMSO 20 % (v/v) | Self-diluted. For freezing of cells. |
| Opti-MEM | | Life Technologies, Carlsbad, US. Reduced serum media for transfection |
| S.O.C. medium | S.O.B. (super optimal broth) medium with glucose supplement: 2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl 2, 10 mM $MgSO_4$, 20 mM glucose. | Invitrogen, Carlsbad, US. For final step in bacteria transformation |
| TAE buffer | 242 g tris (Trisma baze), 57.1 mL acetic acid, 100 mL 0.5 M EDTA (pH 8) ad 1000 mL water for 50X TAE, dilute 1:50 for 1X TAE | Self-diluted. For agarose gel electrophoresis |
| TBE buffer | 54 g tris (Trizma base), 27.5 g boric acid, 20 mL EDTA (0.5 M, pH 8) ad 1000 mL water for 5X TBE, dilute 1:10 for 0.5X TBE | Self-diluted. For electrophoretic mobility shift assay gel electrophoresis |
| TrypLE Express | Trypsin, EDTA | Life Technologies, Carlsbad, US. Basic cell culture enzyme for passaging of cells |

### 4.2.5 Enzymes and master mixes

Table 6: List of enzymes and master mixes

| Name | Manufacturer |
|---|---|
| Fast start advanced Master Mix for TaqMan | Thermo Fisher Scientifc Inc., Waltham, US |
| Fast Start Universal SYBR green Master Rox | Roche, Basel, CH |
| GoTaq G2 Green Master Mix | Promega, Madison, US |
| High Capacity cDNA reverse transciptase Kit | Life Technologies, Carlsbad, US |
| Phusion Hotstart II Mastermix (*Pfu*) | Thermo Fisher Scientifc Inc., Waltham, US |
| PyroMark PCR Kit | Qiagen, Venlo, NL |
| qScript miRNA cDNA Synthesis Set | Quantabio, Beverly, US |
| RNase A | Qiagen, Venlo, NL |
| RNase Free DNAse | Qiagen, Venlo, NL |
| SuperScript VILO Mastermix | Life Technologies, Carlsbad, US |
| *SssI* CpG Methyltransferase | New England Biolabs, Ipswich, US |
| TaqMan advanced miRNA cDNA Synthesis Kit | Life Technologies, Carlsbad, US |

### 4.2.6 Kits

Table 7: List of kits

| Name | Manufacturer and usage |
|---|---|
| $\beta$-Galactosidase Enzyme Assay System | Promega, Madison, US. Photometric assay for normalization of Luciferase signal due to conversion of ONPG (o-nitrophenyl-$\beta$-D-galactopyranoside) by $\beta$-galactosidase. A plasmid coding for this enzyme is co-transfected during Luciferase assay. The substrate is added to the cell lysate. |
| BCA Protein Assay Kit | Life Technologies, Carlsbad, US. Photometric assay to determine protein concentration in a suspension. A serial dilution of BSA suspension is used as standard. |
| Chemiluminescence Nucleic Acid Detection Kit | Life Technologies, Carlsbad, US. For detection of biotinylated nucleic acids on blotting membranes by chemiluminescence. |
| Dual Luciferase Reporter Assay System | Promega, Madison, US. Luminescence assay for gene expression measurement with Firefly and Renilla luciferase. |
| EpiTect Fast DNA Bisulfite Kit | Qiagen, Venlo, NL. Bisulfite conversion of DNA. Unmethylated C will be converted into T, allowing a quantification of DNA methylation. |
| FlashTag Biotin HSR RNA Labeling Kit | Applied Biosystems, Foster City, US. Reagents for preparation and hybridization of RNA on microarray. |
| GeneChip Hybridization Control Kit | Applied Biosystems, Foster City, US. Control oligonucleotides for hybridization, washing and staining procedures of microarray based gene measurement. |
| GeneChip Hybridization, Wash, and Stain Kit | Applied Biosystems, Foster City, US. Ingredients for dilution of hybridization cocktails, staining solutions and washing buffers needed during microarray preparation. |

| | |
|---|---|
| LightShift Chemiluminescence EMSA Kit | Life Technologies, Carlsbad, US. Contains buffer reagents for electrophoretic mobility shift assay, control oligonucleotides and control protein suspensions. Works together with the Chemiluminescence Nucleic Acid Detection Kit. |
| miRNeasy Mini Kit | Qiagen, Venlo, NL. Extraction of whole cell RNA from mammalian cells. |
| miRNeasy Serum/Plasma Advanced Kit | Qiagen, Venlo, NL. Extraction of whole RNA from serum and plasma. |
| miRNeasy Serum/Plasma spike-in control | Qiagen, Venlo, NL. Spike-in miRNA from *Caenorhabditis elegans*. |
| NE-PER Nuclear and Cytoplasmatic Extraction Reagents | Thermo Fisher Scientific Inc., Waltham, US. Extraction of cytoplasmatic and nuclear protein fractions from cells or mammal tissue. |
| PyroMark Q48 advanced CpG Reagents | Qiagen, Venlo, NL. Contains buffer solutions, nucleic acids and enzymes for pyrosequencing. |
| Qiagen Plasmid Midi Kit | Qiagen, Venlo, NL. Extraction of plasmid DNA from 50 mL over-night bacteria culture. |
| QIAmp DNA Mini Kit | Qiagen, Venlo, NL. Extraction of genomic DNA from mammal tissue. |
| QIAprep spin Miniprep Kit | Qiagen, Venlo, NL. Extraction of plasmid DNA from 5 mL over-night bacteria culture. |
| QuantiFluor ONE dsDNA System | Promega, Madison, US. Photometric concentration measurement of DNA. |
| Quick Ligation Kit | New England Biolabs, Ipswich, US. Contains buffer and ligase for ligation of inserts into vectors. |
| TaqMan endogenous control plate | Life Technologies, Carlsbad, US. Pre-pipetted plate containing three times 31 candidate housekeeping genes in duplicates. |
| Wizard SV Gel and PCR clean up System | Promega, Madison, US. Column-based system for clean up of PCR products from reaction solution or agarose gel. |

### 4.2.7 Antibodies

Table 8: List of antibodies

| Name | Manufacturer and description |
|---|---|
| SREBF1 | Thermo Fisher Scientifc Inc., Waltham, US Unconjugated polyclonal IgG antibody from rabbit against hamster, human, mouse and rat SREBF1. Immunogen against a region within 700 to 800 of human SREBF1. |
| SP1 | Thermo Fisher Scientifc Inc., Waltham, US Unconjugated polyclonal IgG antibody from rabbit against human and mouse SP1. Immunogen against a region within amino acids 18 to 303 of human SP1. |
| Stabilized Streptavidin-Horseradish Peroxidase Conjugate (HRP) | Thermo Fisher Scientifc Inc., Waltham, US Part of the chemiluminescence nucleic acid detection module. Detection of biotinylated DNA by chemiluminescence. |

### 4.2.8 Enzyme-linked Immunosorbent Assays (ELISAs)

Table 9: List of ELISAs

| Name | Manufacturer |
|---|---|
| Human Adiponectin | BioVendor, Brno, CZ |
| Human/Canine/Porcine Insulin Quantikine ELISA Kit | R&D Systems, Inc, Minneapolis, US |
| Human C-Peptide Quantikine ELISA Kit | R&D Systems, Inc, Minneapolis, US |

### 4.2.9 Primers, oligonucleotides and TaqMan assays

All primers and oligonculeotides were purchased as custom DNA oligos from Integrated DNA Technologies (IDT), Coralville, US.

Table 10: List of genomic primers with sequence and annealing temperature ($T_a$)

| Name | Sequence | $T_a$ |
|---|---|---|
| hGALgenSp1SREBP_Luci_F | ATT AAG CTT GAG CCA ACC CAT GAA TAG GA | 56 |
| hGALgen1633_Luci_f | ATT AAG CTT AAG TGA CTC ACC CAG GAT GG | - |
| hGALgen1633_Luci_r | AAT CCA TGG AGC CAC TCC AGA TAG GCT GA | 59 $^o$C |
| hABCG1_SNP_f | Biotin - AGT GTT TTC CAG GGC ATC TAT | 61 $^o$C |
| hABCG1_SNP_r | TAA ATT ATG CCC AAA GGA ACT AG | |
| hABCG1_SNP_s | TTA TGC CCA AAG GAA C | |
| hIRS2_genors4547213_F1 | GGG CAG TCA TGG GTG AGG | 55 $^o$C |
| hIRS2_genors4547213_R1 | Biotin - GGG AAG AGC GTG CAC CTA C | |
| hIRS2_genors4547213_S1 | TGG CCA CTG AGA TGA | |

Table 11: List of bisulfite primers with sequence and annealing temperature ($T_a$)

| Name | Sequence | $T_a$ |
|---|---|---|
| hABCG1bis_f | GGT TAG GAG TTT AAA AGG TTG AGT A | 56 $^o$C |
| hABCG1bis_r | ATC CCC AAA ACC TAA AAC CAC CTC AAT AA | |
| hABCG1bis_s | AAA TTA TGT TTA AAG GAA T | |

| | | |
|---|---|---|
| hGALbis1633_f | ATG GTA TTT GAG TTT ATA ATT TAG GGA TAG | 56 $^o$C |
| hGALbis1633_r | Biotin - CT ATA CAA CTT TCC TTC ACT ATA TAC AC | |
| hGALbis1633_s23 | GAT AGA ATT TAT ATT ATT ATT GTA AAA | |
| IRS2_cg12195446_ r4547213_CpG5_f | GTA GGG TAG TTA TGG GTG AGG | 56 $^o$C |
| IRS2_cg12195446_ r4547213_CpG5_r | Biotin - TCC CCA ACA AAA TTT AAA TAC CTC TA | |
| IRS2_cg12195446_ r4547213_CpG5_s | GGG TGG TTA TTG AGA T | |
| IRS2_CpG6_reverse_f | Biotin - GGG GTT AAG AGG AAG TTT TTG AGT AGA | 54 $^o$C |
| IRS2_cpG6_reverse_r | CAC CAC CTC CTA ACC AAT CT | |
| IRS2_cpG6_reverse_s | GGA AGG GAG TTT TTT TAT TT TTA A | |
| hTP53INP1bis_f | GTT TGA AGG TAG AGA GGT TAG TAG | 56 $^o$C |
| hTP53INP1bis_r | Biotin - CCC ACT ATC TCA CTC TCT TAT CAC | |
| hTP53INP1bis_s | GTA GAG AGG TTA GTA GG | |
| hSREBF1bis_f/s | TTT TTT TTG AAG GTA GAT GTA GG | 56 $^o$C |
| hSREBF1bis_r | Biotin - AAA ACA AAT AAA AAA CTC CCT CTT C | |

Table 12: List of genomic primers for mRNA qPCR with sequence. All primers worked with an annealing temperature ($T_a$) of 60 $^o$C.

| **Name** | **Sequence** |
|---|---|
| APOF_qPCR_f1 | GCA GAA CAG GAT CAG GAA GC |
| APOF_qPCR_r1 | GCA TGT CTG GAG CAG AGT AC |
| HsCASC3_cDNA_f | ACC TCG GAA AGG GCT CTT CTT |
| HsCASC3_cDNA_r | ACC TCG GAA AGG GCT CTT CTT |
| hsaCDKN1B_qPCR_f | AAC GTG CGA GTG TCT AAC GG |
| hsaCDKN1B_qPCR_r | CCC TCT AGG GGT TTG TGA TTC T |
| CYP2C19_qPCR_f1 | CAG GAT TGT AAG CAC CCC C |
| CYP2C19_qPCR_r1 | TCC CGG GAA ATA ATC AAT GAT AG |
| HsDIO1_cDNA_f | GTC GTG GGT AAA GTG CTT CTG |
| HsDIO1_cDNA_r | GTT CCG CTT GAC TCT GTC TGG |
| HsDIO2_cDNA_f | TCC AGT GTG GTG CAT GTC TC |
| HsDIO2_cDNA_r | CTG GCT CGT GAA AGG AGG TC |
| HsDIO3_cDNA_f | CTC TCC CTA CAT CAT CCC ACA |
| HsDIO3_cDNA_r | TGA CAT AGA GAC GCT CGA AGTA |
| hsaINSR_qPCR_f1 | AAA ACG AGG CCC GAA GAT TTC |
| hsaINSR_qPCR_r1 | GAG CCC ATA GAC CCG GAA G |
| hSLC10A1_qPCR_f | AAG GAC AAG GTG CCC TAT AAA GG |
| hSLC10A1_qPCR_r | TTG AGG ACG ATC CCT ATG GTG |
| HsSLC16A2_cDNA_f | CCA CGC CTA CGG TAG AGA C |
| HsSLC16A2_cDNA_r | CAG AGT TAT GGA TGC CGA AGA TG |
| hSLCO1C1_qPCR_f | GGA GTT GGA ACA CTG CTC ATT |
| hSLCO1C1_qPCR_r | CTT GAC TCT AGG AGA CAC GGA |
| hTHRA_qPCR_f | AGG TCA CCA GAT GGA AAG CG |
| hTHRA_qPCR_r | AGT GAT AAC CAG TTG CCT TGT C |
| HsTHRB_cDNA_f | TGG GAC AAA CCG AAG CAC TG |

| | |
|---|---|
| HsTHRB2_cDNA_r | TGG CTC TTC CTA TGT AGG CAG |

Table 13: List of genomic primers for miRNA qPCR with sequence. All primers workt with an annealing temperature ($T_a$) of 60 $^o$C.

| **Name** | **Sequence** |
|---|---|
| universal_primer_r | GCA TAG ACC TGA ATG GCG GTA |
| hsa_let_7e_5p_F2 | GCA GTG AGG TAG GAG GTT G |
| cel_miR_39_3p_qPCR | CAC CGG GTG TAA ATC AGC TTG |
| hsa_miR_24_3p_qPCR | TGG CTC AGT TCA GCA GGA ACA |
| hsa_miR_182_5p_qPCR | GCT TTG GCA ATG GTA GAA CTC A |

Table 14: List of genomic oligonucleotide sequences for Luciferase assay or EMSA. The SNP (G/A) is indicated as thick **G**.

| **Name** | **Sequence** |
|---|---|
| IRS2 CpG1/SNP | TGA C**G**G GAT |
| IRS2 CpG1-CpG3 | TGA C**G**G GAT CGG GGG CGG CTT C |
| IRS2 CpG1/SNP_CpG6 | TGA C**G**G GAT CGG GGG CGG CTT CTG CCG CAG ACC CGG GAA TTC ACG TGG TC |
| IRS2 EMSA | ACG GGA TCG GGG GCG GCT TCT GCC GCA GAC CCG GGA ATT CAC GTG GTC |

Table 15: List of TaqMan assays. All TaqMan assays were purchased from Applied Biosystems, Foster City, US.

| **ID** | **Gene/miRNA** | **Dye** |
|---|---|---|
| Hs00153715_m1 | *ACACB* | FAM |
| Hs00201226_m1 | *CASC3* | VIC |
| Hs00765553_m1 | *CCND1* | FAM |
| Hs00167982_m1 | *CYP7A1* | FAM |
| Hs04260088_m1 | *ELOVL6* | FAM |
| Hs01005622_m1 | *FASN* | FAM |
| Hs00231106_m1 | *FOXO1* | FAM |
| Hs00289325_m1 | *GALNT18* | FAM |
| Hs99999905_m1 | *GAPDH* | VIC |
| Hs00275843_s1 | *IRS2* | FAM |
| Hs01092524_m1 | *LDLR* | FAM |
| Hs00233945_m1 | *LRP6* | FAM |
| Hs00545399_m1 | *PCSK9* | FAM |
| Hs01102345_m1 | *RPL37A1* | VIC |
| Hs01682761_m1 | *SCD* | FAM |
| 477935_mir | hsa-miR-182-5p | FAM |
| 477992_mir | hsa-miR-24-3p | FAM |
| 477887_mir | hsa-miR-126-3p | FAM |
| 478056_mir | hsa-miR-361-5p | FAM |
| 478579_mir | hsa-let-7e-5p | FAM |

| | | |
|---|---|---|
| 477855_mir | hsa-miR-122-5p | FAM |
| 477983_mir | hsa-miR-223-3p | FAM |

### 4.2.10 Mimics

Table 16: List of miRNA mimics. All mimics were purchased from Ambion, Life Technologies, Carlsbad, US.

| **ID** | **Mature miRNA** |
|---|---|
| PM12369 | hsa-miR-182-5p |
| PM12304 | hsa-let-7e-5p / mmu-let-7e-5p |

### 4.2.11 Vectors and plasmids

Table 17: List of vectors

| **Name** | **Manufacturer and usage** |
|---|---|
| pCpGL-CMV-FLuc | Gift from (Klug and Rehli, 2006). The backbone of this vector is free of any CpG oligonucleotides and encodes a Firefly Luciferase gene which is under control of a CMV promoter. Respective sequences of interest can be inserted 5' of the CMV promoter by *NcoI* and *HindIII* digestion and subsequent ligation. |
| pRL-SV40 | Promega, Madison, US. This plasmid encondes a Renilla Luciferase which is under control of a SV40 promoter. It serves as signal normalisation during co-transfection with a Firefly Luciferase containing plasmid. |
| pSV-$\beta$-Galactosidase control vector | Promega, Madison, US. This plasmid encondes the gene for $\beta$ galactosidase expression which is under control of a SV40 early promoter and enhancer. It serves for signal normalisation during co-transfection with a Firefly Luciferase containing plasmid. |

Table 18: List of plasmids

| **Name** | **Contents** |
|---|---|
| pCpGL-CMV-RLuc-IRS2-CpG1 | Vector pCpGL-CMV-RLuc with insert *IRS2* intron 1 CpG1 |
| pCpGL-CMV-RLuc-IRS2-SNP | Vector pCpGL-CMV-RLuc with insert *IRS2* intron 1 SNP |
| pCpGL-CMV-RLuc-IRS2-CpG1-CpG3 | Vector pCpGL-CMV-RLuc with insert *IRS2* intron 1 CpG1 until CpG3 |
| pCpGL-CMV_RLuc-IRS2-CpG1_CpG6 | Vector pCpGL-CMV-RLuc with insert *IRS2* intron 1 CpG1 until CpG6 |
| pCpGL-CMV_RLuc-IRS2-SNP_CpG6 | Vector pCpGL-CMV-RLuc with insert *IRS2* intron 1 SNP instead of CpG1, until CpG6 |
| pCpGL-CMV_RLuc-GALNT18-CpG1_CpG3 | Vector pCpGL-CMV-RLuc with insert *GALNT18* intron 3 CpG1 until CpG3 |

| | |
|---|---|
| pCpGL-CMV_RLuc-GALNT18-GCbox_CpG1_CpG3 | Vector pCpGL-CMV-RLuc with insert *GALNT18* intron 3 CpG1 until CpG3 with additional GC box motifs and CpG-1 |

#### 4.2.12 Bacteria

One Shot PIR1 chemically competent *Escherichia coli* bacteria were purchased from Life Technologies, Carlsbad, US. These *E.coli* bacteria are suitable for cloning of plasmids containing the R6K-gamma origin for plasmid replication.

#### 4.2.13 HepG2 cells

HepG2 cells were freshly purchased from American Type Culture Collection (ATCC HB-8065). They originate from a liver hepatocellular carcinoma of a 15 years old male patient Caucasian ethnicity. In cell culture they exhibit an epithelial morphology.

#### 4.2.14 Hardware

Statistics, models and assay designs were performed on Toshiba Satellite Pro A30-C-11G (Intel Core i5-6200U 2.3 GHz, 16 GB RAM, "Ploetze"[†]) and later on Dell XPS 8930 Performance (Intel Core i7-8750H 2.2 GHz, 16 GB RAM, "Ploetze2").

#### 4.2.15 Software

Table 19: List of software

| **Name** | **Manufacturer** | **Version** |
|---|---|---|
| Eclipse | Eclipse Foundation, Inc, Ottawa, CA | Photon Release (4.8.0) |
| GraphPad Prism, statistic and visualization | GraphPad Software Inc, San Diego, US | v7.05 |
| imageJ, image processing | National Institutes of Health, Bethesda, US | v1.51j |
| LibreOffice | The Document Foundation, Berlin, DE | v6.2.0 |
| MATLAB, statistic and modeling | The MathWorks, Natick, US | R2016a - R2018b |
| Office 365 | Microsoft Corporation, Redmond, US | 2016 - 2018 |
| PyroMark Q48 Autoprep | Qiagen, Venlo, NL | v2.4.2 |
| PyroMark Assay Design | Qiagen, Venlo, NL | v2.0.1.15 |
| QuantStudio Desktop qPCR | Applied Biosystems, Foster City, US | v1.3.1 |
| RStudio, statistic and modeling | RStudio, Inc, Boston, US | v1.1.456 |
| Serial Cloner, molecular biology | SerialBasics | v2.6.1 |
| Transcriptome Analysis Console, microarray analysis | Applied Biosystems, Invitrogen, Foster City, US | v4.0.0.25 |
| Unipro UGENE, molceular biology | (Okonechnikov et al., 2012) and NCIT UNIPRO, LLC, Novosibirsk, RU | v1.27.0 |

## 4.3 Methods

### 4.3.1 DNA extraction from liver tissue

Genomic DNA was extracted from approximately 25 mg snap-frozen liver using the QIAmp DNA Mini Kit as indicated by manufacturer with minor changes. Frozen tissue was transferred into 2 mL screw cap micro tubes with five to six ceramic beads and stored on ice until homogenization in 80 $\mu$L ice cold DPBS. Homogenization was performed by Bead Rupter 24 for two times 20 sec shaking. Afterwards, 180 $\mu$L tissue lysis (ALT) buffer and 20 $\mu$L proteinase K included in the extraction kit was added and the suspension was incubated for at least 15 min to 60 min at 56 $^o$C. Homogenized and digested liver lysate was stored for a maximum of 2 months at room temperature prior final DNA extraction.
DNA extraction was performed as indicated by manufacturer. Therefore 200 $\mu$L lysate was used. After every washing step, the flow through and collection tube was discarded. DNA was eluted in 200 $\mu$L AE water included in the extraction kit.
Finally DNA concentrations were determined by QuantiFluor ONE dsDNA System as indicated by manufacturer. Therefore 1 $\mu$L of the DNA eluate was used in 199 $\mu$L fluorescent double-stranded (ds) DNA-binding dye solution and measured by a Quantus fluorometer. DNA was stored at -20 $^o$C until further use.

### 4.3.2 Bisulfite conversion of genomic DNA

The EpiTect Fast DNA Bisulfite Kit was used for bisulfite conversion. Two distinct steps were performed: A conversion of DNA causing unmethylated C to be changes into U (which will be recognized as T later during bisPCR) and a column based clean-up system.
For the conversion step, 20 $\mu$L DNA eluate was used (which varies between 100 and 2000 ng input DNA) in 140 $\mu$L bisulfite conversion reaction as indicated by manufacturer. The EpiTect Fast DNA Bisulfite Kit includes besides a bisulfite treatment solution also a DNA protect solution which reduces strong damages on the DNA which might cause a bias in downstream bisulfite PCR (bisPCR). After preparation, the reaction was denatured for 5 min at 95 $^o$C, followed by a first incubation for 30 min at 60 $^o$C, a second denaturation step for 5 min at 95 $^o$C and a final incubation for 10 min at 60 $^o$C. All reactions were stored at 4 $^o$C until clean-up.
For clean-up, the complete reaction was transferred to respective buffers from the kit and cleaned by columns. In a final step, bisulfite converted DNA (bisDNA) was eluted in 20 $\mu$L EB buffer. BisDNA was stored at -20 $^o$C until further use.

### 4.3.3 Bisulfite polymerase chain reaction (bisPCR)

For bisulfite polymerase chain reaction (bisPCR), specific primers were designed by PyroMark Assay Design software. Most importantly one has to avoid primers which contain themselves a possible CpG oligonucleotide side to reduce the chance of PCR bias inclusion. A complete set contained three primers (forward, reverse and a sequencing) whereby the forward (reverse assay on reverse DNA strand) or reverse primer (normal assay on forward DNA strand) was biotinylated at the 5' end. In a final step of primer design, the binding probability and specificity of the primers to the region of interest were checked and optimized. The conversion of every unmethylated C to a T results in a strong reduction of sequence variability causing a loss in specificity.
After primer design and ordering at IDT, primers were diluted to a 10 $\mu$M working dilution and stored at 4 $^o$C for short time or -20 $^o$C for long time. BisPCR was performed using reagents from the PyroMark PCR kit. The master mix includes a blend from different polymerases and is optimized for bisulfite treated DNA containing long homopolymers of one base. Each 25 $\mu$L reaction contained 12.5 $\mu$L master mix, 2.5 $\mu$L loading dye CoralLoad, 0.5 $\mu$L of each forward and reverse primer, 8 $\mu$L nuclease-free water

and 1 $\mu$L bisDNA. Afterwards the reaction was incubated for 15 min at 95 $^o$C in a initial denaturation step and for 40 to 45 times denatured for 30 sec at 94 $^o$C, incubated for 30 sec for primer annealing at their respective temperature (see Table 11) and incubated for 30 sec at 72 $^o$C for polymerase elongation. A final elongation step was included for 10 min at 72 $^o$C.
Afterwards 5 to 6 $\mu$L bisPCR product was separated by agarose gel electrophoresis using 1 % agarose gel in 1xTAE buffer and SYBR save gel stain as dsDNA intercalator. An empty reaction containing water instead of bisDNA served as control for contamination.

### 4.3.4 Genomic polymerase chain reaction (gPCR)

Two different systems were used for genomic PCR. For complex GC-rich sequences, a *Pfu* polymerase (Phusion Hotstart II) was used. Smaller sequenced with a low melting temperature were amplified by a *Taq* polymerase (GoTaq G2).
For a 20 $\mu$L *Pfu* gPCR reaction, 10 $\mu$L 2X master mix, containing polymerase, buffer and dNTPs, was mixed at room temperature with 1 $\mu$L 10 $\mu$M forward and reverse primer, 10 - 15 ng gDNA and respective amount nuclease-free water. Afterwards the reaction was transferred into a thermocycler and initially denatured for 2 min at 95 $^o$C. The following steps were performed for 35 cycles: Denatured for 15 sec at 95 $^o$C, incubated 15 sec at respective annealing temperature of the used primer and the sequence was elongated for 15 sec (15-30 sec/kb) at 72 $^o$C. In a final elongation step, the reaction was incubated for 5 min at 72 $^o$C. To check for primer specificity and product amplification, 1 - 2 $\mu$L PCR product was mixed with 4 - 5 $\mu$L 6X loading buffer and separated by 1 % agarose gel electrophoresis in 1X TAE buffer.
For a 25 $\mu$L *Taq* gPCR reaction, 12.5 $\mu$L 2X master mix, containing polymerase, buffer, dNTPs and loading dye, was mixed on ice with 0.5 $\mu$L 10 $\mu$M forward and reverse primer, 10 - 15 ng gDNA and respective amount nuclease-free water. For following steps, the reaction was transferred into a thermocycler. An initial denaturation was performed for 5 min at 95 $^o$C. For 35 cycles, the reaction was denatured for 15 sec at 95 $^o$C, incubated for 15 sec at the respective annealing temperature of the primer and the sequence elongated for 30 - 60 sec (60 sec/kb) at 72 $^o$C. In a final elongation step, the reaction was incubated for 5 min at 72 $^o$C. To check for primer specificity and product amplification, 5 - 6 $\mu$L PCR product was separated by 1 % agarose gel electrophoresis in 1X TAE buffer.
Genomic PCR was performed for sequencing of polymorphism or generation of inserts for vector ligation.

### 4.3.5 Bisulfite pyrosequencing

During bisulfite pyrosequencing, the nucleotides ATP, TTP, CTP and GTP were dispensed in a known order and incorporated complementary to a biotinylated template strand as elongation of a sequencing primer. Each addition of a nucleotide is measured by conversion of free pyrophosphatate (PPi) by an enzyme (5'-phosphosulpfate, APS) to ATP. This product serves as a substrate of the conversion of luciferin to oxyluciferin which will result into a measurable light signal.
Due to the low costs, high sensitivity and resulting high resolution, bisulfite pyrosequencing is the gold standard of DNA methylation measurement. Also detection of polymorphisms (genotyping) or *de novo* sequencing is possible with this technique. A limiting factor of this method is the site specific measurement which allows only small gPCR or bisPCR products to be sequenced. Therefore it is not suitable for high throughput whole genome DNA methylation measurement.
The PyroMark Q48 Autoprep pyrosequencer cleans the PCR product automatically from complementary, unbiotinylated DNA strand, not incorporated oligonucleotides and remaining polymerases. Accordingly, all reagents included in the PyroMark Q48 advanced CpG Kit were used. Additionally equipment is needed (disc, absorber stripe) which can be purchased from Qiagen.
For pyrosequencing, 10 $\mu$L PCR product was used. 3 $\mu$L of a suspension containing magnetic beads which are conjugated with streptavidin were added to the PCR product resulting in a strong non-covalent

binding between biotinylated DNA strand and beads which will enable automatic cleaning of the PCR product from nucleotides, primers and proteins by centrifugation.
Assay preparation and evaluation of each run can be created and visualized by PyroMark Q48 Autoprep software. DNA methylation is calculated in %. The accuracy of DNA methylation measurement is about +/- 1 %. Furthermore, each assay was evaluated by sequencing fully, hemi or non methylated control DNA.

### 4.3.6 Whole cell RNA extraction from liver tissue

Whole cell RNA, including miRNA and mRNA, was extracted from approximately 25 mg snap-frozen liver using the miRNeasy Mini Kit as indicated by manufacturer. Therefore frozen tissue was transferred into a 2 mL screw cap micro tube with five to six ceramic beads and stored on ice until homogenization in 700 $\mu$L QIAzol lysis reagent. Homogenization was performed by Bead Rupter 24 for two times 20 sec shaking. The lysates were stored at -80 $^o$C until further extraction. All instruments and tools were previously cleaned with RNase degrading solutions.
For RNA extraction, lysates were thawed at 37 $^o$C and 140 $\mu$L cholorform was added. All reactions were shaken for 15 sec and incubated at room temperature for 2 - 3 min. A separation of two phases (organic phenol phase and an aqueous chloroform phase containing RNA) was achieved by centrifugation for 15 min at 4 $^o$C. Afterwards, the RNA was precipitated by mixing 375 $\mu$L absolute ethanol to 250 $\mu$L aqueous phase. Clean-up of the RNA was performed by columns as indicated by manufacturer. An optional DNase digestion was performed by adding 80 $\mu$L DNase dilution to the columns and incubation for 15 min at room temperature.
RNA was eluated in 50 $\mu$L nuclease-free water and further diluted to achieve concentrations between 500 and 1000 ng/$\mu$L. Concentration of the RNA dilutions was measured photometrically. The quality of the RNA was estimated by a 260 nm/280 nm absorbance ratio. This ratio indicates the purity of a RNA solution, since RNA absorbs at 280 nm and contaminants like proteins or phenol at 260 nm. All RNA dilutions which were used for further experiments exhibit an absorbance ratio $> 2.1$.

### 4.3.7 Whole RNA extraction from serum

The miRNeasy Serum/Plasma Advanced Kit was used to extract RNA molecules from 200 $\mu$L human serum originating from proteins and extracellular vesicles. A spike-in miRNA (miRNeasy Serum/Plasma spike-in control cel-miR-39-3p from *Caenorhabditis elegans*) was added after lysis in fixed amounts ($5.6 * 10^8$ copies/$\mu$L) for normalization of downstream concentration measurements.
All extraction steps were performed as indicated by manufacturer. Serum miRNA were eluted in 25 $\mu$L nuclease-free water. A second elution step was applied using the first eluate as input to concentrate the final amount of extracted miRNA. All miRNA dilutions were stored at -80 $^o$C until further use.

### 4.3.8 Synthesis of mRNA-cDNA for hepatic gene expression measurement

RNA is an unstable molecule which cannot be amplified by PCR. Therefore RNA has to be reverse transcribed into complementary or copy (c)DNA. All subsequent steps were performed on ice to avoid mRNA degradation and all instruments and tools were previously cleaned with RNase degrading solutions.
For cDNA synthesis, 2 $\mu$g RNA was used in a 20 $\mu$L reaction and reverse transcribed by SuperScript VILO master mix with reverse transcriptase (RT). Therefore 14 $\mu$L RNA dilution was mixed with 4 $\mu$L 5X Vilo mix and 2 $\mu$L RT. Moreover, the synthesis included a no enzyme and no template control. Afterwards the reaction was incubated for 10 min at 25 $^o$C, 60 min at 42 $^o$C (cDNA synthesis) and the RT inactivated for 5 min at 85 $^o$C.
The cDNA stock solution was diluted 1:40 to a final concentration of 2.5 ng/$\mu$L for subsequent gene expression measurement by quantitative real time (q)PCR.

### 4.3.9 Synthesis of miRNA-cDNA for hepatic gene expression measurement and concentration measurement of serum miRNAs

Two systems were used for the synthesis of miRNA-cDNA based on the specificity of used assays or primers and the abundance of the transcript of interest.
For highly abundant hepatic miRNAs or cell culture experiments, a TaqMan based assay was used to measure miRNA expression. The respective cDNA synthesis kit (TaqMan advanced miRNA cDNA Synthesis Kit) needed as input 10 ng of whole cell RNA. All subsequent steps of cDNA synthesis were performed as indicated by manufacturer. These steps include a poly-A-tail ligation on the 5' end and an adaptor ligation on the 3' end of the miRNA. In a final step, miRNA-cDNA is amplified to increase the amount of cDNA. The amplified reaction (miR-Amp) was diluted 1:10 for downstream measurements. In the TaqMan qPCR reaction, respective TaqMan assays will include a forward primer recognizing the adaptor sequence, a miRNA specific reverse primer overlapping with the poly-A-tail and a TaqMan probe binding specifically the miRNA region.
For measurement of low abundant hepatic and serum miRNAs, cDNA synthesis was performed using the qScript miRNA Synthesis Set with an input of either 1 $\mu$g hepatic whole cell RNA or 5 $\mu$L serum eluate which corresponds to 40 $\mu$L serum input. The synthesis includes addition of a poly-A-tail by a poly-A-polymerase followed by cDNA synthesis using poly-A-specific primers which add a sequence for universal primer recognition. The miRNA-cDNA stock solution was diluted 1:30 for SYBR green qPCR. Stock miRNA-cDNA and dilutions were stored at -20 $^{o}$C until further use.
Both synthesis included a no reverse transcriptase and no template control.

### 4.3.10 Microarray measurement of whole cell non-coding (nc)RNA

Measurement of low molecular weight (LMW) RNA by GeneChip miRNA 4.0 Array analysis represents a fast and efficient high-throughput method to detect approximately 2.500 mature human miRNA, 2.000 human pre-miRNAs and 2.000 human small nucleolar (sno)RNA or small Cajal body-specific (sca)RNA at once.
The complete cohort was divided into a discovery cohort of about 20 T2D and 20 ND subjects, which differed significantly in HbA1c, HOMA-IR, blood glucose and NAS (Figure 4.6A-D) but were matched for age, BMI, gender and blood lipids (Figure 4.6E-F). 500 ng whole cell RNA was used as input as recommended by manufacturer.
All subsequent steps of probe preparation, hybridization, staining and washing was performed in cooperation with the Institute of Experimental and Clinical Pharmacology, University Hospital Schleswig-Holstein, Campus Kiel, DE, supervised by Prof. Dr. Cascorbi.
Probe preparation was performed by the FlashTag Biotin HSR RNA Labeling kit. Therefore, 500 ng RNA was first used for a 3' poly-A tailing reaction. In a second step, the poly-A-tail was used for adapter ligation of biotin-labeled 3DNA dendrimer, which was afterwards detected by streptavidin conjugated PE during array measurement. Control Oligo B2 from the GeneChip Hybridization Control kit was used besides the RNA probes. Other quality control assays were not performed during array preparation. After 18 h incubation at 48 $^{o}$C and 60 RPM, the hybridization cocktail was discarded. Subsequent washing and staining was performed automatically by the Fluidics Station 450 using the respective protocol (FS450_0003). Scanning was performed by GeneChip Scanner 3000 within 48 h after array preparation.

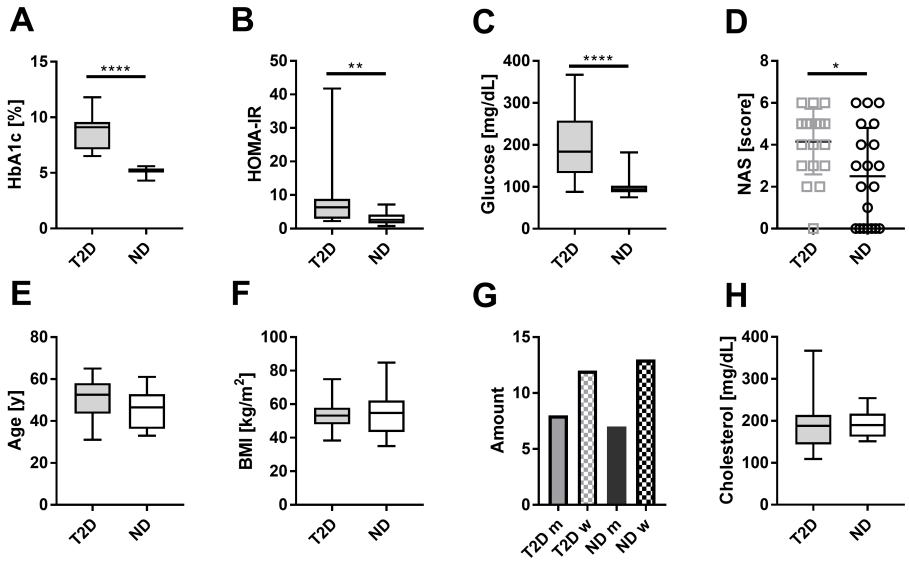


Figure 4.6: **Phenotypic characterization of the discovery cohort (n = 40) used for miRNA microarray analysis.** The T2D (n = 20) and ND (n = 20) subjects differ significantly in HbA1c (A), insulin resistance (B), blood glucose (C) and liver fibrosis and steatosis (D). Both subgroups were matches for age (E), BMI (F), gender (G) and cholesterol (H).

### 4.3.11 TaqMan quantitative real time (q)PCR

Quantitative real time (q)PCR is a technique to estimate gene expression by measuring the amount of mRNA-cDNA or miRNA-cDNA within a sample. Since it uses a revere transcription (RT) step of mRNA into cDNA prior measurement, this method is also referred as RT-qPCR. For normalization by relative quantification, a housekeeping gene or spike-in control is needed. Candidate housekeeping genes can be validated by pre-pipetted endogenous control plates or by literature research and qPCR validation.
TaqMan qPCR needs besides a forward and a reverse primer for target amplification also an intern probe, which is 5' conjugated with a fluorescent dye (mostly FAM or VIC) and 3' conjugated with a nonfluorescent quencher which inhibits an initial fluorescent signal. During the elongation phase of the PCR, the polymerase reaches and cuts the probe which releases the dye and allows a fluorescent signal without quenching. The amount of fluorescence accumulates every amplification cycle until a threshold above background signal is reached. The respective cycle is referred as Ct value. TaqMan assays are validated by manufacturer, therefore they do not need further experimental validation.
For measurement of hepatic mRNA expression, 10 $\mu$L reactions containing 10 ng mRNA-cDNA, 0.5 $\mu$L 20X TaqMan assay, 5 $\mu$L TaqMan advanced master mix and respective nuclease-free water were measured in duplicates. The PCR protocol included an initial denaturation step for 20 sec at 95 $^{o}$C followed for 40 cycles by denaturation at 95 $^{o}$C for 1 sec and an annealing/elongation step at 60 $^{o}$C for 20 sec. All qPCR reactions included a no template control.
Hepatic miRNA expression was measured by using 4 $\mu$L 1:10 dilution of the miR-Amp reaction which is included during miRNA-cDNA synthesis. 20X TaqMan assay, master mix and nuclease-free water was added as described before. The PCR protocol included an initial denaturation step for 20 sec at 95 $^{o}$C followed for 40 cycles by denaturation at 95 $^{o}$C for 3 sec and an annealing/elongation step at 60 $^{o}$C for 30 sec. Measurement of hepatic hsa-miR-182-5p was not reliable by TaqMan assay, therefore a SYBR green approach was used for all low abundant miRNAs which used more input whole cell RNA for miRNA-cDNA synthesis (see sections 4.3.9 and 4.3.12). Candidate housekeeping genes hsa-miR-24-3p, hsa-miR-126-3p and hsa-miR-361-5p were expected to be highly expressed, therefore TaqMan assays were a reliable system for validation.

### 4.3.12 SYBR green quantitative real time (q)PCR

In comparison to TaqMan qPCR, SYBR green qPCR needs only a forward and reverse primer for amplification of a certain cDNA fragment. Each elongation step results in a double stranded DNA product which is needed for intercalation of the fluorescent dye SYBR green. Thereby the amount of fluorescent signal is proportional to the amount of double stranded DNA product. The Ct value states the cycle where the fluorescent signal reaches a value above background similar to TaqMan qPCR. Since SYBR green qPCR uses custom made primer pairs, a standard curve is needed to elucidate the efficiency and specificity of the primers. An efficiency of 2 is desired, meaning a doubling of each PCR product in every amplification cycle. An efficiency above or below 2 indicates a by-product (missing specificity) or insufficient annealing which might be caused by a false annealing temperature. Efficiencies above 1.85 and below 2.15 were considered sufficient.
Each 10 $\mu$L SYBR green reaction contained 200 nM forward and reverse primer, 5 $\mu$L SYBR green master mix, 10 ng mRNA-cDNA and respective nuclease-free water. For miRNA measurement, a universal primer recognizing a DNA tag which was ligated during cDNA synthesis was used instead of a reverse primer. Approximately 4 $\mu$L hepatic miRNA-cDNA or serum miRNA-cDNA of a 1:30 stock dilution was used as qPCR input, which corresponds to either 6.7 ng hepatic whole cell RNA or 267 nL serum. All qPCR reactions contained also a no template control. The PCR protocol included an initial denaturation step for 10 min at 95 $^o$C followed for 40 cycles by denaturation at 95 $^o$C for 15 sec and an annealing/elongation step at 60 $^o$C for 60 sec. Finally, a melt curve analysis of the amplified DNA product was included to ensure specificity during primer annealing and product elongation.

### 4.3.13 Restriction digestion and ligation of plasmid DNA

Inserts for vector ligation were either generated by genomic PCR (gPCR) or dimerization of single stranded oligonucleotides bought from IDT. Inserts have to include a recognition motif for restriction enzymes *NcoI* (5' end) and *HindIII* (3' end) followed by three additional bases for effective digestion. For gPCR, see section 4.3.4.
Dimerization of single stranded oligonucleotides was performed by mixing 100 $\mu$M forward and reverse complement oligonucleotides and incubation for 5 min at 95 $^o$C and gradually decreasing temperature (-1 $^o$C/min).
Restriction digestion was performed by using 500 - 1000 ng DNA in 50 $\mu$L reactions containing 5 $\mu$L 2X buffer, 1 $\mu$L of both enzymes and nuclease-free water. Each reaction was incubated for 80 min at 37 $^o$C and the enzymes were heat inactivated for 20 min at 80 $^o$C. Reactions were cleaned from salt and proteins by Millipore membrane filters.
Ligation was performed by a QuickLigation Kit in 20 $\mu$L reactions on ice. Each reaction contained 50 ng digested vector and respective amounts of insert to yield a 1:3 ratio of vector to insert. After mixing by pipetting, the reaction was incubated for 10 min at room temperature. Until further use, each ligation reaction was stored at -20 $^o$C.

### 4.3.14 Transformation and cloning

Chemically competent *Escherichia coli* bacteria One Shot PIR1 (50 $\mu$L aliquots) were thawed on ice and 5 $\mu$L ligation reaction was added. Cells were incubated for 30 min on ice and a heat shock was performed for 30 sec at 42 $^o$C. After a short cooling on ice, 250 $\mu$L SOC medium was added and bacteria suspensions were incubated for 60 min at 37 $^o$C and 300 rpm.
For selection, transformed bacteria were spread on agar plate containing a selective antibiotic. Therefore only bacteria containing a vector with a gene for respective antibiotics resistance will survive and form colonies.

After an over night incubation at 37 $^{o}$C, colonies can be picked with a pipette tip and transferred into liquid media to grow over night cultures for plasmid extractions. Alternatively, a colony PCR can be performed by dilution of a small sample of the bacteria colony in 15 $\mu$L TE buffer and heating of the suspension for 5 min at 96 $^{o}$C. Afterwards, the suspension was centrifuged for 10 min at 18.000 g and 2 $\mu$L of the supernatant was used in a PCR reaction.

#### 4.3.15 Mini and Midi plasmid preparation

Mini and midi plasmid preparations were performed to extract plasmid DNA from different sizes of overnight bacterial cultures. Therefore bacteria in either 5 mL (mini) or 50 mL (midi) antibiotics containing liquid broth were incubated gently shaking at 37 $^{o}$C. Afterwards, bacteria were pelletized by centrifugation for 15 min at 4 $^{o}$C and 6.000 g. Bacteria were stored at -20 $^{o}$C until further extraction by respective plasmid extraction kits. Subsequent steps were performed as indicated by manufacturer.
Plasmids from mini preparations were eluted in 20 $\mu$L elution buffer (buffer EB) and ready to used for sequencing or luciferase assays. For midi preparations, DNA was eluted in 3 mL elution buffer (buffer QF). After isopropanol precipitation of the DNA and washing in 70 % ethanol, plasmids from midi preparations were suspended in 120 $\mu$L TE buffer and ready to use for further applications.
To ensure a correct sequence for donwstream analysis, plasmids were sequenced at Eurofins GATC (formerly GATC Biotech), Köln, DE.

#### 4.3.16 Cultivation of HepG2 cells

Thawing of HepG2 cells was performed by transferring cells from one cyrotube into 9 mL DMEM low glucose medium with 10 % FBS. Cells were centrifuges for 7 min at 125 g and old medium was discarded. The cell pellet was resolved in 10 mL fresh DMEM low glucose medium with 10 % FBS and the cell suspension was transferred into T-75 tissue culture flask.
Twice a week, cells were subcultivated 1:10 at 80-90 % confluence by incubation for 5 - 10 min with trypsin until cell detachment and dilution in DMEM low glucose medium supplemented with 10 % FBS and 1 % penicillin/streptomycin.

#### 4.3.17 *In vitro* DNA methylation and methylation sensitive Luciferase reportergene assay

Plamid DNA and double stranded oligonucleotides were *in vitro* methylated as indicated by manufacturer. Either 1 $\mu$g plasmid DNA or 100 $\mu$M oligonucleotide were incubated in a 20 $\mu$L reaction with 320 $\mu$M S-Adenosyl methionine (SAM), 2 $\mu$L NEB 2 buffer, 1 $\mu$L *SssI* methylase (4 U) and respective nuclease-free water at 37 $^{o}$C for 4 h. Afterwards, the methylase was heat-inactivated by incubation at 65 $^{o}$C for 20 min. Reactions were cleaned from salt and proteins by Millipore membrane filters.
HepG2 cells were plated in DMEM low glucose medium without antibiotics one day prior transfection at 200.000 cells/mL to reach a confluency between 50 and 80 % after 24 h. Lipofectamin and DNA solutions were prepared in serum-free Opti-MEM using 0.2 $\mu$L Lipofectamin 3000 per reaction (50 - 100 ng pCpGL-CMV-FLuc encoding plasmid containig CpG of interest co-transfected with respectively 5 - 10 ng pRL-SV40 or 500 ng pSV-$\beta$-Galactosidase control vector for transfection normalization). A final volume of 10 $\mu$L Lipofectamin-DNA mixture per 96-well was transfected as triplicates.
Cells were lysed in 50 $\mu$L 1X Passive Lysis Buffer (PLB, included in Dual Luciferase Reporter Assay System) and frozen for at least 30 min at -80 $^{o}$C. Lysates were stored at -20 $^{o}$C until luciferase activity measurement by Dual Luciferase Reporter Assay System. Therefore 20 $\mu$L of each lysate was used with 50 $\mu$L Luciferase Assay Reagent II (LAR II). The Fierfly Luciferase signal was measured for 12 sec and quenched for 3 sec by addition of 50 $\mu$L freshly prepared Stop&Glo solution if normalization was performed by Renilla luciferase signal. Renilla luciferase activity was measured subsequently to Firefly

signal quenching for 12 sec. Division of mean Firefly luciferase signal by mean Renilla luciferase signal was used for signal normalization.
If the CpG containing sequence of interest interfered with pSV40-RLuc activity due to promoter transactivation or the Firefly Luciferase signal was too low due to the sequence of interest, a $\beta$-galactosidase enzyme based normalization assay system was applied as indicates by manufacturer. Therefore 20 $\mu$L cell lysate was used and completed with 30 $\mu$L 1X PLB to a final volume of 50 $\mu$L. After 30 min incubation of the cell lysate with $\beta$-galactosidase substrate at 37 $^o$C, the color reaction was stopped by addition of 150 $\mu$L sodium carbonate and measured photometrically. A standard curve with fixed amounts of $\beta$-galactosidase enzyme served as reference. Division of mean Firefly luciferase signal by amount of $\beta$-galactosidase was used for statistical tests.

#### 4.3.18 Nuclear extract preparation and electrophoretic mobility shif assay (EMSA)

HepG2 cells of one T-75 tissue culture flask, which is equivalent to 50 $\mu$L packed cell volume, were used for extraction of proteins. Nuclear extract was generated by NE-PER Nuclear and Cytoplasmatic extraction reagents as indicated by manufacturer and each buffer was supplemented 1:1000 with proteinase inhibitor. The cytoplasmatic fraction was discarded. Nuclear extracts were stored at -80 $^o$C for three months or subsequently used for electrophoretic mobility shift assays (EMSA).
Protein concentrations were determined photometrically by BCA Protein Assay Kit. For Each EMSA binding reaction with nuclear extract, at least 6.5 $\mu$g nuclear protein was used.
Oligonucleotides were dimerized as described in section 4.3.13 and optional methylated or mock methylated as described in section 4.3.17.
Each 20 $\mu$L binding reaction contained 2 $\mu$L 10X binding buffer from the LightShift Chemiluminescence EMSA Kit (10 mM Tris, 50 mM KCl, 1 mM DTT), 1 ng/$\mu$L Poly dI dC, 10 $\mu$M $ZnSO_4$, 0.2 $\mu$g BSA, 50 fmol biotinylated double stranded oligonucleotide and respective nuclease-free water. Optional binding reactions contained 6.5 $\mu$g nuclear extract or unbiotinylated oligonucleotides (methylated or unmethylated) as serial dilution. For specific protein detection, nuclear extract was prior addition to binding reaction also incubated with a respective antibody on ice for 1 h.
Each binding reaction was incubated for at least 30 min on ice. 4 $\mu$L 6X loading dye was added prior gel loading. Electrophoresis was performed on a 6% DNA retardation gel in 0.5X TBE buffer at 80 - 100 V for 1.5 h (dependent of oligonucleotide size). Afterwards, oligonucleotides were transferred from the DNA retardation gel to a nylon membrane at 100 V for 1 h in 0.5X TBE. Since DNA is negatively charged due to its phosphate backbone, the nylon membrane has to be on the plus side of the transfer sandwich.
The membrane could be dried and stored at room temperature after transfer. Detection of biotinylated oligonucleotides was performed by the Chemilumineschence Nucleic Acid Detection Kit as indicated by manufacturer. Stabilized streptavidin conjugated horseradish peroxidase binds to the biotinylation of the oligonucleotides. After addition of its substrate peroxide with a luminol/enhancer solution, a light signal (luminescence) can be detected indicating the position on the membrane. Though this method is not as sensitive as radioactivity based shift assay, it is sufficient to detect protein biding, changes of protein binding due to antibody incubation and specificity due to competitive unlabeled oligonucleotide binding.

#### 4.3.19 Transfection of miRNA mimics and subsequent analysis

Cell culture experiments were performed to identify and validate possible targets of miRNAs. Therefore, precursor mimics were reverse transfected into HepG2 cells. Precursor mimics are double stranded oligonucleotides similar to endogenous pre-miRNAs. Chemical alterations ensure a correct processing and strand selection for assembly of the RNA-induced silencing complex (RISC).

Reverse transfection was performed by addition of a Lipofectamin 3000 and miRNA precursor mimic mixture subsequently to plating of cells. For each condition, two 6-wells with 2 mL HepG2 cell suspensions (number of cells equivalent to a confluence of 60 % 24 h after plating) were transfected with 200 $\mu$L of a Lipofectamin and mimic dilution in Opti-MEM.
Besides a precursor mimic, negative control transfected cells were used for normalization. Several conditions (24 h or 48 h treatment, as well 50 nM and 10 nM concentrations of mimics and negative controls #1 and #2) were tested. A final mimic concentration of 10 nM in 2.2 mL HepG2 low glucose cultivation medium and an incubation for 48 h showed optimal results. Negative control 1 (nc #1) transfection showed most stable and reproducible results.
Cells were once washed with ice-cold DPBS and lysed in 700 $\mu$L QIAzol. Whole cell RNA was extracted as described for liver tissue. Synthesis of miRNA-cDNA was performed by TaqMan advanced miRNA cDNA Sythesis kit as described for highly abundant hepatic miRNA using 10 ng whole cell RNA. For the synthesis of mRNA-cDNA, the High Capacity cDNA reverse transcriptase kit supplemented with RNase inhibitor was used with an input of 500 ng whole cell RNA. Reverse transcriptase master mix was prepared as indicated by manufacturer. The cDNA stock solution was diluted 1:10 for a final concentration of 2.5 ng/$\mu$L to use in qPCR.

#### 4.3.20 Insulin treatment of HepG2 cells

The treatment of HepG2 cells with high concentrations of insulin under hypergylcemia was used to generate an established hepatic-like *in vitro* cell model of insulin resistance (Huang et al., 2015; Teng et al., 2018).
Therefore HepG2 cells were plated at 200.000 cells/mL in a 6-well plate with HepG2 cultivation medium. After 24 h, medium was discarded and replaced with high glucose (4.5 g/L) medium supplemented with 0.5 % BSA and different concentrations (100 nM and 500 nM) of insulin. HepG2 cells in low glucose (1.5 g/L) medium supplemented with 0.5 % BSA served as control treatment.
Cells were harvested with a cell scraper in ice-cold DPBS 24 h after treatment. Cell pellets were generated by centrifugation for 10 min at 4 $^o$C and 12.700 rpm and stored at -80 $^o$C until further usage.
RNA was isolated as described for HepG2 cells. Synthesis of miRNA-cDNA and miRNA expression analysis was performed as described for mimic transfection by using the TaqMan advanced miRNA cDNA synthesis kit with an input of 10 ng whole cell RNA and TaqMan based qPCR.

#### 4.3.21 Basic statistics

Pearson correlation was performed for continuous variables and Spearman correlation for ordinal (gender, NAS) variables. These correlations indicated also whether a measurement was biased by age, gender or BMI and if these factors should be included in further regression models. Correlations were not corrected for multiple testing since each parameter was assessed by a different analysis system.
A relationship between DNA methylation and metabolic traits was calculated by linear regression models using optional age, gender and BMI as confounding factors (which is referred as adjusted or corrected p value). Unbiased DNA methylation was also analyzed non-parametrically by Mann-Whitney test.
For statistical gene expression analysis, a linear delta Ct value ($\Delta$Ct) was used by calculation of the difference between $Ct_{\text{target gene}}$ - $Ct_{\text{houskeeeping gene}}$. Linear regression models including further cofactors were calculated to estimate a correlation between gene expression and metabolic traits.
Since the cohort was neither matched for age nor gender, corrected p values including also BMI (referred as $p_c$) were always calculated to access an unbiased correlation between explanatory variables and response as described before (Pourhoseingholi et al., 2012).
DNA methylation was plotted as non-parametric box plots indicating besides a mean methylation value also the statistical distribution. A fold change difference was used by calculation of a respective $2^{-\Delta\Delta Ct}$

value and plotted for the comparison between ND and T2D subjects. These values are shown as mean fold change +/- SEM. For plotting of correlations between metabolic traits and gene expression, minumum/maxmimum normalized ΔCt values were plotted which does not change any statistical analysis.
Each hepatic expression analysis was considered independent from each other (difference in time point, assay and plate), therefore p values are not corrected for multiple testing. Results of expression analysis from mimic transfections or microarray analysis were corrected for multiple testing by controlling the false discovery rate (FDR).

#### 4.3.22 Allelic association

In the first attempt it was verified that the population is in Hardy-Weinberg equilibrium (alleles were not influenced by allelic drift or mutations within the cohort) by a chi squared test. Hardy-Weinberg equilibrium was assumed for a $\chi^2 > 3.841$ (one degree of freedom) with an $\alpha = 0.05$. For testing of association between an allele and incidence of T2D, also a chi squared test was performed as described before. Also logistic regression models were tested using the genotype, age, gender and BMI as cofactor to the state of T2D.

#### 4.3.23 Risk group classification by additive methylation risk score (addMRS)

A basic methylation risk score (MRS) was calculated by summing up standardized values (z) of DNA methylation x: $z = \frac{(x - \bar{x})}{\sigma(x)}$; $\text{MRS} = \Sigma(z_{cg}) = z_{cg06500161} + z_{cg11024682}$.
Student's t-test was performed to analyze whether the mean MRS values differ significantly between T2D and ND subjects. Effect sizes of MRS on other outcomes were calculated by linear regression, whereby a change of MRS about 1 score equals a change of DNA methylation at one CpG site about one standard deviation ($\sigma$) if DNA methylation at the other position stays constant. The Area under the curve (AUC) was calculated as goodness of the model. An AUC value above 0.5 indicates an improved classification in comparison to random stratification.

#### 4.3.24 Risk group classification by Random Decision Forest (RDF)

Random Decision Forest (RDF) was performed using the TreeBagger function in MATLAB version R2016a-R2018b. Due to the small sample size, each decision tree was generated by a subset of randomly picked data $\hat{D}$ (bootstrap approach) and tested against out-of-bag data of the whole ensemble D. Each random decision forest (RDF) model contains 200 randomly generated decision trees b, containing several combinations of genetic and epigenetic predictors (SNP genotype, DNA methylation at CpG sites or let-7e expression and serum concentration) and confounding factors age, gender and BMI, that were tested against other predictor sets. For each RDF model, the mean out-of-bag (OOB) error as goodness of the model was calculated. At least 100 models were generated and tested to estimate the limits of precision. No explicit tree depth was used.

#### 4.3.25 Microarray statistics

Generated CEL data was imported into Transcriptome Analysis Console (TAC) 4.0 and pre-processed for further analysis in MATLAB and R. Array data was normalized by Robust Multi-chip Analysis (RMA) algorithm as indicated by manufacturer, which contains background adjustment, log2 transformation and quantile normalization to increase the fold change ratios.
Linear and logistic regression models for continuous responses (metabolic traits like HbA1c, blood glucose and serum triglycerides) or the incidence of T2D were generated using MATLAB. Therefore age, gender, BMI and NAS were used as additional cofactors if not used as response variable. Only truly expressed mature miRNAs (> 50 % of all samples had a detection above background) were considered for statistical

analysis and generation of models. Since FDR adjustment of thereby generated p values resulted into no significantly associated target gene, another approach was used as described in the results section. In short, miRNAs with a detection threshold of log2 counts $> 2$ and an absolute log2 fold change difference between T2D and ND subjects of $> 1.5$ were necessary to be considered as possible candidates. Both, multiple associations to the incidence of T2D or other metabolic traits and missing associations to BMI and age were sufficient to be declared as candidate gene for qPCR validation.
Log2 fold change of gene expression between ND and T2D subjects was assessed after calculation of mean values by an adapted Tukey Biweight Algorithm (see section 10.2.1), which is more robust against outlier.

# 5 Results

The following contents of this thesis covers two different epigenetic mechanisms which might lead to altered hepatic gene expression in human metabolic diseases: Aberrant intragenic DNA methylation and changed miRNA expression.
In a first approach, hepatic expression of metabolic key genes involved in insulin signaling and fatty acid metabolism were identified. Besides insulin resistance and T2D, also alterations in intrahepatic thyroid hormone (TH) signaling was further characterized regarding the manifestation of NAFLD.
Intragenic changes in hepatic DNA methylation was exploratory analyzed for a potential novel T2D target gene which was previously associated to NAFLD in a different human liver cohort. Also recent blood based EWAS and GWAS hits were exploratory analyzed regarding alterations in hepatic intragenic and promoter methylation. In this context, also the suitability of large-scale blood based EWAS hits as tissue and population specific biomarker was critically examined.
Changes in miRNA expression were systematically assessed by microarray measurement and subsequent comprehensive target miRNA validation in the complete human liver cohort and respective cell culture experiments. Moreover, a cooperative mechanism for selective insulin resistance containing genotype-specific changes in DNA methylation and mRNA repression by miRNA overexpression was established and validated by different methods. Lastly, a gene-specific and a tissue-specific approach was used by careful selection of two candidate miRNAs and correlation analysis with metabolic traits or expression of potential target genes.

## 5.1 Hepatic gene expression

Initially, potential hepatic housekeeping genes for qPCR normalization were carefully evaluated to avoid a measurement bias. Afterwards, hepatic expression of metabolically relevant genes was assessed and correlated with blood-based parameters to find suitable candidate genes for subsequent analysis regarding aberrant epigenetic regulation in T2D. Also expression changes of genes contributing zo intrahepatic TH signaling was comprehensively analyzed which is based on the known contribution of hypothyroidism to NAFLD.

### 5.1.1 Hepatic housekeeping gene for qPCR normalization

Gene expression measurement by qPCR needs a stably expressed housekeeping gene for target normalization. Commonly used genes like *β-actin* (*ACTB*) or *glyceraldehyde-3-phosphate* (*GAPDH*) were shown to be influenced by steatosis and cirrhosis which is also observed in liver samples of this cohort (Boujedidi et al., 2012). Additionally, we are interested in a metabolic context, therefore genes directly involved in glucose metabolism (*Phosphoglycerate Kinase 1, PGK1*) need to be rejected. Other selection criteria include a low standard deviation between age, gender, disease state and a sufficient expression in liver.
31 potential housekeeping genes were tested (genes of TaqMan Array Human Controls, Figure 5.1) using 10 ng hepatic cDNA from three T2D subjects (two females, young and old, one male, intermediate age) and three ND subjects (two females, young and old, one male, young). Thermo Fisher cloud software, a Ct value cut-off of $> 30$ and NormFinder algorithm (Andersen et al., 2004) state *GAPDH*, *Ribosomal Protein L37a* (*RPL37A*) and *Cancer Susceptibility Candidate Gene 3 Protein* (*CASC3*) as most stably expressed genes. Additional TaqMan qPCR validation of all three candidate genes in the complete cohort specifies *CASC3* as most suitable hepatic housekeeping gene for subsequent gene expression normalization.

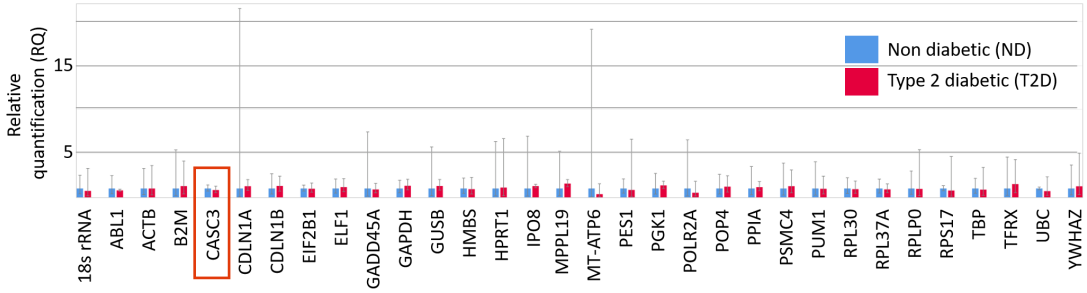


Figure 5.1: **Snaphsot from Thermo Fisher cloud software.** 31 potential housekeeping genes were tested with TaqMan Array Human Controls Plate, using cDNA from three T2D and three ND subjects. *CASC3* (red box) remains as most suitable housekeeping gene due to its stable expression within all conditions.

### 5.1.2 Hepatic gene expression for metabolically relevant genes

Epigenetic changes, which contribute to disease manifestation, result in aberrant gene expression. If this is affecting genes with key functions in glucose or fatty acid metabolism, these epigenetic changes modify the outcome of a certain metabolic trait (like blood glucose concentrations, serum triglyceride concentration, the glycaemic index of hemoglobin or steatosis and fibrosis state of the liver).
In a first attempt it is necessary to identify possible target genes which are dysregulated in T2D and at the same time correlate with a metabolic phenotype. Stratification of the cohort into subjects with T2D (diagnosed by clincial examination and taking medication against T2D) and ND subjects reveals a dysregultion of *Acetyl-CoA Carboxylase Beta* (*ACACB*), *Forkhead Box O1* (*FOXO1*), *LDL Receptor Related Protein 6* (*LRP6*) and *Stearoyl-CoA Desaturase* (*SCD*) in T2D (Figure 5.2A).
As indicated before, there is a significant difference in age and gender between T2D and ND subjects (Figure 4.1B and C). Logistic regression analysis using additional factors will indicate whether gene expression alone is associated to T2D or is influenced as well by age and gender due to the study design. Also a remaining effect of BMI on gene expression has to be considered though the cohort is matched for obese subjects in both subgroups. Expression of *ACACB*, *FOXO1* and *LRP6* is additionally influenced by age ($p_{age} = 0.0005$ for *ACACB*, $p_{age} = 0.0011$ for *FOXO1* and $p_{age} = 0.0142$ for *LRP6*). Furthermore, expression of *IRS2* is significantly influenced by BMI ($p_{BMI} = 0.0477$). After adjustment for respective factors, only *LRP6* remains significantly associated to T2D without bias. These results also indicate an interacting effect between the expression of *ACACB*, *FOXO1* and age as well between *IRS2* and BMI on T2D manifestation. *SCD* expression itself does not correlate with any confounding factor. Further adjustment for age, gender and BMI results into a not significant association of gene expression on the incidence of T2D.
Manifestation of T2D is a result of constant high blood glucose, unresponsiveness of target organs on insulin action and the inability of pancreatic beta cells to produce sufficient insulin to counteract the hyperglycaemic state. Before diabetes manifestation, insulin responsive organs like liver and skeletal muscle underlie an increasing unresponsiveness. One measurement for this insulin resistance is the HOMA-IR (Homeostatic Model Assessment for Insulin Resistance). This score is used as an estimate to describe the functionality of pancreatic insulin secretion by using fasting blood glucose and serum insulin values. Stratification by HOMA-IR (insulin resistance (IR) with HOMA-IR $> 2.5$) reveals further reduced expression of *ACACB*, *FOXO1*, *Insulin Receptor* (*INSR*) and *Cyclin Dependent Kinase Inhibitor 1B* (*CDKN1B*) in the state of insulin resistance (Figure 5.2B).
Besides *ACACB* and *FOXO1*, also *CDKN1B* ($p_{age} = 0.0235$) expression is influenced by age. *INSR* expression is independent of any confounding factor. Further adjustment for age, gender and BMI leads to a not significant association between *INSR* expression and T2D manifestation.
Expression of *Cyclin D1 (CCND1)*, *Cytochrome P450 Family 7 Subfamily A Member 1* (*CYP7A1*),

*GAPDH*, *Low Density Lipoprotein Receptor* (*LDLR*) or *Proprotein Convertase Subtilisin/Kexin Type 9* (*PCSK9*) is not significantly altered in both subgroup comparisons.
Further analysis reveal a largely negative correlation between *ACACB*, *LRP6* and *FOXO1* expression and metabolic traits or liver enzymes (Figure 5.2C). *GALNT18* expression correlates with age and BMI, but also with HbA1c and serum insulin. This gene was previously not significantly altered in T2D or IR. Therefore *GALNT18* is additionally considered for further epigenetic analysis.
There is a strong known association between the incidence of non-alcoholic fatty liver disease (NAFLD) and T2D (Portillo-Sanchez et al., 2015; Hu et al., 2017) . Therefore all potential candidate genes are also analyzed regarding a correlation between expression and the NAFLD activity score (NAS). *GAPDH* ($p_c = 0.0430$) and *SCD* ($p_c = 0.0087$) expression correlates positively and expression of *IRS2* ($p_c = 0.0136$) and *LRP6* ($p_c = 0.0075$) correlates negatively with NAS score. Correlation is independent of age, gender or BMI for those genes. Expression of *FOXO1* ($p = 0.0371$, $p_c = 0.2972$) and *ACACB* ($p = 0.0371$, $p_c = 0.2042$) correlates negatively with NAS but is additionally influenced by age as previously described for T2D.
Using T2D as additional co factor, expression of *IRS2* ($p_{NAS} = 0.0286$), *LRP6* ($p_{NAS} = 0.0283$) and *SCD* ($p_{NAS} = 0.0269$) remain significantly associated to solely NAS.
We focused on identification and characterization of possible epigenetic mechanisms which lead to the observed changes in hepatic gene expression of *ACACB*, *FOXO1*, *IRS2*, *LRP6*, *SCD*, *INSR* and *CDKN1B*. Moreover, associations of hepatic *GALNT18* to insulin and HbA1c levels emphasize a this gene as potential new target gene for T2D.

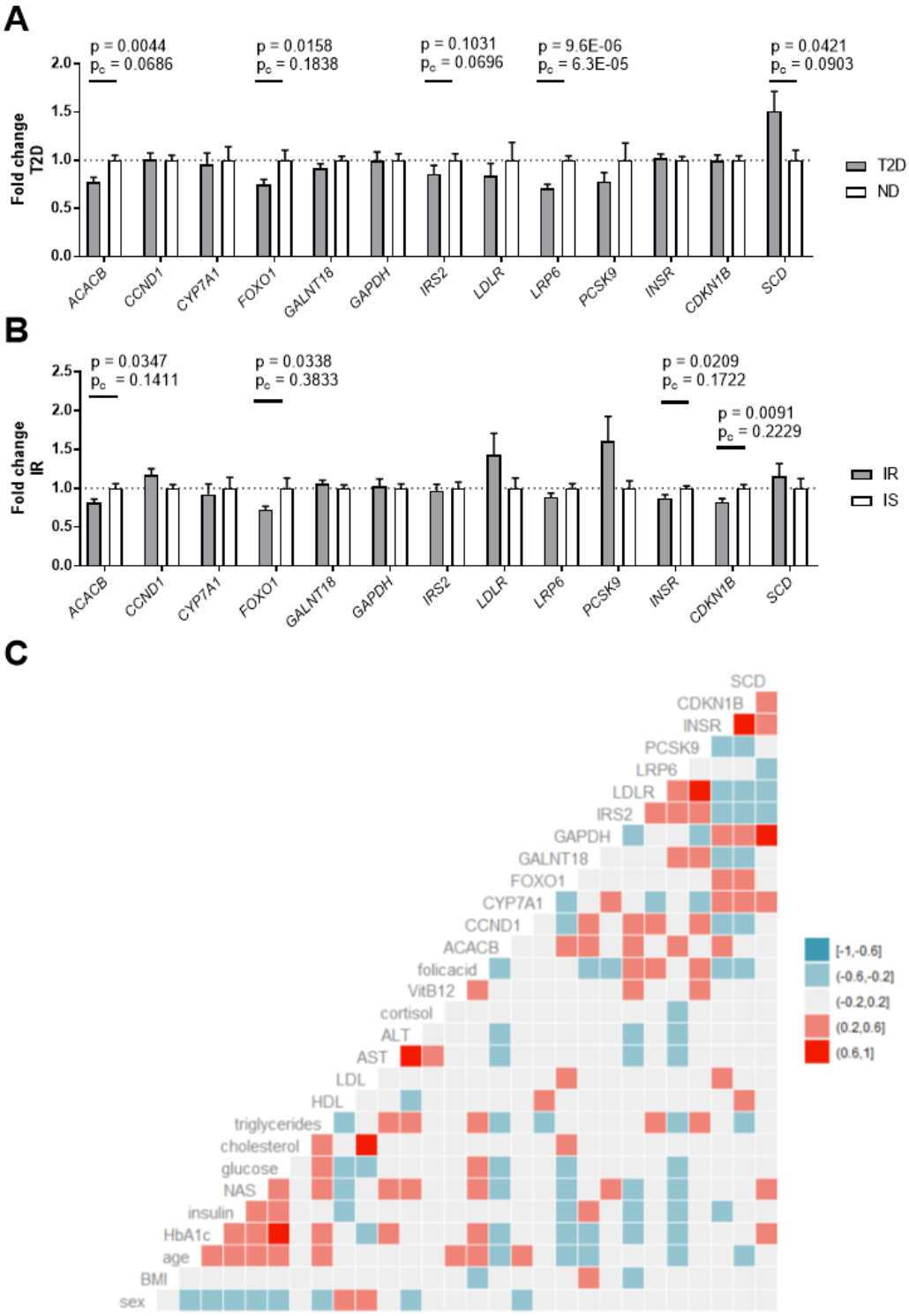


Figure 5.2: **Overview of hepatic gene expression changes in this cohort.** Stratification of the cohort into T2D und ND subjects results into significant repression of *ACACB*, *FOXO1* and *LRP6* and overexpression of *SCD* in T2D (A). Stratification into insulin resistant (IR) and sensitive subjects results into significant repression of *ACACB*, *FOXO1*, *INSR* and *CDKN1B* in IR (B). A correlation plot between gene expression $\Delta$Ct values, metabolic traits, enzymes and vitamins shows a strong interaction within genes as well between gene expression and traits (C). Blue indicates a negative, red a positive correlation. $p_c$ indicates a p value corrected for BMI, age and gender.

### 5.1.3 Hepatic gene expression for thyroid hormone signaling related genes

Besides the association between NAFLD and T2D, thyroid hormone (TH) signaling is known to contribute to the development and progress of NAFLD (Pagadala et al., 2012). The action of intracellular hepatic TH signaling was to date not studied in clinical samples, though compounds targeting Thyroid Hormone Receptor $\beta$ (THRB) were shown to have beneficial effects on cholesterol levels (Sinha et al., 2018). To elucidate whether the association of systemic hypothyroidism, hepatic TH signaling, T2D and fatty liver is mediated by respective hormone transporter or intracellular receptors, hepatic expression of different signaling related genes was measured.

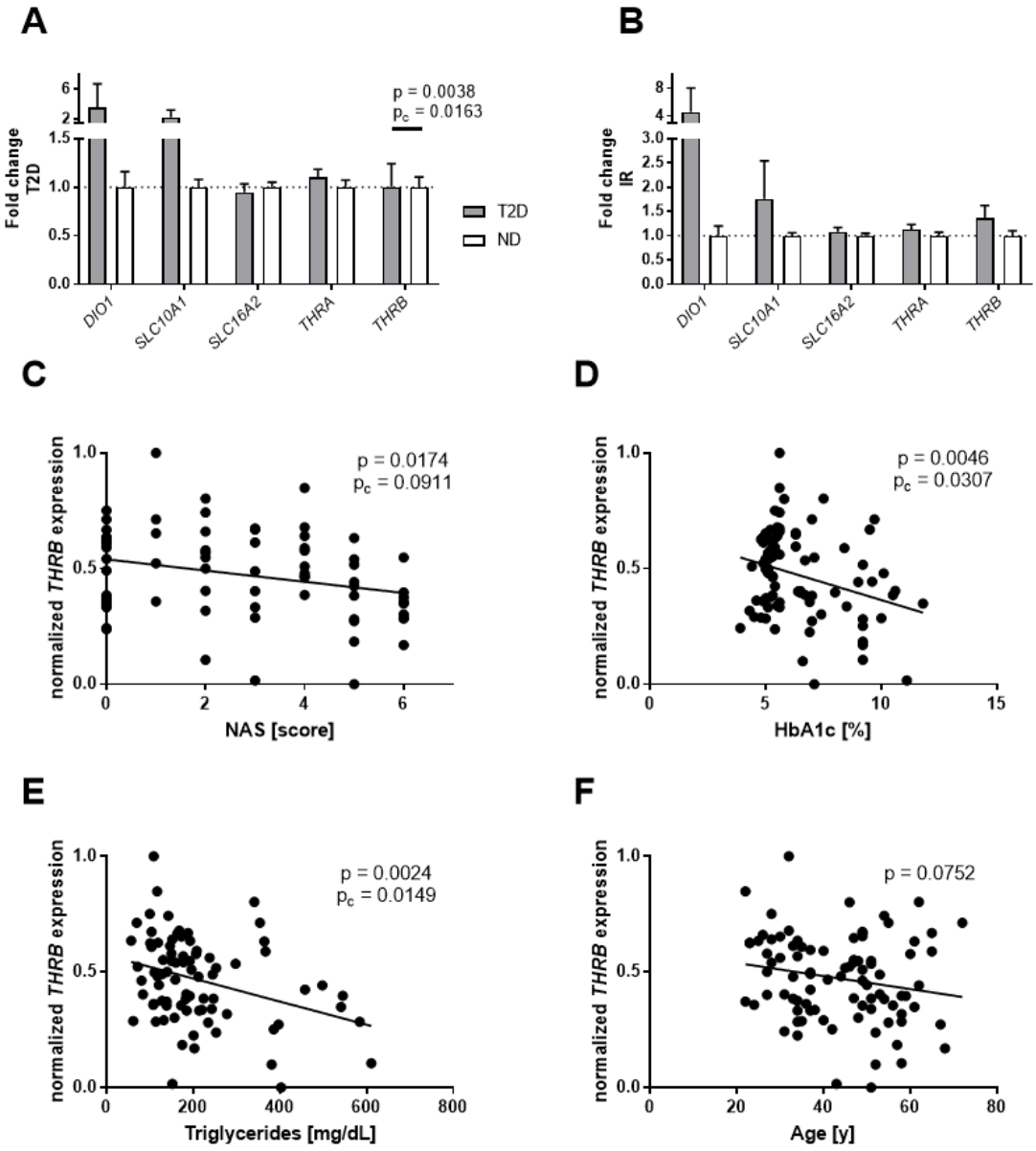


Figure 5.3: **Hepatic gene expression changes in this cohort for thyroid signaling related genes.** Stratification of the cohort into T2D und ND subjects results into a significant change in *THRB* expression with minor impact (A). Stratification into insulin resistant (IR, HOMA-IR > 2.5) and sensitive subjects (HOMA-IR < 2.5) results not into significant differences in TH related gene expression (B). *THRB* expression correlates uncorrected negatively with NAS score (C), HbA1c (D) and serum triglycerides (E). Previous showed correlation between expression and age was erased due to increase in sample size (F).

Gene expression measurement is performed for specific TH transport proteins MCT8 (Monocarboxylate Transporter 8, coded by *SLC16A2*), MCT10 (Monocarboxylate Transporter 10, coded by *SLC16A10*) and OATP1C1 (organic anion transporting polypeptide 1c1, coded by *SLCO1C1*), as well for Type I,II and III Iodothyronine Deiodinases (*DIO1, DIO2* and *DIO3*) and Thyroid Hormone Receptor $\alpha$ and $\beta$ (*THRA* and *THRB*).

The establishment of qPCR primers for SYBRgreen qPCR needs an estimation of amplification efficiency by generation and calculation of a standard cure based on a serial dilution of hepatic stock cDNA. Thereby it is not possible to generate an adequate standard curve for *DIO2*, *DIO3* and *SLCO1C1*. Ct values are permuted approximately at 35 which results into a more than 10 fold lower expression in comparison to other TH signaling related genes. Therefore these genes are regarded as not expressed and excluded for further analysis.

Expression of *DIO1* and *SLC10A1* is not significantly ($p = 0.1375$ and $p_c = 0.1559$ for *DIO1*, $p = 0.6544$ and $p_c = 0.5418$ for *SLC10A1*) increased in T2D and IR (Figure 5.3A and B). Also a correlation analysis between *DIO1* and HbA1c ($p = 0.0579$) or *SLC10A1* and HbA1c ($p = 0.6733$) as measurement of long-term markers for T2D revealed no association. Expression of both genes is not affected by confounding factors. Further regression analysis reveals that *SLC16A2* and *THRA* expression is not altered in any condition. Interestingly, only *THRB* expression is significantly altered between T2D and ND subjects though the fold change difference is extremely small between T2D and ND subjects (about 0.99, Figure 5.3A). No TH signaling related gene is significantly altered after HOMA-IR classification of the cohort into insulin resistant and sensitive subjects (Figure 5.3B).

Additionally, only *THRB* expression decreased with severity of NAFLD (Figure 5.3C), elevation of HbA1c level (Figure 5.3D) and elevation of serum triglycerides (Figure 5.3E). A previously published relationship between age and reduced *THRB* expression is diminished due to a slightly increased sample size (from $n = 82$ to $n = 89$ subjects, Figure 5.3F). BMI (Table 20) and gender ($p = 0.5141$) are also not correlating with *THRB* expression, therefore an uncorrected p value can be applied. Elevated blood glucose and liver enzymes AST and ALT correlate with reduced *THRB* expression besides the already mentioned association to NAS, HbA1c and serum triglycerides (effect sizes see Table 20). All in all, *THRB* expression remains the most interesting target gene for further investigation of hepatic lipid and glucose metabolism by TH.

Table 20: **Linear regression models of hepatic *THRB* expression.** Bold values are considered significant.

| ***THRB*** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | **0.01452** | -45.44 | 0.10419 | -27.50 | 84 |
| Insulin | 0.57820 | -25.31 | 0.97840 | -1.07 | 82 |
| HOMA-IR | 0.38887 | -3.58 | 0.59204 | -2.05 | 77 |
| HbA1c | **0.00463** | -1.48 | **0.03070** | -1.02 | 89 |
| BMI | 0.23448 | -3.71 | | | 89 |
| Age | 0.07524 | -6.43 | | | 89 |
| Gender | 0.51413 | 0.06 | | | 89 |
| Cholesterol | 0.30239 | -12.81 | 0.47128 | -8.97 | 82 |
| Triglycerides | **0.00244** | -110.33 | **0.01490** | -84.77 | 82 |
| HDL | **0.00333** | 10.82 | **0.00663** | 10.18 | 82 |
| LDL | 0.42812 | -8.87 | 0.35077 | -10.62 | 76 |
| AST | **0.00408** | -22.11 | **0.00824** | -20.99 | 84 |
| ALT | **0.00379** | -17.78 | **0.00285** | -17.98 | 84 |
| TSH | 0.70901 | 0.59 | 0.43624 | 1.27 | 79 |

## 5.2 Hepatic DNA methylation

Though *Polypeptide N-Acetylgalactosaminyltransferase 18* (*GALNT18*) expression is not significantly altered between T2D and ND subjects, we observe a significant association between expression, HbA1c and serum insulin levels. Intragenic *GALNT18* methylation and expression was previously shown to be altered in obese patients with NAFLD (Ahrens et al., 2013). Its function in metabolism remains unclear, therefore we aimed to evaluate whether it is a suitable candidate gene for the manifestation of T2D and whether aberrant DNA methylation is a plausible regulatory mechanism.
*Tumor Protein P53 Inducible Nuclear Protein 1* (*TP53INP1*) was chosen as a candidate gene to examine DNA methylation within promoter regions in diabetic liver. A previously performed meta analysis revealed, that a T2D associated polymorphism is influencing *TP53INP1* promoter methylation and thereby affecting gene expression (Xue et al., 2018). We aimed to examine whether this might also affect liver tissue.

### 5.2.1 Intragenic *GALNT18* methylation

The CpG site cg16337763 within intron 3 of *GALNT18* was reported to show an decreased methylation in NAFLD (Ahrens et al., 2013). It is hardly possible to design suitable bisulfite PCR primers for replication of this position, therefore we decided to measure reliable methylation of three 506 bp to 689 bp upstream CpGs instead.
DNA methylation is best characterized for promoter regions and first exons due to its significance and popularity in cancer research. Nevertheless, intragenic methylation, also known as gene body methylation, was shown to follow a non-linear effect on gene expression (Jjingo et al., 2012). Some genes showing hyper or hypo intragenic methylation were associated to extreme high or low mRNA expression. Thereby intragenic promoters are not sufficient to explain this causality. Other approaches indicate a functionality in splice site variation by DNA methylation and subsequent accessibility of recognition motifs. We assume a similar functionality of intragenic DNA methylation affecting *GALNT18* expression in T2D and fatty liver disease.

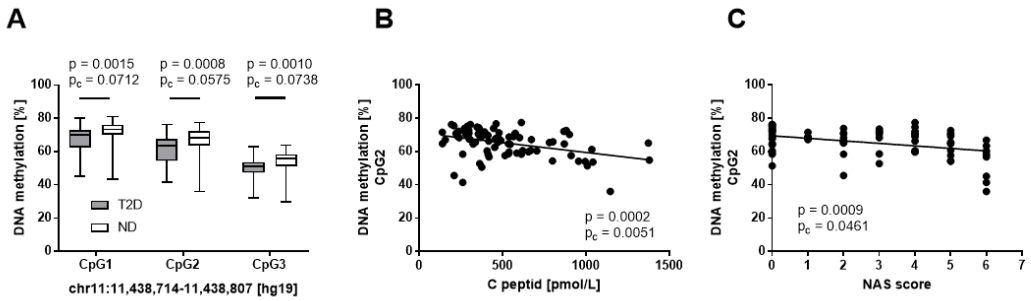


Figure 5.4: **DNA methylation of GALNT18.** DNA methylation within intron 3 of *GALNT18* is significantly altered between T2D and ND subjects, which becomes insignificant after correcting for age, gender and BMI (A). DNA methylation at CpG2 correlates negatively with C peptide (B) and NAS score (C).

DNA methylation at CpG1 to CpG3 is decreased in subjects with T2D (CpG1 5.21 %, CpG2 5.58 % and CpG3 4.28 % absolute difference, Figure 5.4A). Regression analysis revealed a strong association of DNA methylation and aging (Table 21 for CpG2 and Table 28 in supplement), therefore after correction for these confounding factors, the difference in methylation became insignificant. Nevertheless, DNA methylation at all positions correlated significantly with serum C peptide concentrations (Figure 5.4B) and NAS (Figure 5.4C) after ajustment for age, gender and BMI. Only CpG2 is shown as a representative due to the largest difference in DNA methylation.

In previous studies, *GALNT18* was reported to be up-regulated in NAFLD. In this cohort there is no significant change in expression (Figure 5.5A.) Moreover, *GALNT18* expression is influenced by BMI (Figure 5.5B) and age (Figure 5.5C). Additional regression analysis reveals a positive correlation between gene expression and serum insulin concentration (Figure 5.5D) which also results into a significant relationship between enhanced gene expression and insulin resistance indicated by a positive correlation regarding HOMA-IR (Figure 5.5G). Unfortunately, gene expression is neither affected by HbA1c (Figure 5.5E) nor NAS (Figure 5.5F).

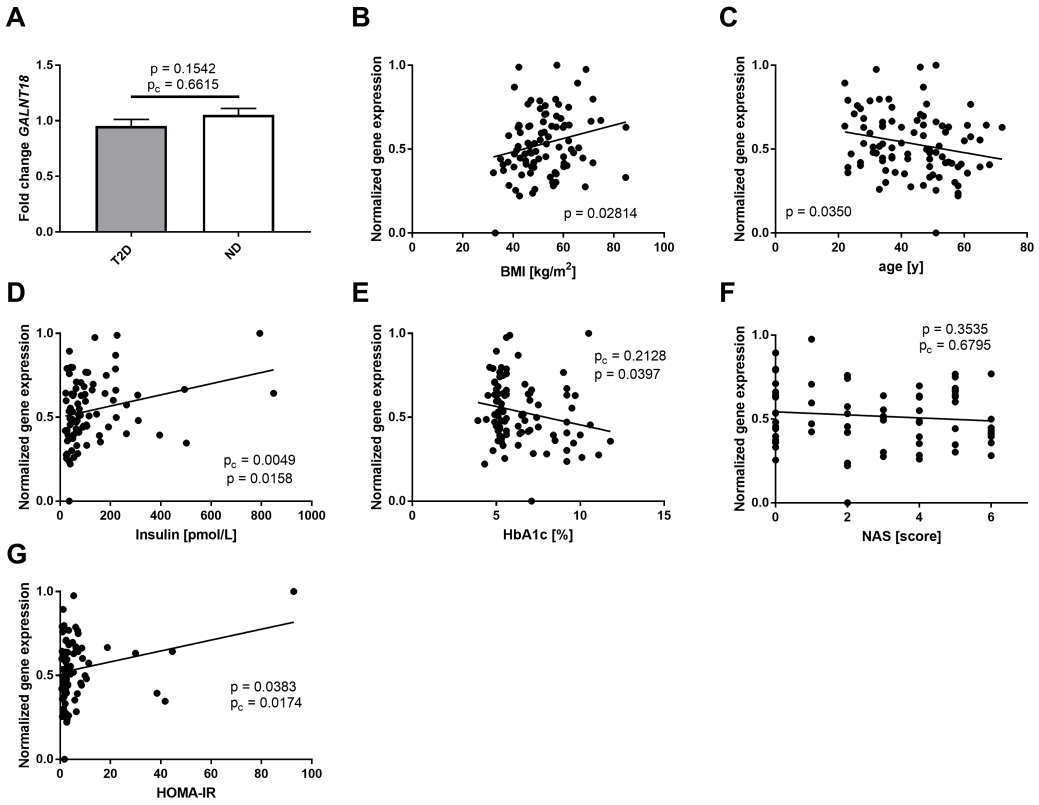


Figure 5.5: **Expression of *GALNT18*.** Gene expression of *GALNT18* is not significantly altered between T2D and ND subjects (A). Expression correlated with BMI (B), age (C) and serum insulin (D). Expression correlated without correction significantly with HbA1c (E). Correlation with NAS score was not reproducible in this cohort (F). Gene expression is enhanced in insulin resistance, indicated by a positive correlation between expression and HOMA-IR without influence of by any confounding factor (G).

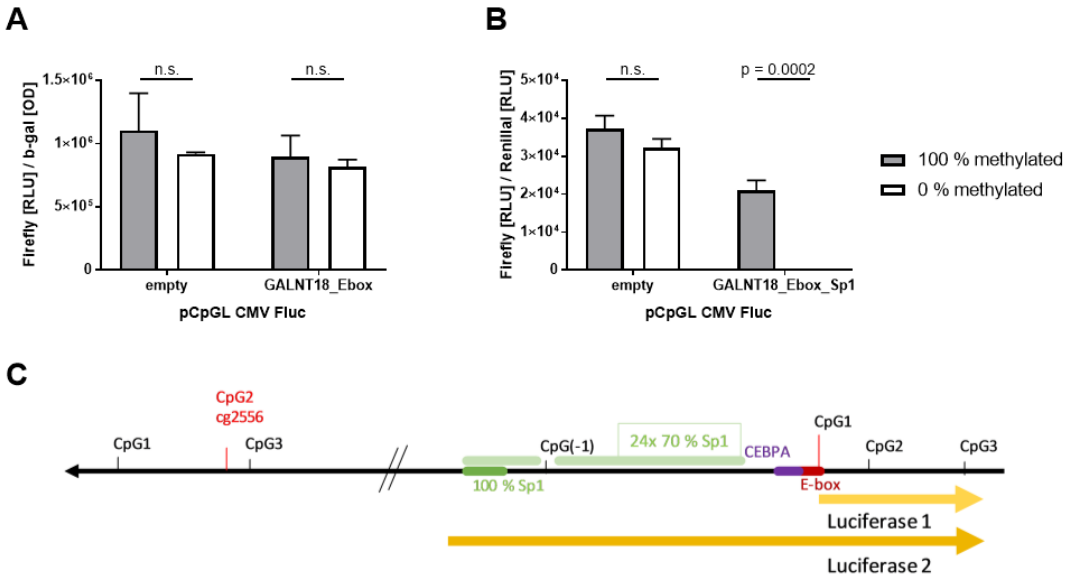


Figure 5.6: **Functional analysis of DNA methylation in intron 3 of *GALNT18*.** Luciferase assay 1 shows no effect by DNA methylation (A). Luciferase assay 2 contains additional Sp1 binding motifs and appears to be drastically influenced by DNA methylation (B). Overview of intron 3 sequences and transcription factor binding motifs included for both Luciferase assays (C).

Since both DNA methylation and gene expression is influenced by aging and DNA methylation correlates with liver health status, we performed a Luciferase reporter gene assay on these positions to analyze whether methylation in intron 3 might have an actual impact on gene expression. We used a plasmid enconding a Firefly Luciferase gene under control of a CMV promoter without CpG sites and integrated our CpG sites of interest downstream of the promoter. Cells were lysed 24 h after transfection and Luciferase signal was measured. The plasmid including solely CpG1 to CpG3 shows no change in Luciferase signal in dependency of methylation (Figure 5.6A). Including a forth CpG site downstream of previous measured CpG oligonucleotides, the effect changes drastically into complete erasement of the Luciferase signal in unmethylated (0 % methylation) state (Figure 5.6B). *In silco* analysis of the nucleotide sequence reveals a putative E box motif for CpG1 and an enrichment for possible Sp1 binding motifs surrounding the forth CpG (namend CpG(-1), Figure 5.6C).
In summary, altered *GALNT18* expression in NAFLD can not be reproduced in our liver cohort. Moreover, aging is strongly influencing both *GALNT18* methylation and gene expression (see additionally Table 28 in supplement and Table 21). DNA methylation at CpG2 is significantly affected by NAFLD and overall liver health status without influence of confounding factors which is indicated by correlations between methylation and NAS score, serum AST and serum ALT levels (Table 21). Gene expression is elevated in insulin resistance without being affected by any confounding factor.

Table 21: **Linear regression models of *GALNT18* methylation at CpG2 and gene expression.** CpG2 was previously shown to have the largest difference in percent DNA methylation. Bold values are considered significant.

| CpG2 | p | estimate | p corrected | estimate | n |
|---|---|---|---|---|---|
| Glucose | 0.07449 | -1.46 | 0.92807 | 0.07 | 82 |
| Insulin | 0.12359 | -3.33 | 0.76653 | 0.58 | 80 |
| HOMA-IR | 0.65830 | -0.09 | 0.13165 | 0.28 | 75 |
| C peptide | **0.00023** | -13.38 | **0.00506** | -10.47 | 82 |
| HbA1c | **0.00207** | -0.07 | 0.14617 | -0.03 | 87 |
| BMI | 0.27048 | -0.16 | | | 87 |

| | | | | | |
|---|---|---|---|---|---|
| Age | **0.00135** | -0.50 | | | 87 |
| Gender | **0.01881** | 0.01 | | | 87 |
| cholesterol | 0.42732 | 0.42 | 0.39564 | 0.49 | 80 |
| triglycerides | **0.00895** | -4.24 | 0.20457 | -2.09 | 80 |
| HDL | **0.01690** | 0.40 | 0.12586 | 0.28 | 80 |
| LDL | 0.07050 | 0.86 | 0.32093 | 0.51 | 74 |
| AST | **0.00035** | -1.21 | **0.00082** | -1.26 | 82 |
| ALT | **0.00000** | -1.22 | **0.00002** | -1.21 | 82 |
| **dCT** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
| Glucose | 0.15346 | -25.08 | 0.62315 | -8.41 | 87 |
| Insulin | **0.01576** | 93.63 | **0.00495** | 97.01 | 85 |
| HOMA-IR | **0.03829** | 7.59 | **0.01736** | 8.69 | 80 |
| C peptide | 0.15282 | 102.73 | 0.36317 | 63.24 | 87 |
| HbA1c | **0.03969** | -0.97 | 0.21280 | -0.53 | 92 |
| BMI | **0.02814** | 6.05 | | | 92 |
| Age | **0.03496** | -6.80 | | | 92 |
| Gender | 0.78024 | 0.03 | | | 92 |
| cholesterol | 0.82587 | -2.51 | 0.50248 | 8.16 | 85 |
| triglycerides | 0.44892 | -25.77 | 0.75485 | 10.70 | 85 |
| HDL | 0.70036 | -1.37 | 0.90801 | -0.44 | 85 |
| LDL | 0.62689 | -4.95 | 0.63981 | -5.15 | 79 |
| AST | 0.38321 | -6.27 | 0.77142 | -2.30 | 87 |
| ALT | 0.43947 | -4.43 | 0.73585 | -2.03 | 87 |

### 5.2.2 Promoter *TP53INP1* methylation

To check whether CGI shore methylation is altered in T2D manifestation, *TP53INP1* promoter methylation was analyzed neighboring two promising target CpG sites cg13393036 and cg20039814. (Kirchner et al., 2016) indicated an up-regulation of *TP53INP1* expression in ND and T2D subjects in comparison to lean subjects. Thereby two CpG sites, cg20039814 and cg18059933, were most affected by changes in DNA methylation. Both sites are located in *TP53INP1* CGI shore within 82 bp. Moreover, methylation at cg13393036 is reported to be significantly influenced by polymorphism which is associated to the incidence of T2D (Xue et al., 2018).

Gene expression regulation by promoter methylation of CGIs is a broadly described principle. It occurs mostly during cell development and CGI hyper and hypo methylation is a feature of most cancer types. In normal healthy cells, CGI are unmethylated resulting in active gene expression. Surrounding regions of CGIs, called CGI shore (up to 2000 bp flanking a CGI from both sides) and shelves (beyond 2000 bp flanking regions) are described to maintain a stable DNA methylation pattern themselves (Edgar et al., 2014).

Bisulfite pyrosequencing of cg13393036 and cg20039814 (CpG7 and CpG8, Figure 5.7) reveals a low DNA methylation of *TP53INP1* CGI shore with no significant difference between T2D and ND subjects. Therefore aberrant promoter methylation was excluded as possible mechanism of hepatic *TP53INP1* expression regulation in this cohort.

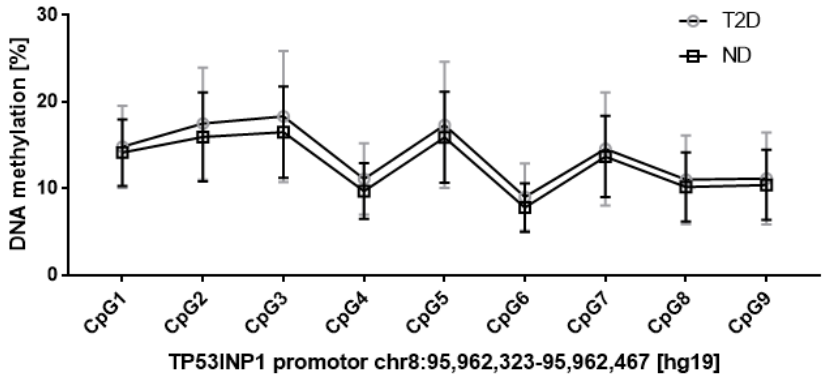


Figure 5.7: **DNA methylation for promoter of *TP53INP1*.** DNA methylation is not altered between T2D and ND subjects within promoter regions of candidate genes. CpG7 is known as cg13393036, CpG8 is known as cg20039814.

## 5.3 DNA methylation and genetics - Example for EWAS validation

High throughput methods for assessment genetic changes, which are associated to a specific phenotype, are a common tool for large scale studies, which started first 2002 for myocardial infarction (Ozaki et al., 2002). These genome-wide association studies (GWASs) are able to detect millions of single nucleotide polymorphisms (SNPs) at once for one individual by microarray measurement. Recent GWAS were able to detect 120 SNPs in association to T2D, still failing to explain most of the heritability of T2D (Prasad and Groop, 2015).

Later on, high throughput measurement of the epigenome acquired greater importance, especially in cancer research. These epigenome wide association studies (EWASs) are able to measure DNA methylation of thousands of CpG oligonucleotides at once. Life style factors are mirrored by changes in DNA methylation and appear to have a larger influence on T2D manifestation and also heritability. Therefore EWAS are also considered for T2D, obesity and respective metabolic traits. Methylation changes of two recurring sites which pass the criteria of association (respective p values) are *Sterol Regulatory Element Binding Transcription Factor 1* (*SREBF1*) cg11024682 and *ATP Binding Cassette Subfamily G Member 1* (*ABCG1*) cg06500161. Due to its easy accessibility, methylation at both positions is mostly measured in leukocytes DNA. These methylation patterns are tissue specific, therefore we are interested whether these markers have a functional background in metabolically relevant tissue and also if they pass the filter of association if the number of subjects is below large scale population based studies.

### 5.3.1 *SREBF1* cg11024682 methylation

The designed assay for measurement of cg11024682 contains three additional CpG sites. Nevertheless, hepatic methylation is not altered between T2D and ND subjects at any position (differences of 0.14 % CpG1, 0.94 % CpG2, 1.38 % CpG3 and 0.01 %, Figure 5.8A). Surprisingly, methylation at cg11024682 correlates weakly with BMI (Figure 5.8B) after correction for age and gender although the cohort itself consists only of obese subjects.

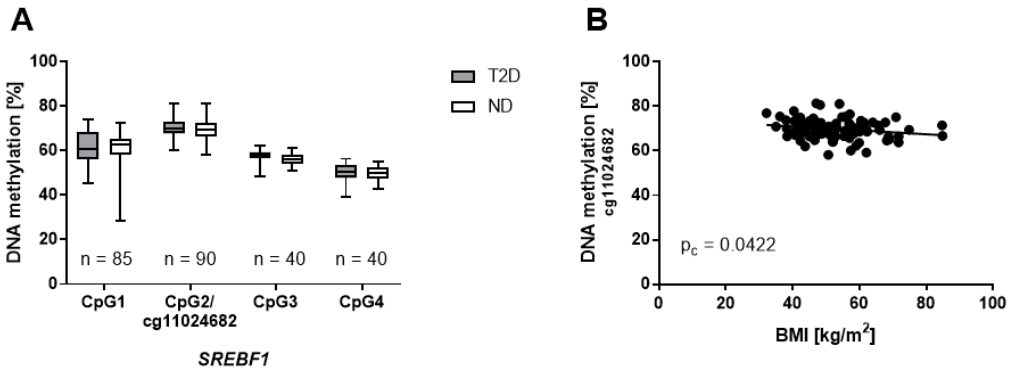


Figure 5.8: **Methylation of *SREBF1* EWAS marker cg11024682.** DNA methylation between EWAS marker cg11024682 and surrounding CpG oligonucleotides is not altered between T2D and ND subjects (A). A weak correlation between BMI and marker methylation is reproducible for this cohort (B).

### 5.3.2 *ABCG1* cg06500161 methylation interacting with rs9982016

Methylation of cg06500161 is inconclusively reported to be affected by a nearby SNP rs9982016 (A>T MAF = 6.63 % on 1000 Genomes), therefore it is necessary to regard the genotype of each subject for a careful evaluation (Hidalgo et al., 2014; Kulkarni et al., 2015; Wahl et al., 2016).
Hepatic methylation at cg06500161 was not altered between T2D and ND subjects (difference of 2.36 %, Figure 5.9A). The used bisulfite pyrosequencing assay is unable to distinguish between CpG methylation and polymorphism for CpG1, therefore only individuals with an A/A genotype are considered for DNA methylation measurement at CpG1, which shows a significant hypermethylation for T2D subjects (4.02 %, Figure 5.9A).
This cohort shows neither a significant impact of rs9982016 on methylation at cg06500161 (Figure 5.9B) nor any influence on disease state. This is further confirmed by a chi squared test on the incidence of T2D (one degree of freedom, $\chi^2 < 3.841$ with $\alpha = 0.05$) after accounting for Hardy-Weinberg equilibrium (polymorphism stays constant across populations). Differences within a genotype range from 2.5 % (A/A) to 2.9 % (A/T) and within the disease state from 5 % (T2D) to 4.7 % (ND). The minor allele T has a more frequent abundance in subjects with T2D (MAF 6.92 %, Figure 5.9C).

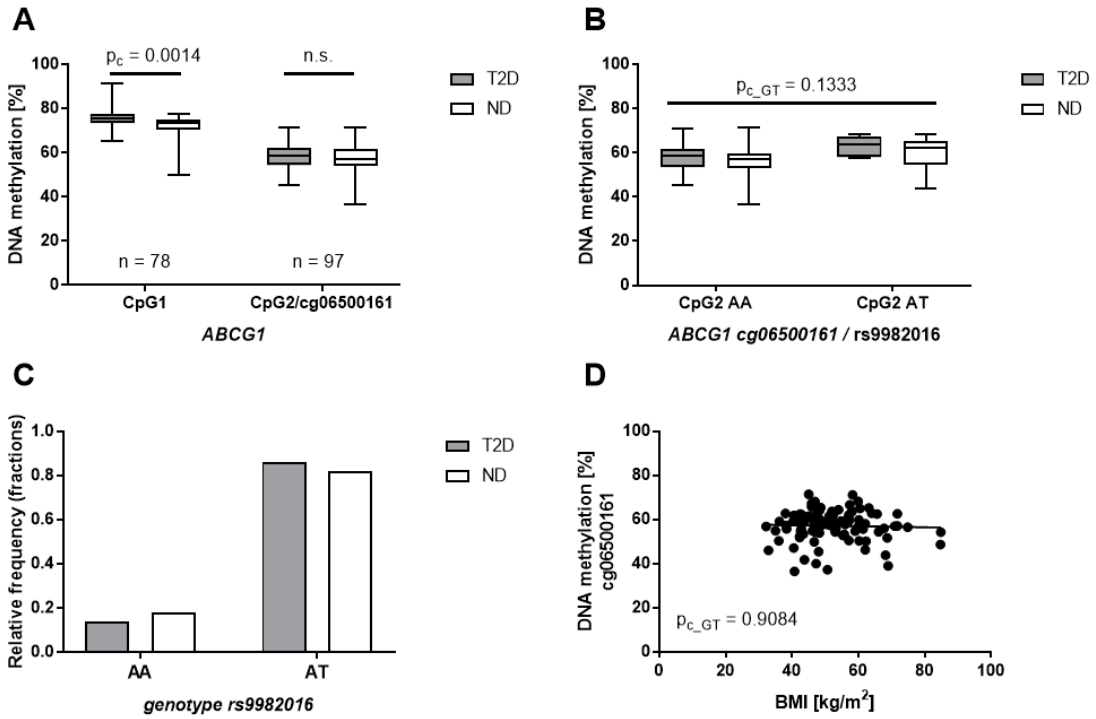


Figure 5.9: **Methylation of *ABCG1* EWAS marker cg06500161.** DNA methylation differs significantly for CpG1 (only n = 16 subjects with A/A genotype at the polymorphism rs9982016) but not the marker cg06500161 itself (CpG2, A). Hepatic DNA methylation of CpG2 is not influenced by the adjacent polymorphism rs9982016 (B). The minor allele T is more prominent in T2D subjects which represents the proposed risk allele (C). Marker methylation does not correlate with BMI (D).

Linear regression analysis of marker methylation with BMI reveals constant methylation independent from age, gender or genotype of the adjacent SNP (Figure 5.9D). Therefore a correlation between methylation at *ABCG1* cg06500161 and BMI is not reproducible in liver regardless of any confounding factor.

### 5.3.3 Clincial relevance: EWAS marker as biomarker

Both markers originate from different blood-based EWAS and were reproduced as lead candidate markers for obesity (correlating with BMI) and incidence of T2D in a large meta analysis based on leukocyte DNA. As indicated before, hepatic methylation is not altered between T2D and ND subjects and only hepatic *SREBF1* methylation correlates weakly with BMI.

T2D is a complex disease, thus a combination of both markers might have the power to stratify T2D and ND subjects. Calculation of an additive score based on methylation (simply called additive methylation risk score or addMRS) results not into a significant difference of mean values (Figure 5.10A). Nevertheless, a Receiver Operating Characteristics (ROC) curve analysis for the addMRS reveals an Area Under Curve (AUC) value of 0.5784 which holds stratification by addMRS results into a better outcome than coin flip (Figure 5.10B).

Lastly, correlation between addMRS and BMI was not significant, regardless of correction for additional confounder age and gender (Figure 5.10C).

In summary, blood-based EWAS marker might not have a functional background in metabolically relevant tissue and therefore challenge the assumption, that blood is a suitable surrogate tissue for large EWAS on T2D. The easy accessibility of blood in comparison to human liver biopsies remains a striking argument but emphasizes the urgent need for marker validation in metabolically relevant tissue like liver, adipose tissue or pancreas.

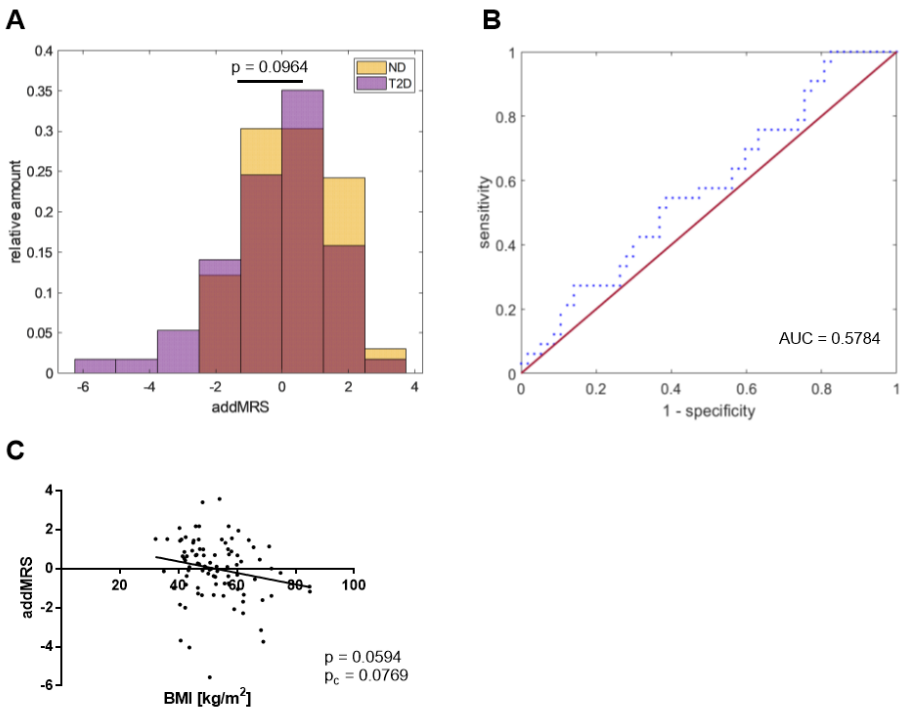


Figure 5.10: **Characterization of an additive methlyation risk score (addMRS) using EWAS markers.** The calculated addMRS does not significantly stratify both cohorts into T2D and ND subjects (A). Nevertheless, stratification improves classification (AUC $> 0.5$, B). The addMRS does not correlate significantly with BMI (C).

## 5.4 miRNA expression

Epigenetic regulation might also affect protein biosynthesis by degradation of mRNA or steric hindrance of the ribosome. This translational repression is mediated by small 21-22 nt, oligonucleotides called micro RNAs (miRNA) which are coded intra- or intergenic in the DNA.

The first miRNAs were identified 1993 in *Caenorhabditis elegans* (*C.elegans*) controlling cell development (Lee et al., 1993; Rougvie, 2001). Since then, contribution to diverse cellular processes especially in disease manifestation were discovered. These discoveries are particularly interesting because miRNAs are able to regulate a broad field of target genes at the same time mirroring the complexity of disease manifestation. Most dysregulated miRNAs emerge from mice and small human studies in the context of NAFLD, non alcoholic steatohepatitis (NASH) or alcoholic steatohepatitis (ASH). We were interested whether aberrant miRNA expression correlating with different metabolic traits (blood glucose, serum triglycerides, serum insulin, HbA1c, serum adiponectin) contributes to the manifestation of T2D. Therefore non-coding (nc)RNA microarray analysis was performed in 40 (20 T2D, 20 ND) subjects, matched for confounders age, gender and BMI. Candidate miRNAs were later measured by qPCR in a cohort of 95 subjects for further replication and validation.

Besides mature miRNAs, also pre-miRNAs (first maturation step of pri-miRNAs), small nucleolar (sno)RNAs (CDBox and H/ACA Box) and small Cajal body-specific (sca)RNAs can be quantified by microarray. Especially snoRNAs are known to have a function in metabolic stress (Michel et al., 2011). These ncRNAs do not target specific genes, thus they are neglected for the moment.

### 5.4.1 Regression analysis for T2D associated miRNAs

Signal intensities which represent respective miRNA expression was assessed by microarray measurement. No array failed quality and labeling control, thus all 40 samples were included for further data generation.

Logistic regression analysis using the incidence of T2D as outcome, adjusted further for age, gender, BMI and NAS score as confounding factors, revealed 37 possible mature miRNAs which are associated to T2D ($p < 0.05$, Figure 5.11A). Hierarchical cluster analysis using euclidean distances is not differentiating both subgroups perfectly into T2D (1) and ND (0) subjects (heatmap classification, see Figure 5.11A). A principle component analysis (PCA) to reduce the dimensions of multi-dimensional expression data reveals that based on all 37 mature miRNA, approximately 49.4 % of the variance between both subgroups is explained (Figure 5.11B). Further filtering by setting the log2 count threshold to 2 results into 28 mature miRNAs (Figure 5.11C, Table 34 in supplement) and approximately 52.9 % explained variance between the subgroups. Both PCA failed to cluster T2D and ND subjects into two distinct groups which indicates a high variability between the individuals within one subgroup.

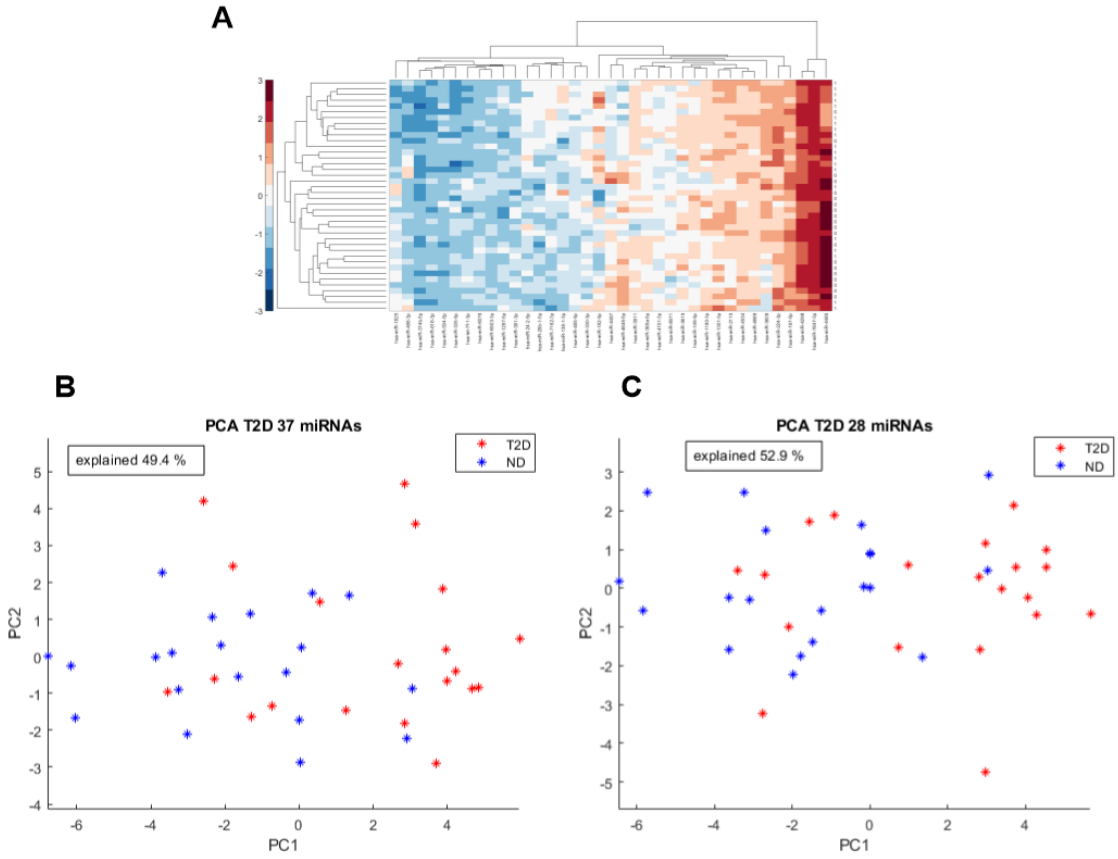


Figure 5.11: **Overview after logistic regression of microarray log2 counts with disease status.** Heatmap of associated ($p < 0.05$) miRNAs und respective hierarchical clustering using euclidean distances of subjects with T2D (1) and ND (0,A). Principle component analysis of miRNAs which are associated ($p < 0.05$) to T2D (B) and which pass the filter of log2 count $> 2$ (C). Plotted were scores for principal component (PC) 1 and PC2 which represent the largest (PC1) and second largest (PC2) variance in the data with respective percentage explained variance.

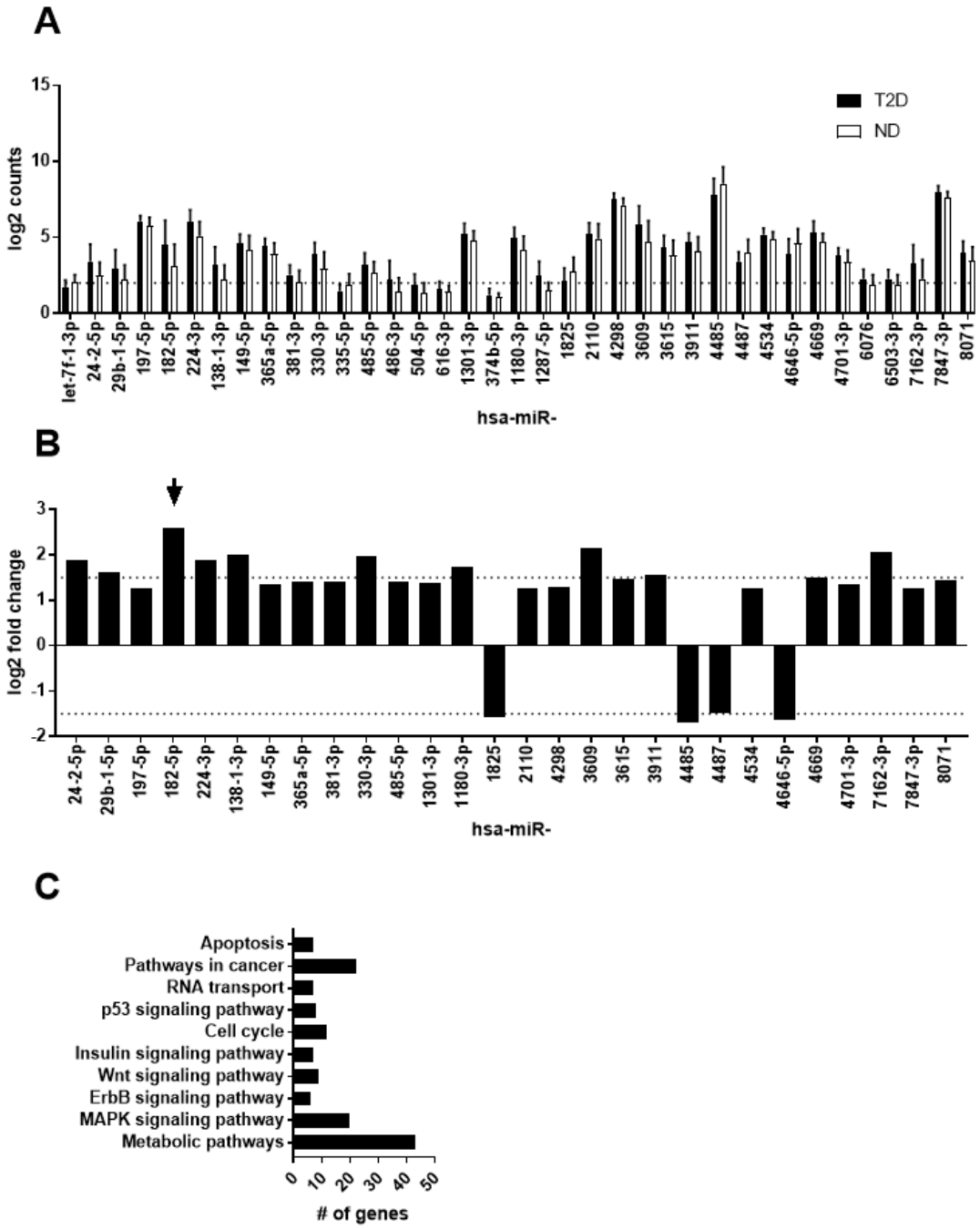


Figure 5.12: **Detailed results of logistic regression of miRNA log2 counts from microarray and disease state.** Most associated signals originate from minor expressed hepatic miRNAs (A). Most miRNAs are up-regulated in T2D. An absolute fold change cut-off (dashed lines) of 1.5 was applicated (B). Hsa-miR-182-5p was chosen as candidate miRNA for qPCR validation (arrow, B). Number of predicted target genes of hsa-miR-182-5p for serveral pathways (C). Predicted target genes were taken from the NetAffx Analysis Center and respective miRNA target gene annotations from the databases microcosm (also known as miRBase Targets) (Griffiths-Jones et al., 2007), MTI (Guo et al., 2015) and miRecords (Xiao et al., 2009). Pathway enrichment was calculated by R using the package KEGGprofile (Shilin Zhao, 2017).

Since the p values from the logistic regression analysis failed to be controlled by FDR, a more strict filter was applied regarding an absolute fold change expression difference (FC) of $> 1.5$ besides a minimum log2 count of 2. Generally more low expressed miRNAs appear to be altered in T2D (Figure 5.12A). Only 13 miRNAs have an absolute FC of $> 1.5$, but hsa-miR-182-5p (arrow, Figure 5.12B) is the most interesting due to a drastic increase in subjects with T2D (FC of 2.6). Pathway analysis of possible target genes of this miRNA by R using KEGGprofile (Shilin Zhao, 2017) reveals some of those genes participate directly in metabolism and insulin signaling, but also in MAPK and p53 signaling which is responsive to insulin and regulates hepatic insulin sensitivity (Figure 5.12C) (Lawan and Bennett, 2017; Geng et al., 2018). Statistical analysis reveals only candidate target genes involved in MAPK signaling pathway ($p = 0.0036$, $p_{adj} < 0.1$), cell cycle ($p = 0.0029$, $p_{adj} < 0.1$) and p53 signaling pathway ($p = 0.0025$, $p_{adj} < 0.1$) are significantly enriched.

### 5.4.2 Regression analysis for miRNAs associated with metabolic traits

To test whether some changes in miRNA expression are not directly associated to T2D but to related metabolic traits, regression analysis was performed for blood glucose, serum triglycerides, serum insulin, HbA1c, HOMA-IR and serum adiponectin. These models included besides log2 values of the miRNA microarray also NAS, age, gender and BMI as confounding factors in response to a specific trait. Additionally, miRNA associated to NAS, age and BMI were considered. A complete list of all miRNA expression associations can be found in supplemental Table 34. Afterwards, the mean change in expression between T2D and ND subjects as FC was calculated. P values of regression models are not controlled by FDR, therefore again a signal threshold of log2 $> 2$ and an absolute FC threshold of $> 1.5$ was applied for further refinement (see Figure 5.13).
All in all, 15 miRNAs remained associated to blood glucose levels (Figure 5.13A), nine to serum triglycerides (Figure 5.13B), three to age (Figure 5.13C), 13 to HbA1c (Figure 5.13D), eight to serum insulin (Figure 5.13E), four to BMI (Figure 5.13F), eight to HOMA-IR (Figure 5.13G), one to serum adiponectin (Figure 5.13H) and two to NAS (Figure 5.13I).
Changes in expression of several identified miRNAs are not only associated to a single metabolic trait (Figure 5.13J). A more accurate division reveals that most miRNAs which were previously shown to be altered in T2D do not correlate with other traits (Figure 5.14). Since there was no direct correction for multiple testing, some false positive hits are expected. Another criteria, multiple associations to metabolic traits, reveals seven unique candidate miRNAs (Table 22) whose expression is altered in T2D and at least one metabolic trait (Figure 5.14A and B) and not associated to age or BMI (Figure 5.14C). Finally hsa-miR-182-5p proved to be a suitable candidate miRNA for qPCR validation due to its drastic increase in T2D in contrast to ND subjects and its association to blood glucose and Hb1Ac levels.

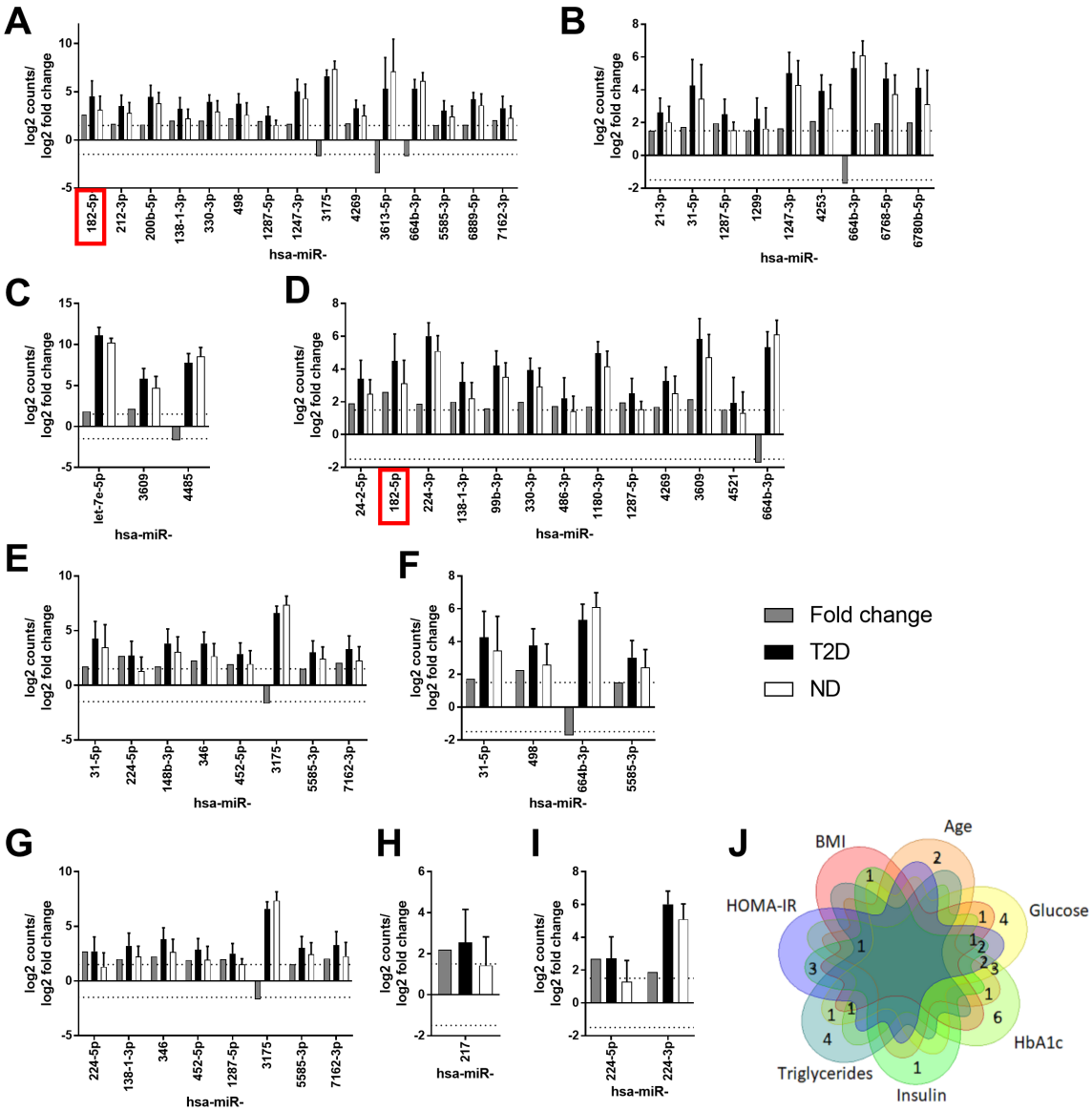


Figure 5.13: **miRNAs associated to different metabolic traits.** Linear regression shows miRNAs associated to blood glucose (A), serum triglycerides (B), age (C), HbA1c (D), serum insulin (E), BMI (F), HOMA-IR (G), serum adiponectin (H) and NAS score (I), which have a log2 count of $> 2$. Some miRNA overlap between several traits (J). Graphs include raw log2 counts for both groups as well a fold change difference between T2D and ND subjects. The dashed line indicates an absolute fold change difference of 1.5.

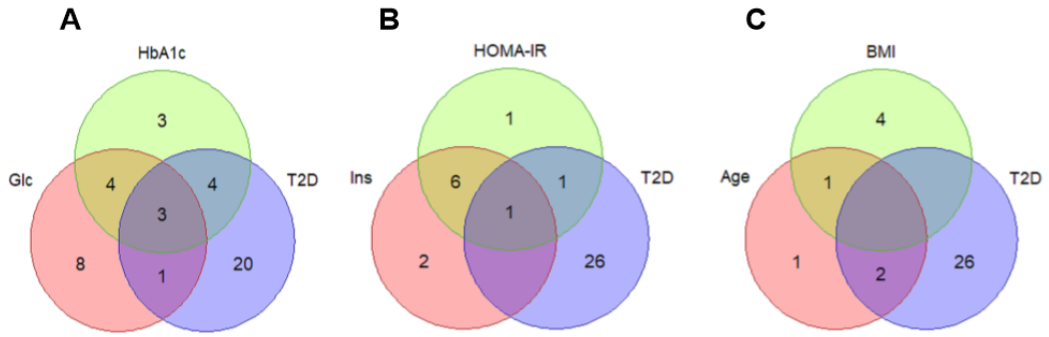


Figure 5.14: **Detailed overview of overlapping miRNA associations between traits.** Glucose related traits (A), insulin related traits (B) and confounder (C) show overlapping miRNAs.

Table 22: **List of overlapping miRNAs from Figure 5.14**

| miRNA | Venn | Trait 1 | Trait 2 | Trait 3 |
|---|---|---|---|---|
| miR-182-5p | A | Glc | HbA1c | T2D |
| miR-330-3p | A | Glc | HbA1c | T2D |
| miR-138-1-3p | A | Glc | HbA1c | T2D |
| miR-7162-3p | A | Glc | | T2D |
| miR-24-2-5p | A | | HbA1c | T2D |
| miR-224-3p | A | | HbA1c | T2D |
| miR-1180-3p | A | | HbA1c | T2D |
| miR-3609 | A | | HbA1c | T2D |
| miR-7162-3p | B | Ins | HOMA-IR | T2D |
| miR-138-1-3p | B | | HOMA-IR | T2D |
| miR-3609 | C | Age | | T2D |
| miR-4485 | C | Age | | T2D |

### 5.4.3 Hepatic qPCR housekeeping miRNA

Literature research was discordantly regarding a suitable hepatic housekeeping miRNA. Most commonly used miRNA species showed at least in serum or adipose tissue a hint towards metabolic dysregulation. Since miRNAs might enable tissue crosstalk, these metabolic effects might also change hepatic miRNA expression. Literature proposed 23 possible housekeeping miRNA genes (Table 23). Three candidates (hsa-miR-126-3p, hsa-miR-24-5p and hsa-miR-361-5p) were chosen for validation by qPCR based on scattering between microarray subgroups and metabolic association (Figure 5.15). All three candidates are measurable in the whole cohort (n = 95) by TaqMan advanced qPCR. Thermo Fisher cloud candidate endogenous control function and NormFinder algorithm states hsa-miR-24-3p as most suitable housekeeping gene. Though the expression of almost all potential housekeeping miRNA genes show either some association to a hepatic cancer phenotype (HCC, fibrosis) or lipid metabolism (NAFLD), all but one miRNA expression did not show any association to any metabolic trait or the incidence of T2D in this analyzed cohort. Hsa-miR-331-3p was significantly ($p = 0.0256$) associated to BMI but was excluded by not reaching the FC threshold (FC 1.39, see Table 35 for BMI associated entries). Nevertheless, a log2 value of 5 to 6 indicates low hepatic expression.

Table 23: **Proposed housekeeping miRNAs from literature and their metabolic association.** Key: T.. true, F.. false, is expr. .. is expressed

| Transcript ID | Mean log2 | is expr. | FC | p | Association (Literature) |
|---|---|---|---|---|---|
| hsa-miR-939-5p | 5.76 | T | 1.17 | 0.04 | HBV (Huang et al., 2016) |
| hsa-miR-181b-5p | 8.66 | T | 1.19 | 0.43 | Liver fibrosis (Kitano and Bloomston, 2016) |
| hsa-miR-199a-3p | 10.03 | T | 1.12 | 0.38 | Fibrosis, HCV (Hayes and Chayama, 2016) |
| hsa-miR-199b-3p | 10.03 | T | 1.12 | 0.55 | HCV (Hayes and Chayama, 2016) |
| hsa-miR-103a-3p | 12.8 | T | 1.2 | 0.41 | NAFLD upregulated (Soronen et al., 2016) |
| hsa-miR-24-3p | 13.14 | T | 1.2 | 0.42 | NAFLD (Ng et al., 2014) |
| hsa-miR-99a-5p | 11.15 | T | -1.17 | 0.05 | NAFLD upregulated (Tan et al., 2014) |
| hsa-miR-191-5p | 12.16 | T | 1.06 | 0.55 | Methylation-sensitive regulation in HCC (Elyakim et al., 2010) |
| hsa-miR-186-5p | 0.87 | F | | | NAFLD downregulated (Leti et al., 2015) |
| hsa-miR-15b-5p | 8.14 | T | 1.12 | 0.44 | HCC upregulated (Yang et al., 2016) |
| hsa-miR-451a | 6.83 | T | 1.07 | 0.63 | NAFLD downregulated (Hur et al., 2015) |
| hsa-miR-361-5p | 10.14 | T | 1.14 | 0.89 | NAFLD upregulated (Zhang et al., 2018b) |
| hsa-miR-423-5p | 8.26 | T | 1.03 | 0.92 | Upregulated in serum of obese men (Ortega et al., 2015) |
| hsa-miR-320a | 12.2 | T | 1.09 | 0.94 | Dysregulated in serum of bypass patients, HCC (Lu et al., 2017; Lirun et al., 2015) |
| hsa-miR-101-5p | 3.04 | T | 1.06 | 0.87 | Affected by weight loss in adipose tissue (Kuryłowicz et al., 2017) |
| hsa-miR-192-5p | 12.05 | T | -1.02 | 0.38 | NAFLD upregulated (Tan et al., 2014) |
| hsa-miR-26b-5p | 4.34 | T | 1.1 | 0.83 | HCC (Ji et al., 2009) |
| hsa-miR-331-3p | 5.44 | T | 1.4 | 0.33 | Serum, NAFLD (Zarrinpar et al., 2016) |
| hsa-miR-484 | 5.25 | T | -1.08 | 0.72 | T2D pancreas downregulated (Williams and Mitchell, 2012) |
| hsa-miR-345-5p | 7.47 | T | 1.05 | 0.87 | HCC (Hayes and Chayama, 2016) |
| hsa-miR-126-3p | 11.51 | T | 1.01 | 0.78 | HCC (Ghosh et al., 2016) |
| hsa-let-7g-5p | 8.8 | T | 1.14 | 0.67 | HCC (Gori et al., 2014) |
| hsa-miR-26a-5p | 13.41 | T | 1.11 | 0.94 | hepatic glucose and lipid metabolism (Fu et al., 2015) |

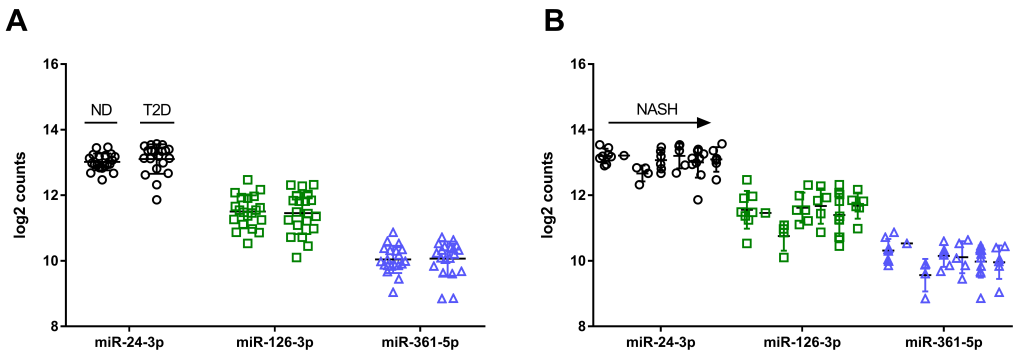


Figure 5.15: **Distribution of log2 counts for all three candidate miRNA housekeeping genes.** Hsa-hiRNA-24-5p shows the smallest scattering among all subgroups after stratification by T2D (A) or NAS score (B).

### 5.4.4 miRNA qPCR validation for hsa-miR-182-5p

Hepatic miRNAs were extracted and reverse transcribed into miRNA-cDNA from 90 subjects. These samples were used for the final validation of the candidate miRNA hsa-miR-182-5p identified in the microarray discovery cohort of 40 subjects.

Gene expression measurement by qPCR reveals a significant up-regulation (FC $>$ 1.5) of hepatic hsa-miR-182-5p in subjects with T2D which is also independent of NAS (Figure 5.16A). Additionally, hepatic miRNA expression correlates positively with HbA1c (Figure 5.16B) and NAS (Figure 5.16C). For other blood parameters, miRNA expression correlates with serum insulin without further adjustment for confounding factors (Figure 5.16D), but with serum triglycerides (Figure 5.16E) and blood glucose after correction (Figure 5.16F).

Additional regression analysis reveals an association between miRNA expression and age, HDL, AST and ALT but not with BMI (see Table 24). Therefore expression is independent from obesity but relies additionally on age. The interaction between aberrant hsa-miR-182-5p expression and metabolism (changes in blood glucose, HbA1c, serum triglycerides, HDL, AST and ALT) are not age driven since they remain significant ($p < 0.05$) after adjustment for age, BMI and gender. The interaction between miRNA expression and serum insulin is influenced by aging.

The microarray results and its calculated associations to T2D, blood glucose and HbA1c are reproducible for hsa-miR-182-5p. There are additional correlations between miRNA expression and lipid traits (NAS, serum triglycerides, HDL) as well between expression and liver enzymes (AST and ALT). Due to the multifaceted associations, hsa-miR-182-5p proved to be a metabolically involved miRNA and emerged as potential key regulatory molecule for T2D and fatty liver.

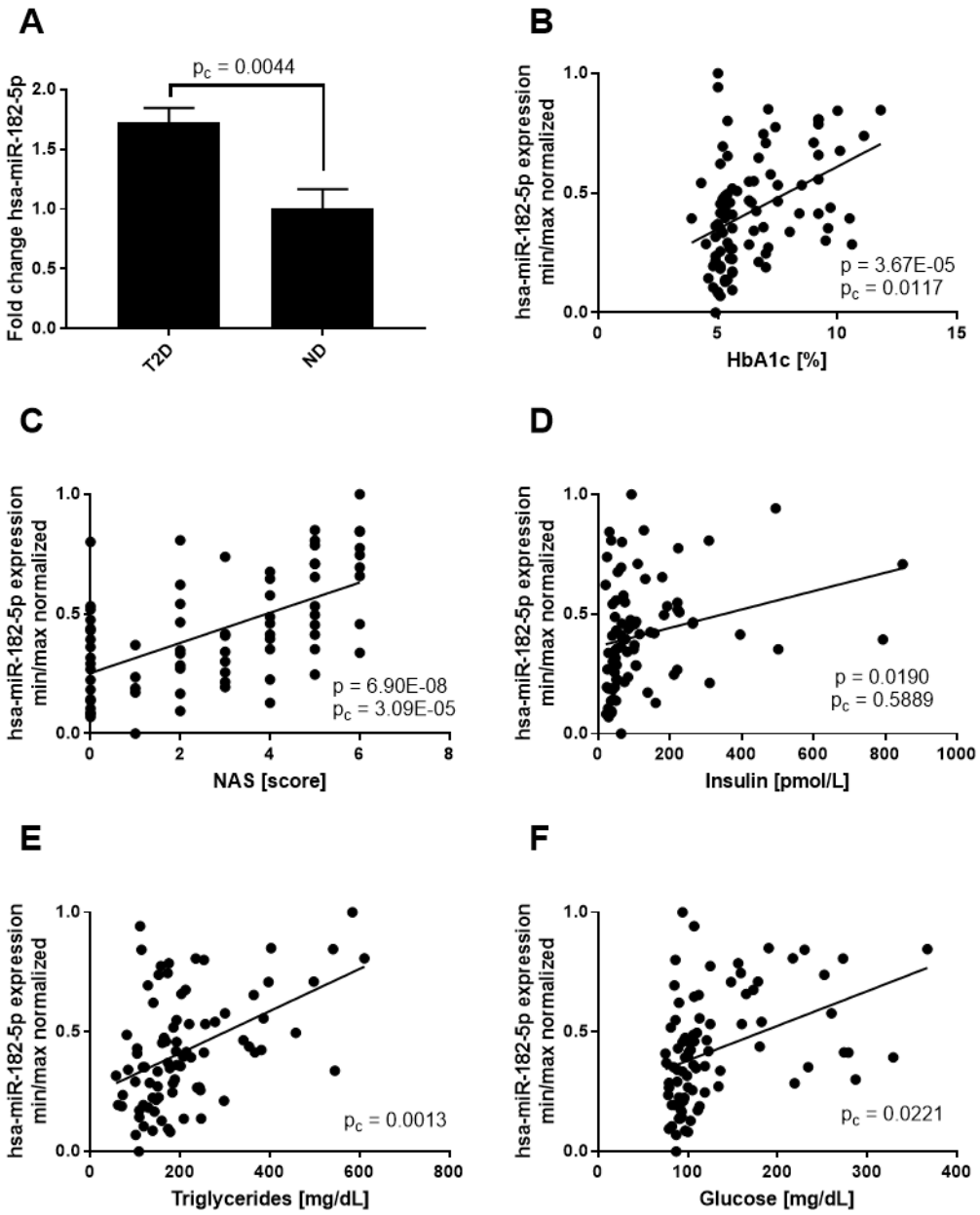


Figure 5.16: **Validation of hsa-miR-182-5p by qPCR.** Significant up-regulation of hsa-miR-182-5p in subjects with T2D which stays significant after additional correction for NAS ($p_{\text{NAS corrected}} = 0.0280$, A). Expression correlates positively with HbA1c (B) and NAS (C). Expession correlates with insulin, which becomes insignificant after correction for age, gender and BMI (D). Expression correlated with serum triglycerides (E) and glucose (F). P values ($p_c$) are corrected for age, gender and BMI.

Table 24: **Linear regression models of hepatic hsa-miR-182-5p.** Bold values are considered significant.

| **hsa-miR-182-5p** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | **0.00015** | 20.61 | **0.02212** | 12.44 | 85 |
| Insulin | **0.01901** | 32.10 | 0.58891 | 6.93 | 83 |
| HOMA-IR | 0.12557 | 1.87 | 0.82811 | -0.26 | 78 |
| HbA1c | **0.00004** | 0.63 | **0.01168** | 0.38 | 90 |
| BMI | 0.74026 | 0.31 | | | 90 |

| | | | | | |
|---|---|---|---|---|---|
| Age | **0.00036** | 3.84 | | | 90 |
| Gender | 0.23445 | -0.37 | | | 90 |
| Cholesterol | 0.89624 | -0.48 | 0.80759 | -0.97 | 83 |
| Triglycerides | **0.00001** | 45.56 | **0.00135** | 34.47 | 83 |
| HDL | **0.00063** | -3.77 | **0.00846** | -3.16 | 83 |
| LDL | **0.04436** | -6.56 | 0.14719 | -5.16 | 77 |
| AST | **0.00039** | 7.90 | **0.00136** | 7.95 | 85 |
| ALT | **0.00013** | 6.71 | **0.00149** | 6.00 | 85 |

### 5.4.5 Transfection of hsa-miR-182-5p mimic for target gene validation

As described earlier, miRNA might bind to the 3'-UTR of several distinct genes and therefore repress different targets at the same time.

Potential metabolic target genes of hsa-miR-182-5p, *ELOVL6*, *FASN*, *FOXO1*, *LRP6*, *SCD*, and *CDKN1B* were identified by data base research for experimentally validated miRNA-mRNA interactions or seed sequence prediction tools (Karagkouni et al., 2018; Chou et al., 2018). Three genes (*SCD*, *ELOVL6* and *FASN*) participate in hepatic de novo lipogenesis and fatty acid metabolism which is reported to be elevated in NAFLD. Moreover, FOXO1 is a transcription factor which activates the expression of key enzymes of gluconeogenesis. *FOXO1* expression was also previously described in this thesis as altered in T2D. A repression of these genes has a favorable course on T2D manifestation.

In contrary, LRP6 is known to play a crucial role in whole body lipid homeostasis by regulating LDLR-mediated hepatic lipid uptake and to activate the Wnt/beta-catenin signaling pathway. A malfunction of pathway regulation causes enhanced hepatic lipid accumulation by increased de novo lipogenesis and triglyceride synthesis.

Lastly, a link between CDKN1B and hepatic insulin resistance is not reported, though it represents a potential link between T2D, liver cirrhosis and cancer due to its contribution to cell cycle regulation.

Additional screening of the 3'UTR of these six potential target genes was performed for the validation of the presence of the respective seed sequence and for assessment of the possible binding mode (Table 25).

Table 25: **3'-UTR screening of possible hsa-miR-182-5p target genes.** Positions are relative to start of 3'-UTR and calculated by a self-written java based unix shell tool (see 10.3.3).

| **Target** | **Transcript ID Ensembl** | **8mer** | **7mer_m8** | **7mer_A1** | **6mer** |
|---|---|---|---|---|---|
| *ELOVL6* | ELOVL6-201 | at 656 bp | false | at 657 bp | at 657 bp |
| *FASN* | FASN-201 | false | false | false | at 365 bp |
| *FOXO1* | FOXO1-201 | false | false | at 264 bp | at 264 bp |
| | | | | at 2139 bp | at 2139 bp |
| *LRP6* | LRP6-201 | false | 1787 bp | false | at 1788 bp |
| | | | | | at 2331 bp |
| *SCD* | SCD-201 | false | false | false | at 1185 bp |
| *CDKN1B* | CDKN1B-201 | false | false | false | at 1290 bp |

Transfection of 10 nM pri-miR-182 mimic into HepG2 cells results into an approximately 450-fold increase in mature miR-182-5p after 48 h incubation (Figure 5.17A). Subsequent expression measurement of possible target genes by qPCR results into a significant decrease of *FOXO1* (p = 0.0022, FDR q = 0.0131), *LRP6* (p = 0.0241, FDR q = 0.0482), *SCD* (p = 0.0068, FDR q = 0.0204) and *CDKN1B* (p = 0.0022, FDR q = 0.0071) expression in mimic transfected HepG2 cells (Figure 5.17B). A control transfection with scrambled oligonucleotides (negative control, nc#1) serves as reference.

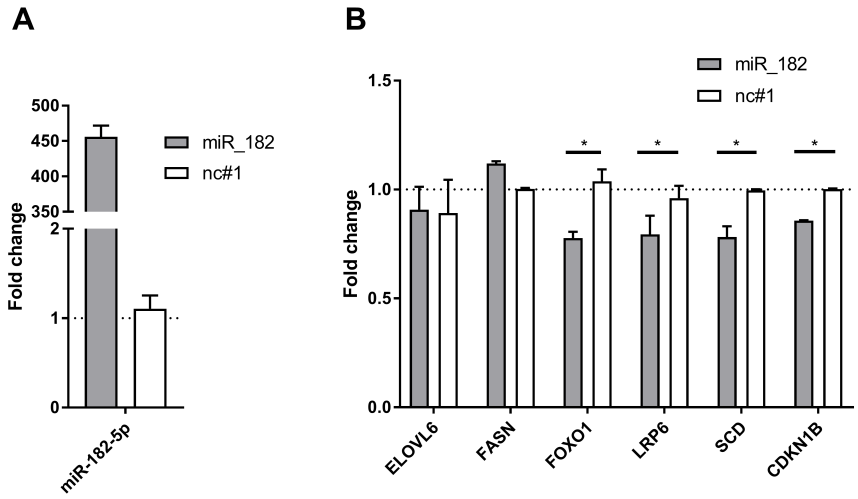


Figure 5.17: **In vitro validation of predicted target genes.** 10 nM Pri-miRNA-182-5p mimic was transfected into HepG2 cells, which shows a large overexpression of miRNA-182-5p after 48 h in comparison to cells transfected with negative control (nc#1, A). Overexpression of mimic caused decreases expression of predicted target genes *FOXO1*, *LRP6*, *SCD* and *CDKN1B* (B). It was corrected for multiple testing by controlling the FDR ($q < 0.05$ was assumed as significant which is indicated by *).

### 5.4.6 Characterization of hsa-miR-182-5p target genes *FOXO1*, *LRP6*, *CDKN1B* and *SCD*

Correlation analysis between miRNA and putative target gene expression reveals a significantly negative correlation between hsa-miR-182-5p and *FOXO1* (r = -0.28, Figure 5.18A) as well for *LRP6* (r = -0.38, Figure 5.18B). *CDKN1B* and miRNA expression do not correlate with each other (p = 0.1081, r = -0.17) though the correlation exhibits a negative trend. Also, both hepatic *SCD* and hsa-miR-182-5p expression is up-regulated in T2D, therefore a degradation of SCD mRNA by this miRNA is unlikely. Additionally, a correlation between *SCD* and miRNA expression is not significant (p = 0.0989, r = 0.3623).

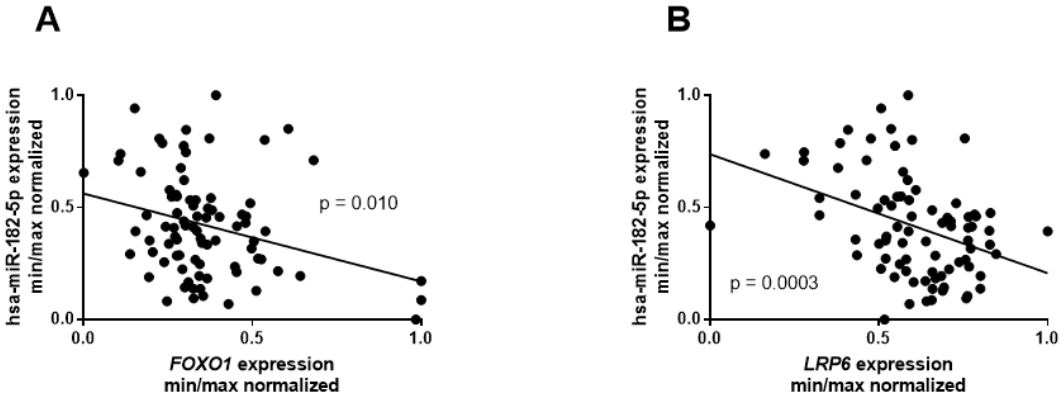


Figure 5.18: **Target genes of hsa-miR-182-5p.** Hepatic hsa-miR-182-5p expression correlates significantly with hepatic *FOXO1* (A) and *LRP6* (B) expression.

Further target gene characterization reveals a negative association of *FOXO1* expression and HbA1c levels (Figure 5.19A) and blood glucose (Figure 5.19B) which is noticeable influenced by age ($p_c > 0.05$). Interestingly, *FOXO1* expression correlates independently of age with blood cholesterol and LDL (Figure 5.19D-E).

*CDKN1B* expression correlates negatively with serum triglycerides (Figure 5.19E) which is also strongly influenced by age. No other significant associations can be found (Table 30 in supplement).

Finally *LRP6* expression correlates negatively with HbA1c levels (5.19F), blood glucose (5.19G) and serum insulin (5.19H). All associations except for blood glucose stay significant after adjustment for confounding factors though *LRP6* expression itself correlated with age. Additional regression models are listed in supplemental Table 31.

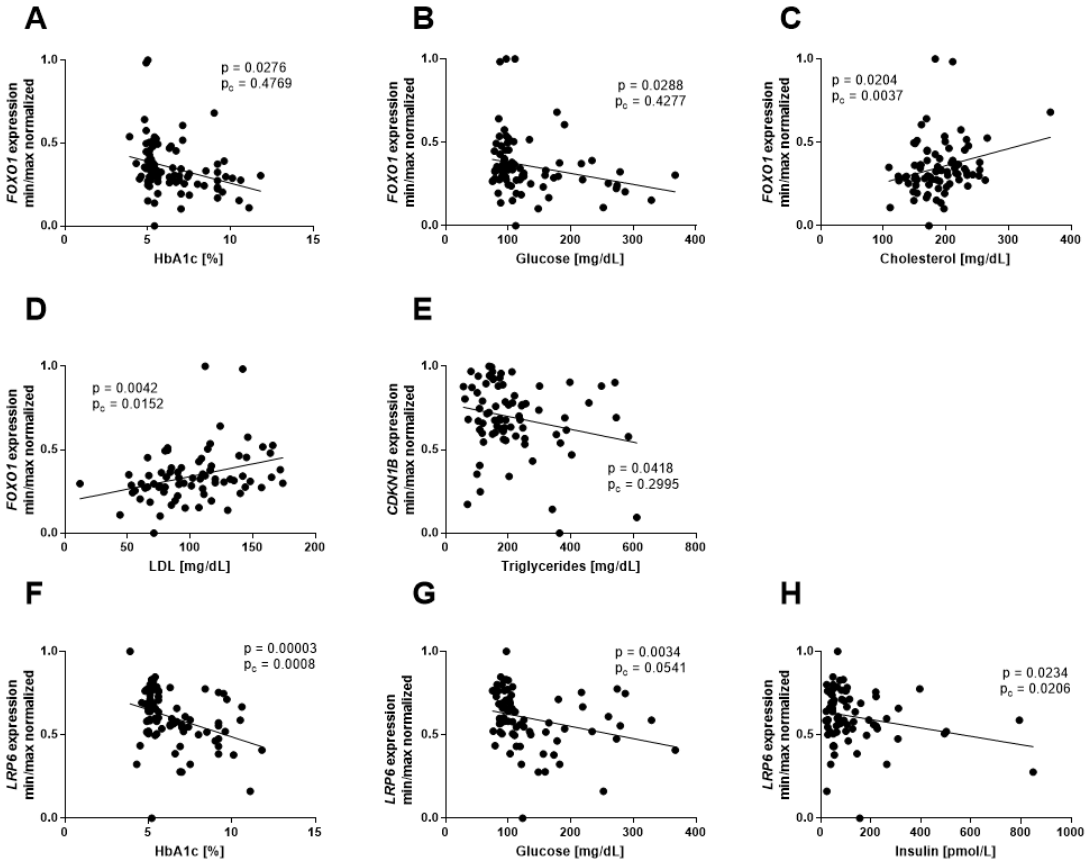


Figure 5.19: **Target gene expression correlated with metabolic traits.** *FOXO1* expression correlates negatively with HbA1c (A) and blood glucose (B), which becomes insignificant after correction for age, BMI and gender. *FOXO1* is stably altered with increased blood lipids cholesterol (C) and LDL (D). *CDKN1B* correlates negatively with serum triglycerides which is also influenced by confounding factors (E). *LRP6* expression correlates negatively with HbA1c (F), blood glucose (G) and serum insulin (H) after correction for age, BMI and gender.

Expression of potential target genes was further analyzed regarding NAS since miRNA expression is also increasingly expressed with a higher degree of liver steatosis and fibrosis. Correlation analysis reveals no changes for *CDKN1B* expression with different scores of NAS (Figure 5.20A) and an age driven negative correlation for *FOXO1* (Figure 5.20B). Expression of *LRP6* is significantly reduced in higher degrees of NAS (Figure 5.20C).

In summary, the most promising target gene of hsa-miR-182-5p remains *LRP6* due to its apparent alterations in expression for several metabolic traits and liver steatosis. On the one hand, expression of *LRP6* and hsa-miR-182-5p is significantly correlating, on the other hand *LRP6* mRNA it is also significantly reduced after overexpression of the miRNA in cell culture.

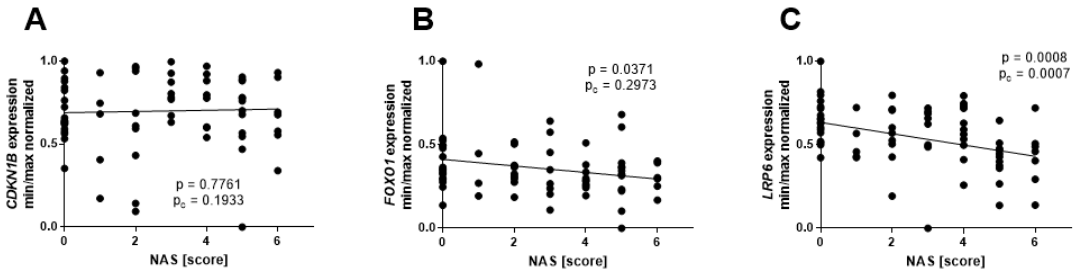


Figure 5.20: **Correlation analysis between expression of proposed target genes and grade of NAFLD.** *CDKN1B* expression does not correlate with liver steatosis and fibrosis (NAS score, A), *FOXO1* expression correlates age-dependent with NAS (B). Only *LRP6* expression is diminished with increasing NAS independently of any confounding factor (C).

### 5.4.7 Clincial relevance: Application as biomarker and induction of hsa-miR-182-5p expression

It is interesting whether serum miRNA levels mirror hepatic miRNA expression since these molecules can be secreted from the liver into the bloodstream and target other metabolically active tissues, like adipose tissue, skeletal muscle or pancreas, as already shown for other miRNAs (Wang et al., 2013b). In addition, serum is an easy accessible body fluid which enables a great potential for T2D biomarker identification and application.

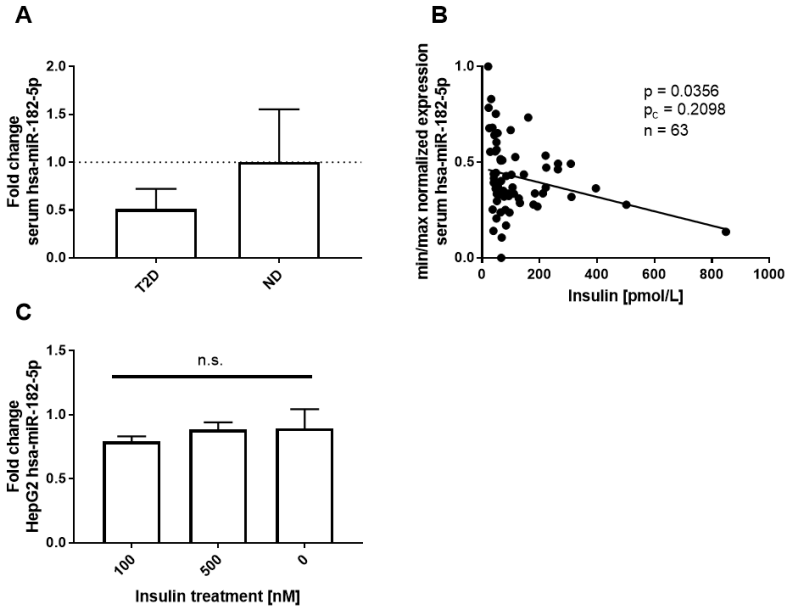


Figure 5.21: **Serum hsa-miR-182-5p expression and induction.** Serum miRNA expression is not significantly reduced in subjects with T2D (A). Due to its low expression, there is a large variance indicated by error bars. Expression correlated independently of age, gender and BMI with serum insulin (B). High levels of insulin and glucose in media is not inducing miR-182-5p expression in vitro (C).

Serum hsa-miR-182-5p concentrations are low in both subgroups (mean FC in T2D 0.51 +/- 0.21 in comparison to ND subjects with an error of +/- 0.55) and exhibit a large variability between all subjects without influence of age, BMI or gender. Subgroup stratification reveal no significant difference between T2D and ND serum miRNA concentrations ($p = 0.9706$, Figure 5.21A). There is also no correlation between serum miRNA concentration and hepatic expression ($p = 0.3284$). Nevertheless, miRNA con-

centrations appear to be decreased in subjects with high serum insulin concentrations (Figure 5.21B). This reduction is mostly driven by the cohort design ($p_c > 0.05$).
To elucidate whether miR-182-5p expression can be induced by insulin and high glucose (mirroring a state of IR), HepG2 cells were treated for 24 h with different concentrations of insulin (100 nM or 500 nM) in high glucose (4.5 g/L) medium in comparison to cells without treatment in low glucose (1.5 g/L) medium. HepG2 miRNA expression is not significantly altered between different concentrations (Figure 5.21C).
Determination of serum hsa-miR-182-5p concentrations proved not to be a suitable T2D biomarker due to the lack of association between serum miRNA concentrations to alterations in metabolic traits. The missing induction of miRNA expression upon insulin treatment in cell culture indicate an increase in hsa-miR-182-5p expression is not an acute process and might be a consequence of IR than a cause.

## 5.5 DNA methylation, genetics and miRNA expression in synergy regulating *IRS2* expression

Previous results depicted different mechanisms of epigenetic regulation. Covalent attachment of methyl groups to a cytosine of CpG oligonucleotides may only effect one specific gene. In contrast, altered expression of miRNAs may consequently change expression of multiple target genes at the same time within the same system. Additionally, a regulatory function of SNPs must be considered for both epigenetic mechanisms. These permanent genetic changes may affect directly CpG sites by eliminating methylation or adding new methylation sites (Aberg et al., 2018; Dayeh et al., 2013; Izzi et al., 2016) and SNPs may also change 3'-UTR recognition motifs for miRNA binding (Moszyńska et al., 2017) or miRNA maturation (Gong et al., 2012).
These described mechanisms may also act in an synergistic manner. The CpG-SNP cg12195446/rs4547213 within intron 1 of *IRS2* was recently associated to T2D (Mahajan et al., 2018). Moreover, observed dysregulation of *IRS2* and correlation with metabolic traits (see previous Figure 5.2) fits into the model of selective IR. Finally, age associated hsa-let-7e-5p targeting *IRS2* was deferentially expressed between T2D and ND subjects of the miRNA microarray discovery cohort (see previous Figure 5.13).
These arguments constitute a careful characterization of *IRS2* expression and regulation in T2D manifestation.

### 5.5.1 Hepatic *IRS2* expression

Hepatic *IRS2* expression correlates positively with BMI ($p = 0.0171$) and is not significantly altered between T2D and ND subjects after correction for all confounding factors age, gender and BMI (Figure 5.22A). Since expression does not correlate with age ($p = 0.4614$) or gender ($p = 0.3842$), correction solely for BMI results into a significant association between *IRS2* expression and presence of T2D ($p = 0.0421$).
Further regression analysis reveals a significant negative correlation between *IRS2* and NAS (Figure 5.22B), HbA1c (Figure 5.22C), AST (Figure 5.22D) and ALT (Figure 5.22E). Correlation between *IRS2* expression and serum insulin fails to reach significance after correction for age, gender and BMI (Figure 5.22F). *IRS2* proved to be a suitable candidate gene based on the correlation between gene expression and NAS score, as well a correlation between expression and HbA1c. Though *IRS2* expression also correlates with BMI, an association to the presence of T2D remains unaffected by obesity. NAS score is a representative for liver steatosis and fibrosis state, whereas high HbA1c values indicate a long-term exposure to high blood glucose concentrations. Furthermore its expression correlates with liver enzymes AST and ALT which represent a surrogate marker for overall liver health status.

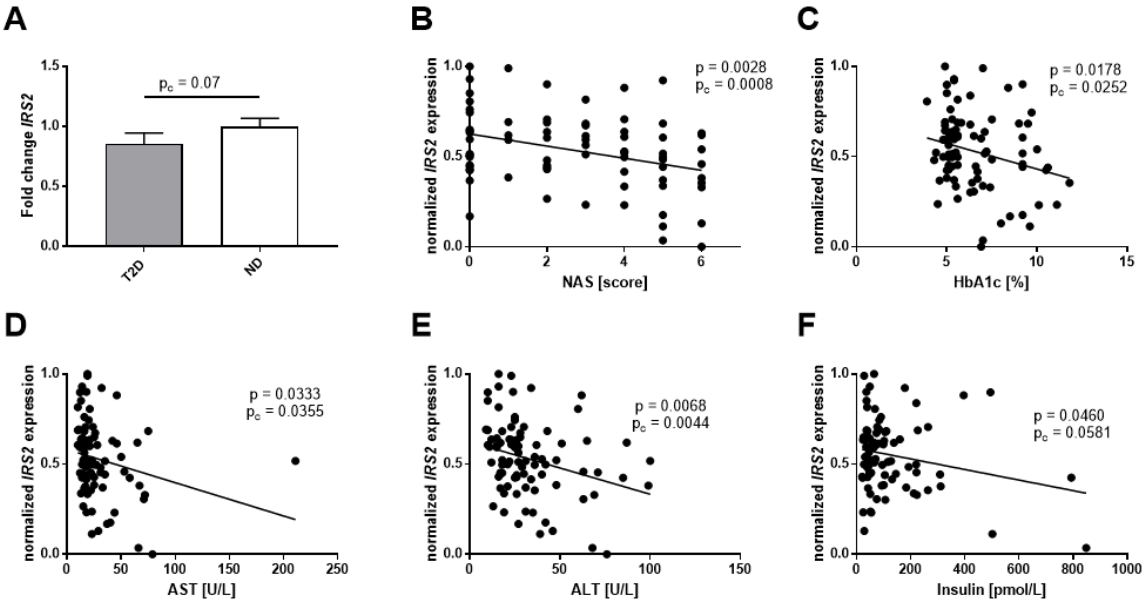


Figure 5.22: **Hepatic *IRS2* expression in dependency of liver health status.** Absolute *IRS2* expression is not significantly altered between T2D and ND subjects (A). Expression correlates negatively with NAS score (B), HbA1c (C), AST (D) and ALT (E) after correction for age, gender and BMI. Expression correlated with serum insulin, which becomes insignificant after correction for age, gender and BMI (F).

### 5.5.2 DNA methylation of *IRS2* intron 1

Bisulfite pyrosequencing of *IRS2* intron 1 reveals two CpG oligonucleotides, namely CpG3 and CpG6, with high variance, though surrounding CpGs appear to be completely methylated (Figure 5.23A). Position 1 represents the CpG-SNP cg12195446/rs454721 and shows variable methylation in dependency of the genotype (full, hemi or no methylation).

CpG6 displays the highest variance, nevertheless CpG6 methylation correlates positively with *IRS2* expression (Figure 5.23B). Although CpG3 methylation correlates negatively with HbA1c values (Figure 5.23C) it fails to reach significance if mean values of T2D and ND subjects are compared (Figure 5.23D). Methylation at CpG6 is significantly decreased (about 3.7 %) in T2D subjects (Figure 5.23E). *In silico* analysis for transcription factor binding motifs within intron 1 of *IRS2* was performed to elucidate a possible coherence between methylation and expression. Thus two DNA regions with high similarity for specificity protein 1 (Sp1) and SREBF1 recognition motifs were found (transcription factor binding profiles from JASPAR see Figure 5.23F and location within intron 1 of *IRS2* see Figure 5.23G). Interestingly, CpG3 and CpG6 which display the highest variability between all subjects (CpG3 81.2 % +/- 5.2 % and CpG6 77.8 % +/- 11.5 %) are located within these two motifs (CpG3 within Sp1 GC box and CpG6 within SREBF1 E box motif).

DNA methylation at any position is neither affected by age, gender or BMI. Therefore correction is applied to correct for differences within both cohorts.

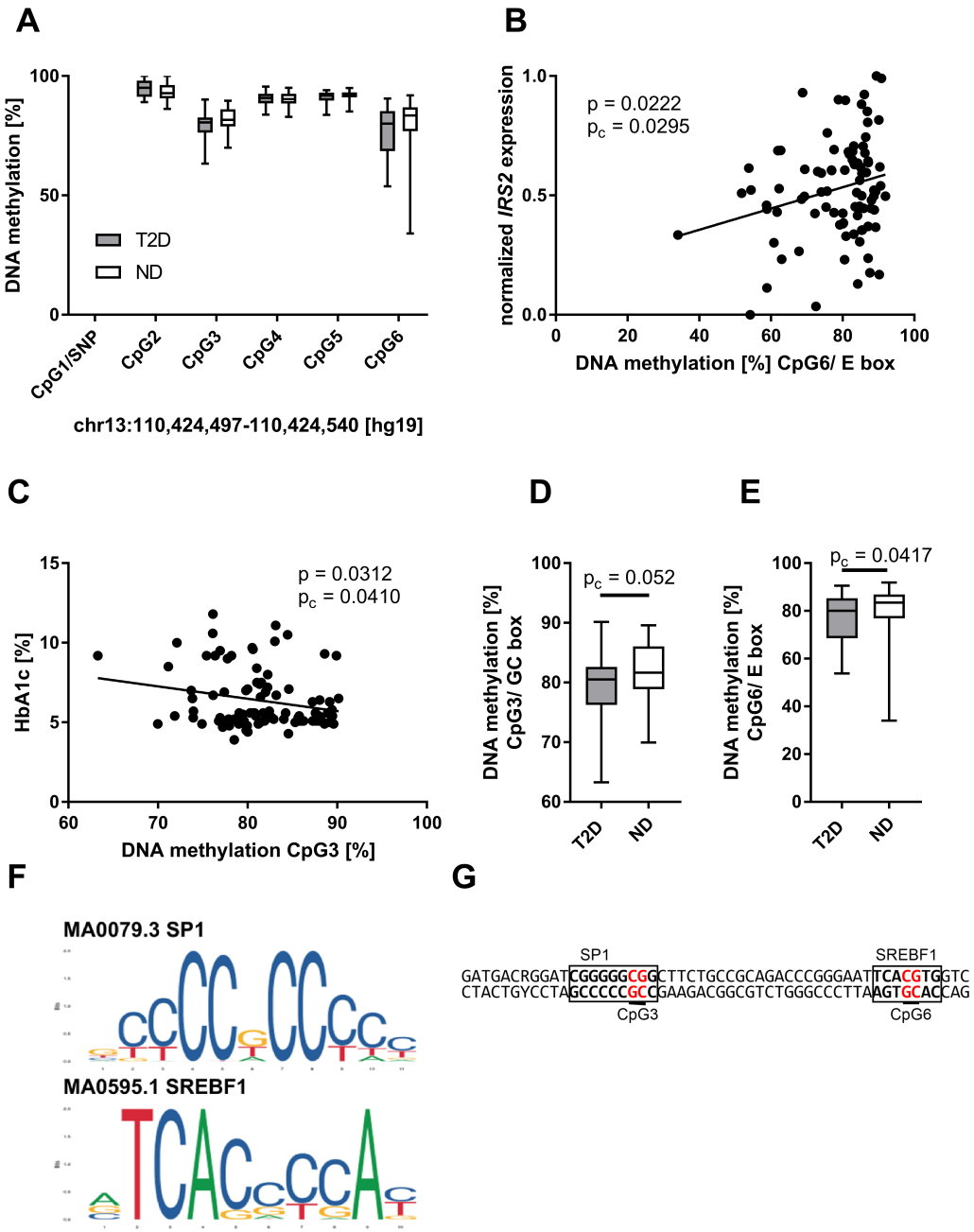


Figure 5.23: **DNA methylation pattern of *IRS2* intron 1.** DNA methylation is dynamic across intron 1 (A). Methylation at CpG6 correlates positively with gene expression (B) and DNA methylation at CpG3 correlates negatively with HbA1c values (C). DNA methylation at CpG3 is not significantly decreased in subjects with T2D (D), but about 3.7 % at CpG6 (E). Two recognition motifs of Sp1 and SREBF1 (F) can be found at CpG3 (Sp1 GC box, G) and at CpG6 (SREBF1 E box, G)

### 5.5.3 Genetics of *IRS2*

As indicated before, the first analyzed position within intron 1 of *IRS2* is the CpG-SNP cg12195446/ rs4547213. The polymorphism rs4547213 changes the guanine of a CpG oligonucleotide into an adenine (G>A, minor allele frequency MAF(A) = 28.2 % on 1000 Genomes) and thereby erasing methylation completely in homozygous carrier of the minor allele A. The MAF of the whole cohort is about 38.0 % (MAF in T2D 30.0 % and MAF in ND 41.3 %).

After accounting for Hardy-Weinberg equilibrium, both a chi squared test and regression analysis reveal

no association of the SNP to the incidence of T2D ($\chi^2 < 3.841$ with $\alpha = 0.05$ and one degree of freedom). The minor allele was more prominent in ND subjects presuming a protective effect (Figure 5.24A). Stratification of all subjects into their respective genotypes reveals an association of genotype and methylation at CpG3 (Figure 5.24B) and CpG6 (Figure 5.24C). Further regression analysis regarding additionally rs4547213 as confounder still leads to significant associations between methylation at CpG3 and HbA1c levels and serum insulin (Table 33 in supplement). Thereby increasing the effect size (estimate) on serum insulin from 4.28 into 4.64 pmol/L per 1 % change in methylation. The effect size on HbA1c is not altered by correction for the genotype.

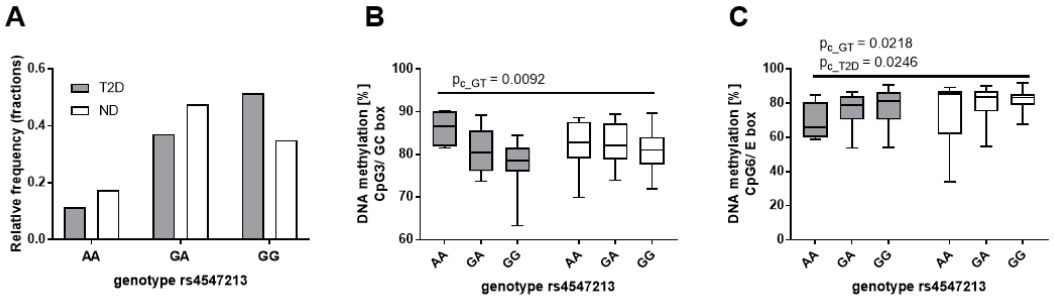


Figure 5.24: **Influence of polymorphism rs4547213 on DNA methylation.** The G allele of rs4547213 is more common in T2D, indicating the G allele as risk allele (A). CpG3 methylation (GC box motif of Sp1) is significantly influenced by the genotype (B). CpG6 methylation (E box motif of SREBF1) is significantly influenced by genotype and disease state (C).

### 5.5.4 Regulation of *IRS2* expression by miRNA hsa-let-7e-5p

Hepatic hsa-let-7e-5p is up-regulated in T2D subjects of the microarray discovery cohort. Measurement of expression by qPCR of this miRNA in the whole cohort revealed a 1.26 up-regulation in T2D subjects after correction for age, gender and BMI (Figure 5.25A). Any association of miRNA expression and age can not be reproduced (Table 32 in supplement).

*IRS2* and hsa-let-7e-5p expression correlates significantly indicating a high miRNA occurrence and presumably concomitant repressed *IRS2* expression (Figure 5.25B).

Hepatic hsa-let-7e-5p expression correlates with NAS (Figure 5.25C) but no other metabolic trait (Table 32 in supplement). A high miRNA expression is denoted by a high grade of liver steatosis and fibrosis.

Besides the ability to repress target genes within the same cell of expression, miRNA can be secreted into the blood stream via exocytosis or other mechanisms (Sohel, 2016). As a consequence, circulating extracellular miRNAs are considered as highly potential biomarkers (Kim, 2015; Santangelo et al., 2017). Extraction of serum miRNAs and consequently qPCR measurement reveals no change in extracellular hsa-let-7e-5p proportions between T2D and ND subjects but a positive correlation between hepatic and serum miRNA amount (Figure 5.25D). Apparently serum miRNA concentration is not able to stratify subjects into their respective disease groups but mirrors the hepatic hsa-let-7e-5p situation.

In summary, *IRS2* is a suitable candidate gene whose expression pattern and correlations fit into T2D manifestation. Aberrant expression might be induced by changes in DNA methylation of two CpG oligonucleotides located within a Sp1 and SREBF1 recognition motif and additionally by overexpression of hsa-let-7e-5p which targets *IRS2* mRNA. These findings are summarized in Figure 5.26. Though methylation at CpG5 appears to correlate with some traits, its hypermethylation ranging within 91 % +/- 2 % causes physiological irrelevance.

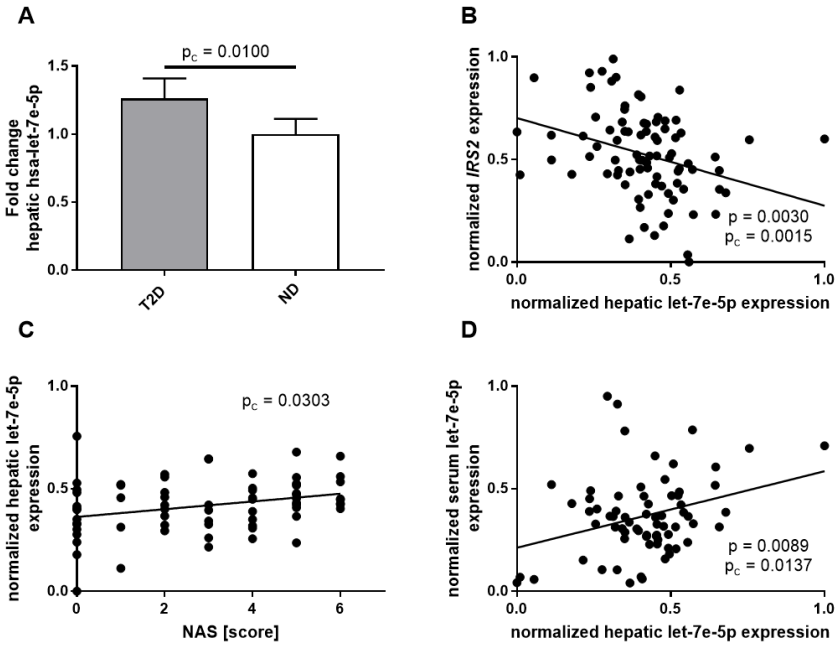


Figure 5.25: **Hsa-let-7e-5p expression in liver and serum.** Hepatic hsa-let-7e-5p expression is significantly upregulated in T2D (A) and correlates negatively with hepatic *IRS2* expression (B). Hsa-let-7e-5p expression correlates positively with NAS score (C). Hsa-let-7e-5p expression in liver is mirrored by serum expression (D).

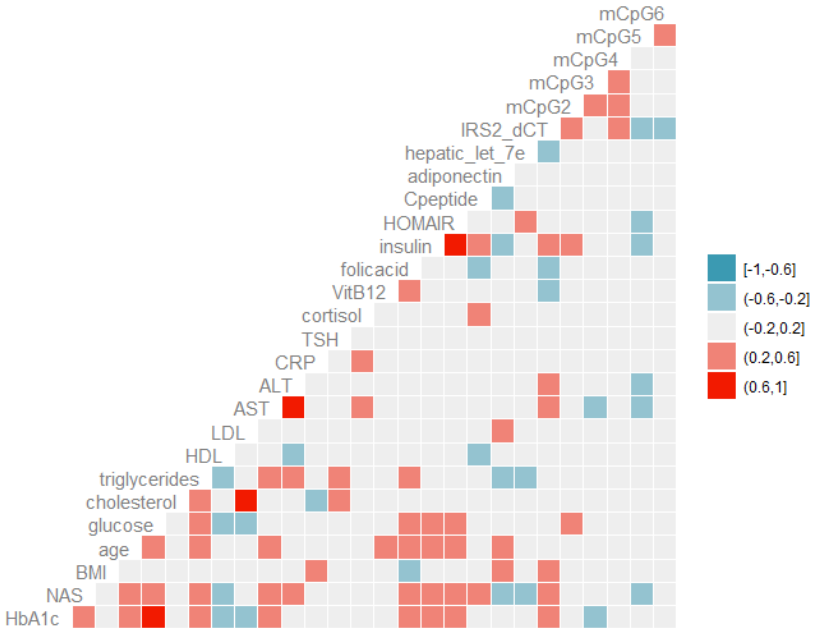


Figure 5.26: **Correlation matrix for hepatic *IRS2* expression, hepatic hsa-let-7e-5p expression and DNA methylation.** Blue indicates a negative, red a positive correlation.

### 5.5.5 Functional analysis of *IRS2* expression modifying changes

*IRS2* intron 1 is aberrantly methlyated between T2D and ND subjects but it remains unclear if this methylation or the SNP itself has an effect on *IRS2* expression and whether proposed transcription factors bind to their predicted binding motifs. Therefore respective genomic parts containing either CpG1 or the polymorphism were cloned into a plasmid containing a Luciferase reporter gene under control of a CMV (cytomegalovirus) promoter. HepG2 cells were transfected in at least triplicates with respective plasmids and lysed after 24 h of incubation. Luciferase signal was subsequently measured.

The SNP rs4547213 itself has no influence on Firefly Luciferase (*FLuc*) gene expression (Figure 5.27A) but full methylation at CpG1 to CpG6 (IRS2 CpG1-6) or CpG2 to CpG6 (IRS2 SNP-CpG6) causes a reduction of *FLuc* signal (Figure 5.27A). Dividing intron 1 for Luciferase plasmid cloning into three parts (CpG1 only, CpG1 to CpG3 and CpG1 to CpG6 completely) revealed a significant repression of *FLuc* signal for solely CpG1 and the complete sequence spanning CpG1 to CpG6 in methylated state (Figure 5.27B).

Previous methylation measurement indicated a lower expression of *IRS2* in hypomethylated state, which is in contrast to Luciferase reporter gene assay results. Nevertheless, this particular region in intron 1 appears to have a regulatory function which is not fully mirrored by our used artificial systems.

An electrophoretic mobility shift assay (EMSA) was used to validate binding of Sp1 and SREBF1 to their predicted recognition sequences. Therefore double stranded oligonucleotides spanning CpG1 to CpG6 were incubated with either nuclear extract from HepG2 cells or nuclear extract previously incubated with specific antibodies against human Sp1 or SREBF1. Gel electrophoresis separates free biotinylated oligonucleotides from protein-bound DNA which is shifted to the top due to its increased size. Additional incubation with different excess of unlabeled oligonucleotides competes for protein binding and deletes a visible shift indicating specificity of protein-DNA-interaction.

Incubation of 50 fmol biotinylated oligonucleotides with HepG2 nuclear extract leads to a shift (Figure 5.27C line 2) in comparison to free oligonucleotide (lane 1). Additional incubation of decreasing concentrations (12.5 pmol to 0.125 pmol) unlabeled oligonucleotides leads to a gradual extinction of this shift (lanes 3-5 for methylated and lanes 6-8 for unmethylated oligonucleotides). Thereby favoring of unmethylated DNA is only vaguely visible for 1.25 pmol excess (lanes 4 and 7).

Incubation of labeled DNA with nuclear extract and specific antibodies leads to a reduction of binding affinity for Sp1 (lane 9) and strong increase of binding affinity for SREBF1 (lane 10). The change in affinity after incubation with respective antibody serves as proof which protein is binding. Lane 11 represents a control for buffer ingredients which might cause a shift and remains negative.

As already indicated by Luciferase assay, the polymorphism does not change protein binding (Figure 5.27D). Lane 1 represents DNA containing SNP at position 1 to CpG6 instead of the CpG1 (lane 3 in comparison) which was incubated with nuclear extract from HepG2 cells. Labeled DNA is shifted in the same pattern for both oligonucleotides and binding is abolished by incubation with unlabeled DNA (lane 2).

Binding of hsa-let-7 family members on *IRS2* mRNA was described previously (Zhu et al., 2011; Gao et al., 2014; Frost and Olson, 2011). Therefore functional studies focused on the induction of hepatic hsa-let-7e-5p. HepG2 cells were treated for 24 h with different concentrations of insulin (100 nM or 500 nM) in high glucose (4.5 g/L) medium to mirror the state of IR. For comparison, cells without treatment (0 nM insulin) were cultivated in low glucose medium for 24 h. Measurement of HepG2 let-7e-5p by qPCR reveals more than 2.5 fold increase of miRNA expression independently of insulin concentration (Figure 5.27 E). These results indicate a high acute responsiveness of hsa-let-7e-5p expression to increasing insulin levels. Up-regulation of hepatic hsa-let-7e-5p might therefore be an early process before T2D manifestation.

Additional cell culture experiments with HepG2 cells failed due to the low expression (CT values around 35) of *IRS2*. Nevertheless, transfection of 10 nM hsa-let-7e-5p mimic into HepG2 cells and incubation for

48 h indicates a repression of HepG2 *IRS2* expression, though it is a questionable effect reasoned on the already low expression (Figure 5.28). Other potential target genes (*CCND1*, *FASN*, *FOXO1*, *GAPDH* or *SCD*) are not affected.

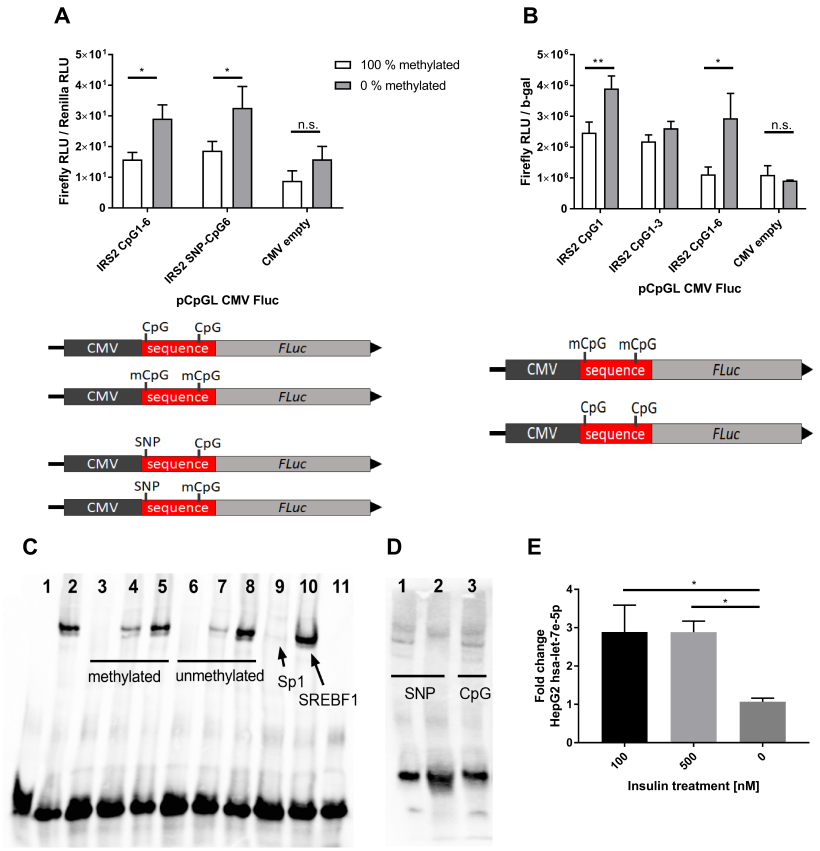


Figure 5.27: **Mechanistical analysis of changes.** A polymorphism and methylation dependent luciferase reporter gene assay revealed no direct influence of rs4547213 on gene expression (A). A methylation sensitive Luciferase reporter gene assay revealed a significant influence of solely CpG1 and CpG1 to CpG6 on gene expression (B). Used luciferase constructs are indicated below. An electrophoretic mobility shift assay (EMSA) revealed neither a differences on protein binding by methylated and unmethylated oligos (C) nor by oligos containing the polymorphism (D). By antibody binding, Sp1 and SREBF1 (arrows, C) are identified as bound proteins. Treatment of HepG2 cells for 24 h with insulin and high glucose medium is able to iduce HepG2 hsa-let-7e-5p expression independently of insulin concentration (E).

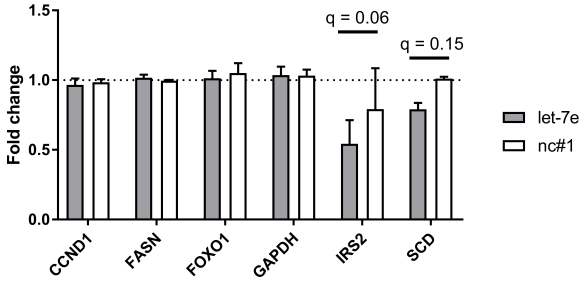


Figure 5.28: **Transfection of hsa-let-7e-5p mimic into HepG2 cells.** Especially *IRS2* expression in HepG2 cells is not sufficient for reliable measurements, indicated by high CT values $> 35$. Nevertheless, fold change analysis implies a repression of *IRS2* expression. Correction for multiple testing is applied by controlling the FDR ($q < 0.05$).

### 5.5.6 Clincial relevance: Synergy of genetics and miRNA as potential biomarker

Identification of molecular pathomechanisms in disease etiology might be useful for further development of T2D treatments by targeting these identified pathways. Since an effective treatment is missing and T2D remains incurable, prevention and intervention is the best strategy against the rising prevalence of the world wide T2D occurrence.

One method of prevention is risk group stratification by biomarker analysis. Therefore reliable biomarkers will be needed which are easy to access and cheap in measurement. We were interested whether our identified regulatory mechanisms, DNA methylation of *IRS2* intron 1 and hepatic hsa-let-7e-5p, might act as potential new biomarker. Information about age, gender and BMI, which are called confounder, were supplemented with information about genotype at rs4547213 and serum hsa-let-7e-5p $\Delta$CT values. This combination is called blood based because both values can be acquired by blood drawing. The genotype is a suitable surrogate marker since DNA methylation at CpG3 is influenced by the genotype and correlates with HbA1c levels which again is a marker for long time exposure of high blood glucose levels. Also serum miRNA expression correlates with hepatic miRNA expression and *in vitro* experiments indicate an induction of acute hepatic expression by high insulin and high glucose concentrations.

In a third setup, confounder and blood based markers were supplemented with additional hepatic DNA methylation information of CpG3 and CpG6. This setup is not favorable due to the low accessibility of hepatic DNA.

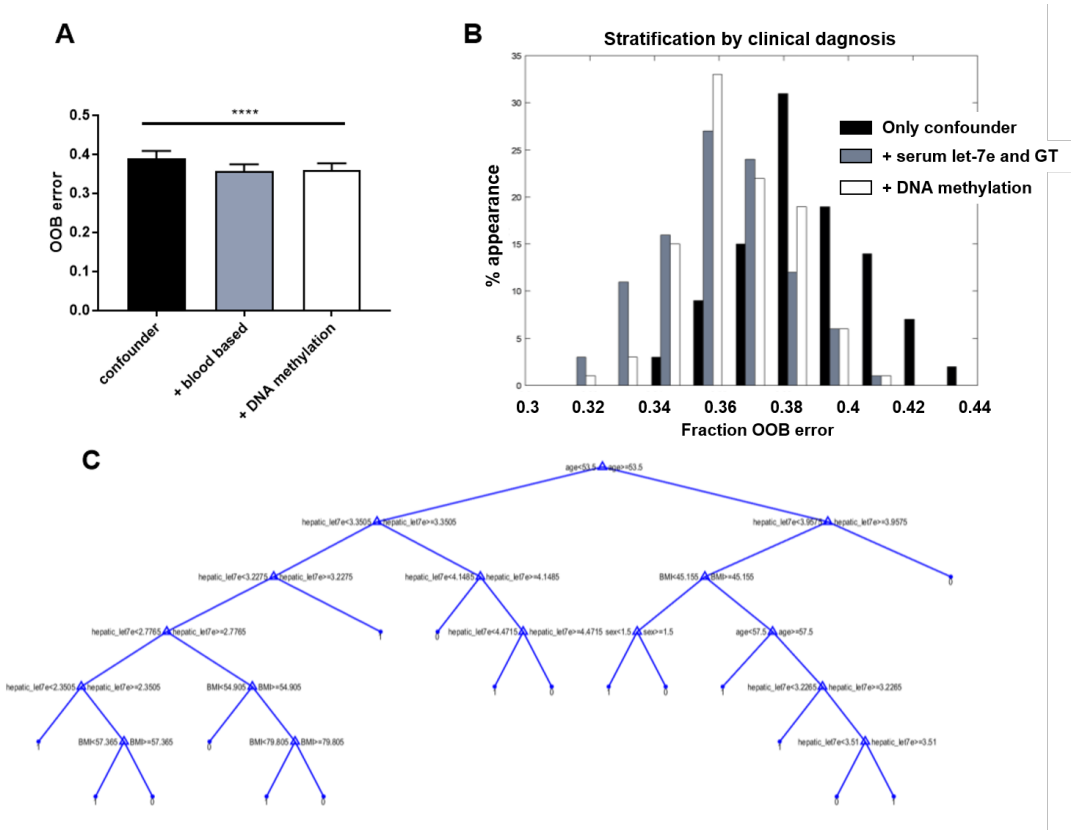


Figure 5.29: **Robust machine learning approaches to generate a biomarker.** Random forest classification revealed an improved classification (reduced out-of-bag error, OOB) of T2D and ND subjects if using blood based markers serum hsa-let-7e-5p expression and genotype at rs4547213 besides confounding factors age, BMI and gender (A). Distribution and limits of OOB error for different models: Only confounder (age, BMI and gender), in addition with blood based marker (serum hsa-let-7e-5p expression and genotype at rs4547213) and in addition with hepatic DNA methylation at CpG3 and CpG6 (B). Expample for a decision tree in random forest model generation (C). For all models, information of n = 84 subjects was used.

Robust machine learning approaches were used for biomarker development and validation. Linear regression models already exist for disease prediction (Fujita et al., 2012; Tabaei and Herman, 2002; Zou et al., 2017), but we hypothesize a non-linear coherence which can be modeled by decision trees. Random decision forest (RDF) containing several decision trees works on small sample sizes with few predictor sets. Bootstrapping approaches for training and test data division increases robustness of this classification approach (Le Teuff et al., 2005).
For crossvalidation of each RDF, an out-of-bag (OOB) error was calculated (mean OOB error, see Figure 5.29A). At least 100 RDFs with b = 200 decision trees were generated. A confounder based classification reaches a mean OOB error of 39.1 %. The mean OOB error decreased to 35.8 % after supplementation of blood based information. Additional information on DNA methylation does not improve classification (mean OOB error of 36.1 %). Equal distributions of the OOB error across all three predictor sets validate the robustness of this method for the application of RDF in T2D prediction (Figure 5.29B). An example decision tree is shown in Figure 5.29C. Each split point assigns a subjects into a respective branch by different threshold values of predictors (for example age $< 53.3$ years) until either a certain depth (fixed number of split points) is achieved or an explicit categorization of the subject into a group (1 for T2D, 0 for ND) is possible. For each subject, this is performed for b = 200 trees. Afterwards, majority voting of 200 decisions determines the final category.
In summary, a robust machine learning approch by RDF was able to classify 64.2 % of the used cohort (n = 84) correctly into subjects with T2D and ND subjects only by using blood based information on epigenetic regulation. The inclusion of more validated predictors might further improve classification to make this approach usable in clinical applications.

## 5.6 Target and tissue matched miRNA

Though a systematic approach by microarray analysis did not reveal any noticeable associations within the discovery subgroup, two specific miRNA were further characterized in the whole liver cohort.
The first miRNA, hsa-miR-223-3p, was identified by data base research which revealed a binding motif within the 3'UTR of *GALNT18* and *FOXO1* (target matched approach). As described before, this gene emerges as an promising metabolically involved candidate gene with unknown function and regulatory mechanism.
As a second candidate, hsa-miR-122-5p was analyzed in the full hepatic cohort. This miRNA is the most prominent and abundant hepatic miRNA (tissue matched approach). Data base research revealed two genes, *INSR* and *ACACB*, as potential target genes. Interestingly, expression of these genes was also shown to be reduced in diseased liver of subjects within our cohort. Therefore we were interested whether miRNA mediated repression is a the underlying regulatory mechanism.

### 5.6.1 Hepatic hsa-miR-223-3p expression correlates with blood lipids and increased age dependently with blood glucose markers

Since *GALNT18* methylation is not influencing its gene expression, also repression by miRNA should be considered as additional regulatory mechanism. Seed analysis reveals hsa-miR-223-3p as possible miRNA by targeting the 3'-UTR of *GALNT18* as an 8mer (see section 10.3.3 for tool). This assumption is supported by several database hits for hsa-miR-223-3p using NetAffx Analysis Center and miRCarta (Backes et al., 2018).
Hepatic hsa-miR-223-3p expression was not significantly altered between T2D and ND subjects ($p = 0.081$ and $p_c = 0.064$ after correction for age, n = 90) or in NAFLD progression (regression with NAS score: $p = 0.2412$ and $p_c = 0.6838$, n = 74) but is higher expressed in liver of older subjects (Table 26). Additionally, there is no correlation between hepatic *GALNT18* and miRNA expression (Table 26). However miRNA expression correlates significantly with serum triglycerides and HDL concentrations

(Table 26). Further database research ascertain *FOXO1* as another only by Luciferase assay validated target (Wu et al., 2013). Correlation analysis reveals no relation between expression of both expressions. Therefore hepatic hsa-miR-223-3p might have an age dependent metabolic function but its target gene remains unclear at the moment. *In vitro* validated or *in silico* predicted target genes are unlikely also *in vivo* hepatic target genes.

Table 26: **Linear regression models of hepatic hsa-miR-223-3p.** Bold values are considered significant.

| **hsa-miR-223-3p** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | **0.01048** | 15.12 | 0.08611 | 9.48 | 85 |
| insulin | 0.37155 | 12.05 | 0.37763 | 12.75 | 84 |
| HOMA-IR | 0.09286 | 2.23 | 0.15964 | 1.99 | 79 |
| HbA1c | **0.02439** | 0.37 | 0.21525 | 0.19 | 90 |
| BMI | 0.09765 | 1.63 | | | 90 |
| Age | **0.02751** | 2.54 | | | 90 |
| Gender | 0.46477 | -0.20 | | | 90 |
| Cholesterol | 0.41202 | -3.18 | 0.28279 | -4.30 | 83 |
| Triglycerides | **0.00189** | 35.08 | **0.01116** | 28.01 | 83 |
| HDL | **0.00064** | -3.99 | **0.00339** | -3.55 | 83 |
| LDL | 0.08165 | -6.30 | 0.14345 | -5.69 | 77 |
| AST | 0.65265 | 1.11 | 0.86320 | 0.45 | 85 |
| ALT | 0.41518 | 1.58 | 0.54039 | 1.20 | 85 |
| *GALNT18* | 0.24302 | 0.04 | 0.22691 | 0.05 | 90 |
| *FOXO1* | 0.68752 | 0.08 | 0.17837 | 0.29 | 87 |

### 5.6.2 Hepatic hsa-miR-122-5p expression is decreased for blood glucose markers

The miRNA hsa-miR-122-5p represents the most abundant miRNA species in liver (Filipowicz and Grosshans, 2011). Therapeutically, this miRNA is best known for the necessity of HCV replication and respective antagomirs targeting hsa-miR-122-5p represent attractive pharmacological fields for drug development (Luna et al., 2015; van der Ree et al., 2016). Physiologically, this miRNA contributes to the regulation of hepatic lipid metabolism (Esau et al., 2006; Wen and Friedman, 2012). In the context of T2D and insulin resistance, raised serum levels of this miRNA appear to predict the incidence of T2D in humans (Willeit et al., 2017).

Additionally, two potential target genes with binding motifs of hsa-miR-122-5p within the 3'UTR have a decreased hepatic expression in T2D (*ACACB*) or IR (*ACACB* and *INSR*, Figure 5.2). Characterization of hepatic hsa-miR-122-5p might explain whether T2D manifests in dependency of miRNA expression alterations by affecting those mentioned target genes.

Hepatic miR-122-5p expression is significantly repressed with increasing NAS (Figure 5.30A). Nevertheless, hepatic miR-122-5p expression varies strongly in diabetic liver (mean FC = 0.73 +/- 0.47), therefore a comparison of its expression between T2D and ND subjects reveals a repression in T2D without significance (Figure 5.30B). Further regression analysis reveals a significant repression of hepatic miRNA expression with increasing blood glucose, HbA1c, AST and ALT (Figure 5.30C-F) independent of age, gender or BMI (Table 27).

Though hepatic hsa-miR-122-5p is not significantly reduced in T2D, its expression correlates negatively with HbA1c. The potential target genes of this miRNA (*INSR* and *ACACB*) have also reduced expression in diabetic liver, making hsa-miR-122-5p binding to either *INSR* or *ACACB* 3'-UTR not plausible. Also correlation analysis between miRNA expression and potential target genes prove no significant association

(Table 27). A reduction of *ACACB* expression as a cause of miR-122-5p knock-down is in accordance to previous studies (Esau et al., 2006).
All in all, alterations in liver-specific miR-122-5p correlating with liver health status and blood glucose markers could be reproduced in this cohort although these alterations were in an opposite direction than indicated from literature. Additionally, potential metabolic target genes of this miRNA which were also altered in the diseased liver are excluded due to lack of significance (*INSR*) or an unplausible correlation direction (*ACACB*).

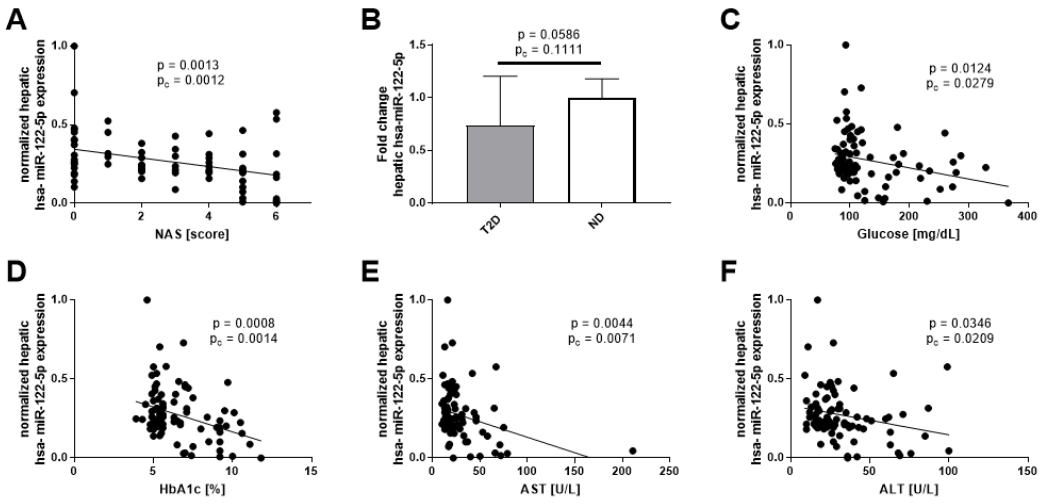


Figure 5.30: **Characterization of hepatic hsa-miR-122-5p expression.** miRNA expression is repressed with increasing NAS (A) but is not deferentially expressed between T2D and ND subjects (B) due to the high variance. Furthermore hepatic miRNA expression is repressed with increasing blood glucose (C), HbA1c (D), AST (E) or ALT (F).

Table 27: **Linear regression models of hepatic hsa-miR-122-5p.** Bold values are considered significant.

| **hsa-miR-122-5p** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | **0.01238** | -19.91 | **0.02788** | -15.58 | 84 |
| Insulin | 0.19444 | -25.03 | 0.05925 | -30.58 | 82 |
| HOMA-IR | 0.15828 | -2.45 | 0.07290 | -2.79 | 77 |
| HbA1c | **0.00082** | -0.73 | **0.00143** | -0.61 | 89 |
| BMI | 0.40610 | -1.10 | | | 89 |
| Age | 0.20006 | -1.97 | | | 89 |
| Gender | 0.73426 | -0.07 | | | 89 |
| Cholesterol | 0.49593 | -3.54 | 0.74069 | -1.70 | 82 |
| Triglycerides | 0.43349 | -12.20 | 0.70238 | -5.50 | 82 |
| HDL | 0.70964 | 0.60 | 0.73649 | 0.53 | 82 |
| LDL | 0.92159 | -0.44 | 0.96183 | -0.22 | 76 |
| AST | **0.00437** | -9.20 | **0.00706** | -8.82 | 84 |
| ALT | **0.03462** | -5.46 | **0.02088** | -5.79 | 84 |
| *INSR* | 0.11939 | -0.10 | 0.07003 | -0.12 | 85 |
| *ACACB* | 0.06715 | 0.13 | 0.14946 | 0.10 | 86 |

# 6 Discussion

It is hypothesized that key metabolic genes show aberrant expression in diabetic liver due to changes in epigenetic regulation.

The identified potential underlying mechanism includes changes in DNA methylation within transcription factor binding motifs and aberrant hepatic miRNA expression. DNA and whole cell RNA was extracted from liver biopsies of 101 obese (BMI > 32) human subjects with different stages of NAFLD (indicated as NAS score), insulin resistance and diagnosed T2D. Based on a known connection between TH an NAFLD, respective genes contributing to hepatic signaling were additionally analyzed. Since liver biopsies were taken only once, this is a cross-sectional snap shot of a hepatic dysregulation in T2D and NAFLD and cannot answer a cause or consequence coherence. Nevertheless, the large sample size of human hepatic tissue and the BMI optimized cohort allows a solid basis to identify molecular mechanisms of T2D manifestation without a bias by obesity, which is an epigenetic key driver in metabolic diseases.

The link between metabolic tissue specific changes in DNA methylation and alterations in gene expression in the manifestation of T2D is already described for pancreatic beta cells and skeletal muscle (Volkmar et al., 2012; Nitert et al., 2012; Lindholm et al., 2014). Also epigenome-wide studies for adipose tissue were performed (Nilsson et al., 2015). Three studies for hepatic alterations in DNA methylation and gene expression were performed but only few target genes were identified (Ahrens et al., 2013; Nilsson et al., 2015; Kirchner et al., 2016). However, changes in hepatic DNA methylation is still a plausible cause for T2D manifestation. Two candidate methylation sites which emerged from such hepatic EWAS, promoter methylation of *TP53INP1* and intragenic methylation of *GALNT18*, were chosen for reproduction in this large hepatic cohort. We were also the first to show a connection between hepatic *GALNT18* methylation, gene expression and glucose metabolism. Moreover, two most promising blood-based EWAS markers for T2D and obesity, namely *SREBF1* and *ABCG1*, were carefully evaluated in human metabolically active tissue.

A second regulatory mechanism based on changes in miRNA expression pattern was further assumed. Previously, most analysis regarding aberrant miRNA expression in T2D manifestation and progression was performed either in human blood-based samples (serum, plasma or peripheral blood mononuclear cells), human pancreatic islet cells and adipose tissue or in obese rodent liver (He et al., 2017; Li et al., 2009; Trajkovski et al., 2011; Yang et al., 2016). This thesis describes the results of the first systematic approach which identifies aberrant hepatic miRNA expression in humans independent of liver steatosis and obesity leading towards T2D. Therefore hepatic miRNA expression from 40 samples was measured by microarray and these results were correlated with metabolic traits (blood glucose, serum insulin, serum triglycerides) or the incidence of T2D. The candidate miRNA hsa-miR-182-5p was identified by several regression models and carefully characterized. Possible target genes were identified by computational methods and validated by cell culture experiments. Hepatic expression of these candidate target genes, *FOXO1*, *SCD*, *CDKN1B* and *LRP6* was further analyzed. Moreover, expression of the hepato-specific miRNA hsa-miR-122-5p was analyzed regarding their contribution to the manifestation of T2D or NAFLD. This miRNA also potentially targets the 3'UTR of *ACACB* and *INSR*, which are reduced expressed in diseases liver. In a gene-specific approach, expression of hsa-miR-223-3p was measured and analyzed regarding a possible repression of *GALNT18* and *FOXO1*, which are two potential target genes with recognition motifs for miRNA binding within their 3'UTR.

These regulatory mechanism can also act in synergy to target specific genes. We were able to show that a combination of reduced DNA methylation which correlates with an adjacent genotype and overexpression of hepatic hsa-let-7e-5p is responsible for *IRS2* dysregulation. This might contribute to the phenotype of selective insulin resistance in humans.

Lastly, the identification of underlying molecular mechanisms and key molecules in T2D represent possible new target genes or pathways for drug development. Nevertheless, prevention of disease manifestation

stays the most effective therapy against the increasing prevalence. Therefore a computational attempt by using machine learning models was performed to translate our results into a clinical application. These models include the generation of a predictive risk scoring system which is either based on linear methylation risk scores emerged from EWAS markers or which is based on a non-linear model and robust machine learning principles for the complex network of *IRS2* regulation.

## 6.1 Thyroid hormone signaling in NAFLD and T2D

The action of intrahepatic TH signaling on hepatic cholesterol metabolism is not well characterized in clinical human samples. Expression measurement of several genes encoding proteins of hepatic TH signaling revealed a significant inverse correlation between hepatic *THRB* expression and NAS, HbA1c and serum triglycerides. In addition, expression of two key enzymes of hepatic cholesterol clearance, *LDLR* and *CYP7A1*, was not significantly altered within this cohort. This is the first time to prove decreased *THRB* expression in human liver samples as main driver of hepatic TH resistance in NAFLD. TH signaling itself is not directly associated to hepatic or systemic insulin resistance in humans (Bos et al., 2017). Nevertheless, due to TH related control of cholesterol metabolism, there is a strong link between TH signaling and NAFLD, and therefore also between TH and T2D (Pagadala et al., 2012; Portillo-Sanchez et al., 2015). Hepatic TH signaling also regulates serum lipid clearance by induction of hepatic LDLR and subsequent uptake of cholesterol. Afterwards, this cholesterol is converted to bile acids by TH induced increased expression of *CYP7A1*, a key enzyme in bile acid synthesis (Ness, 1991; Ness and Lopez, 1995). Nevertheless, blood lipids (total cholesterol, LDL and HDL) were not different between T2D and ND subjects, therefore the observed missing alteration of *LDLR* and *CYP7A1* expression in T2D or NAFLD is an expected result for this cohort. Another mechanism of TH signaling was proposed as cause for the altered lipid metabolism and different stages of liver steatosis and fibrosis. All levels of hepatic TH action (transporter mediated uptake, conversion of TH, transcription factors) were analyzed for a complete characterization.

*SLCO1C1* (also *OATP1C1*) encodes a transporter protein (Solute carrier organic anion transporter family member 1C1 or also organic anion transporting polypeptide 1C1) which mediates TH flux across the blood-brain barrier (Roberts et al., 2008). Therefore no hepatic gene expression is detectable.

*DIO2* encodes a protein called Iodothyronine Deiodinase 2 which catalyzes the conversion of the precursor hormone T4 to the bioactive T3. Though *DIO2* is ubiquitary expressed, mRNA levels are more prominent in different regions of the brain (Guadaño-Ferraz et al., 1997). *DIO3* encodes a third type of deiodinases, which inactivates T3 and T4 in fetal tissues (Richard et al., 1998; Peeters et al., 2003). This tissue and time dependency might explain a missing traceability in adult liver.

Other studies observed by immunohistochemistry a decrease of DIO1 and increase of DIO3 in NAFLD liver biopsies which results in a higher hepatic TH inactivation (Bohinc et al., 2014). As mentioned, DIO3 was not measurable in the liver cohort of this thesis. On the contrary, *DIO1* was not significantly elevated in liver of T2D subjects. DIO1 catalysis the activation of pro hormone and inactivation of T3 as previously described. A higher conversion of T4 to the bioactive T3 estimated by higher serum T3/T4 ratios was also reported in the context of NAFLD (Bilgin and Pirgon, 2014). Hepatic *DIO1* expression neither correlated significantly with HbA1c, TSH nor NAS. Also an age effect is diminishing with increasing sample size. Nevertheless, a higher conversion rate was reported in obese children and a smaller sample size already showed age-related repression of *DIO1*. Furthermore, studies in old rodents support this aging effect (Silvestri et al., 2008; Visser et al., 2016).

*THRA* and *THRB* are nuclear receptors whereby *THRB* is the predominant hepatic isofrom (Vatner et al., 2013; Sinha et al., 2018). THR$\beta$ is a TH activated transcription factor which co-localizes with repressors on TH responsive elements on the DNA (Apriletti et al., 1998). Upon binding of T3, THR$\beta$ co-localizes with transcriptional activators and facilitates transcription of respective target genes. Targets

of TH induced gene expression are involved in *de novo* lipogenesis, which synthesize free fatty acids (FFA) from glucose for vLDL synthesis (Sinha et al., 2018). TH triggers additional adipose lypolysis which releases FFA to fuel hepatic lipid synthesis and storage. Though *de novo* lipogenesis is activated, hepatic triglyceride content decreases upon TH signaling due to increased lipid metabolism for energy gain (Oppenheimer et al., 1991). This coherence between THR$\beta$ activation and reduction of hepatic lipid accumulation was the initial motivation to treat NAFLD patients with THR$\beta$ agonists. After treatment with an liver-specific agonist, hepatic steatosis was decreased at the expense of hepatic insulin sensitivity (Vatner et al., 2013). The observed hepatic repression of *THRB* in this liver cohort is in accordance to a reduced fatty acid metabolism and enhanced hepatic triglyceride content. Since there was only a minor or no significant difference in gene expression between T2D and ND subjects, a negative correlation between *THRB* expression and HbA1c might emerge as comorbidity from altered hepatic lipid content due to NAFLD, which explains the discrepancy between the effects of THRB agonists on the impairment hepatic insulin resistance.
Whether reduced transcription factor expression mirrors also a reduction of activated protein and downstream changes in key genes of lipogenesis and fatty acid metabolism needs to be elucidated. Since no hypo- or hyperthyroidism was observed for any subject, decreased T3 action might indicate a hepatic resistance to TH signaling.

## 6.2 Alterations of hepatic DNA methylation for pre selected genes

Recent studies on hepatic DNA methylation changes in metabolic diseases revealed alteration of methylation and gene expression of two candiate genes, *Tumor Protein P53 Inducible Nuclear Protein 1* (*TP53INP1*) and *Polypeptide N-Acetylgalactosaminyltransferase 18* (*GALNT18*). Promoter methylation measurement of *TP53INP1* was chosen due to previously observed changes of two promoter CpG sites in T2D (CpG7 cg13393036 and CpG8 cg20039814), which could not be reproduced in this liver cohort. Intragenic DNA methylation was exploratory analyzed by measurement of three CpG sites within *GALNT18* intron 3. Changes in *GALNT18* methylation and expression were previously associated to different stages of NAFLD and NAS. This is the first time to prove such methylation alterations of three CpG sites within intron 3 of *GALNT18* age- and gender dependent for T2D. These changes also mirror the overall liver health status by inverse correlation with liver enzymes AST and ALT. Moreover, *GALNT18* expression was shown to be inversely altered for HbA1c levels in an age- and BMI dependent manner in our cohort. Subsequent functional analysis of DNA methylation on gene expression revealed a possible contribution of a forth intragenic CpG site and three potential transcription factor binding sites (Sp1, SREBF1 and CEBPA). These results emphasize *GALNT18* as interesting new target gene for T2D manifestation.
Aberrant promoter methylation is an unlikely cause for dysregulation of genes involved in metabolism and a more common observation in cancer. For an exploratory approach, nine CpG sites within the promoter of *TP53INP1* were chosen due to previous results on single sites within this sequence. Additionally, *TP53INP1* was 1.2 fold upregulated in a comparable cohort of liver samples of T2D obese subjects in comparison to glucose tolerant lean subjects (Kirchner et al., 2016). Due to the low sample size of 11 to 13 subjects per group, these results were only exploratory hints for reproduction by missing significance. Nevertheless, *TP53INP1* is an interesting candidate gene based on the discovery of SNPs within the gene associated to T2D (Voight et al., 2010; Xue et al., 2018). One of those polymorphisms (rs10097617, MAF = 0.51, C>T) might also have a functional epigenetic background by influencing promoter methylation at cg13393036 (Xue et al., 2018). Therefore it was proposed that CGI methylation is actually enhancing *TP53INP1* transcription by inhibiting of repressor protein binding to the promoter.
This discrepancy might originate from the difference in cohort design and sample size. The association found in literature originates from more than 650.000 individuals (9 % T2D subjects) and all associations

were calculated by summary-data-based Mendelian randomization analysis without validation by site specific laboratory techniques. Nevertheless, the large sample size allows the identification of variants with small impact on the incidence of T2D. The usability of large scale hits in the context of personalized medicine will be discussed further.

Though the association between T2D and promoter methylation was not reproducible, this gene remains an interesting metabolic target gene with implications in obesity and resulting low-grade inflammation (Seillier et al., 2015). Knock-out mice lacking global *Tp53inp1* expression were more prone to high-fat diet (HFD) induced obesity and insulin resistance based on chronic oxidative stress and a dysfuction in mitophagy. Interestingly, this paper also found an increase of hepatic *TP53INP1* expression in steatotic human liver which was highly discussed upon the review progress to fit the mice model which indicates a decrease of *TP53INP1* expression as pathogenic driver. One explanation might be a protective response to fat accumulation in obesity. More functional studies without being influenced by obesity would be needed to elucidate the emerging discrepancy between mouse model and human expression data.

*GALNT18* was previously described as novel metabolic target gene in liver steatosis in a large North German cohort (Ahrens et al., 2013). In their study, including 45 subjects with different stages of metabolic liver diseases (NAFLD and NASH) and 18 lean control subjects, *GALNT18* expression was observed to be increased in obese subjects, whereby methylation of cg16337763 (also located within intron 3) was decreased. Subjects with a NASH phenotype had even higher expression and lower methylation in comparison to obese subjects with a healthy liver (Ahrens et al., 2013).

It was possible to reproduce a positive correlation between BMI and *GALNT18* gene expression in our cohort. Comparing obese T2D and obese ND subjects, gene expression was not altered. Also a negative correlation between gene expression and HbA1c levels was age driven. The lower expression in T2D subjects can be explained by the positive relationship between age and HbA1c within this cohort. In contrast, DNA methylation within intron 1 of *GALNT18* differed ($> 4$ %) between T2D and ND subjects prior adjustment for age and correlated significantly with status of NAFLD. DNA methylation at all three analyzed CpG sites was also influenced by age. A lack of influence of DNA methylation at these specific positions on gene expression was validated in Luciferase reporter gene assay.

The first CpG was predicted to be located within an E box motif for SREBF1 binding. A drastic effect of methylation by decreasing the Luciferase signal was observed after elongation of the analyzed sequence. This new designed sequence setting contained several low consensus binding motifs for the SREBF1 cofactor Sp1 and an additional CpG site within a high consensus Sp1 motif (Reed et al., 2008). Therefore inclusion of this CpG site (CpG-1) downstream of CpG1 could be responsible for this effect. Unfortunately it is not possible to design a suitable bisulfite assay to analyze the hepatic *in vivo* situation due to the isolated position of CpG(-1) within diverse T-homo-polymers.

*In silico* analysis raveled an additional binding motif for CCAAT/enhancer-binding protein alpha 8-16 nt downstream of the first measured CpG site (CEBPA, 5'-ATTGCAAAA, JASPAR sequence logo MA0102.3). This transcription factor is known to be involved into activation gluconeogenic genes by cooperation with FOXO1 and is predicted to interact with SREBF1 (Sekine et al., 2007; Bridges et al., 2014). A hepatic *Cebpa* knock-out mouse was shown to have decreased hepatic lipogenesis caused by decreased expression of *Fasn* and *Scd* (Matsusue et al., 2004). Whether intergenic binding of this transcriptional enhancer also regulate *GALNT18* and contributes to hepatic lipogenesis by activation of *GALNT18* needs to be elucidated. Interestingly, hypomethylation of respective recognition motifs is associated to transcriptional activation by CEBPA (Rishi et al., 2010). This is in accordance our results where hypomethylation is observed for lower expressed *GALNT18* in type 2 diabetic liver.

The metabolic function of *GALNT18* itself needs to be elucidated. The protein is a member of the UDP-Nacetyl-$\alpha$-D-galactosamine polypeptide N-acetylgalactosaminyltransferases Y-subfamily and expressed in various tissues (Raman et al., 2012). Other family members are responsible for the synthesis of mucin-type O-glycans. GALNT18 itself appears to have a catalytic and no transferase effect (Li et al., 2012).

Due to the high similarity between murine and human protein (99 % of shared amino acids stated from the Cancer Genome Anatomy Project from the National Cancer Institute, National Institute of Health), the metabolic function of GALNT18 could be explored by generation and careful characterization of respective knock-out mice.

## 6.3 Evaluation of blood-based EWAS markers in metabolically active tissue

Blood methylation at *SREBF1* cg11024682 and *ABCG1* cg06500161 are highly associated to metabolic traits, still a careful evaluation whether there might be a mechanistically involvement in the manifestation of hepatic insulin resistant is still missing. To elucidate a hepatic function, methylation at both markers was measured in our cohort with a carefully validated bisulfite pyrosequencing assay (Krause et al., 2019). Though *SREBF1* cg11024682 correlated weakly with BMI, no difference between T2D and ND subjects was observed. Methylation at *ABCG1* cg06500161 was not altered between T2D and ND subjects, nor was the polymorphism rs99822016 associated to T2D or influencing hepatic methylation. Methylation at both markers did not correlate with any serum parameter.

Epigenome wide association studies gain in importance to elucidate epigenetic mechanism in disease manifestation for T2D (Volkmar et al., 2012; Hidalgo et al., 2014; Nilsson et al., 2014; Chambers et al., 2015; Kulkarni et al., 2015; Nilsson et al., 2015; Kirchner et al., 2016; Kriebel et al., 2016; Wahl et al., 2016; Sala et al., 2017; Volkov et al., 2017; Walaszczyk et al., 2018). These studies are also a common tool for a diverse set of complex, non-Mendelian inherited diseases, like obesity, inflammation, psychiatric or cardiovascular diseases (Demerath et al., 2015; Liu et al., 2013b; Mill et al., 2008; Liu et al., 2018; Nakatochi et al., 2017). One advantage is the easy extraction and stability of DNA methylation, therefore it is more suitable for high throughout screening than miRNA exosome isolation and expression measurement. Nevertheless, most studies use blood related DNA as input, though DNA methylation is highly tissue dependent (Kitamura et al., 2007; Jones, 2012; Bernstein et al., 2007; Rakyan et al., 2011). *SREBF1* cg11024682 and *ABCG1* cg06500161 are two recurring methylation sites, which emerged from different blood based EWAS to elucidate possible markers for T2D and obesity. Besides the lack of mechanistic studies, the suitability of single markers emerged from large-scale studies to estimate the risk of T2D in the frame of personalized medicine is questionable due to their mostly low impact or effect size. Additionally, whether or not a correlation between methylation of leukocyte DNA and DNA from metabolically active tissue is significant or in the same direction depends strongly on the analyzed population and sample size (Chambers et al., 2015; Wahl et al., 2016; Pfeiffer et al., 2015; Dayeh et al., 2016; Krause et al., 2019).

Neither gene expression nor any mechanistic experiments were performed since no significant associations regarding marker methylation were detected. However, both genes remain interesting metabolic targets in liver.

SREBF1 is a known insulin-responsive key transcription factor in hepatic lipid metabolism (Kim et al., 1998; Harris et al., 2007; Xu et al., 2013). Though gene expression is activated by insulin, it is not known whether a proteolytic cleavage of the precursor protein is also regulated by an insulin-dependent manner. Nevertheless, a cholesterol dependent mechanism for processing of SREBF1 in an inhibitory feedback loop was already shown (Goldstein et al., 2002).

Interestingly, *SREBF1c* expression is significantly decreased in adipose tissue of obese human subjects in comparison to lean subjects (Oberkofler et al., 2002). Adipose tissue *SREBF1* mRNA increased again calorie intake dependent in subjects which underwent bariatric surgery and experienced thereby weight loss (Oberkofler et al., 2002). SREBF1c proved to be the main isoform of cholesterol and fatty acid biosynthesis in adipose tissue and liver (Shimomura et al., 1997; Xu et al., 2013). Especially the reversible diet and obesity dependent expression pattern of *SREBF1* is a strong indicator for an epigenetic mechanism for insulin responsive expression activation or repression.

Besides DNA methylation, also SREBF1 mRNA degradation by miR-33 might be a potential mechanism which was shown in mice (Horie et al., 2013). Knock-out of this miRNA causes enhanced *SREBF1* expression which results into liver steatosis and obesity which is in contrast to the human situation in adipose tissue. Unfortunately miR-33 was not expressed in our human liver cohort and other studies in hepatocarcinoma cell lines were unable to validate these mouse data (Goedeke et al., 2013). Interestingly, hsa-miR-224-5 is reported to regulate *SREBF1* expression (Kameswaran et al., 2014). This miRNA was associated to NAS in our miRNA microarray discovery cohort and was more than 1.5 fold upregulated in subjects with T2D though it was expressed only slightly above background signal.
Another mouse study revealed a more prominent effect of solely hepatic *SREBF1* overexpression on systemic *de novo* lipogenesis by increasing also visceral adipose tissue mass independent of the diet (Knebel et al., 2012). Therefore hepatic *SREBF1* expression appears to have a higher weight on obesity than the expression in adipose tissue itself.
The transmembrane transporter ABCG1 is a member of the ATB binding casette (ABC) superfamily which regulates the efflux of cellular cholesterol to HDL in hepatocytes and macrophages (Kennedy et al., 2005). At the same time, this transporter regulates and coordinates intracellular sterol localization (Tarling and Edwards, 2011). Experiments in mice revealed a function in adipocyte triglyceride storage and a specific inhibition of adipocyte Abcg1 reversed weight gain (Frisdal et al., 2015). Interestingly, the genotype of two SNPs (rs1378577 and rs1893590 in linkage disequilibrium) within the *ABCG1* promoter is influencing adipose tissue *ABCG1* expression in obese subjects. The polymorphism rs1378577 changes an AG genotype into a CG genotype whereby the AG genotype was associated to higher *ABCG1* expression and increased BMI (Frisdal et al., 2015). DNA methylation was not analyzed at this position.
Alterations in leukocyte methylation at cg06500161 was previously associated to serum HDL or triglycerides (Pfeiffer et al., 2015; Dayeh et al., 2016). A respective EMSA for this position revealed a stronger binding affinity to unmethylated DNA for unspecified proteins of a nuclear extract acquired from a monocyte (THP1) cell line (Pfeiffer et al., 2015). There is no reported correlation between adipose tissue or liver methylation and respective blood methylation for *ABCG1* cg06500161. Nevertheless, significantly elevated DNA methylation at this position is contradictory reported in adipose tissue of T2D subjects (Dayeh et al., 2016; Krause et al., 2019).
EWAS still represent a good approach to find novel targets to predict T2D. Blood based markers might not have a metabolically involvement but might sense changes in blood glucose and insulin levels. Respective EWAS in liver are needed to elucidate the mechanisms of hepatic insulin resistance due to the tissue specificity of DNA methylation. The accessibility of liver tissue remains a bottleneck to perform large scale hepatic studies in humans. Furthermore, large scale studies often discover markers with low effect sizes but small p values. Therefore a careful evaluation of these CpG sites by good quality bisulfite pyrosequencing assays and other functional assays (for example Luciferase assays whether there is a promoter regulatory effect or EMSA to detect protein binding) are needed.

## 6.4 Liver glucose and lipid metabolism is altered by aberrant hepatic miRNA expression

Three approaches were chosen to find metabolically involved hepatic miRNA: The first approach was a systematic attempt by high throughput microarray analysis with a smaller age, gender and BMI matched discovery cohort of n = 20 T2D and n= 20 ND subjects. Afterwards, the emerged candidate miRNA hsa-miR-182-5p was validated in the full cohort and its contribution to insulin resistance was extensively characterized. *In silico* analysis and cell culture experiments revealed *FOXO1* and *LRP6* as plausible target genes. This leads to the conclusion that hepatic overexpression of hsa-miR-182-5p reduced gluconeogenesis but fuels hepatic steatosis by repression of *LRP6* and sustained activation of the non-canonical Wnt signaling pathway. The second attempt was a gene specific approach, whereby the miRNA

hsa-miR-223-3p was chosen by *in silico* analysis of the 3'-UTR of *GALNT18* and *FOXO1* for further characterization. In the last attempt, the liver specific hsa-miR-122-5p was characterized in the full cohort. This choice is reasoned by its known association to hepatic lipid metabolism, its potential to act as serum biomarker for T2D and its ability to target the 3'-UTR of two genes, *ACACB* and *INSR*, which were shown to have diminished mRNA in T2D or IR. These approaches and results will be discussed further in the next three subsections.

Due to technical issues and a discrepancy in literature, the choice of a hepatic housekeeping miRNA for expression normalization needs further discussion. Expression of established housekeeping genes was evaluated by micorarray and three candiate miRNA were validated by subsequent TaqMan qPCR. Hsa-miR-24-3p represents the most stable expressed hepatic miRNA in NAFLD and T2D. Most publications use ribosomal or other non-coding RNA for expression normalization. This is not compatible with TaqMan based miRNA-qPCR due to a specific miRNA-cDNA synthesis system. Nevertheless, the suitability of non-coding RNA, like RNU6, is argued in literature based on different chemical properties between ribosomal and micro RNA (Benz et al., 2013; Chugh and Dittmer, 2012). These non-coding RNA species might differ in extraction efficiency, cDNA synthesis and PCR amplification from miRNAs (Schwarzenbach et al., 2015). Also most potential housekeeping miRNA have a metabolic implication (listed in Table 23) but none was reproducible in several cohorts or studies. The identification of a stably expressed miRNA is mandatory for correct data normalization. The measurements of this cohort represent the first systematic analysis for the identification of hepatic housekeeping miRNAs.

### 6.4.1 Systematic approach by microarray analysis

Stratification only for T2D was not sufficient to identify adequate candidate miRNA which contribute to disease manifestation. Though subjects of the discovery cohort were carefully chosen and matched for diverse factors, a large variability within the results was observed in the PCA clustering of the microarray signals. This variability was further confirmed by a lack of significance after correction for multiple testing. T2D is a multi-factorial disease, therefore subgroup stratification was completed by addition of linear regression analysis using different clinical parameters as response which are altered in T2D (serum triglycerides, serum insulin, blood glucose, HbA1c). Preliminary exclusion criteria represent significant associations to BMI or age. One candidate miRNA, hsa-miR-182-5p, was associated to the presence of T2D, but also to changes of its hallmark markers blood glucose and HbA1c in the smaller microarray discovery cohort. Though other miRNA also fit into this scheme, hsa-miR-182-5p remained the most promising candidate due to a considerable overexpression in T2D and diverse potential target genes in metabolically relevant pathways which dysregulation might contribute to hepatic insulin resistance.

Expression measurement in our full cohort revealed a more than 1.5 fold overexpression of hsa-miR-182-5p in diabetic liver independent of NAS score. Expression correlated also with HbA1c and blood glucose levels, which was already observed in our discovery cohort. Besides, miRNA expression correlated also positively NAS, serum insulin, serum triglycerides and liver enzymes AST and ALT.

Another study of human liver biopsies from 30 subjects with liver steatosis and different fibrosis states revealed an upregulation of miR-182-5p upon increase in fibrosis (Leti et al., 2015). The cohort was matched for confounding factors and T2D, therefore a contribution of insulin resistance was not analyzed. The high abundance of hsa-miR-182-5p for inceasing NAS is reproducible in this cohort. Thereby the functionality of this miRNA in hepatic insulin resistance and the significant correlations between expression and alterations of related traits (HbA1c, serum triglycerides and glucose) was proven for the first time. These observations in the human situation were further supported by studies from rodent liver. Mice and rats with fatty liver showed in two independent studies increased hepatic miR-182 expression (Dolganiuc et al., 2009; Nie et al., 2018).

Hsa-miR-182-5p is transcribed as a highly conserved polycistronic transcript including miR-183/96/182

(Dambal et al., 2015). Besides small changes in pre-miRNA sequence, cluster structure and seed sequence of the mature miRNA is conserved between human and mouse. Studies in murine liver indicate a transcriptional regulation by SREBB2 which might be further conserved (Jeon et al., 2013). Interestingly in our hepatic cohort, only enhanced abundance of hsa-miR-182-5p was associated to either T2D or any metabolic trait and other cluster members were not expressed at all.

Overexpression of hsa-miR-182-5p is also observed in different types of cancer, like gliomal tumorgenesis, myeloid leukemia or oral squamous cell carcinoma (Xue et al., 2016; Zhang et al., 2018a; Li et al., 2018). In a hepatic context, $\beta$-catenin induced overexpression of the miR-183/96/182 cluster results into an increased severity of HCC by reducing *FOXO1* expression (Leung et al., 2015).

*In silico* analysis of the 3'UTR of metabolic key genes and subsequent cell culture experiments revealed *LRP6*, *SCD*, *FOXO1* and *CDKN1B* as significantly repressed target genes after overexpression of hsa-miR-182-5p. Therefore expression of these genes was analyzed in our complete cohort.

Expression of *LRP6* correlated significantly with age but was independent of BMI or gender. After considering the age dependency, decreased expression remained significantly associated to elevated HbA1c, serum insulin, NAS and levels of the liver enzyme AST. In addition, gene expression correlated noticeable with hsa-miR-182-5p expression. Screening of the 3'UTR revealed two possible binding sites which are recognized in different modes (a 7mer_A1 and a 6mer). To elucidate which recognition site is preferred in miRNA binding, respective Luciferase Assays and motif mutations have to be performed (Jin et al., 2013; Clément et al., 2015).

LRP6 is a transmembrane protein which co localized on the cellular membrane with a receptor called Frizzled (FZD) and fulfills essential functions in hepatic Wnt/$\beta$-catenin signaling (Tamai et al., 2000; Thompson and Monga, 2007). Loss of LRP6 leads towards the activation of the non-canonical Wnt signaling pathway and enhanced hepatic *de novo* lipogenesis (Go, 2015). Several loss-of-function mutations are known within *LRP6*, which cause hallmark phenotypes of metabolic syndrome, like T2D and increased blood glucose, hypertension, coronary artery diseases, NAFLD or pathologically altered blood lipids (Wang et al., 2018). Interestingly, even in mutation carriers an age effect was observed (Mani et al., 2007). Repression of *LRP6* by miRNA binding and reduced amount of protein might act equal to reduced amount of functional protein.

Initially *SCD* was thought to be repressed by hsa-miR-182-5p overexpression which was here validated in cell culture experiments. The *in vivo* situation is not mirroring this relationship. There still might be a regulatory effect of this miRNA on SCD by stalling the ribosome and inhibiting protein translation instead of mRNA degradation, which could partly explain the paradox between mRNA and protein content (Gu and Kay, 2010). This mode of action was already observed for the miRNA let-7a (Nottrott et al., 2006). A concrete validation of this assumption is not possible due to the limited access to liver tissue for protein extraction. Respective studies have to be performed in murine liver after verification of sequence homology between human and mouse. Since SCD is a key enzyme of hepatic *de novo* lipogenesis, which is activated by the non-canonical Wnt signaling pathway, our observation of increased hepatic expression might be a direct consequence of *LRP6* repression. In conclusion, low *LRP6* mRNA appears to have a larger influence on *SCD* expression and enhanced *de novo* lipogenesis than repression by miRNA binding.

*FOXO1* was identified as possible target gene of hsa-miR-182-5p, which was strongly indicated by prior literature research. FoxO1 is member of the Forkhead transcription factor subfamily which is involved in diverse pathways of metabolism, cellular proliferation, survival and differentiation (Accili and Arden, 2004). Hepatic FoxO1 appears to have a dual function: Enhanced protein abundance improves insulin sensitivity via stimulation of Akt phosphorylation but causes decreased fatty acid oxidation leading to enhanced liver steatosis (Matsumoto et al., 2006). This is in contrast to expression studies performed in liver specific $Irs1^{-/-}$/ $Irs2^{-/-}$ double knock out mice or hepatic $Insr^{-/-}$ knock out mice which resemble a simple murine model of T2D. An additional deletion of hepatic *FOXO1* expression resulted into markedly improvement of blood glucose and insulin and even reversion to healthy levels in those animals (Dong

et al., 2008; O-Sullivan et al., 2015). This amelioration might be based on the missing activation of enzymes involved in gluconeogenesis by FoxO1 and subsequent reduction of hepatic glucose production. In our human liver cohort, reduced *FOXO1* mRNA is associated with elevated HbA1c and blood glucose. Diet induced obese mice without mentioned knock out of key proteins of hepatic insulin signaling exhibit a similar phenotype like this human cohort upon *FOXO1* deletion (Zhang et al., 2012). The discrepancy between murine T2D models and human T2D subjects after a reduction of *FOXO1* expression might originate from other developmental compensations after *Irs1* or *Insr* deletion. Indeed, *FOXO1* gene expression is dependent of IRS2 mediated insulin stimulation. A decrease of *IRS2* mRNA was also observed in T2D which will be discussed later. Besides the missing activation of *FOXO1* transcription, hsa-miR-182-5p binding might additionally act as a protective mechanism against sustained hyperglycemia which is a consequence of continuously activated glyconeogenesis. Interestingly, a decrease of *FOXO1* expression results into elevated cholesterol synthesis which boosts a decline in liver steatosis. These observations are in accordance to previous reports on a relationship between FoxO1 and NAFLD (Dong, 2017).
FoxO1 is indirectly involved in cell cycle regulation by enhancing *CDKN1B* expression (van der Horst and Burgering, 2007). An observed repression of HepG2 *CDKN1B* by miRNA overexpression might be the cause of enhanced *FOXO1* mRNA degradation.
*CDKN1B* itself encodes a protein called p27, which is an inhibitor of the cyclin D1/CDK4 complex which in turn regulates cell cycle progression (Ray et al., 2009). It has been observed that this cyclin D1/CDK4 complex is also involved in in the regulation of glucose metabolism, hence the regulation of its suppression represents a potential pathway for the development of hyperglycemia (Lee et al., 2014). Cyclin D1/CDK4 is activated upon insulin stimulation to suppress hepatic glucose production. The loss of hepatic insulin sensitivity would cause a loss of activation leading to sustained hepatic glucose production and hyperglycemia. Diminished p27 caused by decreased *CDKN1B* expression is rather a consequence of IR since repression is not mandatory. Further prove of the hypothesis that diminished *CDKN1B* expression is not directly linked to disease manifestation is given by a lack of association to any metabolic trait or a missing correlation between gene and miR-182-5p expression. A dysregulation of this signaling axis would still make sense in the context of sustained hyperglycemia and IR leading to T2D.
Expression of another component, *CCND1* encoding cyclin D1, was also measured in this thesis. Cyclin D1 is not only involved in this cell cycle regulation but also in repression of diverse key enzymes initiating lipogenic and gluconeogenic pathways (Núñez et al., 2017). Unfortunately, hepatic *CCND1* expression was only associated to a gender and age related effect and did not correlate with any metabolic trait.
In summary, hsa-miR-182-5p was identified as candidate miRNA involved in metabolic dysregulation by multiple associations to T2D, aberrant HbA1c and blood glucose. These associations besides others (changes in serum insulin, serum triglycerides, HDL, AST, ALT and NAS) could be reproduced in a larger cohort setup strengthen the theory of a metabolic implication. Two *in silico* and *in vitro* identified target genes *LRP6* and *FOXO1* remain plausible. Repression of *FOXO1* counteracts hepatic glucose production at the expense of enhanced de novo lipogenesis by activation of the non-canonical Wnt pathway after T2D manifestation. This assumption is further supported by a lack of acute expression induction after insulin treatment and the observation in rodent liver where silencing of hsa-miR-182-5p after treatment with an absorption inhibitor for cholesterol resulted in loss of glucose control (Sedgeman et al., 2018). Considering this miRNA as a consequence of T2D manifestation, its application as a biomarker to predict disease onset is not suitable.
Especially the regulation of transcription factors is not only driven by alterations in gene expression. Therefore cell culture experiments after hsa-miR-182-5p overexpression regarding changes in protein abundance and phosphorylation status across the insulin signaling axis are further required.

### 6.4.2 Target specific miRNA selection

The expression of hepatic hsa-miR-223-3p correlates significantly with serum triglycerides and HDL after adjustment for age, gender an BMI within the cohort. In literature, expression and cellular functions of the miRNA hsa-miR-223-3p is mostly associated to hematopoietic cells (Johnnidis et al., 2008). Nevertheless, a study of 83 human subjects revealed that serum hsa-miR-223-3p is decreased in overweight and obese subjects (Kilic et al., 2015). Furthermore, this miRNA might be a key regulator of adipocyte differentiation and inflammation (Sliwinska et al., 2017). Targeting *FOXO1* was described as one of the possible mechanisms behind these associations. Hepatic miRNA expression correlated neither with hepatic *FOXO1* nor hepatic *GALNT18* expression.
Interestingly, this miRNA is also reported to inhibit hepatic HDL uptake which was shown in a hepatocellular model (Wang et al., 2013a). This is in accordance to a negative correlation between hepatic miRNA expression and serum HDL concentrations which was observed within the cohort of this thesis. A reduction of hepatic scavenger receptor class B type I (*SR-BI*) mRNA was indicated, which was not further analyzed in the context of this thesis. It would be interesting whether these *in vitro* experiments actually mirror the hepatic *in vivo* situation. Nevertheless, there is a close relationship between systemic insulin resistance and lowering of blood HDL concentrations (Fossati and Romon-Rousseaux, 1987; Rashid et al., 2003). There might be an indirect and not obvious connection (for example a time dependency) between hepatic miRNA expression and T2D manifestation. This hypothesis is supported by another study of 160 subjects which revealed an association of decreased plasma miR-223-3p concentrations on the early onset of T2D manifestation (Zampetaki et al., 2010).
*In vitro* experiments with Oleate-Palmitate treated HepG2 cells, which resembles a cellular model of NAFLD, revealed a desregulation of cellular hsa-miR-223-3p expression upon induction of cellular steatosis (Di Mauro et al., 2016). This relation could not be reproduced in the *in vivo* situation.
In summary, a key function of hsa-miR-223-3p in regulating hepatic cholesterol metabolism is implied by these results though this might be independent of NAFLD progression and not directly dependent of T2D. This regulatory mechanism is also independent of a reduction in hepatic *GALNT18* and *FOXO1* mRNA. The positive correlation between hepatic miRNA expression and age implies also a regulatory mechanism of the aging liver.

### 6.4.3 Liver specific miRNA selection

Expression of the liver-specific hsa-miR-122-5p correlates negatively with NAS, fasting glucose, HbA1c, AST and ALT after adjustment for age, gender and BMI within this cohort. This miRNA is highly conserved across several species and accounts for more than 70 % of all hepatic miRNAs (Jopling, 2012; Filipowicz and Grosshans, 2011). Though this miRNA is mainly associated with HCV infection and disease progression, its physiological function is established in hepatic fatty acid and cholesterol synthesis (Krützfeldt et al., 2005; Esau et al., 2006; Jopling et al., 2008; Luna et al., 2015).
Emerged from human studies regarding chronic hepatitis C, which is a part of disease progression during HCV infection, serum levels of miR-122-5p were increased in subjects with NAFLD (n = 34) in comparison to healthy subjects. Serum miRNA concentration also correlated positively with liver enzymes, lipids and stage of fibrosis and inflammation (Cermelli et al., 2011). This is in accordance to studies on murine liver, where a specific inhibition of this miRNA by antisense oligos results into an improvement of liver steatosis in diet-induced obese mice (Esau et al., 2006). Furthermore, this *in vivo* study also revealed a reduction of plasma cholesterol level as well a reduction of hepatic fatty acid and cholesterol synthesis upon inhibition. Since *ACACB* also synthesises malonyl-CoA, which is a rate limiting step in fatty acid synthesis, a connection between miR-122-5p and *ACACB* expression would represent an additional plausible mechanism for these observations (Wakil and Abu-Elheiga, 2009). Nevertheless, *ACACB* is known to rather regulate fatty acid oxidation than synthesis (Abu-Elheiga, 2001). It is even

more controversial when considering the results of this thesis that hepatic insulin resistance and NAFLD is reported to be associated to enhanced *ACACB* expression, as shown for a diet induced NAFLD rat model (Savage, 2006).
Interestingly, total cholesterol does not correlate with alterations in hepatic miR-122-5p expression for this thesis. This might be due to the obesity matched cohort and less variance within the blood lipids. Instead, hepatic miRNA expression correlates negatively with blood glucose and HbA1c, which indicates reduced miRNA expression contributes to a hepatic phenotype of T2D. Interestingly, hepatic miRNA expression correlates also negatively with NAS, which is again in contrast to literature results from human serum concentrations or studies in mice. A different study of human liver biopsies (n = 67) from various stages of NASH revealed that serum miRNA concentrations might mirror the hepatic situation (Miyaaki et al., 2014). Also an enhancement of expression upon severity of fibrosis and steatosis is reported in this mentioned study. An interesting difference between the described human cohort and the liver cohort of this thesis remains the overweight (BMI of 28.5 +/- 4.2) and not obese (50.5 +/- 9.1 for T2D and 53.3 +/- 11.1 for ND subjects) BMI as well the missing stratification for T2D. Therefore this miRNA might have a general enhanced expression in obesity but is still altered in T2D within obese subjects.
In a prospective approach, increased serum miR-122-5p was able to predict metabolic syndrome and T2D in a large human cohort (n = 810, 15 years follow-up) (Willeit et al., 2017). The BMI cohort was again overweight (26 +/- 3.9) and not obese. A correlation between BMI and circulating miRNA was detected, which was neither found in the previously mentioned nor hepatic cohort of this thesis.
All in all, our results are in contrast to previously published associations of hepatic and serum miR-122-5p in humans which is elevated in T2D. The differences might emerge from the T2D focused cohort design including only obese subjects of this thesis. Hepatic miRNA repression is consistently associated with increasing glucose markers (blood glucose and HbA1c) and overall liver fibrosis and steatosis health status (NAS, AST and ALT). Correlation analysis did not reveal any potential repression of *INSR* or *ACACB* mRNA by miR-122-5p. In fact, both genes were also repressed in T2D or IR, therefore this result was expected.

## 6.5 Synergistic regulation of insulin signaling pathways by DNA methylation and miRNA repression causes decreased hepatic *IRS2* expression in obese subjects with T2D

We were able to show that hepatic *IRS2* expression is repressed in T2D and NAFLD by DNA hypermethylation and high serum insulin stimulated overexpression of hepatic hsa-let-7e-5p. *In silco* analysis revealed a SREBF1 recognition motif (E box, CpG6) in proximity to a Sp1 recognition motif (GC box, CpG3) within intron 1 of *IRS2*. A shift assay validated both candidate transcription factors by respective antibody binding. DNA methylation within the E box motif of intron 1 is about 3.7 % reduced in T2D subjects and correlates significantly with a decreased *IRS2* gene expression. This principle influence of methylation on gene expression was further validated by a Luciferase reporter gene assay.
Insulin receptor substrate (IRS) 1 and 2 are key molecular adapter for transmitting hepatic insulin signaling after activation of the respective receptor by insulin. A liver specific knock-out of $Irs2^{-/-}$ in mice leads to a phenotype called selective insulin resistance, which is characterized by sustained lipogenesis and gluconeogenesis in hyperglycemic state (Kubota et al., 2008). This manifests as increased insulin resistance and hepatic fat accumulation. Interestingly, a knock-out of *Irs1* results into a milder insulin resistance due to different distributions of both proteins between the PP and PV zone and additional compensation via *Irs2* upregulation in both zones (Kubota et al., 2016).
Regulation of *IRS2* gene expression by Forkhead transcription factors which bind the insulin responsive element (IRE) of the promoter is well described (Lu et al., 2012). A proximal sterol regulatory element (SRE) leads to competitive binding between SREBF1 and FOXO transcription factors. Thereby SREBF1

acts as a translational repressor (Ide et al., 2004). It is further known, that SREBF1 needs Sp1 or Nuclear transcription factor Y (NFY) as cofactor for DNA binding (Reed et al., 2008). Binding of both identified transcription factors SREBF1 and Sp1 is repressed by DNA methylation (Zhu et al., 2003; Douet et al., 2007; Ehara et al., 2012). Also the measured difference in methylation of CpG6 is in the same range as a previously reported difference of about 3 % within a CGI of *IRS2* exon 1 (cg05514401) described in subcutaneous adipose tissue of insulin resistant obese women concomitant with decreased IRS2 gene expression (Arner et al., 2016). The E box methylation (CpG6) does not correlate with any other metabolic trait, therefore decreased methylation appears to be a long-term mechanism for *IRS2* repression. Interestingly, GC box (CpG3) methylation correlates positively with serum insulin levels and negatively with HbA1c and AST though no significant difference in mean methylation values between T2D and ND subjects is observed. GC box methylation appears to be a short to medium term mechanism to enhance SREBF1 binding and *IRS2* repression.

The genotype at the neighboring SNP rs4547213 influences DNA methylation at both CpG sites. The minor allele A has protective function by being associated with increased methylation within the GC box. This protective function is in accordance to a previously published meta study of about 900.000 subjects European ancestry (p = 6.80E-07, effect size minor allele A $E_A = 0.967$) (Mahajan et al., 2018). DNA methylation within the E box motif is also influenced by this polymorphism but in the opposite manner. Luciferase reporter gene and shift assay reveal no functional association of the SNP itself to gene expression nor protein binding. Its effect is solely transmitted by influencing DNA methylation.

*IRS2* promoter activity is reduced by decreased DNA methylation causing enhanced binding of SREBF1 and Sp1. This relationship between DNA methylation and promoter activity could not be reproduced *in vitro* by the Luciferase reporter gene assay. It should be noted that the Luciferase assay uses a CMV promoter in an artificial system and might not mirror the mechanism of competing transcription factor binding, which was already shown for SREBF1 and FOXO1/3. Nevertheless, the Luciferase assay remains a prove for a mechanistic effect of this sequence snippet.

Besides T2D, aberrant *IRS2* expression is also found in the context liver fibrosis and steatosis (NAS). A different study of liver biopsies of 51 non-diabetic obese subjects reveald a significant decrease in *IRS2* with concomitant increased expression of enzymes involved in gluconeogenesis (PEPCK, G6Pase) in NAFLD livers (Honma et al., 2018). Controversially, another study of 71 non-diabetic obese subjects stated an increase of *IRS2* expression with severity of NAFLD state caused by increased *FOXO1* expression and activating FOXO protein phosphorylation (Rametta et al., 2013). Increased IRS2 activation is thought to mediate insulin stimulation and persistent *de novo* lipogenesis leading to a worsening of NAFLD. Expression of *IRS2* appears to be dynamically regulated in dependency of the progression of insulin resistance and NAFLD. This emphasizes an underlying epigenetic mechanism in *IRS2* regulation and the dependency of the exposure time to hyperglycemia.

Liver-specific $Irs1^{-/-}$, $Irs2^{-/-}$ or double knock-out mice show distinct activation patterns of associated signaling pathways in dependency of fasting or feeding of mice. *IRS2* expression is mostly upregulated during fasting conditions, leading to increased PI3K/AKT signaling, whereby IRS1 is dominantly active during refeeding. The knockout of Irs2 leads to insulin resistance during fasting. Our human liver biopsies were taken in a fasted state with IRS2 being a central regulatory molecule for insulin responsiveness.

Correlation analysis favors a transcriptional repression of IRS2 by the miRNA hsa-let-7e-5p as cofactor for development of insulin resistance in NAFLD, since expression of both genes correlate with liver NAS score (IRS2 is downregulated and hsa-let-7e-5p is upregulated with higher NAS) while displaying an altered gene expression between T2D and ND subjects. Additionally, expression of *IRS2* and hsa-let-7e-5p correlate negatively with each other. Binding of this miRNA to the 3'-UTR of *IRS2* was previously established (Zhu et al., 2011; Boudreau et al., 2014; Gao et al., 2014). Targeting *IRS2* mRNA by let-7e mimic transfection was not possible due to the low expression of *IRS2* in HepG2 cells.

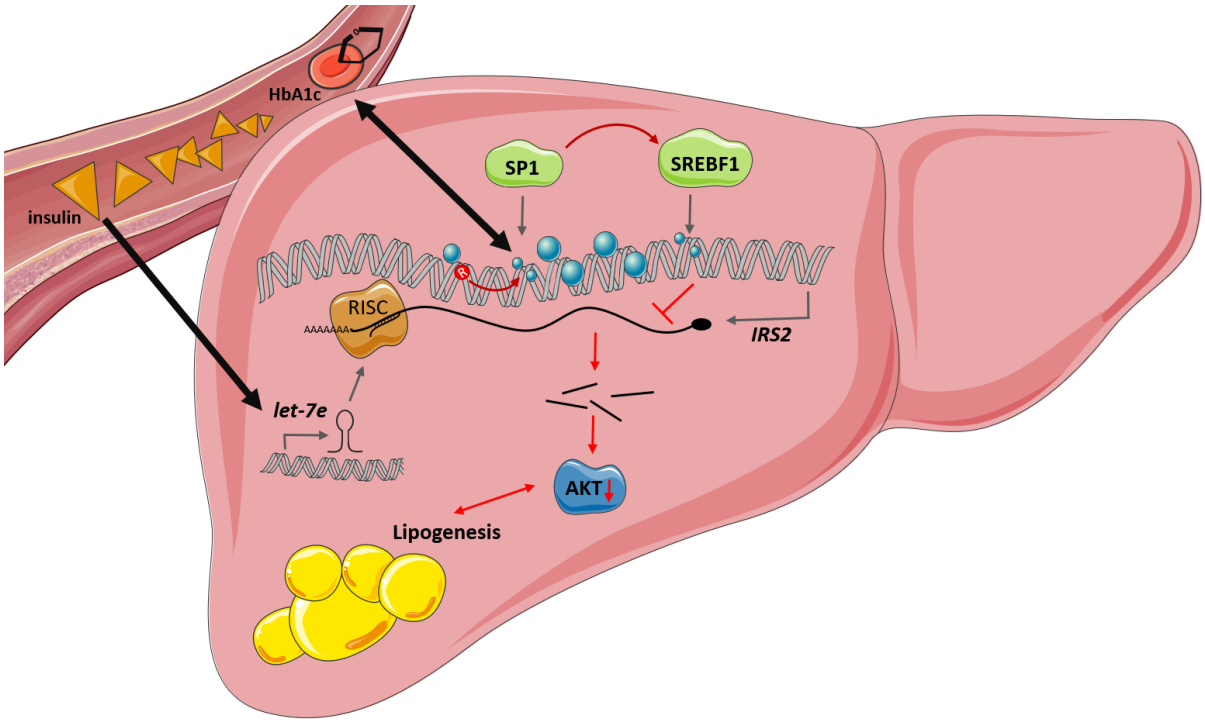


Figure 6.1: **Proposed mechanism of hepatic *IRS2* expression in T2D and NAFLD.** Long-term exposure to hyperglycemia reduces DNA methylation within a SREBF1 recognition motif. A bad blood glucose control results into high HbA1c which leads to additional de-methylation of a Sp1 recognition motif, which enhances binding and the repressing effect of SREBF1 on *IRS2* transcription. The methylated CG genotype of a GWAS-SNP (CG>CA) leads to enhanced methylation of the Sp1 recognition motif which has a protective effect by regional hypermethylation. Acute high levels of serum insulin leads to expression of the hepatic miRNA hsa-let-7e, which binds to the 3'-UTR and degrades transcribed IRS2 mRNA. Altogether, reduced *IRS2* expression leads to reduced signal transmission, resulting into a decreased AKT activation and enhanced lipogenesis, favoring the progression of NAFLD.

Treatment of HepG2 cells with high glucose and insulin concentrations was able to induce more than 2.5 fold miRNA expression independently of the used insulin concentration. Though hepatic hsa-let-7e-5p expression appears to be highly responsive to serum insulin, no difference in serum miRNA concentration was found between T2D and ND subjects. Nevertheless, serum miRNA concentration correlated significantly with hepatic miRNA expression, therefore considering let-7e-5p appears to be a promising approach for biomarker development. Additionally, miRNAs can be secreted to the bloodstream to influence genes in other tissues and to act in a systemic way (Wang et al., 2013b). Other studies also indicate a function of let-7e upregulation in disease progression of hepatocellular carcinoma (HCC) which is often a consequence of NAFLD (Shi et al., 2017). Reducing hepatic hsa-let-7e-5p expression by antagonizing molecules represents a potential therapeutic target for metabolic liver diseases.
The complex mechanism of DNA methylation, genotype and miRNA binding to regulate *IRS2* expression is summarized in Figure 6.1.

## 6.6 Predictive risk scoring and disease stratification based on hepatic markers

Two attempts of supervised learning methods were used for the prediction of T2D development based on different training assumptions. One assumption was about a linear correlation between predictors and outcome, therefore a methylation risk score (MRS) and linear regression models were used. The

other assumption was a non-linear relation between predictors and outcome, therefore a robust machine learning approach called random decision forest (RDF) was used. Bot approaches are able to categorize T2D and ND subjects into their respective subgroups but retain a high error rate.

The principle of a MRS is similar to genetic risk scores (GRS) or risk allele counting using genotype information for T2D prediction (Meigs et al., 2008; van Hoek et al., 2008; Cornelis et al., 2009; De Jager et al., 2009; Lin et al., 2009; Yarwood et al., 2015). Addition of small scale methylation changes might strengthen the effect on an outcome like BMI or the incidence of T2D. Additionally, the calculation and use of a MRS is already common practice to estimate lethality in prostate cancer (Ahmad et al., 2016). Thereby we assumed a linear negative relationship between DNA methylation and obesity. To achieve comparability with methylation results of other tissues (mostly with blood based EWAS), a standardized hepatic methylation value was used for generation of a MRS.

Though the mean MRS values of T2D and ND subjects do not differ significantly and the regression analysis concerning BMI as output variable is not significant, the MRS proofs to have a marginal better sensitivity and specificity than coin flip (AUC $> 0.5$). The usage of truly associated tissue-specific methylation markers as input might improve further stratification. Also markers are needed which work on the individual and not on a whole population. Most proposed markers originate from large-scale population based or meta studies and are rarely evaluated in metabolically active tissue as discussed before. In comparison, most GRS in addition to common factors (age, gender, BMI) improve risk stratification by increasing the AUC by only 0.01 (AUC in the range of 0.6 to 0.8) (van Hoek et al., 2008; Lin et al., 2009). Other approaches calculate an odds ratio based on allele information which cannot be applied to this cohort due to a missing follow-up.

A similar study regarding a blood based MRS for T2D prediction was performed previously by combination of CpG loci within *ABCG1*, *SREBF1*, *SOCS3* (encoding Suppressor of cytokine signaling 3), *TXNIP* (encoding Thioredoxin Interacting Protein) and *PHOSPHO1* (encoding Phosphoethanolamine/Phosphocholine Phosphatase) which is not resulting into future T2D prediction in n = 258 subjects (Dayeh et al., 2016). Solely methylation at *ABCG1* cg06500161 is associated to an 9 % increased risk for T2D (Dayeh et al., 2016).

A bottleneck of this approach is the use of only one training data set and the lack of further validation. Besides a training and testing cohort (discovery cohort), a second cohort (replication cohort) is needed to validate each approach independently of training-induced bias. Due to the resulting low sample size it is not possible to split this cohort into a meaningful training and testing subset. Therefore the MRS was only characterized by linear regression analysis regarding a correlation between score and BMI. Furthermore, the ROC analysis and the statistical test whether the mean scores differ between T2D and ND subjects is only exploratory for this cohort. DNA methylation is also tissue specific and blood proved to be an unfavorable surrogate tissue to mirror the hepatic situation. Liver tissue is rarely accessible which makes it questionable if a hepatic MRS is an useful approach to predict T2D or obesity.

RDFs are used if a non-linear relationship between incidence of a phenotype and its predictive variables is assumed. This is also called a non-parametric model. An example is the relationship between age and T2D. High age is a know risk factor for T2D manifestation but younger people also develop T2D and are more prone for disease manifestation if they are obese. On the other hand, obese subjects can be glucose tolerant. So there is no linear rule between age, obesity and the incidence of T2D for one individual. This interplay was assumed between hepatic DNA methylation of *IRS2* intron 1 which was represented by the SNP rs4547213 and hepatic hsa-let-7e-5p expression which was represented by serum hsa-let-7e-5p expression. Both mechanisms, DNA methylation and 3'-UTR binding by miRNA, repress hepatic *IRS2* expression independently of each other.

RDFs were previously used in other fields to predict transplant rejection or the detection of Alzheimer's disease (Lebedev et al., 2014; Shaikhina et al., 2017). Other concepts on type 2 diabetes prediction and subject stratification are mostly based on support vector machines (SVMs) (Chen et al., 2014; Han et al.,

2015; Wu et al., 2009; Barakat et al., 2010). Nevertheless, RDF are better in dealing with small sample sizes and unbalanced distribution within the data, as well with data that contains more predictors than observations (Liu et al., 2013a; Boulesteix et al., 2012).
The RDF also uses a bootstrapping approach (also called bagging), which is more robust against outlier by choosing randomly predictor variables and training subsets (Breiman, 1996a; Le Teuff et al., 2005). Due to its randomness this method reduces a training-induced bias for disease prediction of subjects which were not included in model training. The out-of-bag error (OOB) is an estimate to measure the prediction error and is often used to avoid validation of a RDF in an independent data subset of a replication cohort (Breiman, 1996b). Nevertheless, the OOB error tends to overestimate the true prediction error when an RDF is used on small data sets with low variance and a large number of predictors (Janitza and Hornung, 2018). This overestimation is especially important when comparing different sets of predictors, number of trees or tree depth during RDF modeling and thereby choosing the best settings. Our generated models RE not well balanced (large variance between the predictors) and only few (three to seven) predictors are used. Therefore only the small sample subset (n = 84) and a small correlation between each predictor set might affect the true prediction error.
The RDF which includes besides age, gender and BMI also blood based markers (genotype at *IRS2* intron 1 rs4547213 and serum hsa-let-7e-5p expression) is the best prediction model with 35.8 % classification error. In comparison, other RDFs are using only age, gender and BMI (OOB error 39.1 %) or additional information on hepatic DNA methylation at *IRS2* intron 1 (OOB error 36.1 %). Although addition of genotype information and miRNA expression improves significantly disease prediction, it is not reach a specificity of $> 90$ % as other published methods used with SVMs. Nevertheless, these markers are metabolically well evaluated and might change during disease progression by reflecting the actual health state of the liver. Since this analysis is performed in a cross-sectional study, this hypothesis can not be addressed.
Another algorithm for unsupervised learning should be mentioned, which was not used for T2D prediction on any here generated data. Deep learning algorithms used in artificial neural networks (ANNs) proved to reduce prediction error in T2D risk estimation (Rahimloo and Jafarian, 2016; Gholipour et al., 2018). These models need large sets of input data for training, testing and validation.
Epigenetic marks are known to be inherited form both parents, as well as certain predisposition towards glucose intolerance and obesity (Ravelli et al., 1998; Vaag et al., 2012; Tomar et al., 2015). Addition of such epigenetic marks might further improve early risk estimation, despite genetic risk scores and the early existence of several regression models based on blood parameters and the ongoing development (Welborn et al., 1984; Abdul-Ghani et al., 2011; Gholipour et al., 2018; Tabaei and Herman, 2002).

## 6.7 Conclusion, limitations and advantages

Distinct putative regulatory mechanism were identified in human liver which cause altered expression of specific genes in T2D. Whether these mechanism are causal to T2D manifestation, a consequence of sustained exposure to hyperglycemia, a protective mechanism to counteract metabolic dysregulation or even a driver of the parthenogenesis remains unclear for now. Nevertheless, it is mandatory to first identify dysregulated pathways and its key genes in humans as especially the glucose metabolism can vary greatly between mice and humans. This thesis represents a solid base for further research on T2D development regarding the regulation of *IRS2*, *LRP6* and *FOXO1* by miRNAs, as well concerning the regulation of *IRS2* and *GALNT18* by DNA methylation (Figure 6.2). Besides these genes, also genes involved in hepatic TH signaling were further analyzed regarding the progression of NAFLD.

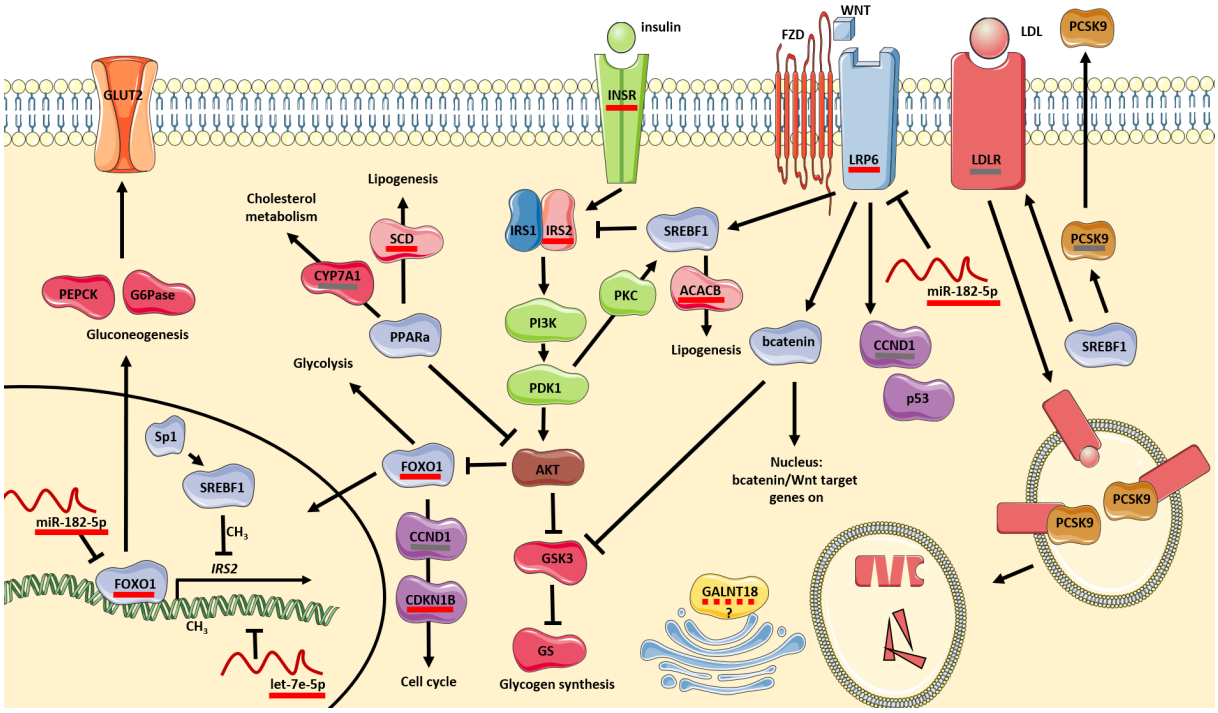


Figure 6.2: **Summary of all analyzed genes and interaction between them.** Besides their interaction, also newly identified epigenetic regulatory mechanisms (DNA methylation $CH_3$ for *IRS2* and miRNA repression for *IRS2, FOXO1* and *LRP6*) are indicated. Red lines indicate genes whose expression showed a metabolic association, grey lines indicate genes with were analyzed but did not show any correlation to metabolic parameters. *GALNT18* showed T2D and liver steatosis dependent DNA methylation changes but its metabolic function remains unclear. For a detailed overview regarding the proposed complex regulatory mechanism of *IRS2* see Figure 6.1.

Some genes (namely *ACACB*, *INSR*, *SCD* and *THRB*) are dysregulated in T2D and/or NAFLD but its mechanism remains unclear. An approach of miRNA regulation was not plausible for now.

Additionally, another important question was not addressed by this thesis. How is the change in miRNA expression or introduction of aberrant DNA methylation itself regulated? Especially DNA methylation regulation and function is highly discussed in literature (Robertson, 2005; He et al., 2011; Ziller et al., 2013; Chen and Riggs, 2011).

An important limitation in human studies remain the influence of medications. Metformin is a commonly prescribed anti-diabetic drug which was shown to increase hepatic insulin sensitivity (Tiikkainen et al., 2004). Effects of Metformin on gene expression might act as confounding factor.

One superior strength of the setting of this thesis is the relatively large sample size of metabolically relevant human tissue (instead of blood) and the species. Though the human and murine genome share common genes, promoter, exon coding sequences and physiology, small differences with high impact on metabolic functions are observed (Perlman, 2016). For example the whole sequence analysis of *IRS2* intron 1 cannot be reproduced in a murine model due to a lack of motifs and polymorphisms. Also, as proved and discussed within this thesis, the usage of easy to assess tissue or cells might not represent a suitable surrogate marker to mimic hepatic alterations in T2D. Lastly, a liver cohort of morbidity obese human subjects represents the best model to study manifestation of metabolic diseases independent of BMI.

# 7 Outlook

Due to ethical reasons, longitudinal approaches to study epigenome wide changes in human liver are not possible. These longitudinal studies would include liver biopsies of healthy subjects which were taken without medical indication. Instead, alterations of (gene and miRNA) expression and DNA methylation patterns can be accessed in samples pre and post bariatric surgery.

One form of bariatric surgery, the Roux-en-Y gastric bypass, which is a drastic, irreversible reduction of the stomach volume, proves to be an effective treatment against obesity which can cause also a sustained reversion of T2D (Madsen et al., 2019). Significant effects of gastric bypass on hepatic DNA methylation and gene expression was already shown for NAFLD related diseases and might be induced by a rapid weight loss (Ahrens et al., 2013). Nevertheless, the molecular mechanism of T2D reversion remains unclear and need to be elucidated to understand disease etiology further.

Also the discrepancy between "good" responders and non-responders, who are subjects lacking long-term weight loss and T2D reversion, cannot be explained to date. In regard to the irreversibly of this serious interference, predictive markers whether this surgery would improve one's metabolic health are urgently needed. Blood based markers would be most suitable in order to depict such a prospective model. Recent human studies identified several serum, plasma or exosomal miRNAs, which show alterations pre and post bariatric surgery, which includes also liver-specific miR-122-5p (Lopez et al., 2017; Atkin et al., 2019; Bae et al., 2019). Different questions need to be elucidated in future work: Are concentration changes of these miRNAs are reproducible in larger cohorts and different ethnicities, from which tissue do they originate and which tissues and metabolic pathways do they affect?

The obesity optimized cohort design, tissue collection (serum and liver) and a large sample size of metabolic tissue biopsies facilitate a solid basis to address some of these pending issues for hepatic alterations in future work.

# 8 Literature


Abdul-Ghani, M. A., T. Abdul-Ghani, M. P. Stern, J. Karavic, T. Tuomi, I. Bo, R. A. DeFronzo, and L. Groop
2011. Two-Step Approach for the Prediction of Future Type 2 Diabetes Risk. *Diabetes Care*, 34(9):2108–2112.

Aberg, K. A., A. A. Shabalin, R. F. Chan, M. Zhao, G. Kumar, G. v. Grootheest, S. L. Clark, L. Y. Xie, Y. Milaneschi, B. W. J. H. Penninx, and E. J. C. G. v. d. Oord
2018. Convergence of evidence from a methylome-wide CpG-SNP association study and GWAS of major depressive disorder. *Translational Psychiatry*, 8(1):162.

Abu-Elheiga, L.
2001. Continuous Fatty Acid Oxidation and Reduced Fat Storage in Mice Lacking Acetyl-CoA Carboxylase 2. *Science*, 291(5513):2613–2616.

Accili, D. and K. C. Arden
2004. FoxOs at the crossroads of cellular metabolism, differentiation, and transformation. *Cell*, 117(4):421–426.

Ahmad, A. S., N. Vasiljević, P. Carter, D. M. Berney, H. Møller, C. S. Foster, J. Cuzick, and A. T. Lorincz
2016. A novel DNA methylation score accurately predicts death from prostate cancer in men with low to intermediate clinical risk factors. *Oncotarget*, 7(44).

Ahrens, M., O. Ammerpohl, W. von Schönfels, J. Kolarova, S. Bens, T. Itzel, A. Teufel, A. Herrmann, M. Brosch, H. Hinrichsen, W. Erhart, J. Egberts, B. Sipos, S. Schreiber, R. Häsler, F. Stickel, T. Becker, M. Krawczak, C. Röcken, R. Siebert, C. Schafmayer, and J. Hampe
2013. DNA Methylation Analysis in Nonalcoholic Fatty Liver Disease Suggests Distinct Disease-Specific and Remodeling Signatures after Bariatric Surgery. *Cell Metabolism*, 18(2):296–302.

Akinlade, K., S. Agbebaku, S. Rahamon, and W. Balogun
2015. Vitamin B12 levels in patients with type 2 diabetes mellitus on Metformin. *Annals of Ibadan Postgraduate Medicine*, 13(2):79–83.

Akinyemiju, T., A. N. Do, A. Patki, S. Aslibekyan, D. Zhi, B. Hidalgo, H. K. Tiwari, D. Absher, X. Geng, D. K. Arnett, and M. R. Irvin
2018. Epigenome-wide association study of metabolic syndrome in African-American adults. *Clinical Epigenetics*, 10(1):49.

American Diabetes Association
2018. 2. Classification and Diagnosis of Diabetes: *Standards of Medical Care in Diabetes—2018. Diabetes Care*, 41(Supplement 1):S13–S27.

Andersen, C. L., J. L. Jensen, and T. F. Ørntoft
2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*, 64(15):5245–5250.

Andersson, E. A., K. H. Allin, C. H. Sandholt, A. Borglykke, C. J. Lau, R. Ribel-Madsen, T. Sparso, J. M. Justesen, M. N. Harder, M. E. Jorgensen, T. Jorgensen, T. Hansen, and O. Pedersen
2013. Genetic Risk Score of 46 Type 2 Diabetes Risk Variants Associates With Changes in Plasma Glucose and Estimates of Pancreatic -Cell Function Over 5 Years of Follow-Up. *Diabetes*, 62(10):3610–3617.

Apriletti, J. W., R. C. Ribeiro, R. L. Wagner, W. Feng, P. Webb, P. J. Kushner, B. L. West, S. Nilsson, T. S. Scanlan, R. J. Fletterick, and J. D. Baxter
1998. Molecular and Structural Biology of Thyroid Hormone Receptors. *Clinical and Experimental Pharmacology and Physiology*, 25(S1):S2–S11.

Araki, O., H. Ying, X. G. Zhu, M. C. Willingham, and S. Y. Cheng
2009. Distinct dysregulation of lipid metabolism by unliganded thyroid hormone receptor isoforms. *Molecular Endocrinology (Baltimore, Md.)*, 23(3):308–315.

Arner, P., A.-S. Sahlqvist, I. Sinha, H. Xu, X. Yao, D. Waterworth, D. Rajpal, A. K. Loomis, J. M. Freudenberg, T. Johnson, A. Thorell, E. Näslund, M. Ryden, and I. Dahlman
2016. The epigenetic signature of systemic insulin resistance in obese women. *Diabetologia*, 59(11):2393–2405.

Atkin, S. L., V. Ramachandran, N. A. Yousri, M. Benurwar, S. C. Simper, R. McKinlay, T. D. Adams, S. H. Najafi-Shoushtari, and S. C. Hunt
2019. Changes in Blood microRNA Expression and Early Metabolic Responsiveness 21 Days Following Bariatric Surgery. *Frontiers in Endocrinology*, 9.

Backes, C., T. Fehlmann, F. Kern, T. Kehl, H.-P. Lenhof, E. Meese, and A. Keller
2018. miRCarta: a central repository for collecting miRNA candidates. *Nucleic Acids Research*, 46(D1):D160–D167.

Bae, Y.-U., Y. Kim, H. Lee, H. Kim, J. S. Jeon, H. Noh, D. C. Han, S. Ryu, and S. H. Kwon
2019. Bariatric Surgery Alters microRNA Content of Circulating Exosomes in Patients with Obesity: Exosomal miRNA Profiling in Patients with Obesity. *Obesity*, 27(2):264–271.

Barakat, N. H., A. P. Bradley, and M. N. H. Barakat
2010. Intelligible support vector machines for diagnosis of diabetes mellitus. *IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society*, 14(4):1114–1120.

Bartel, D. P.
2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116(2):281–297.

Bartel, D. P.
2009. MicroRNAs: target recognition and regulatory functions. *Cell*, 136(2):215–233.

Benz, F., C. Roderburg, D. Vargas Cardenas, M. Vucur, J. Gautheron, A. Koch, H. Zimmermann, J. Janssen, L. Nieuwenhuijsen, M. Luedde, N. Frey, F. Tacke, C. Trautwein, and T. Luedde
2013. U6 is unsuitable for normalization of serum miRNA levels in patients with sepsis or liver fibrosis. *Experimental & Molecular Medicine*, 45(9):e42–e42.

Bernstein, B. E., A. Meissner, and E. S. Lander
2007. The mammalian epigenome. *Cell*, 128(4):669–681.

Bilgin, H. and z. Pirgon
2014. Thyroid Function in Obese Children with Non-Alcoholic Fatty Liver Disease. *Journal of Clinical Research in Pediatric Endocrinology*, Pp. 152–157.

Bohinc, B. N., G. Michelotti, G. Xie, H. Pang, A. Suzuki, C. D. Guy, D. Piercy, L. Kruger, M. Swiderska-Syn, M. Machado, T. Pereira, A. M. Zavacki, M. Abdelmalek, and A. M. Diehl
2014. Repair-Related Activation of Hedgehog Signaling in Stromal Cells Promotes Intrahepatic Hypothyroidism. *Endocrinology*, 155(11):4591–4601.

Bos, M. M., R. A. J. Smit, S. Trompet, D. van Heemst, and R. Noordam
2017. Thyroid Signaling, Insulin Resistance, and 2 Diabetes Mellitus: A Mendelian Randomization Study. *The Journal of Clinical Endocrinology & Metabolism*, 102(6):1960–1970.

Boudreau, R. L., P. Jiang, B. L. Gilmore, R. M. Spengler, R. Tirabassi, J. A. Nelson, C. A. Ross, Y. Xing, and B. L. Davidson
2014. Transcriptome-wide Discovery of microRNA Binding Sites in Human Brain. *Neuron*, 81(2):294–305.

Boujedidi, H., L. Bouchet-Delbos, A.-M. Cassard-Doulcier, M. Njiké-Nakseu, S. Maitre, S. Prévot, I. Dagher, H. Agostini, C. S. Voican, D. Emilie, G. Perlemuter, and S. Naveau
2012. Housekeeping gene variability in the liver of alcoholic patients. *Alcoholism, Clinical and Experimental Research*, 36(2):258–266.

Boulesteix, A.-L., S. Janitza, J. Kruppa, and I. R. König
2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics: Random forests in bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6):493–507.

Braun, K. V. E., K. Dhana, P. S. de Vries, T. Voortman, J. B. J. van Meurs, A. G. Uitterlinden, A. Hofman, F. B. Hu, O. H. Franco, and A. Dehghan
2017. Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study. *Clinical Epigenetics*, 9(1):15.

Breiman, L.
1996a. Bagging Predictors. *Machine Learning*, 24(2):123–140.

Breiman, L.
1996b. Out-of-bag estimation. *Techni-cal report, Dept. of Statistics, Univ. of Calif., Berke-ley*, P. 13.

Bridges, J. P., A. Schehr, Y. Wang, L. Huo, V. Besnard, M. Ikegami, J. A. Whitsett, and Y. Xu
2014. Epithelial SCAP/INSIG/SREBP Signaling Regulates Multiple Biological Processes during Perinatal Lung Maturation. *PLoS ONE*, 9(5).

Brown, M. S. and J. L. Goldstein
2008. Selective versus Total Insulin Resistance: A Pathogenic Paradox. *Cell Metabolism*, 7(2):95–96.

Brunt, E. M., D. E. Kleiner, L. A. Wilson, P. Belt, and B. A. Neuschwander-Tetri
2011. The NAS and The Histopathologic Diagnosis in NAFLD: Distinct Clinicopathologic Meanings. *Hepatology (Baltimore, Md.)*, 53(3):810–820.

Campanella, G., M. J. Gunter, S. Polidoro, V. Krogh, D. Palli, S. Panico, C. Sacerdote, R. Tumino, G. Fiorito, S. Guarrera, L. Iacoviello, I. A. Bergdahl, B. Melin, P. Lenner, T. M. C. M. de Kok, P. Georgiadis, J. C. S. Kleinjans, S. A. Kyrtopoulos, H. B. Bueno-de Mesquita, K. A. Lillycrop, A. M. May, N. C. Onland-Moret, R. Murray, E. Riboli, M. Verschuren, E. Lund, N. Mode, T. M. Sandanger, V. Fiano, M. Trevisan, G. Matullo, P. Froguel, P. Elliott, P. Vineis, and M. Chadeau-Hyam
2018. Epigenome-wide association study of adiposity and future risk of obesity-related diseases. *International Journal of Obesity*, 42(12):2022–2035.

Cedar, H. and Y. Bergman
2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics*, 10(5):295–304.

Cermelli, S., A. Ruggieri, J. A. Marrero, G. N. Ioannou, and L. Beretta
2011. Circulating MicroRNAs in Patients with Chronic Hepatitis C and Non-Alcoholic Fatty Liver Disease. *PLoS ONE*, 6(8):e23937.

Chambers, J. C., M. Loh, B. Lehne, A. Drong, J. Kriebel, V. Motta, S. Wahl, H. R. Elliott, F. Rota, W. R. Scott, W. Zhang, S.-T. Tan, G. Campanella, M. Chadeau-Hyam, L. Yengo, R. C. Richmond, M. Adamowicz-Brice, U. Afzal, K. Bozaoglu, Z. Y. Mok, H. K. Ng, F. Pattou, H. Prokisch, M. A. Rozario, L. Tarantini, J. Abbott, M. Ala-Korpela, B. Albetti, O. Ammerpohl, P. A. Bertazzi, C. Blancher, R. Caiazzo, J. Danesh, T. R. Gaunt, S. de Lusignan, C. Gieger, T. Illig, S. Jha, S. Jones, J. Jowett, A. J. Kangas, A. Kasturiratne, N. Kato, N. Kotea, S. Kowlessur, J. Pitkäniemi, P. Punjabi, D. Saleheen, C. Schafmayer, P. Soininen, E.-S. Tai, B. Thorand, J. Tuomilehto, A. R. Wickremasinghe, S. A. Kyrtopoulos, T. J. Aitman, C. Herder, J. Hampe, S. Cauchi, C. L. Relton, P. Froguel, R. Soong, P. Vineis, M.-R. Jarvelin, J. Scott, H. Grallert, V. Bollati, P. Elliott, M. I. McCarthy, and J. S. Kooner
2015. Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *The Lancet Diabetes & Endocrinology*, 3(7):526–534.

Chen, H., C. Tan, Z. Lin, and T. Wu
2014. The diagnostics of diabetes mellitus based on ensemble modeling and hair/urine element level analysis. *Computers in Biology and Medicine*, 50:70–75.

Chen, Z.-x. and A. D. Riggs
2011. DNA Methylation and Demethylation in Mammals. *Journal of Biological Chemistry*, 286(21):18347–18353.

Chou, C.-H., S. Shrestha, C.-D. Yang, N.-W. Chang, Y.-L. Lin, K.-W. Liao, W.-C. Huang, T.-H. Sun, S.-J. Tu, W.-H. Lee, M.-Y. Chiew, C.-S. Tai, T.-Y. Wei, T.-R. Tsai, H.-T. Huang, C.-Y. Wang, H.-Y. Wu, S.-Y. Ho, P.-R. Chen, C.-H. Chuang, P.-J. Hsieh, Y.-S. Wu, W.-L. Chen, M.-J. Li, Y.-C. Wu, X.-Y. Huang, F. L. Ng, W. Buddhakosai, P.-C. Huang, K.-C. Lan, C.-Y. Huang, S.-L. Weng, Y.-N. Cheng, C. Liang, W.-L. Hsu, and H.-D. Huang
2018. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Research*, 46(D1):D296–D302.

Chugh, P. and D. P. Dittmer
2012. Potential pitfalls in microRNA profiling: Potential pitfalls in microRNA profiling. *Wiley Interdisciplinary Reviews: RNA*, 3(5):601–616.

Clément, T., V. Salone, and M. Rederstorff
2015. Dual luciferase gene reporter assays to study miRNA function. *Methods in Molecular Biology (Clifton, N.J.)*, 1296:187–198.

Cornelis, M. C., L. Qi, C. Zhang, P. Kraft, J. Manson, T. Cai, D. J. Hunter, and F. B. Hu
2009. Joint effects of common genetic variants on the risk for type 2 diabetes in US men and women of European ancestry. *Annals of internal medicine*, 150(8):541–550.

Crider, K. S., T. P. Yang, R. J. Berry, and L. B. Bailey
2012. Folate and DNA Methylation: A Review of Molecular Mechanisms and the Evidence for Folate's Role2. *Advances in Nutrition*, 3(1):21–38.

Dambal, S., M. Shah, B. Mihelich, and L. Nonn
2015. The microRNA-183 cluster: the family that plays together stays together. *Nucleic Acids Research*, 43(15):7173–7188.

Dayeh, T., T. Tuomi, P. Almgren, A. Perfilyev, P.-A. Jansson, V. D. de Mello, J. Pihlajamäki, A. Vaag, L. Groop, E. Nilsson, and C. Ling
2016. DNA methylation of loci within *ABCG1* and *PHOSPHO1* in blood DNA is associated with future type 2 diabetes risk. *Epigenetics*, 11(7):482–488.

Dayeh, T., P. Volkov, S. Salö, E. Hall, E. Nilsson, A. H. Olsson, C. L. Kirkpatrick, C. B. Wollheim, L. Eliasson, T. Rönn, K. Bacos, and C. Ling
2014. Genome-Wide DNA Methylation Analysis of Human Pancreatic Islets from Type 2 Diabetic and Non-Diabetic Donors Identifies Candidate Genes That Influence Insulin Secretion. *PLOS Genetics*, 10(3):e1004160.

Dayeh, T. A., A. H. Olsson, P. Volkov, P. Almgren, T. Rönn, and C. Ling
2013. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia*, 56(5):1036–1046.

De Jager, P. L., L. B. Chibnik, J. Cui, J. Reischl, S. Lehr, K. C. Simon, C. Aubin, D. Bauer, J. F. Heubach, R. Sandbrink, M. Tyblova, P. Lelkova, E. Havrdova, C. Pohl, D. Horakova, A. Ascherio, D. A. Hafler, and E. W. Karlson
2009. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. *The Lancet Neurology*, 8(12):1111–1119.

Dekkers, K. F., M. van Iterson, R. C. Slieker, M. H. Moed, M. J. Bonder, M. van Galen, H. Mei, D. V. Zhernakova, L. H. van den Berg, J. Deelen, J. van Dongen, D. van Heemst, A. Hofman, J. J. Hottenga, C. J. H. van der Kallen, C. G. Schalkwijk, C. D. A. Stehouwer, E. F. Tigchelaar, A. G. Uitterlinden, G. Willemsen, A. Zhernakova, L. Franke, P. A. C. 't Hoen, R. Jansen, J. van Meurs, D. I. Boomsma, C. M. van Duijn, M. M. J. van Greevenbroek, J. H. Veldink, C. Wijmenga, E. W. van Zwet, P. E. Slagboom, J. W. Jukema, and B. T. Heijmans
2016. Blood lipids influence DNA methylation in circulating cells. *Genome Biology*, 17(1):138.

Demerath, E. W., W. Guan, M. L. Grove, S. Aslibekyan, M. Mendelson, Y.-H. Zhou, s. K. Hedman, J. K. Sandling, L.-A. Li, M. R. Irvin, D. Zhi, P. Deloukas, L. Liang, C. Liu, J. Bressler, T. D. Spector, K. North, Y. Li, D. M. Absher, D. Levy, D. K. Arnett, M. Fornage, J. S. Pankow, and E. Boerwinkle
2015. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Human Molecular Genetics*, 24(15):4464–4479.

Dhana, K., K. V. E. Braun, J. Nano, T. Voortman, E. W. Demerath, W. Guan, M. Fornage, J. B. J. van Meurs, A. G. Uitterlinden, A. Hofman, O. H. Franco, and A. Dehghan
2018. An Epigenome-Wide Association Study of Obesity-Related Traits. *American Journal of Epidemiology*, 187(8):1662–1669.

Di Mauro, S., M. Ragusa, F. Urbano, A. Filippello, A. Di Pino, A. Scamporrino, A. Pulvirenti, A. Ferro, A. Rabuazzo, M. Purrello, F. Purrello, and S. Piro
2016. Intracellular and extracellular miRNome deregulation in cellular models of NAFLD or NASH: Clinical implications. *Nutrition, Metabolism and Cardiovascular Diseases*, 26(12):1129–1139.

Dick, K. J., C. P. Nelson, L. Tsaprouni, J. K. Sandling, D. Aïssi, S. Wahl, E. Meduri, P.-E. Morange, F. Gagnon, H. Grallert, M. Waldenberger, A. Peters, J. Erdmann, C. Hengstenberg, F. Cambien, A. H. Goodall, W. H. Ouwehand, H. Schunkert, J. R. Thompson, T. D. Spector, C. Gieger, D.-A. Trégouët, P. Deloukas, and N. J. Samani
2014. DNA methylation and body-mass index: a genome-wide analysis. *The Lancet*, 383(9933):1990–1998.

Dolganiuc, A., J. Petrasek, K. Kodys, D. Catalano, P. Mandrekar, A. Velayudham, and G. Szabo
2009. MicroRNA expression profile in Lieber-DeCarli diet-induced alcoholic and methionine choline deficient diet-induced nonalcoholic steatohepatitis models in mice. *Alcoholism, Clinical and Experimental Research*, 33(10):1704–1710.

Dong, X. C.
2017. FOXO transcription factors in non-alcoholic fatty liver disease. *Liver Research*, 1(3):168–173.

Dong, X. C., K. D. Copps, S. Guo, Y. Li, R. Kollipara, R. A. DePinho, and M. F. White
2008. Inactivation of hepatic Foxo1 by insulin signaling is required for adaptive nutrient homeostasis and endocrine growth regulation. *Cell metabolism*, 8(1):65–76.

Douet, V., M. B. Heller, and O. Le Saux
2007. DNA methylation and Sp1 binding determine the tissue-specific transcriptional activity of the mouse Abcc6 promoter. *Biochemical and biophysical research communications*, 354(1):66–71.

Edgar, R., P. P. C. Tan, E. Portales-Casamar, and P. Pavlidis
2014. Meta-analysis of human methylomes reveals stably methylated sequences surrounding CpG islands associated with high gene expression. *Epigenetics & Chromatin*, 7(1):28.

Ehara, T., Y. Kamei, M. Takahashi, X. Yuan, S. Kanai, E. Tamura, M. Tanaka, T. Yamazaki, S. Miura, O. Ezaki, T. Suganami, M. Okano, and Y. Ogawa
2012. Role of DNA Methylation in the Regulation of Lipogenic Glycerol-3-Phosphate Acyltransferase 1 Gene Expression in the Mouse Neonatal Liver. *Diabetes*, 61(10):2442–2450.

Elyakim, E., E. Sitbon, A. Faerman, S. Tabak, E. Montia, L. Belanis, A. Dov, E. G. Marcusson, C. F. Bennett, A. Chajut, D. Cohen, and N. Yerushalmi
2010. hsa-miR-191 Is a Candidate Oncogene Target for Hepatocellular Carcinoma Therapy. *Cancer Research*, 70(20):8077–8087.

Esau, C., S. Davis, S. F. Murray, X. X. Yu, S. K. Pandey, M. Pear, L. Watts, S. L. Booten, M. Graham, R. McKay, A. Subramaniam, S. Propp, B. A. Lollo, S. Freier, C. F. Bennett, S. Bhanot, and B. P. Monia
2006. miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. *Cell Metabolism*, 3(2):87–98.

Esteller, M.
2002. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene*, 21(35):5427–5440.

Filipowicz, W. and H. Grosshans
2011. The liver-specific microRNA miR-122: biology and therapeutic potential. *Progress in Drug Research. Fortschritte Der Arzneimittelforschung. Progres Des Recherches Pharmaceutiques*, 67:221–238.

Fossati, P. and M. Romon-Rousseaux
1987. Insulin and HDL-cholesterol metabolism. *Diabete & Metabolisme*, 13(3 Pt 2):390–394.

Frisdal, E., S. Le Lay, H. Hooton, L. Poupel, M. Olivier, R. Alili, W. Plengpanich, E. F. Villard, S. Gilibert, M. Lhomme, A. Superville, L. Miftah-Alkhair, M. J. Chapman, G. M. Dallinga-Thie, N. Venteclef, C. Poitou, J. Tordjman, P. Lesnik, A. Kontush, T. Huby, I. Dugail, K. Clement, M. Guerin, and W. Le Goff
2015. Adipocyte ATP-Binding Cassette G1 Promotes Triglyceride Storage, Fat Mass Growth, and Human Obesity. *Diabetes*, 64(3):840–855.

Frost, R. J. A. and E. N. Olson
2011. Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs. *Proceedings of the National Academy of Sciences*, 108(52):21075–21080.

Fu, X., B. Dong, Y. Tian, P. Lefebvre, Z. Meng, X. Wang, F. Pattou, W. Han, X. Wang, F. Lou, R. Jove, B. Staels, D. D. Moore, and W. Huang
2015. MicroRNA-26a regulates insulin sensitivity and metabolism of glucose and lipids. *The Journal of Clinical Investigation*, 125(6):2497–2509.

Fujita, N., S. Tsujii, H. Kuwata, R. Kurokawa, S. Matsunaga, S. Okamura, T. Mashitani, M. Furuya, M. Kitatani, and H. Ishii
2012. Predictor variables and an equation for estimating HbA1c attainable by initiation of basal supported oral therapy†. *Journal of Diabetes Investigation*, 3(2):164–169.

Gao, Y., F. Wu, J. Zhou, L. Yan, M. J. Jurczak, H.-Y. Lee, L. Yang, M. Mueller, X.-B. Zhou, L. Dandolo, J. Szendroedi, M. Roden, C. Flannery, H. Taylor, G. G. Carmichael, G. I. Shulman, and Y. Huang
2014. The H19/let-7 double-negative feedback loop contributes to glucose metabolism in muscle cells. *Nucleic Acids Research*, 42(22):13799–13811.

Gebert, L. F. R., M. A. E. Rebhan, S. E. M. Crivelli, R. Denzler, M. Stoffel, and J. Hall
2014. Miravirsen (SPC3649) can inhibit the biogenesis of miR-122. *Nucleic Acids Research*, 42(1):609–621.

Geng, S., W. Zhu, S. Wang, C. Xie, X. Li, J. Wu, Y. Li, Y. Chen, X. Wang, Y. Meng, Q. Zhang, J. Chen, and C. Zhong
2018. P53 modulates hepatic insulin sensitivity through NF-kB and p38/ERK MAPK pathways. *Biochemical and Biophysical Research Communications*, 495(3):2139–2144.

Gholipour, K., M. Asghari-Jafarabadi, S. Iezadi, A. Jannati, and S. Keshavarz
2018. Modelling the prevalence of diabetes mellitus risk factors based on artificial neural network and multiple regression. *Eastern Mediterranean Health Journal*, 24(08):770–777.

Ghosh, A., A. Ghosh, S. Datta, D. Dasgupta, S. Das, S. Ray, S. Gupta, S. Datta, A. Chowdhury, R. Chatterjee, S. K. Mohapatra, and S. Banerjee
2016. Hepatic miR-126 is a potential plasma biomarker for detection of hepatitis B virus infected hepatocellular carcinoma. *International Journal of Cancer*, 138(11):2732–2744.

Go, G.-w.
2015. Low-Density Lipoprotein Receptor-Related Protein 6 (LRP6) Is a Novel Nutritional Therapeutic Target for Hyperlipidemia, Non-Alcoholic Fatty Liver Disease, and Atherosclerosis. *Nutrients*, 7(6):4453–4464.

Goedeke, L., F. M. Vales-Lara, M. Fenstermaker, D. Cirera-Salinas, A. Chamorro-Jorganes, C. M. Ramirez, J. A. Mattison, R. de Cabo, Y. Suarez, and C. Fernandez-Hernando
2013. A Regulatory Role for MicroRNA 33* in Controlling Lipid Metabolism Gene Expression. *Molecular and Cellular Biology*, 33(11):2339–2352.

Goldstein, J. L., R. B. Rawson, and M. S. Brown
2002. Mutant mammalian cells as tools to delineate the sterol regulatory element-binding protein pathway for feedback regulation of lipid synthesis. *Archives of Biochemistry and Biophysics*, 397(2):139–148.

Gong, J., Y. Tong, H.-M. Zhang, and A.-Y. Guo
2012. miRNASNP: a database of miRNA related SNPs and their effects on miRNA function. *BMC Bioinformatics*, 13(18):A2.

Gong, Z. and R. H. Muzumdar
2012. Pancreatic Function, Type 2 Diabetes, and Metabolism in Aging.

Gori, M., M. Arciello, and C. Balsano
2014. MicroRNAs in Nonalcoholic Fatty Liver Disease: Novel Biomarkers and Prognostic Tools during the Transition from Steatosis to Hepatocarcinoma.

Griffiths-Jones, S., H. K. Saini, S. van Dongen, and A. J. Enright
2007. miRBase: tools for microRNA genomics. *Nucleic Acids Research*, 36(Database):D154–D158.

Gu, S. and M. A. Kay
2010. How do miRNAs mediate translational repression? *Silence*, 1:11.

Guadaño-Ferraz, A., M. J. Obregón, D. L. St Germain, and J. Bernal
1997. The type 2 iodothyronine deiodinase is expressed primarily in glial cells in the neonatal rat brain. *Proceedings of the National Academy of Sciences of the United States of America*, 94(19):10391–10396.

Guo, Z.-W., C. Xie, J.-R. Yang, J.-H. Li, J.-H. Yang, and L. Zheng
2015. MtiBase: a database for decoding microRNA target sites located within CDS and 5'UTR regions from CLIP-Seq and expression profile datasets. *Database: The Journal of Biological Databases and Curation*, 2015.

Han, L., S. Luo, J. Yu, L. Pan, and S. Chen
2015. Rule extraction from support vector machines using ensemble learning approach: an application for diagnosis of diabetes. *IEEE journal of biomedical and health informatics*, 19(2):728–734.

Harris, T. E., T. A. Huffman, A. Chi, J. Shabanowitz, D. F. Hunt, A. Kumar, and J. C. Lawrence
2007. Insulin controls subcellular localization and multisite phosphorylation of the phosphatidic acid phosphatase, lipin 1. *The Journal of Biological Chemistry*, 282(1):277–286.

Hayes, C. N. and K. Chayama
2016. MicroRNAs as Biomarkers for Liver Disease and Hepatocellular Carcinoma. *International Journal of Molecular Sciences*, 17(3).

He, X.-J., T. Chen, and J.-K. Zhu
2011. Regulation and function of DNA methylation in plants and animals. *Cell Research*, 21(3):442–465.

He, Y., Y. Ding, B. Liang, J. Lin, T.-K. Kim, H. Yu, H. Hang, and K. Wang
2017. A Systematic Study of Dysregulated MicroRNA in Type 2 Diabetes Mellitus. *International Journal of Molecular Sciences*, 18(3).

Hedman, s. K., M. M. Mendelson, R. E. Marioni, S. Gustafsson, R. Joehanes, M. R. Irvin, D. Zhi, J. K. Sandling, C. Yao, C. Liu, L. Liang, T. Huan, A. F. McRae, S. Demissie, S. Shah, J. M. Starr, L. A. Cupples, P. Deloukas, T. D. Spector, J. Sundström, R. M. Krauss, D. K. Arnett, I. J. Deary, L. Lind, D. Levy, and E. Ingelsson
2017. Epigenetic Patterns in Blood Associated With Lipid Traits Predict Incident Coronary Heart Disease Events and Are Enriched for Results From Genome-Wide Association Studies. *Circulation: Cardiovascular Genetics*, 10(1).

Helman, A., D. Avrahami, A. Klochendler, B. Glaser, K. H. Kaestner, I. Ben-Porath, and Y. Dor
2016. Effects of ageing and senescence on pancreatic beta-cell function. *Diabetes, Obesity and Metabolism*, 18(S1):58–62.

Hidalgo, B., M. R. Irvin, J. Sha, D. Zhi, S. Aslibekyan, D. Absher, H. K. Tiwari, E. K. Kabagambe, J. M. Ordovas, and D. K. Arnett
2014. Epigenome-Wide Association Study of Fasting Measures of Glucose, Insulin, and HOMA-IR in the Genetics of Lipid Lowering Drugs and Diet Network Study. *Diabetes*, 63(2):801–807.

Hong, K.-W., M. Chung, and S. B. Cho
2014. Meta-analysis of genome-wide association study of homeostasis model assessment beta cell function and insulin resistance in an East Asian population and the European results. *Molecular Genetics and Genomics*, 289(6):1247–1255.

Honma, M., S. Sawada, Y. Ueno, K. Murakami, T. Yamada, J. Gao, S. Kodama, T. Izumi, K. Takahashi, S. Tsukita, K. Uno, J. Imai, E. Kakazu, Y. Kondo, K. Mizuno, N. Kawagishi, T. Shimosegawa, and H. Katagiri
2018. Selective insulin resistance with differential expressions of IRS-1 and IRS-2 in human NAFLD livers. *International Journal of Obesity*.

Horie, T., T. Nishino, O. Baba, Y. Kuwabara, T. Nakao, M. Nishiga, S. Usami, M. Izuhara, N. Sowa, N. Yahagi, H. Shimano, S. Matsumura, K. Inoue, H. Marusawa, T. Nakamura, K. Hasegawa, N. Kume, M. Yokode, T. Kita, T. Kimura, and K. Ono
2013. MicroRNA-33 regulates sterol regulatory element-binding protein 1 expression in mice. *Nature Communications*, 4(1):2883.

Hu, M., F. Phan, O. Bourron, P. Ferré, and F. Foufelle
2017. Steatosis and NASH in type 2 diabetes. *Biochimie*, 143:37–41.

Huang, J.-Y., H.-L. Chen, and C. Shih
2016. MicroRNA miR-204 and miR-1236 inhibit hepatitis B virus replication via two different mechanisms. *Scientific Reports*, 6.

Huang, Q., L. Chen, H. Teng, H. Song, X. Wu, and M. Xu
2015. Phenolic compounds ameliorate the glucose uptake in HepG2 cells' insulin resistance via activating AMPK: Anti-diabetic effect of phenolic compounds in HepG2 cells. *Journal of Functional Foods*, 19:487–494.

Hur, W., J. H. Lee, S. W. Kim, J.-H. Kim, S. H. Bae, M. Kim, D. Hwang, Y. S. Kim, T. Park, S.-J. Um, B.-J. Song, and S. K. Yoon
2015. Downregulation of microRNA-451 in non-alcoholic steatohepatitis inhibits fatty acid-induced proinflammatory cytokine production through the AMPK/AKT pathway. *The International Journal of Biochemistry & Cell Biology*, 64:265–276.

Ide, T., H. Shimano, N. Yahagi, T. Matsuzaka, M. Nakakuki, T. Yamamoto, Y. Nakagawa, A. Takahashi, H. Suzuki, H. Sone, H. Toyoshima, A. Fukamizu, and N. Yamada
2004. SREBPs suppress IRS-2-mediated insulin signalling in the liver. *Nature Cell Biology*, 6(4):351–357.

Izzi, B., M. Pistoni, K. Cludts, P. Akkor, D. Lambrechts, C. Verfaillie, P. Verhamme, K. Freson, and M. F. Hoylaerts
2016. Allele-specific DNA methylation reinforces PEAR1 enhancer activity. *Blood*, 128(7):1003–1012.

Jacobs, E., A. Hoyer, R. Brinks, A. Icks, O. Kuß, and W. Rathmann
2017. Healthcare costs of Type 2 diabetes in Germany. *Diabetic Medicine*, 34(6):855–861.

Jang, H. S., W. J. Shin, J. E. Lee, and J. T. Do
2017. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes*, 8(6).

Janitza, S. and R. Hornung
2018. On the overestimation of random forest's out-of-bag error. *PLOS ONE*, 13(8):e0201904.

Jeon, T.-I., R. M. Esquejo, M. Roqueta-Rivera, P. E. Phelan, Y.-A. Moon, S. S. Govindarajan, C. C. Esau, and T. F. Osborne
2013. An SREBP-responsive microRNA operon contributes to a regulatory loop for intracellular lipid homeostasis. *Cell Metabolism*, 18(1):51–61.

Ji, J., J. Shi, A. Budhu, Z. Yu, M. Forgues, S. Roessler, S. Ambs, Y. Chen, P. S. Meltzer, C. M. Croce, L.-X. Qin, K. Man, C.-M. Lo, J. Lee, I. O. L. Ng, J. Fan, Z.-Y. Tang, H.-C. Sun, and X. W. Wang
2009. MicroRNA expression, survival, and response to interferon in liver cancer. *The New England Journal of Medicine*, 361(15):1437–1447.

Jin, Y., Z. Chen, X. Liu, and X. Zhou
2013. Evaluating the MicroRNA Targeting Sites by Luciferase Reporter Gene Assay. *Methods in molecular biology (Clifton, N.J.)*, 936:117–127.

Jjingo, D., A. B. Conley, S. V. Yi, V. V. Lunyak, and I. K. Jordan
2012. On the presence and role of human gene-body DNA methylation. *Oncotarget*, 3(4):462–474.

Johnnidis, J. B., M. H. Harris, R. T. Wheeler, S. Stehling-Sun, M. H. Lam, O. Kirak, T. R. Brummelkamp, M. D. Fleming, and F. D. Camargo
2008. Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. *Nature*, 451(7182):1125–1129.

Jones, P. A.
2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews. Genetics*, 13(7):484–492.

Jopling, C.
2012. Liver-specific microRNA-122: Biogenesis and function. *RNA Biology*, 9(2):137–142.

Jopling, C. L., S. Schütz, and P. Sarnow
2008. Position-Dependent Function for a Tandem MicroRNA miR-122-Binding Site Located in the Hepatitis C Virus RNA Genome. *Cell Host & Microbe*, 4(1):77–85.

Kameswaran, V., N. Bramswig, L. McKenna, M. Penn, J. Schug, N. Hand, Y. Chen, I. Choi, A. Vourekas, K.-J. Won, C. Liu, K. Vivek, A. Naji, J. Friedman, and K. Kaestner
2014. Epigenetic Regulation of the DLK1-MEG3 MicroRNA Cluster in Human Type 2 Diabetic Islets. *Cell Metabolism*, 19(1):135–145.

Karagkouni, D., M. D. Paraskevopoulou, S. Chatzopoulos, I. S. Vlachos, S. Tastsoglou, I. Kanellos, D. Papadimitriou, I. Kavakiotis, S. Maniou, G. Skoufos, T. Vergoulis, T. Dalamagas, and A. G. Hatzigeorgiou
2018. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA–gene interactions. *Nucleic Acids Research*, 46(D1):D239–D245.

Kennedy, M. A., G. C. Barrera, K. Nakamura, n. Baldán, P. Tarr, M. C. Fishbein, J. Frank, O. L. Francone, and P. A. Edwards
2005. ABCG1 has a critical role in mediating cholesterol efflux to HDL and preventing cellular lipid accumulation. *Cell Metabolism*, 1(2):121–131.

Kilic, I. D., Y. Dodurga, B. Uludag, Y. I. Alihanoglu, B. S. Yildiz, Y. Enli, M. Secme, and H. E. Bostancı
2015. microRNA -143 and -223 in obesity. *Gene*, 560(2):140–142.

Kim, J. B., P. Sarraf, M. Wright, K. M. Yao, E. Mueller, G. Solanes, B. B. Lowell, and B. M. Spiegelman
1998. Nutritional and insulin regulation of fatty acid synthetase and leptin gene expression through ADD1/SREBP1. *The Journal of Clinical Investigation*, 101(1):1–9.

Kim, Y.-K.
2015. Extracellular microRNAs as Biomarkers in Human Disease. *Chonnam Medical Journal*, 51(2):51–57.

Kirchner, H., I. Sinha, H. Gao, M. A. Ruby, M. Schönke, J. M. Lindvall, R. Barrès, A. Krook, E. Näslund, K. Dahlman-Wright, and J. R. Zierath
2016. Altered DNA methylation of glycolytic and lipogenic genes in liver from obese and type 2 diabetic patients. *Molecular Metabolism*, 5(3):171–183.

Kitamura, E., J. Igarashi, A. Morohashi, N. Hida, T. Oinuma, N. Nemoto, F. Song, S. Ghosh, W. A. Held, C. Yoshida-Noro, and H. Nagase
2007. Analysis of tissue-specific differentially methylated regions (TDMs) in humans. *Genomics*, 89(3):326–337.

Kitano, M. and P. M. Bloomston
2016. Hepatic Stellate Cells and microRNAs in Pathogenesis of Liver Fibrosis. *Journal of Clinical Medicine*, 5(3).

Klug, M. and M. Rehli
2006. Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. *Epigenetics*, 1(3):127–130.

Knebel, B., J. Haas, S. Hartwig, S. Jacob, C. Köllmer, U. Nitzgen, D. Muller–Wieland, and J. Kotzka
2012. Liver-Specific Expression of Transcriptionally Active SREBP-1c Is Associated with Fatty Liver and Increased Visceral Fat Mass. *PLoS ONE*, 7(2):e31812.

Kota, S. K., S. K. Kota, S. Jammula, S. V. S. Krishna, and K. D. Modi
2012. Hypertriglyceridemia-induced recurrent acute pancreatitis: A case-based review. *Indian Journal of Endocrinology and Metabolism*, 16(1):141–143.

Krause, C., H. Sievert, C. Geißler, M. Grohs, A. T. El Gammal, S. Wolter, O. Ohlei, F. Kilpert, U. M. Krämer, M. Kasten, C. Klein, G. E. Brabant, O. Mann, H. Lehnert, and H. Kirchner
2019. Critical evaluation of the DNA-methylation markers ABCG1 and SREBF1 for Type 2 diabetes stratification. *Epigenomics*, Pp. epi–2018–0159.

Kriebel, J., C. Herder, W. Rathmann, S. Wahl, S. Kunze, S. Molnos, N. Volkova, K. Schramm, M. Carstensen-Kirberg, M. Waldenberger, C. Gieger, A. Peters, T. Illig, H. Prokisch, M. Roden, and H. Grallert
2016. Association between DNA Methylation in Whole Blood and Measures of Glucose Metabolism: KORA F4 Study. *PLOS ONE*, 11(3):e0152314.

Krotkiewski, M.
2000. Thyroid hormones and treatment of obesity. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, 24 Suppl 2:S116–119.

Krützfeldt, J., N. Rajewsky, R. Braich, K. G. Rajeev, T. Tuschl, M. Manoharan, and M. Stoffel
2005. Silencing of microRNAs in vivo with 'antagomirs'. *Nature*, 438(7068):685–689.

Kubota, N., T. Kubota, S. Itoh, H. Kumagai, H. Kozono, I. Takamoto, T. Mineyama, H. Ogata, K. Tokuyama, M. Ohsugi, T. Sasako, M. Moroi, K. Sugi, S. Kakuta, Y. Iwakura, T. Noda, S. Ohnishi, R. Nagai, K. Tobe, Y. Terauchi, K. Ueki, and T. Kadowaki
2008. Dynamic Functional Relay between Insulin Receptor Substrate 1 and 2 in Hepatic Insulin Signaling during Fasting and Feeding. *Cell Metabolism*, 8(1):49–64.

Kubota, N., T. Kubota, E. Kajiwara, T. Iwamura, H. Kumagai, T. Watanabe, M. Inoue, I. Takamoto, T. Sasako, K. Kumagai, M. Kohjima, M. Nakamuta, M. Moroi, K. Sugi, T. Noda, Y. Terauchi, K. Ueki, and T. Kadowaki
2016. Differential hepatic distribution of insulin receptor substrates causes selective insulin resistance in diabetes and obesity. *Nature Communications*, 7:12977.

Kulkarni, H., M. Z. Kos, J. Neary, T. D. Dyer, J. W. Kent, H. H. Göring, S. A. Cole, A. G. Comuzzie, L. Almasy, M. C. Mahaney, J. E. Curran, J. Blangero, and M. A. Carless
2015. Novel epigenetic determinants of type 2 diabetes in Mexican-American families. *Human Molecular Genetics*, 24(18):5330–5344.

Kuryłowicz, A., Z. Wicik, M. Owczarz, M. I. Jonas, M. Kotlarek, M. Świerniak, W. Lisik, M. Jonas, B. Noszczyk, and M. Puzianowska-Kuźnicka
2017. NGS Reveals Molecular Pathways Affected by Obesity and Weight Loss-Related Changes in miRNA Levels in Adipose Tissue. *International Journal of Molecular Sciences*, 19(1).

Lango, H., the U.K. Type 2 Diabetes Genetics Consortium, C. N. Palmer, A. D. Morris, E. Zeggini, A. T. Hattersley, M. I. McCarthy, T. M. Frayling, and M. N. Weedon
2008. Assessing the Combined Impact of 18 Common Genetic Variants of Modest Effect Sizes on Type 2 Diabetes Risk. *Diabetes*, 57(11):3129–3135.

Lawan, A. and A. M. Bennett
2017. Mitogen-Activated Protein Kinase Regulation in Hepatic Metabolism. *Trends in Endocrinology & Metabolism*, 28(12):868–878.

Le Teuff, G., C. Quantin, A. Venot, E. Walter, and J. Coste
2005. Improving model robustness with bootstrapping – application to optimal discriminant analysis for ordinal responses (ODAO). *Methods of Information in Medicine*, 44(5):704–711.

Lebedev, A., E. Westman, G. Van Westen, M. Kramberger, A. Lundervold, D. Aarsland, H. Soininen, I. Kłoszewska, P. Mecocci, M. Tsolaki, B. Vellas, S. Lovestone, and A. Simmons
2014. Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage : Clinical*, 6:115–125.

Lee, R. C., R. L. Feinbaum, and V. Ambros
1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75(5):843–854.

Lee, Y., J. E. Dominy, Y. J. Choi, M. Jurczak, N. Tolliday, J. P. Camporez, H. Chim, J.-H. Lim, H.-B. Ruan, X. Yang, F. Vazquez, P. Sicinski, G. I. Shulman, and P. Puigserver
2014. Cyclin D1–Cdk4 controls glucose metabolism independently of cell cycle progression. *Nature*, 510(7506):547–551.

Leti, F., I. Malenica, M. Doshi, A. Courtright, K. Van Keuren-Jensen, C. Legendre, C. D. Still, G. S. Gerhard, and J. K. DiStefano
2015. High-throughput sequencing reveals altered expression of hepatic microRNAs in nonalcoholic fatty liver disease-related fibrosis. *Translational Research: The Journal of Laboratory and Clinical Medicine*, 166(3):304–314.

Leung, W. K. C., M. He, A. W. H. Chan, P. T. Y. Law, and N. Wong
2015. Wnt/beta-Catenin activates MiR-183/96/182 expression in hepatocellular carcinoma that promotes cell invasion. *Cancer Letters*, 362(1):97–105.

Li, N., C.-C. Nan, X.-Y. Zhong, J.-Q. Weng, H.-D. Fan, H.-P. Sun, S. Tang, L. Shi, and S.-X. Huang
2018. miR-182-5p Promotes Growth in Oral Squamous Cell Carcinoma by Inhibiting CAMK2n1. *Cellular Physiology and Biochemistry*, 49(4):1329–1341.

Li, S., X. Chen, H. Zhang, X. Liang, Y. Xiang, C. Yu, K. Zen, Y. Li, and C.-Y. Zhang
2009. Differential expression of microRNAs in mouse liver under aberrant energy metabolic status. *Journal of Lipid Research*, 50(9):1756–1765.

Li, X., J. Wang, W. Li, Y. Xu, D. Shao, Y. Xie, W. Xie, T. Kubota, H. Narimatsu, and Y. Zhang
2012. Characterization of ppGalNAc-T18, a member of the vertebrate-specific Y subfamily of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology*, 22(5):602–615.

Lin, X., K. Song, N. Lim, X. Yuan, T. Johnson, A. Abderrahmani, P. Vollenweider, H. Stirnadel, S. S. Sundseth, E. Lai, D. K. Burns, L. T. Middleton, A. D. Roses, P. M. Matthews, G. Waeber, L. Cardon, D. M. Waterworth, and V. Mooser
2009. Risk prediction of prevalent diabetes in a Swiss population using a weighted genetic score—the CoLaus Study. *Diabetologia*, 52(4):600–608.

Lindholm, M. E., F. Marabita, D. Gomez-Cabrero, H. Rundqvist, T. J. Ekström, J. Tegnér, and C. J. Sundberg
2014. An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training. *Epigenetics*, 9(12):1557–1569.

Lirun, K., M. Sewe, and W. Yong
2015. A Pilot Study: The Effect of Roux-en-Y Gastric Bypass on the Serum MicroRNAs of the Type 2 Diabetes Patient. *Obesity Surgery*, 25(12):2386–2392.

Liu, C., C. Jiao, K. Wang, and N. Yuan
2018. DNA Methylation and Psychiatric Disorders. In *Progress in Molecular Biology and Translational Science*, volume 157, Pp. 175–232. Elsevier.

Liu, K. W., L. K. Dai, and W. Jean
2006. Metformin-related vitamin B12 deficiency. *Age and Ageing*, 35(2):200–201.

Liu, M., M. Wang, J. Wang, and D. Li
2013a. Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical*, 177:970–980.

Liu, Q., S. Bengmark, and S. Qu
2010. The role of hepatic fat accumulation in pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Lipids in Health and Disease*, 9:42.

Liu, Y., M. J. Aryee, L. Padyukov, M. D. Fallin, E. Hesselberg, A. Runarsson, L. Reinius, N. Acevedo, M. Taub, M. Ronninger, K. Shchetynsky, A. Scheynius, J. Kere, L. Alfredsson, L. Klareskog, T. J. Ekström, and A. P. Feinberg
2013b. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nature Biotechnology*, 31(2):142–147.

Lopez, Y. N., P. Coen, B. Goodpaster, and A. Seyhan
2017. Gastric bypass surgery with exercise alters plasma microRNAs that predict improvements in cardiometabolic risk. *International journal of obesity (2005)*, 41(7):1121–1130.

Loria, P., A. Lonardo, and F. Anania
2013. Liver and diabetes. A vicious circle: Liver and T2d. *Hepatology Research*, 43(1):51–64.

Lu, C., Z. Liao, M. Cai, and G. Zhang
2017. MicroRNA-320a downregulation mediates human liver cancer cell proliferation through the Wnt/beta-catenin signaling pathway. *Oncology Letters*, 13(2):573–578.

Lu, M., M. Wan, K. F. Leavens, Q. Chu, B. R. Monks, S. Fernandez, R. S. Ahima, K. Ueki, C. R. Kahn, and M. J. Birnbaum
2012. Insulin regulates liver metabolism in vivo in the absence of hepatic Akt and Foxo1. *Nature Medicine*, 18(3):388–395.

Luna, J. M., T. K. H. Scheel, T. Danino, K. S. Shaw, A. Mele, J. J. Fak, E. Nishiuchi, C. N. Takacs, M. T. Catanese, Y. P. de Jong, I. M. Jacobson, C. M. Rice, and R. B. Darnell
2015. Hepatitis C virus RNA functionally sequesters miR-122. *Cell*, 160(6):1099–1110.

Madsen, L. R., L. M. Baggesen, B. Richelsen, and R. W. Thomsen
2019. Effect of Roux-en-Y gastric bypass surgery on diabetes remission and complications in individuals with type 2 diabetes: a Danish population-based matched cohort study. *Diabetologia*, 62(4):611–620.

Mah, S. M., C. Buske, R. K. Humphries, and F. Kuchenbauer
2010. miRNA*: a passenger stranded in RNA-induced silencing complex? *Critical Reviews in Eukaryotic Gene Expression*, 20(2):141–148.

Mahajan, A., D. Taliun, M. Thurner, N. R. Robertson, J. M. Torres, N. W. Rayner, A. J. Payne, V. Steinthorsdottir, R. A. Scott, N. Grarup, J. P. Cook, E. M. Schmidt, M. Wuttke, C. Sarnowski, R. Mägi, J. Nano, C. Gieger, S. Trompet, C. Lecoeur, M. H. Preuss, B. P. Prins, X. Guo, L. F. Bielak, J. E. Below, D. W. Bowden, J. C. Chambers, Y. J. Kim, M. C. Y. Ng, L. E. Petty, X. Sim, W. Zhang, A. J. Bennett, J. Bork-Jensen, C. M. Brummett, M. Canouil, K.-U. E. Kardt, K. Fischer, S. L. R. Kardia, F. Kronenberg, K. Läll, C.-T. Liu, A. E. Locke, J. Luan, I. Ntalla, V. Nylander, S. Schönherr, C. Schurmann, L. Yengo, E. P. Bottinger, I. Brandslund, C. Christensen, G. Dedoussis, J. C. Florez, I. Ford, O. H. Franco, T. M. Frayling, V. Giedraitis, S. Hackinger, A. T. Hattersley, C. Herder, M. A. Ikram, M. Ingelsson, M. E. Jørgensen, T. Jørgensen, J. Kriebel, J. Kuusisto, S. Ligthart, C. M. Lindgren, A. Linneberg, V. Lyssenko, V. Mamakou, T. Meitinger, K. L. Mohlke, A. D. Morris, G. Nadkarni, J. S. Pankow, A. Peters, N. Sattar, A. Stančáková, K. Strauch, K. D. Taylor, B. Thorand, G. Thorleifsson, U. Thorsteinsdottir, J. Tuomilehto, D. R. Witte, J. Dupuis, P. A. Peyser, E. Zeggini, R. J. F. Loos, P. Froguel, E. Ingelsson, L. Lind, L. Groop, M. Laakso, F. S. Collins, J. W. Jukema, C. N. A. Palmer, H. Grallert, A. Metspalu, A. Dehghan, A. Köttgen, G. R. Abecasis, J. B. Meigs, J. I. Rotter, J. Marchini, O. Pedersen, T. Hansen, C. Langenberg, N. J. Wareham, K. Stefansson, A. L. Gloyn, A. P. Morris, M. Boehnke, and M. I. McCarthy
2018. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nature Genetics*, 50(11):1505.

Mani, A., J. Radhakrishnan, H. Wang, A. Mani, M.-A. Mani, C. Nelson-Williams, K. S. Carew, S. Mane, H. Najmabadi, D. Wu, and R. P. Lifton
2007. LRP6 mutation in a family with early coronary disease and metabolic risk factors. *Science (New York, N.Y.)*, 315(5816):1278–1282.

Matsumoto, M., S. Han, T. Kitamura, and D. Accili
2006. Dual role of transcription factor FoxO1 in controlling hepatic insulin sensitivity and lipid metabolism. *The Journal of Clinical Investigation*, 116(9):2464–2472.

Matsusue, K., O. Gavrilova, G. Lambert, H. B. Brewer, J. M. Ward, Y. Inoue, D. LeRoith, and F. J. Gonzalez
2004. Hepatic CCAAT/enhancer binding protein alpha mediates induction of lipogenesis and regulation of glucose homeostasis in leptin-deficient mice. *Molecular Endocrinology (Baltimore, Md.)*, 18(11):2751–2764.

Matthews, D. R., J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher, and R. C. Turner
1985. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*, 28(7):412–419.

Maunakea, A. K., R. P. Nagarajan, M. Bilenky, T. J. Ballinger, C. D'Souza, S. D. Fouse, B. E. Johnson, C. Hong, C. Nielsen, Y. Zhao, G. Turecki, A. Delaney, R. Varhol, N. Thiessen, K. Shchors, V. M. Heine, D. H. Rowitch, X. Xing, C. Fiore, M. Schillebeeckx, S. J. M. Jones, D. Haussler, M. A. Marra, M. Hirst, T. Wang, and J. F. Costello
2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, 466(7303):253–257.

Meigs, J. B., P. Shrader, L. M. Sullivan, J. B. McAteer, C. S. Fox, J. Dupuis, A. K. Manning, J. C. Florez, P. W. Wilson, R. B. D'Agostino, and L. A. Cupples
2008. Genotype Score in Addition to Common Risk Factors for Prediction of Type 2 Diabetes. *New England Journal of Medicine*, 359(21):2208–2219.

Meijer, H. A., E. M. Smith, and M. Bushell
2014. Regulation of miRNA strand selection: follow the leader? *Biochemical Society Transactions*, 42(4):1135–1140.

Michel, C. I., C. L. Holley, B. S. Scruggs, R. Sidhu, R. T. Brookheart, L. L. Listenberger, M. A. Behlke, D. S. Ory, and J. E. Schaffer
2011. Small nucleolar RNAs U32a, U33 and U35a are critical mediators of metabolic stress. *Cell metabolism*, 14(1):33–44.

Mill, J., T. Tang, Z. Kaminsky, T. Khare, S. Yazdanpanah, L. Bouchard, P. Jia, A. Assadzadeh, J. Flanagan, A. Schumacher, S.-C. Wang, and A. Petronis
2008. Epigenomic Profiling Reveals DNA-Methylation Changes Associated with Major Psychosis. *The American Journal of Human Genetics*, 82(3):696–711.

Miyaaki, H., T. Ichikawa, Y. Kamo, N. Taura, T. Honda, H. Shibata, M. Milazzo, F. Fornari, L. Gramantieri, L. Bolondi, and K. Nakao
2014. Significance of serum and hepatic microRNA-122 levels in patients with non-alcoholic fatty liver disease. *Liver International: Official Journal of the International Association for the Study of the Liver*, 34(7):e302–307.

Moran, S., C. Arribas, and M. Esteller
2015. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics*, 8(3):389–399.

Moszyńska, A., M. Gebert, J. F. Collawn, and R. Bartoszewski
2017. SNPs in microRNA target sites and their potential role in human disease. *Open Biology*, 7(4).

Nakatochi, M., S. Ichihara, K. Yamamoto, K. Naruse, S. Yokota, H. Asano, T. Matsubara, and M. Yokota
2017. Epigenome-wide association of myocardial infarction with DNA methylation sites at loci related to cardiovascular disease. *Clinical Epigenetics*, 9(1):54.

Ness, G. C.
1991. Thyroid hormone. Basis for its hypocholesterolemic effect. *The Journal of the Florida Medical Association*, 78(6):383–385.

Ness, G. C. and D. Lopez
1995. Transcriptional regulation of rat hepatic low-density lipoprotein receptor and cholesterol 7 alpha hydroxylase by thyroid hormone. *Archives of Biochemistry and Biophysics*, 323(2):404–408.

Núñez, K. G., J. Gonzalez-Rosario, P. T. Thevenot, and A. J. Cohen
2017. Cyclin D1 in the Liver: Role of Noncanonical Signaling in Liver Steatosis and Hormone Regulation. *The Ochsner Journal*, 17(1):56–65.

Ng, R., H. Wu, H. Xiao, X. Chen, H. Willenbring, C. J. Steer, and G. Song
2014. Inhibition of microRNA-24 expression in liver prevents hepatic lipid accumulation and hyperlipidemia: NG ET AL. *Hepatology*, 60(2):554–564.

Nie, J., C. Li, J. Li, X. Chen, and X. Zhong
2018. Analysis of non-alcoholic fatty liver disease microRNA expression spectra in rat liver tissues. *Molecular Medicine Reports*.

Nilsson, E., P. A. Jansson, A. Perfilyev, P. Volkov, M. Pedersen, M. K. Svensson, P. Poulsen, R. Ribel-Madsen, N. L. Pedersen, P. Almgren, J. Fadista, T. Rönn, B. Klarlund Pedersen, C. Scheele, A. Vaag, and C. Ling
2014. Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes*, 63(9):2962–2976.

Nilsson, E., A. Matte, A. Perfilyev, V. D. de Mello, P. Käkelä, J. Pihlajamäki, and C. Ling
2015. Epigenetic Alterations in Human Liver From Subjects With Type 2 Diabetes in Parallel With Reduced Folate Levels. *The Journal of Clinical Endocrinology & Metabolism*, 100(11):E1491–E1501.

Nitert, M. D., T. Dayeh, P. Volkov, T. Elgzyri, E. Hall, E. Nilsson, B. T. Yang, S. Lang, H. Parikh, Y. Wessman, H. Weishaupt, J. Attema, M. Abels, N. Wierup, P. Almgren, P.-A. Jansson, T. Rönn, O. Hansson, K.-F. Eriksson, L. Groop, and C. Ling
2012. Impact of an Exercise Intervention on DNA Methylation in Skeletal Muscle From First-Degree Relatives of Patients With Type 2 Diabetes. *Diabetes*, 61(12):3322–3332.

Nottrott, S., M. J. Simard, and J. D. Richter
2006. Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nature Structural & Molecular Biology*, 13(12):1108–1114.

O-Sullivan, I., W. Zhang, D. H. Wasserman, C. W. Liew, J. Liu, J. Paik, R. A. DePinho, D. B. Stolz, C. R. Kahn, M. W. Schwartz, and T. G. Unterman
2015. FoxO1 integrates direct and indirect effects of insulin on hepatic glucose production and glucose utilization. *Nature communications*, 6:7079.

Oberkofler, H., N. Fukushima, H. Esterbauer, F. Krempler, and W. Patsch
2002. Sterol regulatory element binding proteins: relationship of adipose tissue gene expression with obesity in humans. *Biochimica Et Biophysica Acta*, 1575(1-3):75–81.

Okonechnikov, K., O. Golosova, and M. Fursov
2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8):1166–1167.

Oppenheimer, J. H., H. L. Schwartz, J. T. Lane, and M. P. Thompson
1991. Functional relationship of thyroid hormone-induced lipogenesis, lipolysis, and thermogenesis in the rat. *The Journal of Clinical Investigation*, 87(1):125–132.

Orozco, L. D., C. Farrell, C. Hale, L. Rubbi, A. Rinaldi, M. Civelek, C. Pan, L. Lam, D. Montoya, C. Edillor, M. Seldin, M. Boehnke, K. L. Mohlke, S. Jacobsen, J. Kuusisto, M. Laakso, A. J. Lusis, and M. Pellegrini
2018. Epigenome-wide association in adipose tissue from the METSIM cohort. *Human Molecular Genetics*, 27(10):1830–1846.

Ortega, F. J., M. I. Cardona-Alvarado, J. M. Mercader, J. M. Moreno-Navarrete, M. Moreno, M. Sabater, N. Fuentes-Batllevell, E. Ramírez-Chávez, W. Ricart, J. Molina-Torres, E. L. Pérez-Luque, and J. M. Fernández-Real
2015. Circulating profiling reveals the effect of a polyunsaturated fatty acid-enriched diet on common microRNAs. *The Journal of Nutritional Biochemistry*, 26(10):1095–1101.

Ozaki, K., Y. Ohnishi, A. Iida, A. Sekine, R. Yamada, T. Tsunoda, H. Sato, H. Sato, M. Hori, Y. Nakamura, and T. Tanaka
2002. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nature Genetics*, 32(4):650–654.

Pagadala, M. R., C. O. Zein, S. Dasarathy, L. M. Yerian, R. Lopez, and A. J. McCullough
2012. Prevalence of hypothyroidism in nonalcoholic fatty liver disease. *Digestive Diseases and Sciences*, 57(2):528–534.

Palmer, A. K., T. Tchkonia, N. K. LeBrasseur, E. N. Chini, M. Xu, and J. L. Kirkland
2015. Cellular Senescence in Type 2 Diabetes: A Therapeutic Opportunity. *Diabetes*, 64(7):2289–2298.

Peeters, R. P., P. J. Wouters, E. Kaptein, H. van Toor, T. J. Visser, and G. Van den Berghe
2003. Reduced Activation and Increased Inactivation of Thyroid Hormone in Tissues of Critically Ill Patients. *The Journal of Clinical Endocrinology & Metabolism*, 88(7):3202–3211.

Perlman, R. L.
2016. Mouse models of human disease. *Evolution, Medicine, and Public Health*, 2016(1):170–176.

Peterson, S. M., J. A. Thompson, M. L. Ufkin, P. Sathyanarayana, L. Liaw, and C. B. Congdon
2014. Common features of microRNA target prediction tools. *Frontiers in Genetics*, 5.

Pfeiffer, L., S. Wahl, L. C. Pilling, E. Reischl, J. K. Sandling, S. Kunze, L. M. Holdt, A. Kretschmer, K. Schramm, J. Adamski, N. Klopp, T. Illig, s. K. Hedman, M. Roden, D. G. Hernandez, A. B. Singleton, W. E. Thasler, H. Grallert, C. Gieger, C. Herder, D. Teupser, C. Meisinger, T. D. Spector, F. Kronenberg, H. Prokisch, D. Melzer, A. Peters, P. Deloukas, L. Ferrucci, and M. Waldenberger
2015. DNA Methylation of Lipid-Related Genes Affects Blood Lipid Levels. *Circulation: Cardiovascular Genetics*, 8(2):334–342.

Pheiffer, C., R. T. Erasmus, A. P. Kengne, and T. E. Matsha
2016. Differential DNA methylation of microRNAs within promoters, intergenic and intragenic regions of type 2 diabetic, pre-diabetic and non-diabetic individuals. *Clinical Biochemistry*, 49(6):433–438.

Portillo-Sanchez, P., F. Bril, M. Maximos, R. Lomonaco, D. Biernacki, B. Orsak, S. Subbarayan, A. Webb, J. Hecht, and K. Cusi
2015. High Prevalence of Nonalcoholic Fatty Liver Disease in Patients With Type 2 Diabetes Mellitus and Normal Plasma Aminotransferase Levels. *The Journal of Clinical Endocrinology and Metabolism*, 100(6):2231–2238.

Pourhoseingholi, M. A., A. R. Baghestani, and M. Vahedi
2012. How to control confounding effects by statistical analysis. *Gastroenterology and Hepatology from Bed to Bench*, 5(2):79–83.

Prasad, R. B. and L. Groop
2015. Genetics of Type 2 Diabetes—Pitfalls and Possibilities. *Genes*, 6(1):87–123.

Rahimloo, P. and A. Jafarian
2016. Prediction of Diabetes by Using Artificial Neural Network, Logistic Regression Statistical Model and Combination of Them. *Bulletin de la Société Royale des Sciences de Liège*, 85:17.

Rakyan, V. K., T. A. Down, D. J. Balding, and S. Beck
2011. Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics*, 12(8):529–541.

Raman, J., Y. Guan, C. L. Perrine, T. A. Gerken, and L. A. Tabak
2012. UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases: completion of the family tree. *Glycobiology*, 22(6):768–777.

Rametta, R., E. Mozzi, P. Dongiovanni, B. M. Motta, M. Milano, G. Roviaro, S. Fargion, and L. Valenti
2013. Increased insulin receptor substrate 2 expression is associated with steatohepatitis and altered lipid metabolism in obese subjects. *International Journal of Obesity*, 37(7):986–992.

Rashid, S., T. Watanabe, T. Sakaue, and G. F. Lewis
2003. Mechanisms of HDL lowering in insulin resistant, hypertriglyceridemic states: the combined effect of HDL triglyceride enrichment and elevated hepatic lipase activity. *Clinical Biochemistry*, 36(6):421–429.

Ravelli, A., J. van der Meulen, R. Michels, C. Osmond, D. Barker, C. Hales, and O. Bleker
1998. Glucose tolerance in adults after prenatal exposure to famine. *The Lancet*, 351(9097):173–177.

Ray, A., M. K. James, S. Larochelle, R. P. Fisher, and S. W. Blain
2009. p27kip1 Inhibits Cyclin D-Cyclin-Dependent Kinase 4 by Two Independent Modes. *Molecular and Cellular Biology*, 29(4):986–999.

Reed, B. D., A. E. Charos, A. M. Szekely, S. M. Weissman, and M. Snyder
2008. Genome-Wide Occupancy of SREBP1 and Its Partners NFY and SP1 Reveals Novel Functional Roles and Combinatorial Regulation of Distinct Classes of Genes. *PLoS Genetics*, 4(7):e1000133.

Ribel-Madsen, R., M. F. Fraga, S. Jacobsen, J. Bork-Jensen, E. Lara, V. Calvanese, A. F. Fernandez, M. Friedrichsen, B. F. Vind, K. Højlund, H. Beck-Nielsen, M. Esteller, A. Vaag, and P. Poulsen
2012. Genome-Wide Analysis of DNA Methylation Differences in Muscle and Fat from Monozygotic Twins Discordant for Type 2 Diabetes. *PLOS ONE*, 7(12):e51302.

Richard, K., R. Hume, E. Kaptein, J. P. Sanders, H. van Toor, W. W. de Herder, J. C. den Hollander, E. P. Krenning, and T. J. Visser
1998. Ontogeny of Iodothyronine Deiodinases in Human Liver [1]. *The Journal of Clinical Endocrinology & Metabolism*, 83(8):2868–2874.

Rishi, V., P. Bhattacharya, R. Chatterjee, J. Rozenberg, J. Zhao, K. Glass, P. Fitzgerald, and C. Vinson
2010. CpG methylation of half-CRE sequences creates C/EBPalpha binding sites that activate some tissue-specific genes. *Proceedings of the National Academy of Sciences of the United States of America*, 107(47):20311–20316.

Roberts, L. M., K. Woodford, M. Zhou, D. S. Black, J. E. Haggerty, E. H. Tate, K. K. Grindstaff, W. Mengesha, C. Raman, and N. Zerangue
2008. Expression of the Thyroid Hormone Transporters Monocarboxylate Transporter-8 ( *SLC16A2* ) and Organic Ion Transporter-14 ( *SLCO1C1* ) at the Blood-Brain Barrier. *Endocrinology*, 149(12):6251–6261.

Robertson, K. D.
2005. DNA methylation and human disease. *Nature Reviews Genetics*, 6(8):597.

Roglic, G. and W. H. Organization, eds.
2016. *Global report on diabetes*. Geneva, Switzerland: World Health Organization. OCLC: ocn948336981.

Rougvie, A. E.
2001. Control of developmental timing in animals. *Nature Reviews Genetics*, 2:690.

Ryaboshapkina, M. and M. Hammar
2017. Human hepatic gene expression signature of non-alcoholic fatty liver disease progression, a meta-analysis. *Scientific Reports*, 7(1):12361.

Sala, P., R. S. M. de Miranda Torrinhas, D. C. Fonseca, G. R. Ravacci, D. L. Waitzberg, and D. Giannella-Neto
2017. Tissue-specific methylation profile in obese patients with type 2 diabetes before and after Roux-en-Y gastric bypass. *Diabetology & Metabolic Syndrome*, 9(1).

Samuel, V. and G. Shulman
2012. Mechanisms for Insulin Resistance: Common Threads and Missing Links. *Cell*, 148(5):852–871.

Santangelo, A., A. Tamanini, G. Cabrini, and M. C. Dechecchi
2017. Circulating microRNAs as emerging non-invasive biomarkers for gliomas. *Annals of Translational Medicine*, 5(13).

Savage, D. B.
2006. Reversal of diet-induced hepatic steatosis and hepatic insulin resistance by antisense oligonucleotide inhibitors of acetyl-CoA carboxylases 1 and 2. *Journal of Clinical Investigation*, 116(3):817–824.

Sayols-Baixeras, S., I. Subirana, C. Lluis-Ganella, F. Civeira, J. Roquer, A. Do, D. Absher, A. Cenarro, D. Muñoz, C. Soriano-Tárraga, J. Jiménez-Conde, J. M. Ordovas, M. Senti, S. Aslibekyan, J. Marrugat, D. K. Arnett, and R. Elosua
2016. Identification and validation of seven new loci showing differential DNA methylation related to serum lipid profile: an epigenome-wide approach. The REGICOR study. *Human Molecular Genetics*, 25(20):4556–4565.

Schübeler, D.
2015. Function and information content of DNA methylation. *Nature*, 517(7534):321–326.

Schwarzenbach, H., A. M. da Silva, G. Calin, and K. Pantel
2015. Data Normalization Strategies for MicroRNA Quantification. *Clinical Chemistry*, 61(11):1333–1342.

Sedgeman, L. R., C. Beysen, R. M. Allen, M. A. Ramirez Solano, S. M. Turner, and K. C. Vickers
2018. Intestinal bile acid sequestration improves glucose control by stimulating hepatic miR-182-5p in type 2 diabetes. *American Journal of Physiology. Gastrointestinal and Liver Physiology*.

Seillier, M., L. Pouyet, P. N'Guessan, M. Nollet, F. Capo, F. Guillaumond, L. Peyta, J.-F. Dumas, A. Varrault, G. Bertrand, S. Bonnafous, A. Tran, G. Meur, P. Marchetti, M. A. Ravier, S. Dalle, P. Gual, D. Muller, G. A. Rutter, S. Servais, J. L. Iovanna, and A. Carrier
2015. Defects in mitophagy promote redox-driven metabolic syndrome in the absence of TP53inp1. *EMBO Molecular Medicine*, 7(6):802–818.

Sekine, K., Y.-R. Chen, N. Kojima, K. Ogata, A. Fukamizu, and A. Miyajima
2007. Foxo1 links insulin signaling to C/EBPalpha and regulates gluconeogenesis during liver development. *The EMBO journal*, 26(15):3607–3615.

Shaikhina, T., D. Lowe, S. Daga, D. Briggs, R. Higgins, and N. Khovanova
2017. Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. *Biomedical Signal Processing and Control*.

Shayevitch, R., D. Askayo, I. Keydar, and G. Ast
2018. The importance of DNA methylation of exons on alternative splicing. *RNA (New York, N.Y.)*, 24(10):1351–1362.

Shi, W., Z. Zhang, B. Yang, H. Guo, L. Jing, T. Liu, Y. Luo, H. Liu, Y. Li, and Y. Gao
2017. Overexpression of microRNA let-7 correlates with disease progression and poor prognosis in hepatocellular carcinoma. *Medicine*, 96(32).

Shilin Zhao, Y. G.
2017. KEGGprofile.

Shimomura, I., H. Shimano, J. D. Horton, J. L. Goldstein, and M. S. Brown
1997. Differential expression of exons 1a and 1c in mRNAs for sterol regulatory element binding protein-1 in human and mouse organs and cultured cells. *Journal of Clinical Investigation*, 99(5):838–845.

Silvestri, E., A. Lombardi, P. de Lange, L. Schiavo, A. Lanni, F. Goglia, T. J. Visser, and M. Moreno
2008. Age-related changes in renal and hepatic cellular mechanisms associated with variations in rat serum thyroid hormone levels. *American Journal of Physiology-Endocrinology and Metabolism*, 294(6):E1160–E1168.

Singh, R. G., H. D. Yoon, S. D. Poppitt, L. D. Plank, and M. S. Petrov
2017. Ectopic fat accumulation in the pancreas and its biomarkers: A systematic review and meta-analysis. *Diabetes/Metabolism Research and Reviews*, 33(8).

Sinha, R. A., B. K. Singh, and P. M. Yen
2018. Direct effects of thyroid hormones on hepatic lipid metabolism. *Nature reviews. Endocrinology*, 14(5):259–269.

Sjoberg, R. J. and G. S. Kidd
1989. Pancreatic diabetes mellitus. *Diabetes Care*, 12(10):715–724.

Sladek, R., G. Rocheleau, J. Rung, C. Dina, L. Shen, D. Serre, P. Boutin, D. Vincent, A. Belisle, S. Hadjadj, B. Balkau, B. Heude, G. Charpentier, T. J. Hudson, A. Montpetit, A. V. Pshezhetsky, M. Prentki, B. I. Posner, D. J. Balding, D. Meyre, C. Polychronakos, and P. Froguel
2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*, 445(7130):881–885.

Sliwinska, A., M. A. Kasinska, and J. Drzewoski
2017. MicroRNAs and metabolic disorders – where are we heading? *Archives of Medical Science*, 4:885–896.

Sohel, M. H.
2016. Extracellular/Circulating MicroRNAs: Release Mechanisms, Functions and Challenges. *Achievements in the Life Sciences*, 10(2):175–186.

Soronen, J., H. Yki-Järvinen, Y. Zhou, S. Sädevirta, A.-P. Sarin, M. Leivonen, K. Sevastianova, J. Perttilä, P.-P. Laurila, A. Sigruener, G. Schmitz, and V. M. Olkkonen
2016. Novel hepatic microRNAs upregulated in human nonalcoholic fatty liver disease. *Physiological Reports*, 4(1):e12661.

Speliotes, E. K., C. J. Willer, S. I. Berndt, K. L. Monda, G. Thorleifsson, A. U. Jackson, H. L. Allen, C. M. Lindgren, J. Luan, R. Mägi, J. C. Randall, S. Vedantam, T. W. Winkler, L. Qi, T. Workalemahu, I. M. Heid, V. Steinthorsdottir, H. M. Stringham, M. N. Weedon, E. Wheeler, A. R. Wood, T. Ferreira, R. J. Weyant, A. V. Segrè, K. Estrada, L. Liang, J. Nemesh, J.-H. Park, S. Gustafsson, T. O. Kilpeläinen, J. Yang, N. Bouatia-Naji, T. Esko, M. F. Feitosa, Z. Kutalik, M. Mangino, S. Raychaudhuri, A. Scherag, A. V. Smith, R. Welch, J. H. Zhao, K. K. Aben, D. M. Absher, N. Amin, A. L. Dixon, E. Fisher, N. L. Glazer, M. E. Goddard, N. L. Heard-Costa, V. Hoesel, J.-J. Hottenga, s. Johansson, T. Johnson, S. Ketkar, C. Lamina, S. Li, M. F. Moffatt, R. H. Myers, N. Narisu, J. R. B. Perry, M. J. Peters, M. Preuss, S. Ripatti, F. Rivadeneira, C. Sandholt, L. J. Scott, N. J. Timpson, J. P. Tyrer, S. van Wingerden, R. M. Watanabe, C. C. White, F. Wiklund, C. Barlassina, D. I. Chasman, M. N. Cooper, J.-O. Jansson, R. W. Lawrence, N. Pellikka, I. Prokopenko, J. Shi, E. Thiering, H. Alavere, M. T. S. Alibrandi, P. Almgren, A. M. Arnold, T. Aspelund, L. D. Atwood, B. Balkau, A. J. Balmforth, A. J. Bennett, Y. Ben-Shlomo, R. N. Bergman, S. Bergmann, H. Biebermann, A. I. F. Blakemore, T. Boes, L. L. Bonnycastle, S. R. Bornstein, M. J. Brown, T. A. Buchanan, F. Busonero, H. Campbell, F. P. Cappuccio, C. Cavalcanti-Proença, Y.-D. I. Chen, C.-M. Chen, P. S. Chines, R. Clarke, L. Coin, J. Connell, I. N. M. Day, M. d. Heijer, J. Duan, S. Ebrahim, P. Elliott, R. Elosua, G. Eiriksdottir, M. R. Erdos, J. G. Eriksson, M. F. Facheris, S. B. Felix, P. Fischer-Posovszky, A. R. Folsom, N. Friedrich, N. B. Freimer, M. Fu, S. Gaget, P. V. Gejman, E. J. C. Geus, C. Gieger, A. P. Gjesing, A. Goel, P. Goyette, H. Grallert, J. Gräßler, D. M. Greenawalt, C. J. Groves, V. Gudnason, C. Guiducci, A.-L. Hartikainen, N. Hassanali, A. S. Hall, A. S. Havulinna, C. Hayward, A. C. Heath, C. Hengstenberg, A. A. Hicks, A. Hinney, A. Hofman, G. Homuth, J. Hui, W. Igl, C. Iribarren, B. Isomaa, K. B. Jacobs, I. Jarick, E. Jewell, U. John, T. Jørgensen, P. Jousilahti, A. Jula, M. Kaakinen, E. Kajantie, L. M. Kaplan, S. Kathiresan, J. Kettunen, L. Kinnunen, J. W. Knowles, I. Kolcic, I. R. König, S. Koskinen, P. Kovacs, J. Kuusisto, P. Kraft, K. Kvaløy, J. Laitinen, O. Lantieri, C. Lanzani, L. J. Launer, C. Lecoeur, T. Lehtimäki, G. Lettre, J. Liu, M.-L. Lokki, M. Lorentzon, R. N. Luben, B. Ludwig, P. Manunta, D. Marek, M. Marre, N. G. Martin, W. L. McArdle, A. McCarthy, B. McKnight, T. Meitinger, O. Melander, D. Meyre, K. Midthjell, G. W. Montgomery, M. A. Morken, A. P.

Morris, R. Mulic, J. S. Ngwa, M. Nelis, M. J. Neville, D. R. Nyholt, C. J. O'Donnell, S. O'Rahilly, K. K. Ong, B. Oostra, G. Paré, A. N. Parker, M. Perola, I. Pichler, K. H. Pietiläinen, C. G. P. Platou, O. Polasek, A. Pouta, S. Rafelt, O. Raitakari, N. W. Rayner, M. Ridderstråle, W. Rief, A. Ruokonen, N. R. Robertson, P. Rzehak, V. Salomaa, A. R. Sanders, M. S. Sandhu, S. Sanna, J. Saramies, M. J. Savolainen, S. Scherag, S. Schipf, S. Schreiber, H. Schunkert, K. Silander, J. Sinisalo, D. S. Siscovick, J. H. Smit, N. Soranzo, U. Sovio, J. Stephens, I. Surakka, A. J. Swift, M.-L. Tammesoo, J.-C. Tardif, M. Teder-Laving, T. M. Teslovich, J. R. Thompson, B. Thomson, A. Tönjes, T. Tuomi, J. B. J. van Meurs, G.-J. van Ommen, V. Vatin, J. Viikari, S. Visvikis-Siest, V. Vitart, C. I. G. Vogel, B. F. Voight, L. L. Waite, H. Wallaschofski, G. B. Walters, E. Widen, S. Wiegand, S. H. Wild, G. Willemsen, D. R. Witte, J. C. Witteman, J. Xu, Q. Zhang, L. Zgaga, A. Ziegler, P. Zitting, J. P. Beilby, I. S. Farooqi, J. Hebebrand, H. V. Huikuri, A. L. James, M. Kähönen, D. F. Levinson, F. Macciardi, M. S. Nieminen, C. Ohlsson, L. J. Palmer, P. M. Ridker, M. Stumvoll, J. S. Beckmann, H. Boeing, E. Boerwinkle, D. I. Boomsma, M. J. Caulfield, S. J. Chanock, F. S. Collins, L. A. Cupples, G. D. Smith, J. Erdmann, P. Froguel, H. Grönberg, U. Gyllensten, P. Hall, T. Hansen, T. B. Harris, A. T. Hattersley, R. B. Hayes, J. Heinrich, F. B. Hu, K. Hveem, T. Illig, M.-R. Jarvelin, J. Kaprio, F. Karpe, K.-T. Khaw, L. A. Kiemeney, H. Krude, M. Laakso, D. A. Lawlor, A. Metspalu, P. B. Munroe, W. H. Ouwehand, O. Pedersen, B. W. Penninx, A. Peters, P. P. Pramstaller, T. Quertermous, T. Reinehr, A. Rissanen, I. Rudan, N. J. Samani, P. E. H. Schwarz, A. R. Shuldiner, T. D. Spector, J. Tuomilehto, M. Uda, A. Uitterlinden, T. T. Valle, M. Wabitsch, G. Waeber, N. J. Wareham, H. Watkins, J. F. Wilson, A. F. Wright, M. C. Zillikens, N. Chatterjee, S. A. McCarroll, S. Purcell, E. E. Schadt, P. M. Visscher, T. L. Assimes, I. B. Borecki, P. Deloukas, C. S. Fox, L. C. Groop, T. Haritunians, D. J. Hunter, R. C. Kaplan, K. L. Mohlke, J. R. O'Connell, L. Peltonen, D. Schlessinger, D. P. Strachan, C. M. van Duijn, H.-E. Wichmann, T. M. Frayling, U. Thorsteinsdottir, G. R. Abecasis, I. Barroso, M. Boehnke, K. Stefansson, K. E. 2010. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics*, 42(11):937–948.

Sterling, R. K., E. Lissen, N. Clumeck, R. Sola, M. C. Correa, J. Montaner, M. S. Sulkowski, F. J. Torriani, D. T. Dieterich, D. L. Thomas, D. Messinger, M. Nelson, and APRICOT Clinical Investigators
2006. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology*, 43(6):1317–1325.

Szabo, G. and S. Bala
2013. MicroRNAs in liver disease. *Nature reviews. Gastroenterology & hepatology*, 10(9):542–552.

Tabaei, B. P. and W. H. Herman
2002. A Multivariate Logistic Regression Equation to Screen for Diabetes: Development and validation. *Diabetes Care*, 25(11):1999–2003.

Tamai, K., M. Semenov, Y. Kato, R. Spokony, C. Liu, Y. Katsuyama, F. Hess, J.-P. Saint-Jeannet, and X. He
2000. LDL-receptor-related proteins in Wnt signal transduction. *Nature*, 407(6803):530–535.

Tamayo, T., R. Brinks, A. Hoyer, O. Kuß, and W. Rathmann
2016. The Prevalence and Incidence of Diabetes in Germany. *Deutsches Ärzteblatt International*, 113(11):177–182.

Tan, Y., G. Ge, T. Pan, D. Wen, and J. Gan
2014. A Pilot Study of Serum MicroRNAs Panel as Potential Biomarkers for Diagnosis of Nonalcoholic Fatty Liver Disease. *PLOS ONE*, 9(8):e105192.

Tarling, E. J. and P. A. Edwards
2011. ATP binding cassette transporter G1 (ABCG1) is an intracellular sterol transporter. *Proceedings of the National Academy of Sciences*, 108(49):19719–19724.

Teng, W., W. Yin, L. Zhao, C. Ma, J. Huang, and F. Ren
2018. Resveratrol metabolites ameliorate insulin resistance in HepG2 hepatocytes by modulating IRS-1/AMPK. *RSC Advances*, 8(63):36034–36042.

The BLUEPRINT Consortium, C. Bock, F. Halbritter, F. J. Carmona, S. Tierling, P. Datlinger, Y. Assenov, M. Berdasco, A. K. Bergmann, K. Booher, F. Busato, M. Campan, C. Dahl, C. M. Dahmcke, D. Diep, A. F. Fernández, C. Gerhauser, A. Haake, K. Heilmann, T. Holcomb, D. Hussmann, M. Ito, R. Kläver, M. Kreutz, M. Kulis, V. Lopez, S. S. Nair, D. S. Paul, N. Plongthongkum, W. Qu, A. C. Queirós, F. Reinicke, G. Sauter, T. Schlomm, A. Statham, C. Stirzaker, R. Strogantsev, R. G. Urdinguio, K. Walter, D. Weichenhan, D. J. Weisenberger, S. Beck, S. J. Clark, M. Esteller, A. C. Ferguson-Smith, M. F. Fraga, P. Guldberg, L. L. Hansen, P. W. Laird, J. I. Martín-Subero, A. O. H. Nygren, R. Peist, C. Plass, D. S. Shames, R. Siebert, X. Sun, J. Tost, J. Walter, and K. Zhang
2016. Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. *Nature Biotechnology*, 34(7):726–737.

Thompson, M. D. and S. P. S. Monga
2007. WNT/beta-catenin signaling in liver health and disease. *Hepatology (Baltimore, Md.)*, 45(5):1298–1305.

Tiikkainen, M., A.-M. Hakkinen, E. Korsheninnikova, T. Nyman, S. Makimattila, and H. Yki-Jarvinen
2004. Effects of Rosiglitazone and Metformin on Liver Fat Content, Hepatic Insulin Resistance, Insulin Clearance, and Gene Expression in Adipose Tissue in Patients With Type 2 Diabetes. *Diabetes*, 53(8):2169–2176.

Tomar, A. S., D. S. P. Tallapragada, S. S. Nongmaithem, S. Shrestha, C. S. Yajnik, and G. R. Chandak
2015. Intrauterine Programming of Diabetes and Adiposity. *Current Obesity Reports*, 4(4):418–428.

Trajkovski, M., J. Hausser, J. Soutschek, B. Bhat, A. Akin, M. Zavolan, M. H. Heim, and M. Stoffel
2011. MicroRNAs 103 and 107 regulate insulin sensitivity. *Nature*, 474(7353):649–653.

Truong, V., S. Huang, J. Dennis, M. Lemire, N. Zwingerman, D. Aïssi, I. Kassam, C. Perret, P. Wells, P.-E. Morange, M. Wilson, D.-A. Trégouët, and F. Gagnon
2017. Blood triglyceride levels are associated with DNA methylation at the serine metabolism gene PHGDH. *Scientific Reports*, 7(1):11207.

Vaag, A. A., L. G. Grunnet, G. P. Arora, and C. Brøns
2012. The thrifty phenotype hypothesis revisited. *Diabetologia*, 55(8):2085–2088.

van der Horst, A. and B. M. T. Burgering
2007. Stressing the role of FoxO proteins in lifespan and disease. *Nature Reviews. Molecular Cell Biology*, 8(6):440–450.

van der Ree, M. H., A. J. van der Meer, A. C. van Nuenen, J. de Bruijne, S. Ottosen, H. L. Janssen, N. A. Kootstra, and H. W. Reesink
2016. Miravirsen dosing in chronic hepatitis C patients results in decreased microRNA-122 levels without affecting other microRNAs in plasma. *Alimentary Pharmacology & Therapeutics*, 43(1):102–113.

van Hoek, M., A. Dehghan, J. C. Witteman, C. M. van Duijn, A. G. Uitterlinden, B. A. Oostra, A. Hofman, E. J. Sijbrands, and A. C. J. Janssens
2008. Predicting Type 2 Diabetes Based on Polymorphisms From Genome-Wide Association Studies: A Population-Based Study. *Diabetes*, 57(11):3122–3128.

Vatner, D. F., D. Weismann, S. A. Beddow, N. Kumashiro, D. M. Erion, X.-H. Liao, G. J. Grover, P. Webb, K. J. Phillips, R. E. Weiss, J. S. Bogan, J. Baxter, G. I. Shulman, and V. T. Samuel
2013. Thyroid hormone receptor-beta agonists prevent hepatic steatosis in fat-fed rats but impair insulin sensitivity via discrete pathways. *American Journal of Physiology-Endocrinology and Metabolism*, 305(1):E89–E100.

Veloso, A., K. S. Kirkconnell, B. Magnuson, B. Biewen, M. T. Paulsen, T. E. Wilson, and M. Ljungman
2014. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Research*, 24(6):896–905.

Visser, W. E., C. R. Bombardieri, C. Zevenbergen, S. Barnhoorn, A. Ottaviani, I. van der Pluijm, R. Brandt, E. Kaptein, R. van Heerebeek, H. van Toor, G. A. Garinis, R. P. Peeters, M. Medici, W. van Ham, W. P. Vermeij, M. C. de Waard, R. R. de Krijger, A. Boelen, J. Kwakkel, J. J. Kopchick, E. O. List, J. P. M. Melis, V. M. Darras, M. E. T. Dollé, G. T. J. van der Horst, J. H. J. Hoeijmakers, and T. J. Visser
2016. Tissue-Specific Suppression of Thyroid Hormone Signaling in Various Mouse Models of Aging. *PLOS ONE*, 11(3):e0149941.

Voight, B. F., L. J. Scott, V. Steinthorsdottir, A. P. Morris, C. Dina, R. P. Welch, E. Zeggini, C. Huth, Y. S. Aulchenko, G. Thorleifsson, L. J. McCulloch, T. Ferreira, H. Grallert, N. Amin, G. Wu, C. J. Willer, S. Raychaudhuri, S. A. McCarroll, C. Langenberg, O. M. Hofmann, J. Dupuis, L. Qi, A. V. Segrè, M. van Hoek, P. Navarro, K. Ardlie, B. Balkau, R. Benediktsson, A. J. Bennett, R. Blagieva, E. Boerwinkle, L. L. Bonnycastle, K. Bengtsson Boström, B. Bravenboer, S. Bumpstead, N. P. Burtt, G. Charpentier, P. S. Chines, M. Cornelis, D. J. Couper, G. Crawford, A. S. F. Doney, K. S. Elliott, A. L. Elliott, M. R. Erdos, C. S. Fox, C. S. Franklin, M. Ganser, C. Gieger, N. Grarup, T. Green, S. Griffin, C. J. Groves, C. Guiducci, S. Hadjadj, N. Hassanali, C. Herder, B. Isomaa, A. U. Jackson, P. R. V. Johnson, T. Jørgensen, W. H. L. Kao, N. Klopp, A. Kong, P. Kraft, J. Kuusisto, T. Lauritzen, M. Li, A. Lieverse, C. M. Lindgren, V. Lyssenko, M. Marre, T. Meitinger, K. Midthjell, M. A. Morken, N. Narisu, P. Nilsson, K. R. Owen, F. Payne, J. R. B. Perry, A.-K. Petersen, C. Platou, C. Proença, I. Prokopenko, W. Rathmann, N. W. Rayner, N. R. Robertson, G. Rocheleau, M. Roden, M. J. Sampson, R. Saxena, B. M. Shields, P. Shrader, G. Sigurdsson, T. Sparsø, K. Strassburger, H. M.

Stringham, Q. Sun, A. J. Swift, B. Thorand, J. Tichet, T. Tuomi, R. M. van Dam, T. W. van Haeften, T. van Herpt, J. V. van Vliet-Ostaptchouk, G. B. Walters, M. N. Weedon, C. Wijmenga, J. Witteman, R. N. Bergman, S. Cauchi, F. S. Collins, A. L. Gloyn, U. Gyllensten, T. Hansen, W. A. Hide, G. A. Hitman, A. Hofman, D. J. Hunter, K. Hveem, M. Laakso, K. L. Mohlke, A. D. Morris, C. N. A. Palmer, P. P. Pramstaller, I. Rudan, E. Sijbrands, L. D. Stein, J. Tuomilehto, A. Uitterlinden, M. Walker, N. J. Wareham, R. M. Watanabe, G. R. Abecasis, B. O. Boehm, H. Campbell, M. J. Daly, A. T. Hattersley, F. B. Hu, J. B. Meigs, J. S. Pankow, O. Pedersen, H.-E. Wichmann, I. Barroso, J. C. Florez, T. M. Frayling, L. Groop, R. Sladek, U. Thorsteinsdottir, J. F. Wilson, T. Illig, P. Froguel, C. M. van Duijn, K. Stefansson, D. Altshuler, M. Boehnke, M. I. McCarthy, MAGIC investigators, and GIANT Consortium
2010. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature Genetics*, 42(7):579–589.

Volkmar, M., S. Dedeurwaerder, D. A. Cunha, M. N. Ndlovu, M. Defrance, R. Deplus, E. Calonne, U. Volkmar, M. Igoillo-Esteve, N. Naamane, S. Del Guerra, M. Masini, M. Bugliani, P. Marchetti, M. Cnop, D. L. Eizirik, and F. Fuks
2012. DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients: DNA methylation profiling of type 2 diabetic islets. *The EMBO Journal*, 31(6):1405–1426.

Volkov, P., K. Bacos, J. K. Ofori, J. L. S. Esguerra, L. Eliasson, T. Rönn, and C. Ling
2017. Whole-Genome Bisulfite Sequencing of Human Pancreatic Islets Reveals Novel Differentially Methylated Regions in Type 2 Diabetes Pathogenesis. *Diabetes*, 66(4):1074–1085.

Wahl, S., A. Drong, B. Lehne, M. Loh, W. R. Scott, S. Kunze, P.-C. Tsai, J. S. Ried, W. Zhang, Y. Yang, S. Tan, G. Fiorito, L. Franke, S. Guarrera, S. Kasela, J. Kriebel, R. C. Richmond, M. Adamo, U. Afzal, M. Ala-Korpela, B. Albetti, O. Ammerpohl, J. F. Apperley, M. Beekman, P. A. Bertazzi, S. L. Black, C. Blancher, M.-J. Bonder, M. Brosch, M. Carstensen-Kirberg, A. J. M. de Craen, S. de Lusignan, A. Dehghan, M. Elkalaawy, K. Fischer, O. H. Franco, T. R. Gaunt, J. Hampe, M. Hashemi, A. Isaacs, A. Jenkinson, S. Jha, N. Kato, V. Krogh, M. Laffan, C. Meisinger, T. Meitinger, Z. Y. Mok, V. Motta, H. K. Ng, Z. Nikolakopoulou, G. Nteliopoulos, S. Panico, N. Pervjakova, H. Prokisch, W. Rathmann, M. Roden, F. Rota, M. A. Rozario, J. K. Sandling, C. Schafmayer, K. Schramm, R. Siebert, P. E. Slagboom, P. Soininen, L. Stolk, K. Strauch, E.-S. Tai, L. Tarantini, B. Thorand, E. F. Tigchelaar, R. Tumino, A. G. Uitterlinden, C. van Duijn, J. B. J. van Meurs, P. Vineis, A. R. Wickremasinghe, C. Wijmenga, T.-P. Yang, W. Yuan, A. Zhernakova, R. L. Batterham, G. D. Smith, P. Deloukas, B. T. Heijmans, C. Herder, A. Hofman, C. M. Lindgren, L. Milani, P. van der Harst, A. Peters, T. Illig, C. L. Relton, M. Waldenberger, M.-R. Järvelin, V. Bollati, R. Soong, T. D. Spector, J. Scott, M. I. McCarthy, P. Elliott, J. T. Bell, G. Matullo, C. Gieger, J. S. Kooner, H. Grallert, and J. C. Chambers
2016. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature*, 541(7635):81–86.

Wakil, S. J. and L. A. Abu-Elheiga
2009. Fatty acid metabolism: target for metabolic syndrome. *Journal of Lipid Research*, 50(Supplement):S138–S143.

Walaszczyk, E., M. Luijten, A. M. W. Spijkerman, M. J. Bonder, H. L. Lutgers, H. Snieder, B. H. R. Wolffenbuttel, and J. V. van Vliet-Ostaptchouk
2018. DNA methylation markers associated with type 2 diabetes, fasting glucose and HbA1c levels: a systematic review and replication in a case–control sample of the Lifelines study. *Diabetologia*, 61(2):354–368.

Wang, L., X.-J. Jia, H.-J. Jiang, Y. Du, F. Yang, S.-Y. Si, and B. Hong
2013a. MicroRNAs 185, 96, and 223 Repress Selective High-Density Lipoprotein Cholesterol Uptake through Posttranscriptional Inhibition. *Molecular and Cellular Biology*, 33(10):1956–1964.

Wang, Y.-c., Y. Li, X.-y. Wang, D. Zhang, H. Zhang, Q. Wu, Y.-q. He, J.-y. Wang, L. Zhang, H. Xia, J. Yan, X. Li, and H. Ying
2013b. Circulating miR-130b mediates metabolic crosstalk between fat and muscle in overweight/obesity. *Diabetologia*, 56(10):2275–2285.

Wang, Y.-S., W.-W. Chou, K.-C. Chen, H.-Y. Cheng, R.-T. Lin, and S.-H. H. Juo
2012. microRNA-152 Mediates DNMT1-Regulated DNA Methylation in the Estrogen Receptor alpha Gene. *PLoS ONE*, 7(1).

Wang, Z.-M., J.-Q. Luo, L.-Y. Xu, H.-H. Zhou, and W. Zhang
2018. Harnessing low-density lipoprotein receptor protein 6 (LRP6) genetic variation and Wnt signaling for innovative diagnostics in complex diseases. *The Pharmacogenomics Journal*, 18(3):351–358.

Welborn, T. A., M. Knuiman, V. McCann, K. Stanton, and I. J. Constable
1984. Clinical macrovascular disease in Caucasoid diabetic subjects: logistic regression analysis of risk variables. *Diabetologia*, 27(6):568–573.

Wen, J. and J. R. Friedman
2012. miR-122 regulates hepatic lipid metabolism and tumor suppression. *The Journal of Clinical Investigation*, 122(8):2773–2776.

Whiting, D. R., L. Guariguata, C. Weil, and J. Shaw
2011. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Research and Clinical Practice*, 94(3):311–321.

Willeit, P., P. Skroblin, A. R. Moschen, X. Yin, D. Kaudewitz, A. Zampetaki, T. Barwari, M. Whitehead, C. M. Ramírez, L. Goedeke, N. Rotllan, E. Bonora, A. D. Hughes, P. Santer, C. Fernández-Hernando, H. Tilg, J. Willeit, S. Kiechl, and M. Mayr
2017. Circulating MicroRNA-122 Is Associated With the Risk of New-Onset Metabolic Syndrome and Type 2 Diabetes. *Diabetes*, 66(2):347–357.

Williams, M. D. and G. M. Mitchell
2012. MicroRNAs in Insulin Resistance and Obesity. *Experimental Diabetes Research*, 2012.

Willmer, T., R. Johnson, J. Louw, and C. Pheiffer
2018. Blood-Based DNA Methylation Biomarkers for Type 2 Diabetes: Potential for Clinical Applications. *Frontiers in Endocrinology*, 9.

Wu, J., Y.-B. Diao, M.-L. Li, Y.-P. Fang, and D.-C. Ma
2009. A semi-supervised learning based method: Laplacian support vector machine used in diabetes disease diagnosis. *Interdisciplinary Sciences, Computational Life Sciences*, 1(2):151–155.

Wu, L. H., Q. Q. Cai, Y. W. Dong, R. Wang, B. M. He, B. Qi, C. J. Xu, and X. Z. Wu
2013. Decoy oligonucleotide rescues IGF1r expression from MicroRNA-223 suppression. *PloS One*, 8(12):e82167.

Xiao, F., Z. Zuo, G. Cai, S. Kang, X. Gao, and T. Li
2009. miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Research*, 37(Database issue):D105–110.

Xu, X., J.-S. So, J.-G. Park, and A.-H. Lee
2013. Transcriptional control of hepatic lipid metabolism by SREBP and ChREBP. *Seminars in liver disease*, 33(4):301–311.

Xue, A., Y. Wu, Z. Zhu, F. Zhang, K. E. Kemper, Z. Zheng, L. Yengo, L. R. Lloyd-Jones, J. Sidorenko, Y. Wu, eQTLGen Consortium, A. F. McRae, P. M. Visscher, J. Zeng, and J. Yang
2018. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nature Communications*, 9(1):2941.

Xue, J., A. Zhou, Y. Wu, S.-A. Morris, K. Lin, S. Amin, R. Verhaak, G. Fuller, K. Xie, A. B. Heimberger, and S. Huang
2016. miR-182-5p Induced by STAT3 Activation Promotes Glioma Tumorigenesis. *Cancer Research*, 76(14):4293–4304.

Yang, W.-M., K.-H. Min, and W. Lee
2016. MicroRNA expression analysis in the liver of high fat diet-induced obese mice. *Data in Brief*, 9:1155–1159.

Yarwood, A., B. Han, S. Raychaudhuri, J. Bowes, M. Lunt, D. A. Pappas, J. Kremer, J. D. Greenberg, R. Plenge, Rheumatoid Arthritis Consortium International (RACI), J. Worthington, A. Barton, and S. Eyre
2015. A weighted genetic risk score using all known susceptibility variants to estimate rheumatoid arthritis risk. *Annals of the Rheumatic Diseases*, 74(1):170–176.

Younossi, Z. M., A. B. Koenig, D. Abdelatif, Y. Fazel, L. Henry, and M. Wymer
2016. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology (Baltimore, Md.)*, 64(1):73–84.

Zampetaki, A., S. Kiechl, I. Drozdov, P. Willeit, U. Mayr, M. Prokopi, A. Mayr, S. Weger, F. Oberhollenzer, E. Bonora, A. Shah, J. Willeit, and M. Mayr
2010. Plasma MicroRNA Profiling Reveals Loss of Endothelial MiR-126 and Other MicroRNAs in Type 2 Diabetes. *Circulation Research*, 107(6):810–817.

Zarrinpar, A., S. Gupta, M. R. Maurya, S. Subramaniam, and R. Loomba
2016. Serum microRNAs explain discordance of non-alcoholic fatty liver disease in monozygotic and dizygotic twins: a prospective study. *Gut*, 65(9):1546–1554.

Zhang, K., L. Li, Y. Qi, X. Zhu, B. Gan, R. A. DePinho, T. Averitt, and S. Guo
2012. Hepatic Suppression of Foxo1 and Foxo3 Causes Hypoglycemia and Hyperlipidemia in Mice. *Endocrinology*, 153(2):631–646.

Zhang, S., Q. Zhang, G. Shi, and J. Yin
2018a. MiR-182-5p regulates BCL2l12 and BCL2 expression in acute myeloid leukemia as a potential therapeutic target. *Biomedicine & Pharmacotherapy*, 97:1189–1194.

Zhang, Z., X. Liu, H. Xu, X. Feng, Y. Lin, Y. Huang, Y. Peng, and M. Gu
2018b. Obesity-induced upregulation of miR-361-5p promotes hepatosteatosis through targeting Sirt1. *Metabolism - Clinical and Experimental*, 88:31–39.

Zhao, J. V., C. M. Schooling, and J. X. Zhao
2018. The effects of folate supplementation on glucose metabolism and risk of type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. *Annals of Epidemiology*, 28(4):249–257.e1.

Zhu, H., N. Shyh-Chang, A. V. Segrè, G. Shinoda, S. P. Shah, W. S. Einhorn, A. Takeuchi, J. M. Engreitz, J. P. Hagan, M. G. Kharas, A. Urbach, J. E. Thornton, R. Triboulet, R. I. Gregory, DIAGRAM Consortium, MAGIC Investigators, D. Altshuler, and G. Q. Daley
2011. The Lin28/let-7 axis regulates glucose metabolism. *Cell*, 147(1):81–94.

Zhu, W.-G., K. Srinivasan, Z. Dai, W. Duan, L. J. Druhan, H. Ding, L. Yee, M. A. Villalona-Calero, C. Plass, and G. A. Otterson
2003. Methylation of Adjacent CpG Sites Affects Sp1/Sp3 Binding and Activity in the p21cip1 Promoter. *Molecular and Cellular Biology*, 23(12):4056–4065.

Ziller, M. J., H. Gu, F. Müller, J. Donaghey, L. T.-Y. Tsai, O. Kohlbacher, P. L. De Jager, E. D. Rosen, D. A. Bennett, B. E. Bernstein, A. Gnirke, and A. Meissner
2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature*, 500(7463):477–481.

Ziller, M. J., K. D. Hansen, A. Meissner, and M. J. Aryee
2015. Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nature Methods*, 12(3):230–232.

Zou, D., Y. Ye, N. Zou, and J. Yu
2017. Analysis of risk factors and their interactions in type 2 diabetes mellitus: A cross-sectional survey in Guilin, China. *Journal of Diabetes Investigation*, 8(2):188–194.

# 9 Supplemental Tables

Table 28: **Linear regression models of *GALNT18* methylation at CpG1 and CpG3.**

| **CpG1** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | 0.02098 | -2.06 | 0.58717 | -0.47 | 82 |
| Insulin | 0.01803 | -5.54 | 0.38928 | -1.79 | 80 |
| HOMA-IR | 0.07018 | -0.38 | 0.78641 | -0.05 | 75 |
| C peptide | 0.01545 | -9.82 | 0.11856 | -6.40 | 82 |
| HbA1c | 0.00050 | -0.09 | 0.06205 | -0.04 | 87 |
| BMI | 0.71319 | -0.06 | | | 87 |
| Age | 0.00076 | -0.58 | | | 87 |
| Gender | 0.01735 | 0.01 | | | 87 |
| cholesterol | 0.22796 | 0.70 | 0.16587 | 0.85 | 80 |
| triglycerides | 0.00463 | -5.00 | 0.14567 | -2.59 | 80 |
| HDL | 0.02809 | 0.41 | 0.15587 | 0.28 | 80 |
| LDL | 0.01900 | 1.25 | 0.12725 | 0.88 | 74 |
| AST | 0.00300 | -1.11 | 0.00989 | -1.07 | 82 |
| ALT | 0.00005 | -1.18 | 0.00051 | -1.08 | 82 |
| **CpG3** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
| Glucose | 0.04305 | -2.10 | 0.71256 | -0.37 | 82 |
| Insulin | 0.05729 | -5.27 | 0.67013 | -1.03 | 80 |
| HOMA-IR | 0.28304 | -0.27 | 0.58886 | 0.13 | 75 |
| C peptide | 0.00315 | -13.79 | 0.03371 | -9.90 | 82 |
| HbA1c | 0.00163 | -0.09 | 0.09281 | -0.05 | 87 |
| BMI | 0.58671 | -0.10 | | | 87 |
| Age | 0.00274 | -0.60 | | | 87 |
| Gender | 0.02446 | 0.02 | | | 87 |
| cholesterol | 0.23292 | 0.81 | 0.21226 | 0.89 | 80 |
| triglycerides | 0.04324 | -4.22 | 0.50132 | -1.39 | 80 |
| HDL | 0.06392 | 0.40 | 0.31423 | 0.23 | 80 |
| LDL | 0.04681 | 1.19 | 0.22123 | 0.78 | 74 |
| AST | 0.00167 | -1.37 | 0.00469 | -1.34 | 82 |
| ALT | 0.00005 | -1.38 | 0.00042 | -1.26 | 82 |

Table 29: **Linear regression models of hepatic *FOXO1* expression.**

| ***FOXO1*** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | 0.02757 | -25.28 | 0.47689 | -7.82 | 82 |
| Insulin | 0.01939 | -62.84 | 0.11845 | -37.73 | 80 |
| HOMA-IR | 0.03066 | -5.19 | 0.21245 | -2.86 | 75 |
| C peptide | 0.26154 | -56.08 | 0.54431 | -29.71 | 82 |
| HbA1c | 0.00412 | -0.91 | 0.15836 | -0.43 | 87 |
| BMI | 0.70369 | -0.74 | | | 87 |
| Age | 0.00115 | -6.98 | | | 87 |
| Gender | 0.15476 | 0.21 | | | |
| Cholesterol | 0.02035 | 17.14 | 0.00369 | 21.83 | 80 |
| Triglycerides | 0.38011 | -19.92 | 0.58087 | 12.14 | 80 |
| HDL | 0.81498 | -0.53 | 0.38743 | -1.99 | 80 |
| LDL | 0.00421 | 18.55 | 0.01517 | 17.19 | 74 |
| AST | 0.30602 | -4.84 | 0.62211 | -2.50 | 82 |
| ALT | 0.13074 | -5.50 | 0.36044 | -3.47 | 82 |
| miR-182-5p | 0.01064 | 0.54 | | | 85 |

Table 30: **Linear regression models of hepatic *CDKN1B* expression.**

| ***CDKN1B*** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | 0.42539 | -10.12 | 0.58275 | 6.33 | 82 |
| Insulin | 0.09749 | -49.40 | 0.77137 | -7.59 | 80 |
| HOMA-IR | 0.24721 | -3.16 | 0.80127 | 0.65 | 75 |
| C peptide | 0.20488 | -69.13 | 0.58139 | -29.51 | 82 |
| HbA1c | 0.56515 | -0.20 | 0.39559 | 0.26 | 87 |
| BMI | 0.98213 | 0.05 | | | 87 |
| Age | 0.02347 | -5.26 | | | 87 |
| Gender | 0.13443 | 0.21 | | | 87 |
| Cholesterol | 0.17669 | 10.80 | 0.14151 | 11.73 | 80 |
| Triglycerides | 0.04180 | -49.82 | 0.29945 | -24.12 | 80 |
| HDL | 0.07416 | 4.49 | 0.21371 | 3.12 | 80 |
| LDL | 0.09344 | 12.09 | 0.33360 | 7.15 | 74 |
| AST | 0.91773 | -0.54 | 0.74589 | 1.75 | 82 |
| ALT | 0.55965 | -2.34 | 0.97583 | 0.12 | 82 |
| miR-182-5p | 0.10977 | 0.37 | | | 85 |

Table 31: **Linear regression models of hepatic *LRP6* expression.**

| ***LRP6*** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | 0.00339 | -37.52 | 0.05410 | -23.21 | 82 |
| Insulin | 0.02338 | -73.14 | 0.02059 | -64.64 | 80 |
| HOMA-IR | 0.11529 | -4.58 | 0.18941 | -3.54 | 75 |
| C peptide | 0.09546 | -98.11 | 0.13285 | -83.40 | 82 |
| HbA1c | 0.00003 | -1.46 | 0.00081 | -1.08 | 87 |
| BMI | 0.22907 | -2.65 | | | 87 |
| Age | 0.01419 | -5.99 | | | 87 |
| Gender | 0.64291 | 0.06 | | | 87 |
| Cholesterol | 0.31714 | -8.75 | 0.56474 | -5.08 | 80 |
| Triglycerides | 0.01341 | -64.07 | 0.06911 | -45.24 | 80 |
| HDL | 0.08888 | 4.56 | 0.09232 | 4.51 | 80 |
| LDL | 0.75893 | 2.35 | 0.83820 | 1.60 | 74 |
| AST | 0.00190 | -16.18 | 0.00538 | -15.38 | 82 |
| ALT | 0.04265 | -8.31 | 0.05551 | -8.03 | 82 |
| miR-182-5p | 0.00008 | 0.92 | | | 85 |

Table 32: **Linear regression models of *IRS2* methylation, *IRS2* expression and hsa-let-7e-5p expression.**

| **CpG3** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
|---|---|---|---|---|---|
| Glucose | 0.27913 | -1.41 | 0.38814 | -0.98 | 93 |
| Insulin | 0.06188 | 4.81 | 0.04793 | 4.28 | 89 |
| HOMA-IR | 0.22271 | 0.32 | 0.23160 | 0.29 | 85 |
| C peptide | 0.32909 | 5.59 | 0.35267 | 4.96 | 92 |
| HbA1c | 0.03121 | -0.08 | 0.04095 | -0.06 | 97 |
| BMI | 0.98051 | -0.01 | | | 97 |
| Age | 0.41996 | -0.20 | | | 97 |
| Gender | 0.96408 | 0.05 | | | 97 |
| cholesterol | 0.49732 | -0.58 | 0.56764 | -0.47 | 91 |
| triglycerides | 0.87037 | -0.41 | 0.90439 | 0.28 | 91 |
| HDL | 0.75673 | -0.08 | 0.66263 | -0.11 | 91 |
| LDL | 0.82680 | -0.16 | 0.82003 | -0.17 | 85 |
| AST | 0.00207 | -1.57 | 0.00297 | -1.52 | 93 |
| ALT | 0.10693 | -0.67 | 0.10203 | -0.66 | 93 |
| **CpG6** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
| Glucose | 0.59045 | -0.32 | 0.38818 | -0.45 | 96 |

| | | | | | |
|---|---|---|---|---|---|
| Insulin | 0.54548 | -0.85 | 0.36254 | -1.09 | 92 |
| HOMA-IR | 0.41641 | -0.10 | 0.28817 | -0.12 | 88 |
| C peptide | 0.39227 | 2.25 | 0.22158 | 3.01 | 95 |
| HbA1c | 0.41702 | -0.01 | 0.20142 | -0.02 | 100 |
| BMI | 0.39061 | -0.08 | | | 100 |
| Age | 0.55329 | 0.07 | | | 100 |
| Gender | 0.88064 | 0.38 | | | 100 |
| cholesterol | 0.90926 | 0.04 | 0.80223 | -0.09 | 94 |
| triglycerides | 0.31526 | 1.13 | 0.39058 | 0.89 | 94 |
| HDL | 0.79772 | -0.03 | 0.65321 | -0.05 | 94 |
| LDL | 0.66168 | -0.14 | 0.50717 | -0.22 | 88 |
| AST | 0.72036 | -0.08 | 0.61476 | -0.12 | 96 |
| ALT | 0.79293 | -0.05 | 0.78052 | -0.05 | 96 |
| **Hepatic *IRS2* expression** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
| Glucose | 0.20644 | -11.53 | 0.44620 | -6.32 | 86 |
| Insulin | 0.04600 | -44.51 | 0.05809 | -36.88 | 84 |
| HOMA-IR | 0.08640 | -3.49 | 0.09814 | -3.07 | 79 |
| C peptide | 0.26625 | -45.36 | 0.78085 | -10.76 | 86 |
| HbA1c | 0.01782 | -0.59 | 0.02521 | -0.51 | 91 |
| BMI | 0.01714 | -3.52 | | 0.00 | 91 |
| Age | 0.46136 | -1.27 | | 0.00 | 91 |
| Gender | 0.38421 | 0.16 | | | 91 |
| cholesterol | 0.77623 | -1.69 | 0.73346 | -2.02 | 84 |
| triglycerides | 0.37414 | -15.78 | 0.55775 | -9.78 | 84 |
| HDL | 0.52941 | -1.15 | 0.25398 | -2.08 | 84 |
| LDL | 0.91322 | 0.58 | 0.92432 | -0.51 | 78 |
| AST | 0.03320 | -7.85 | 0.03552 | -7.99 | 86 |
| ALT | 0.00679 | -7.88 | 0.00433 | -8.18 | 86 |
| **Hepatic hsa-let-7e-5p expression** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
| Glucose | 0.44412 | 7.88 | 0.48941 | 6.43 | 81 |
| Insulin | 0.21018 | -30.39 | 0.62322 | -10.21 | 79 |
| HOMA-IR | 0.07172 | -3.86 | 0.19158 | -2.58 | 75 |
| C peptide | 0.94199 | -3.24 | 0.47244 | 28.73 | 81 |
| HbA1c | 0.22218 | 0.35 | 0.20125 | 0.32 | 85 |
| BMI | 0.98098 | 0.04 | | 0.00 | 85 |
| Age | 0.58356 | 1.10 | | 0.00 | 85 |
| Gender | 0.17566 | 0.23 | | | 85 |
| cholesterol | 0.15985 | 9.19 | 0.38042 | 5.69 | 79 |
| triglycerides | 0.06427 | 36.62 | 0.05362 | 35.32 | 79 |
| HDL | 0.37439 | -1.85 | 0.16737 | -2.80 | 79 |
| LDL | 0.54959 | 3.39 | 0.82200 | 1.30 | 73 |
| AST | 0.39901 | 3.58 | 0.45724 | 3.23 | 81 |
| ALT | 0.35585 | 3.09 | 0.18634 | 4.34 | 81 |
| **Serum hsa-let-7e-5p concentration** | **p** | **estimate** | **p corrected** | **estimate** | **n** |
| Glucose | 0.46838 | -2.44 | 0.80675 | -0.75 | 80 |
| Insulin | 0.01955 | -19.41 | 0.13337 | -10.97 | 82 |
| HOMA-IR | 0.04346 | -1.46 | 0.20393 | -0.86 | 78 |
| C peptide | 0.35705 | -14.85 | 0.77169 | -4.48 | 84 |
| HbA1c | 0.32197 | -0.10 | 0.66193 | -0.04 | 84 |
| BMI | 0.94717 | 0.04 | | | 84 |
| Age | 0.31495 | -0.70 | | | 84 |
| Gender | 0.06090 | 0.90 | | | 84 |
| cholesterol | 0.66888 | -1.06 | 0.38541 | -2.11 | 78 |
| triglycerides | 0.65503 | -3.01 | 0.81385 | 1.51 | 78 |

| | | | | | |
|---|---|---|---|---|---|
| HDL | 0.19209 | 0.93 | 0.51711 | 0.46 | 78 |
| LDL | 0.64763 | -0.94 | 0.27962 | -2.24 | 74 |
| AST | 0.98228 | 0.03 | 0.82049 | 0.35 | 80 |
| ALT | 0.81711 | 0.27 | 0.36486 | 1.03 | 80 |

Table 33: **Linear regression models of *IRS2* methylation at CpG3 in dependency of the genotype at rs4547213**

| **CpG3** | **p corrected CpG3** | **estimate CpG3** | **p corrected GT** | **estimate GT** | **n** |
|---|---|---|---|---|---|
| Glucose | 0.44259 | -0.91 | 0.83314 | 1.81 | 93 |
| Insulin | 0.03940 | 4.64 | 0.53520 | 10.20 | 89 |
| HOMA-IR | 0.20324 | 0.32 | 0.63056 | 0.84 | 85 |
| C peptide | 0.44301 | 4.27 | 0.63711 | -19.59 | 92 |
| HbA1c | 0.04398 | -0.06 | 0.82277 | -0.05 | 97 |
| cholesterol | 0.61684 | -0.43 | 0.87252 | 1.02 | 91 |
| triglycerides | 0.85416 | -0.44 | 0.28910 | -18.72 | 91 |
| HDL | 0.99150 | 0.00 | 0.14648 | 2.86 | 91 |
| LDL | 0.93418 | -0.06 | 0.65842 | 2.48 | 85 |
| AST | 0.00610 | -1.46 | 0.68063 | 1.56 | 93 |
| ALT | 0.12107 | -0.66 | 0.95913 | 0.16 | 93 |

Table 34: **Linear regression models of hepatic miRNA expression from microarray associated to T2D.** Only associations of truly expressed miRNA with an $p < 0.05$ and an log2 intensity $> 2$ are considered. Linear regression models were adjusted for age, BMI, gender and NAS.

| **T2D** | **Transcript** | **p** | **FC** | **log2 T2D** | **log2 ND** |
|---|---|---|---|---|---|
| 20500149 | hsa-miR-24-2-5p | 0.0229 | 1.89 | 3.41 | 2.49 |
| 20500187 | hsa-miR-29b-1-5p | 0.0342 | 1.61 | 2.94 | 2.25 |
| 20500394 | hsa-miR-197-5p | 0.0485 | 1.25 | 6.06 | 5.74 |
| 20500450 | hsa-miR-182-5p | 0.0447 | 2.60 | 4.50 | 3.12 |
| 20500490 | hsa-miR-224-3p | 0.0234 | 1.89 | 6.01 | 5.09 |
| 20500777 | hsa-miR-138-1-3p | 0.0242 | 2.01 | 3.22 | 2.22 |
| 20500780 | hsa-miR-149-5p | 0.0491 | 1.35 | 4.58 | 4.15 |
| 20501209 | hsa-miR-365a-5p | 0.0426 | 1.41 | 4.41 | 3.91 |
| 20501249 | hsa-miR-381-3p | 0.0144 | 1.40 | 2.54 | 2.05 |
| 20501276 | hsa-miR-330-3p | 0.0095 | 1.98 | 3.93 | 2.94 |
| 20503103 | hsa-miR-485-5p | 0.0353 | 1.41 | 3.19 | 2.69 |
| 20504572 | hsa-miR-1301-3p | 0.0259 | 1.38 | 5.21 | 4.74 |
| 20506712 | hsa-miR-1180-3p | 0.0468 | 1.72 | 4.96 | 4.17 |
| 20509071 | hsa-miR-1825 | 0.0323 | -1.56 | 2.14 | 2.78 |
| 20511549 | hsa-miR-2110 | 0.0389 | 1.27 | 5.24 | 4.90 |
| 20517680 | hsa-miR-4298 | 0.0081 | 1.29 | 7.50 | 7.13 |
| 20517816 | hsa-miR-3609 | 0.0158 | 2.16 | 5.83 | 4.72 |
| 20517824 | hsa-miR-3615 | 0.0175 | 1.46 | 4.37 | 3.83 |
| 20518432 | hsa-miR-3911 | 0.0428 | 1.57 | 4.71 | 4.06 |
| 20518879 | hsa-miR-4485 | 0.0377 | -1.69 | 7.79 | 8.54 |
| 20518881 | hsa-miR-4487 | 0.0473 | -1.47 | 3.41 | 3.98 |
| 20518935 | hsa-miR-4534 | 0.0149 | 1.25 | 5.18 | 4.86 |
| 20519425 | hsa-miR-4646-5p | 0.0105 | -1.65 | 3.90 | 4.62 |
| 20519467 | hsa-miR-4669 | 0.0344 | 1.49 | 5.30 | 4.72 |
| 20519518 | hsa-miR-4701-3p | 0.0242 | 1.34 | 3.82 | 3.39 |
| 20526885 | hsa-miR-7162-3p | 0.0493 | 2.05 | 3.29 | 2.26 |
| 20529139 | hsa-miR-7847-3p | 0.0150 | 1.27 | 7.98 | 7.63 |
| 20529781 | hsa-miR-8071 | 0.0322 | 1.44 | 4.01 | 3.49 |

Table 35: **Linear regression models of hepatic miRNA expression from microarray associated to metabolic traits.** Only associations of truly expressed miRNA with an $p < 0.05$ and an log2 intensity $> 2$ are considered. Linear regression models were adjusted for age, BMI and gender (if not response value). FC .. Fold change difference between T2D and ND subjects; GLC .. blood glucose; INS .. serum insulin; TRI .. serum triglycerides

| **AGE** | **Transcript** | **estimate** | **p** | **FC** | **log2 T2D** | **log2 ND** |
|---|---|---|---|---|---|---|
| 20500119 | hsa-let-7d-5p | -5.83 | 0.0476 | 1.31 | 11.81 | 11.43 |
| 20500121 | hsa-let-7e-5p | -4.41 | 0.0286 | 1.84 | 11.11 | 10.23 |
| 20502235 | hsa-miR-18b-5p | -3.39 | 0.0353 | 1.28 | 2.61 | 2.25 |
| 20503103 | hsa-miR-485-5p | -4.64 | 0.0435 | 1.41 | 3.19 | 2.69 |
| 20503877 | hsa-miR-501-5p | -3.82 | 0.0360 | 1.32 | 3.32 | 2.92 |
| 20504379 | hsa-miR-629-5p | -3.67 | 0.0263 | 1.44 | 4.16 | 3.63 |
| 20517680 | hsa-miR-4298 | -8.11 | 0.0310 | 1.29 | 7.50 | 7.13 |
| 20517816 | hsa-miR-3609 | -2.73 | 0.0204 | 2.16 | 5.83 | 4.72 |
| 20517948 | hsa-miR-3687 | 2.02 | 0.0499 | 1.46 | 7.01 | 6.46 |
| 20518879 | hsa-miR-4485 | 3.08 | 0.0286 | -1.69 | 7.79 | 8.54 |
| **BMI** | **Transcript** | **estimate** | **p** | **FC** | **log2 T2D** | **log2 ND** |
| 20500164 | hsa-miR-31-5p | -1.99 | 0.0492 | 1.73 | 4.26 | 3.47 |
| 20500722 | hsa-miR-27b-5p | -4.95 | 0.0322 | 1.25 | 5.47 | 5.15 |
| 20501249 | hsa-miR-381-3p | -5.00 | 0.0463 | 1.40 | 2.54 | 2.05 |
| 20501293 | hsa-miR-331-3p | -3.63 | 0.0256 | 1.39 | 5.64 | 5.16 |
| 20503815 | hsa-miR-498 | 3.02 | 0.0476 | 2.24 | 3.76 | 2.59 |
| 20504559 | hsa-miR-1224-5p | 3.26 | 0.0420 | 1.40 | 4.68 | 4.20 |
| 20518783 | hsa-miR-378e | -3.21 | 0.0387 | -1.35 | 3.51 | 3.95 |
| 20518788 | hsa-miR-378f | -6.40 | 0.0412 | -1.27 | 8.18 | 8.52 |
| 20518794 | hsa-miR-378g | -5.83 | 0.0259 | -1.35 | 5.87 | 6.31 |
| 20518800 | hsa-miR-4428 | 3.84 | 0.0461 | 1.33 | 3.14 | 2.73 |
| 20519417 | hsa-miR-4640-5p | 4.24 | 0.0436 | 1.33 | 5.03 | 4.62 |
| 20519429 | hsa-miR-4649-5p | 5.61 | 0.0343 | 1.28 | 7.61 | 7.25 |
| 20521811 | hsa-miR-664b-3p | -3.92 | 0.0423 | -1.71 | 5.33 | 6.11 |
| 20521825 | hsa-miR-5585-3p | 3.71 | 0.0237 | 1.51 | 3.02 | 2.43 |
| 20525684 | hsa-miR-6861-5p | 4.39 | 0.0133 | 1.47 | 5.13 | 4.58 |
| **HbA1c** | **Transcript** | **estimate** | **p** | **FC** | **log2 T2D** | **log2 ND** |
| 20500149 | hsa-miR-24-2-5p | 0.66 | 0.0276 | 1.89 | 3.41 | 2.49 |
| 20500150 | hsa-miR-25-5p | 0.76 | 0.0274 | 1.47 | 4.25 | 3.69 |
| 20500188 | hsa-miR-29b-3p | 0.61 | 0.0494 | 1.38 | 2.68 | 2.21 |
| 20500394 | hsa-miR-197-5p | 1.38 | 0.0351 | 1.25 | 6.06 | 5.74 |
| 20500450 | hsa-miR-182-5p | 0.47 | 0.0171 | 2.60 | 4.50 | 3.12 |
| 20500490 | hsa-miR-224-3p | 0.80 | 0.0466 | 1.89 | 6.01 | 5.09 |
| 20500777 | hsa-miR-138-1-3p | 0.75 | 0.0047 | 2.01 | 3.22 | 2.22 |
| 20501177 | hsa-miR-99b-3p | 0.89 | 0.0101 | 1.61 | 4.21 | 3.53 |
| 20501249 | hsa-miR-381-3p | 1.20 | 0.0043 | 1.40 | 2.54 | 2.05 |
| 20501276 | hsa-miR-330-3p | 0.93 | 0.0009 | 1.98 | 3.93 | 2.94 |
| 20503103 | hsa-miR-485-5p | 1.35 | 0.0004 | 1.41 | 3.19 | 2.69 |
| 20503789 | hsa-miR-491-5p | 1.12 | 0.0223 | 1.28 | 3.87 | 3.52 |
| 20504379 | hsa-miR-629-5p | 0.82 | 0.0099 | 1.44 | 4.16 | 3.63 |
| 20504569 | hsa-miR-1271-5p | 0.97 | 0.0313 | 1.29 | 5.26 | 4.90 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 20504572 | hsa-miR-1301-3p | 0.97 | 0.0203 | 1.38 | 5.21 | 4.74 |
| 20506712 | hsa-miR-1180-3p | 0.78 | 0.0438 | 1.72 | 4.96 | 4.17 |
| 20517680 | hsa-miR-4298 | 1.95 | 0.0053 | 1.29 | 7.50 | 7.13 |
| 20517726 | hsa-miR-4269 | 0.67 | 0.0367 | 1.69 | 3.27 | 2.51 |
| 20517816 | hsa-miR-3609 | 0.56 | 0.0169 | 2.16 | 5.83 | 4.72 |
| 20519467 | hsa-miR-4669 | 1.12 | 0.0130 | 1.49 | 5.30 | 4.72 |
| 20521811 | hsa-miR-664b-3p | -0.75 | 0.0325 | -1.71 | 5.33 | 6.11 |
| 20525659 | hsa-miR-6849-5p | 0.90 | 0.0231 | 1.44 | 3.61 | 3.08 |
| 20529139 | hsa-miR-7847-3p | 1.88 | 0.0127 | 1.27 | 7.98 | 7.63 |

| **GLC** | **Transcript** | **estimate** | **p** | **FC** | **log2 T2D** | **log2 ND** |
|---|---|---|---|---|---|---|
| 20500188 | hsa-miR-29b-3p | 25.00 | 0.0219 | 1.38 | 2.68 | 2.21 |
| 20500450 | hsa-miR-182-5p | 15.05 | 0.0235 | 2.60 | 4.50 | 3.12 |
| 20500469 | hsa-miR-212-3p | 22.01 | 0.0292 | 1.65 | 3.51 | 2.79 |
| 20500555 | hsa-miR-200b-5p | 23.22 | 0.0157 | 1.60 | 4.47 | 3.80 |
| 20500777 | hsa-miR-138-1-3p | 27.05 | 0.0039 | 2.01 | 3.22 | 2.22 |
| 20501249 | hsa-miR-381-3p | 31.03 | 0.0440 | 1.40 | 2.54 | 2.05 |
| 20501276 | hsa-miR-330-3p | 23.99 | 0.0188 | 1.98 | 3.93 | 2.94 |
| 20503103 | hsa-miR-485-5p | 30.12 | 0.0340 | 1.41 | 3.19 | 2.69 |
| 20503815 | hsa-miR-498 | 21.84 | 0.0235 | 2.24 | 3.76 | 2.59 |
| 20504379 | hsa-miR-629-5p | 25.49 | 0.0234 | 1.44 | 4.16 | 3.63 |
| 20506839 | hsa-miR-1247-3p | 17.59 | 0.0199 | 1.64 | 5.01 | 4.29 |
| 20515603 | hsa-miR-3175 | -32.69 | 0.0277 | -1.69 | 6.61 | 7.36 |
| 20515610 | hsa-miR-3180-3p | 45.54 | 0.0159 | 1.32 | 7.13 | 6.73 |
| 20517680 | hsa-miR-4298 | 51.36 | 0.0441 | 1.29 | 7.50 | 7.13 |
| 20517726 | hsa-miR-4269 | 23.41 | 0.0333 | 1.69 | 3.27 | 2.51 |
| 20517820 | hsa-miR-3613-5p | -7.19 | 0.0312 | -3.44 | 5.33 | 7.11 |
| 20521811 | hsa-miR-664b-3p | -37.56 | 0.0016 | -1.71 | 5.33 | 6.11 |
| 20521825 | hsa-miR-5585-3p | 21.63 | 0.0430 | 1.51 | 3.02 | 2.43 |
| 20525023 | hsa-miR-6511a-5p | 25.80 | 0.0279 | 1.36 | 3.72 | 3.27 |
| 20525539 | hsa-miR-6789-5p | 49.97 | 0.0312 | 1.34 | 9.35 | 8.93 |
| 20525659 | hsa-miR-6849-5p | 32.83 | 0.0189 | 1.44 | 3.61 | 3.08 |
| 20525739 | hsa-miR-6889-5p | 21.77 | 0.0485 | 1.58 | 4.23 | 3.57 |
| 20526885 | hsa-miR-7162-3p | 17.18 | 0.0431 | 2.05 | 3.29 | 2.26 |
| 20529781 | hsa-miR-8071 | 29.29 | 0.0269 | 1.44 | 4.01 | 3.49 |

| **INS** | **Transcript** | **estimate** | **p** | **FC** | **log2 T2D** | **log2 ND** |
|---|---|---|---|---|---|---|
| 20500164 | hsa-miR-31-5p | -19.47 | 0.0475 | 1.73 | 4.26 | 3.47 |
| 20501291 | hsa-miR-148b-3p | 28.72 | 0.0228 | 1.72 | 3.84 | 3.05 |
| 20501314 | hsa-miR-346 | 48.96 | 0.0001 | 2.26 | 3.83 | 2.66 |
| 20502235 | hsa-miR-18b-5p | 44.58 | 0.0206 | 1.28 | 2.61 | 2.25 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 20515603 | hsa-miR-3175 | -53.08 | 0.0245 | -1.69 | 6.61 | 7.36 |
| 20517835 | hsa-miR-3621 | 68.01 | 0.0386 | 1.26 | 8.44 | 8.10 |
| 20519463 | hsa-miR-4668-5p | -34.52 | 0.0494 | -1.34 | 7.01 | 7.44 |
| 20519474 | hsa-miR-4674 | 68.23 | 0.0357 | 1.30 | 8.84 | 8.46 |
| 20519615 | hsa-miR-371b-5p | 21.31 | 0.0288 | 1.47 | 3.61 | 3.06 |
| 20521825 | hsa-miR-5585-3p | 38.16 | 0.0122 | 1.51 | 3.02 | 2.43 |
| 20525509 | hsa-miR-6774-5p | 31.27 | 0.0385 | 1.41 | 3.90 | 3.41 |
| 20525539 | hsa-miR-6789-5p | 72.00 | 0.0293 | 1.34 | 9.35 | 8.93 |
| 20525659 | hsa-miR-6849-5p | 62.37 | 0.0022 | 1.44 | 3.61 | 3.08 |
| 20526885 | hsa-miR-7162-3p | 30.80 | 0.0183 | 2.05 | 3.29 | 2.26 |
| 20529568 | hsa-miR-7977 | -32.75 | 0.0476 | -1.29 | 7.43 | 7.80 |
| 20529774 | hsa-miR-8064 | 29.88 | 0.0417 | 1.41 | 3.19 | 2.70 |
| 20529783 | hsa-miR-8073 | 26.13 | 0.0474 | 1.43 | 4.29 | 3.78 |
| **HOMA - IR** | **Transcript** | **estimate** | **p** | **FC** | **log2 T2D** | **log2 ND** |
| 20500777 | hsa-miR-138-1-3p | 3.53 | 0.0232 | 2.01 | 3.22 | 2.22 |
| 20501314 | hsa-miR-346 | 3.79 | 0.0005 | 2.26 | 3.83 | 2.66 |
| 20515603 | hsa-miR-3175 | -4.46 | 0.0227 | -1.69 | 6.61 | 7.36 |
| 20515610 | hsa-miR-3180-3p | 5.16 | 0.0229 | 1.32 | 7.13 | 6.73 |
| 20517821 | hsa-miR-3613-3p | -3.04 | 0.0183 | -1.42 | 6.81 | 7.32 |
| 20517835 | hsa-miR-3621 | 6.90 | 0.0100 | 1.26 | 8.44 | 8.10 |
| 20519463 | hsa-miR-4668-5p | -3.55 | 0.0134 | -1.34 | 7.01 | 7.44 |
| 20519474 | hsa-miR-4674 | 6.45 | 0.0157 | 1.30 | 8.84 | 8.46 |
| 20519615 | hsa-miR-371b-5p | 1.98 | 0.0136 | 1.47 | 3.61 | 3.06 |
| 20519663 | hsa-miR-4436b-5p | 3.01 | 0.0279 | 1.41 | 4.39 | 3.89 |
| 20521825 | hsa-miR-5585-3p | 3.60 | 0.0038 | 1.51 | 3.02 | 2.43 |
| 20525539 | hsa-miR-6789-5p | 7.56 | 0.0047 | 1.34 | 9.35 | 8.93 |
| 20525659 | hsa-miR-6849-5p | 5.79 | 0.0005 | 1.44 | 3.61 | 3.08 |
| 20526885 | hsa-miR-7162-3p | 2.51 | 0.0208 | 2.05 | 3.29 | 2.26 |
| 20529568 | hsa-miR-7977 | -2.91 | 0.0332 | -1.29 | 7.43 | 7.80 |
| 20529774 | hsa-miR-8064 | 2.76 | 0.0224 | 1.41 | 3.19 | 2.70 |
| **TRI** | **Transcript** | **estimate** | **p** | **FC** | **log2 T2D** | **log2 ND** |
| 20500142 | hsa-miR-21-3p | 50.06 | 0.0489 | 1.50 | 2.61 | 2.02 |
| 20500164 | hsa-miR-31-5p | 35.37 | 0.0054 | 1.73 | 4.26 | 3.47 |
| 20500780 | hsa-miR-149-5p | -53.93 | 0.0422 | 1.35 | 4.58 | 4.15 |
| 20504324 | hsa-miR-550a-5p | 65.15 | 0.0158 | 1.39 | 3.78 | 3.31 |
| 20504559 | hsa-miR-1224-5p | 50.96 | 0.0136 | 1.40 | 4.68 | 4.20 |
| 20506839 | hsa-miR-1247-3p | 38.97 | 0.0093 | 1.64 | 5.01 | 4.29 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 20515623 | hsa-miR-3187-3p | 46.78 | 0.0304 | 1.29 | 4.06 | 3.69 |
| 20517710 | hsa-miR-4253 | 49.20 | 0.0025 | 2.08 | 3.93 | 2.87 |
| 20517745 | hsa-miR-4286 | -28.25 | 0.0242 | -1.44 | 6.31 | 6.84 |
| 20518800 | hsa-miR-4428 | 66.25 | 0.0065 | 1.33 | 3.14 | 2.73 |
| 20518881 | hsa-miR-4487 | 63.83 | 0.0201 | -1.47 | 3.41 | 3.98 |
| 20518935 | hsa-miR-4534 | 100.16 | 0.0238 | 1.25 | 5.18 | 4.86 |
| 20519554 | hsa-miR-4721 | 42.73 | 0.0380 | 1.41 | 4.23 | 3.74 |
| 20521786 | hsa-miR-5572 | 55.06 | 0.0131 | 1.29 | 5.03 | 4.66 |
| 20521811 | hsa-miR-664b-3p | -49.64 | 0.0446 | -1.71 | 5.33 | 6.11 |
| 20523018 | hsa-miR-6086 | 55.23 | 0.0438 | 1.47 | 4.69 | 4.13 |
| 20525444 | hsa-miR-6741-5p | 66.18 | 0.0314 | 1.28 | 4.86 | 4.50 |
| 20525497 | hsa-miR-6768-5p | 44.13 | 0.0282 | 1.95 | 4.69 | 3.73 |
| 20525513 | hsa-miR-6776-5p | 67.15 | 0.0136 | 1.28 | 4.88 | 4.52 |
| 20525633 | hsa-miR-6780b-5p | 43.01 | 0.0011 | 2.01 | 4.13 | 3.12 |
| 20525683 | hsa-miR-6860 | 61.96 | 0.0370 | 1.25 | 4.89 | 4.57 |
| 20525703 | hsa-miR-6871-5p | 57.69 | 0.0240 | 1.32 | 3.25 | 2.85 |

# 10 Appendix

## 10.1 Vector maps

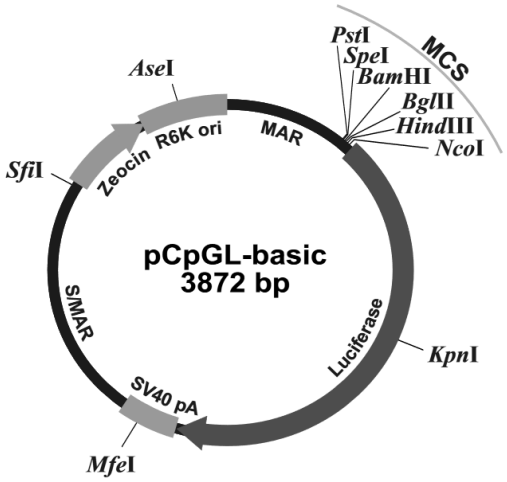


Figure 10.1: **Vector map of the basic pCpGL vector.** This vector was used for the analysis whether CpG methylation might affect upstream CMV promoter activation (not shown). The CMV promoter itself was previously inserted into the basic vector using the *PstI* and *HindIII* restriction sites. CpG containing sequences of interest were cloned downstream the CMV promoter within the *HindIII* and *NcoI* restriction enzymes recognition sites. This vector was a gift from (Klug and Rehli, 2006).

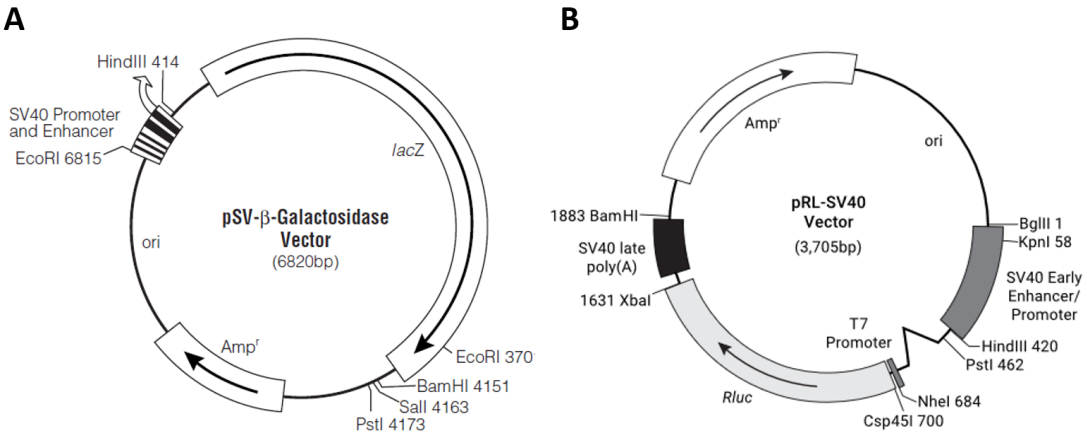


Figure 10.2: **Vector maps of control plasmids, co-transfected during Luciferase reporter gene assay.** Maps of the pSV-$\beta$-Galactosidase vector (A) and pRL-SV40 (B) originate from Promega, Madison, US. Both plasmids served for normalization of Firefly luciferase signal by assessing the transfection efficiency.

## 10.2 MATLAB code

### 10.2.1 Tukey Biweight Algorithm

Adapted for MATLAB from Bioconductor's sscore: S-Score Algorithm for Affymetrix Oligonucleotide Microarrays:

```
function [ value ] = tukeyBiweight( x,c,epsilon )
    medianValue = median(x);
    mad = median(abs(x - medianValue))*c + epsilon;
    n = length(x);
    difference = x - medianValue;
    u = difference ./ mad;
    uSquare = u.^2;
    oneMinusUSquare = 1.-uSquare;

    for i=1:length(x)
        if abs(u(i)) < 1
            weightedSumNumer(i) = differenceerence(i) * oneMinusUSquare(i);
        else
            weightedSumNumer(i) = 0;
        end
    end

    weightedSumNumer2 = sum(weightedSumNumer);

    for i=1:length(x)
        if abs(u(i)) < 1
            weightedSumDenom(i) = oneMinusUSquare(i)^2;
        else
            weightedSumDenom(i) = 0;
        end
    end

    weightedSumDenom2 = sum(weightedSumDenom);

    if (weightedSumDenom2 ~= 0)
        value = medianValue + weightedSumNumer2/weightedSumDenom2;
    else
        value = 0;
    end

end
```

Tuning values of c = 5 and epsilon = 0.0001 (default value to avoid division by zero) were chosen. The Tukey Biweight Algorithm was used to determine a robust average value of microarray expression which is unaffected by outliers. Therefore the median of a data set was determined. All data points contributed to the final average depending on the distance to the median value. In other words, data points with larger distances to the median are weighted less for final mean calculation. Outliers are characterized by a large distance and therefore contribute only marginally to the result.

## 10.3 JAVA code

### 10.3.1 Class MIRNA

```
/**
* Class to generate a miRNA bbject containing a miRNA core
* @author christin
*/
public class MIRNA{
    private int length;
    private String sequence;
    private String full_seqence;
    private String origSeq;

    /**
    * Constructor of miRNA core
    * @param seq String containing core miRNA
    */
    public MIRNA(String seq){
        this.sequence = translateUtoT(seq);
        this.length = this.setLength();
    }

    /**
    * Second constructor for full sequence of miRNA to extract seed which is
    * normally at position 2
    * @param full_seq String with full miRNA sequence
    * @param start_seed Int with start of 8 nt seed sequence
    */
    public MIRNA(String full_seq, int start_seed) {
        this.full_seqence = full_seq;
        String orig =
            translateUtoT(full_seq.substring(start_seed-2, start_seed+6));
        this.origSeq = orig;
        this.sequence = reverseComplement(origSeq);
        this.length = this.setLength();
    }

    private String translateUtoT(String s) {
        String result = "";
        for(int i = 0; i < s.length(); i++) {
            if(s.charAt(i) =='U') {
                result += 'T';
            }
        }
            else {
                result += s.charAt(i);
            }
        }
        return result;
    }

    public String reverseComplement(String s) {
        String result = "";
        for(int i = s.length()-1; i >= 0; i--) {
            if(s.charAt(i) == 'A') {
                result += 'T';
            }
            if(s.charAt(i) == 'T') {
                result += 'A';
            }
            if(s.charAt(i) == 'G') {
                result += 'C';
            }
            if(s.charAt(i) == 'C') {
```

```
                result += 'G';
            }
        }
        return result;
    }

    public String getSeq(){
        return this.sequence;
    }

    public String getFull() {
        return this.full_seqence;
    }

    public int getLength(){
        return this.length;
    }

    public int setLength(){
        if(this.sequence.equals("")){
            this.length = 0;
            return 0;
        }
        this.length = this.sequence.length();
        return this.length;
    }
}
```

### 10.3.2 Class MRNA

```
/**
* Class to generate a mRNA Target sequence
* @author christin
*
*/
public class MRNA{
    private int length;
    private String sequence;

    /**
    * Constructor of mRNA
    * @param seq String with target sequence
    */
    public MRNA(String seq){
        this.sequence = seq;
        this.length = seq.length();
    }

    public String getSeq(){
        return this.sequence;
    }

    public int getLength(){
        return this.length;
    }

    public void setLength(){
        this.length = this.sequence.length();
    }

    public String getSeqAt(int i){
        return this.sequence.substring(i,i+8);
    }
}
```

### 10.3.3 Class LocateSeed

```
1  /**
2  * Class to find a seed sequence of miRNA in a target mRNA object
3  * @author christin
4  * @version 0.2
5  *
6  */
7  public class LocateSeed{
8      /**
9      * 7mer-m8: A perfect match from nucleotides 2-8 of the miRNA seed.
10     * @param miRNA
11     * @param target
12     * @return boolean
13     */
14     public static boolean is7mer_m8(MIRNA miRNA, MRNA target) {
15         boolean flag = false;
16         for(int i = 0; i < target.getLength()-miRNA.getLength(); i++) {
17             String temp1 = target.getSeq().substring(i,i+miRNA.getLength()-1);
18             String temp2 = miRNA.getSeq().substring(0,miRNA.getLength()-1);
19             if(temp1.equals(temp2)) {
20                 if(target.getSeq().charAt(i+miRNA.getLength()-1) != 'A') {
21                     System.out.println(temp1 + " at position "+i);
22                     flag = true;
23                 }
24             }
25 
26         }
27         return flag;
28     }
29 
30     /**
31     * 7mer-A1: A perfect match from nucleotides 2-7 of the miRNA seed in
32     * addition to an A across from the miRNA nucleotide 1.
33     * @param miRNA
34     * @param target
35     * @return boolean
36     */
37     public static boolean is7mer_A1(MIRNA miRNA, MRNA target) {
38         boolean flag = false;
39         for(int i = 1; i < target.getLength()-miRNA.getLength(); i++) {
40             String temp1 = target.getSeq().substring(i,i+miRNA.getLength()-2);
41             String temp2 = miRNA.getSeq().substring(1,miRNA.getLength()-1);
42             if(temp1.equals(temp2)) {
43                 if(target.getSeq().charAt(i+miRNA.getLength()-2) == 'A') {
44                     flag = true;
45                     System.out.println(temp1 + " at position "+i);
46                 }
47             }
48         }
49         return flag;
50     }
51 
52     /**
53     * 8mer: A perfect match from nucleotides 2-8 of the miRNA seed in addition
54     * to an A across from the miRNA nucleotide 1.
55     * @param miRNA
56     * @param target
57     * @return boolean if sequence contains 8mer
58     */
59     public static boolean is8mer(MIRNA miRNA, MRNA target) {
60         boolean flag = false;
```

```
61          for(int i = 1; i < target.getLength()-miRNA.getLength(); i++) {
62              String temp1 = target.getSeq().substring(i,i+miRNA.getLength()-1);
63              String temp2 = miRNA.getSeq().substring(0,miRNA.getLength()-1);
64              if(temp1.equals(temp2)) {
65                  if(target.getSeq().charAt(i+miRNA.getLength()-1) == 'A') {
66                      System.out.println(temp1 + " at position "+i);
67                      flag = true;
68                  }
69              }
70          }
71          return flag;
72      }
73
74      /**
75      * 6mer: A perfect match between the miRNA seed and mRNA for six
76      * nucleotides.
77      * @param miRNA
78      * @param target
79      * @return boolean if sequence contains 6mer
80      */
81      public static boolean is6mer(MIRNA miRNA, MRNA target) {
82          boolean flag = false;
83          for(int i = 1; i < target.getLength()-miRNA.getLength(); i++) {
84              String temp1 = target.getSeq().substring(i,i+miRNA.getLength()-2);
85              String temp2 = miRNA.getSeq().substring(1,miRNA.getLength()-1);
86              if(temp1.equals(temp2)) {
87                  flag = true;
88                  System.out.println(temp1 + " at position "+i);
89              }
90          }
91          return flag;
92      }
93
94      /**
95      * Main method to start the programm
96      * @param args
97      */
98      public static void main (String[] args){
99          MIRNA miRNA = new MIRNA(args[0], Integer.parseInt(args[1]));
100         MRNA target = new MRNA(args[2]);
101
102         System.out.println("Seed Sequence of miRNA: " + miRNA.getSeq());
103         System.out.println(miRNA.getLength());
104         System.out.println("- - - - - - - - - - - - - - -");
105         System.out.println("mRNA contains 8mer: " + is8mer(miRNA, target));
106         System.out.println("- - - - - - - - - - - - - - -");
107         System.out.println("mRNA contains 6mer: " + is6mer(miRNA, target));
108         System.out.println("- - - - - - - - - - - - - - -");
109         System.out.println("mRNA contains 7mer_m8: " + is7mer_m8(miRNA, target));
110         System.out.println("- - - - - - - - - - - - - - -");
111         System.out.println("mRNA contains 7mer_A1: " + is7mer_A1(miRNA, target));
112     }
113 }
```

# 11 Abbreviations

**ADA** American Diabetes Association
**addMRS** additive MRS
**ALD** Alcoholic Liver Disease
**ALT** Alanine Aminotransferase
**ARIC** The Atherosclerosis Risk in Communities
**AST** Aspartate Aminotransferase
**BMI** Body mass index
**bisPCR** bisulfite DNA PCR
**bp** base pair
**cDNA** copy DNA
**CGI** CpG island
**CMV** Cytomegalovirus
**CRP** C-reactive protein
**DMNT** DNA-methyltransferases
**DMR** Differentially methylated region
**DNA** Deoxyribonucleic Acid
**ELISA** Enzyme-linked Immunosorbent Assay
**EWAS** Epigenome-wide association study
**FLuc** Firefly Luciferase
**gDNA** genomic, untreated DNA
**GLC** Glucose, blood glucose
**GOLDN** Genomics of post-prandial lipidomic phenotypes in the Genetics of Lipid lowering Drugs and Diet Network
**GWAS** Genome-wide association study
**HbA1c** Glycated hemoglobin, hemoglobin A1c
**HCC** hepatocellular cancer
**HDL** High-density lipoprotein
**HOMA-IR** Homeostatic Model Assessment for Insulin Resistance
**INS** Insulin, serum insulin
**IR** Insulin Resistance
**KORA** Kooperative Gesundheitsforschung in der Region Augsburg
**LDL** Low-density lipoprotein
**LMW** Low molecular weight
**LOLIPOP** The London Life Sciences Prospective Population
**MARTHA** MARseille THrombosis Association
**METSIM** The Metabolic Syndrome in Men
**MBD** Methyl-CpG-binding domain
**miRNA** micro RNA
**mRNA** messenger RNA
**MRS** Methylation Risk Score
**MuTHER** Multiple Tissue Human Expression Resource
**n** Sample size
**ncRNA** non-coding RNA
**NAFLD** Non-alcoholic fatty liver disease
**NAS** NAFLD Activity Score
**NASH** Non-alcoholic steatohepatitis
**ND** Non-diabetic
**OOB** out-of-bag
**PCR** Polymerase Chain Reaction
**qPCR** quantitative real time PCR
**RDF** Random Decision Forest
**REGICOR** Registre Gironí del COR

**RISC** RNA-induced silencing complex
**RLuc** Renilla Luciferase
**RNA** Ribonucleic acid
**RT** Reverse transcriptase, reverse transcription
**SAM** S-adenosyl Methionine
**SNP** Single nucleotide polymorphism
**siRNA** short-interfering RNA
**T1D** Type 1 diabetes
**T2D** Type 2 diabetes, Type 2 diabetic
**TET** ten-eleven translocation
**TF** Transcription factor
**TH** Thyroid hormone
**TRI** Tryglycerides, serum triglycerides
**TSH** Thyroid-stimulating Hormone
**TSS** Transcriptional start side
**UTR** Untranslated Region
**WGBS** Whole genome bisulfite sequencing

## 11.1 Gene names

**ABCG1** ATP Binding Cassette Subfamily G Member 1
**ACACB** Acetyl-CoA Carboxylase Beta
**APOF** Apolipoprotein F
**CASC3** CASC3 Exon Junction Complex Subunit, Cancer Susceptibility Candidate 3
**CCND1** Cyclin D1
**CDKN1B** Cyclin Dependent Kinase Inhibitor 1B
**CYP2C19** Cytochrome P450 Family 2 Subfamily C Member 19
**CYP7A1** Cytochrome P450 Family 7 Subfamily A Member 1
**DIO1, DIO2, DIO3** Iodothyronine Deiodinase 1, 2 and 3
**ELOVL6** Elongation Of Very Long Chain Fatty Acids Protein 6
**FASN** Fatty Acid Synthase
**FOXO1** Forkhead Box O1
**GALNT18** Polypeptide N-Acetylgalactosaminyltransferase 18
**GAPDH** Glyceraldehyde-3-Phosphate Dehydrogenase
**INSR** Insulin Receptor
**IRS1, IRS2** Insulin Receptor Substrate 1 and 2
**LDLR** Low Density Lipoprotein Receptor
**LRP6** LDL Receptor Related Protein 6
**PCSK9** Proprotein Convertase Subtilisin/Kexin Type 9
**SCD** Stearoyl-CoA Desaturase
**SLC10A1** Solute Carrier Family 10 Member 1, encoding a Sodium/Bile Acid Cotransporter
**SLC16A2** Solute Carrier Family 16 Member 2, encoding the TH transporther Monocarboxylate Transporter 7
**SLC16A10** Solute Carrier Family 16 Member 10, encdoing Aromatic Amino Acid Transporter 1
**SLCO1C1** Solute Carrier Organic Anion Transporter Family Member 1C1, encoding a Thyroxine Transporter
**SREBF1** Sterol Regulatory Element Binding Transcription Factor 1
**THRA, THRB** Thyroid Hormone Receptor Alpha and Beta
**TP53INP1** Tumor Protein P53 Inducible Nuclear Protein 1

# 12 Figures and Tables

## List of Figures

# List of Tables

# 13 Acknowledgments

This was a triumph
I'm making a note here, HUGE SUCCESS
It's hard to overstate my satisfaction
Aperture Science
We do what we must because we can
For the good of all of us
Except the ones who are dead
But there's no sense crying over every mistake
You just keep on trying till you run out of cake
And the science gets done
And you make a neat gun
For the people who are still alive
I'm not even angry
I'm being so sincere right now
Even though you broke my heart and killed me
And tore me to pieces
And threw every piece into a fire
As they burned it hurt because
I was so happy for you
Now these points of data make a beautiful line
And we're out of beta, we're releasing on time
So I'm GLaD I got burned, think of all the things we learned
For the people who are still alive
Go ahead and leave me
I think I prefer to stay inside
Maybe you'll find someone else to help you
Maybe Black Mesa
That was a joke, haha, FAT CHANCE
It's so delicious and moist
Look at me still talking when there's science to do
When I look out there it makes me GLaD I'm not you
I've experiments to run, there is research to be done
On the people who are still alive
And believe me I am still alive
I'm doing science and I'm still alive
I feel fantastic and I'm still alive
While you're dying I'll be still alive
And when you're dead I'll be still alive
Still alive.
STILL ALIVE!

Still Alive (from the video game Portal), by Ellen McLain in 2007
Songwriter: Jonathan Coulton

# 14 Curriculum Vitae

## Personal Data


| | |
|---|---|
| Name | Christin Krause |
| Adress | Trendelenburgstr. 19, 23562 Lübeck |
| Tel. | 0151 - 6142 07 42 |
| E-Mail | christin.krause55@gmail.com |
| Date and Place of Birth | 19.01.1991 in Geesthacht, Germany |


## Education

| | |
|---|---|
| Since October 2018 | **University of Lübeck**, *Lübeck*, Bachelors degree program Computer Science. |
| June 2016 – Mai 2019 | **University of Lübeck**, *Lübeck*, PhD student.<br>**Title of PhD thesis:** Gene-expression in Liver is Altered by DNA-methylation and miRNAs in Obese Subjects.<br>Assigned and supervised by Dr. Henriette Kirchner, PhD at the 1st Dept. of Medicine |
| November 2013 – April 2016 | **University of Lübeck**, *Lübeck*, masters degree program Molecular Life Science, M.Sc (1,5).<br>**Title of master thesis:** Genetische Untersuchungen des *ADCY5*-Gens bei Patienten mit Dyskinesien sowie funktionelle Studien an Neuronen aus induzierten pluripotenten Stammzellen von Patienten mit X-chromosomal vererbten Dystonie-Parkinson-Syndrom.<br>Assigned and supervised by Prof. Dr. Christine Klein, MD, at the Institute of Neurogenetics |
| October 2010 – November 2013 | **University of Lübeck**, *Lübeck*, Bachelors degree program Molecular Life Science, B.Sc (2,1).<br>**Title of bachelor thesis:** Generierung und experimentelle Charakterisierung von BVDV Replikonderivaten für Life Cell Imaging und Proteomics.<br>Assigned and supervised by Prof. Dr. Norbert Tautz, PhD, at the Institute of Virology and Cell Biology |
| October 2001 – Mai 2010 | **Gymnasium Schwarzenbek**, *Schwarzenbek*, Abitur (2,3).<br>Advanced level subject biology and mathematics |

## Internships

April – June 2015 **Univerity of Lübeck**, *Lübeck*, internship at the Institute of Integrative und Experimental Genomics, consolidation in bioinformatics.
**Topic:** Stratifkation und Risikoabschätzung des Myokardinfarktes durch genetische Marker klassischer Riskofaktoren. Assigned and supervised by Prof. Dr. Jeanette Erdmann, PhD.

November 2014 – February 2015 **University of Lübeck**, *Lübeck*, internship at the Institute of Neurogenetics, consolidation in cell culture, stem cell biology and molecular genetics.
**Topic:** Charakterisierung humaner induzierter pluripotenter Stammzellen von Patienten mit X-chromosomalem Dystonie-Parkinson-Syndrom und philippinischen Kontrollen. Assigned and supervised by Prof. Dr. Christine Klein, MD.

## Work Experience

Since June 2019 **University of Lübeck**, *Lübeck*, 1st Dept. of Medicine, Postdoc Researcher.

October 2015 - March 2016 **University Clinic of Schleswig-Holstein (UKSH)**, *Lübeck*, Institute of Neurogenetics, student assistant.
Tasks: DNA and serum/plasma extraction from blood samples for the special research field (*Sonderforschungsbereich* SFB) 134 of the clinic of neurology

November 2013 – July 2014 **Universität zu Lübeck**, *Lübeck*, Institut für Theoretische Informatik (TCS), academic assistant.
Tasks: Supervision of tutorials in the context of two bachelor courses *Informatik für MLS - 1* and *Informatik für MLS - 2*

October 2011 – September 2015 **University of Lübeck**, *Lübeck*, Zentrale Hochschulbibliothek (ZHB), student assistant.
Tasks: Evening service at the lending agent

## Certifications

September 2019 RNA-Seq Data Analysis Workshop: Quality control, Read Mapping, Visualization and Downstream Analyses (ecSeq Bioinformatics)

November 2018 DNA Methylation Data Analysis: How to use bisulfite-treated sequencing to study DNA methylation (ecSeq Bioinformatics)

## Publications

2019 Krause, C., Sievert, H., Grohs, M., Geißler, C., El Gammal, A., Wolter, S., Ohlei, O., Kilpert, F., Krämer, U., Kasten, M., Klein, C., Brabant, GE, Mann, O., Lehnert, H., Kirchner, H. (2019) Critical evaluation of the DNA-methylation markers ABCG1 and SREBF1 for Type 2 diabetes stratification. Epigenomics (Epub ahead of print).

2018 Krause, C., Grohs, M., El Gammal, A.T., Wolter, S., Lehnert, H., Mann, O., Mittag, J., and Kirchner, H. (2018). Reduced expression of thyroid hormone receptor $\beta$ in human nonalcoholic steatohepatitis. Endocrine Connections 1448–1456.