Causes and consequences of hyperglycemia and glycosuria.

Ottosson Laakso, Emilia

2016

Document Version:
Other version

Causes and consequences of hyperglycemia and glycosuria

Causes and consequences of hyperglycemia and glycosuria

Emilia Ottosson Laakso

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at CRC Aula, Jan Waldenströms gata 35, Skåne University Hospital, Malmö, on May 31st at 9.00.

Faculty opponent
docent Ingrid Dahlman, Department of Medicine, Huddinge
Karolinska Institutet

Organization LUND UNIVERSITY	Document name	
	Date of issue	
Author(s)	Sponsoring organization	
Title and subtitle		
Abstract		
Key words		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		Language
ISSN and key title		ISBN
Recipient's notes	Number of pages	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sourcespermission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature _____Date _____

Causes and consequences of hyperglycemia and glycosuria

Emilia Ottosson Laakso

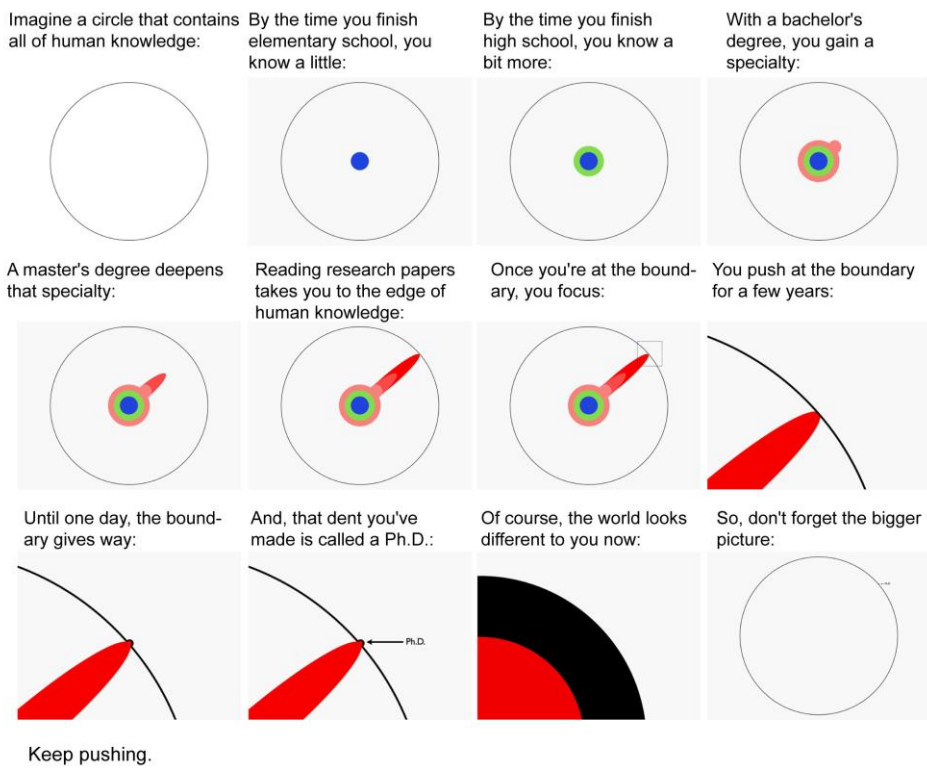
LUND
UNIVERSITY

Cover photo by Jasmina Kravic

Faculty of Medicine| Department of Clinical Sciences, Malmö

ISBN 978-91-7619-295-5
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2016



Matt Might (http://matt.might.net/), a professor in Computer Science at the University of Utah, created The Illustrated Guide to a Ph.D (http://matt.might.net/articles/phd-school-in-pictures/. to explain what a Ph.D. is to new and aspiring graduate students.

Content

Abbreviations

A	adenine
C	cytosine
cDNA	complementary DNA
CNV	copy number variation
DNA	deoxyribonucleic acid
eQTL	expression quantitative trait loci
ER	endoplasmic reticulum
FRG	familial renal glycosuria
G	guanine
GLUT	glucose transporter
GWAS	genome wide association study
HbA1c	glycated haemoglobin
IFG	impaired fasting glucose
IGT	impaired glucose tolerance
INS	insulin
MAF	minor allele frequency
MODY	maturity-onset diabetes of the young
MPDZ	multiple PDZ domain protein
NGS	next generation sequencing
NGT	normoglycemia
OGTT	oral glucose tolerance test
qPCR	quantitative real time PCR
RNA	ribonucleic acid

ROS	reactive oxygen species
SGLT	sodium glucose co-transporters
SI	stimulatory index
SLC5A2	solute carrier family 5 member 2
SNP	single nucleotide polymorphism
sQTL	splicing quantitative trait loci
T	thymine
T2D	type 2 diabetes
WGS	whole genome sequencing

Abstract

Type 2 diabetes is one of the world's major challenges today. Over 400 million people around the globe are diagnosed with diabetes and the consequences to the individual patient and health care systems are significant. Diabetes is a chronic disease and the standard treatment including lifestyle changes, metformin and insulin is in many cases not effective enough. New anti-diabetic drugs have been developed aiming to lower blood glucose by inhibition of the glucose re-uptake via the sodium glucose co-transporter 2 (SGLT2) from urine, leading to excessive glycosuria. The potential effects of this treatment on the prevention of glucose intolerance have not been addressed in previous studies. The SGTL2-inhibitors mimic a condition where mutations in the *SLC5A2* gene, which encodes the SGLT2 transporter, give rise to glycosuria. In study 1 we investigated the effect of chronic glycosuria on glucose tolerance over time in a family with mutations in the *SLC5A2* gene and found that despite their life-long "SGLT2-inhibition" there was no effect of chronic glycosuria on the development of glucose intolerance.

There is a strong genetic component of diabetes that has been investigated using linkage studies and genome-wide association studies but only a small part of the estimated heritability has been explained by these findings. In families with high incidence of diabetes we can search for part of the missing heritability in the form of rare, family specific mutations. We show in study 2 that a previous linkage on chromosome 9 in families enriched with type 2 diabetes might in part be explained by rare variants in the multiple PDZ domain containing protein (*MPDZ*) gene. Knock down of the gene in an insulin secreting rat beta cell line resulted in impaired glucose-stimulated insulin secretion.

A vital part of glucose homeostasis is the regulation of blood glucose by insulin. The ultimate characteristic of type 2 diabetes is the failure of the pancreas to produce insulin in response to increased glucose demands. Studies of gene expression in islets of Langerhans may provide an answer to why this occurs. Study 3 compared the global gene expression in islets from human type 2 diabetic and non-diabetic donors. Over 1500 genes were differentially expressed in diabetic islets, e.g. the RAS guanyl releasing protein 1 (*RASGRP1*) was negatively associated with diabetes and positively associated with insulin secretion. Of the genes associated with diabetes we found that the expression of 35 genes was influenced by genetic variants and silencing of the gene tetraspanin 33 (*TSPAN33*), 5'-nucleotidase, ecto (*NT5E*), transmembrane emp24

protein transport domain containing 6 (*TMED6*) and p21 protein activated kinase 7 (*PAK7*) in a rat beta cell line resulted in reduced glucose-stimulated insulin secretion.

Glucose is a potent regulator of gene expression and the regulation of genes by elevated glucose might further impair insulin secretion. In study 4 we aimed to separate the causes from the consequences of hyperglycemia on islet gene expression by global transcriptome analysis of islets exposed to short-term glucose (18.9 mmol/l glucose for 24 hours). We then compared the changes in gene-expression seen in patients with chronic hyperglycemia (diabetes) with genes regulated by short-term glucose exposure with the assumption that genes whose expression change after short-time hyperglycemia may reflect consequences rather than causes of hyperglycemia. This resulted in about 400 genes likely to be pathogenically involved in the development of hyperglycemia. For example the ERO1-Like Beta (*ERO1LB*) gene was down-regulated in islets from diabetic donors and correlated positively with insulin secretion whereas the transmembrane protein 132C (*TMEM132C*) gene that was up-regulated in islets from diabetic donors and correlated negatively with insulin secretion.

Populärvetenskaplig sammanfattning

Diabetes är idag en av världens stora hälsoutmaningar. Världsdiabetesorganisationen har uppskattat att över 400 miljoner människor är drabbade världen över vilket har enorma konsekvenser inte bara för den enskilda patienten men även för sjukvården. Typ 2 diabetes är den vanligaste formen av diabetes och den drabbar ofta vuxna som en följd av övervikt, låg fysisk aktivitet och genetiska faktorer som leder till förhöjda nivåer av socker i blodet. Insulin är ett livsviktigt hormon som signalerar till kroppen att den ska ta upp socker ur blodet. Försämrad funktion eller utsöndring (insulinsekretion) gör att sockret inte längre kan tas upp i t ex muskler och lever där det används som energi. Typ 2 diabetes är en kronisk sjukdom som i första hand behandlas med livsstilsförändringar, metformin och ibland insulin. Tyvärr är detta i många fall inte är tillräckligt effektivt för att nå behandlingsmålen. Därför söker man ständigt efter nya behandlingsmetoder och nyligen godkändes de så kallade SGLT2-inhibitorerna som behandling av typ 2 diabetes. De syftar till att sänka sockernivåerna i blodet genom att orsaka utsöndring av sockret via urinen, så kallad glykosuri.

I det första arbetet i denna avhandling har vi försökt svara på frågan om SGLT2 inhibition kan användas för att förhindra utveckling eller försämring av typ 2 diabetes över tid. SGLT2 inhibitorerna härmar ett naturligt förekommande tillstånd där förändringar i arvsmassan, så kallade mutationer, i den glukostransportör i njuren som normalt ser till att allt socker stannar i kroppen försämrar funktionen av transportören vilket leder till att man istället kissar ut delar av sockret. Det är ett ärftligt tillstånd och genom att studera en stor familj och jämföra de med och utan mutation kan vi se vad naturlig, kronisk ”SGLT2 inhibition” har för effekter på typ 2 diabetes över lång tid. Vi fann att det i vårt studiematerial inte fanns någon påverkan av glykosuri på utvecklingen av förhöjda blodsockernivåer och typ 2 diabetes.

Nästa arbete i avhandlingen syftade till att hitta de genetiska orsakerna till typ 2 diabetes i en annan stor familj med många drabbade familjemedlemmar. Vi har tidigare sett att en viss del av kromosom 9 är kopplad till diabetes i familjen, men vi kunde på grund av tekniska begränsningar inte identifiera vilken gen i detta område som är orsaken till typ 2 diabetes. Med hjälp av ny sekvenseringsteknik där vi kan titta på nästan alla mutationer i området har vi hittat genetiska varianter i en gen kallad *MPDZ* som är associerade med diabetes i familjen. Om man ändrar nivån av genen i ett modellsystem ser man att insulinsekretionen påverkas vilket pekar på att genen är inblandad i utvecklingen av diabetes.

De Langerhanska öarna i bukspottkörteln producerar allt kroppens insulin och deras funktion är essentiell för regleringen av blodsockernivåerna. Att förstå vad som går fel i öarna när diabetes uppstår är viktigt för att kunna hitta nya sätt att behandla eller förhindra sjukdomen. Socker är en nyckelsignal för de Langerhanska öarna och försämringar i deras förmåga att svara på signalen är en av de viktigaste orsakerna till diabetes. I studie tre och fyra undersöker vi hur sockret påverkar öarnas gen-uttryck på kort och lång sikt, som vid diabetes, och försöker ta reda på vilka gener som är förändrade som följd av och vilka som bidrar till förhöjda blodsockernivåer. Vi försökte också förstå hur genetiska varianter kan påverka genernas uttryck och hur detta är kopplat till förändrat genuttryck vid diabetes. Socker ändrade flera tusen geners nivåer i öarna som svar på kortvarig exponering och många av dessa var även förändrade vid diabetes. Vissa av de gener vars uttryck var förändrat vid diabetes var också kopplade till öarnas förmåga att utsöndra insulin, vilket kan ligga bakom den försämrade regleringen av blodsockernivåerna som leder till diabetes.

Uttrycket “alla vägar leder till Rom” är sant även för typ 2 diabetes. Flera olika faktorer såsom genetiska och miljöfaktorer har visat sig vara involverade i utvecklingen av diabetes, men det finns fortfarande mycket att lära. Denna avhandling har bidragit med ny kunskap om potentiella mekanismer bakom typ 2 diabetes och förhoppningsvis kan den vara till hjälp i framtida studier.

List of Publications

Scientific Papers Contributing to this Thesis

I. **Ottosson-Laakso E.**, Tuomi T., Forsén B., Gullström M., Groop P-H., Groop L., Vikman P. *Influence of Familial Renal Glycosuria due to Mutations in the SLC5A2 Gene on Changes in Glucose Tolerance Over Time*, Plos ONE 2016 Jan 6;11(1):e0146114.

II. **Ottosson-Laakso E.**, Bennet H., Fex M., Mattisson H., Sarelin L., Dwivedi OP., Kvist A., Almgren A., Isomaa B., Groop L., Vikman P. *Contribution of variants in the MPDZ gene to familial clustering of abnormal glucose intolerance*, Manuscript.

III. Fadista J., Vikman P., **Ottosson Laakso E.**, Mollet I., Esguerra J., Taneera J., Storm P., Osmark P., Ladenvall C., Prasad R., Hansson K., Finotello F., Uvebrant K., Ofori J., Di Camillo B., Krus U., Cilio C., Hansson O., Eliasson L., Rosengren A., Renström E., Wollheim C., Groop L. *Global Genomic and Transcriptomic Analysis of Human Pancreatic Islets Reveals Novel Genes Influencing Glucose Metabolism.* PNAS 2014 Sep 23;111(38):13924-9.

IV. **Ottosson-Laakso E.**, Krus U., Storm P., Oskolkov N., Ahlqvist E., Fadista J., Hansson O., Groop L., Vikman P. *Glucose-induced Changes in Gene Expression in Human Pancreatic Islets – Causes or Consequences of Chronic Hyperglycemia.* Manuscript.

Publications not Included in this Thesis

Bennet H, Mollet IG., Balhuizen A., Medina A., Nagorny C., Bagge A., Fadista J., **Ottosson-Laakso E.,** Vikman P., Dekker-Nitert M., Eliasson L., Wierup N., Artner I., Fex M.. *Serotonin (5-HT) receptor 2b activation augments glucose-stimulated insulin secretion in human and mouse islets of Langerhans.* Diabetologia. 2016 Apr;59(4):744-54.

Bennet H., Balhuizen A., Medina A., Dekker Nitert M., **Ottosson Laakso E.**, Essén S., Spégel P., Storm P., Krus U., Wierup N., Fex M. *Altered serotonin (5-HT) 1D and 2A receptor expression may contribute to defective insulin and glucagon secretion in human type 2 diabetes.* Peptides. 2015 Sep;71:113-20.

Nagaraj V., King B., Storm P., Vikman P., **Ottosson-Laakso E.**, Blom AM., Renström E. *Complement inhibitor CD55 governs the integrity of membrane rafts in pancreatic beta cells, but plays no role in insulin secretion.* Biochem Biophys Res Commun. 2015 May 8;460(3):518-24.

Taneera J., Fadista J., Ahlqvist E., Atac D., **Ottosson-Laakso E.**, Wollheim CB., Groop L. *Identification of novel genes for glucose metabolism based upon expression pattern in human islets and effect on insulin secretion and glycemia.* Hum Mol Genet. 2015 Apr 1;24(7):1945-55.

Zhou Y., Park SY., Su J., Bailey K., **Ottosson-Laakso E.**, Shcherbina L., Oskolkov N., Zhang E., Thevenin T., Fadista J., Bennet H., Vikman P., Wierup N., Fex M., Rung J., Wollheim C., Nobrega M., Renström E., Groop L., Hansson O. *TCF7L2 is a master regulator of insulin production and processing.* Hum Mol Genet. 2014 Dec 15;23(24):6419-31.

Krus U., King BC., Nagaraj V., Gandasi NR., Sjölander J., Buda P., Garcia-Vaz E., Gomez MF., **Ottosson-Laakso E.**, Storm P., Fex M., Vikman P., Zhang E., Barg S., Blom AM., Renström E. *The complement inhibitor CD59 regulates insulin secretion by modulating exocytotic events.* Cell Metab. 2014 May 6;19(5):883-90.

Introduction

Diabetes Mellitus – an overview

Diabetes mellitus is a complex disease that was probably first described by the ancient Egyptians around 1500 BC as ”too great emptying of the urine” [1]. With the discovery of the importance of the pancreas in the development of diabetes and the identification of insulin as the key hormone necessary to avoid diabetes, an effective treatment for the disease was finally developed in the twentieth century. This meant that the diagnosis no longer translated into certain death. Today we know that there are many forms of diabetes, the two most common ones are type 1 diabetes and type 2 diabetes, and there are several treatments available depending on the type. However, diabetes is still a severe disease that is affecting millions of people around the globe and every 6 seconds someone dies as a consequence of diabetes (IDF diabetes atlas 7th edition).

The International Diabetes Federation (IDF) estimates that there are 415 million people with diabetes today, and that this will increase to 642 million by 2040 (IDF diabetes atlas 7th edition). This means that one in every 10 adults has diabetes and almost half of these are undiagnosed. Due to the number of patients and the complications of diabetes, treatment and prevention takes a substantial part of the world health expenditure; global costs are estimated to 673 billion USD. The consequences are detrimental to the society and the patients; the complications the patient is facing include nephropathy that can culminate in end-stage renal disease, neuropathy and peripheral artery disease that might lead to amputation, retinopathy that could cause blindness, and an increased risk of cardiovascular disease. It was estimated by the IDF that 15 million people died from diabetes and its complications in 2015. The reasons why diabetes incidence has exploded in all parts of the world in only a couple of decades are partly because of sedentary lifestyle and diet. However, the individual risk is also determined by genetic variants. The genetic component of diabetes has been investigated using linkage studies and genome-wide association studies, but only a small part of the estimated heritability can be explained by the findings from these studies. Using of families with high incidence of diabetes, we can search for part of the missing heritability that is believed to be hiding in rare, family specific mutations.

Pathogenesis of diabetes

Symptoms and diagnosis

Diabetes often presents with increased thirst, polyuria, glycosuria, blurred vision and weight loss. T1D and some other forms may also induce ketoacidosis that can lead to coma and, in the worst case, death. Ketoacidosis is not so common in type 2 diabetes since the onset of the disease is slow and defects in glucose metabolism may be present for years before diagnosis [7]. The diagnostic criteria for diabetes are the same for all diabetes types and the classification into T1D, T2D, other specified types and GDM is then based on both clinical stages and aetiology. Diagnostic criteria for diabetes according to the World Health Organization [7, 8] are summarized in Table 2. Two prediabetic stages, impaired fasting glucose (IFG) and impaired glucose tolerance (IGT), are also diagnosed based on the glycemic measurements (Table 2). WHO recommends the use of an oral glucose tolerance test (OGTT) to assess the blood glucose at fasting and 2 hours post glucose load to allow for the detection of patients with IGT although the measurement of glycated haemoglobin A1c (HbA1c) is also acceptable [7, 8]. The diagnosis should in asymptomatic patients always be confirmed with subsequent glucose measurements.

Table 2. Diagnostic criteria for diabetes mellitus, IFG and IGT.

	Fasting (mmol/l)	2 h post glucose load (mmol/l)	HbA1c (%)
Impaired Fasting Glucose (IFG)	≥6.1 and <7.0	<7.8 (if measured)	-
Impaired Glucose Tolerance (IGT)	<7.0 (if measured)	≥7.8 and <11.1	-
Diabetes Mellitus	≥7.0	≥11.1	6.5

The values are referring to venous plasma samples

Genes and the environment

Type 2 diabetes is commonly associated with obesity and physical inactivity [9, 10]. The changes in lifestyle and increasing age of the population are the main reasons why T2D has gone from being a relatively uncommon disease to one of the largest health issues in the world. It is also clear that the development of diabetes is strongly influenced by the genetic factors. One hypothesis is that our genes have been under selection pressure in an environment where food was sparse and where we had to spend a considerable amount of energy to hunt and scavenge for food. This would result in genes that favor energy storage and uptake. Today when we instead have an abundance of food and a sedentary lifestyle our genes would work against us and predispose us to metabolic diseases such as T2D [11].

Insulin resistance and beta cell failure

Insulin resistance is a keystone in the development of T2D. Insulin action is impaired when target tissues such as liver, skeletal muscle and adipose tissue fail to respond to insulin. The consequence is decreased glucose uptake of the peripheral tissues and insufficient suppression of endogenous glucose production in the liver [12]. In the early stages of disease the blood glucose levels are controlled by a compensatory hyperinsulinemia. The compensatory insulin secretion will have to fail in order for overt T2D to develop. A reduction in beta cell function is already present early in the progression to T2D. Groups with increased risk of T2D such as first-degree relatives of diabetics and elderly people show beta cell dysfunction. Beta cell function is also heritable which leads to differences in progression and risk of T2D in different ethnic groups [13].

Pharmacological treatment of diabetes

As the incidence of diabetes has changed so has the need to prevent and treat diabetes. The only effective treatment of insulin-deficient diabetes is with injections of biosynthetic insulin. However, Harold Percival Himsworth published a paper in Lancet 1936 demonstrating that there were two types of diabetes patients; the insulin-sensitive and the insulin-insensitive one. He administrated intravenous insulin to his patients simultaneously with oral glucose and measured the blood glucose concentrations finding two very different types of curves in different patients. He also demonstrated that this was due to impaired glucose disposal in the peripheral tissues in the insulin-insensitive patients [2]. The insulin-insensitive patient proposes a new challenge for the therapy; simply administering insulin will not be enough to relieve the symptoms.

Drugs targeting different mechanisms and complications have been developed along with methods for home glucose monitoring, surgical methods for islet and kidney transplantation, laser therapy for retinopathy, and bariatric surgery to control weight and induce disease remission. Among the first oral anti-diabetic drugs were the biguanides that are derived from guanidine. The French lilac was said to decrease the frequent urination seen in diabetes mellitus, and it turned out to be the isoamylene guanidine that was the active substance [3]. Guanidine hydrochloride was in 1918 tested in rabbits by Watanabe for its hypoglycemic effect with positive results [4], but due to the toxicity of guanidine it could not be used as a treatment. The biguanides were developed, and metformin was released for clinical use in the 1950s. This is still today one of the most common treatments of diabetes even though the precise mechanism of action is debated. The hypoglycemic effect of sulfonylureas was discovered as a side effect when the chemist Marcel Janbon was studying the effect of sulfonamide compounds on typhoid fever [5]. This was developed into the drug

tolbutamide that was released in the United States in 1956 [1]. Later other oral drugs have also been added to the medical toolkit for treatment of diabetes, such as the SGLT2 inhibitors that will be discussed later (Table 1) [6].

Table 1. Drugs used in the treatment of diabetes mellitus.

Drug class (compound names)	Mechanism of action	Biological effect
Insulin	Activates insulin receptor	• Increase glucose uptake in peripheral tissue • Decrease hepatic glucose production
Biguanides (metformin)	Activates AMP-kinase, ?	Decrease hepatic glucose production
Sulfonylurea (Glipizide, Glimepiride…)	Interacts with K_{ATP} channels on the beta cell	Increase insulin secretion
Megalitinides (Repaglinide, Nateglinide)	Interacts with K_{ATP} channels on the beta cell	Increase insulin secretion
Thiazolidinediones (Pioglitazone, Rosiglitazone)	Activates PPAR-gamma	Increase insulin sensitivity
α-glucosidase inhibitors (Acarbose, Miglitole, Voglibose)	Inhibits α-glucosidase	Decrease intestinal carbohydrate uptake
DPP-4 inhibitors (Sitagliptin, Vildagliptin…)	In hibits DPP-4 activity -> increasing GLP-1 and GIP concentrations	• Increase insulin secretion • Decrease glucagon secretion
Bile acid sequestrants (Colesevelam)	Binds bile acids in the intestinal tract -> increased bile acid production	?
Dopamine-2 agonists (Bromocriptine)	Activates dopaminergic receptors	• Increase insulin sensitivity • Regulates hypothalamic effects on metabolism • ?
GLP1-receptor agonists (Exenatide, Liraglutide)	Activate GLP1-receptors	• Increase insulin secretion • Decrease glucagon secretion • Increase satiety
Amylin agonists (Pramlintide)	Activates amylin receptors	• Increase glucose uptake in periferal tissue • Decrease heptic glucose production
SGLT2 inhibitors (Canagliflozin, Dapagliflozin, Empagliflozin)	Inhibits glucose reuptake in the kidney -> glucose excretion through urine	Decrease blood glucose concentration

Next generation sequencing and genetics

The human genome consists of over 3 billion base pairs. These base pairs contain the code for the building of our bodies. The code is written in the form of deoxyribonucleic acid (DNA) with only four different molecules (nucleotides); adenine (A), cytosine (C), guanine (G) and thymine (T). The sequence of the nucleotides determines how the genome will be read and what proteins will be produced. Being able to read and translate the code and the variations responsible for differing phenotypes is of great importance in understanding and treating disease. The human genome project was started in 1990 with the aim to map the genome and provide a guide, or a reference, to the sequence of the DNA [14], which then could be used to identify variations or mutations in the sequence that would lead to disease. The project was announced complete in 2003, presenting the sequence of 99% of the human genome as well as the location of all the genes known at that time [15]. Reading the sequence of base pairs was mostly done at this time using Sanger sequencing, where fragments of DNA were copied *in vitro* using labeled nucleotides. A well-trained technician could produce around 2000 base pair worth of data per day from a single machine (https://www.genome.gov/10001477). With the advent of next generation sequencing (NGS) we can today produce 1.8×10^{12} bases in less than three days (http://www.illumina.com/systems/hiseq-x-sequencing-system/performance-specifications.html). This means that we can now sequence a human genome 562 times on one single machine in three days. This technique can be used not only to sequence DNA but also the ribonucleic acid (RNA) molecules. RNA is the basics unit of our transcriptome and the first step from genotype to phenotype. It partly regulates the response to environmental factors that might be perturbed in disease. The transcriptome is highly dynamic in order to be able to respond to the needs of the body. NGS can measure these changes in a global way, which can provide vital information on how an organism responds to changes in the environment or even how variations in the genome affect the transcription of genes.

Candidate genes

One of the earliest approaches to genetics in human disease was the candidate gene approach. It is based on association of alleles of a gene with the disease in question. This approach has certain pre-requisites; 1. Knowledge about biologically relevant genes for the disease or a linked region, 2. Knowledge about genetic variation in the gene, and 3. A cohort of related or unrelated people with the disease (cases) and matched people without the disease (controls) [16]. In T2D there are six genes identified using this method that have been consistently associated with the disease table 3 [17]. One example is the wolframine ER transmembrane glycoprotein (*WFS1*)

gene that was first linked to the Wolfram Syndrome [18, 19] (MIM:222300), which is characterized by juvenile diabetes and optic atrophy (affecting sight). The diabetes in the Wolfram Syndrome is insulin-requiring and selective destruction of beta cells in patients with the disease pointed to a function of *WFS1* in beta cell survival and mutations cause beta cell death [20]. Therefore it was suggested that this gene might harbor common genetic variants that also increased the risk of T2D. Association studies of variation in *WFS1* did indeed provide evidence of genetic susceptibility to T2D in carriers [21, 22].

Table 3. Genes associated with T2D identified in candidate gene studies

Gene ID	Gene name	Variant(s)	OR* T2D
HNF1B	HNF1 homeobox A	rs1920792, rs7501939, rs757210, rs4430796	1.1-1.17
IRS1	Insulin receptor substrate 1	rs2943641, rs7578326	1.09-1.12
KCNJ11	Potassiuminwardly-rectifying channel, subfamily J, member 11	rs5219	1.09-1.14
PPARG	Peroxisome proliferator-activated receptor gamma	rs1801282, rs13081389	1.14-1.24
WFS1	Wolframine ER transmembrane glycoprotein	rs10010131, rs6446482, rs1801214	1.11-1.13
HNF1A	HNF1 homeobox A	Rs2071190, rs7957197	1.07-1-14

*OR – Odds ratio

Linkage studies

Another method to discover chromosomal regions likely to contain causal genes is to perform linkage analysis in affected pedigrees by covering the entire genome with a few hundred markers. A region is linked if markers (often microsatellites or single nucleotide polymorphisms/SNPs) are inherited together from the same parent more often than would be expected [23]. If the markers are shared by affected family members this is a sign that the region contains a disease-causing variant, which is linked to the markers. In T2D only two loci have been identified using linkage analysis; calpain 10 (*CAPN10*) [24, 25] and transcription factor 7 like 2 (*TCF7L2*) [26-28].

Genome-Wide Association Studies

With the introduction of affordable and fast genotyping of thousands to millions of markers simultaneously a new type of analysis could be performed, the genome-wide association study (GWAS). This method is population based and thousands of

affected as well as healthy controls are genotyped on micro-array platforms. The single nucleotide polymorphisms (SNPs) are then tested for association with disease in a genome-wide manner [29]. The advantage of GWAS is that there is no need for prior knowledge about disease causing loci. The use of common variation found across different populations allows for meta-analysis of GWAS results, increasing sample sizes and thereby power to detect association. The combined efforts of GWAS and meta-analysis have provided novel information about over 80 genetic loci associated with T2D [30].

Targeted resequencing and whole genome resequencing

While most of the common variation associated with disease can be captured by GWAS, the rare and novel variants are missed due to the low frequency. The best way to find these mutations is by whole genome sequencing (WGS). WGS is designed to cover almost the entire genome, which will provide information about all the variants in that genome. Even though the cost for sequencing has plummeted since the introduction of NGS it is still often prohibitively expensive to do WGS in a large number of individuals. By targeting specific, smaller regions, such as the coding parts of the genome or previously identified linkage regions, larger numbers of individuals can be sequenced at a lower cost. One example where NGS was successfully used to provide new information about disease is the discovery of rare loss-of-function variants in the solute carrier family 30 member 8 (*SLC30A8*) that were shown to be protective against T2D [31].

Whole transcriptome sequencing

Another important tool in the dissection of molecular causes of diabetes is whole transcriptome (RNA) sequencing. NGS allows for unbiased quantification of transcripts without any prior knowledge about what genes are transcribed, their sequence or splice patterns [32]. The most common use of RNA sequencing is global gene expression analysis, simultaneously measuring all transcripts in the sample, but since the technique also provides information about the nucleotide sequence of the transcripts it can be used to detect SNPs, splice variation, allelic imbalance and RNA editing events, where post transcriptional editing changes the nucleotide composition of the RNA molecule. RNA sequencing is also used in combination with genotype information to find expression quantitative trait loci (eQTL) where the genotype influences the expression of genes close by (cis eQTL) or further away (trans eQTL).

The genetics of diabetes

Diabetes is a complex polygenic disease. The genetic component of T2D is relatively strong. Studies in twins have shown that monozygotic twins have about 70% concordance rate (both twins being affected) of T2D whereas it is only 20-30% in dizygotic twins [33-35]. In addition, the estimated heritability (the proportion of variation in a phenotype that can be explained by the variation in genotype) of T2D is between 64-79% [33, 36, 37] and the lifetime risk of developing T2D if one or both parents are affected is 40% and 70%, respectively [38, 39]. Despite the identification of over 80 loci associated with T2D and another 40 or so loci associated with related traits such as fasting glucose, insulin levels and HbA1c the explained heritability is less than 15% [40]. The missing heritability might be due to a number of factors discussed below [40].

Monogenic forms of diabetes

In some cases of diabetes the cause is a mutation in a single gene. These cases are important to identify, because insulin may not be the best option to stabilize glycemia in these patients. However, it is estimated that a large number of these patients are misdiagnosed with T1D or T2D [41]. There are two major forms of monogenic diabetes, neonatal (NDM) and maturity-onset diabetes of the young (MODY).

If diagnosis of diabetes has been done before the age of 6 months it is called neonatal diabetes (NDM) [41]. The majority of NDM can be explained by single gene mutations. Some of the genes affected in NDM are ATP binding cassette subfamily C member 8 (*ABCC8*) coding for SUR1, potassium voltage-gated channel subfamily J member 11 (*KCNJ11*) coding for Kir6.2 and insulin (*INS*). [42].

MODY is a dominantly inherited form of diabetes with an age of onset typically below 25 years [42]. Most of the genes mutated in MODY affect beta cell function, and are summarized in Table 4 [43].

Table 4. The different forms of MODY, the genes responsible and function of the genes.

MODY	Gene (gene symbol)	Function
1	Hepatocyte nuclear factor 4 alpha (*HNF4A*)	Transcription factor – decreased insulin secretion
2	Glucokinase (*GCK*)	Phosphorylates glucose – decreased glucose sensitivity and glycogen storage
3	HNF1 homeobox A (*HNF1A*)	Transcription factor – decreased insulin secretion
4	Pancreatic and duodenal homeobox 1 (*PDX1*)	Transcriptional activator – impaired pancreas development
5	HNF1 homeobox B (*HNF1B*)	Transcription factor – decreased insulin secretion
6	Neuronal differentiation 1 (*NEUROD1*)	Transcription factor – abnormal development of beta cell function
7	Kruppel-like factor 11 (*KLF11*)	Tumor-supressor gene – decreased glucose sensitivty of beta cells
8	Carboxyl ester lipase (*CEL*)	Lipase – Decreased endocrine function
9	Paired box 4 (*PAX4*)	Transcription factor – apoptosis and proliferation of beta cells
10	Insulin (*INS*)	Defect insulin secretion
11	BLK proto-oncogene, Scr family tyrosine kinase (*BLK*)	Defect insulin secretion
12	ATP binding cassette subfamily C member 8 (*ABCC8*)	Ion channel subunit – decreased insulin secretion
13	Potassium voltage-gated channel subfamily J member 11 (*KCNJ11*)	Ion channel subunti – decreased insulin secretion

The common genetic variation in diabetes

Due to the nature of GWAS design all of the variants identified by GWAS are common and have a minor allele frequency (MAF) above 1%. The most strongly associated gene is *TCF7L2,* which was first identified in linkage studies where the risk allele confers a relative risk of approximately 1.4 compared to non-carriers of the allele [44]. The first results from GWAS of T2D were published in 2007, where hematopoietically expressed homeobox (*HHEX*) and solute carrier family 30 member 8 (*SLC30A8*) were identified as novel susceptibility loci [45]. The *HHEX* locus also contains the gene encoding *IDE* (insulin degrading enzyme). This was the start of a sequence of publications from around Europe that added cyclin-dependent kinase inhibitor 2A/2B (*CDKN2A*/*CDKN2B*), insulin like growth factor 2 mRNAbinding protein 2 (*IGF2BP2*) and CDK5 regulatory subunit associated protein 1 like 1 (*CDKAL1*) to the list of novel loci [46-49]. For a complete review of the GWAS and meta-analysis loci see Ahlqvist et al and Brunetti et al [17, 30].

The missing heritability – the role of rare variants, epistasis, epigenetics and the environment

Despite the large number of genetic loci identified to be involved in T2D these variants explain less than 15% of the estimated heritability. The still unknown genetic factors influencing the risk of T2D are called the missing heritability [50]. It has been proposed that the heritability is in fact overestimated due to genetic interactions where one gene interacts with another in a non-additive way (epistasis) [51]. It is however not likely to explain all of the missing heritability and other explanations include rare variants, structural mutations, parent-of-origin effects and gene-environmental interactions [40].

The common variants captured by GWAS usually have a low effect size, since they otherwise would most likely have been removed from the gene pool by selection, while more rare variants (MAF<5%) may have larger effect size. It is more difficult to discover these variants and larger sample sizes are needed. Structural variants such as copy number variations (CNVs) are also poorly captured by the conventional genotyping arrays and may hide some of the missing heritability together with undiscovered rare variants.

Parent-of-origin effects occur when the phenotypic effect of an allele is dependent on whether it is inherited from the mother or from the father [52]. The risk of T2D in the offspring is greater if the mother is affected than the father, as has been shown for variants in *KCNQ1* [53] and *KLF14* [54]. If the allele inherited from the other parent is protective the net result in a population study will be no effect at all.

Part of the heritability might also be hidden in gene-environmental interactions. They occur when the response to the environment of two alleles are different [55]. If, for example, one allele at a specific locus would increase the risk of obesity only if the person would eat a high-fat diet, this effect would be seen in one environment but not another, complicating the discovery of the risk allele.

Glycosuria

Renal glycosuria is characterized by decreased reabsorption of glucose in the kidney. Glucose is normally filtered out from the blood into the urine by the functional unit of the kidney, the nephron, together with other small molecules such as amino acids and salts. There are thousands of nephrons in the kidneys and around 180 l of plasma is filtered every day, most of this never make it to the urine thanks to an intricate interplay of reabsorption (passive or active) and secretion. The reabsorption of glucose in the proximal tubule by sodium glucose transporters (SGLT) is part of this system

and about 180 g of glucose is filtered and then transported back into the bloodstream per day, leaving essentially no glucose in the urine. When the glucose concentration in the filtrate exceeds the capacity for reabsorption glucose excretion in the urine occurs and glycosuria ensues. This might be the consequence of elevated blood glucose levels (as seen in diabetes), which in effect increases the amount of glucose in the kidney filtrate, or it can be caused by reduced kidney function in which case it is referred to as renal glycosuria. In a small number of cases of renal glycosuria there is a genetic cause behind the condition. This familial renal glycosuria (FRG) often displays a co-dominant inheritance pattern with incomplete penetrance [56] and most published cases are due to mutations in the gene coding for the main glucose transporter in the kidney; solute carrier family 5 member 2 (*SLC5A2*) [56-61]. The gene product of *SLC5A2* is the sodium-glucose co-transporter 2 (SGLT2), which together with SGLT1 is responsible for almost all of the glucose reabsorption in the kidney.

Sodium-glucose co-transporters in the kidney

SGLT1

The sodium-glucose co-transporter 1 (SGLT1) was the first of the SGLT proteins to be discovered. It is an active glucose transporter that uses the sodium gradient to co-transport glucose against the glucose concentration gradient. For every two sodium molecules, one glucose molecule is co-transported. The substrate affinity is high ($K_{0.5}$=0.5 mmol/l) while the transport capacity is low [62]. SLGT1 is predominately expressed in the small intestine where it is involved in glucose and galactose absorption from food, but it is also involved in the reabsorption of glucose in the kidney. The gene *SLC5A1* codes for the protein and mutations in the gene give rise to intestinal glucose-galactose malabsorption (GGM, MIM:606824) resulting in watery diarrhea and dehydration. The expression in the kidney is located to the proximal tubule where it transports around 10% of the filtered glucose [63].

SGLT2

The major glucose transporter in the kidney is the SGLT2 protein that is located before SGLT1 in the proximal tubule where it is responsible for 90% of the re-absorption of glucose. SGLT2 is in contrast to SGLT1 a high capacity and low affinity ($K_{0.5}$= 2 mmol/l) glucose transporter that co-transports one glucose per sodium (Figure 1) [63]. Familial renal glycosuria (FRG, MIM:233100) arises as a consequence of mutations in the gene coding for SGLT2 (*SLC5A2*). The renal threshold for glucose decreases due to dysfunctional transport. This means that even at normal filtered glucose concentrations some of the glucose will not be reabsorbed but excreted in the urine. FRG is generally considered to be a benign condition

although increased frequency of urinary tract infections, activation of the renin-angiotensin aldosterone system and polyuria have been reported in mutation carriers [64-66].

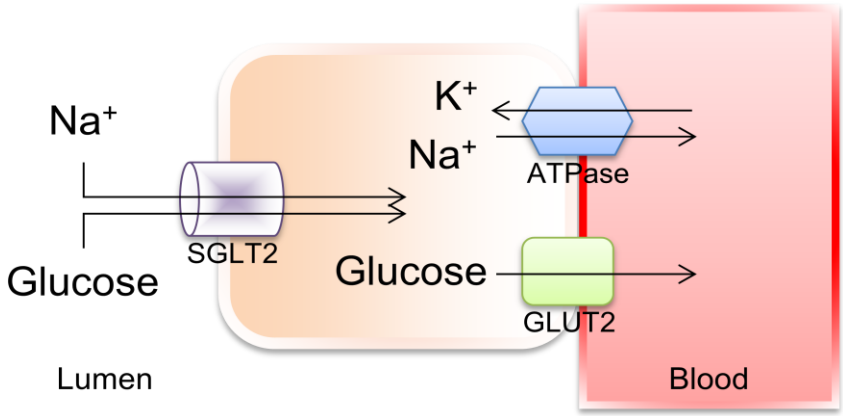


Figure 1. Glucose reabsorption in the proximal tubule by SGLT2
Glucose is co-transported with sodium (Na^+) against a concentration gradient via SGLT2 into the cells in the proximal tubule of the kidney. Sodium is pumped out of the cell into the blood via sodium-potassium (K^+) ATPase pump to maintain a downhill concentration gradient for sodium. The glucose then diffuse through the passive glucose transporter (GLUT2) into the blood.

SGLT2 inhibitors

SGLT2 inhibitors are, due to the expected low risk of side effects, been of interest for the treatment of diabetes and hyperglycemia. The drug-induced blocking of SGLT2 leads to decreased renal threshold for glucose and blood glucose levels as the glucose is excreted via the urine. Among the first studies on SLGT2 inhibition were experiments done in rats using phlorizin. They reported improvement of glucose tolerance by the inhibition of the renal glucose reabsorption by phlorizin, which prompted the introduction of SGLT2 inhibitors as means to lower glucose in patients with type 2 diabetes. Furthermore, it was shown that T2D patients had an increased SGLT2 expression and increased glucose uptake in the proximal tubule [67]. One of the benefits of SGLT2 inhibition for the treatment of hyperglycemia is that it is insulin independent, i.e. the insulin secreting capacity of the patient does not matter for the effect of the drugs. Today there are several SGLT2 inhibitors on the market for the treatment of T2D, for example Dapagliflozin, Canagliflozin, Empagliflozin and Ipragliflozin [68]. They have a beneficial effect on HbA1c levels, a low risk of causing hypoglycemia, some weight reducing effect due to the loss of energy through the urine and possibly a blood pressure lowering effect [68]. The side effects reported are generally mild and include genital mycotic infections and benign urinary tract

infections [69, 70]. Lately reports of increased rate of ketoacidosis in patients treated with SGLT2 inhibitors have emerged [71].

Human knock out models

As the SLGT2 inhibitors are relatively new on the market, studies on the potential long-term effects on prevention of diabetes are lacking. One possibility to address this question without testing the drug in a non-diabetic population is to study non-diabetic carriers with FRG for their propensity to develop diabetes or deterioration of glucose tolerance. As the carriers have been under "SGLT2 inhibition" since birth it is possible to follow them over a long period of time to investigate the effects of chronic glycosuria. This will give an indication of the long-term effects of the drug in the general population.

The function of islets of Langerhans in type 2 diabetes

As was mentioned before, the main clinical manifestations of diabetes are elevated blood glucose levels and impaired glucose metabolism. This occurs when the body is incapable of removing glucose from the circulation by uptake into tissues where it is supposed to be used as fuel. Insulin is the key hormone involved in the two factors contributing to hyperglycemia; insulin resistance and the incapability of the islet of Langerhans in the pancreas to secrete enough insulin [12]. However, the genetic evidence that tells us that it is the islets dysfunction that is the main driving force of T2D is quite compelling. In a review by Ahlqvist et al. aggregating the genetic information about T2D it was pointed out that the majority of associated loci modulate the risk of diabetes probably through an effect on insulin secretion [17]. The same was commented in the review by Grarup et al [72]. It is therefore not unfounded to turn to functional studies of the pancreatic islets to understand the molecular mechanisms behind the disease.

Physiology of human islets

A human pancreas consists of endocrine and exocrine tissue. The endocrine tissue is organized in clusters throughout the organ, which are called islets of Langerhans. The main task of the islets is to control glucose homeostasis. There are more than 3 million islets in a healthy pancreas making up to 4.5% of the total pancreatic volume [73]. The human islets are composed of 4 main endocrine cell-types [74]; alpha, beta, delta and PP cells. They produce specific hormones, insulin in the beta cell, glucagon

in the alpha cell, somatostatin in the delta cell and pancreatic polypeptide in the PP cell. The beta cells are the most abundant, making up slightly more than 50% of the islet cell population in humans, slightly more than 30% are alpha cells and around 10% are delta cells [73]. The islets are closely associated with blood vessels and they release their hormones directly into the blood stream in response to different stimuli. The hormones also have paracrine and autocrine effects. Insulin is secreted in response to elevated blood glucose levels and has an inhibitory effect on the secretion of glucagon. Glucagon is on the other hand secreted at low blood glucose levels to promote glucose production (gluconeogenesis) in the liver and the kidney and glycogen breakdown (glycogenolysis) [75]. Glucagon has a stimulatory paracrine effect on the beta cell. Furthermore, somatostatin is stimulated by increased glucose and it inhibits both insulin and glucagon release possibly to fine tune the response to nutrients of the pancreas [76]. The pancreatic polypeptide is released after a meal and it has been implicated in appetite control, having a negative effect on energy intake [77, 78].

Insulin secretion

The main function of the beta cell is to secrete insulin in response to elevated glucose and other nutrient concentrations after a meal. There are two main pathways in insulin secretion. The triggering pathway is relatively well studied and it is initiated with the metabolism of glucose in the beta cell. Glucose is taken up from the blood into the cells via GLUT2 by means of facilitated diffusion. The metabolism of glucose occurs mainly by glycolysis. Pyruvate is the end product of glycolysis and its metabolism in the mitochondria by the citric acid cycle increases the ATP:ADP ratio, which in turn closes the ATP sensitive potassium channels (K_{ATP} channels). This reduces the efflux of K^+, triggering depolarization of the plasma membrane and the opening of the voltage-dependent Ca^{2+} channels (VDCC) leading to elevated intracellular Ca^{2+}. This, in turn, triggers the exocytosis of insulin [79]. The amplifying pathway is less well characterized. It stimulates insulin secretion independently of VDCC, and putative signaling molecules generated by glucose metabolism include NADPH and ATP [110].

Gene regulation and expression

Transcription is controlled by the structure of the DNA and the ability of the RNA polymerase to bind and transcribe the gene. This is in turn controlled by transcription factors (TF), epigenetic modifications and chromatin structures [80]. The TFs include activators, repressors, enhancers and silencers that all affect the recruitment and binding of other TFs and the transcriptional machinery. Other molecules such as

glucose can also influence gene expression. Glucose has been shown to directly and indirectly control the expression of many genes, especially in the islets. The insulin gene is for example transcribed in response to glucose due to phosphorylation of pancreatic and duodenal homeobox 1 (PDX1) by Phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K) which leads to the translocation of PDX1 to the cell nucleus where it binds the insulin gene promoter and initiates transcription [81]. Another transcription factor, carbohydrate response element binding protein (ChREBP) is also important in glucose regulated gene expression [82].

By using global gene expression profiling in glucose treated rat islets it was shown that thousands of genes are dynamically altered in response to glucose [83] and many genes are altered by chronic hyperglycemia (diabetes) although it is not clear if these changes are the cause or the consequence of the hyperglycemia [84].

Hyperglycemia – the cause and the consequence of T2D

In patients with T2D there is a progressive decline in islet function, a condition often ascribed to the consequences of elevated glucose, so called glucotoxicity which can stimulate apoptosis and lead to reduced beta cell mass [85-88]. Prolonged exposure to hyperglycemia can also induce endoplasmic reticulum (ER) stress and production of reactive oxygen species (ROS) [89], which can impair islet function and the ability of the islets to adjust the insulin secretion to meet the increased demands imposed by insulin resistance and obesity [90]. In this setting the islet dysfunction, which together with insulin resistance will cause progression to type 2 diabetes, is a consequence of the hyperglycemia. And in addition to glucose induced islet dysfunction other factors such as genetic variants can reduce the islets function and are then the cause of the hyperglycemia through decreased ability of the islets to secret insulin in response to glucose.

Glucotoxicity

The detrimental effect of prolonged hyperglycemia on the beta cell can be ascribed to altered gene expression and increased beta cell apoptosis leading to impaired insulin secretion [91, 92]. The mechanisms behind this have not been fully elucidated but it is clear that different cellular stresses, such as oxidative and ER stress, hypoxia and inflammatory pathways, are involved [91].

Oxidative stress markers have been found to be increased in islets from donors with T2D compared to non-diabetic controls and the markers also had a negative correlation with glucose-stimulated insulin secretion (GSIS) and insulin gene

expression [93, 94]. Incubation with anti-oxidants partly reversed the dysfunction and increased *INS* expression. Oxidative markers were also shown to be induced by glucose, which could be prevented by gliclazide, a sulphonylurea drug with anti-oxidant properties [95]. Taken together these observations provide evidence for the induction of oxidative stress by glucose and the negative consequences of oxidative stress on beta cell function.

Aim of this thesis

The overall aim of this thesis was to explore the role of genetic factors, gene expression and the effect of genetic variation on gene expression in the development of T2D. The latest development in sequencing, NGS, was used to find causes and consequences of hyperglycemia and glycosuria.

Specific aims of the studies

Study I	The aim of study I was to define the genetic cause for FRG in a family with clustering of T2D and to evaluate the effects of chronic glycosuria on changes in glucose tolerance over time.
Study II	The aim of study II was to fine-map a region previously linked to T2D on chromosome 9 [96] and to identify genetic variants that could explain the linkage.
Study III	The aim of study III was to find genes whose expression was associated with glycemic levels in human islets and to identify genetic variation that would influence gene expression. Furthermore, we aimed to explore allele specific expression and splice patterns in the islets.
Study IV	The aim of study IV was to build on the work in study III and to try to unravel causes and consequences of hyperglycemia. We did this by comparing the genes regulated by short and long-term glucose stimulation in islets from donors with normal and abnormal glucose tolerance.

Methodology

Human pancreatic islets

The pancreas and the islet of Langerhans are obvious target tissues for T2D research. Dysfunctional insulin secretion is a vital part of the pathogenesis of diabetes [12] and the study of isolated human islets is important to understand disease mechanisms. All projects in this thesis have used data from human pancreatic islets. The Nordic Network for Clinical Islet Transplantation in Uppsala, Sweden provided the islets via the Human Tissue Laboratory, Lund University Diabetes Center, Malmö, Sweden. All islet donors had given consent to donate organs for medical research and all procedures are approved by the ethical committees at Uppsala and Lund University. The isolation of the islets was done as previously described using a digestion-filtration method [97]. For study IV we incubated the islets in normal (5.5 mmol/l) and high (18.9 mmol/l) glucose for 24 hours to capture the physiological gene expression changes in islets (Figure 2).

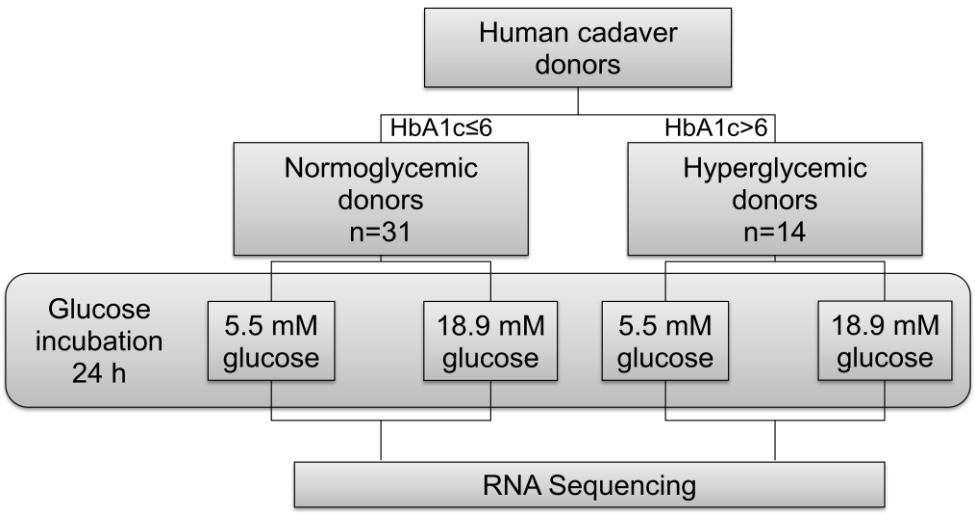


Figure 2. Glucose treatment of human pancreatic islets in study IV.
Islets from one donor were divided into two pools that were incubated in parallel with normal (5.5 mmol/l) or high (18.9 mmol/l) glucose for 24 hours before RNA was extracted.

Next generation sequencing

The advances in sequencing technology over the past few years have been essential for the work in this thesis [32]. The methodology here is based on the Illumina technology that we used for the capture of global and local genetic variation and/or gene expression in all four studies.

Sample preparation

The preparation of RNA and DNA samples for sequencing using the Illumina technology follows the same basic principles. After RNA has been extracted and mRNA has been selected using poly A-capture it is converted to complementary DNA (cDNA). The DNA/cDNA is then fragmented to a suitable size for sequencing, often around 300 bp for paired-end sequencing. After end repair of the fragments and 3' adenylation, adaptors are ligated to both ends of the fragments. The adaptors contain sequences necessary for the amplification and hybridization to the flow cell and primer binding sites used in the sequencing reaction. The adaptors also contain unique sequences called barcodes or indexes that are specific to each sample to allow for multiplexing, i.e. pooling of samples prior to the sequencing. In the case of DNA sequencing there can be a selection step prior to the sequencing that targets regions of interest.

Target selection

We used whole exome sequencing in study I, II and III to capture variation in the coding region of the genome. A common use for DNA sequencing is targeted sequencing of interesting regions, for example all the known exons in the genome allowing sequencing of more samples at a lower cost. The selection is performed using complementary sequences, probes, to the regions of interest and the probe-sample hybridization is pulled down using streptavidin-coated magnetic beads. The fragments containing enough complementary sequence to the capture sequence are selected and the beads are collected with a magnet while the rest of the un-hybridized fragments are washed away.

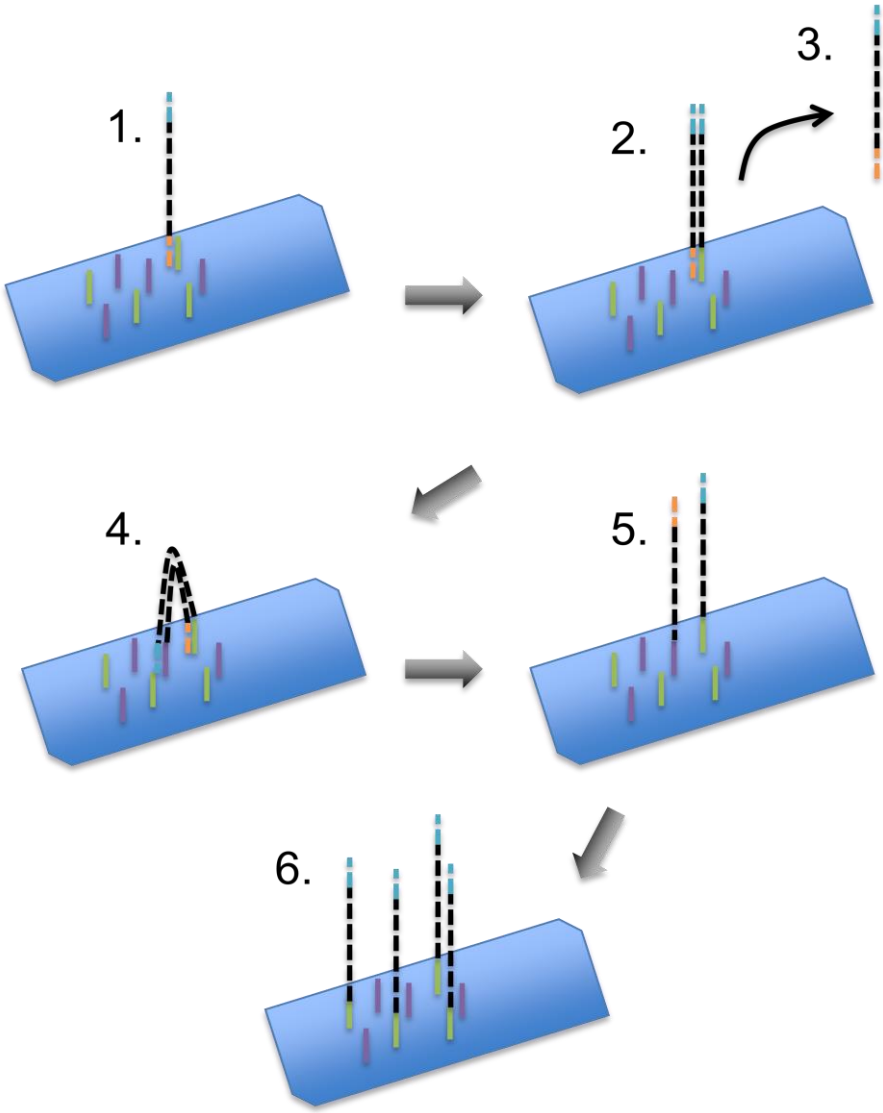


Figure 3. Cluster generation on an Illumina flow cell
1. The fragment binds to the grafting primer (green) attached on the surface of the flow cell. 2. The primer is extended to form a covalently bound complementary strand and the 3. hybridized fragment is denatured and washed away. 4. The 3' end of the attached fragment hybridzes to another grafting primer (purple) to form a bridge for amplification. 5. This results in two complementary strands covalently bound to the surface of the flow cell. Steps 4 and 5 are repeated until a cluster of complementary strands originating from the first fragment is formed. 6. The complement strand is cleaved off and washed away. This is simultaneously done for millions of fragments over the surface of the flow cell resulting in millions of unique clusters for sequencing.

Cluster generation

The clustering generation step was done using a cBot (Illumina Inc., CA, USA). The sequence libraries are immobilized on a solid glass surface in a so-called flow cell where the individual fragments are amplified into clusters that can be detected by the machine. On the surface of the flow cell millions of grafting primers are attached that will bind the adaptors ligated to the fragments (Figure 3 step 1). The amount of DNA added to the flow cell will determine the number of clusters created. Once bound there is an extension of the hybridized graft primer that results in a covalently bound sequence molecule complementary to the fragment (Figure 3 step 2 and 3). The next step is bridge amplification where the fragment bends over to hybridize to a close grafting primer forming a bridge, the new primer is extended and a copy of the first sequence is made (Figure 3 step 4 and 5). This step is repeated and in the end the complement strand is cleaved off and washed away until a cluster of identical sequences is bound to the flow cell surface (Figure 3 step 6). The cluster generation is necessary to produce a strong enough signal for the sequencer to detect.

Sequencing by Synthesis

The sequencing in this thesis was performed on a HiSeq 2000 (Illumina Inc., CA, USA) that employs the sequencing by synthesis method. The fragments that are bound to the flow cell are hybridized with a sequencing primer that is complementary to the part of the adaptor closest to the original fragment. This primer is extended by a polymerase that incorporates fluorescently labeled reversible terminator-bound dNTPs one nucleotide at a time. The dNTP have specific fluorescent signals associated with them depending on if they contain A, C, G or T. The signal from millions of clusters is detected in parallel. After the signal has been detected the reversible terminator is removed and a new nucleotide corresponding to the next base in the sequence can be incorporated (Figure 4). The cycle of incorporation, detection and de-blocking is repeated until the desired read length is obtained. We used 100 bp reads in this thesis.

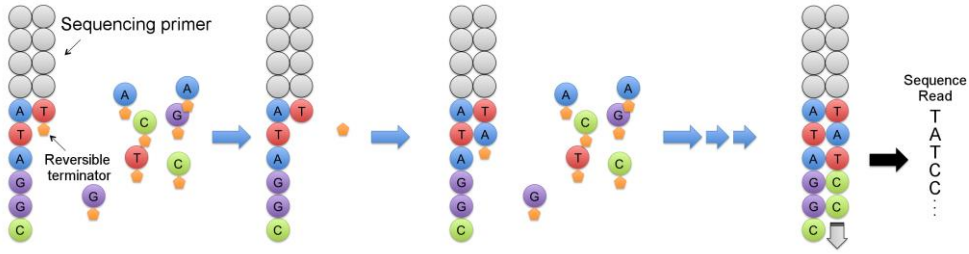


Figure 4. Sequencing by synthesis
The reaction starts with the binding of the sequencing primer to the adaptor sequence (grey) and the incorporation of the complementary reversible terminator-bound base and the corresponding colour is detected. The reversible terminator is the removed and the next base can be incorporated and detected. These steps are repeated until the desired read lenght has been obtained and the sequence of colour is then translated into the sequence of bases.

Data analysis

In order to analyze the sequencing data the fluorescence signal needs to be converted to base calls. We used CASAVA (Illumina) to compile the sequence and quality information of the fragments into FASTQ-files [111] that be used as input for downstream analysis.

DNA sequence analysis

The next step in the analysis pipeline is to align the sequence fragments either to a reference genome. The Burrows-Wheeler Aligner (BWA) [112] was used for alignment of DNA samples in this thesis. The resulting alignment is stored in the binary BAM-file format, which contains both the raw sequence and the alignment coordinates [113]. Genome Analysis Tool Kit (GATK) [114] was used to compare aligned reads to the human reference genome and produce high quality variant calls. First, duplicate reads that most likely are a technical artifact from amplification were marked and discarded. Secondly, the sequence fragments were realigned around known indels to reduce the number of miss matches to the reference that can occur at these spots. Thirdly, a step of base quality score recalibration was performed as this accounts for a number of variables that affect the quality of a base call and report more accurate quality scores. Lastly GATK was used for the actual variant calling in two steps where the variants were first identified against a reference genome and then were filtered according to the likelihood that the variant was true or not, based on sequencing depth and base call quality. The variants were stored in the variant call format (VCF) that contains customizable meta-information lines and each following line contains information about a genomic position. For more information about the file formats visit https://github.com/samtools/hts-specs.

RNA sequence analysis

The analysis of RNA sequence data is different from the DNA counterpart and computationally much harder, due to splicing. The sequence reads are mapped to the genome and the alignment benefit greatly from taking gene annotations into account. We used the Spliced Transcripts to a Reference (STAR) aligner [115] in study IV, mapping the RNA sequence reads against the hg19 genome using GENCODE [116] gene annotations to guide the alignment. FeatureCounts [117] was then used to count the number of reads mapped to genes or exons, providing quantification. Before statistical inference the counts were normalized to account for variables such as gene length and sequencing depth to make expression values comparable across samples. This since longer genes produce more reads, and the deeper a transcriptome is sequenced, the more reads will map to each transcript, which creates biases [32]. We normalized our data by expressing it as counts per million (CPM) mapped reads by dividing the reads mapped to a transcript by the reads mapped in total and multiplying that by 10^6. The mean/variance trend was then identified using voom [118] to correct for high dispersion at low expression levels and allow for correct linear modeling. The effect of different reagent batches was corrected for with ComBat [119]. The data was then used to detect differential expression between sample groups by fitting a linear model using limma [120].

Quantitative real time PCR

The expression of *SLC5A2* and *MPDZ* gene expression was measured using quantitative real time PCR (qPCR) in study I and II. We used two methods commonly employed; the probe based TaqMan™ gene expression assay (Thermo Fisher Scientific, MA, USA) and the double stranded DNA binding dye SYBR® Green (Thermo Fisher Scientific, MA, USA).

Taqman gene expression

The *MPDZ* expression was measured using an off-the-shelf TaqMan gene expression assay (assay ID: Rn01455039_m1). The expression is detected using primers that bind over exon boundaries in the target transcript and a specific probe binding to the PCR product. This design allows for a more specific detection of target mRNA over DNA. The probe has a fluorescent molecule in one end and a quencher in the other that suppresses the signal from the fluorescent molecule. The probe binds the cDNA and when the polymerase amplifies the cDNA it cleaves off the molecules at the ends, which removes the quencher from the fluorescent molecule allowing the detection of the signal.

SYBR Green

To detect *SLC5A2* expression we used the SYBR Green PCR Mastermix (Thermo Fisher Scientific, MA, USA). The SYBR Green dye gives off a fluorescent signal when it binds to double stranded DNA. As the target cDNA is amplified with transcript specific primers the SYBR Green binds and the signal increases until it can be detected and expression can be calculated.

Glucose uptake in SGLT2-transfected cell lines

In study I we wanted to investigate the effect of a missense mutation in SLGT2 on glucose transport. We did this by synthesizing the wild type (wt) sequence of SGLT2 cDNA and inducing the mutation (mutant) using site-directed mutagenesis (GenScript USA Inc., NJ, USA). The sequence was then inserted into the pCMV6-Neo vector (OriGene Technologies, Inc., MD, USA) that contains neomycin resistance, which allows for selection using neomycin that normally is toxic to the cells.

Transfection and generation of stable cell lines

The wt and mutant containing vectors were transfected into a HEK293 cell line since they do not express the natural SGLT2 protein [98, 99]. We used a cationic lipid, Lipofectamine® LTX (Thermo Fisher Scientific, MA, USA) as transfection reagent. The positively charged cation part of the molecule interacts with the DNA and forms lipid/DNA complexes that are thought to be delivered into the cells by endocytosis [100]. The transfected cells were selected for using 0.8 mg/ml G418 (Merk Millipore, MA, USA), an antibiotic often used to select for neomycin resistant cells, for 14 days. The level of *SLC5A2* expression was measured using qPCR for use in the glucose uptake experiments.

Glucose uptake cell based assay

To measure the impact of the mutation on the glucose transport capability of the SGLT2 protein we measured the glucose uptake of the transfected HEK293 cells with the Glucose Uptake Cell based Assay Kit (Cayman Chemical, MI, USA). The cells were incubated in assay buffer for 30 minutes containing 150 µg/ml of fluorescent glucose analogue 2-deoxy-2[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-D-Glucose (2-NBDG) that is taken up by the cells where it cannot be metabolized by glycolysis [101]. This allows for the accumulation of the fluorescent molecule in the cell that can be detected by conventional spectrofluorometry. The relative uptake of the mutant SGLT2 transfected cells versus the wt SGLT2 transfected cells were calculated by dividing the uptake in mutant cells with the uptake in wt cells normalized against

the *SLC5A2* expression. To control for unspecific glucose uptake untransfected cells were treated with the glucose uptake inhibitor Apigenin before measurement of the 2-NBDG uptake.

Results

Study I

Influence of Familial Renal Glycosuria Due to Mutations in the *SLC5A2* Gene on Changes in Glucose Tolerance Over Time

In study I we aimed to investigate the effect of chronic glycosuria on changes in glucose tolerance over time.

SGLT2 inhibitors are used to treat diabetes but the potential effect of glycosuria on the prevention of glucose intolerance has not been investigated. By following a family with familial renal glycosuria and diabetes over several years, with a mean follow up time of 10.5 years, we could analyze the progression of glucose intolerance in glycosuric individuals.

Familial renal glycosuria is most often caused by mutations in the *SLC5A2* gene encoding the SGLT2 glucose transporter [102]. We first sequenced the exome of 23 family members and found a 6 bp deletion (c.300-303+2del) in the *SLC5A2* gene that truncates the protein by introducing a premature stop codon. The mutation was found in 11 family members who all excreted glucose in the urine (Figure 5). Furthermore, we detected a second missense mutation (p.A343V) in family member 45 who had the most severe form of glycosuria. The missense mutation was passed down to her children and grand children, who all but one also had mild forms of glycosuria.

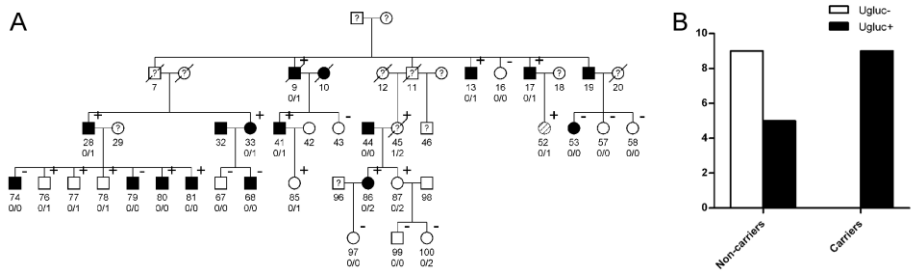


Figure 5. A 6 bp mutation in the *SLC5A2* gene was associated with glycosuria in the pedigree.
A) The pedigree of the family with glycosuria. Black symbols indicate that the person has IFG, IGT or T2D (squares = men, circles = women), white symbols indicate normal glucose tolerance. Diagonally hatched symbol indicates T1D (not included in the analysis). Plus and minus indicates positive or negative test for glycosuria. First number under symbols indicates ID number. Genotyping of the deletion (c300-303 +2del) showed that 10 of 22 genotyped family members are heterozygous for the mutation. The carriers of the deletion (c300-303+2del) are denoted with 0/1 and carriers of the reference allele are denoted with 0/0 while carriers of the missense mutation (p.A343V) are denoted with 0/2. B) glycosuria in carriers and non-carriers of the mutation deletion (c300-303+2del). All but one of the non-carriers displaying glycosuria had impaired fasting glucose, impaired glucose tolerance or type 2 diabetes,

Next we followed the glucose tolerance over time in the mutation carriers and non-carriers to see if the glycosuria would balance out potential shifts in glucose tolerance and thus be protective against glucose intolerance and T2D. We used the area under the curve during oral glucose tolerance test (AUC_{OGTT}) to assess the glucose tolerance at the first and last follow-up visits (mean follow-up time was 10.5 years, range 3-22 years). The difference in AUC_{OGTT} between the first and last visit was compared between the carriers and non-carriers of the *SLC5A2* deletion and we found no difference between the groups (Figure 6, p=0.67). We therefore believe that chronic glycosuria did does not affect the glucose tolerance over time in this family.

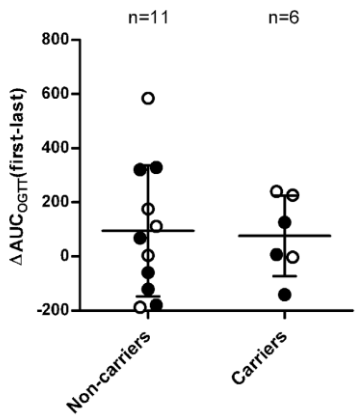


Figure 6. Changes in glucose tolerance over time in mutation carriers.
Changes in area under the glucose curve from the first to the last visit was not different in carriers of the 6 bp deletion compared to non-carriers (p=0.64). Filled dots represent family members with T2D and empty dots represent non-diabetic family members.

Study II

Contribution of variants in the *MPDZ* gene on familial clustering of abnormal glucose tolerance

In 2002 Lindgren at al [96] published a suggestive linkage to T2D on chromosome 9 in 58 families from western Finland. Due to limitations in technology no causal gene could be found at the time. We have now used NGS to fine-map the region in an effort to find the causal gene.

The initial linkage was reanalyzed in four families using Illumina's linkage panel IVb. This redefined the peak from 12.8 to 16.9 Mb on chromosome 9 revealed that only one family contributed to the peak (Figure 7).

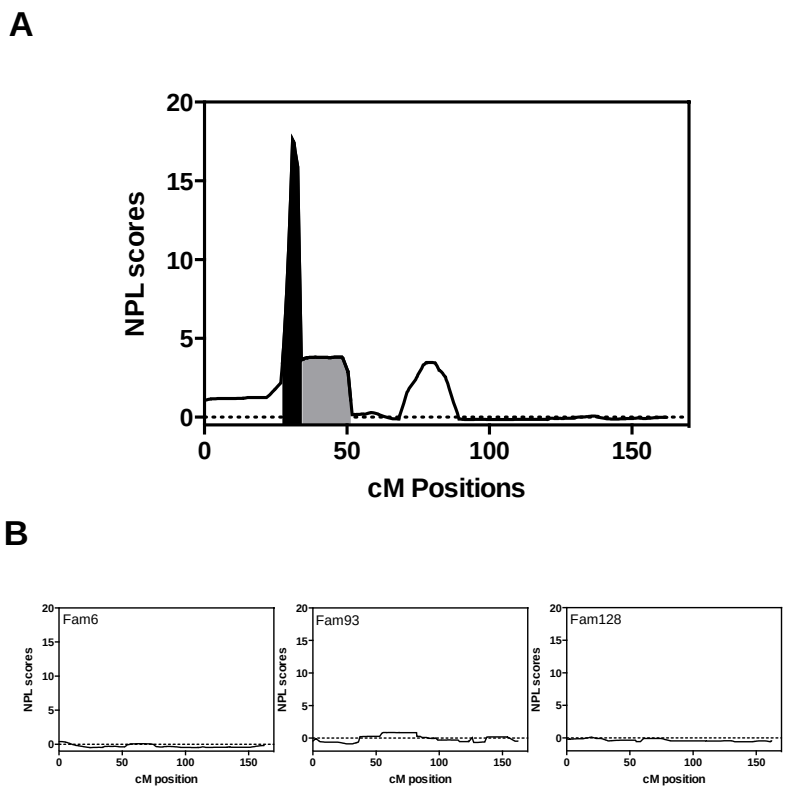


Figure 7. Linkage on chromosome 9 in four families using Illumina Linkage Panel IVb genotype data.
A) The NPL score for family 225 over chromosome 9. The black area was covered by targeted sequencing and ranges from 13 to 16 Mb. The grey and black area was covered with exome sequencing and ranges from 16 to 27 Mb. B) The NPL score for the four additional families included in the GWS with Illuminas Linkage Panel IVb.

We therefore defined the haplotype structure in this family using dense genotyping arrays, showing a 8.3 Mb large haplotype from 7.0 to 15.3 Mb shared by seven of nine affected family members used in the linkage analysis and by none of the unaffected individuals.

To fine-map the linked region we used targeted and whole exome sequencing in 10 affected and 9 unaffected family members. We then filtered for novel variants, primarily found in T2D affected individuals, located within the boundary of genes. Selected variants where subsequently genotyped in additional family members and in 343 samples from the same geographical region as the original family. Two intronic and one synonymous variant in the multiple PDZ domain protein (*MPDZ*) within the disease-associated haplotype were associated with T2D in the family (chr9_13164682, chr9_13241460 and rs144481155, p=0.02, Figure 8) and in the replication material (p=0.02). All carriers of the haplotype except one were diagnosed with T2D. As would be expected from family specific variants the haplotype frequency was much lower in the replication material, 1.2%, compared to 20.5% in the original family. We also genotyped the variants in a large Finnish population consisting of 6738 samples from the Botnia PPP study material [103] and DIREVA (Diabetes registry in Vasa) finding only 8 individuals carrying the variants (0.1%).

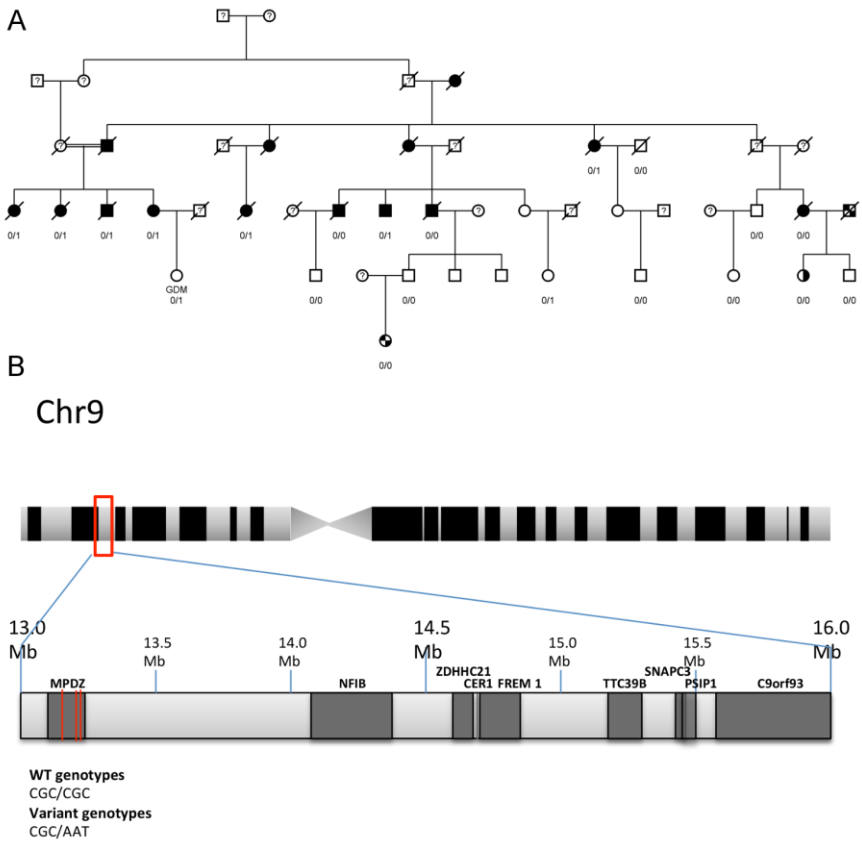


Figure 8. Three variants in *MPDZ* were associated with T2D in the original familiy.
A) The pedigree of family 225. Square symbols represent males and circles females. Black symbols are T2D affected individuals, chequered symbols are individuals with IGT and half-filled symbols are T1D affected individuals. GDM – Gestational Diabetes Mellitus. Genotypes for the three variants are denoted as numbers under the symbols for genotyped individuals – 0/0=CGC/CGC, 0/1=CGC/AAT. B)The genes under the linkage peak. The area of 13-16 Mb on chromosome 9 corresponds to the linkage peak. The dark boxes represents the genes in this area. *MPDZ* were found to harbour three varians (red lines) part of a haplotype associated with T2D (p=0.02). The WT genotypes are the reference genotypes of the variants in the haplotype while the variant genotypes are the genotypes of the T2D associated variants in the haplotype.

The *MPDZ* gene has not been implicated in T2D before so the mechanism by which this gene might confer risk to disease is unknown. We found that the gene was expressed in human pancreatic islets and was down regulated in islets from T2D donors (relative expression=0.93, FDR corrected p-value=0.002, n=31) and one possibility was thus that the expression of the gene would influence beta cell function. We therefore silenced *MPDZ* in an insulin secreting cell line by siRNA-mediated knockdown. An 80% reduction in *MPDZ* expression resulted in a significant 24% reduction in glucose-stimulated insulin secretion (Figure 9, p=0.007).

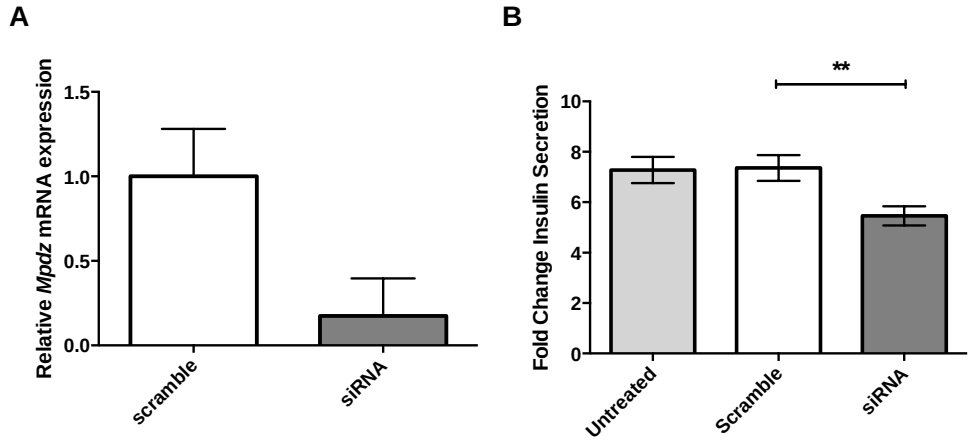


Figure 9. Silencing of MPDZ in a clonal rat beta cell line.
A) The expression of MPDZ was reduced after siRNA mediated silencing by 80% compared to scramble sequence transfected cells. B) The glucose -timulated insulin secretion was reduced by 24% in the MPDZ silenced cells (dark grey bar) compared to the scramble sequence transfected cells (white bar). There was no effect of scramble sequence transfection on GSIS compared to control cells (light grey bar). n=16 in all groups. **p<0.01. The data is represented as the mean and the bars show the SEM.

Study III

Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism

The islet transcriptome is perturbed in T2D and this might be regulated by both genomic variation and environmental factors. By combining RNA sequencing with exome sequencing and genotyping of pancreatic islets from 89 human donors we provide a global analysis of the islet gene expression and regulation by genomic variation in health and disease.

The pancreatic islets expressed 92% of known genes in RefSeq [104] although most are expressed at low levels. Islet specific genes, such as insulin and glucagon were among the most highly expressed genes (Figure 10). We compared the expression of all genes in islets donors with normoglycemia (NGT; HbA1c < 6%, n=51), impaired glucose tolerance (IGT; 6% ≤ HbA1c < 6.5%, n=15) and type 2 diabetes (T2D; HbA1c ≥ 6.5%, n=12) using a linear regression model adjusted for age and sex. We found 1619 genes that were associated with glycemic status, 70 of which were also associated with *in vitro* insulin secretion from the islets, positioning them as putatively involved in islet dysfunction. As an example RAS guanyl releasing protein 1 (*RASGRP1*) was associated with lower glycemic status and increased insulin secretion,

which is in line with previous studies suggesting a role for *RASGRP1* in the regulation of insulin secretion [105].

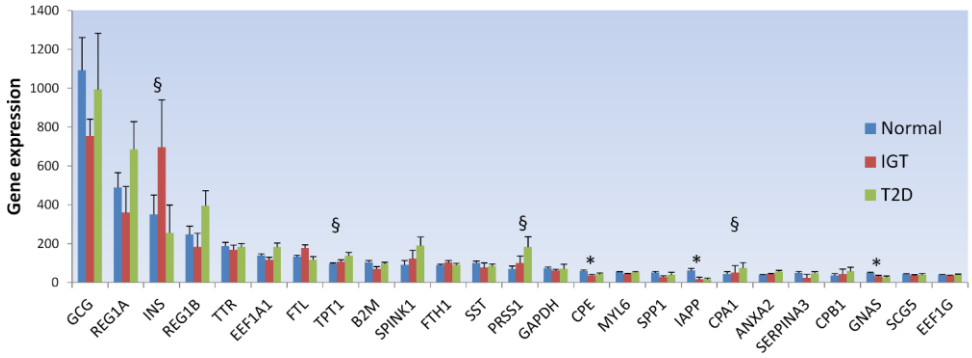


Figure 10. Genes expressed in human pancreatic islets stratified by glucose tolerance status.
RNA-seq normalized median expression of the top 25 nonribosomal genes expressed in islets. Genes are ordered by decreasing median expression in all 89 islet donors. Normal represents normoglycemic donors (HbA1c < 6%; n = 51), IGT represents impaired glucose-tolerant donors (6% ≤ HbA1c < 6.5%; n = 15), and T2D represents diabetic donors (HbA1c ≥ 6.5%; n = 12). Error bars represent SEM values. *Genes that show significant association between expression and glucose tolerance status detected by expression arrays and RNA-seq with both nominal and per- mutation P values < 0.05 (after performing 10,000 permutations). §Additional genes that show significant expression association with glucose tolerance status detected only with RNA-seq (at permuted P value < 0.05).

Many of the SNPs found to be associated with T2D in GWAS are located in intergenic and non-coding regions so the effect on T2D risk might be conferred through regulation of gene expression of nearby genes (cis eQTL). To investigate the regulation of gene expression by genetic variants in the islets we combined DNA sequence data and imputed genotype data with RNA sequencing expression data. At a genome-wide significance level 616 cis eQTLs were detected (Figure 11). Of these, 35 were eQTLs for genes associated with glycemic levels such as tetraspanin 33 (*TSPAN33*), 5'-nucleotidase, ecto (*NT5E*), transmembrane emp24 protein transport domain containing 6 (*TMED6*) and p21 protein activated kinase 7 (*PAK7*). Silencing of these genes in an insulin secreting rat beta cell line resulted in a significant reduction of GSIS (Figure 12). One of the advantages of RNA sequence data is the possibility to detect splicing events [32], which allowed us to find 371 genetic variants affecting the splicing of genes (sQTL) in the islets.

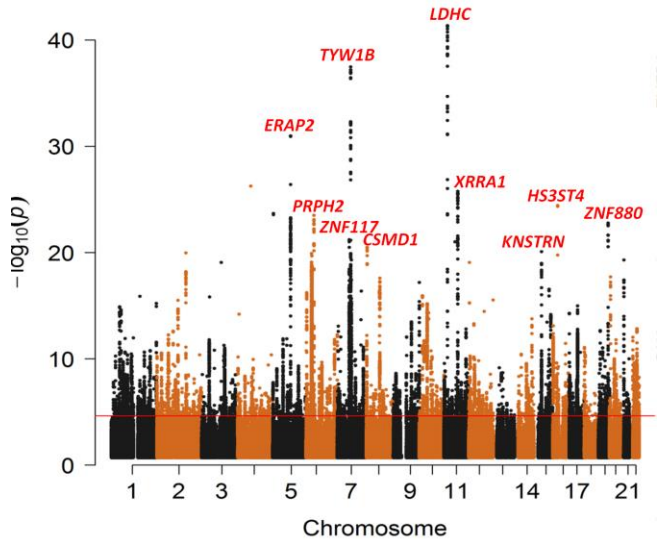


Figure 11. Landscape of cis-eQTLs (≤250 kb) in 89 human pancreatic islets
Manhattan plot of the best P value per SNP, showing the top 10 eQTL genes (FDR < 1% line drawn in black at P value = 2.267e-05).

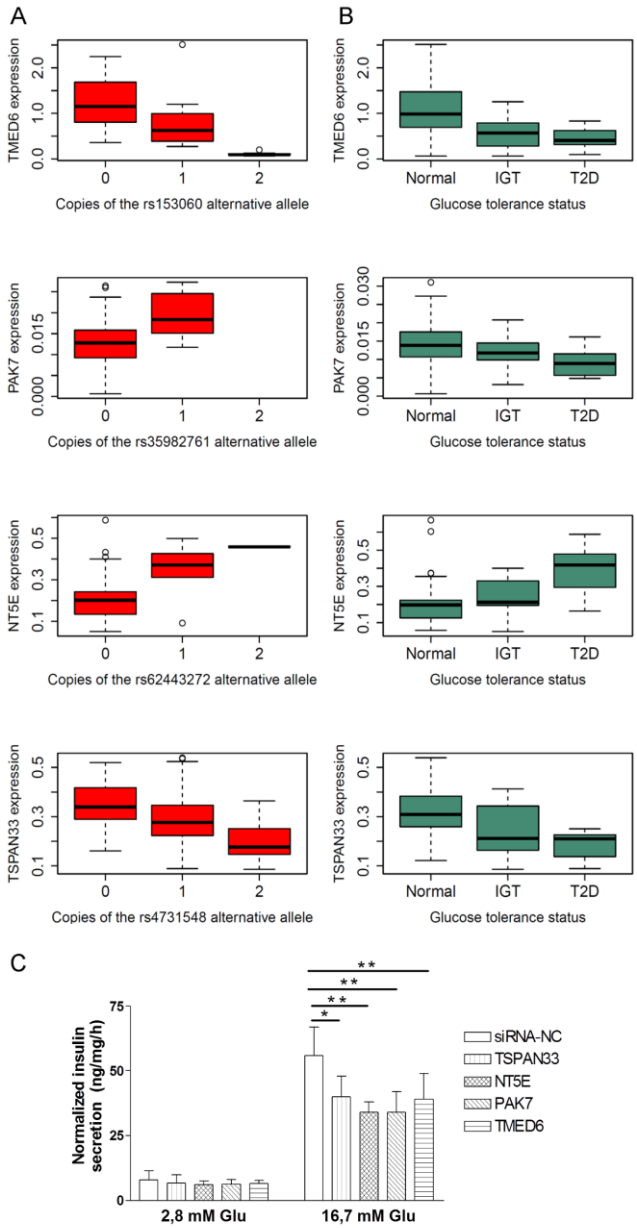


Figure 12. eQTL genes validated to interfere with glucose metabolism and insulin secretion.
A) eQTL SNP genotype versus eQTL gene expression (n = 89) of *TMED6*, *PAK7*, *NT5E* and *TSPAN33* in human islets. B) eQTL gene expression stratified by glucosetolerance status [normal, HbA1c < 6% (n = 51); IGT, 6% ≤ HbA1c < 6.5% (n = 15); T2D, HbA1c ≥ 6.5% (n = 12)]. C) Insulin secretion in response to 2.8 and 16.7 mM glucose 72 h after siRNA transfection, as measured during 1 h static incubation. Data are shown from three independent experiments for each siRNA. Data are normalized for protein content. Bars represent mean ± SD. **P < 0.01 and *P < 0.05 versus control siRNA.

Study IV

Glucose-induced Changes in Gene Expression in Human Pancreatic Islets – Causes or Consequences of Chronic Hyperglycemia.

In study III we identified gene expression changes associated with chronic hyperglycemia (IGT and T2D) but we cannot separate the changes that cause hyperglycemia from those that are a consequence of hyperglycemia. We therefore set up study IV with the assumption that genes changed as a consequence of glucose would be regulated by acute glucose exposure. By filtering these genes out from those changed in chronic hyperglycemia we are presumably left with genes contributing to the cause of hyperglycemia.

We started by comparing the gene expression, measured by RNA sequencing, between islets from human cadaver donors who had normoglycemia (NGT; HbA1c < 6%, n=81), impaired glucose tolerance (IGT; 6% ≤ HbA1c < 6.5%, n=19) and T2D (HbA1c ≤ 6%, n=16), and found 717 differentially expressed genes (Figure 13).

Then we incubated islets from 31 normoglycemic and 14 hyperglycemic human cadaver donors for 24 hours in normal (5.5 mmol/l) or high (18.9 mmol/l) glucose to find genes regulated by acute glucose exposure. A large number of genes were affected by acute glucose; 4658 genes were either induced or suppressed by high glucose for 24 hours (Figure 13).

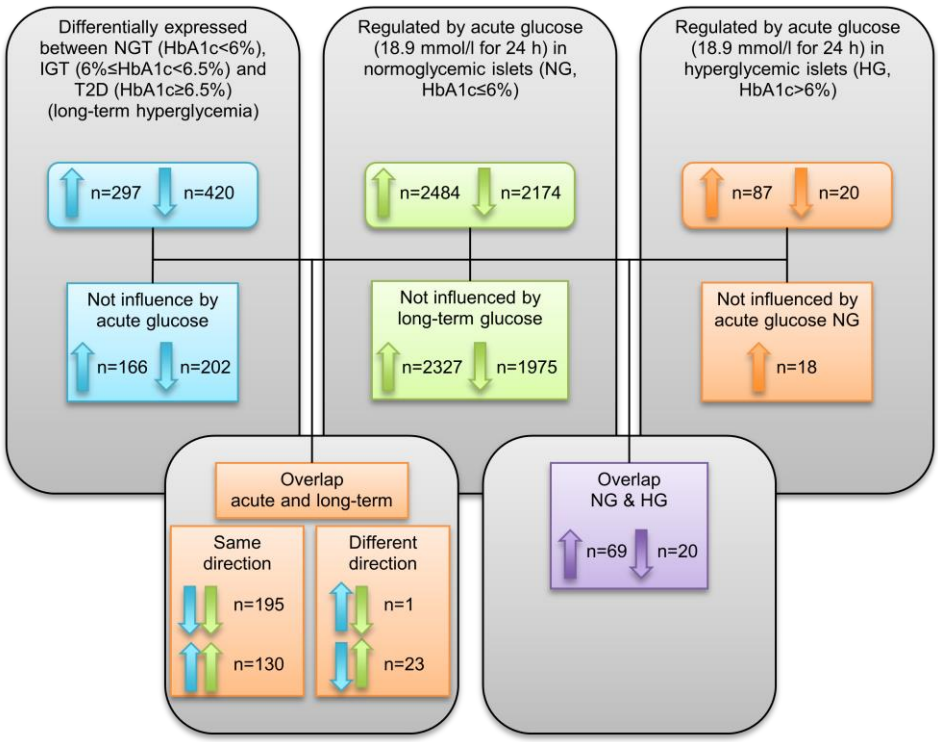


Figure 13. Genes changed by long-term or acute glucose exposure.
The number of genes whose expression differed between islets from donors with normal (n=81) and abnormal glucose tolerance (n=19 IGT and n=16 T2D) at FDR < 5%, and the number of genes whose expression changed after exposure to high glucose for 24 hours in 31 islets from donors with normal (normoglycemic, NG) and 14 donors with abnormal glucose tolerance (hyperglycemic, HG) at FDR < 5%.

We next compared the changes induced by long-term and acute glucose. Those genes that were differentially expressed in NGT, IGT and T2D and not changed by acute glucose in the same direction were considered to be potentially causal of hyperglycemia, while the genes changed by both long-term and acute glucose were considered as potential consequences of hyperglycemia (glucotoxic effects). To investigate the possible effects of the gene expression changes on islet function we correlated gene expression with *in vitro* insulin secretion as measured by stimulatory index (SI). In total 92 of the potentially causal genes were correlated with SI. Five of them were also found to harbor loss-of-function variants nominally associated with *in vivo* insulin secretion in a large Finnish cohort (n=3720, The Botnia PPP Study [103]); *TMEM132C*, *ERO1LB*, *DOCK10*, *PRR14L* and *IGSF11* (Figure 14).

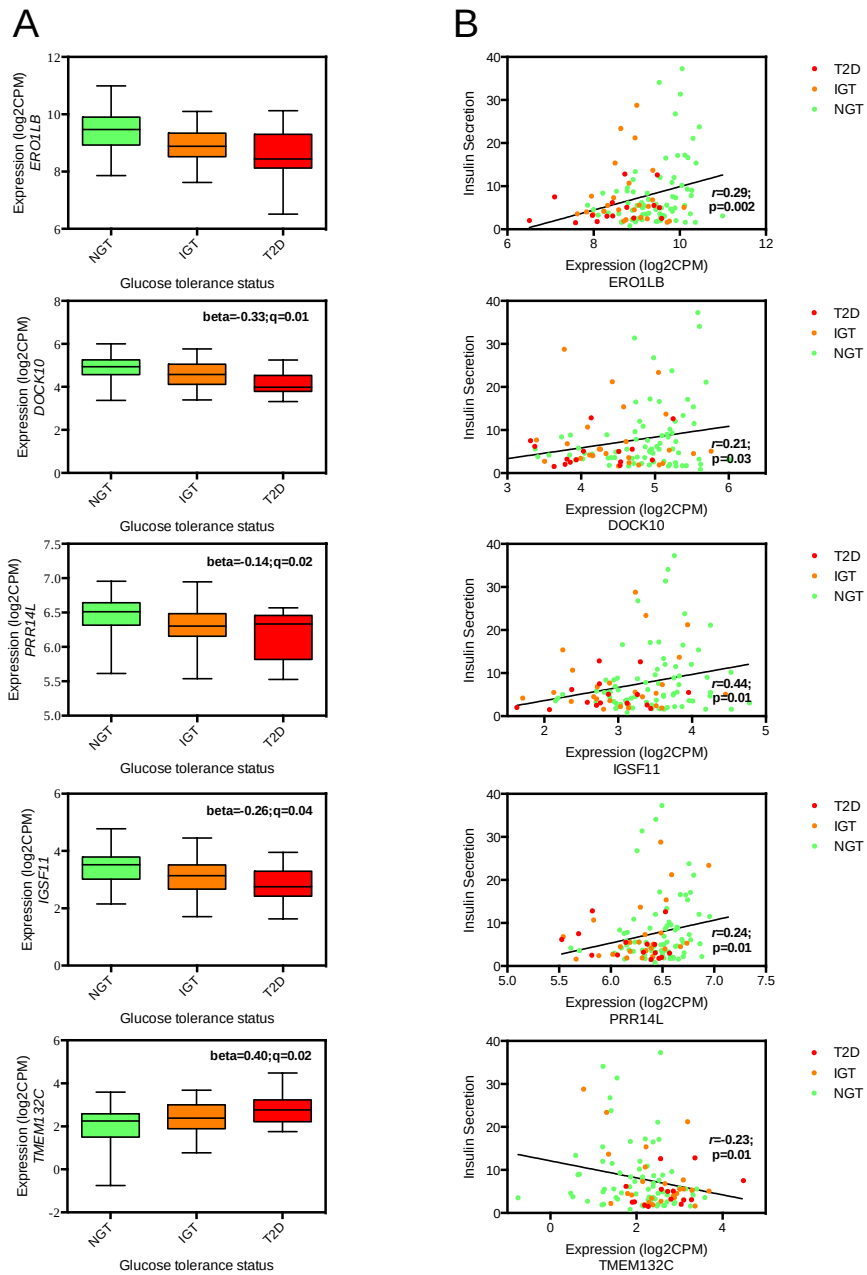


Figure 14. Genes whose expression correlated with insulin secretion that were differentially expressed in NGT, IGT and T2D islets but that were not changed, or changed in the opposite direction, by acute high glucose.
A) Expression of the five genes that also had potential loss of function variants associated with in vivo insulin secretion in NGT, IGT and T2D and B) correlation of the gene expression with in vivo insulin secretion (SI). Four of the genes (*ERO1LB*, *DOCK10*, *PRR14L* and *IGSF11*) were down-regulated in T2D and *TMEM132C* were up-regulated. The bars show the min and max values.

Among the 325 genes that we considered as potentially changed as a consequence of hyperglycemia 73 genes also had an expression that correlated with *in vitro* insulin secretion. The 5 genes (*SIPA1L2*, *HRK*, *TMED132D*, *MBP* and *CPEB1*) that also responded to acute glucose in hyperglycemic islets could be the most sensitive to glucose, and expression of *TMED132D* and *MBP* were correlated with *in vitro* insulin secretion (Figure 15). These two genes might be involved in the islet dysfunction seen after prolonged and repeated periods of hyperglycemia

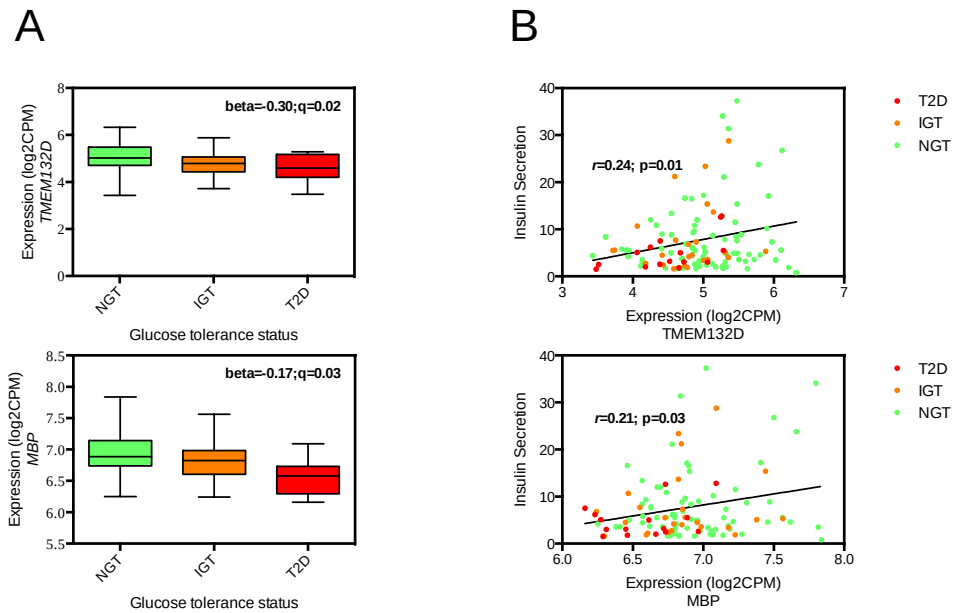


Figure 15. Potentially glucotoxic genes that were differentially expressed in AGT islets, and were changed by acute high glucose whose expression correlated with insulin secretion.
A) Expression of the two genes (*TMEM132D* and *MBP*) in NGT, IGT and T2D that also were changed by acute high glucose in islets from hyperglycemic donors (n=14) and correlated with *in vitro* insulin secretion that might represent genes that are the most sensitive to the toxic effects of glucose. The bars show the min and max values. D) The correlation of the gene expression with *in vitro* insulin secretion (SI) of the two genes above.

Discussion

Human genetics in disease have rapidly evolved over the past couple of decades thanks to new techniques such as dense genotyping arrays and NGS. The shift from microarray gene expression analysis to RNA sequencing offers an unbiased view of the transcriptome and allows for the detection of splice events, allelic imbalance effects and RNA editing events [32]. The work in this thesis could not have been performed without NGS as it was used to 1) fine map and try to explain a previous linkage to T2D on chromosome 9, 2) explain the genetic cause for glycosuria to help explore the preventive effect of chronic glycosuria and 3) to provide global maps of the pancreatic islet transcriptome and its regulation by glucose.

T2D treatment is often not straightforward and requires different approaches such as lifestyle intervention and metformin to reach glycemic goals. The SGLT2 inhibitors were recently approved for the treatment of T2D as new tool for the management of hyperglycemia in an inulin independent way [106]. We wanted to investigate the effect of SGLT2 inhibition on glucose tolerance by following a large family with FRG with respect to metabolic phenotypes. We found that the molecular cause for FRG in this family was a novel 6 bp deletion that caused a truncated protein. All carriers of the mutation had glycosuria and comparing the change in glucose tolerance of these carriers with non-carriers in the family showed that chronic glycosuria had no effect on the glucose tolerance over time. It is important to keep in mind that this was only tested in one, although large, pedigree and we cannot generalize this observation to all carriers of mutations in the *SLC5A2* gene or patients being treated with SGLT2 inhibitors. This is however the first study of this kind, and it provides initial insight into the role of SLGT2 inhibitors in the prevention of T2D.

A previous study by Lindgren et al [96] almost 15 years ago found a region on chromosome 9 linked to T2D that could not be explained at the time. With the help of denser genotypes and sequencing we wanted to fine-map and find causal genes under the linkage peak. The linkage was first redefined to chromosome 9 p21-p23 under which we found a long haplotype containing variants in the *MPDZ* gene associated with T2D in the pedigree and in a replication cohort of 343 samples ($p<0.05$). All carriers of the *MPDZ* variants except one were diagnosed with T2D in this material, but when genotyping a more heterogeneous cohort of Finnish samples we found 8 carriers, whereof 6 were non-diabetic controls. It is possible they had not yet developed T2D since they were significantly younger than the affected carriers

(38.9 ± 13.7 years in control carriers vs. 63.3 ± 10.0 in T2D carriers, p-value= 0.0002). The control carrying the variant in the main pedigree was 49 years. It is also possible that the haplotype in the more genetically distant samples is missing a vital part that contributes to the dysregulation of MPDZ leading to diabetes in the family. *MPDZ* expression was shown to be down-regulated in islets from donors with T2D and silencing of *MPDZ* in an insulin secreting rat beta cell line resulted in impaired glucose-stimulated insulin secretion. Dysregulation of the gene can therefore contribute to islet dysfunction and T2D.

The islet transcriptome is perturbed in T2D and this can be regulated by genomic variation in addition to environmental factors. The Genotype-Tissue Expression (GTEx) Project has shown that there are thousands of eQTLs in the human genome [107]. The eQTLs might explain part of the missing heritability in T2D and it is important to search for eQTLs in specific tissues, such as the islets of Langerhans [108]. We found over 600 genes to have at least one eQTL in in the islets. Expression levels of 35 of those genes were associated with the donors' glycemic levels and five of the previously known GWAS SNPs for T2D were also shown to be eQTLs in the islets. We further show that the expression of over 1600 genes differed between islets from donors with NGT, IGT and T2D in study III. The expression is associated with glycemic levels of the donors but we cannot say whether the changes are a cause or a consequence of the hyperglycemia. We addressed this in study IV where we compared short- and long-term induced gene regulation and how this correlated with insulin secretion *in vitro*. Among the 92 potentially causal genes was *ERO1LB* that was down regulated in T2D islets and the expression correlated positively with insulin secretion. A loss of function variant in this gene was associated with impaired *in vivo* insulin secretion in an independent cohort, suggesting that impaired ERO1LB function contributes to islet dysfunction in humans. The *ERO1LB* gene encodes for an oxidoreductase that is involved in the folding of pro-insulin in the ER and knockdown of *ERO1LB* in the insulin misfolding-prone Akita mouse resulted in islet destruction and development of diabetes [109]. Four other of these putative causative genes (*TMEM132C*, *DOCK10*, *PRR14L* and *IGSF11*) also harbored potential loss of function variants associated with *in vivo* insulin secretion. The only gene with increased expression in T2D and a negative correlation between expression and *in vitro* insulin secretion was transmembrane protein 132C (*TMEM132C*). Little is known about its role in islet function. Taken together, we have found evidence for potential T2D causative genes and this catalogue of gene expression in human pancreatic islets after acute and chronic exposure to glucose can serve as a resource for the dissection of molecular mechanisms leading to or protecting from T2D.

Summary and conclusions

The work in this thesis has contributed to the knowledge about the role of genetic factors, gene expression and genetic variation influencing gene expression in T2D. This was done by;

1. Showing that chronic glycosuria does not affect changes in glucose tolerance over time in a large family with familial renal glycosuria caused by a 6 bp deletion in *SLC5A2*.
2. Identifying variants in the *MPDZ* gene that may contribute to the clustering of T2D in a pedigree and showing that *MPDZ* expression levels affect glucose-stimulated insulin secretion.
3. Providing a comprehensive catalog of novel genetic variants influencing gene expression in human pancreatic islets and metabolic phenotypes to facilitate diabetes research.
4. Making a catalogue of changes in gene expression in human pancreatic islets after acute and chronic exposure to glucose that can serve as a resource for the dissection of molecular mechanisms leading to or protecting from T2D.

There is a saying that "all roads lead to Rome". This is particularly true for T2D where a host of different factors including genetic and environmental have been identified in the pathogenesis of T2D. There is however still much to learn about the mechanisms that lead to diabetes and this thesis may provide additional tools for future studies of T2D in general and islet dysfunction in particular.

Acknowledgements

This thesis would not have been possible to write without the help of so many others.

First I want to express my gratitude to my supervisor, Leif. Thank you for giving me the opportunity to do my PhD in your lab. The possibilities to learn new and exciting things in science and genetics have truly been endless thanks to you and the time with you has shown me that nothing is impossible. It is inspiring to see your enthusiasm for science and all the collaborations you've created over the course of your career. I believe that is how great things are accomplished.

Many thanks also to my supervisor, Petter. You have been such a great support for me over the years! Without your guidance I would not have made it through and despite your "great" humor I've really enjoyed working with you. Your help and your endless supply of anecdotes were always well appreciated ☺.

Thanks to all my collaborators and co-authors. Without the input and work of you, nothing in this thesis would have been possible.

A big thanks to all my co-workers in the lab. To Esa, Malin, Maria, Gabriella, Helena and Thea. Not only did you always help me with technical things by the lab bench, you were also great company and always had time for a chat.

To my other co-workers at the unit of Diabetes and Endocrinology. Peter A, Rashmi, and Om, thanks for all discussions on science and unrelated things. A special thanks to Ulrika B-N who always has a helping hand for any problem you can think of. And to Ola H who "tricked" me into undertaking this PhD by being a great supervisor during my master thesis, and who's continued to help me during these past years as well. Yang, thanks for all the good talks and the time we spent together in Boston. And of course a special shout out to my lunch company in house 60; Emma, Jasmina, Jenny, Johan, Mattias, Mikael, Petter S, Ulrika K and Åsa. You always brighten the day with your great discussions and funny comments ☺.

I would also like to thank those co-workers who already moved on. João and Nikolay for all the help with science and bioinformatics. Mona and Anna B who were there for me my first years in the lab but retired before I managed to finish this thesis.

Yuedan, I've missed you during this last time of my PhD. Thank you for being my friend and support in everything from science to relationships. You might be crazy, but I love you for it! Never ever lose that spark, enthusiasm and energy of yours.

Thanks also to all the other students that I have met at CRC throughout the years. To Anna L for your company and our chats. To Anna J who always answered my stupid questions about anything or everything. To Nina, Liliya, Mia, Hedvig, Elin H, Tasnim, Israa, Geeti, Olof, Ola E, Pradeep, Robin, Daemon, Agnes and all the rest of you that I have had the pleasure of meeting.

Also thanks to everyone at CRC for making it a great place to work and do science at.

Thank you Claes W, for the help with reading parts of this thesis and for our collaboration. It is always a pleasure to talk to you.

And thank you Malin F for all the collaboration and the advice you have given me over the years.

To Nina and Peter O. Thanks for all the dinners and good conversations. Thanks Nina for sticking with me since our days in Kalmar, I hope you will stick with me for many many years to come.

And from the bottom of my heart, thanks to my family! Despite not knowing exactly what it is I've been doing here in Skåne you all have been very keen to listen and to support me. Thanks to mom and dad for always having a bed for me to sleep in when I needed to get away for a bit. Thanks to my siblings Erika, Elin, Elof and Elias for your visits and fun times during our "syskonkvällar". Thanks to my brothers- and sisters-in-law Emanuel, Anders, Beckan and Elin for being part of this family and making it what it is. Thanks to my nephews Josh, Måns and Milo for taking my mind of unimportant things in life and being the cutest nephews ever. I love you all!

And lastly, but mostly, thank you Barry. You mean the world to me and without you I would have given up a long time ago. You always make me feel special and you give me the energy to go on and try harder when things are rough. Thank you for being the best part of me, I love you!

List of References

[1] Poretsky L (ed) (2010) Principles of Diabetes Mellitus. Springer Science & Buisness Media, New York, USA

[2] HIMSWORTH HP (1936) Diabetes Mellitus: its differentiation into insulin-sensitive and insulin-insensitive types. Lancet 1:127-130

[3] Witters LA (2001) The blooming of the French lilac. J Clin Invest 108:1105-1107

[4] Watanabe C (1918) Studies in the metabolic changes induced by administration of guanidine bases. I. Influence of injected guanidine hydrochloride upon blood sugar content. J Biol Chem :253-265

[5] Janbon M, Chaptal J, Vedel A, Schaap J (1942) Accidents hypoglycémiques graves par un sulfamidothiodiazol (le VK 57 ou 2254 RP). Montpellier Med :21-22

[6] Inzucchi SE, Bergenstal RM, Buse JB, et al. (2012) Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care 35:1364-1379

[7] World Health Organization, Dept. of Noncommunicable Disease Surveillance (1999) Definition, Classification and Diagnosis of Diabetes Mellitus and its Complications. Part 1: Diagnosis and Classification of Diabetes Mellitus.

[8] World Health Organization (2011) Use of Glycated Haemoglobin (HbA1c) in the diagnosis of diabetes mellitus
Abbreviated report of a WHO consultation.

[9] Chan JM, Rimm EB, Colditz GA, Stampfer MJ, Willett WC (1994) Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. Diabetes Care 17:961-969

[10] Colditz GA, Willett WC, Rotnitzky A, Manson JE (1995) Weight gain as a risk factor for clinical diabetes mellitus in women. Ann Intern Med 122:481-486

[11] NEEL JV (1962) Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet 14:353-362

[12] Stumvoll M, Goldstein BJ, van Haeften TW (2005) Type 2 diabetes: principles of pathogenesis and therapy. Lancet 365:1333-1346

[13] Kahn SE, Cooper ME, Del Prato S (2014) Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. Lancet 383:1068-1083

[14] Lander ES, Linton LM, Birren B, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409:860-921

[15] International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. Nature 431:931-945

[16] Kwon JM, Goate AM (2000) The candidate gene approach. Alcohol Res Health 24:164-168

[17] Ahlqvist E, Ahluwalia TS, Groop L (2011) Genetics of type 2 diabetes. Clin Chem 57:241-254

[18] Polymeropoulos MH, Swift RG, Swift M (1994) Linkage of the gene for Wolfram syndrome to markers on the short arm of chromosome 4. Nat Genet 8:95-97

[19] Inoue H, Tanizawa Y, Wasson J, et al. (1998) A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome). Nat Genet 20:143-148

[20] Karasik A, O'Hara C, Srikanta S, et al. (1989) Genetically programmed selective islet beta-cell loss in diabetic subjects with Wolfram's syndrome. Diabetes Care 12:135-138

[21] Minton JA, Hattersley AT, Owen K, et al. (2002) Association studies of genetic variation in the WFS1 gene and type 2 diabetes in U.K. populations. Diabetes 51:1287-1290

[22] Sandhu MS, Weedon MN, Fawcett KA, et al. (2007) Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet 39:951-953

[23] Dawn Teare M, Barrett JH (2005) Genetic linkage studies. Lancet 366:1036-1044

[24] Hanis CL, Boerwinkle E, Chakraborty R, et al. (1996) A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. Nat Genet 13:161-166

[25] Horikawa Y, Oda N, Cox NJ, et al. (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 26:163-175

[26] Duggirala R, Blangero J, Almasy L, et al. (1999) Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans. Am J Hum Genet 64:1127-1140

[27] Reynisdottir I, Thorleifsson G, Benediktsson R, et al. (2003) Localization of a susceptibility gene for type 2 diabetes to chromosome 5q34-q35.2. Am J Hum Genet 73:323-335

[28] Grant SF, Thorleifsson G, Reynisdottir I, et al. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 38:320-323

[29] Bush WS, Moore JH (2012) Chapter 11: Genome-wide association studies. PLoS Comput Biol 8:e1002822

[30] Brunetti A, Chiefari E, Foti D (2014) Recent advances in the molecular genetics of type 2 diabetes mellitus. World J Diabetes 5:128-140

[31] Flannick J, Thorleifsson G, Beer NL, et al. (2014) Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. Nat Genet 46:357-363

[32] Vikman P, Fadista J, Oskolkov N (2014) RNA sequencing: current and prospective uses in metabolic research. J Mol Endocrinol 53:R93-101

[33] Kaprio J, Tuomilehto J, Koskenvuo M, et al. (1992) Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. Diabetologia 35:1060-1067

[34] Newman B, Selby JV, King MC, Slemenda C, Fabsitz R, Friedman GD (1987) Concordance for type 2 (non-insulin-dependent) diabetes mellitus in male twins. Diabetologia 30:763-768

[35] Medici F, Hawa M, Ianari A, Pyke DA, Leslie RD (1999) Concordance rate for type II diabetes mellitus in monozygotic twins: actuarial analysis. Diabetologia 42:146-150

[36] Almgren P, Lehtovirta M, Isomaa B, et al. (2011) Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. Diabetologia 54:2811-2819

[37] Lehtovirta M, Pietilainen KH, Levalahti E, et al. (2010) Evidence that BMI and type 2 diabetes share only a minor fraction of genetic variance: a follow-up study of 23,585 monozygotic and dizygotic twins from the Finnish Twin Cohort Study. Diabetologia 53:1314-1321

[38] Köbberling J, Tillil H (1982) Empirical risk figures for first degree relatives of non-insulin-dependent diabetics. The Genetics of Diabetes Mellitus, Academic Press, London :201-209

[39] Groop L, Forsblom C, Lehtovirta M, et al. (1996) Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. Diabetes 45:1585-1593

[40] Groop L, Pociot F (2013) Genetics of diabetes - Are we missing the genes or the disease? Mol Cell Endocrinol

[41] Thomas CC, Philipson LH (2015) Update on diabetes classification. Med Clin North Am 99:1-16

[42] Naylor RN, Greeley SA, Bell GI, Philipson LH (2011) Genetics and pathophysiology of neonatal diabetes mellitus. J Diabetes Investig 2:158-169

[43] Anik A, Catli G, Abaci A, Bober E (2015) Maturity-onset diabetes of the young (MODY): an update. J Pediatr Endocrinol Metab 28:251-263

[44] Tong Y, Lin Y, Zhang Y, et al. (2009) Association between TCF7L2 gene polymorphisms and susceptibility to type 2 diabetes mellitus: a large Human Genome Epidemiology (HuGE) review and meta-analysis. BMC Med Genet 10:15-2350-10-15

[45] Sladek R, Rocheleau G, Rung J, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445:881-885

[46] Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, et al. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316:1331-1336

[47] Scott LJ, Mohlke KL, Bonnycastle LL, et al. (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316:1341-1345

[48] Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. (2007) A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 39:770-775

[49] Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661-678

[50] Manolio TA, Collins FS, Cox NJ, et al. (2009) Finding the missing heritability of complex diseases. Nature 461:747-753

[51] Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci U S A 109:1193-1198

[52] Lawson HA, Cheverud JM, Wolf JB (2013) Genomic imprinting and parent-of-origin effects on complex traits. Nat Rev Genet 14:609-617

[53] Kong A, Steinthorsdottir V, Masson G, et al. (2009) Parental origin of sequence variants associated with complex diseases. Nature 462:868-874

[54] Small KS, Hedman AK, Grundberg E, et al. (2011) Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat Genet 43:561-564

[55] Ottman R (1996) Gene-environment interaction: definitions and study designs. Prev Med 25:764-770

[56] Santer R, Kinner M, Lassen CL, et al. (2003) Molecular analysis of the SGLT2 gene in patients with renal glucosuria. J Am Soc Nephrol 14:2873-2882

[57] van den Heuvel LP, Assink K, Willemsen M, Monnens L (2002) Autosomal recessive renal glucosuria attributable to a mutation in the sodium glucose cotransporter (SGLT2). Hum Genet 111:544-547

[58] Calado J, Soto K, Clemente C, Correia P, Rueff J (2004) Novel compound heterozygous mutations in SLC5A2 are responsible for autosomal recessive renal glucosuria. Hum Genet 114:314-316

[59] Francis J, Zhang J, Farhi A, Carey H, Geller DS (2004) A novel SGLT2 mutation in a patient with autosomal recessive renal glucosuria. Nephrol Dial Transplant 19:2893-2895

[60] Magen D, Sprecher E, Zelikovic I, Skorecki K (2005) A novel missense mutation in SLC5A2 encoding SGLT2 underlies autosomal-recessive renal glucosuria and aminoaciduria. Kidney Int 67:34-41

[61] Yu L, Lv J-, Zhou X-, Zhu L, Hou P, Zhang H (2011) Abnormal expression and dysfunction of novel SGLT2 mutations identified in familial renal glucosuria patients. Hum Genet 129:335-344

[62] Wright EM, Loo DDF, Hirayama BA (2011) Biology of human sodium glucose transporters. Physiol Rev 91:733-794

[63] Wright EM, Hirayama BA, Loo DF (2007) Active sugar transport in health and disease. J Intern Med 261:32-43

[64] Scholl-Burgi S, Santer R, Ehrich JH (2004) Long-term outcome of renal glucosuria type 0: the original patient and his natural history. Nephrol Dial Transplant 19:2394-2396

[65] Calado J, Sznajer Y, Metzger D, et al. (2008) Twenty-one additional cases of familial renal glucosuria: absence of genetic heterogeneity, high prevalence of private mutations and further evidence of volume depletion. Nephrol Dial Transplant 23:3874-3879

[66] Lee H, Han KH, Park HW, et al. (2012) Familial renal glucosuria: a clinicogenetic study of 23 additional cases. Pediatr Nephrol :1-5

[67] Rahmoune H, Thompson PW, Ward JM, Smith CD, Hong G, Brown J (2005) Glucose transporters in human renal proximal tubular cells isolated from the urine of patients with non-insulin-dependent diabetes. Diabetes 54:3427-3434

[68] Scheen AJ (2015) Pharmacodynamics, efficacy and safety of sodium-glucose co-transporter type 2 (SGLT2) inhibitors for the treatment of type 2 diabetes mellitus. Drugs 75:33-59

[69] Wilding JP (2014) The role of the kidneys in glucose homeostasis in type 2 diabetes: clinical implications and therapeutic significance through sodium glucose co-transporter 2 inhibitors. Metabolism 63:1228-1237

[70] Abdul-Ghani MA, DeFronzo RA (2014) Lowering plasma glucose concentration by inhibiting renal sodium-glucose cotransport. J Intern Med 276:352-363

[71] Taylor SI, Blau JE, Rother KI (2015) SGLT2 Inhibitors May Predispose to Ketoacidosis. J Clin Endocrinol Metab 100:2849-2852

[72] Grarup N, Sparso T, Hansen T (2010) Physiologic characterization of type 2 diabetes-related loci. Curr Diab Rep 10:485-497

[73] Ionescu-Tirgoviste C, Gagniuc PA, Gubceac E, et al. (2015) A 3D map of the islet routes throughout the healthy human pancreas. Sci Rep 5:14634

[74] Cabrera O, Berman DM, Kenyon NS, Ricordi C, Berggren PO, Caicedo A (2006) The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. Proc Natl Acad Sci U S A 103:2334-2339

[75] Quesada I, Tuduri E, Ripoll C, Nadal A (2008) Physiology of the pancreatic alpha-cell and glucagon secretion: role in glucose homeostasis and diabetes. J Endocrinol 199:5-19

[76] Hauge-Evans AC, King AJ, Carmignac D, et al. (2009) Somatostatin secreted by islet delta-cells fulfills multiple roles as a paracrine regulator of islet function. Diabetes 58:403-411

[77] Asakawa A, Inui A, Yuzuriha H, et al. (2003) Characterization of the effects of pancreatic polypeptide in the regulation of energy balance. Gastroenterology 124:1325-1336

[78] Batterham RL, Le Roux CW, Cohen MA, et al. (2003) Pancreatic polypeptide reduces appetite and food intake in humans. J Clin Endocrinol Metab 88:3989-3992

[79] Koster JC, Permutt MA, Nichols CG (2005) Diabetes and insulin secretion: the ATP-sensitive K+ channel (K ATP) connection. Diabetes 54:3065-3072

[80] Phillips T (2008) Regulation of transcription and gene expression in eukaryotes. 1(1):199 :2016-04-14

[81] Vaulont S, Vasseur-Cognet M, Kahn A (2000) Glucose regulation of gene transcription. J Biol Chem 275:31555-31558

[82] Iizuka K, Horikawa Y (2008) ChREBP: a glucose-activated transcription factor involved in the development of metabolic syndrome. Endocr J 55:617-624

[83] Bensellam M, Van Lommel L, Overbergh L, Schuit FC, Jonas JC (2009) Cluster analysis of rat pancreatic islet gene mRNA levels after culture in low-, intermediate- and high-glucose concentrations. Diabetologia 52:463-476

[84] Fadista J, Vikman P, Laakso EO, et al. (2014) Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. Proc Natl Acad Sci U S A 111:13924-13929

[85] U.K. Prospective Diabetes Study Group (1995) U.K. prospective diabetes study 16. Overview of 6 years' therapy of type II diabetes: a progressive disease. U.K. Prospective Diabetes Study Group. Diabetes 44:1249-1258

[86] Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, Butler PC (2003) Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. Diabetes 52:102-110

[87] Ferrannini E, Gastaldelli A, Miyazaki Y, Matsuda M, Mari A, DeFronzo RA (2005) beta-Cell function in subjects spanning the range from normal glucose tolerance to overt diabetes: a new analysis. J Clin Endocrinol Metab 90:493-500

[88] Marchetti P, Bugliani M, Boggi U, Masini M, Marselli L (2012) The pancreatic beta cells in human type 2 diabetes. Adv Exp Med Biol 771:288-309

[89] Prentki M, Nolan CJ (2006) Islet beta cell failure in type 2 diabetes. J Clin Invest 116:1802-1812

[90] Alejandro EU, Gregg B, Blandino-Rosano M, Cras-Meneur C, Bernal-Mizrachi E (2014) Natural history of beta-cell adaptation and failure in type 2 diabetes. Mol Aspects Med

[91] Jonas JC, Bensellam M, Duprez J, Elouil H, Guiot Y, Pascal SM (2009) Glucose regulation of islet stress responses and beta-cell failure in type 2 diabetes. Diabetes Obes Metab 11 Suppl 4:65-81

[92] Del Prato S (2009) Role of glucotoxicity and lipotoxicity in the pathophysiology of Type 2 diabetes mellitus and emerging treatment strategies. Diabet Med 26:1185-1192

[93] Del Guerra S, Lupi R, Marselli L, et al. (2005) Functional and molecular defects of pancreatic islets in human type 2 diabetes. Diabetes 54:727-735

[94] Lupi R, Del Guerra S, Mancarella R, et al. (2007) Insulin secretion defects of human type 2 diabetic islets are corrected in vitro by a new reactive oxygen species scavenger. Diabetes Metab 33:340-345

[95] Del Guerra S, Grupillo M, Masini M, et al. (2007) Gliclazide protects human islet beta-cells from apoptosis induced by intermittent high glucose. Diabetes Metab Res Rev 23:234-238

[96] Lindgren CM, Mahtani MM, Widen E, et al. (2002) Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: the Botnia study. Am J Hum Genet 70:509-516

[97] Goto M, Eich TM, Felldin M, et al. (2004) Refinement of the automated method for human islet isolation and presentation of a closed system for in vitro islet culture. Transplantation 78:1367-1375

[98] Graham FL, Smiley J, Russell WC, Nairn R (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol 36:59-74

[99] Kanwal A, Singh SP, Grover P, Banerjee SK (2012) Development of a cell-based nonradioactive glucose uptake assay system for SGLT1 and SGLT2. Anal Biochem 429:70-75

[100] Zuhorn IS, Kalicharan R, Hoekstra D (2002) Lipoplex-mediated transfection of mammalian cells occurs through the cholesterol-dependent clathrin-mediated pathway of endocytosis. J Biol Chem 277:18021-18028

[101] WICK AN, DRURY DR, NAKADA HI, WOLFE JB (1957) Localization of the primary metabolic block produced by 2-deoxyglucose. J Biol Chem 224:963-969

[102] Calado J, Santer R, Rueff J (2011) Effect of kidney disease on glucose handling (including genetic defects). Kidney Int 79:S7; S13

[103] Isomaa B, Forsen B, Lahti K, et al. (2010) A family history of diabetes is associated with reduced physical fitness in the Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study. Diabetologia 53:1709-1713

[104] Pruitt KD, Brown GR, Hiatt SM, et al. (2014) RefSeq: an update on mammalian reference sequences. Nucleic Acids Res 42:D756-63

[105] Taneera J, Lang S, Sharma A, et al. (2012) A systems genetics approach identifies genes and pathways for type 2 diabetes in human islets. Cell Metab 16:122-134

[106] Tahrani AA, Barnett AH, Bailey CJ (2013) SGLT inhibitors in management of diabetes. Lancet Diabetes Endocrinol 1:140-151

[107] GTEx Consortium (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 348:648-660

[108] Torres JM, Gamazon ER, Parra EJ, et al. (2014) Cross-tissue and tissue-specific eQTLs: partitioning the heritability of a complex trait. Am J Hum Genet 95:521-534

[109] Zito E, Chin KT, Blais J, Harding HP, Ron D (2010) ERO1-beta, a pancreas-specific disulfide oxidase, promotes insulin biogenesis and glucose homeostasis. J Cell Biol 188:821-832

[110] Henquin JC (2000) Triggering and amplifying pathways of regulation of insulin secretion by glucose. Diabetes 49:1751-1760

[111] Cock PJ, Fields CJ, Goto N, Heuer ML, Rice PM (2010) The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. Nucleic Acids Res 38:1767-1771

[112] Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-1760

[113] Li H, Handsaker B, Wysoker A, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079

[114] McKenna A, Hanna M, Banks E, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297-1303

[115] Dobin A, Davis CA, Schlesinger F, et al. (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21

[116] Harrow J, Frankish A, Gonzalez JM, et al. (2012) GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 22:1760-1774

[117] Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930

[118] Law CW, Chen Y, Shi W, Smyth GK (2014) voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15:R29-2014-15-2-r29

[119] Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8:118-127

[120] Ritchie ME, Phipson B, Wu D, et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47