Positional cloning of type 2 diabetes genes derived from the polygenic New Zealand Obese (NZO) mouse model

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Delsi Altenhofen
aus Itapiranga-SC
Brasilien

Düsseldorf, April 2020

aus dem Institut für Klinische Biochemie und Pathobiochemie
des Deutschen Diabetes-Zentrums (DDZ)
Leibniz-Zentrum für Diabetesforschung an
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Hadi Al-Hasani
2. Prof. Dr. Axel Gödecke

Tag der mündlichen Prüfung: 28.08.2020

Summary

The pathogenesis of type 2 diabetes (T2D) is strongly driven by individual lifestyle and its corresponding genetic adaptations. Despite the fact that hundreds of genes have already been linked to T2D development, they still account for only a small percentage of the total heritability, estimated to explain up to 80 % of the disease risk. Linkage studies in mouse inbred strains have been proven efficient in identifying novel disease-associated genes. Taking advantage of this method, a novel diabetes quantitative trait locus (QTL) was identified on chromosome 15, termed *Nbg15* (NZO Blood glucose on chromosome 15, LOD 6.6 13 cM). *Nbg15* was identified in an outcross population (N_2(C3HxNZO)), generated from the two parental inbred strains New Zealand obese (NZO), a T2D mouse model, and the C3HeB/FeJ (C3H), a strain known to be lean and diabetes resistant. *Nbg15* carriers of the C3H alleles were less susceptible to develop T2D in comparison to NZO carriers of *Nbg15*. The aim of this study was to investigate the pathophysiological mechanisms and the gene variants underlying the *Nbg15*-linked phenotype. For this purpose, three major experimental strategies were combined. First, to better define a QTL confidence interval, additional genetic markers were mapped to the *Nbg15* locus and revealed a proximal (*Nbg15*p, 6.9 LOD 18.5 cM) and a distal (*Nbg15*d, 5.6 LOD and 27 cM) *Nbg15* sub-locus, both associated with lower blood glucose levels in the C3H allele carriers. Secondly, recombinant congenic mouse strains (RCS) were generated by introgression of the entire (*Nbg15*c), the proximal (*Nbg15*p) or the distal (*Nbg15*d) *Nbg15* locus from the C3H strain into the NZO genome. In accordance with the results observed in the N_2(C3HxNZO) population, homozygous allele carries of the full *Nbg15* locus (*Nbg15*-$c^{C3H/C3H}$) displayed lower blood glucose levels in combination with elevated plasma insulin levels and thus a strikingly lower diabetes prevalence in comparison to the control mice (*Nbg15*$^{NZO/NZO}$). Moreover, histological analysis of the pancreata of these mice revealed that the consomic and the distal C3H allele carriers (*Nbg15*-$c^{C3H/C3H}$ and *Nbg15*-$d^{C3H/C3H}$) of the *Nbg15* locus had an increased number of pancreatic islets compared to *Nbg15*-$c^{NZO/NZO}$ control mice. Hence, the pathophysiological mechanism underlying the diabetes-protective *Nbg15* phenotype targets the pancreatic islets of Langerhans and was exclusively phenocopied in mice carrying the C3H alleles at the distal sublocus of *Nbg15* (*Nbg15*-$d^{C3H/C3H}$). Furthermore, a third strategy consisted of a gene candidacy scanning to unravel the most likely *Nbg15* associated genes. The combination of gene expression profiling in the parental NZO and C3H strains and the backcross (N_2(C3HxNZO)) population, haplotype analysis and *in silico* predictions of gene/protein functional impact were applied. The results led to the identification of the gene *Kdelr3* (*Endoplasmic reticulum protein retention receptor 3*) as a major candidate gene affecting pancreatic islet function and/ or integrity. *Kdelr3* mRNA expression levels were higher in isolated primary NZO islets compared to islets of Langerhans from the C3H strain. Moreover, the identification of a potentially disruptive mutation in the *Kdelr3* gene variant from NZO was associated with higher T2D-associated mortality, increased blood glucose levels and lower plasmatic levels of insulin in the N_2(C3HxNZO) backcross population. Future studies aim for the investigation of KDELR3 potential function in the pathophysiology of T2D in humans.

Zusammenfassung

Die Pathogenese des Typ-2-Diabetes (T2D) wird stark vom dem individuellen Lebensstil und den entsprechenden genetischen Anpassungen bestimmt. Obwohl bereits Hunderte von Gene mit der Entwicklung des T2Din Verbindung gebracht wurden, können diese bisher nur einen kleinen Prozentsatz der gesamten Erblichkeit, die schätzungsweise bis zu 80 % des Krankheitsrisikos ausmacht, erklären. In der Vergangenheit, haben sich Kopplungsstudien mit Hilfe von Maus-Inzuchtstämmen für die Identifizierung von Krankheits-assoziierten Genen als effizient erwiesen. Mittels dieser vorteilhaften Methode wurde ein neuer quantitatives Merkmals-Locus (*Quantitative trait loci*; QTL) auf Chromosom 15 identifiziert, welcher zu *Nbg15* (NZO-Blutglukose auf Chromosom 15, LOD 6,6 13 cM) benannt wurde. *Nbg15* wurde in einer Rückkreuzungs-Population (N2 (C3HxNZO)) identifiziert, welche aus den beiden Mausinzuchtstämmen New Zealand Obese (NZO), einem populären Mausmodell für T2D, und dem als T2D-resistenten C3HeB/FeJ (C3H) Stamm erzeugt wurde. C3H-Allelträger für *Nbg15* wiesen ein geringeres T2D-Risiko im Vergleich zu NZO-Allelträgern auf. Das Ziel dieser Studie war es, die pathophysiologischen Mechanismen und Genvarianten zu identifizieren, die dem *Nbg15*-assoziierten Phänotyp zugrunde liegen. Zu diesem Zweck wurden drei experimentelle Strategien kombiniert. Um das Konfidenzintervall des QTL besser zu definieren, wurden zunächst neue zusätzliche genetische Marker für den*Nbg15*-Locus verwendet. Die genetische Analyse ergab einen proximalen (*Nbg15*p, 6,9 LOD 18,5 cM), sowie einen distalen (*Nbg15*d, 5,6 LOD und 27 cM) *Nbg15*-Sublocus, welche jeweils mit einem niedrigeren Blutglukosespiegel in C3H Allelträgern assoziiert waren. Zweitens wurden rekombinante kongene Mausstämme (RCS) durch Introgression des gesamten (*Nbg15*c), des proximalen (*Nbg15*p) oder des distalen (*Nbg15*d) *Nbg15*-Locus vom C3H-Stamm in das NZO-Genom, erzeugt. Übereinstimmend mit den in der N2(C3HxNZO) Population beobachteten Ergebnissen, zeigten homozygote Allelträger des vollständigen *Nbg15*-Locus (*Nbg15*-$c^{C3H/C3H}$) niedrigere Blutglukose Kombiniert mit erhöhten Plasmainsulin-Konzentrationen und demnach eine auffallend niedrigere T2D-Prävalenz im Vergleich zu den Kontrollmäusen (*Nbg15*$^{NZO/NZO}$). Weitere histologische Analysen des Pankreas der jeweiligen Mauslinien zeigten, dass die die C3H-Allelträger für den konsomischen und distalen Locus (*Nbg15*-$c^{C3H/C3H}$ und *Nbg15*-$d^{C3H/C3H}$), eine erhöhte Anzahl von pankreatischen Inseln im Vergleich zu NZO-Kontrollmäuse (*Nbg15*-$c^{NZO/NZO}$) aufwiesen. Basierend auf diesen Ergebnissen, wirkt der dem *Ngb15*-assoziierte Phänotyp zugrundeliegende Mechanismus vermutlich in den pankreatischen Langerhans-Inseln und konnte in C3H-Allel-Trägern des distalen Sublokus von *Nbg15* (*Nbg15*-$d^{C3H/C3H}$) phänokopiert werden. Das Ziel der dritten Strategie war die Identifizierung von Kandidatengenen, welche dem *Nbg15*-assoziierten Phänotyp zugrunde liegen könnten. Dazu wurden Genexpressionsanalysen in den Parentalstämmen NZO und C3H und der Rückkreuzungspopulation (N2 (C3HxNZO)), sowie eine Haplotypen-Analyse mit *in silico* Gen / Proteinfunktionens-Vorhersagen kombiniert. Die Ergebnisse der Analysen führten zur Identifizierung des Gens *Kdelr3* (*Endoplasmic reticulum protein retention receptor 3*) als Hauptkandidatengen, welches möglicherweise die Funktion und/oder Integrität der Langerhans-Inseln beeinflusst. Die *Kdelr3*-mRNA-Expressionsmenge in isolierten primären Langerhans-Inseln war höher im NZO-, als im C3H Stamm. Außerdem, konnte eine als „schädlich“ klassifizierte Mutation in der *Kdelr3*-Genvariante von NZO mit einer höheren T2D-assoziierten Mortalität, einem erhöhten Blutglukosespiegel, sowie einer niedrigeren Plasmainsulin-Konzentration in der

N2(C3HxNZO) Rückkreuzungspopulation assoziiert werden. In zukünftige Studien soll ein möglicher Zusammenhang der Funktion von KDELR3 mit der Pathophysiologie des T2D im Menschen untersucht werden.

Printed with the support of the German Academic Exchange Service –DAAD

Table of contents

1 INTRODUCTION

1.1 Diabetes is a worldwide disease burden

Diabetes is a chronic disease affecting millions of people worldwide. The latest estimation present a global prevalence of 425 million people living with diabetes in 2017 and among those, 98 million people over the age of 65. However, the prevalence has tremendously increased among younger ages. Among the age of 20-64, 327 million people are estimated to be affected worldwide, which represents 7.8 % of the adult population. Moreover, it is predicted that the number of diabetes affected adults will increase by 48 % until 2045, corresponding to a total of 629 million affected people in the near future (IDF 2019).

Diabetes is mainly diagnosed via the detection of elevated fasting blood glucose levels. Levels under 100 mg/dl (5.6 mmol/L) are considered normal, while levels from 100 to 125 mg/dl (5.6 to 6.9 mmol/L) are categorized as a prediabetic state while values over 126 mg/dl (7 mmol/L) result in the diagnosis of diabetes. The hyperglycemic state in diabetes originates from two major pathophysiological mechanisms by which different subtypes of diabetes can be distinguished (MayoClinic 2019).

Type 1 (T1D) diabetes accounts for about 10 % of all diabetes cases and mostly affects young adults and children. The causal risk factors leading to the manifestation of the disease have not been completely elucidated so far, but a consensus of an autoimmune-driven destruction of insulin-producing beta cells was established as one of the major underlying mechanisms. The primary lack of insulin results in a severe chronic state of hyperglycemia due to impaired glucose uptake in insulin sensing cells. Therefore, patients with T1D diabetes depend on daily injections of insulin (IDF 2019; Tan, et al. 2019).

On the other hand, 90 % of diabetes cases are classified as type 2 diabetes (T2D). The incidence is higher among the elderly population than in younger people, however, increased prevalence of obesity and unhealthy life style raised the incidence among young adults and children during the last decades. Chronic hyperglycemia as it occurs in T2D, is the result of a progressively reduced insulin sensitivity of insulin-target tissues combined with secondarily deficient insulin secretion. The prediabetic state which is characterized by an impaired insulin sensitivity can be transiently

compensated by elevated insulin secretion, thus leading to a state of not only hyperglycemia but also hyperinsulinemia. However, during progression of this pathophysiological state, beta cell failure and presumably also transdifferentiation of beta cells result in a severe depletion of circulating insulin (hypoinsulinemia). This leads to body incapacity to maintain glucose homeostasis and, consequently the onset of the disease (IDF 2019; Olokoba 2012; Rorsman and Braun 2013).

1.2 Pathophysiology of type-2 diabetes

T2D development is tightly associated to obesity. High caloric food intake combined with a sedentary lifestyle are determinant environmental factors for obesity development. Obesity is classified according to the body mass index (BMI). A BMI over 25 Kg/m^2 is classified as overweight whereas individuals with a BMI exceeding 30 Kg/m^2 are categorized as obese. The energy imbalance characteristic in an obese condition leads to the expansion of the adipose tissue, where the excess of energy is primarily stored as triglycerides. Especially intra-abdominal fat accumulation has been shown to correlate with a reduced insulin sensitivity. The central subcutaneous fat deposition is also associated to alterations in the insulin effectiveness. However, the buildup of visceral fat is believed to confer the increased risk to metabolic diseases (Chadt 2018; DeFronzo, et al. 2015; Kahn, et al. 2006).

The reduced insulin sensitivity in the adipose tissue leads to the rise of circulating triglycerides and cholesterol. High levels of triglycerides in the blood trigger a phenomenon known as “lipotoxicity”. This promotes the non-adipose tissue intra-organ accumulation of fat that disrupts the insulin signaling. One organ where the accumulation of lipids occurs is the liver, consequently leading to the development of non-alcoholic fatty liver disease (NAFLD). In this condition, the impaired insulin signaling for glycogenesis, give rise to the glucose production via glycogenolysis and gluconeogenesis, impairing the maintenance of glucose homeostasis.

However, once the peripheral response to insulin is impaired, compensatory mechanisms are activated leading to an increase insulin synthesis and release. Nonetheless, chronic exposure to energy imbalance results in the intensification of the peripheral insulin resistance and, combined to beta cell exhaustion in producing insulin, leads to the metabolic condition of irreversible hyperglycemia which represents a major feature of TD2 (Scheen 2003). Apart from environmental factors, T2D also has a strong genetic background. Individual susceptibility to develop insulin resistance and

T2D is determined by hundreds gene variants characterizing the polygenic fate of T2D (DeFronzo et al. 2015).

1.3 Molecular mechanisms in type 2 diabetes development

1.3.1 Maintenance of glucose homeostasis in a healthy state

Glucose represents the main body fuel. The maintenance of glucose homeostasis is very crucial to ensure a healthy metabolism. Thus, the regulation of systemic glycemia requires a complex network of mechanisms, the two hormones insulin and glucagon playing a key role in this process (Figure 1).

In a condition of high glucose availability, for instance directly after a meal (postprandial state), beta cells in the pancreatic islets sense the elevated blood glucose levels and activate an intracellular signaling cascade that culminates in the enhanced secretion of insulin into the blood stream. Increased levels of insulin act to stimulate glucose disposal by inducing glucose uptake in insulin sensitive cells in a variety of organs. In skeletal muscle, heart and fat, for instance, glucose uptake is activated upon insulin binding to its receptor followed by an autophosphorylation of the insulin receptor (IR). This process promotes the anchorage and phosphorylation of the insulin receptor substrate family of proteins (IRS). There are four known IRS isoforms in mammals, but IRS1 and IRS2 seem to govern the regulation of metabolic homeostasis, with certain tissue specificity. The whole body knockout of IRS1 results in insulin resistance in skeletal muscle and grown retardation while IRS2 knockout causes beta cell failure in a strain specific manner (Saltiel and Kahn 2001; Taniguchi, et al. 2006). The tyrosine phosphorylation of IRS activates the docking of phosphoinositide 3-kinase (PI3K).

This kinase converts phosphatidylinositol-4,5-diphosphate (PIP2) to phosphatidylinositol-3,4,5-trisphosphate (PIP3). Following this, PIP3 recruits the phosphoinositide-dependent kinase 1 (PDK1) that phosphorylates and activates the main effector protein kinase B (PKB or AKT). The phosphorylation of the Rab GTPase-activating proteins TBC1D1/D4/AS160 by AKT leads to its dissociation from the GLUT4 containing vesicles (GSV). As this occurs, downstream Rab GTPases proteins remain in their GTP-bound active state. This event mediates the trafficking of GSV to the plasma membrane (Leto and Saltiel 2012; Mafakheri, et al. 2018). Moreover, exercise can mediate GLUT4 translocation to the plasma membrane

independent from insulin action. In this scenario, the 5' AMP-activated protein kinase (AMPK) is activated and promote the dissociation of TBC1D1/D4/AS160 from GSVs (Larance, et al. 2005; Mafakheri et al. 2018; Peck, et al. 2006).

In the basal state, GLUT4 is mainly located intracellularly. Upon insulin stimulus, the rate of glucose uptake can increase up to 40 fold and during exercise this rate can reach up to 100 fold in muscle tissue. This specialized mechanism of glucose transport is the key metabolic component in the control of whole-body glycemia (Bryant, et al. 2002; Hou and Pessin 2007). Apart from GLUT4 that is translocated to the membrane upon insulin stimulus or muscle-contraction activated pathway, another 14 members of the family of GLUT transporters facilitate the uptake of glucose in various tissues independent from insulin action (Charron 2003).

In addition to its function as direct energy substrate via the glycolysis and oxidative phosphorylation pathway, glucose is used to store energy in the form of glycogen in the liver and, to a lower extent in the muscle. This process, is also activated by the insulin signaling cascade. The mechanism by which glycogenesis takes place involves the activation of the serine/threonine protein kinase glycogen synthase kinase 3 (GSK-3). GSK-3 dephosphorylation enables the activation of the glycogen synthase (GS). As a consequence, GS allows the glucose, bound to uridine diphosphate (UDP-Glc), to be transferred to the glycogen chain via 1-4 glycosidic bonds. In sum, these processes permit the maintenance of basal blood glucose upon high glucose availability (Cohen and Frame 2001).

In contrast, in a situation of low glucose availability, for instance during the night or during periods of prolonged fasting, several pathways are activated in order to promote the release of glucose into the blood stream. The mechanisms underlying this process are mainly activated by the insulin-counteracting hormone glucagon. Glucagon is secreted by another cell type of the pancreatic islets of Langerhans, the so-called “alpha cells”. Upon detection of low glucose concentrations followed by elevated intracellular calcium influx and, as a final event, the fusion of glucagon-containing vesicles to the cell membrane is promoted (Roder, et al. 2016).

Glucagon acts in maintaining glucose homeostasis via two major mechanisms that mainly target the liver. Firstly, glucagon binds to G-protein coupled receptors (GPCR) in hepatocytes which subsequently initiate a signaling cascade followed by activation of adenylate cyclase (AC). AC converts ATP directly into cyclic AMP (cAMP). Elevated cAMP in turn activates protein kinase A (PKA), culminating in

the activation of glycogen phosphorylase, an enzyme that promotes the final breakdown of glycogen molecules to glucose monomers that are consequently released in the blood stream (Han, et al. 2016). Secondly, glucose can be produced by the liver from non-carbohydrate substrates like pyruvate and lactate. In this process, named gluconeogenesis, phosphoenolpyruvate carboxykinase (PECK) represents the major regulatory enzyme that promotes the conversion of pyruvate back to glucose. Via these two mechanisms of insulin and glucagon counteraction, glucose levels are maintained in a physiological range during fed and fasted states (Figure 1) (Han et al. 2016; Roder et al. 2016).

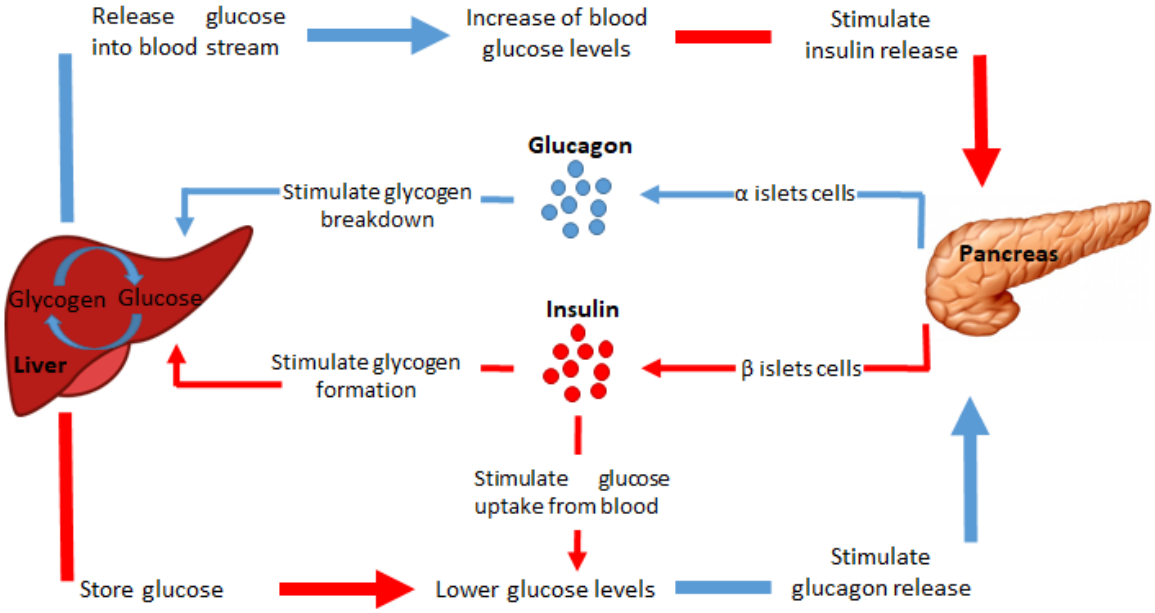


Figure 1. Hormonal maintenance of glucose homeostasis. Glucose homeostasis is mainly maintained by a coordinated counteraction of the two hormones insulin and glucagon. After a meal, elevated blood glucose levels stimulate insulin secretion from beta cells in the pancreatic islets of Langerhans. Insulin acts by promoting cellular glucose uptake both, as a direct source of energy and, in addition, as energy storage in the form of glycogen. During the fasting state, glucagon is secreted from alpha cells in the pancreatic islets of Langerhans, promoting glucose release to the blood stream via either breaking down glycogen (glycogenolysis) or via the production of glucose from non-carbohydrate precursors (gluconeogenesis), thereby maintaining glucose levels in a healthy range. Source: Modified from (Yu 2016).

1.3.2 Regulation of insulin and glucagon synthesis and secretion

Insulin, one of the key hormones regulating whole-body glycemia, is structurally composed of 51 amino acids, which are distributed in two chains (A and B) connected by two disulfide bounds. In humans, insulin is encoded by the *INS* gene, while in rodents two genes, *Ins1* and *Ins2,* have been identified to code for the insulin molecule (Rorsman and Braun 2013).

The synthesis of insulin (Figure 2) is started upon binding of distinct transcription factors to the promoter region of the *INS* gene. *Pdx1*, *MafA* and *NeuroD1* are major elements promoting the transcription of the *INS* gene into the immature preproinsulin hormone. The preproinsulin contains a signal peptide that is cleaved by a signal peptidase eliciting its transport into the lumen of the endoplasmic reticulum (ER). In the ER, the prohormone (proinsulin) is folded into its tertiary structure composed by three disulfide bounds. This process requires a complex interaction between chaperone proteins in the ER, which also participate in the transport of the immature hormone to the Golgi network, where the proinsulin is packed into vesicles (Fu, et al. 2013; Tokarz, et al. 2018).

Moreover, in the Golgi, the proinsulin undergoes a maturation process that, in a first step, requires the convertase enzymes known as PC1/3 and PC2, followed by carboxypeptidase E action. These enzymes cleave the prehormone into its mature A and B chains, thus separating them from the remaining C-peptide. Furthermore, six insulin molecules are coupled with Zn^{2+}, forming a hexamer whose transfer into mature granules is mediated by related zinc transporters like ZnT8. The mature granules are stored in the cytoplasm of the beta cells in the pancreatic islets of Langerhans and are docked to the plasma membrane upon specific stimulus triggered by several elements, being glucose the major sensitizer (Fu et al. 2013; Tokarz et al. 2018).

Glucose enters the beta cells via the facilitative glucose transporter 2 (GLUT2) in rodents or GLUT1 in humans. GLUT2 is abundantly expressed in the plasma membrane. This characteristic enable the efficient transport of glucose in fasted and feed states being the transport rate not a limitation factor in the glucose utilization. On the other hand, experimental observations in *Saccharomyces cerevisiae* postulated that GLUT1 km for glucose is around 3.4 mM. In contrast, the kinetics of GLUT1 in human is not very well elucidated so far. Apparently, GLUT1 undergoes specific exoplasmic conformational changes that enable glucose binding and transport. (Augustin 2010; Galochkina, et al. 2019; Kasahara and Kasahara 1996; Mueckler and Thorens 2013). In the cytoplasm of the beta cells, the glucose is converted to glucose-6-phosphate (G6P) by the glucokinase (GK). G6P undergoes the glycolysis pathway which generates pyruvate, feeding the citric acid cycle for the production of ATP via coupled oxidative phosphorylation. Increased cytoplasmic levels of ATP lead to the blockage of ATP-dependent potassium channels (K_{ATP}) eliciting the depolarization of the cell membrane. As a consequence, calcium influx is enhanced by the opening of

voltage-dependent calcium channels (VDCCs). Increased levels of intracellular calcium is a key component in the process of docking and fusion of the insulin granules to the plasma membrane. Furthermore, fenestrated vessels surrounding the pancreatic islets allow the rapid distribution of insulin to its target tissues (Fu et al. 2013; Tokarz et al. 2018).

The secretion of insulin in known as biphasic. Upon raised blood glucose levels a transient secretion of insulin occurs to rapidly activate the glucose disposal machinery. This first phase last for only a few minutes and is followed by a sustained second phase of secretion (Hou, et al. 2009). In humans, the first phase is more prominent than in rodents. The first phase results from the release of 1 % from the insulin granules that are organized in a ready releasable pool (RRP), already docked to the plasma membrane and immediately released upon the glucose stimulus (Rorsman, et al. 2000; Rorsman and Renstrom 2003). On the other hand, 99 % of the granules remain in a reserve pool that requires cytoskeletal, more precisely F-actin, remodeling for activation and mobilization of the granules to the RRP (Fu et al. 2013; Seino, et al. 2011). This second phase is also referred to as amplifying phase and requires the activation of small GTPases, most importantly the cell division cycle 2 (Cdc42) and the Ras-related C3 botulinum toxin substrate 1 (Rac1). Ultimately, the interaction of insulin granules with the vesicle-associated membrane protein 2 (VAMP-2) and the N-ethylmaleimide sensitive factor attachment receptor (SNARE) proteins located at the plasma membrane. These, also known as syntaxins and SNAPs, permit the formation of a complex that allows the fusion of the vesicles to the plasma membrane and consequently, the release of the insulin cargo (Kalwat and Thurmond 2013).

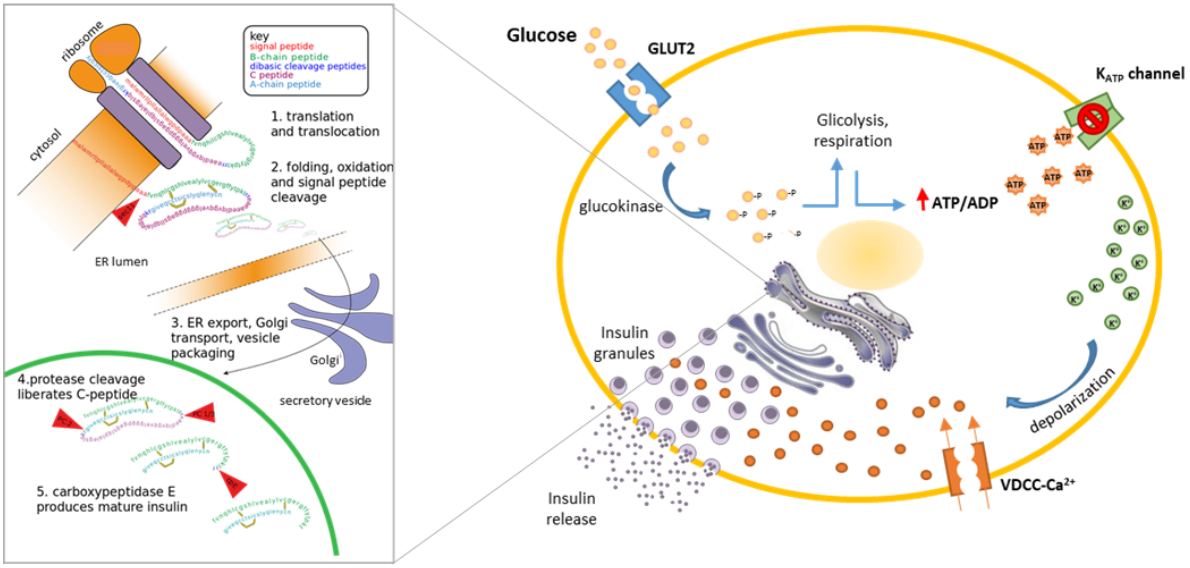


Figure 2. Insulin synthesis and secretion. Insulin is synthetized as a preprohormone. The preprohormone signaling peptide is first cleaved by a peptidase that enables the transport of the pro-hormone to the Golgi apparatus to undergo maturation processes. In the Golgi, further cleavage steps catalyzed by carboxipeptidases form the final mature insulin with A and B peptide chains connected through two disulfide bonds and release the hormone from the previous bound C-peptide. Furthermore, insulin is packed into vesicles as a zinc hexamer and transported to the cytoplasm. Increased blood glucose levels are the major stimulus for enhanced insulin secretion. Intracellular metabolism of glucose leads to increase the ratio ATP/ADP. Consequently, the closure of ATP-dependent potassium channels (K_{ATP}) triggers cell membrane depolarization, an event that promotes the opening of calcium channels dependent on voltage (VDCC-Ca^{+2}). The resulted elevation in intracellular levels of calcium is responsible for inducing the translocation and fusion of insulin-containing granules to the plasma membrane. Source: Modified from: (BCBC 2019, Yonemoto 2014).

Glucagon is encoded by the *Gcg* gene and, similar to insulin, is also synthetized as a preprohormone in the alpha cells of the pancreatic islets of Langerhans. This preprohormone gives rise to several other hormones like the glucagon like-peptide 1 and 2 (GLP-1 and GLP-2), with glucagon-opposing effects on glucose metabolism. The generation of glucagon, however, depends on the action of tissue-specific prehormone convertases (PC), enzymes that determine the cleavage of the preprohormone to glucagon (Gaisano, et al. 2012; Gromada, et al. 2018).

Glucagon is a 29 amino acid peptide overexpressed upon binding of the transcription factors Pax6 and Foxa1 and 2. The machinery by which alpha-cells stimulate glucagon secretion is very similar to the beta cells secreting insulin. Electrical pulses promote the opening of calcium channels that trigger the releasing of the glucagon-containing vesicles. However, this process is activated by lower levels of glucose, while high glucose concentrations inhibit glucagon release. One of the underlying mechanisms in this regulation involves the blockage of the ATP-dependent potassium channels. High levels of glucose in the beta cells produce elevated intracellular levels of ATP, consequently leading to the closure of K_{ATP} channels. On

the other hand, in alpha cells, lower levels of ATP appear to block K_{ATP} channels and elicit the activation of the glucagon secretion machinery (Muller, et al. 2017; Scott and Bloom 2018). Furthermore, intrinsic regulation of glucagon secretion also appears to be triggered by Na^+. Apart from GLUT1/2 glucose transporter, alpha cells sense the glucose fluctuations via the sodium-dependent glucose cotransporter 2 (SGLT2). The co-transport of Na^+ via these channels contribute to enhanced membrane conductance and therefore trigger the mechanisms that inhibit glucagon secretion (Briant, et al. 2016). Additionally, insulin and somatostatin, secreted by the beta and gamma cells respectively, compose the paracrine component of glucagon secretion regulation.

In a healthy individual, the secretion of insulin and glucagon is coordinated by several complex mechanisms that permit glucose or energy supply in both fasted and fed conditions. However, the chronic imbalance between energy intake and consumption can disrupt these mechanisms leading to the development of sustained hyperglycemia characteristic of T2D onset.

1.4 Molecular mechanisms of insulin resistance

The mechanisms by which impaired insulin signaling occurs are very complex and involve a multifactorial network of events that are not completely understood. However, some hypothesis that include the lipids overflow, low-grade inflammation and oxidative stress have been postulated as hallmarks in the development of insulin resistance (Chadt 2018; Haeusler, et al. 2018; Samuel and Shulman 2012).

1.4.1 Lipid-induced insulin resistance

The adipose tissue have been for long time referred as the main organ for energy storage. However, it is nowadays known that apart from adipocytes, the adipose tissue is composed by several cell types that secrete several metabolic mediators directly affecting the glucose metabolism. These include immune cells like macrophages and lymphocytes (Ouchi, et al. 2011).

In a normal condition, the adipocytes store energy in the form of triacylglycerol (TAG), upon insulin-stimulus for lipogenesis. On the other hand, during fasting, the glucagon stimulates the lipolysis of TAG which leads to the generation of dyacylglycerol (DAG) and free fatty acids (FFA). Those are energy fuels that are further oxidized to generate ATP in the citric acid cycle (Duncan, et al. 2007; Habegger, et al. 2010).

Conversely, in obesity, the constant excess of energy supply promotes the expansion of adipocytes in order to increase the TAG storage. However, this expansion requires significant changes in the structure of the adipocytes itself and the whole surrounding vascular network. The adipose tissue is a large endocrine organ secreting many metabolic modulators known as adipokines that modulate the differentiation and grow of adipocytes. In obesity, the grown of adipocytes is generally associated to a higher metabolic risk than the expansion of the cells known as adipogenesis (Ghaben and Scherer 2019). In this matter, immune cells as well as the adipocytes undergo phenotypic changes that promote the blockage of insulin-inducing lipogenesis in an attempt to avoid the tissue damage (Ouchi et al. 2011).

Furthermore, the impaired insulin signaling in adipocytes give rise to counteracting adaptations that lead to the release of TAG to the bloodstream known as lipolysis. The circulating lipids tend, as a consequence, to accumulate within organs like the liver, skeletal muscle and the pancreas. In these, the metabolism of TAG leads to the generation of high levels of DAG. DAG acts as an intracellular signaling molecule promoting in a first step the activation of the novel protein kinase C (PKC). PKC is a serine/threonine kinase that can potentially inactivate the insulin receptor and its substrate IRS-1 via phosphorylation. Moreover, the FFA metabolism can also generate ceramides. Ceramides are able to inhibit the insulin cascade by directly dephosphorylating Akt2, not participating on upstream regulation of insulin signaling, although evidences suggest an interaction with the PKC as well (Samuel and Shulman 2012).

1.4.2 Inflammation markers of insulin resistance

In addition to the ectopic accumulation of fat, cytokines released by immune cells that are thought to infiltrate the adipose tissue are also known to act as metabolic modulators of insulin signaling. It has been demonstrated that obese and T2D patients have increased adipose tissue infiltration of macrophages (Kratz, et al. 2014; Lu, et al. 2019; McLaughlin, et al. 2017). This event appears to correlate with the hypertrophy of the adipocytes during obesity. Recent evidence suggests that already at early stages of obesity development, the proliferation of resident macrophages is induced. During obesity progression, monocytes are also recruited from the blood stream and infiltrate the adipose tissue where the differentiation to macrophages occurs (Zheng, et al. 2018). The majority or around 50 % of the immune cells in the obese adipose tissue

are macrophages, while in lean adipose tissue, the macrophages represent only 10 % of this fraction (Lu et al. 2019; Xu, et al. 2003).

The infiltration of immune cells in the adipose tissue appears to be a result of the first metabolic adaptation of adipose tissue towards obesity. Once this is followed by the generation of new adipocytes and neovascularization, the maintenance of normal energy supply is guaranteed and no metabolic distress is observed. However, with the exacerbation of adipose tissue hypertrophy, a space distress is generated leading to scarce vascular supply of nutrients as well as oxygen, known as hypoxia. This condition is believed to promote a metabolic switch in the immune cells fate from the anti-inflammatory towards a pro-inflammatory state. Thus, the adipocytes as well as macrophages increase the release of pro-inflammatory cytokines that can potentially disrupt the insulin signaling cascade (Hsieh 2017; Kratz et al. 2014; Lu et al. 2019).

One of these mediators is the Interleukin-6 (IL-6). High plasmatic levels IL-6 are found in obese mice and human, and the blockage of IL-6 receptor has been shown to improve insulin signaling in obese patients. The mechanism by which IL-6 can alter the insulin signaling may result from the reduced phosphorylation of the IRS in hepatocytes (Senn, et al. 2002). In adipose tissue it can activate lipolysis thereby contribute to the ectopic accumulation of fat and prompt the infiltration of pro-inflammatory macrophages (Kraakman, et al. 2015; Trujillo, et al. 2004).

Additionally, infiltrated macrophages and in less extend adipocytes secrete the tumor necrosis factors-*α* (TNF-α). TNF-α is a major effector of reduced insulin sensitivity in the adipose tissue (Weisberg, et al. 2003). The intracellular signaling induced by TNF-α triggers the activation of the c-Jun N-terminal kinase-1 (JNK1) a direct inhibitor of IRS. The branch TNF-α/JNK1 is a major mechanism associating the inflammatory signals to insulin resistance (Cawthorn and Sethi 2008; Makki, et al. 2013; Samuel and Shulman 2012). Furthermore, TNF-α as well as free-fatty acids via the toll-like receptor 4 (TLR4), activate the intracellular IκB kinase (IKK)/ nuclear factor kappa B *(NF-kB*) promoting the blockage of insulin signaling and thereby insulin resistance, together with the nuclear induction of inflammatory genes in adipose tissue and immune cells (Tornatore, et al. 2012).

Furthermore, similar mechanisms contribute to the development of insulin resistance in the skeletal muscle and hepatocytes. However, it is not clearly known if insulin resistance in the liver and skeletal muscle is primary initiated by the ectopic

accumulation of fat in these tissues, or the low-grade inflammation initiated in the adipose tissue is what drives a crosstalk between insulin resistance in liver and skeletal muscle (Tornatore et al. 2012).

1.4.3 Unfolded Protein Response (UPR) and the pathogenesis of insulin resistance

The UPR or endoplasmic reticulum stress (ER) is also recognized as a major component in the pathogenesis of insulin resistance. The ER is the main organelle involved in protein maturation and folding. Therefore, the stress generated by certain metabolic conditions like obesity can drastically affect its function by promoting the accumulation of misfolded proteins. A major mechanism activated in this condition is the so-called “unfolded protein response (UPR)” which is considered a compensatory pathway to prevent the overload of misfolded proteins. This response is composed by several proteins that can trigger cell apoptosis upon failure of proper adaptive signals (Flamment, et al. 2012).

It was demonstrated that UPR proteins are overexpressed in obese mice and humans, reason that linked UPR response to insulin resistance and the pathogenesis of T2D. In an obese individual, the activation of UPR appears to be beneficial for the restoration of normal ER function. However, a chronic signal for its activation results in inefficient stabilization of the ER and, as a consequence, the release of misfolded proteins as well as the activation of cell apoptosis. The chronic low-grade inflammation in obesity development is proposed to work as a signal triggering the UPR response. However, both, hypoxia and depletion of intracellular calcium (Ca^{+}) seem to be important additional signal effectors (Khan and Wang 2014; Ozcan, et al. 2009).

The UPR response is composed by three major upstream effectors, the protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK), the serine/threonine-protein kinase/endoribonuclease inositol-requiring enzyme 1 (IRE-1) and the activating transcription factor 6 (ATF6). Under healthy conditions, the UPR response is maintained inactive due to its association with the chaperone immunoglobulin heavy chain-binding protein/glucose regulated protein 78 (BiP/GRP78). The generation of ER stress in turn dissociates the BiP/GRP78 eliciting it binding to misfolded proteins, thereby activating IRE1, PERK, and ATF6 upon its specific signals (Ozcan et al. 2009; Sundar Rajan, et al. 2007)

IRE-1 and, in a condition of chronic stress, PERK and ATF6 can activate the response mediate*d* by NF-kB and JNK that, as mentioned above, block insulin signaling via phosphorylation of IRS-1 on specific serine residues. These appear to be the major mechanisms by which insulin resistance is installed upon ER stress in the liver. Moreover, in adipose tissue, the direct induction of IL-6 secretion is believed to be one of the major mechanisms. In muscle, no concrete evidences exist to relate ER stress as causal factor for insulin resistance, being the accumulation of fatty acids a factor that is believed to precede the ER stress (Deng, et al. 2004; Hu, et al. 2006; Urano, et al. 2000; Yamazaki, et al. 2009).

1.5 Molecular mechanism driving beta cell dysfunction in T2D

A major compensatory mechanism to overcome the peripheral insulin resistance occurring during the pathophysiology of T2D is the increase of insulin synthesis and secretion by the endocrine pancreas. Obese patients with pronounced insulin resistance do not necessarily develop overt T2D if the compensatory mechanism of insulin secretion is maintained. Hence, the failure in insulin secretion is another hallmark in the pathogenesis of T2D (Prentki and Nolan 2006).

The activation of UPR response via ER stress mediated mechanisms is also a major factor contributing to the beta cell dysfunction in type 2 diabetes. The increased demand of insulin during obesity activates UPR response via IRE-1 in order to increase insulin gene transcription, thus also promoting cell survival. This appear to happen via induction and splicing of the X-box binding protein 1 (XBP-1). Conversely, the chronic exposure to high levels of lipids and glucose, termed lipo and glucotoxicity appear to trigger different modulatory features, via the same signaling effectors, that on the other hand can activate cell death signals (Fonseca, et al. 2011). Lipotoxicity activated apoptosis appears to be initiated by the CCAAT/-enhancer-binding protein homologous protein (CHOP) and the activating transcription factor 4 (ATF4). Moreover, intrinsic apoptosis can be activated independent of CHOP/ATF4, but by the tumor protein p53 and the Bcl-2-binding component 3 (BBC3 or PUMA) upon palmitate-stimulated PERK activation. Nonetheless, lipotoxicity seem to only directly affect beta cell mass, being the glucotoxicity the main mediator in the failure of the insulin secretory machinery (Biden, et al. 2014; Cantley and Ashcroft 2015; Halban, et al. 2014; Poitout and Robertson 2002).

Additionally, chronic exposure to high glucose levels increase mitochondrial metabolism and thereby the generation of oxidative reactive species (ROS). In a healthy state, the generation of ROS is neutralized by antioxidant defenses that involve the superoxide dismutase (SOD) enzyme. However, once SOD scavenger capacity is exceeded, ROS can be released and act in many ways promoting DNA damage and triggering cell apoptosis. *Pdx1*, an important transcription factor for the proinsulin synthesis is known to be directly affected by ROS, subsequently leading to beta-cell demise (Drews, et al. 2010; Gerber and Rutter 2017). In addition, beta cells are also primary targets for oxidative stress due to their low antioxidative capacity (Drews et al. 2010; Lenzen 2008).

Furthermore, it was demonstrated that isolated islets from diabetic patient have an accumulation of amyloid aggregates of the insulin amyloid polypetide (IAPP), also known as Amylin. IAPP is a 37 residue peptide co-secreted with insulin by the pancreatic beta cells. IAPP complements the insulin effect on the regulation of glucose homeostasis by suppressing glucagon secretion, slowing gastric emptying and promoting satiety via central nervous system signals (Mietlicki-Baase 2016; Xi, et al. 2019). Its beneficial effects on glucose homeostasis reflect the monomer form of the protein. On the other hand, hypersecretion of IAPP, as occurs in diabetes, favors the aggregation into oligomers follow by the formation of fiber-like deposits (Akter, et al. 2016; Asthana, et al. 2018; Qiao, et al. 2019). These non-degradable amyloid deposits exert cytotoxic effect and have been repeatedly associated to beta cell apoptosis (Kanatsuka, et al. 2018; Nedumpully-Govindan and Ding 2015). Moreover, amyloidogenic forms of the protein or the release of misfolded IAPP might be associated to variable risk to and/or favor the progression of beta cell dysfunction in T2D (Kanatsuka et al. 2018; Wu and Shea 2013).

Together with IAPP, the infiltration of immune cells, including macrophages, are described as major inflammatory mediators that promote beta cell damage. Therapy with anti-inflammatory drugs has been demonstrated to be able to preserve beta cell function (Donath 2014; Halban et al. 2014).

In summary, many mechanisms converge together to trigger adaptive responses to beta cell overload. However, the responsiveness to these adaptions significantly vary among diabetic patients due to mechanisms that are not completely understood. However, individual genetic predisposition also appears to be a pivotal

factor in the regulation of metabolism between adaptive and pro-apoptotic fates (Ahlqvist, et al. 2011).

1.6 Genetic factors in the pathogenesis of type-2 diabetes

It is recognized that a complex interaction between lifestyle and genetic factors are major determinants of diabetes susceptibility. With regard to diabetes progression, the genetic background appears to be crucial in the maintenance of adaptive responses (Diamond 2003).

One of the first links between genetics and T2D development could be observed in the Nauru population, a small island in the Pacific Ocean. The discovery of a commercially fertilizer in the rock of this island completely reverted the lifestyle of Nauruans from genetically adapted to scarce access to food or long period of starvation to a sedentary and abundant food availability. These changes resulted in an epidemic incidence of diabetes, affecting 2 out of 3 adults over the age of 55 (Dean L 2004).

Further studies in human populations have postulated that the risk of T2D development during life course is 70 % greater in individuals whose parents are diagnosed with the disease. If only one of the parents is diabetic, the risk is still 40 % greater and it seems to raise if the mother is the one affected. Additionally, between monozygotic twins, the T2D incidence can reach 70 % and is reduced to 20-30 % in dizygotic tweens (Groop, et al. 1996; Kaprio, et al. 1992). These results were reproduced in other studies with general predictions ranging from 25-80 % (Prasad and Groop 2015).

Among different ethnic groups both, the diabetes incidence as well as the progression have a strong genetic predisposition. A study performed among immigrants from the Middle East living in Sweden demonstrated that this subpopulation had a 2-3-fold greater chance to develop diabetes compared to the native Swedish population. Moreover, compared to the Swedish population, Middle East residents were also diagnosed with type-2 diabetes earlier than Swedish and had generally lower levels of C-peptide. In addition, the T2D heritability also seems to be higher among Middle East patients (Glans, et al. 2008). Given these findings, it is recognizable that environmental factors are a key component during T2D development, however, genetic factors are also decisive in this context since specific genetic modulations upon environmental changes are often determinants in the diagnosis and progression of the disease (Ahlqvist et al. 2011; Das and Elbein 2006). Furthermore, the genetic

modulation of T2D does not only originate from familial monogenic defective genes but can be spontaneously inherited due to epigenetic modulations over several generations (Van Baak, et al. 2018).

The advances in the genetic field including the completion of the human genome project in 2001 enable a better understanding of the genetic architecture of T2D (Naidoo, et al. 2011). Using these approaches, linkage studies have been conducted to dissect the genetic causality of T2D with higher penetrance. As a result, a class of monogenic forms of diabetes affecting young adults was identified termed the Maturity-Onset-Diabetes of the Young (MODY). Those were the first non-autoimmune and non-obese associated monogenetic forms of diabetes described. More than 10 genes have been found so far to affect this specific type of diabetes, the hepatocyte nuclear factor-1A (*HNF1A*) and the glucokinase (*GCK*) gene representing the most common ones. Mutations in *HNF1A* and *GCK* were found to be extremely deleterious resulting in acute hyperglycemia (Ali 2013).

On the other hand, apart from rare gene mutations, T2D is commonly influenced by a multiplicity of genes distributed across the entire genome, each of them generally displaying a low degree of penetrance. Therefore, further advances in the genetic field including high-throughput SNP genotyping technology and the development of the human haplotype map (HapMap) facilitated the investigation of gene variants associated to T2D susceptibility. Using these approaches, human genomic association studies (GWAS) were conducted. In these studies, single nucleotide polymorphisms (SNPs) are used to identify gene variants frequently associated with a certain trait like diabetes for example. More recently, whole genome sequence is also being used as input in the search for novel gene variants associated to diseases (Tam, et al. 2019).

Already identified by linkage studies, *TCF7L2* was the first gene also identified in GWAS with a strong link to T2D (Duggirala, et al. 1999). TCF7L2 is a transcription factor that directly affects insulin secretion (Sladek, et al. 2007). Diabetic patients were found to overexpress a risk variant of *TCF7L2* in Langerhans islets that was demonstrated to correlate with decreased insulin secretion (Gjesing, et al. 2011; Schafer, et al. 2007). Hematopoietically expressed homeobox (*HHEX)* is another gene repeatedly reproduced by many GWAs to be associated with T2D (Li, et al. 2012). Associated to islet disruption, recent publication demonstrated that *HHEX* downregulation in mouse liver seems to drive the insulin secretion deficiency in islets (Lu, et al. 2018). Alterations in beta cell function have also been associated with the

zinc transporter solute carrier family 30 (*SLC30A8*) and insulin-like growth factor 2 mRNA binding protein 2 (*IGF2BP2*) in several studies (Ali 2013; Strawbridge, et al. 2011; Xu, et al. 2012). Another example, a SLC16A11 gene was identified as a T2D risk variant among Mexicans. Alteration of fatty acid and lipid metabolism were identified as major metabolic features addressing this risk (Consortium, et al. 2014; Tam et al. 2019).

Additionally, GWAS studies on T2D subjects identified several obesity genes, being fat mass and obesity-associated (*FTO)* gene and Melanocortin-4 receptor (*MC4R*) the most prominent ones. *FTO* was found to be associated with the trait BMI in diverse genomic association studies. MCR4R on the other hand, plays a role in regulating energy intake and expenditure. Although strongly associated with obesity development, these genes only have a mild influence on T2D manifestation (Fawcett and Barroso 2010; Hess and Bruning 2014; Qi, et al. 2008; Scuteri, et al. 2007).

1.7 The identification of new T2D diabetes associated genes

Despite of the discovery of several T2D gene variants, these still account for only a small percent of the overall T2D heritability, estimated to be between 20-70 % depending on family disease circumstances as aforementioned (Ahlqvist et al. 2011; Wheeler and Barroso 2011). Thus, many gene variants still remain unknown. Moreover, most of the GWA studies have been very efficient in detecting gene variants associated to beta cell dysfunction and thereby implicated in the causality of the disease (Ali 2013). However, only weak signals were found to affect insulin sensitivity. In addition, regardless of the discovery of these genes, most of the polymorphisms associated to T2D matched to intergenic location thereby shadowing the molecular link to the causal variants (Bonnefond and Froguel 2015).

On the other hand, the use of GWAS in finding disease-associated genetic variants have some further disadvantages as no control of environmental factors and lifestyle of the subjects is covered. Therefore, confounding factors are frequently affecting the resolution of the analysis. Consequently, the results of GWA studies do not imply causation and the ultimate confirmation of causality requires the functional characterization of a gene or gene variants (Broman 2009; Joost and Schurmann 2014).

Mouse models are frequently used to overcome these adversities as environmental factors can be tightly controlled. In addition, a better control of the

genetic composition is possible by selecting inbred strains, which are strains resulted from brother-sister breeding and therefore considered genetically identical. Therefore, using inbred strains for genetic linkage studies results in an association that have stronger evidence for being causal. This is achieved by the magnification of genetic association when judiciously choosing the mice to cross. In addition, genetic variation can be effectively accomplished in the generated progeny. Moreover, an important aspect to be considered in mouse studies is that mouse models that highly differ in the metabolic trait of interest can be chosen to allow a better resolution of the genetic association analysis (Broman 2009).

1.8 Mouse models for the study of genetic susceptibility of T2D

About 99 % of the mouse genes share homology to the human genome, being therefore a gold standard in the study of human chronic genetic-associated disorders (Guenet 2005). The generation of laboratory mice started in the early twentieth century with the aim to breed mice for different coat colors and comportments. As a consequence, inbred strains that are currently used in research consist of a genetic recombination of at least four subspecies of the ancestral *Mus musculus*: *M. m. musculus* and *M. m. domesticus*, *M. m. castaneus* and *M. m. molossinus* (Petkov, et al. 2004; Wade, et al. 2002). Years of inbreeding resulted in mouse strains very diverse in their genetic susceptibility towards T2D development, some of those closely resembling the development of the human disease. Spontaneous gene mutations occurred during inbreeding and gave rise for instance to the known T2D-prone db/db mouse strain, in which the leptin receptor deficiency produces a severe energy imbalance leading to diabetes onset at early stages of life (Castracane 2007).

Apart from point mutations, the majority of the generated inbred mouse strains have a specific T2D genetic susceptibility that result from a combination of several gene variants, thereby also resembling the human polygenic background of T2D. For example, the New Zealand obese mouse (NZO) strain is a known polygenic mouse model of T2D. In contrast, C3HeB/FeJ and C57BL/6J strains are examples for mouse lines known to be resistant towards high-fat diet-induced T2D.

Figure 3. Mouse strains in diabetes research. New Zealand Obese Mouse (NZO) (A), C3HeB/FeJ (B) and C57BL/6J (C). Source: (Laboratory 2019).

1.8.1 New Zealand obese mouse (NZO) strain

The NZO mouse strain was initially generated in 1948 by M. Bielschowsky. After several breeding steps that included a common origin from the New Zealand Black mice (NZB), part of the progeny started to develop severe obesity at early time points of life, with body fat exceeding 40 % of the total body weight in later stages of life. These mice were further bred as NZO and later used as a polygenic model of obesity (Bielschowsky and Goodall 1970; Haskell, et al. 2002). Furthermore, the NZO strain was soon also been attributed as a model for obesity-associated insulin resistance and T2D. About 50 % of NZO males on a standard chow diet develop severe hyperglycemia and overt diabetes that is associated to a rapidly impaired beta cell function, while some mice develop severe obesity without hyperglycemia. Interestingly, some substrains of NZO females are also protected from obesity associated T2D (Junger, et al. 2002; Leiter and Reifsnyder 2004; Leiter, et al. 1998).

Similar to the human metabolic syndrome, which is characterized by the combination of several metabolic abnormalities such as high blood pressure, elevated blood glucose levels, elevated BMI, and abnormal cholesterol or triglyceride levels, the high-fat diet intervention in NZO mice produces a rapid gain of weight, accompanied by evident accumulation of visceral fat. In addition, these mice develop severe insulin resistance with persistent hyperglycemia, hypercholesterolemia and hypertension. These characteristics elected NZO a common model for the study of the complex genetic predisposition of T2D (Jurgens, et al. 2006). Although the phenotype associated to NZO obesity have been partially attributed to the existence of a polymorphic leptin receptor in NZO that is associated to hyperleptinemia and resistance to leptin action, the genes that contribute to T2D susceptibility in these mice remain to be discovered (Igel, et al. 1997).

Furthermore, NZO has been used as model for the identification of obesity and diabetogenic genes via linkage analysis. The genes TBC1 domain family member 1

(*Tbc1d1)* and the zinc finger protein 69 (*Zfp69)* were identified as obesity associated type 2 diabetes risk in an outcross population generated with the NZO and the lean SJL mouse. Thus, NZO remains as a major mouse model for the genetic dissection of the human metabolic syndrome (Chadt, et al. 2008; Joost 2012; Laboratory 2019; Scherneck, et al. 2009).

1.8.2 C3HeB/FeJ and C57BL/6J mouse strains

The C3HeB/FeJ mouse strain was generated in 1948 by insemination of C57BL/6 with C3H/HeJ fertilized ova. On the other hand, C57BL/6J mice were generated 27 years earlier from Lathrop's stock and are currently the most widely used inbreed strain in research (Lyon 1996; Tucker, et al. 1992). Comparing to C57BL/6J, C3HeB/FeJ mice display elevated food intake and body fat mass. Additionally, C3HeB/FeJ fasting glucose levels are higher compared to C57BL/6J. However, it was demonstrated that the C3HeB/FeJ strain demonstrates improved glucose tolerance compared to C57BL/6J mice. Although C3HeB/FeJ mice develop phenotypic characteristics of the human metabolic syndrome with a combination of elevated plasma glucose, cholesterol and triglyceride levels, they appear to have better glucose tolerance than the C57BL/6J mice (Champy, et al. 2008). This phenotype has been attributed to the fact that the C3HeB/FeJ mice show higher insulin secretion, presumably as a result of an improved pancreatic islets integrity and compensatory machinery. In contrast, the C57BL/6J mice carry a mutation in the nicotinamide nucleotide dehydrogenase (NADH) that confers impaired mitochondrial metabolism in beta cells and consequently reduces islets integrity (Freeman, et al. 2006).

Both strains also diverge on their sensitivity towards diabetes-induction by streptozotocin (STZ) treatment. While the C3HeB/FeJ mice appear to be resistant to a single dose of 150 mg/Kg body weight of STZ, C57BL/6J mice develop hyperglycemia (Kaku 1989). Nevertheless, both strains show a high degree of resistance to develop T2D diabetes upon high-fat diet intervention and are protected from beta cell dysfunction in contrast to the NZO strain (Champy et al. 2008).

1.9 Mouse linkage analysis and positional cloning as strategy for the identification of novel diabetes susceptibility genes

The use of mouse models to uncover diabetes susceptibility genes is being used as a straightforward approach and offers some advantages in comparison to GWAS as aforementioned (Clee and Attie 2007; Cox and Church 2011).
Linkage studies using mouse models start with the selection of two mouse models that naturally differ in the trait(s) of interest, like blood glucose and body weight.

To identify the genes variants that underlie these quantitative metabolic traits, an outcross population with the inbreed strains is generally generated by repetitive intercrosses and/or backcrosses (Figure 4). Backcrosses are commonly chosen as the generated progeny have therefore only two possible genotypes (Flint and Eskin 2012; Schwenk, et al. 2013).

By using the backcross system, two chosen parental mice (A and B) are crossed for the generation of the filial F1 progeny. In the sequence, F_1 individuals are backcrossed with one of the parental mice generating thereby a backcross population that is designated N2. These approach enable the generation of a huge number of animals that represent a genetic mosaic of the parental mice. Furthermore, genotyping markers can be used to determine which of the parental alleles have been inherited throughout the whole genome in a distance of ideally 10-20 centiMorgans (cM), which represents a unit of recombination frequency used to estimate a genetic distance (Broman 2001; Broman 2009). A final goal consists of applying the known linkage analysis to reveal possible quantitative trait loci (QTL), which is defined as a genomic region or locus that cosegregate with a given phenotype. A numerical degree of association is thereby determined by the calculation of logarithm of the odds (LOD) (Broman 2001; Su, et al. 2010).

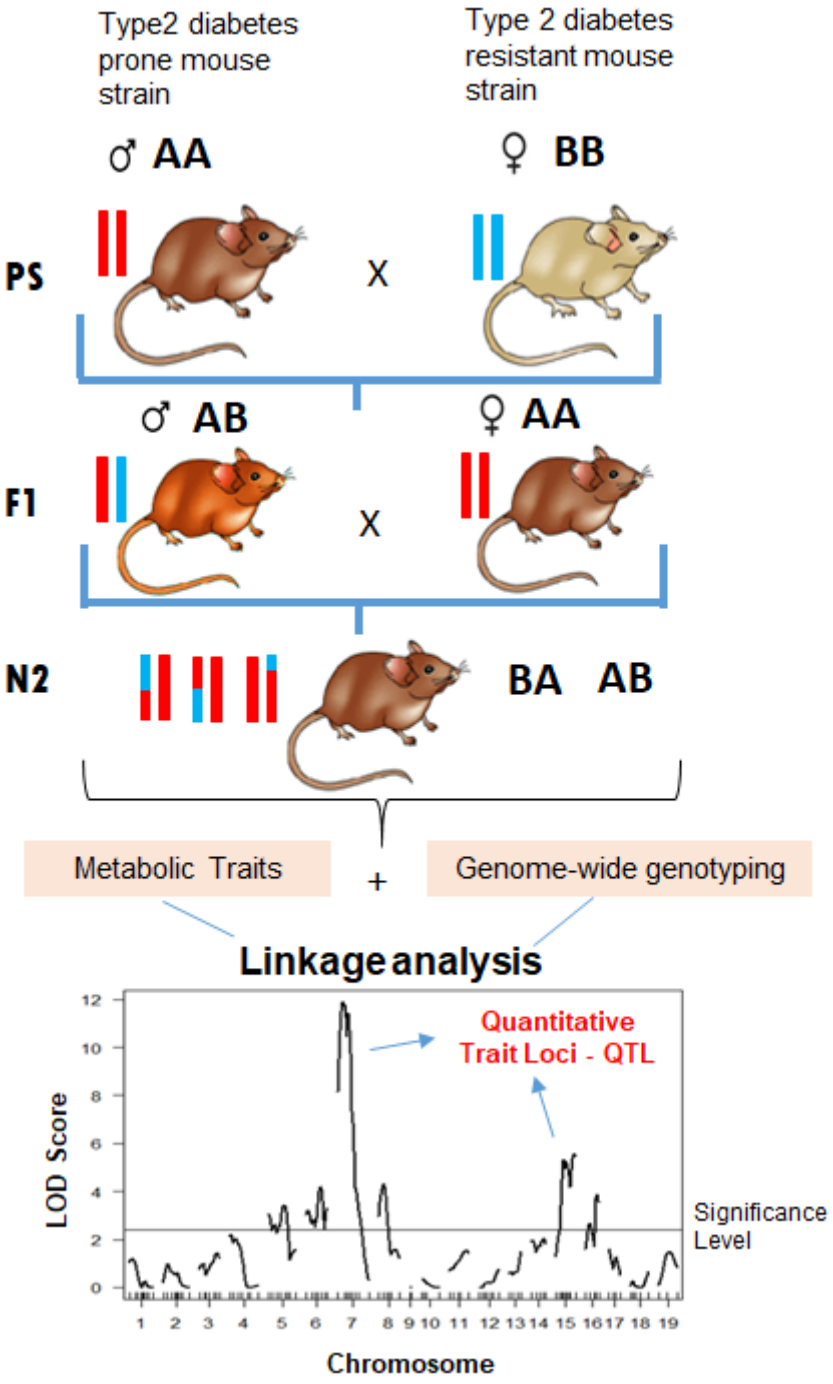


Figure 4. Design of mouse Genome-wide Linkage Study. In a first step two inbreed mouse strains that diverge in certain phenotype are chosen (AA and BB). Furthermore, an outcross population is generated by a first crossbreeding (F1) followed by a backcross (N2), resulting in a generation of mice that represent a genetic mosaic of the parental mice. In the sequence, metabolic quantitative traits of the generated N2 population are combined with genome-wide genetic markers using *in silico* linkage analysis to predict and determine a Quantitative trait loci (QTL). A QTL represents a signal of significant association between metabolic traits and genotypes, indicated by a logarithm of the odds (LOD) score. Source: Modified from (Appleby and Ramsdell 2003).

Conversely, the identification of a QTL alone does not reveal the underlying gene variants that result in the phenotypic variations. In addition, the identification of QTL for common metabolic disorders with low penetrance have lower mapping

resolution and comprehend therefore large genomic regions that contain usually hundreds of genes. In addition, the pathogenesis of a given phenotype is not revealed by the determination of linked genomic regions. Therefore, further approaches are commonly necessary to be addressed to identify the causal gene variant(s) for a QTL as well as the pathophysiological mechanism mediated by the gene(s). One of this approaches comprises the generation of congenic strains. This approach follows the transferring or introgression of given positive or negative QTL-chromosomal segment from one mouse strain onto the genome of the other strain showing an opposite phenotype. This is accomplished by repetitive backcrosses from which individual genotypes can be judiciously selected for QTL-haplotypes and therefore speed the number of generations required to reduce further background fragments. Further, an intercross of heterozygous mice allows the introgression or generation of homozygous QTL-haplotypes blocks from the strain AA onto the strain BB. In the sequence, a phenotypic characterization can be performed which ultimately permits the segmentation of important phenotypic-associated gene variants from other QTL-neighboring genes (Brockmann 2012; Nadeau 2000).

1.10 Strategies for the identification of QTL causative gene variants

The generation of congenic strains enable the dissection of the QTL fragments that harbor the genes underlying a QTL effect. This strategy can therefore substantially narrow down the candidate genes. However, additional strategies can be applied to further select the genes with the highest likelihood to be functionally causal.

1.10.1 Investigation of gene expression analysis

Once a genomic region associated to a certain trait has been identified, the investigation of mRNA expression levels of genes located in the critical QTL region can be applied. This strategy is based on the assumption that the two parental inbred strains show differential gene expression in the candidate gene-of-interest. In this matter, the use of microarray chips facilitates the analysis of a variety of genes in different tissues. The use of microarrays already contributed to the discovery of several diabetes-associated genes in many cross studies (Lowe and Reddy 2015). Moreover, studies conducted in human populations also revealed several loci associated to gene expression regulation. In a genomic study performed in an European and an African

cohort, approximately 83 % of the analyzed genes were differentially expressed among individuals of these populations (Gerber and Rutter 2017; Storey, et al. 2007).

In addition, associations between genetic markers and gene expression have facilitated the search for QTL causal gene variants. For this purpose, gene expression levels are used as a quantitative trait to identify expression quantitative trait loci (eQTL). The identification of a gene eQTL uncovers the locus that explains the genetic variance of the gene expression and represents an useful tool when candidate genes cannot be directly linked to a trait-associated disease or do not match the exact genomic region that is in linkage disequilibrium (Lowe and Reddy 2015). A practical example of this phenomenon is the genomic region linked to body mass index where the *FTO* gene has been found (Fawcett and Barroso 2010). Despite of the fact that *FTO* is highly associated to BMI and matches to the locus of strongest association, a later study using eQTL analysis demonstrated that the gene iroquois-related homeobox 3 (*IRX3*) located in the proximity genomic region also explains the pathophysiological mechanisms of the body mass index variations (Ronkainen, et al. 2015; Smemo, et al. 2014).

Moreover, another study among European and Asian populations found 1097 genes out of 4197 were differentially expressed and the phenotypic variation associated to this genes were mostly found to be *cis*-regulators (Spielman, et al. 2007). According to the genomic regulatory region, an eQTL can empirically be subdivided into *cis* and *trans* eQTL. If the identified transcript regulatory region lies within the gene location or in the close proximity, the respective eQTL is termed *cis*-eQTL. On the other hand, *trans*-eQTL refers to a genomic region that lies far downstream or upstream or even in another chromosome. Therefore, the identification of a *cis*-eQTL is generally associated to a stronger indication of regulatory control and therefore often used as criterion to narrow down candidate genes underlying a given QTL. This is also attributed to the fact that the identification of *trans* regulatory regions lack statistical confidence (Lowe and Reddy 2015; Nica and Dermitzakis 2013; Schadt, et al. 2008; Wray 2007).

1.10.2 Identification of strain-dependent polymorphic genes and *in silico* predictions of functional impact

Although the gene expression analysis has a great importance in the identification of new disease susceptibility genes, several studies have demonstrated

that the origin of gene expression can reside in different gene polymorphisms among distinct populations. Thus, the identification of gene variants is of great importance to detect the underlying mechanisms that account for the metabolic changes associated to a given QTL. Especially when considering linkage studies of inbred mouse lines, the identification of allelic variations between different strains can substantially improve the identification of QTL causative genes (Lowe and Reddy 2015; Spielman et al. 2007).

The identification of polymorphic gene variants takes into consideration that inbred strains have a common ancestor and that gene polymorphisms are therefore likely inherited across the generations (Figure 5). Hence, progenitor-specific polymorphic regions that contain non-identical by descendent polymorphisms are potentially harboring gene variants that could explain phenotypic variations (Browning and Browning 2011; McLaren, et al. 2016).

The development of high-throughput DNA sequencing technologies have enabled the completion of the mouse genome sequencing that facilitated the generation of maps of informative genomic markers or single nucleotide polymorphisms (SNP) among several inbred strains. However, the identification of polymorphic SNPs among inbred strains is not per se a strong indication that the resulted gene variants have divergent functional features or a resulted disease associated functional impact. This is because polymorphic SNPs can code for synonymous amino acid exchanges or lead to amino acid substitutions with similar properties.

To overcome this challenge, *in silico* prediction of functional impact can be applied before further experiments are performed. This tool is straightforward in selecting genes carrying non-synonymous amino acids exchanges that are predicted to influence the function of the transcribed and translated protein and/or interfere in the binding of transcription factors that regulate gene expression before transcriptional events take place. In summary, combining the identification of strain-specific SNPs and the predictions of functional impact bring up the strongest evidence to further select candidate genes as QTL causative targets (Lowe and Reddy 2015; Spielman et al. 2007).

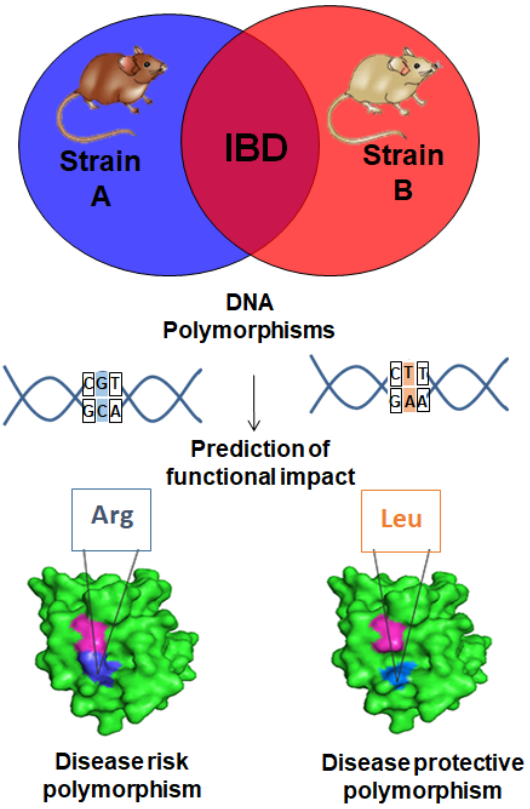


Figure 5. Identification of Identical regions by descendent (IBD). The identification of IBD regions is a straightforward approach to uncover QTL causative genes. This method is based on the assumption that mouse inbred strains share a common ancestor and thereby share genomic regions with higher identity. These are therefore regions that do not contain gene polymorphisms and consequently less likely to harbor QTL causative genes. Once single nucleotide polymorphisms (SNP) are identified, further *in silico* analysis can be applied to predict the impact of respective polymorphisms in a gene or protein and associate those with metabolic features. Source Modified from (RADHAKRISHNAN 2015).

1.11 Identification of Quantitative Trait Loci in the C3HxNZO crossbreeding

Using the aforementioned approaches for the identification of new T2D susceptibility loci, an outcross population was generated by crossbreeding the obese NZO with the lean C3H mouse strain (PhD of Tanja Schallschmidt, 2018; Schallschmidt, 2018) (Figure 5). The generated filial F1 generation was backcrossed with the parental NZO in order to generate a genetically heterogeneous N2(C3HxNZO) backcross population. Thus, the generated N2 mice presented a genetic mosaic of the parental mouse lines, which was confirmed by a genotyping scan using a genome-wide informative SNP panel.

Furthermore, N2 mice underwent deep phenotypic characterization for a diversity of metabolic traits, e.g. blood glucose, body weight, body fat and lean mass and plasma insulin levels. In the sequence, an *in silico* linkage analysis was performed which revealed 60 loci associated to several diabetes associated traits. Three major blood

glucose quantitative trait loci (QTL) were identified on chromosome 4, 7 and 15. The mice carrying the C3H alleles on chromosome 15 and 7 were widely protected from T2D. Conversely, on chromosome 4, the NZO alleles addressed the T2D protection effect. Furthermore, the combination of the three protective loci resulted in additive effect on blood glucose levels. The mice carrying the three protective alleles had an average of 238 mg/dl lower blood glucose levels compared to mice carrying the three risk alleles (Schallschmidt et al. 2018).

Aforementioned approaches were used for the identification of causative gene variants underlying the newly detected QTL. Recombinant congenic strains (RCS) generated by the introgression of the protective C3H allele on chromosome 7 into the diabetes-susceptible NZO background strain revealed significantly lower blood glucose levels and elevated levels of plasma insulin in the C3H allele carriers group. Further gene expression profiling combined with the haplotype analysis revealed three genes as major candidates underlying the diabetes protective effect of the locus on chromosome 7, termed *Nbg7* (NZO blood glucose on chromosome 7). Further experiments conducted in insulinoma cells (Min6) associated *Pop4* and *Atp4a* to the phenotypic effects observed in the recombinant congenic lines. The knockdown of these genes in Min6 cells also corroborate with the results obtained in the mouse lines leading to alterations on glucose-stimulated insulin secretion (GSIS) (Schallschmidt, 2018).

A locus located distally (119 Mb) on chromosome 4, termed *Nbg4* (NZO blood glucose on chromosome 4), accounted for differences in blood glucose that reached 109 mg/dl. Similarly, the novel T2D associated locus on chromosome 15 (LOD6.6), termed *Nbg15* (NZO blood glucose on chromosome 15) spans a region of 30 Mb around 13 cM peak (Figure 6). The genotype segregation of the N2(C3HxNZO) revealed that he *Nbg15* carriers of the C3H alleles displayed 100 mg/dl lower blood glucose in comparison to the carriers of the NZO alleles at *Nbg15*. However, the physiological interplay between different tissues and the genetic regulation underlying these effects on both *Nbg15* and *Nbg4* remain to be determined.

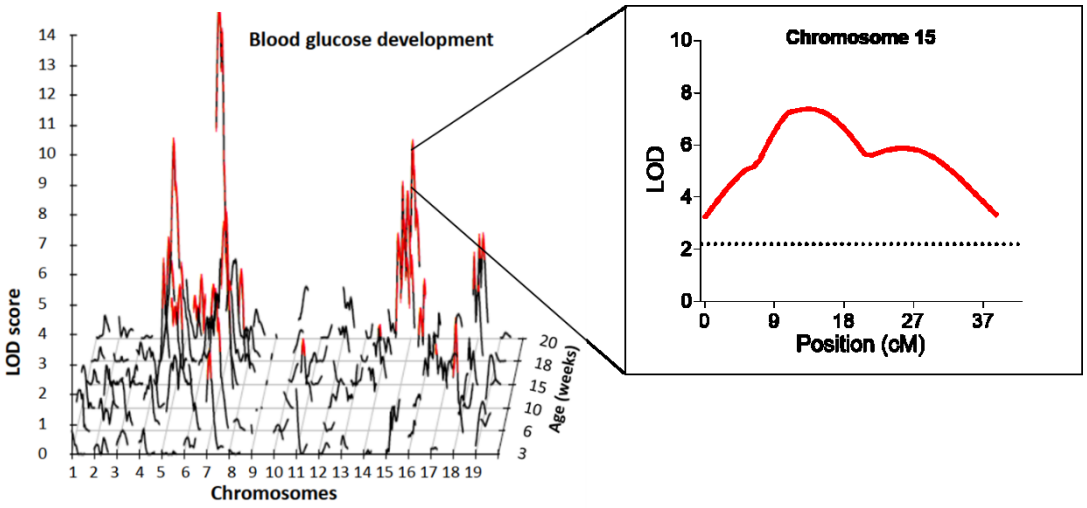


Figure 6. Identification of Quantitative Trait Loci in the N2(C3HxNZO) outcross population. An N2(C3HxNZO) population was generated by breeding the diabetes prone NZO with the diabetes resistant C3H strain. Diabetes related quantitative traits of these mice were combined with their genotypes in the linkage analysis. The results revealed three major QTL on chromosome 4, 7 and 15 that regulate blood glucose levels. Significant QTL are depicted in red and QTL on chromosome 15, focus of this thesis, magnified on the right. Adapted from: Tanja Schallschmidt PhD thesis, 2018.

1.12 Aim of the project

A previous linkage study performed in the research group with an outcross population using T2D resistant CH3 mice and the T2D susceptible NZO strain led to identification of a major QTL for blood glucose on chromosome 15 (*Nbg15*), contributing to different features of diabetes susceptibility (Figure 6) (Schallschmidt, 2018). The blood glucose levels as well as the diabetes prevalence were strikingly lower in the C3H alleles carrier of the *Nbg15* locus in comparison to the NZO allele carriers of the same locus in a body weight independent manner.

The aim of this study was to investigate the pathophysiological mechanisms and the genes underlying the T2D-proctetive *Nbg15* locus. For this purpose the following approached were applied:

- ✓ Investigation of additional genomic markers underlying the QTL on chromosome 15 to better map the contribution of the respective genomic regions in the observed phenotype.
- ✓ Application of positional cloning for the generation of mouse lines carrying different *Nbg15*-C3H-haploblock into the diabetes prone NZO genome or into the diabetes protective C57BL/6J genome.
- ✓ Gene expression profiling in the parental NZO and C3H mouse lines, followed by the determination of gene expression in the N2(C3HxNZO) and application of expression QTL analysis.

- ✓ Investigation of polymorphic genes variants and *in silico* predictions of functional impact for the identification of *Nbg15* associated candidate genes.

2 MATERIAL AND METHODS

2.1 Materials

2.1.1 Mouse strains and Diets

The parental mouse lines used for the generation of the recombinant congenic lines are listed in Table 1.

Table 1. Mouse strains

Mouse strain	Supplier
NZO/Hl	Lieselotte Herberg (Herberg and Coleman, 1977), DDZ Düsseldorf
C3HeB/FeJ	Helmholtz Center Munich
C57BL/6J	Charles River Laboratories, Sulzfeld, Germany

The standard diet (SD) was purchased from Ssniff (Soest, Germany) and high-fat diets (HFD) were purchased from Research Diets Inc. (New Brunswick, NJ, USA). The composition of the diets is summarized in Table 2.

Table 2. List of mouse diets

Diet	Components (gm %/ kcal %)		
	Protein	Carbohydrate	Fat
Standard diet (#V1126 M-Z Extrudat)	22.1 /36.0	53.3 / 53	4.4 / 11.0
HFD 45 % calories from fat (#D12451)	24.0 / 20.0	41.0 / 35.0	24.0 / 45.0
HFD 60 % calories from fat (#D12492)	26.2 /20.0	26.3 / 20.0	34.9 / 60.0

2.1.2 Primers

Hexanucleotide primers for cDNA synthesis were purchased from Roche (Mannheim, Germany). For quantitative Real-Time PCR experiments, SYBR Green primers were designed using the NCBI Primer Blast online tool (see references) and produced by Eurogentec (Liège, Belgium). The primer sequences are detailed in Table 3.

Table3. List of designed primer pairs

Gene	Locus ID		Primer sequence	Start	Stop	Exon junction
Apol7a	NM_029419.2	F	TACAATGGCACTAGGATTCCAGA	42	64	273/274
		R	CCAGGTGACCAGGAGTTTTCT	286	266	
Csf2rb2	NM_007781.3	F	ACCAGCTAGACAAGATTCAATCAG	496	519	508/509
		R	TGTGGATGTCCTTGGGAGGG	726	707	
Fam83f	NM_145986.2	F	GGGGTGGACAGACACCAACT	415	434	541/542
		R	AGCAATGACCTTCTGGGCCTG	550	530	
Apol6	NM_001163621.1	F	TGGCCCAGTTGGTGAATGTT	17	36	97/98
		R	CTGTCTGTCCTCCCTCTCCA	110	91	
Mchr1	NM_145132.2	F	ATGGATCTGCAAGCCTCGTT	276	295	357/358
		R	AGGTGGCCCCGCCAA	365	351	
Slc25a17	NM_011399.3	F	AAGGCCTCCTGGCACCATAC	211	230	217/218
		R	ACCTTTGACCCACACTGCTT	323	304	
Tst	NM_009437.4	F	GCCGGATATAGTAGGACTAGACT	608	630	621/622
		R	TCCTGGAATATGGCACGCAG	730	711	
Tmem184b	NM_001253820.1	F	GGCTCCTGAGCCTTTGTCATCTAT	46	69	55/56
		R	CAGCACGTGCCGTACATACA	173	154	
Slc30a8	NM_172816.3	F	CCAAAACTGGACAGCGCATC	273	292	404/405
		R	GGCTTCTGTCGAAGTTCTCTGT	422	401	
Has2	NM_008216.3	F	AAGCATTTAAGGCTCCCCCA	372	391	553/554
		R	CAATGCATCTTGTTCAGCTCCT	561	540	
Ly6d	NM_010742.1	F	GTAAGAACCCTCAGGTCTGCC	102	122	158/159
		R	TTCAGAGGCTCCACTGAGGTGA	171	150	
Adck5	NM_172960.3	F	ATGTTCTGGTGCGGAAAGGT	1218	1237	1299/1300
		R	AGGATCGGTCCTTCTCGTCC	1310	1291	
Lrrc24	NM_198119.2	F	CTGGGCTTTCCTCCTTAGCC	766	785	873/874
		R	GGGTTCTCTGTGAGGCGTAG	880	861	
Cdc42ep1	NM_027219.3	F	TCAACCAGGCCACCTATGAC	715	734	804/805
		R	AGACTCCAGGCCGTAACCAG	815	796	
Kdelr3	NM_134090.2	F	GGCGGGCGGGACCAT	130	144	334/335
		R	GAAAACCACCTTCATCACTGTGT	343	321	

2.1.3 SNP genotyping makers

All primers for genotyping using the KASP assay were purchased from LGC genomics (LGC group, Teddington, United Kingdom). The primers used are summarized in Table 4.

Table 4. Genotyping primers for chromosome 15

Primer	SNP ID	Position (mb)
15-005994001-M	rs3715343	6,12
15-033125499-M	rs3720676	32,963563
15-041919640-N	rs3090057	41,569377
15-049283927-M	rs3719583	48,898332
15-059192855-M	rs3653368	58,603788
15-063976267-N	rs3023419	63,349565
15-071973707-N	rs3088506	71,351481
15-074947367-M	rs3724474	74,316084
15-080641980-N	rs4230879	79,746551
15-087100507-M	rs3690689	86,131385
15-093195380-M	rs3707104	92,148538
15-098621793-N	rs3023427	97,542351
15-102788257-N	rs3023429	101,641254

2.1.4 Reaction Kits

2.1.4.1 Reaction Kits for molecular biology analysis

Table 5. Reaction kits molecular biology

Kit description	Application	Company
Invisorb® Genomic DNA Kit II	DNA isolation	STRATEC Molecular GmbH, Berlin, Germany
KASP V4.0 2X Master mix, high Rox /low Rox	KASP Genotyping	LGC group, Teddington, United Kingdom
RNeasy Mini Kit	RNA isolation	QIAGEN, Hilden, Germany
RNAse free DNAse Set	RNA isolation	QIAGEN, Hilden, Germany
GoScript® Reverse Transcription System	cDNA synthesis	Promega, Madison, USA
GoTaq® qPCR Master Mix	qPCR	Promega, Madison, USA

2.1.4.2 Reaction Kits for biochemical measurements

Table 6. Biochemical reaction kits

Kit	Application	Supplier
Cholesterol Liquicolor kit	Plasma cholesterol determination	Human Diagnostics, Wiesbaden, Germany
Glucose Liquicolor kit	Plasma- and tissue glycogen determination	Human Diagnostics, Wiesbaden, Germany
Insulin ELISA Ultrasensitive (Mouse)	Plasma insulin determination	DRG, Marburg, Germany
NEFA –HR (2) Assay Kit	Plasma free fatty acid determination	Wako Chemicals, Neuss, Germany
BCA Protein Assay kit	Total protein determination	Pierce Rockford, IL, USA
Triglycerides (TRIGS) GPO PAP kit	Plasma- and tissue triglyceride determination	RANDOX Laboratories Ltd. (Ardmore, UK)
Bio-Plex Pro Mouse Cytokine 23-plex assay Kit	Plasma cytokines determination	Biorad Laboratories, Düsseldorf, Germany

2.1.5 Chemical reagents

Table 7. List of chemicals reagents

Chemical / Solution	Supplier
Chloroform (pure)	AppliChem, Darmstadt, Germany
Deoxynucleoside triphosphates (dNTPs)	Promega, Madison, USA
D-Glucose (C6H12O6)	Sigma-Aldrich, Steinheim, Germany
Ethanol (absolute, pure)	AppliChem, Darmstadt, Germany
Hexanucleotide primer	Roche, Mannheim, Germany
Insulin Actrapid® HM Penfill®	Novo Nordisk Pharma GmbH, Mainz, Germany
Isoflurane	Piramal Healthcare, Morpeth, UK
Isopropanol (≥ 99. 5 %)	AppliChem, Darmstadt, Germany
Potassium hydroxide (KOH)	Merck, Darmstadt, Germany
QIAzol® reagent	QIAGEN, Hilden, Germany
Random primer	Roche Diagnostics, Mannheim, Germany
Potassium hydroxide (KOH)	Merck, Darmstadt, Germany
QIAzol® reagent	QIAGEN, Hilden, Germany
Random primer	Roche Diagnostics, Mannheim, Germany
Sodium chloride (NaCl)	Roth, Karlsruhe, Germany
Sterile glucose 20 %	Braun (Melsungen, Germany)
TRIzol® reagent	Thermo Fisher Scientific, Darmstadt, Germany

Acetic acid	Merck, Darmstadt, Germany
Aniline	Merck, Darmstadt, Germany
Aniline blue	Merck, Darmstadt, Germany
Eosin yellowish	Merck, Darmstadt, Germany
Ethanol absolute (EMSURE ®)	Merck, Darmstadt, Germany
Hematoxylin	Merck, Darmstadt, Germany
Isopropanol (≥ 99.5 %)	AppliChem, Darmstadt, Germany
Paraformaldehyde (PFA) extra pure	Merck, Darmstadt, Germany
Shandon EZ-Mount	Thermo Fisher Scientific, Darmstadt, Germany
Sodium chloride (NaCl)	Roth, Karlsruhe, Germany
Sodium iodate	AppliChem, Darmstadt, Germany
Xylene	AppliChem, Darmstadt, Germany
Wax (Paraplast®PLUS)	Roth, Karlsruhe, Germany

2.1.6 Expendable materials

Table 8. List of expendable materials

Material	Supplier
Cellstar Multiwell Plate	Oehmen, Essen, Germany
Combitips advanced	Eppendorf, Hamburg, Germany
CRYO tubes 1.8 mL	Thermo Fisher Scientific, Darmstadt, Germany
Disposable scalpels # 12 +22	BBraun, Melsungen, Germany
Glucose test stripes Contour Next	Bayer HealthCare, Leverkusen, Germany
MaXtract High Density	QIAGEN, Hilden, Germany
MicroAmp® Fast Optical 96-well reaction plates	Applied Biosystems, Foster City, USA
MicroAmp® Optical adhesive film	Applied Biosystems, Foster City, USA
Microvette CB 300 µL Lithium-Heparin	Sarstedt, Nümbrecht, Germany
PCR-Plates (FrameStar® 384-Well)	4titude, Berlin, Germany
QIAshreder	QIAGEN, Hilden, Germany
Rotilabo microtest plates (96 well)	Carl Roth, Karlsruhe, Germany
Rotilabo®-aluminium foils	Carl Roth, Karlsruhe, Germany
Stainless steel beads (5 mm)	Qiagen, Hilden, Germany
Sterican® needle G24x1" / ø 0,55 x 25 mm	BBraun, Melsungen, Germany
Sterican® needle G26x1" / ø 0,45 x 25 mm	BBraun, Melsungen, Germany
Syringe Omnifix® 1 mL	BBraun, Melsungen, Germany

2.1.7 Devices

Table 9. List of devices

Device	Supplier
Agilent 2100 Bioanalyzer	Agilent Technologies, Waldbronn, Germany
Centrifuge GPKR	Beckman Coulter, Krefeld, Germany
Compact centrifuge SPROUTTM	Biozym Scientific, Hessisch Oldendorf, Germany
Cooling table centrifuge 5425 R	Eppendorf, Hamburg, Germany
Electronic scale	Sartorius, Göttingen, Germany
FeedTime	TSE-Systems, Bad Homburg, Germany
IKA® Vortex 4 basic	IKA®-Werke, Staufen, Germany
iMark™ Microplate reader	Bio-Rad, Munich, Germany
Infinite® 200 PRO Series Multimode Reader	Tecan Trading AG, Männedorf, Switzerland
Mastercycler	Eppendorf, Hamburg, Germany
Multipipette® E3	Eppendorf, Hamburg, Germany
NanoDrop 2000	Thermo Fisher Scientific, Darmstadt, Germany
NMR	EchoMRI, Houston, USA
PCR Plate Spinner	VWR, Radnor, USA
QuantStudio 7 Flex	Applied Biosystems, Foster City, USA
StepOnePlusTM System	Applied Biosystems, Foster City, USA
Thermomixer ® comfort	Eppendorf, Hamburg, Germany
TissueLyser II	Qiagen, Hilden, Germany
Uniprep-Gyrator	UniEquip, Planegg, Germany
Shandon™Excelsior™ Tissue Processor	Thermo Fisher Scientific, Schwerte, Germany
Rotary microtome HM360	Thermo Microm, Schwerte, Germany
Camera Olympus DP73	Olympus Corporation, Tokyo, Japan

2.2 Methods

2.2.1 Animal Experiments

2.2.1.1 Animals housing

All mouse experiments were approved by the Ethics Committee of the State Agency for Nature, Environment and Consumer Protection (LANUV State of North Rhine-Westphalia, Germany; reference numbers: 84-02.04.2013.A118 and 84-02.04.2015.A354) and performed accordingly to the guidelines of the National Institutes of Health (NIH). A maximal of 6 animals per cage (Macrolon type III) were housed in acclimatized rooms (22 °C) and under a 12 h light-dark cycle (lights on at 6 a.m.) with *ad libitum* access to food and water. Mice were weaned at the age of 3 weeks and either maintained on standard chow diet for further breeding or subjected to a high-fat diet (HFD, 45 or 60 % Kcal fat) intervention.

2.2.1.2 Diets

Parental C3HeB/FeJ, NZO/Hl and C57BL/6J were subjected to 45 Kcal% fat diet for 3 weeks and 18 weeks after weaning. Recombinant congenic lines generated in a NZO/Hl background (NZO.C3H-*Nbg15)* were subjected to the same diet for 12 weeks after weaning while the lines generated in a C57BL/6J background (B6.C3H-*Nbg15)* received HFD 60% Kcal fat (Table2).

2.2.1.3 Genotyping

2.2.1.3.1 DNA isolation for genotyping

Mouse tail tips were used for DNA isolation. Approximately 2 mm of tail tip biopsy was collected after mice weaning for subsequent isolation of genomic DNA using Invisorb® Genomic DNA Kit II. The method included a first enzymatic digestion step which allowed the inactivation of DNAses, followed by binding of DNA to high-affinity silica particles. The concentration of the DNA was measured using a Spectrophotometer Nanodrop 2000 device at 260 nm (NanoDrop ND-2000, PEQLAB Biotechnology, Erlangen, Germany).

2.2.1.3.2 Genotyping - Kompetitive Allele Specific PCR (KASP)

For the generation of RCS, all mice were genotyped using the KASP (Kompetitive Allele Specific PCR) method developed by the company LGC Genomic (LGC group, Teddington, United Kingdom). The mice from backcrosses breedings were genotyped for the whole genome using an informative Single polynucleotide polymorphisms (SNP) panel from LGC. Mice from backcross generation N4 to N6 and intercross generations were genotyped using SNPs only for chromosome 15. In total, 14 SNP markers, in a distance of maximal 20 Mb were designed for genotyping of the lines carrying *Nbg15* from C3HeB/FeJ into the NZO/Hl background. For the introgression of *Nbg15* from C3HeB/FeJ into the C57BL/6J background 10 SNP markers covering full chromosome 15 were used (Table 4).

The KASP assay specifically identifies SNP and therefore is able to differentiate the genomic regions between the mouse strains. The KASP master mix contains two forward primers containing the allele-specific target sequence that is bound to a non-complementary sequence referred as tail sequence. Each primer binds

to its specific DNA target. In a first PCR round, DNA is elongated and the tail sequence remains bound. In a second PCR step the tail sequence is therefore elongated. The presence of the elongated tail sequence in the final PCR product allows the binding of the Fret cassettes, designed to match these specific tails. One of the Fret cassettes is labelled with FAM™ dye (520 nm emission) and the other with HEX™ (556 nm emission) or VIC™ dye (554 nm emission). Therefore, as unquenched, the fluorescence emission occurs and permits the identification of the different alleles and as a consequence the genotype of the mouse (Figure 7).

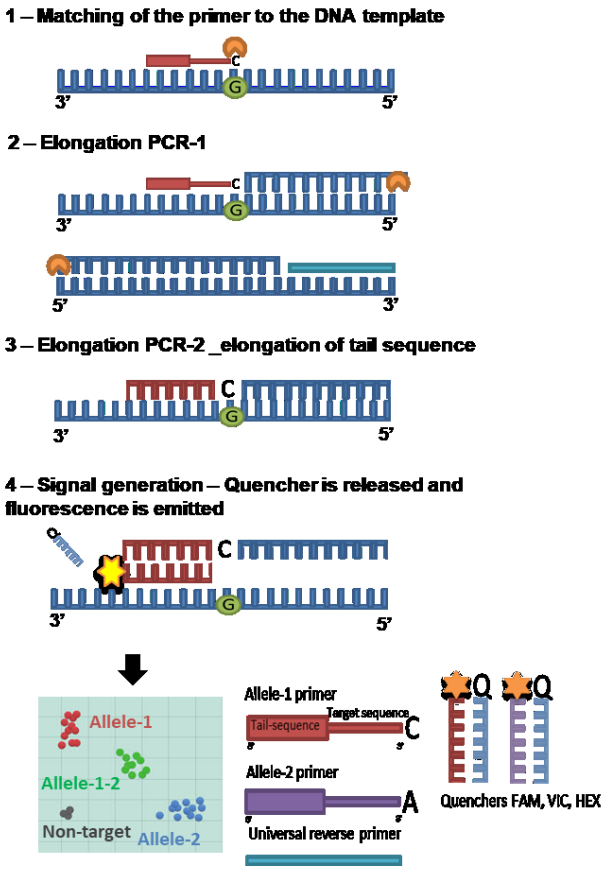


Figure 7. Genotyping – Kompetitive Allele Specific PCR (KASP). KASP genotyping reagents is composed by two allele specific primers and two Fret cassetes bound to different dyes that match to the tail sequence of the primers. This enable the release of the quencher from the Fret cassettes and emission of fluorescence that results in a color code that identifies the genotype accordingly. Modified from https://www.biosearchtech.com, 20.07.2019.

2.2.1.4 Generation of Recombinant Congenic Strains (RCS)

2.2.1.4.1 Generation of backcross N_2(NZOxC3H) population

NZO/HI females were bred with C3HeB/FeJ males for the generation of the F1 generation. In the sequence, a backcross generation (N2(NZOxC3H)) was bred by

crossing males from F1 generation with parental NZO/Hl females (PhD thesis Tanja Schallschmidt).

2.2.1.4.2 Generation of RCS lines carrying the *Nbg15* locus from C3HeB/FeJ

The generation of RCS carrying different fragments of the *Nbg15* locus was carried out following the previously performed generation of an N2 population from a backcross of F1(NZOxC3H) males on NZO females or from a new breeding of C3HeB/FeJ males with C57BL/6J females (PhD thesis Tanja Schallschmidt, 2018). Males from the generated F1 generation were backcrossed with parental females NZO/Hl or C57BL/6J to generate the N_2(NZOxC3H) and N_2(B6XC3H) respectively. Subsequently, backcrosses were performed for another 3-4 generations in order to increase the background of the parental mice. Genome-wide genotyping was performed (KASP assay, Table 4) to judiciously select mice carrying the heterozygous C3H alleles only on chromosome 15. Applying this method backcross generations N5 and N6 presented already a background of 95 % of the parental mice. Subsequently, animals carrying the heterozygous C3H alleles on chromosome 15 were intercrossed in order to generate homozygous carriers of the C3H allele for chromosome 15 on either NZO or C57BL/6J background. Following this procedure, in total 4 lines were generated carrying the C3H alleles onto the NZO background. The first line consisted of mice carrying the full chromosome 15 from C3H (NZO.C3H-*Nbg15*$^{C3H/C3H}$). The second and third line, respectively, carried either the proximal or the distal LOD score peak of the *Nbg15* locus (NZO.C3H-*Nbg15p*$^{C3H/C3H}$, NZO.C3H-*Nbg15d*$^{C3H/C3H}$). A fourth line carrying the NZO alleles (NZO.C3H-Nbg15$^{NZO/NZO}$) at *Nbg15* locus was generated and used as control. Furthermore, a line carrying the full chromosome 15 from C3H onto the B6 (B6.C3H-*Nbg15*$^{C3H/C3H}$) background was generated and compared to a line homozygous for the *Nbg15* B6 allele (B6.C3H-*Nbg15*$^{B6/B6}$).

2.2.1.5 Metabolic characterization of the RCS mice

2.2.1.5.1 Measurements of blood glucose levels

Random blood glucose was measured weekly (from 8 a.m. to 10 a.m.) using a Contour XT glucometer (Bayer, Germany). The total blood was collected by pricking the mouse tail vein. A diabetes threshold was defined to random blood glucose levels exceeding 300 mg/dl for 3 weeks in a row.

2.2.1.5.2 Measurements of body weight and body composition

Body weight was accessed weekly using electronic scales (Table 9). Body composition was determined using a nuclear magnetic resonance (NMR) spectrometer (NMR; Whole Body Composition Analyzer, Echo MRI, Houston, TX, USA).

2.2.1.5.3 Fasting and refeeding

For the fasting/refeeding experiments, mice were fasted for 16 h overnight (O/N) and refed for 2 h *ad libitum* in order to determine hormones and circulating metabolites. Total tail blood was collected both in the fasting and the refed state in a micro tube coated with calcium-disodium EDTA and centrifuged at 9000 x g for 5 min and 4 °C for insulin measurements and 2348 x g at room temperature for FFA and triglycerides measurements. In the sequence, the plasma was collected in a micro tube and stored at -80°C until measurements were conducted.

2.2.1.5.4 Intraperitoneal glucose tolerance test (i.p. GTT)

For the i.p.GTT, mouse lines generated in a C57BL/6J background were fasted for 6h (from 8 a.m. to 2 p.m.) whereas mouse lines on a NZO background were fasted for 16h (overnight). After fasting, a small portion of the tail was cut to assess tail vein blood for basal measurement of blood glucose using a glucometer (Contour XT Bayer, Germany) (time point 0). Next, mice received an intraperitoneal injection of a 20 % sterile glucose solution (2 g/kg body weight). Blood glucose was assessed at time points 15, 30, 60, 120 and 240 min after glucose injection. Furthermore, during all time points, tail blood was collected in a micro tube coated with Lithium-Heparin by gently massaging the tail. In the sequence the blood was centrifuged at 9000 x g for 5 min and 4 °C for the subsequent determination of plasma insulin levels using ELISA kit (Table 6). Until the measurements, plasma was stored at -80 °C.

2.2.1.5.5 Intraperitoneal insulin tolerance test (i.p. ITT)

For the i.p.ITT, mouse lines generated in a NZO background were fasted for 6 h (from 8 a.m. to 2 p.m.) whereas mouse lines on a C57BL/6J background underwent i.p.ITT randomly fed. Blood glucose was subsequently measured by pricking the tail vein in the fasted state, before receiving an injection of insulin 1 U/kg body weight prepared in sterile saline (0.1U/ml)) (time point 0). After insulin injection, blood glucose

was measured at time points 15, 30, and 60 min using a Contour XT glucometer (Bayer, Germany).

2.2.1.5.6 Tissue collection

For end point measurements, mice were fasted in the morning for 6h (3 a.m to 9 a.m.). Blood glucose was measured in the fasted state using a glucometer Contour XT glucometer (Bayer, Germany). Moreover, blood was collected for subsequent measurements of insulin, triglycerides, cholesterol, FFA and inflammatory markers in the plasma (Table 6). Subsequently, mice were sacrificed by cervical dislocation and immediately after, gonadal white adipose tissue, subcutaneous adipose tissue, liver, spleen, pancreas, skeletal muscle (quadriceps) and brown adipose tissue were harvested, snap-frozen in liquid nitrogen and stored at -80 °C for later use.

2.2.1.5.7 Histological analysis of pancreatic tissue

The dissected pancreata were carefully placed in a cassette and immersed in 4 % paraformaldehyde (PFA) solution in PBS buffer (ph 7.4). Further, the pancreata were fixated in PFA for 24 h before embedding protocol. The embedding procedure was performed using the Shandon Excelsior Tissue Processor (Thermo Fisher Scientific, Schwerte, Germany). First, the dehydratation of the pancreata was performed by submerging the tissues for 1h in 70 % ethanol and increasing ethanol concentration to 96 % and 100 % thereafter. Final steps included immersion in Xylene for 3 h followed by 3h immersion in pre-warmed paraffin at 60 °C. The construction of cuttable blocks was obtained by immersion of the pancreas in wax (Table 7). A rotary microtome (Thermo Microm, Schwerte, Germany) was used to first trim the block until a representative pancreas profile was reached. Thereafter, a series of 15 consecutive 5 µm sections were selected before an additional level of 75 µm was discarded. This method permitted a better analysis of pancreatic islets located within each level as an estimated average size of an islet corresponds to 150 µm (defined by Cellular Core Unit), therefore matching the cut section together with discarded levels. For the determination of pancreatic islet size and density, 8 section per level were selected for hematoxylin vs Eosin staining. In the sequence, morphometric analysis of the pancreas was conducted which included the determination of the total pancreatic area and islets area.

2.2.1.5.7.1 Determination of pancreatic and islets area

For the determination of pancreatic area the point-counting method was used (Weibel 1969). For this purpose, an overhead projector was used to project a stained section onto a point grid. Thus, all points located within the pancreas were counted. Each point corresponded to one square milimiter considering a 28 mm distance between each point on the grid and the 28-fold magnification through the projection (established by the Cellular Core Unit).

For the quantification of pancreatic islet area the stained sections were scanned and every islet was imaged with a DP73 camera (Olympus, Tokyo, Japan). The determination of the area was calculated by using the Wacom Tablet and the polygon tool of the CellSens Dimension 1.9 software (Olympus) after manually surrounding the perimeters of each islet.

2.2.2 Biochemical measurements and molecular biology methods

2.2.2.1 Plasma Insulin levels

For insulin measurement the mouse ultrasensitive ELISA kit (Table 6) was used. All samples were single measured. The insulin ELISA is an enzyme-linked immunoassay detecting insulin in several biological samples based on its binding to a snit-insulin antibody coated in a 96 well plate. A secondary antibody horse radish-peroxidase (HRP)-coupled oxidizes 3, 3', 5, 5'-tetramethylbenzidine (TMB), a substrate that is added in the reaction tube and produces the color detected at 450 nm (iMark Microplate Absorbance Reader, Bio-Rad, Hercules, CA, USA) which intensity is proportional to the insulin concentration in each sample.

2.2.2.2 Measurement of plasma triglycerides, total cholesterol and plasma glucose

The plasma triglycerides measurements were performed using the colorimetric kit listed in Table 6. The method is based on a first hydrolysis reaction by lipases followed by oxidation of Glycerol-3-phosphate. This reaction results in the release of hydrogen peroxide that in presence of 4-aminoantipyrine phenol and peroxidase produces the indicator quinoneimine detected at 510 nm. For the reaction, 1.5 µl of each sample was incubated with 150 µl of reagent for 10 min at room temperature. For the total cholesterol determination similar colorimetric kit was used (Table 6), using

quinoneimine as final indicator. 5 µl of each samples was incubated with 200 µl of reagent and read at 500 nm after 20 min incubation at room temperature. For the determination of plasma glucose, 2 µl of samples were added to the reaction well and incubated with the kit reagent (Table 5) for 5 min at 37°C before determination in the reader (iMark Microplate Absorbance Reader, Bio-Rad, Hercules, CA, USA) at 505 nm. The results were expressed as mg/dl.

2.2.2.3 Free fatty acids (NEFA)

For the determination of plasma NEFA the NEFA-HR (2) Assay Kit (Table 6) was used, which consists in the formation of Acyl-CoA as intermediate reagent that is further oxidized for the formation of the violet complex color measured at 560 nm. For the measurements, the plasma samples were diluted 1:2 and 5 µl was incubated for 15 min at RT with 100 µl of the first reagent followed by incubation with 50 µl of the second reagent during 15 min at RT. The results were expressed as mmol/l.

2.2.2.4 Multiplexing cytokines panel

The measurements of plasma cytokines was conducted using the Bio-Plex Pro Mouse Cytokine 23-plex assay Kit (Table 6). The presence of fluorescently dye beads are covalently bound to desired biomarkers. After several washing steps the unbound targets are removed and a biotinylated antibody is added. The final addition of streptavidin-phycoerythrin conjugate produce the specific indicator for each biomarker. The results were expressed as pg/ml.

2.2.3 Molecular biological methods

2.2.3.1 RNA isolation from mouse tissues

The RNA isolation from frozen adipose tissues (WAT, GWAT and BAT) and islets was performed using the RNA extraction Kit miReasy-Mini Kit (Table 5).Using the referred kit, around 100mg of tissue was transferred to 2ml Safelock Eppendorf tube with a 5 mm Stainless Steel Bead and 700 µl of QIAzol lysis reagent (Table 7). The tissue was further lysed for 5 min at 25 1/s frequency using the TissueLyser II (Qiagen, Hilden, Germany). After 5 min incubation at room temperature (RT) the lysate was transferred to a spin column (QIAshredder, Qiagen, Hilden, Germany) and

centrifuge at 18000 x g for 2 minutes at 4 °C. The flow-through was collected in a pre-centrifuged MAXtract High Density tubes (Qiagen, Hilden, Germany). In addition, 140 µl of chloroform was added and after inversion the tubes were additionally incubated at RT for 5 min, before centrifugation at 12000 x g for 15 minutes at 4 °C. The separated aqueous phase containing the RNA was collected in a new 1.5 ml RNAse free eppendorf tube. 1.5X isopropanol was added to precipitate the RNA. Furthermore, the mixed solution was transferred to the miRNeasy-Mini columns containing a membrane that retains the RNA. The following protocol included a DNAse digestion step RNase-free DNAse Set (Table 4) prepared according to manufacture's instructions. After, the RNA was washed once with 500 µl of RWT buffer and twice with 350 µl of RPE buffer provided by the manufacture. In the last step, the miRNeasy-Mini columns were dried by centrifugation and the RNA was eluted with Nuclease free water in a new RNAse free eppendorf tube for storage at -80°C.

2.2.3.2 Determination of RNA concentration and cDNA synthesis

The RNA concentration was measured using the NanoDrop ND-2000, (PEQLAB Biotechnology, Erlangen, Germany). The concentration was photometrically determined at 260 nm. Furthermore, the isolated RNA was used for the synthesis into complementary DNA (cDNA). For this procedure, 1 or 2 µg of RNA was pipetted into an eppendorf tube where 1 µl of dNTPs and 2 µl of pre-diluted (1:10) hexanucleotide primers was added (Table 7). The samples were further spun down before pre-incubation at 65 °C for 5 minutes in the thermocycler (T100 Thermal Cycler, Bio-Rad, Hercules, CA, USA) followed by another 1 min incubation at 4 °C to permit primer annealing. Subsequently, 7 µl of master mix containing the reverse transcriptase and buffers according to manufactures were added to the reaction tube (Table 5). After spun down, the samples were placed in the thermocycler using the following protocol.

Table 10. cDNA synthesis protocol

Reaction	Temperature	Duration
Annealing	25 °C	5 minutes
Synthesis	42 °C	60 minutes
Reverse transcriptase inactivation	70 °C	15 minutes
Hold	+4 °C	

After completion, the resulting cDNA was diluted in nuclease free water at a ratio of 1:20 for 1 µg RNA input and 1:40 for 2 µg RNA for further gene expression analysis by qPCR.

2.2.3.3 Quantitative Real-Time PCR (qPCR)

For gene expression analysis quantitative Real-Time PCR (qPCR) was performed. The qPCR method enables the quantification of transcripts abundance based on fluorescence detection. Forward and reverse primers specifically target the gene of interest and the presence of reverse transcriptase enzyme permits the exponential amplification of target DNA in every PCR cycle. The quantification is achieved by adding special DNA binding dyes to the reaction tube, generally SYBR Green. SYBR Green in solution only displays low fluorescence and upon DNA binding the fluorescence is increased up to 1000-fold, allowing therefore the precise quantification of amplified DNA over the PCR cycles. During the amplification, the fluorescence increases proportionally to the formation of the new transcript copies. Therefore the determination of a threshold at this point enable the calculation of a value (Ct) that corresponds to DNA or relative mRNA abundance.

The experimental setup started with the addition of 4 µl of pre-diluted cDNA (see item 2.2.3.2) mixed with 5 µl of GoTaq qPCR Master Mix (Table 5) and 0.5 µl of pre-diluted (1:10 in Nuclease free water) forward and reverse SYBR Green primers designed for each specific target (Table 3). The reaction was performed in 384 well plate (Table 7) densely covered with MicroAmp adhesive films (Thermo Scientific, Waltham, MA, USA). In the sequence, the plate was shortly centrifuged and placed in the QuantStudio 7 Flex Real-Time PCR system device (Applied Biosystems, Foster City, USA) following the reaction protocol described as follow (Table 10).

Table 11. qPCR protocol

Reaction	Temperature	Duration	Cycles
Hot Start	95 °C	2 minutes	1x
Denaturation	95 °C	15 seconds	40x
Annealing/ Extension	60 °C	60 seconds	
Dissociation	60-95 °C	90 seconds	1x

For the quantification of transcripts abundance, the threshold of exponential phase was set as 0.2 and resulted calculated Ct values were exported and further normalized to a reporter gene (housekeeping gene-ΔCt) followed by normalization to a reference experimental group (2ΔCt). Final results were therefore expressed as relative normalized gene expression (2-ΔΔCt).

2.2.3.4 Microarray analysis

RNA samples from Liver, skeletal muscle (quadriceps), gonadal white adipose tissue, brown adipose tissue and pancreatic islets of the parental C3HeB/FeJ, NZO/HI and C57BL/6J were collected by Tanja Schallschmidt for Microarray analysis. The RNA quality was controlled using the Agilent 2100 Bioanalyzer (Table 9). The microarray analysis was performed by Dr. Birgit Knebel (Institute for Clinical Biochemistry and Pathobiochemistry, German Diabetes Center, Düsseldorf) using the Affymetrix-Chip (GeneChip® Mouse Genome 430A 2.0 Array). Furthermore, the results were further analyzed by Dr. Axel Rasche (Max Planck Institute for Molecular Genetics, Berlin).

2.2.4 *In silico* analysis

2.2.4.1 Haplotype analysis

The free available DNA sequence of parental C3HeB/FeJ, NZO/HI and C57BL/6J mice was obtained from the Mouse Genome Project of the Welcome Trust Sanger Institute Database (https://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1303), release GRCm38 1303. The closely related C3HeB/HeJ substrain was used instead of C3HeB/FeJ, not available in the database. C3H non-IBD region were determined by comparing its DNA sequence with the NZO sequence. Further, the number of SNPs between C3H and NZO was counted for an interval of 150 base pairs. For the determination of total SNPs number, the reference sequence from C57BL/6J was used and total annotated SNPs counted for the same abovementioned interval.

2.2.4.2 Variant effect predictor analysis (VEP)

Following the haplotype analysis, the identification number (rs number) of every C3H-non IBD SNPs was filtered for the QTL peak on chromosome 15 (50-80

mb) and used as input for prediction of functional impact using VEP analysis. The VEP is a free available tool from Ensembl that performs annotation and analysis of variations in the DNA in coding and non-coding regions and predict the impact on transcripts and protein level. For the analysis, web VEP interface was used (https://www.ensembl.org/Tools/VEP). Predictions were made by comparing the selected SNP-IDs to the reference C57BL/6J (GRCm38.p6 assembly) using Ensembl/GENCODE and RefSeq transcripts data bases with standard filters options (McLaren et al. 2016).

2.2.5 Statistical analysis

All results in this thesis are represented as mean ± SEM (Standart error of the mean). The number of biological replicates (n) is described in the figures legend. Significant differences were reported by two-tailed student´s t-test or one/ two-way analysis of variance (ANOVA) followed by post hoc Sidak's test. For all statistical tests applied, significant results were considered when $p \leq 0.05$. Moreover, p values between 0.05-0.08 were sometimes highlighted to indicate a tendency for an effect. Specific statistical test applied for every individual experiment are also indicated in figures legend. All tests were performed using the GraphPad Prism 7 software.

2.2.5.1 Quantitative trait loci (QTL) and expression eQTL analysis

The QTL analysis was performed with the data generated in the doctorate of Tanja Schallschmidt and combined with a new input of genome-wide genotyping markers. For the QTL analysis, the R software (version i386 3.3.2) with installed R/qtl 1.40-8 package was used (Broman 2009). Phenotyping traits were tested for normality distribution using D'Agostino-Pearson omnibus test and logarithmized (log2) in case of no fulfil of normality criteria. Furthermore, a genome map was generated with AntMap 1.1 software which indicated a calculated genomic distance between genotyping markers represented by the unit centimorgan (cM). In addition, the genotypes and phenotype traits of the N_2(NZOxC3H) outcross population were combined in an excel file and converted to .csv format before running the QTL analysis. The R/qtl package is a free available method developed to identify association between genotypes and phenotypes eliciting therefore the identification of genomic regions containing genes that are linked to disease susceptibility across an experimental mouse cross for example. Associations were determined by interval mapping using

expectation-maximization algorithm (EM), Haley-Knott regression and multiple imputation. Single-QTL scans were graphically represented by the Logarithm of Odds (LOD) score and the significance threshold indicated under the null hypothesis calculated using 100 permutations test. Furthermore, for the expression QTL (eQTL) analysis, relative gene expression levels (2-ΔCT) measured in different tissues of the N_2(NZOxC3H) population were used as phenotypic trait and combined with genome wide scans for the identification of the genomic regions regulating the target gene expression using a non-parametric QTL analysis (Lander and Kruglyak 1995; Lander and Botstein 1989).

2.2.5.2 Linear regression analysis

Expression mRNA levels of target genes performed in the N_2(NZOxC3H) population were correlated to metabolic traits using linear regression with GraphPad Prism 7.0. Significant correlations were indicated when $p<0.05$ and by R-squared (R^2) values.

3 RESULTS

This session is divided in three main chapters. The first chapter presents the results of the refined analysis of the T2D-linkage on chromosome 15. The second chapter refers to the phenotypic characterization of the *Nbg15* locus using the recombinant congenic mouse lines RCS.NZO.C3H-*Nbg15* and RCS.B6.C3H-*Nbg15*. In the third part, the results of the investigation of *Nbg15* associated genes are presented, which include the gene expression profiling and the *in silico* haplotype and variant effect predictor analysis.

3.1 Identification of *Nbg15* as diabetes protective locus on chromosome 15

A genomic region contributing to blood glucose regulation was previously identified on chromosome 15 in a backcross population of the diabetes-prone NZO with the diabetes resistant C3H mouse model N_2(C3HxNZO) (PhD thesis Tanja Schallschmidt, 2018; Schallschmidt, 2018). Within this backcross population, mice carrying the C3H alleles on chromosome 15 displayed lower blood glucose levels than NZO allele carriers. Therefore, in order to further dissect the genomic regions contributing to blood glucose regulation, additional new genomic markers were analyzed and mapped to the previously identified QTL.

The updated linkage analysis revealed that a large genomic region on chromosome 15 is predicted to significantly contribute to blood glucose homeostasis. Thus, a critical region was defined by statistically calculating the confidence interval for the observed QTL. As a result, the interval of 14 to 26 cM (50-80 mb) was determined as the confidence interval and, as a consequence, most likely to contain the causal genes for the identified QTL. This locus was thereby named as *Nbg15* (NZO blood glucose on chromosome 15). *Nbg15* consists two major loci or LOD score peaks that are statistically contributing to blood glucose regulation to a comparable extend. A locus located on proximal chromosome 15, therefore named "*Ng15-P*" was identified at 17 cM (*rs3023419,* 63Mb) with a LOD score value of 6.9, while a distal locus, or "*Nbg15-D"* was identified at 24 cM (*rs4230879*, 79Mb) with a LOD score value of 5.9 (Figure 8).

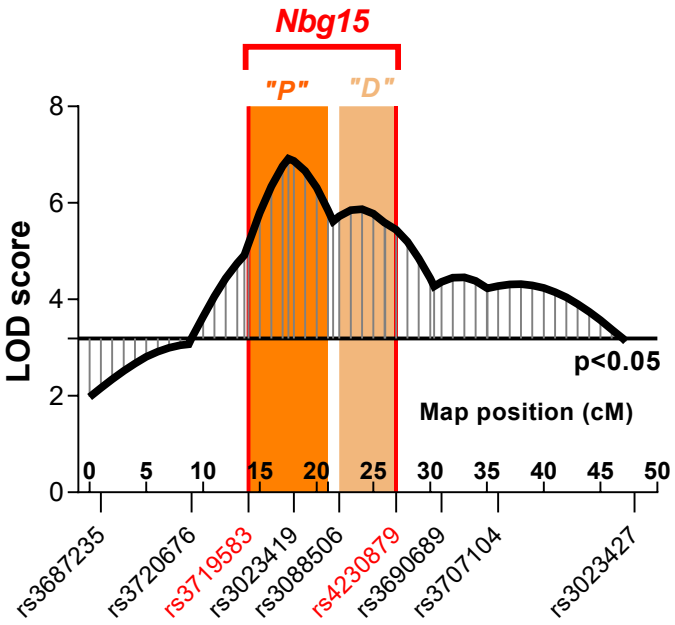


Figure 8. Identified QTL for blood glucose levels on chromosome 15. Calculation of linkage between genotypes and blood glucose levels of N2 males from the N_2(NZOxC3H) backcross population revealed a major QTL with two defined LOD score peaks within a critical region named *Nbg15*. For the determination of the updated QTL, 9 genotyping markers were used indicated in the bottom (rs-number). Permutation test (100 permutations) was used for the calculation of statistical significance, *P*: proximal; *D*: distal.

The segregation of the whole N2 backcross population based on the proximal peak (chromosome 15 C3H/NZO-P) conferred lower blood glucose levels in heterozygous C3H allele carriers compared to NZO allele carriers, reaching a maximal effect size of 99.56 mg/dl at 15 weeks of age (Figure 9A). Furthermore, the segregation based on the distal locus (*Nbg15-D*) revealed that heterozygous C3H allele carriers at the peak marker position (chromosome 15 C3H/NZO-P) displayed lower blood glucose levels reaching a maximal effect size of 85.26 mg/dl compared to the NZO-allele carriers at the same position (chromosome 15 NZO/NZO) (Figure 9B). For both subloci, blood glucose levels were significantly lower in C3H allele carriers already at 7 weeks of age, reaching its maximal effect at 15 weeks of age.

Moreover, N2 mice carrying the *Nbg15* C3H alleles either at the proximal or the distal peak markers were less susceptible to develop T2D. At final stages (21 weeks of age), 49.65 % of mice carrying the C3H allele at *Nbg15-P* developed T2D, whereas T2D incidence reached 77.25 % in mice carrying the NZO allele at the same position (Figure 9C). Mice carrying the *Nbg15-D* C3H allele demonstrated a diabetes prevalence of 52.48 % in 21 weeks of age compared to 76.34 % of the NZO allele carriers at this locus (Figure 9D).

Similarly, the survival rate was improved in *Nbg15* C3H allele carriers. While 97.26 % of C3H/NZO-P carriers (Figure 9E) and 96.96 % of C3H/NZO-D carriers (Figure 9F) survived the HFD intervention period of 21 weeks, the survival rate was 84.8 % and 85.1 % in the chromosome 15 NZO/NZO carriers, respectively.

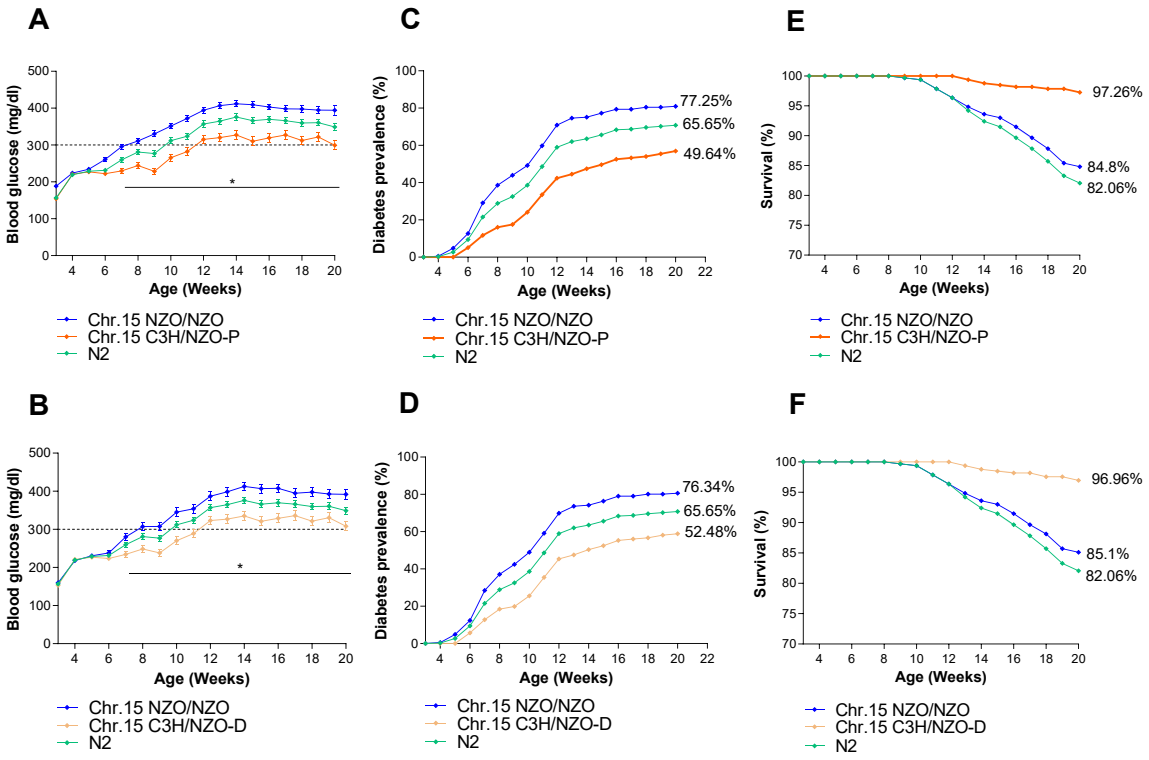


Figure 9. Blood glucose effect size of males from the N2(NZO(C3HxNZO) backcross population. Mice were segregated into heterozygous C3H or NZO allele carriers for *Nbg15* based on a marker located at the proximal peak (*rs3023419*, 63Mb) (A) and the distal peak (*rs4230879,*79Mb) (B). *p<0.05, 100 permutation test, n=140-189.

3.1.1 Effects of *Nbg15* locus linked to plasma insulin levels

Previous linkage analysis (PhD thesis Tanja Schallschmidt, 2018) did not detect a significant linkage between insulin levels and the *Nbg15* locus (Figure 10A). However, insulin levels were only assessed at 21 weeks of age. At that point, the mortality rate in *Nbg15* NZO-allele carriers reached 15 %, thereby decreasing the statistical strength of the linkage analysis, as well as the phenotypic effect, by naturally selecting the healthiest individuals. For this reason, a recalculation of the linkage analysis for plasma insulin was performed by deducing an average diabetic insulin level for dead mice at 21 weeks of age. Thus, mice with measured random blood glucose over 600 mg/dl at 21 weeks of age were selected from the whole N2 population for the calculation of average plasma insulin levels, as those were considered highly diabetic and represented the closest phenotype to the missing mice. The results

revealed that these mice had an average of 1.34 µg/L insulin (log2=0.42). By deducing this value for the missing mice, a significant linkage was detected on chromosome 15 matching the *Nbg15* locus position and the previous linkage analysis conducted with the available data (Figure 10A-B).

A segregation of the N2 backcross population based on the marker located in the proximal peak revealed that heterozygous C3H allele carriers exhibited higher plasma insulin levels compared to NZO allele carriers at this locus (chromosome 15 NZO/NZO 2.1 ± 0.15 µg/L; chromosome 15 C3H/NZO 2.7 ± 0.13 µg/L) (Figure 10C). A similar effect was observed for the distal locus, where C3H allele carriers displayed average plasma insulin levels of 2.7 ± 0.13 µg/L compared to 2.2 ± 0.16 µg/L in NZO allele carriers (Figure 10D). The average insulin levels for both the groups in Figure 10C and 10D are represented only with measured values, not considering the estimated insulin levels.

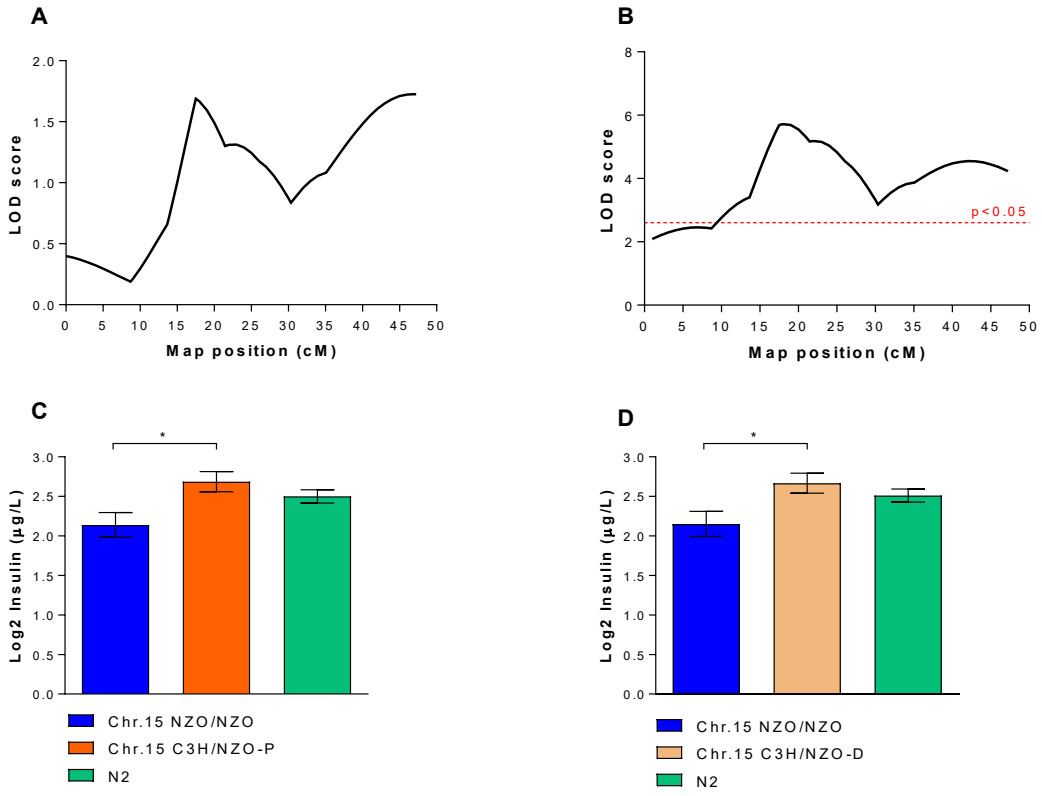


Figure 10. QTL for plasma insulin on chromosome 15. Identified QTL on chromosome 15 for plasma insulin (A) and re-calculation of QTL by deduction of missing data points (B). Effect size of insulin levels in N_2(C3HxNZO) backcross population. Male mice were segregated into heterozygous C3H or NZO allele carriers, respectively, for the *Nbg15* locus, based on the proximal (*rs3023419*) (A) and the distal peak (*rs4230879*) markers (B). *p<0.05, one-way ANOVA followed by Sidak's post-hoc test. n=140-189. Significant QTL was calculated using permutation test with 100 permutations.

3.1.2 Effects of *Nbg15* locus linked to body weight development

Interestingly, both the significant linkage of blood glucose as well as plasma insulin appear to be independent from the body weight development, since no significant linkage to body weight was detected on chromosome 15. The segregation of the N2 population using genotyping markers at proximal and distal peaks of *Nbg15* showed that mice carrying the C3H allele had a tendency to higher body weight compared to the NZO allele carriers, however this effect was not significant. (Figure 11A, 11B).

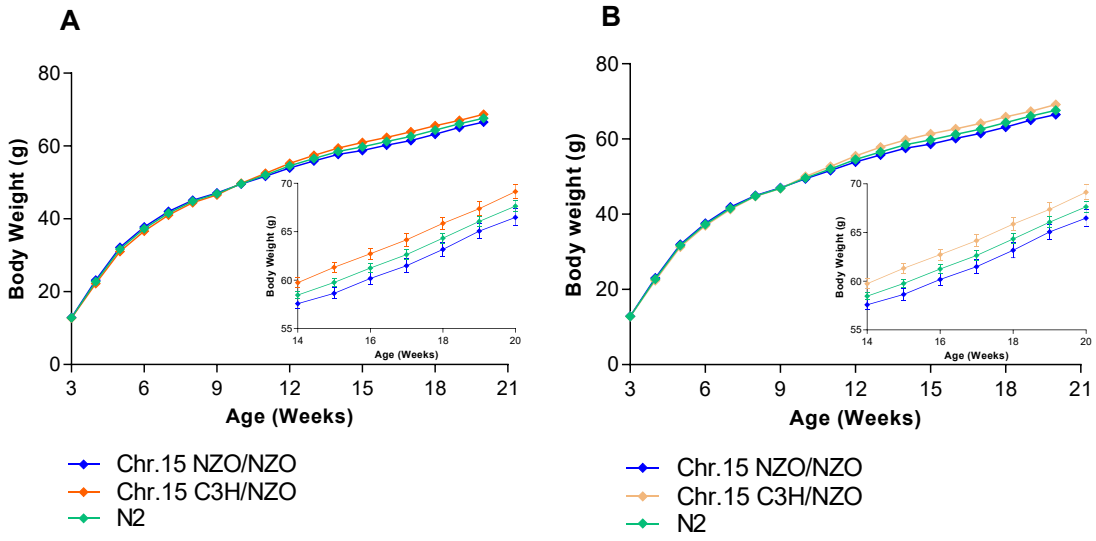


Figure 11. Body weight development of males from the N2(C3HxNZO) backcross population. Male mice were segregated into heterozygous C3H or NZO allele carriers for the *Nbg15* locus, based on the proximal (*rs3023419*) (A) and the distal peak (*rs4230879*) markers (B).

3.2 Development of strategies for the identification of causal gene variants underlying *Nbg15* locus

Considering that the N2 population was composed of mice carrying a high degree of genetic recombination of the parental C3H and NZO, further approaches were used to identify and characterize the pathophysiological mechanisms responsible for the observed effect on blood glucose levels within the N2 population driven by a genomic region located on chromosome 15. For this purpose, recombinant congenic lines were generated carrying different fragments of the C3H chromosome 15 into the NZO and B6 background.

Moreover, additional strategies were used to identify the candidate genes contributing to the blood glucose phenotype observed in the N2 population. Available genomic database (NCBI Genome Data viewer) showed that a total of 1620 genes are

annotated on murine chromosome 15. *Nbg15* harbors 503 annotated genes. Hence, two additional strategies were used to narrow down the number of potential candidate genes. Firstly, expression of all annotated genes within *Nbg15* were assessed in gonadal white adipose tissue, brown adipose tissue, liver, muscle and pancreatic islets of the parental C3H and NZO mice. Differentially expressed genes were further analyzed by quantitative real time PCR for validation purposes.

In the sequence, validated genes were further selected for the determination of transcripts abundance in the N2 backcross population for the corresponding tissues (except islets). Next, obtained gene expression data was used as a quantitative trait and therefore combined with available genotyping to perform the gene expression QTL analysis (eQTL). The results of the eQTL analysis revealed the genomic regions underlying the variation of gene expression. Genes whose genomic region matched to the locus where the gene is located (*cis*-QTL) were further selected as possible candidates for *Nbg15*.

Additionally, available information of Single Nucleotide Polymorphisms (SNP) obtained from Welcome Sanger Institute genome database was used for the identification of C3H polymorphic regions. SNPs located at polymorphic regions were further used as input for *in silico* calculation of protein impact. Genes containing C3H-specific SNP and likely to alter protein behavior were considered as the strongest candidates for *Nbg15* (Figure 12). The results of these approaches are presented in the following.

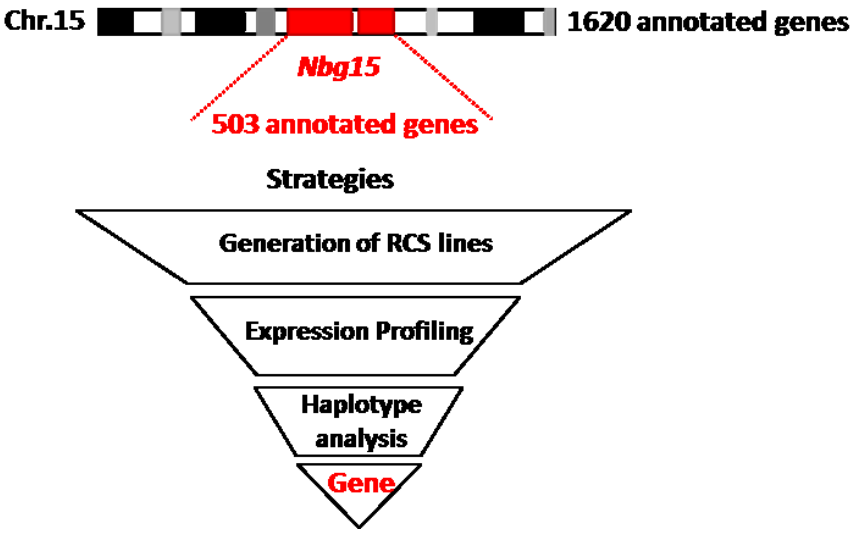


Figure 12. Strategies to dissect *Nbg15* causal genetic variants. *Nbg15* was identified as critical region harboring genes that contribute to C3H-diabetes protective phenotype in the N2 backcross population. Strategies to dissect the causal gene variant included the generation of Recombinant Congenic Strains (RCS) carrying different fragments of C3H genome into NZO and B6 background, followed by gene expression profiling of parental mice and identification of C3H polymorphic region combined with *in silico* predictions of protein impact. Source: NCBI genome viewer: www.ncbi.nlm.nih.gov/genome. Mouse assembly GRCm38.p6, 2019.

3.3 Metabolic characterization of Recombinant Congenic Strains RCS.NZO.C3H-*Nbg15* and RCS.B6.C3H-*Nbg15*

3.3.1 Generation of RCS.NZO.C3H-*Nbg15* and RCS.B6.C3H-*Nbg15*

For the introgression of *Nbg15* from C3H in the NZO and B6 genome, first the N_2(NZOxC3H) or N_2(B6xC3H) males were backcrossed with NZO and B6 females respectively, until the generation N5 and N6. SNP-based genome-wide genotyping enabled specific selection of mice containing maximal proportions of the respective recipient genome as genetic background for further breading in a so called “speed congenic” approach. Using this method, a theoretically calculated recipient background over 95 % can be reached at generation N5 and N6. For the generated backcross mouse lines, a recipient NZO and B6 background of 91.5 % and 86.5 %, respectively, was achieved already at backcross generation N3 (McBride, et al. 2004). Furthermore, repeated intercrosses were performed (brother-sister breeding) in the generation N5 and N6 to generate mice carrying the homozygous C3H alleles of *Nbg15* on NZO or B6 genomic background, respectively. These lines were consomic and therefore named as *Nbg15c*. Additionally, two RCS carrying either the proximal or the distal fragment of the *Nbg15* locus (*Nbg15p* and *Nbg15d*) were generated on the NZO

genomic background. Generations N5F3-4 N6F3 originated from the initial NZOxC3H crossbreeding and generation N6F3 originated from a parallel B6xC3H crossbreeding were chosen for metabolic characterization (Figure 13).

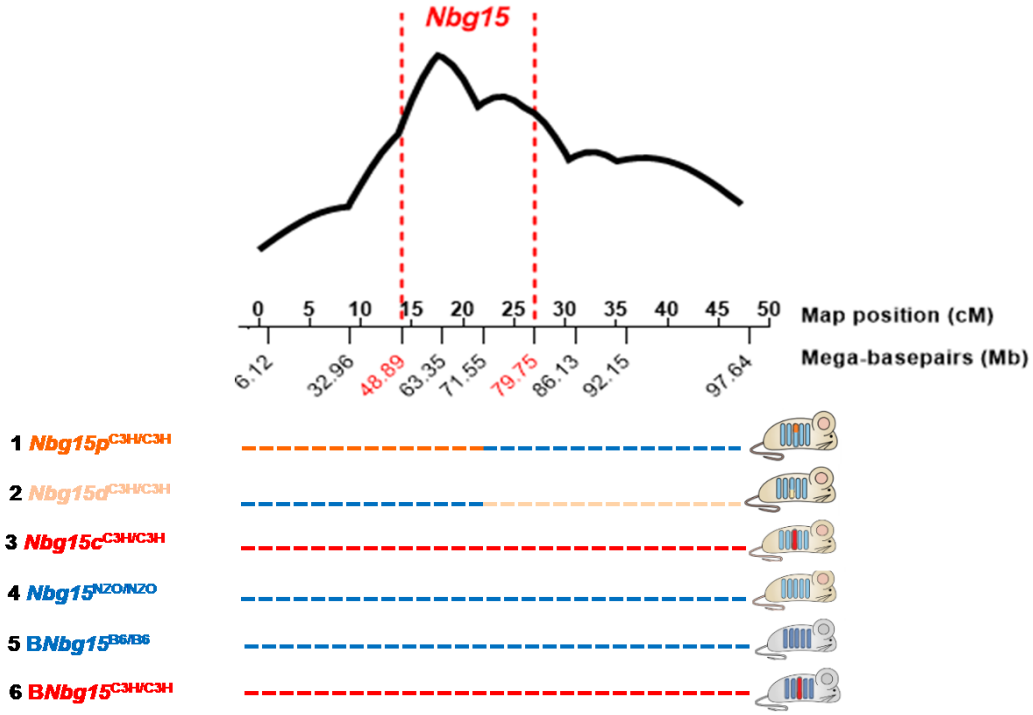


Figure 13. Generation of recombinant congenic strains. For the generation of RCS, N2 backcross generations from N_2(NZOxC3H) or N_2(B6xC3H) males were breed with the respective parental NZO and B6 females with the aim to increase the genetic recipient background of the parental mice. In the following, mice carrying heterozygous C3H alleles on chromosome 15 were further intercrossed for the generation of the lines carrying the homozygous C3H alleles for the different fragments of *Nbg15* locus. Thus, the generated *Nbg15p*$^{C3H/C3H}$ (1-orange dashed line) carried the homozygous C3H allele from position 6.12 to 71 Mb which comprehended the proximal region of *Nbg15*. *Nbg15d*$^{C3H/C3H}$ (2-beige dashed line) carried the C3H alleles at 71.55 to 97.64 Mb position harboring therefore the distal fragment of *Nbg15*. Consomic lines carrying the full chromosome 15 C3H alleles were generated in a NZO and B6 background (3, 6 –red dashed line), representing the lines carrying the complete *Nbg15 locus.* Two lines carrying the respective homozygous NZO and B6 (4, 5-blue dashed lines) alleles on chromosome 15 were also generated for comparison purposes.

3.3.2 Phenotyping Schedule

All experimental mice received a high-fat diet directly after weaning at the age of 3 weeks. The proportion of lipids in the respective diet was adapted to the genetic background strain, NZO.C3H.*Nbg15* mice received a diet containing 45 % kcal and B6.C3H.*Nbg15* a diet containing 60 % kcal fat. Furthermore, different phenotype schedules were performed for these lines as phenotypic differences were expected to appear at later time point for the B6.C3H.*Nbg15* line. Thus, NZO.C3H.*Nbg15* lines were phenotyped until week-of-life 15 while characterization of the B6.C3H.*Nbg15* line was finished at 21 weeks. Fasting/Refeeding, intraperitoneal glucose tolerance tests

and insulin tolerance tests were performed at different time points according to the schedule depicted in Figure 14.

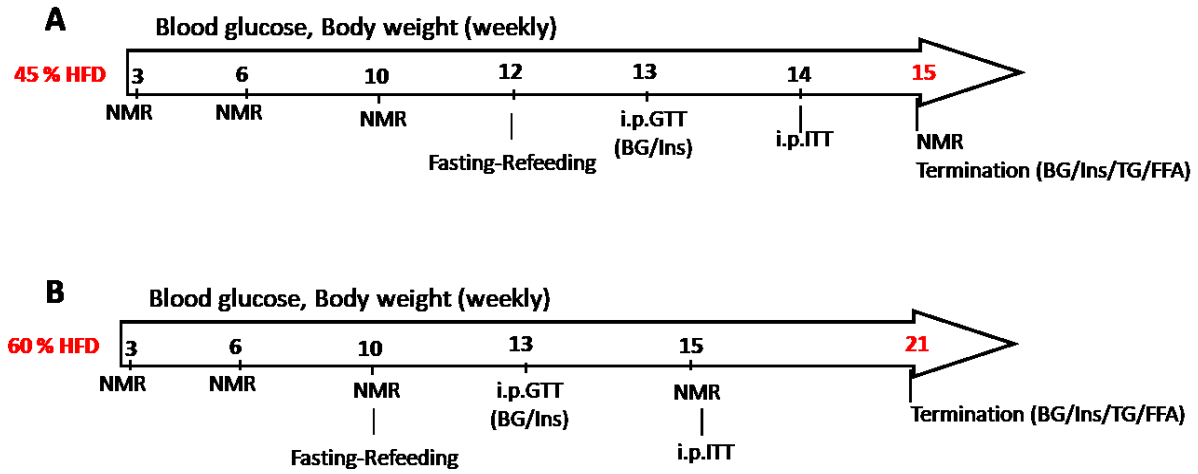


Figure 14. Phenotyping schedule of Recombinant Congenic Strains. Diet intervention was started at 3 weeks of age for both RCS.NZO.C3H-*Nbg15* (A) and RCS.B6.C3H-*Nbg15* (B). Body composition was assessed by nuclear magnetic resonance (NMR) spectroscopy at 3, 6, 10 and 15 weeks of age. Fasting-refeeding was conducted at 12 and 10 weeks of age for RCS.NZO.C3H-Nbg15 and RCS.B6.C3H-Nbg15 respectively. Intraperitoneal glucose tolerance tests (i.p.GTTs) were performed at 13 weeks for both genotypes. Intraperitoneal insulin tolerance tests (i.p.ITTs) were performed at 14 and 15 weeks of age for RCS.NZO.C3H-Nbg15 and RCS.B6.C3H-Nbg15, respectively, followed by sacrifice of the animals at 15 and 21 weeks of age.

3.3.3 Metabolic characterization of RCS.NZO.C3H-*Nbg15*

3.3.3.1 Measurements of random blood glucose levels, body weight and body composition

All lines carrying the homozygous C3H alleles at *Nbg15* displayed lower random blood glucose levels when compared to the NZO allele carriers. However, the strongest effect was observed for *Nbg15c*$^{C3H/C3H}$ mice. After 9 weeks of age, *Nbg15c*$^{C3H/C3H}$ mice exhibited significantly lower random blood glucose levels compared to *Nbg15*$^{NZO/NZO}$. No differences were observed at earlier stages (Figure 15A). The major effect was observed at 13 weeks of age when the size effect for blood glucose levels reached an average of 159 mg/dl (*Nbg15*$^{NZO/NZO}$ 401.8 ± 32.7 mg/dl, *Nbg15c*$^{C3H/C3H}$ 242.4 ± 22.5 mg/dl, $p<0.001$). *Nbg15p*$^{C3H/C3H}$ displayed a maximal lowering on blood glucose levels of 105 mg/dl at 12 weeks of age (*Nbg15*$^{NZO/NZO}$ 417.8 ± 30.7 mg/dl, *Nbg15p*$^{C3H/C3H}$ 312.7 ± 36 mg/dl, $p<0.01$). However, this effect diminished over time. Moreover, *Nbg15d*$^{C3H/C3H}$ presented a smaller effect size on blood glucose levels over time, but in contrast showed a strong drop of blood glucose levels at final

stages, reaching a size effect of 136 mg/dl at 14 weeks of age (*Nbg15*$^{NZO/NZO}$ 393.6 ± 34 mg/dl, *Nbg15d*$^{C3H/C3H}$ 256.9 ± 36.6 mg/dl, p<0.001) (Figure 15B).

Moreover, diabetes prevalence was strikingly lower for *Nbg15c*$^{C3H/C3H}$ animals compared to *Nbg15*$^{NZO/NZO}$ mice (*Nbg15*$^{NZO/NZO}$ 64.5%, *Nbg15c*$^{C3H/C3H}$ 22.7%). At 15 weeks of age, 20 out of 31 mice were diabetic within the *Nbg15*$^{NZO/NZO}$ cohort, while only 5 mice out of 22 reached the diabetes threshold in the *Nbg15c*$^{C3H/C3H}$ group. Similarly, mice carrying the *Nbg15p*$^{C3H/C3H}$ allele had a lower diabetes prevalence (40 %) compared to *Nbg15*$^{NZO/NZO}$ animals, only 8 out of 20 mice were classified as diabetic at 15 weeks of age. Interestingly, although mice carrying the *Nbg15d*$^{C3H/C3H}$ allele had lower random blood glucose levels, 60% or 9 out of 15 mice were diabetic at week 15 (Figure 15C). This effect resulted from a less severe hyperglycemia observed in *Nbg15d*$^{C3H/C3H}$ cohort. The results for the random blood glucose at 15 weeks of age showed that severe hyperglycemia was more frequently observed within the *Nbg15*$^{NZO/NZO}$ and the *Nbg15p*$^{C3H/C3H}$ genotype compared to *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice. While 11 out of 38 in the *Nbg15*$^{NZO/NZO}$ group and 8 out of 25 *Nbg15p*$^{C3H/C3H}$ mice reached random blood glucose levels over 500 mg/dl , only one mouse out of 24 for the *Nbg15c*$^{C3H/C3H}$ group and one out of 25 *Nbg15d*$^{C3H/C3H}$ were severely hyperglycemic (Figure 15D).

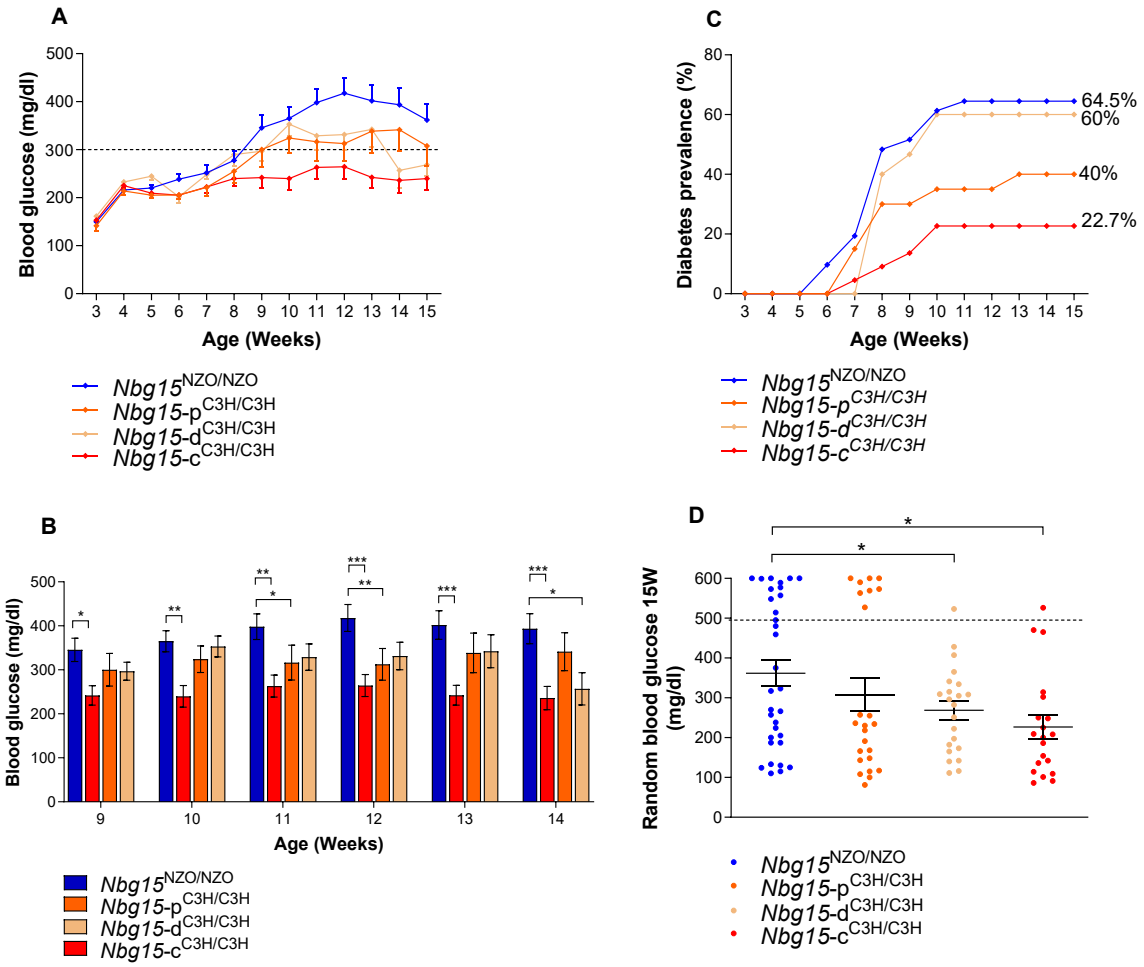


Figure 15. Blood glucose development in RCS.NZO.C3H-*Nbg15* mice. Random blood glucose development (A, B), diabetes prevalence (C) and final random blood glucose (D) in RCS.NZO.C3H-*Nbg15* genotypes on a 45 % kcal high-fat diet. Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =31-38 mice; *Nbg15*$^{C3H/C3H}$ = 15-25 mice. 1-way ANOVA (D) or 2-way ANOVA (A, B, C) followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

Body weight development was monitored on a weekly basis in the experimental cohorts. During the first weeks of high-fat-diet intervention, no differences in body weight were observed between the genotypes (Figure 16A). Nevertheless, at 12 weeks of age *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ exhibited higher body weight compared to *Nbg15*$^{NZO/NZO}$. This effect remained until final stages of age when the average difference in body weight reached 6.7 g and 7.8 g, respectively (*Nbg15*$^{NZO/NZO}$ 54.6 ± 1.17 g, *Nbg15c*$^{C3H/C3H}$ 61.3 ± 1.35 g, *Nbg15d*$^{C3H/C3H}$ 62.5 ± 1.5 g p<0.001) (Figure 16B). On the other hand, *Nbg15p*$^{C3H/C3H}$ mice exhibited a significantly lower body weight than the NZO allele carriers. This difference was observed only from 10 to 13 weeks of age, with a maximal effect size of 3.4 g in week 11 (*Nbg15*$^{NZO/NZO}$ 53.4 ± 0.7 g, *Nbg15p*$^{C3H/C3H}$ 50 ± 0.6 g, p<0.05) (Figure 16B).

Generally, no significant effects on gain of weight were observed during the first weeks of age (3 to 11 weeks of age) when comparing *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice with *Nbg15*$^{NZO/NZO}$ (Figure 16C). However, these genotypes significantly gain weight at later stages (12 to 15 weeks of age) when compared to the *Nbg15*$^{NZO/NZO}$ group. However, this effect could be a result of diabetes-driven loss of weight within the *Nbg15*$^{NZO/NZO}$ group as at this stage 64 % of the mice in this groups were severely diabetic (Figure 13C). In contrast, *Nbg15p*$^{C3H/C3H}$ mice gained significantly less weight compared to *Nbg15*$^{NZO/NZO}$ mice. This difference was observed from 5 to 13 weeks of age. At final stages, this effect was attenuated by loss of weight in the *Nbg15*$^{NZO/NZO}$ genotype (Figure 16D).

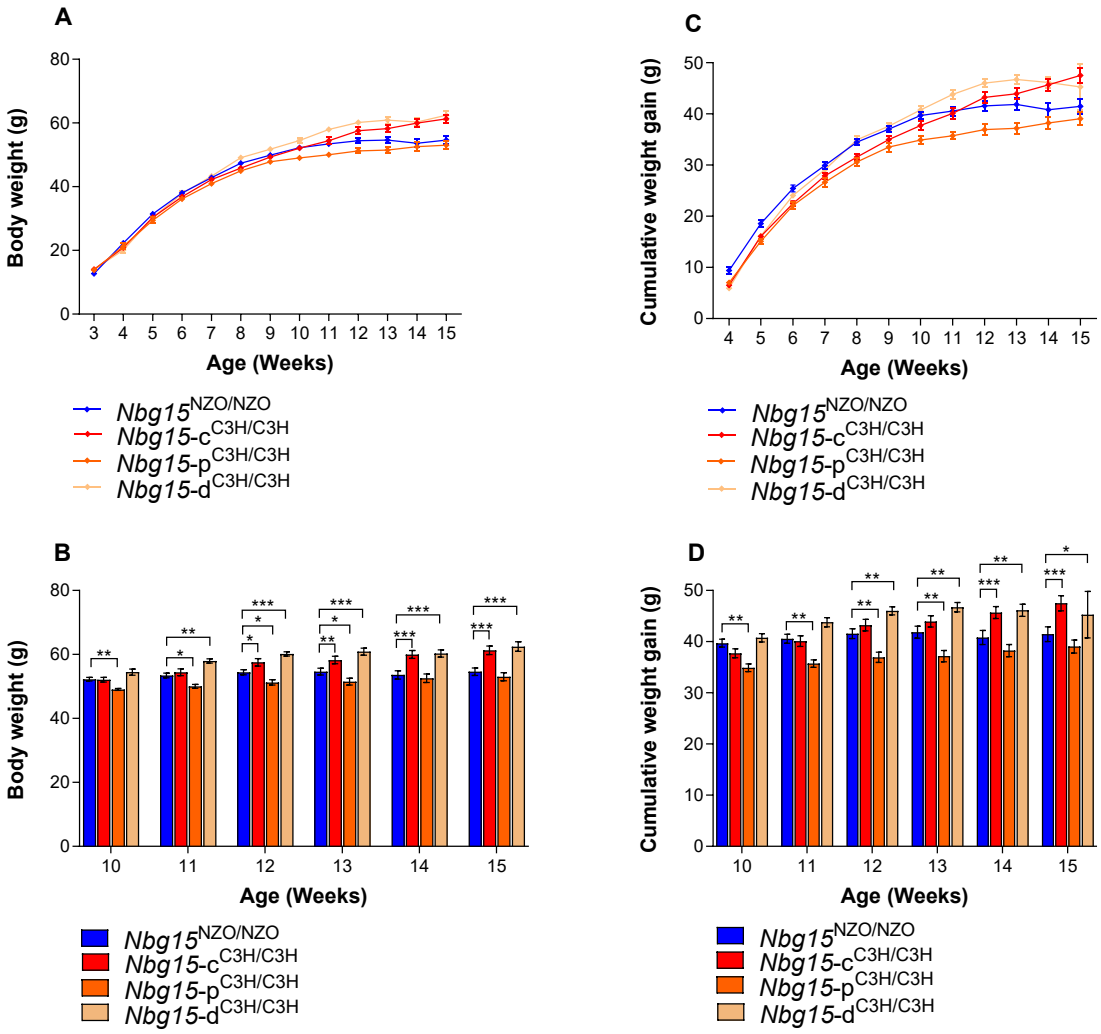


Figure 16. Body weight development in RCS.NZO.C3H-*Nbg15* mice. Body weight development (A, B), cumulative weight gain (C, D) in RCS NZO.C3H-*Nbg15* genotypes fed a 45 % kcal high-fat diet. Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =31-38 mice; *Nbg15*$^{C3H/C3H}$ = 15-25 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

The variations observed in body weight resulted mainly from differences in body fat content. At 15 weeks of age, *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice exhibited significantly higher fat content compared to the *Nbg15*$^{NZO/NZO}$ animals (*Nbg15*$^{NZO/NZO}$ 19.1 ± 1.2 g, *Nbg15c*$^{C3H/C3H}$ 26.3 ± 1.4 g, *Nbg15d*$^{C3H/C3H}$ 25.5 ± 1.4 g, $p<0.001$). On the other hand, no fat content differences were observed in *Nbg15p*$^{C3H/C3H}$ mice in comparison to *Nbg15*$^{NZO/NZO}$ (Figure 17A). As seen in the body weight development, the effects observed in body fat composition could also not be distinguished from those resulted from severe hyperglycemia-driven complications between the genotypes.

In contrast, the variations in body weight observed in *Nbg15p*$^{C3H/C3H}$ mice can be partially explained by the lower lean mass observed in this mice in comparison to *Nbg15*$^{NZO/NZO}$ genotype. *Nbg15p*$^{C3H/C3H}$ mice exhibited lower body lean mass compared to *Nbg15*$^{NZO/NZO}$ at already 6 weeks of age, effect that increased until the 15 weeks of age, when *Nbg15p*$^{C3H/C3H}$ mice had an average of 1.4 g lower lean mass than the Nbg15$^{NZO/NZO}$ mice (*Nbg15*$^{NZO/NZO}$ 33 ± 0.3 g, *Nbg15p*$^{C3H/C3H}$ 31,1 ± 0.31 g, $p<0.05$). Nonetheless, this represented a small effect accounting for only 3 % of total body mass in *Nbg15p*$^{C3H/C3H}$ mice (Figure 17B).

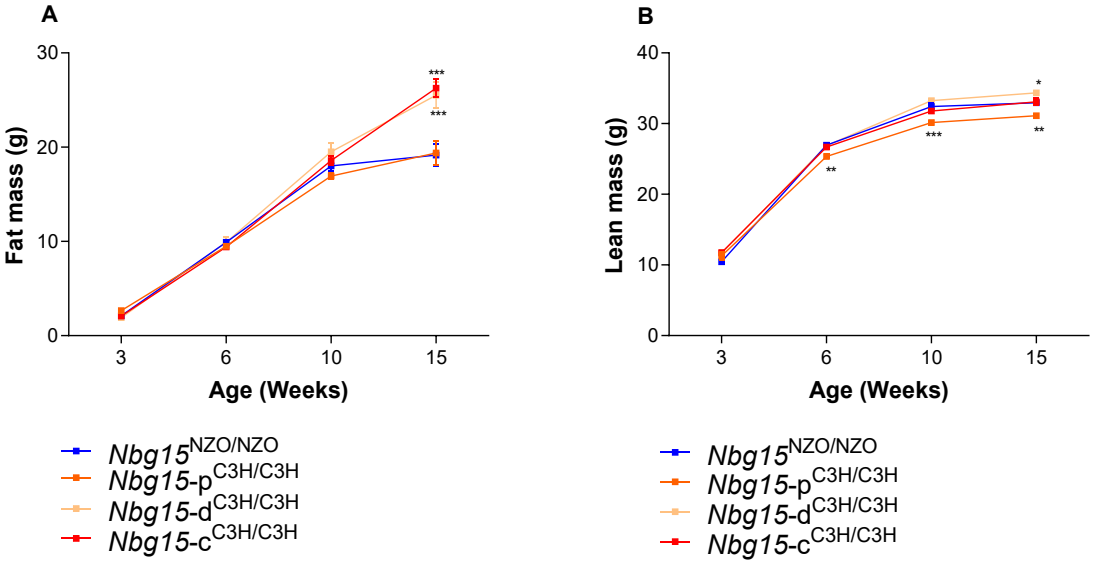


Figure 17. Body composition in RCS.NZO.C3H-*Nbg15* mice. Fat mass (A), lean mass (B) in RCS.NZO.C3H-*Nbg15* mice fed a 45 % kcal high-fat diet. Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =31-38 mice; *Nbg15*$^{C3H/C3H}$ = 15-25 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

3.3.3.2 Metabolic characterization of RCS.NZO.C3H-*Nbg15* during fasting and refeeding

For the fasting and refeeding measurements mice with an age of 12 weeks feed a high fat diet (45 % kcal of fat) were fasted overnight for 16 h. Subsequently, the mice were refeed for 2h with the HFD. The tail blood was collected during fasting and refeeding for plasma insulin, free-fatty acids and triglycerides determination.

Blood glucose assessed during fasting state demonstrated that the *Nbg15c*$^{C3H/C3H}$ mice had on average 116.9 mg/dl lower blood glucose levels compared to *Nbg15*$^{NZO/NZO}$ genotype (*Nbg15*$^{NZO/NZO}$ 226.6 ± 30 mg/dl, *Nbg15c*$^{C3H/C3H}$ 109.6 ± 7.2 mg/dl, $p<0.05$). However, no differences were observed in the *Nbg15p*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ cohort compared to *Nbg15*$^{NZO/NZO}$. During the refeeding state, *Nbg15c*$^{C3H/C3H}$ mice exhibited blood glucose levels that were on average 176.4 mg/dl lower than in the *Nbg15*$^{NZO/NZO}$ mice (*Nbg15*$^{NZO/NZO}$ 390 ± 33.6 mg/dl, *Nbg15c*$^{C3H/C3H}$ 213.7 ± 25.9 mg/dl, $p<0.001$). Moreover, the *Nbg15d*$^{C3H/C3H}$ mice exhibited on average 121 mg/dl lower blood glucose levels compared to the *Nbg15*$^{NZO/NZO}$ (*Nbg15*$^{NZO/NZO}$ 390 ± 33.6 mg/dl, *Nbg15c*$^{C3H/C3H}$ 269 ± 27.5 mg/dl, $p<0.05$). Blood glucose levels in the *Nbg15p*$^{C3H/C3H}$ group compared to *Nbg15*$^{NZO/NZO}$ did not change, neither during fasting or refeeding state (Figure 18A). After refeeding, blood glucose levels in the *Nbg15*$^{NZO/NZO}$ and *Nbg15p*$^{C3H/C3H}$ genotype were significantly increased while no significant differences were observed in blood glucose of fasting to refeeding state in the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice (Figure 18A).

The results of plasma insulin measurements during fasting revealed no differences. However, the *Nbg15c*$^{C3H/C3H}$ mice had significantly higher insulin levels after refeeding, when compared to the *Nbg15*$^{NZO/NZO}$ mice (*Nbg15*$^{NZO/NZO}$ 10.8 ± 1.5 μg/L, *Nbg15c*$^{C3H/C3H}$ 14.6 ± 1.8 μg/L, $p<0.05$). In addition, the *Nbg15d*$^{C3H/C3H}$ genotype had on average 3.5 μg/L higher insulin levels during refeeding compared to the *Nbg15*$^{NZO/NZO}$ group. However, the differences failed to reach statistical significance (*Nbg15*$^{NZO/NZO}$ 10.8 ± 1.5 μg/L, *Nbg15d*$^{C3H/C3H}$ 14.3 ± 1.8 μg/L, $p<0.09$). Moreover, comparisons within the genotypes of blood glucose values during fasting to refeeding showed a significant increase in insulin levels in all lines. (Figure 18B).

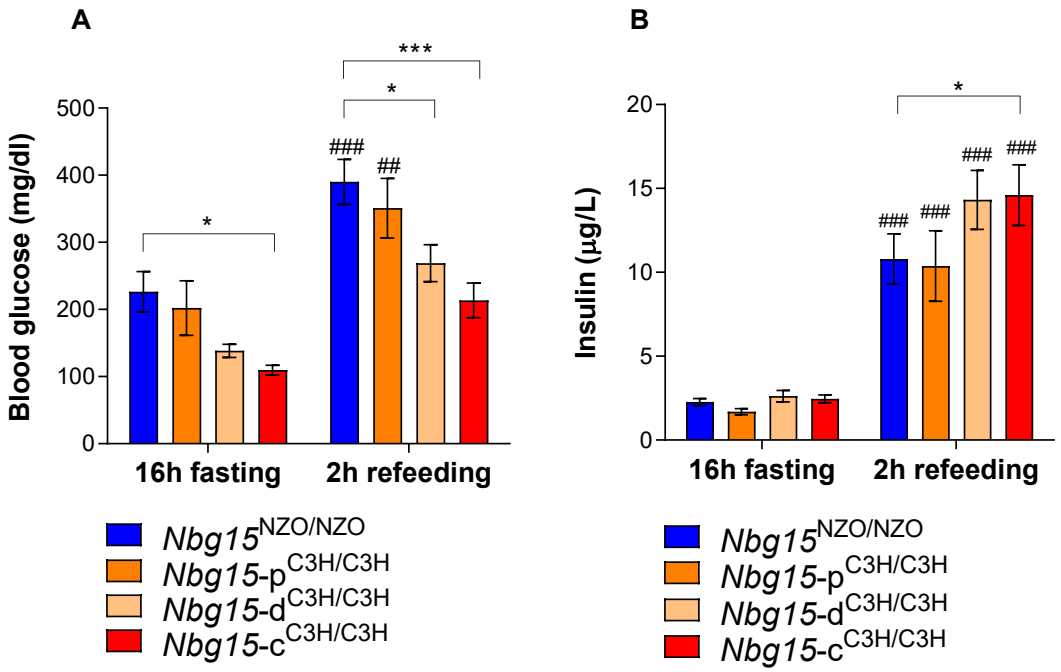


Figure 18. Blood glucose and plasma insulin levels during fasting and refeeding in RCS.NZO.C3H-*Nbg15*. Blood glucose (A), plasma insulin (B) in RCS.NZO.C3H-*Nbg*15 mice fed a 45 % kcal high-fat diet. Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =31 mice; *Nbg15*$^{C3H/C3H}$ = 16-20 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001. Comparisons within the same genotype during fasting to the refeeding are represented as #p<0.05, ##p<0.01, ###p<0.001, 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test.

Furthermore, the plasma triglycerides levels were not altered between the genotypes after 16h fasting. Nevertheless, the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice exhibited significantly lower plasma triglycerides after 2h refeeding in comparison to the *Nbg15*$^{NZO/NZO}$ group, while the *Nbg15p*$^{C3H/C3H}$ group showed similar plasma triglycerides levels compared to the *Nbg15*$^{NZO/NZO}$ mice (*Nbg15*$^{NZO/NZO}$ 303.4 ± 7.8 mg/dl, *Nbg15c*$^{C3H/C3H}$ 237.8 ± 14.6 mg/dl, *Nbg15d*$^{C3H/C3H}$ 183.5 ± 9.5 mg/dl, *Nbg15p*$^{C3H/C3H}$ 427 ± 96.6 mg/dl p<0.05) (Figure 19A). Additionally, comparisons within the genotypes revealed that the *Nbg15*$^{NZO/NZO}$ and *Nbg15p*$^{C3H/C3H}$ mice had significantly increased plasma triglycerides levels after refeeding compared to the fasting state, whereas no differences were observed in the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice for the same comparisons. Plasma free-fatty-acids (FFA) assessed during fasted and refeed state were not different when comparing the homozygous C3H allele carriers mice (*Nbg15c*$^{C3H/C3H}$, *Nbg15p*$^{C3H/C3H}$, *Nbg15d*$^{C3H/C3H}$) with the NZO allele carriers (*Nbg15*$^{NZO/NZO}$) on chromosome 15 exhibited no

differences during fasting and refeeding. However, within all groups, a significant reduction of FFA levels after refeeding was observed. (Figure 19B).

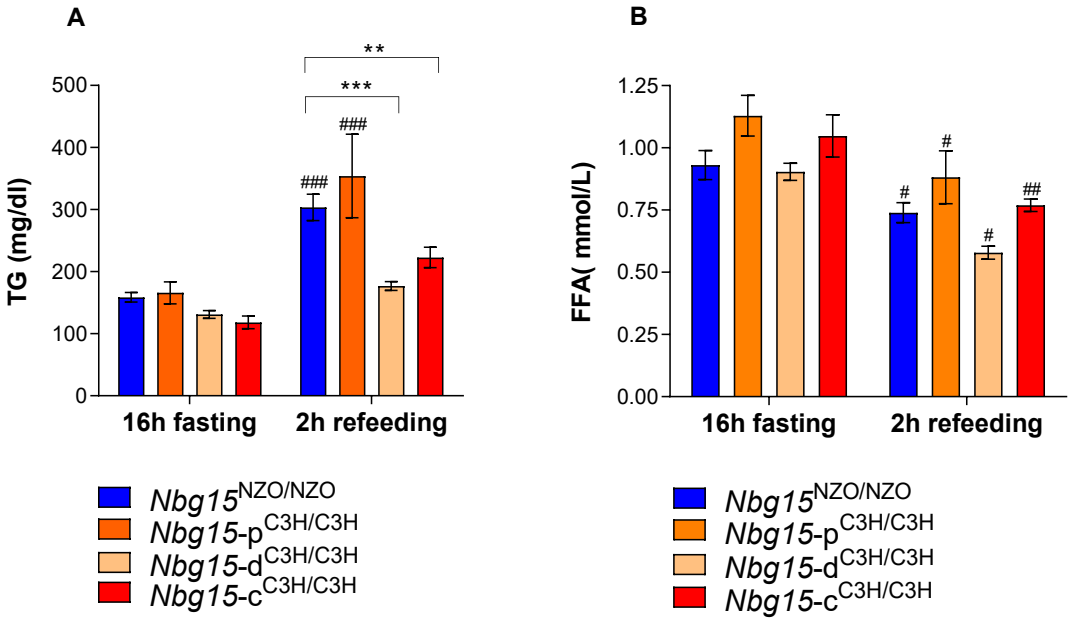


Figure 19. Plasma triglycerides and free-fatty acids measurements during fasting and refeeding in RCS.NZO.C3H-*Nbg15* mice. Plasma triglycerides (A), free-fatty-acids (B) in RCS.NZO.C3H-*Nbg15* mice fed a 45 % kcal high-fat diet. Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =31 mice; *Nbg15*$^{C3H/C3H}$ = 16-20 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001. Comparisons within the same genotype during fasting to the refeeding are represented as #p<0.05, ##p<0.01, ###p<0.001, 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test.

3.3.3.3 Intraperitoneal glucose tolerance test (i.p.GTT)

For the i.p.GTT, mice with an age of 13 weeks were fasted overnight for 16 h. Tail blood was collected and the blood glucose was measured after fasting period and 15, 30, 60, 120, 240 min after injection of a high glucose dose (2g/Kg body weight). Due to the fact that the glucometer has a limitation to measure blood glucose levels over 600 mg/dl, the plasma blood glucose was further measured using a glucose oxidase colorimetric reaction kit (see item 2.2.2.2). As demonstrated in Figure 20A, only the *Nbg15c*$^{C3H/C3H}$ mice had improved glucose tolerance. Significantly lower blood glucose levels were observed for this group at 30 and 60 min after glucose injection when compared to the *Nbg15*$^{NZO/NZO}$ mice. This effect could also be observed in the calculated area under the curve (AUC) (Figure 20B). *Nbg15c*$^{C3H/C3H}$ mice exhibited significantly lower AUC compared to *Nbg15*$^{NZO/NZO}$. However, although basal levels were not statistically different, a tendency was observed showing that the *Nbg15c*$^{C3H/C3H}$ mice had a tendency for lower fasting blood glucose levels, which could

have resulted in the improved glucose clearance seen over time in this genotype (Figure 20C).

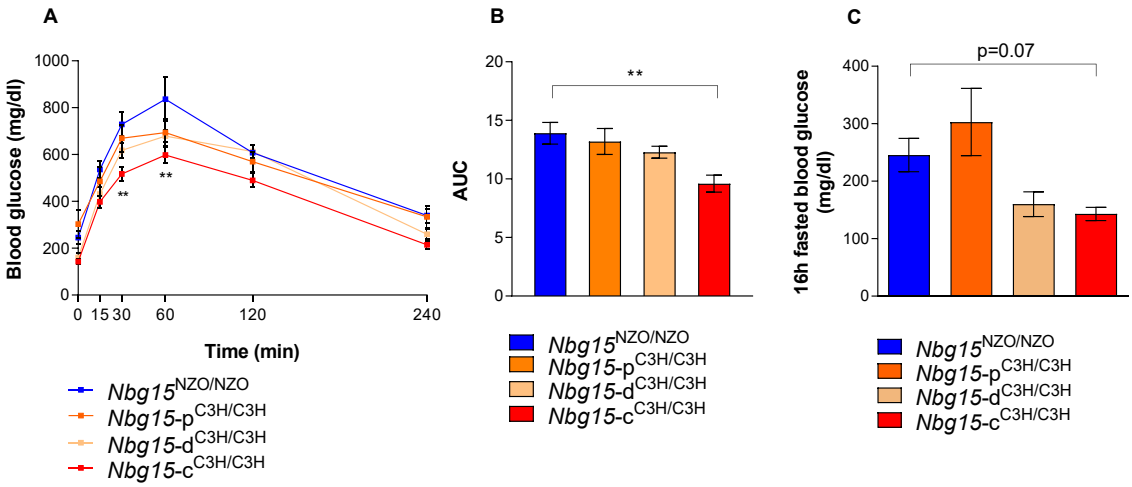


Figure 20. Intraperitoneal glucose tolerance test (i.p.GTT) in RCS.NZO.C3H-*Nbg15*. Blood glucose levels after glucose overload (A), area under the curve (B), 16h fasted blood glucose (C). Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =31 mice; *Nbg15*$^{C3H/C3H}$ = 16-20 mice. 1 way ANOVA (B, C) or 2-way ANOVA (A) followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

Moreover, insulin levels were assessed during i.p.GTT. After glucose injection, no significant differences were observed between the genotypes carrying the C3H alleles on chromosome15 compared to the carriers of the NZO alleles (Figure 21A). Additionally, the calculated area under the curve confirmed these results (Figure 21B). However, the analysis of basal insulin levels alone showed that the *Nbg15c*$^{C3H/C3H}$ mice had significantly higher insulin levels when compared to the *Nbg15*$^{NZO/NZO}$ group (Figure 21C).

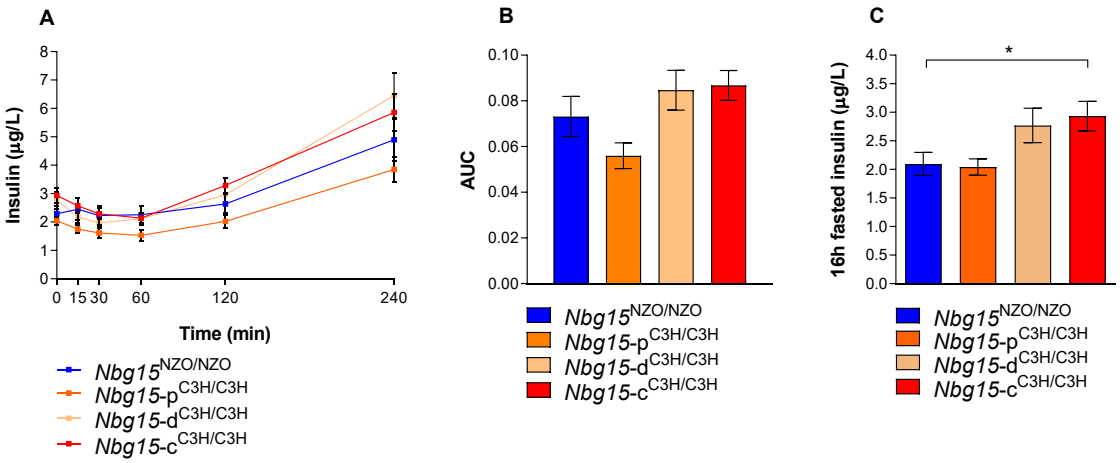


Figure 21. Insulin during the intraperitoneal glucose tolerance test (i.p.GTT) in RCS.NZO.C3H-*Nbg15*. Plasma insulin levels after glucose injection (2g/kg body weight) (A), area under the curve (B), basal insulin levels (C). Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =31 mice; *Nbg15*$^{C3H/C3H}$ = 16-20 mice. 1 way ANOVA (B, C) or 2-way ANOVA (A) followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

3.3.3.4 Intraperitoneal insulin tolerance test (i.p.ITT)

For the i.p.ITT, mice at an age of 14 weeks were fasted for 6h before receiving an intraperitoneal injection of insulin (1IU/Kg of body weight). Blood glucose was measured at fasting state and 15, 30, 60 min after insulin injection. The fasting blood glucose levels during i.p.ITT in the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice was significantly lower compared to *Nbg15*$^{NZO/NZO}$ (*Nbg15*$^{NZO/NZO}$ 351.7 ± 38.4 mg/dl, *Nbg15c*$^{C3H/C3H}$ 201.8 ± 37.3 mg/dl, *Nbg15d*$^{C3H/C3H}$ 203 ± 40 mg/dl, p≤ 0.05) (Figure 22A). For this reason, the results of i.p.ITT were normalized to basal levels. Thus, no differences could be observed between the genotypes (Figure 22B).

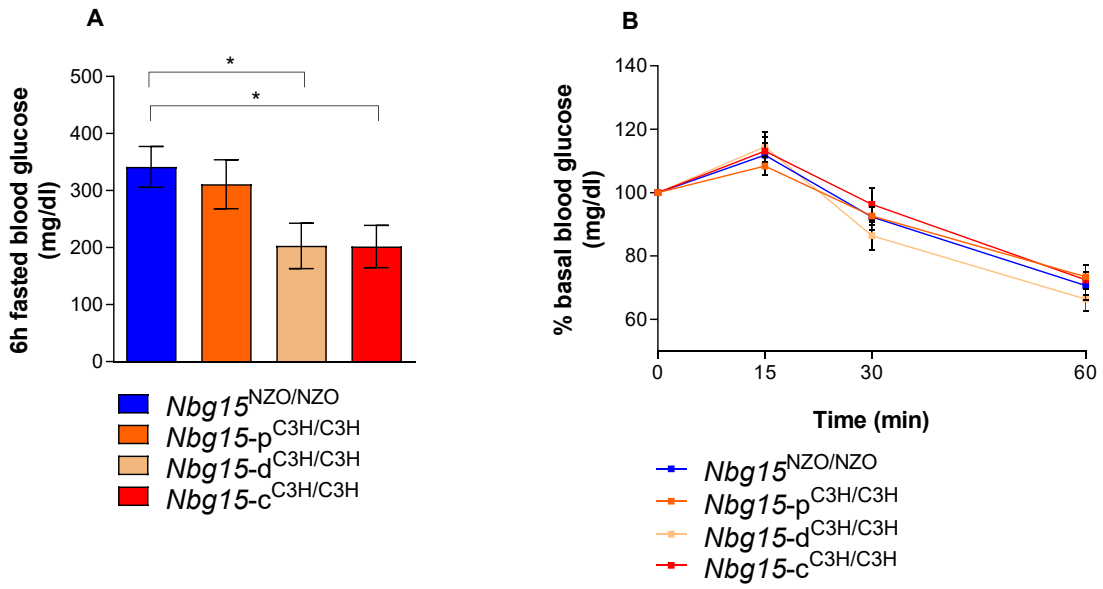


Figure 22. Insulin tolerance test (i.p.ITT) in RCS.NZO.C3H-*Nbg15*. 6 h fasted basal blood glucose levels (A), blood glucose during i.p.ITT (B). Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =~27 mice; *Nbg15*$^{C3H/C3H}$ = 16-20 mice. 1way ANOVA (A) or 2-way ANOVA (B) followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

3.3.3.5 Final fasting measurements

At an age of 15 weeks, mice were fasted in the morning for 6h before sacrificing. Blood was collected for final fasting measurements that included the determination of plasma insulin levels. Moreover, organs were collected for further metabolic characterization.

3.3.3.5.1 Final blood glucose and plasma insulin measurements

The final fasting measurements of blood glucose revealed that the consomic and distal homozygous C3H allele carriers had lower blood glucose levels compare to the NZO allele carriers on *Nbg15*. The *Nbg15c*$^{C3H/C3H}$ mice had 168 mg/dl lower fasting blood glucose compared to the *Nbg15*$^{NZO/NZO}$ mice (*Nbg15*$^{NZO/NZO}$ 358.3 ± 32.1 mg/dl, *Nbg15c*$^{C3H/C3H}$ 190.5 ± 23.4 mg/dl, p≤ 0.001). The *Nbg15d*$^{C3H/C3H}$ mice presented the highest difference when comparing to the *Nbg15*$^{NZO/NZO}$ genotype. The blood glucose levels were on average 185 mg/dl lower compared to the *Nbg15*$^{NZO/NZO}$ mice (*Nbg15*$^{NZO/NZO}$ 358.3 ± 32.1 mg/dl, *Nbg15d*$^{C3H/C3H}$ 173.6 ± 13 mg/dl, p≤ 0.001). The Nbg15p$^{C3H/C3H}$ mice exhibited the lowest effect with average effect size of 73.4 mg/dl compared to *Nbg15*$^{NZO/NZO}$ genotype. Although this represents huge differences, statistical comparison between these groups failed to reach the significance threshold

(*Nbg15*$^{NZO/NZO}$ 358.3 ± 32.1 mg/dl, *Nbg15p*$^{C3H/C3H}$ 285.6 ± 40.4 mg/dl, p=0.22) (Figure 23A).

The observed lower fasting blood glucose levels in *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice could be explained by the elevated plasma insulin levels. The *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice had respectively an average 6.9 µg/L and 9.4 µg/L elevated plasma insulin compared to the *Nbg15*$^{NZO/NZO}$ group (*Nbg15*$^{NZO/NZO}$ 5.9 ± 0.89 µg/L, *Nbg15d*$^{C3H/C3H}$ 15.2 ± 2.3 µg/L, *Nbg15c*$^{C3H/C3H}$ 12.8 ± 1.9 µg/L, p≤ 0.01). Contrary, no differences on plasma insulin levels were observed in the *Nbg15p*$^{C3H/C3H}$ mice compared to the *Nbg15*$^{NZO/NZO}$ genotype, suggesting that the effects on blood glucose addressed by the proximal locus probably derived from other pathophysiological mechanisms (*Nbg15*$^{NZO/NZO}$ 5.9 ± 0.89 µg/L, *Nbg15p*$^{C3H/C3H}$ 4.1 ± 0.6 µg/L, p≤ 0.01) (Figure 23B).

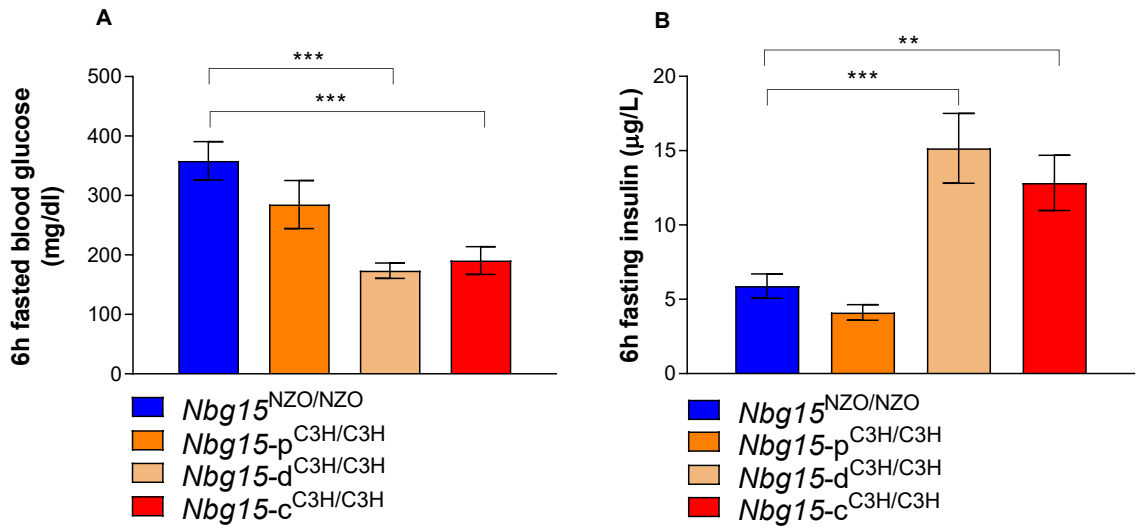


Figure 23. Final blood glucose and insulin levels in RCS.NZO.C3H-*Nbg15* mice. 6 h fasted blood glucose levels measured at 15 weeks of age (A), final measurement of plasma insulin levels (B). Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =~27 mice; *Nbg15*$^{C3H/C3H}$ = 16-20 mice. 1-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

3.3.3.5.2 Inflammatory markers

Diabetes type-2 is associated to low grade inflammation. To determine whether the phenotypic characteristics of the RCS.NZO.C3H-*Nbg15* resulted from differences in circulating pro-inflammatory markers, plasma was collected for multiplexing measurements of cytokines levels. The results are summarized in Table 12 and showed that pro-inflammatory cytokines were significantly reduced in mice carrying the homozygous C3H alleles on *Nbg15* locus compared to the *Nbg15* NZO

allele carriers. A tendency for reduced levels of IL-1b was observed in all *Nbg15* lines carrying the C3H alleles compared to the NZO allele carriers. In addition, MIP-1b levels were significantly reduced in the *Nbg15c*$^{C3H/C3H}$, *Nbg15p*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice compared to the *Nbg15*$^{NZO/NZO}$ group. Moreover, levels of TNF-alpha were significantly lower in the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice, while a tendency for reduction was observed in the *Nbg15p*$^{C3H/C3H}$ group.

Moreover, the *Nbg15c*$^{C3H/C3H}$ mice exhibited significantly lower levels of IL-12(p40), G-CSF, KC and MCP-1, while a tendency for reduction was observed for RANTES levels. On the other hand, in the *Nbg15d*$^{C3H/C3H}$ group the plasmatic levels of RANTES were significantly reduced.

Table 12. Determination of circulation cytokines in RCS.NZO.C3H-*Nbg15* mice. Plasma cytokines were measured in plasma of 6h fasted mice at an age of 15 weeks fed a 45 % kcal HFD. Significant results were considered as p<0.05 indicated in bold, unpaired ttest. n=5-24. IL-1a (Interleukin 1 alpha), IL-1b (Interleukin 1 beta), IL-12(p40) (Interleukin 12 subunit p40), G-CSF (Granulocyte-Colony Stimulating Factor), KC (Keratinocyte chemoattractant), MCP-1 (Monocyte chemoattractant protein-1), MIP-1b (Macrophage Inflammatory Protein-1 beta), RANTES (Regulated and Normal T cell Expressed and Secreted), TNF-α (Tumor necrosis factor-α)

Cytokines	*Nbg15*$^{NZO/NZO}$		*Nbg15c*$^{C3H/C3H}$			*Nbg15p*$^{C3H/C3H}$			*Nbg15d*$^{C3H/C3H}$		
	Mean	**SEM**	**Mean**	**SEM**	**p-value**	**Mean**	**SEM**	**p-value**	**Mean**	**SEM**	**p-value**
IL-1a	31.66	4.79	20.97	3.95	0.104	20.86	3.87	0.097	29.18	4.79	0.722
IL-1b	249.28	17.95	77.08	14.16	**0.063**	83.44	18.37	**0.074**	88.83	17.95	**0.083**
IL-12(p40)	324.07	98.96	229.67	28.53	**0.023**	336.10	30.13	0.763	367.48	98.96	0.678
G-CSF	336.29	68.54	57.05	3.72	**0.034**	71.88	11.10	0.934	189.98	68.54	0.501
KC	27.10	29.81	10.65	0.58	**0.001**	34.49	17.62	0.697	55.40	29.81	0.369
MCP-1	86.23	22.51	42.84	6.76	**0.027**	56.69	11.00	0.160	94.45	22.51	0.775
MIP-1b	24.89	1.40	4.07	0.68	**0.009**	7.33	2.76	**0.031**	6.77	1.40	**0.022**
RANTES	27.05	1.52	19.59	2.20	**0.052**	23.08	4.58	0.475	17.24	1.52	**0.006**
TNF-a	286.13	14.45	63.36	5.63	**0.017**	113.54	47.48	**0.088**	97.51	14.45	**0.041**

3.3.3.5.3 Organs weight

At 15 weeks of age the mice were sacrificed and liver, gonadal white adipose tissue and subcutaneous adipose tissue were weighted as markers for fatty liver-associated phenotypes as well as differentiation of adipose depots between the genotypes. The results revealed no differences in liver weight comparing all groups (Figure 24A). However, gonadal fat was significantly increase in the *Nbg15d*$^{C3H/C3H}$ mice compared to *Nbg15*$^{NZO/NZO}$ group *(Nbg15*$^{NZO/NZO}$ 2.4 ± 0.2 g, *Nbg15d*$^{C3H/C3H}$ 3.14 ± 0.2 g, p≤ 0.05). In addition, the *Nbg15d*$^{C3H/C3H}$ as well as *Nbg15c*$^{C3H/C3H}$ mice exhibited higher amount of subcutaneous fat compared to *Nbg15*$^{NZO/NZO}$ genotype. (*Nbg15*$^{NZO/NZO}$ 2.8 ± 0.25 g, *Nbg15d*$^{C3H/C3H}$ 4.6 ± 0.2 g, p≤ 0.001) (Figure 24B).

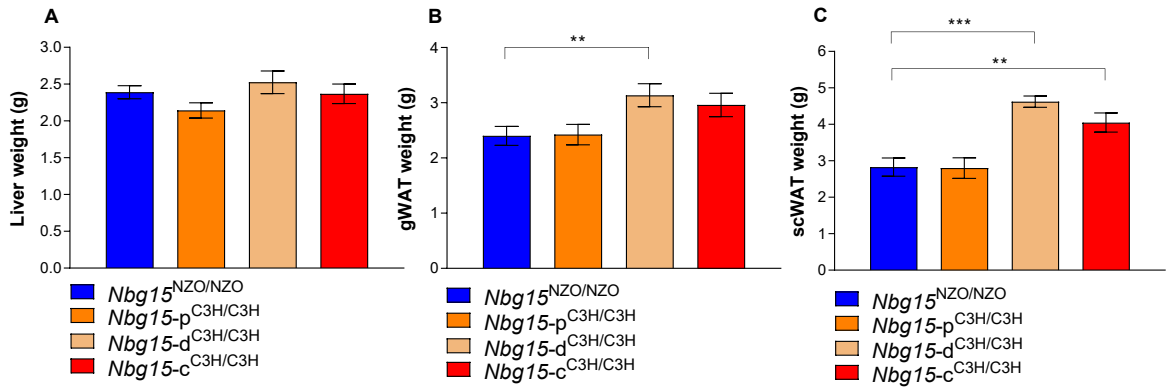


Figure 24. Final organs weight in RCS.NZO.C3H-*Nbg15* mice. Liver weight measured at 15 weeks of age (A), gonadal white adipose tissue weight (gWAT) (B), subcutaneous adipose tissue (scWAT) (C). Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =~25 mice; *Nbg15*$^{C3H/C3H}$ = 15-20 mice. 1-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

3.3.3.5.4 Histological analysis of pancreatic islets

Final measurements of blood glucose and plasma insulin indicated that mice carrying the homozygous C3H alleles at *Nbg15* locus could be protected from diabetes. To investigate whether these effects were linked to an islet associated phenotype the pancreata of mice fed a HFD and at an age of 15 weeks were collected for histological analysis. For this purpose, the pancreas was sectioned at different levels and 8 sections per pancreas were selected for further analysis.

The results demonstrated that, in comparison to the *Nbg15*$^{NZO/NZO}$ mice, the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice had on average higher number of islets representing respectively an increase of 58 % and 40 %. However, these results failed to reach statistical significance (Figure 25A). Nevertheless, when islets were segregated within a cluster size, it was observed that the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice exhibited significantly higher islets number within the cluster size 0-5000 μm^2 when compared to the *Nbg15*$^{NZO/NZO}$ group (Figure 25B). Moreover, the average pancreas area did not differ between the genotypes, although a tendency for smaller pancreas area was observed in the *Nbg15c*$^{C3H/C3H}$ line when compared to *Nbg15*$^{NZO/NZO}$ mice (Figure 25C). Calculation of islet density did not reveal significant differences between the genotypes. However, the *Nbg15d*$^{C3H/C3H}$ mice seemed to have a tendency for increased islets density (Figure 25D) (Master thesis Jenny Khuong, 2019).

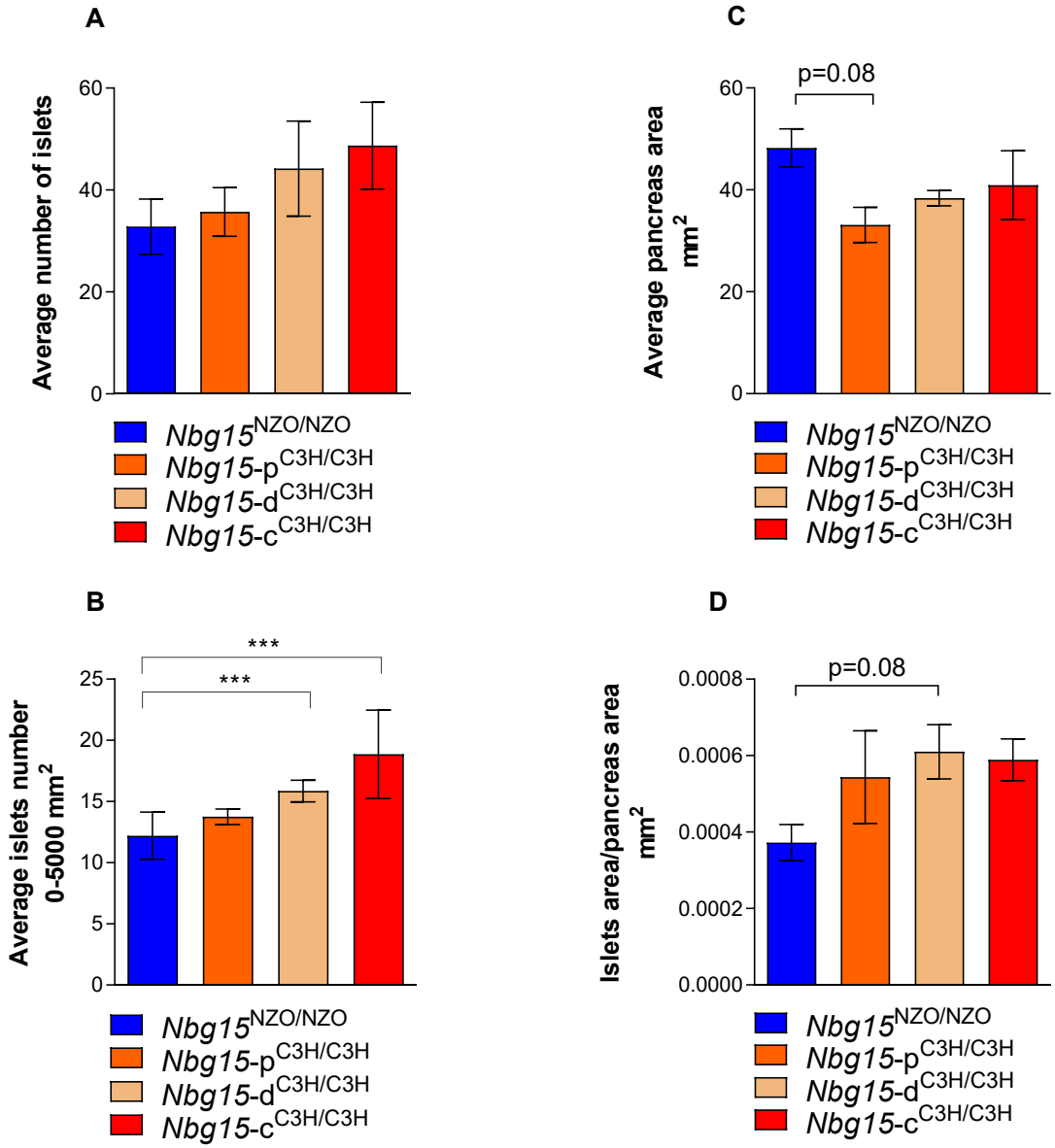


Figure 25. Histological analysis of pancreatic tissue from RCS.NZO.C3H-*Nbg15* mice. Total number of islets (A), total islets number within cluster (0-5000 mm^2) (B), average pancreas area (C), islets density (D). Results are represented as MEAN ± SEM of *Nbg15*$^{NZO/NZO}$ =7 mice; *Nbg15*$^{C3H/C3H}$ = 3 mice. 1-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

3.4 Metabolic characterization of RCS.B6.C3H-*Nbg15*

3.4.1 Random blood glucose, body weight and body composition

In contrast to what was observed in RCS lines generated on a NZO background, the *Nbg15*-C3H allele carriers on a B6 background did not address any effect on random blood glucose measurements as well as body weight development

(Figure 26 A-B). However, differences in body composition could be observed. The B*Nbg15*$^{C3H/C3H}$ mice had reduced body fat content at 15 weeks of age when compared to B*Nbg15*$^{B6/B6}$ mice (B*Nbg15*$^{B6/B6}$ 11.09 ± 1.2 g, B*Nbg15*$^{C3H/C3H}$ 7.8 ± 0.8 g, p<0.001). Moreover, the B*Nbg15*$^{C3H/C3H}$ mice exhibited lean mass content that was on average 2.5 g higher than for the B*Nbg15*$^{B6/B6}$ mice at an age of 15, although significantly higher lean mass was already observed at 10 weeks of age (B*Nbg15*$^{B6/B6}$ 26.5 ± 1.1 g, B*Nbg15*$^{C3H/C3H}$ 29.8 ± 0.4 g, p<0.05) (Figure 26C-D).

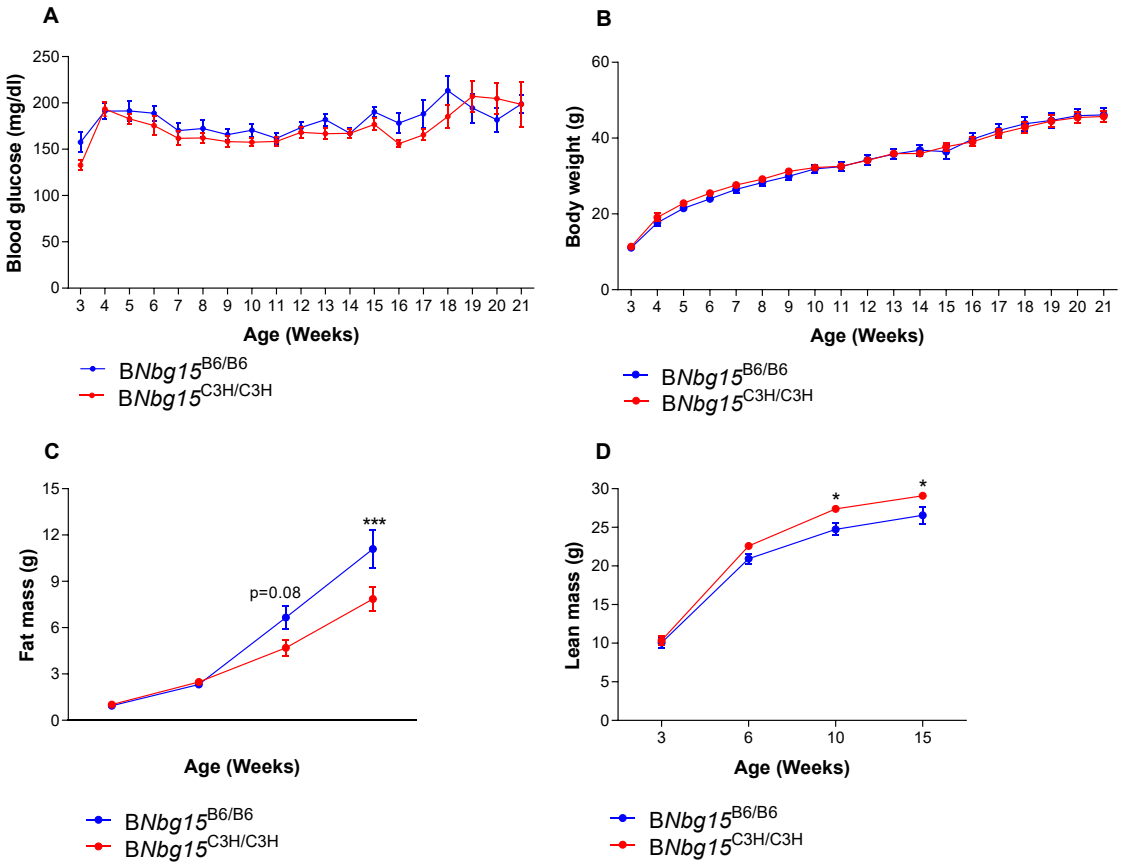


Figure 26. Blood glucose and body weight development and body composition in RCS.B6.C3H-*Nbg15* mice. Random blood glucose development (A), body weight development (B), fat mass (C) and lean mass (D) in RCS.B6.C3H-*Nbg15* fed a 60 % kcal high-fat diet. Results are represented as MEAN ± SEM of B*Nbg15*$^{B6/B6}$ =13 mice; B*Nbg15*$^{C3H/C3H}$ = 19 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001.

3.4.2 Metabolic characterization of RCS.B6.C3H-*Nbg15* during fasting and refeeding

At 10 weeks of HFD intervention the RCS.B6.C3H-*Nbg15* mice were fasted for 16h and refeed for 2h for measurements of blood glucose, plasma triglycerides, plasma insulin and free-fatty acids. The results revealed that the fasting blood glucose levels

in the B*Nbg15*$^{C3H/C3H}$ mice were significantly lower when compared to the B*Nbg15*$^{B6/B6}$ group (B*Nbg*15$^{B6/B6}$ 99 ± 4 mg/dl, B*Nbg15*$^{C3H/C3H}$ 84 ± 3.8 mg/dl, p<0.05). Nevertheless, no differences on blood glucose were observed after refeeding. However, in both lines, the blood glucose levels were significantly increased after refeeding when compared to their fasting levels (Figure 27A).

Furthermore, fasted plasma insulin was not different in B*Nbg15*$^{C3H/C3H}$ mice compared to B*Nbg15*$^{B6/B6}$. After 2h refeeding, the B*Nbg15*$^{C3H/C3H}$ mice exhibited increased plasma insulin levels compared to the B*Nbg15*$^{B6/B6}$ group (B*Nbg15*$^{B6/B6}$ 2.5 ± 0.24 µg/L, B*Nbg15*$^{C3H/C3H}$ 3.3 ± 0.34 µg/L, p<0.05). Additionally, increased insulin levels were observed after refeeding in both lines when compared to the fasted levels (Figure 27B). Measurements of triglycerides during fasting and refeeding revealed that the B*Nbg15*$^{C3H/C3H}$ mice had reduced plasma triglycerides compared to B*Nbg15*$^{B6/B6}$ mice with an effect size of 65.6 mg/dl at refeeding state (B*Nbg15*$^{B6/B6}$ 208.1 ± 9.7 mg/dl, B*Nbg15*$^{C3H/C3H}$ 142.5 ± 12.4 mg/dl, p<0.001). On the other hand, no differences were observed in plasma triglycerides I the feasted state. However, for both groups an increased level of plasma TG was observed when comparing the fasted to refeeding state (Figure 27C).

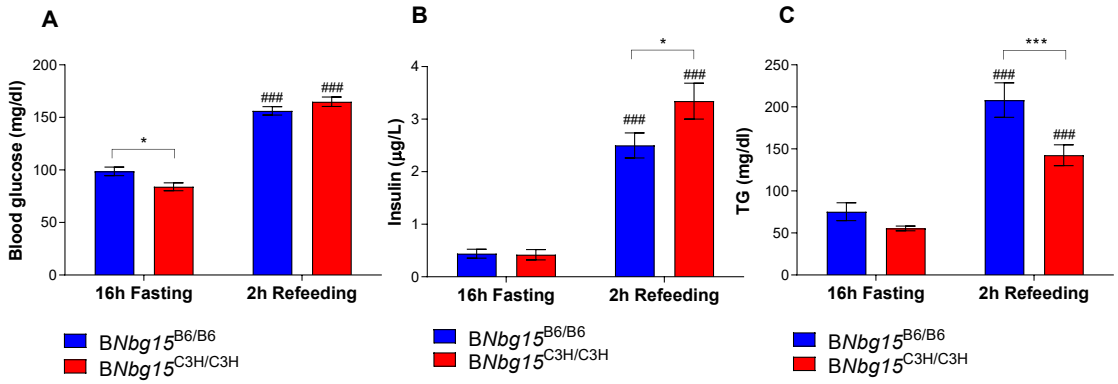


Figure 27. Metabolic characterization of RCS.B6.C3H-*Nbg15* mice during fasting and refeeding. Blood glucose levels (A), plasma insulin (B), plasma triglycerides (C) during fasting and refeeding in RCS.B6.C3H-*Nbg15* genotypes fed a 60 % kcal high-fat diet. Results are represented as MEAN ± SEM of B*Nbg15*$^{B6/B6}$ =13 mice; B*Nbg15*$^{C3H/C3H}$ = 19 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test, *p<0.05, **p<0.01, ***p<0.001. Comparisons within fasted and refeed measurements are represented as #p<0.05, ##p<0.01, ###p<0.001, 2- way ANOVA followed by Sidak's multiple comparison post-hoc-test.

Furthermore, the measurements of circulating free-fatty acids revealed no differences between the B*Nbg15*$^{C3H/C3H}$ mice compared to B*Nbg15*$^{B6/B6}$ mice neither at fasting nor at refeeding state (Figure 28A). However, in both groups, FFA levels were significantly decreased after refeeding when compared to the respective fasted values,

whereas the ratio fasted/refeed in B*Nbg15*$^{C3H/C3H}$ was significantly higher compared to B*Nbg15*$^{B6/B6}$ (Figure 28B).

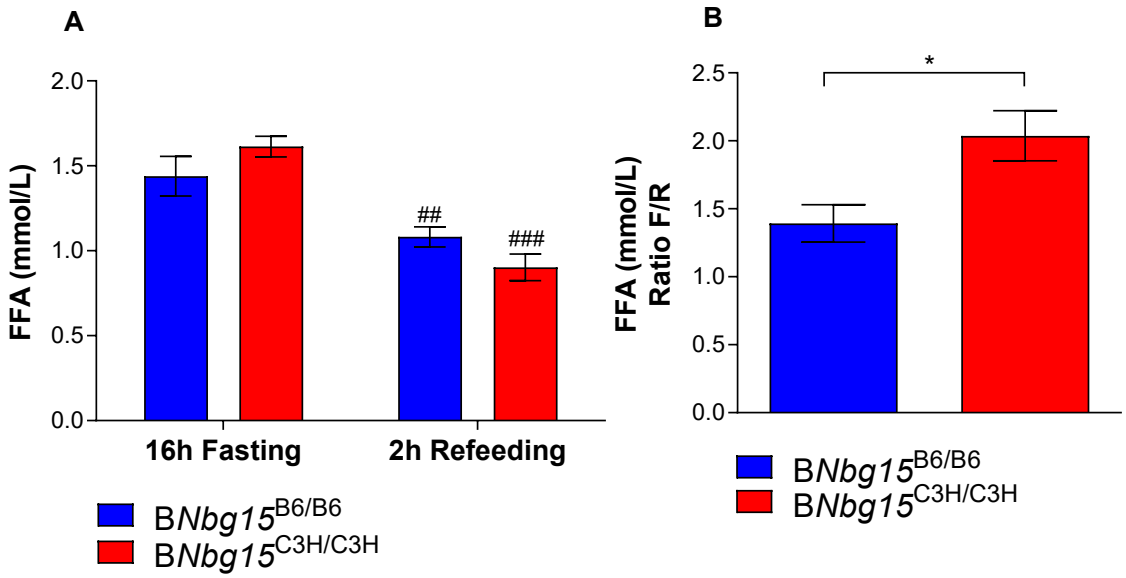


Figure 28. Metabolic characterization of RCS.B6.C3H-*Nbg15* mice during fasting and refeeding. Plasma free-fatty acids levels (A), Free-fatty acids ratio fasting/refeeding (B) in RCS.B6.C3H-*Nbg15* fed a 60 % kcal high-fat diet. Results are represented as MEAN ± SEM of B*Nbg15*$^{B6/B6}$ =13 mice; B*Nbg15*$^{C3H/C3H}$ = 19 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test and unpaired ttest, *p<0.05, **p<0.01, ***p<0.001. Comparisons within fasted and refeed measurements are represented as #p<0.05, ##p<0.01, ###p<0.001, 2- way ANOVA followed by Sidak's multiple comparison post-hoc-test.

3.4.3 Intraperitoneal glucose tolerance test (i.p.GTT)

Glucose tolerance test was performed after 13 weeks of HFD intervention. Mice were fasted for 6 h before receiving glucose intraperitoneal injected. The results revealed that the B*Nbg15*$^{C3H/C3H}$ mice significantly improved glucose tolerance compared to the B*Nbg15*$^{B6/B6}$ group. Blood glucose measured 30 and 60 min after glucose injection was respectively reduced by 92 % and 83 % in addition to decreased area under the curve when compared to B*Nbg15*$^{B6/B6}$ (Figure 29A-B). The improved glucose tolerance was not addressed by improved secreted insulin as no differences in plasma insulin levels measured during the GTT could be observed, result that was also demonstrated in the calculated AUC (Figure 29C-D).

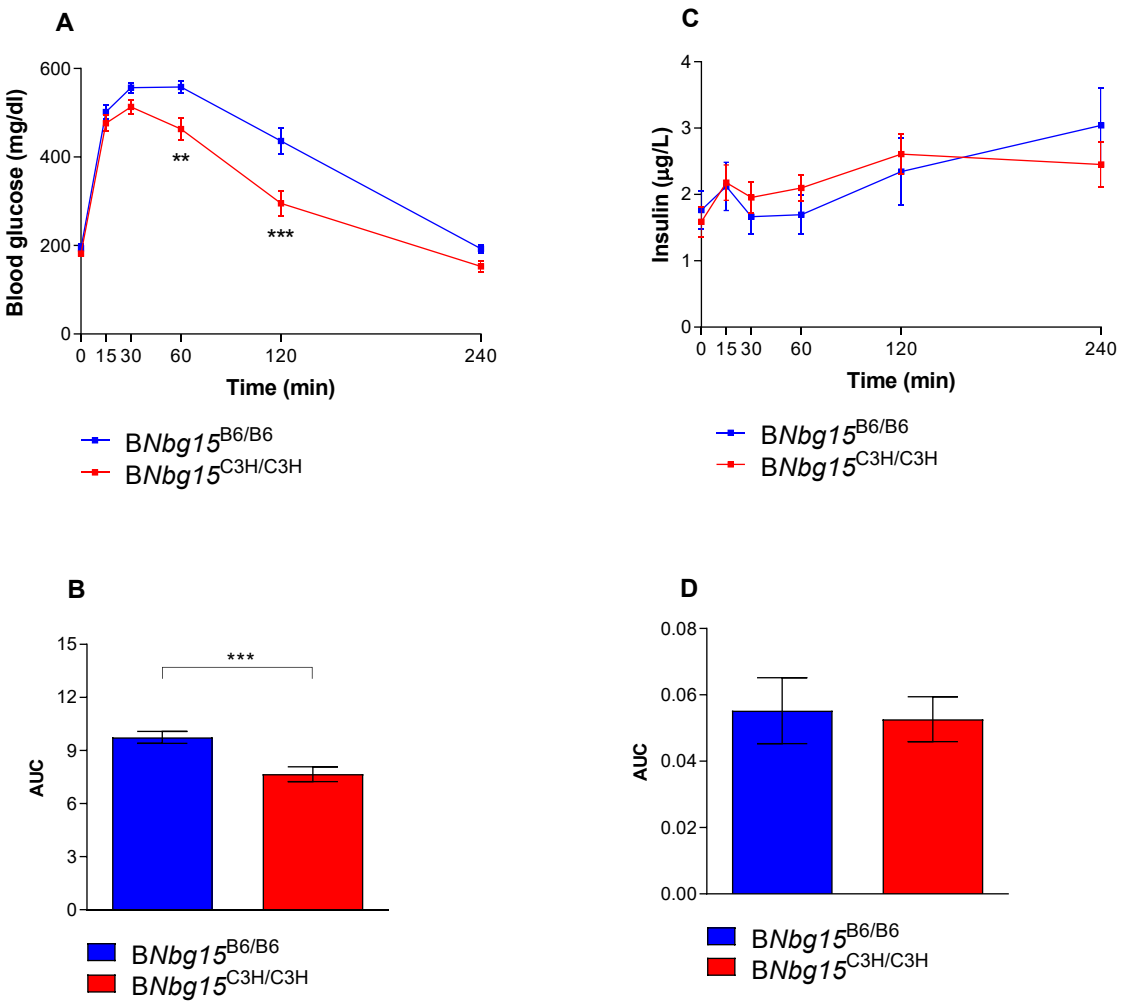


Figure 29. Intraperitoneal glucose tolerance test i.p.GTT (A), i.p.GTT area under the curve (B), insulin levels during i.p.GTT (C) and insulin AUC (D) in the RCS.B6.C3H-*Nbg15* mice fed a 60 % kcal high-fat diet. Results are represented as MEAN ± SEM of B*Nbg15*$^{B6/B6}$ =13 mice; B*Nbg15*$^{C3H/C3H}$ = 19 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test or unpaired ttest, *p<0.05, **p<0.01, ***p<0.001.

3.4.4 Intraperitoneal insulin tolerance test (i.p.ITT)

Non-fasted mice at an age of 15 weeks were injected with insulin (0.1 U/mL) and blood glucose was monitored at different time points after injection to assess the insulin tolerance between the genotypes. The results showed that, in comparison to the *BNbg15*$^{B6/B6}$ mice, *BNbg15*$^{C3H/C3H}$ exhibited significantly lower blood glucose already at 15 min after insulin injection, effect that remained until 60 min alongside to a smaller area under the curve (Figure 30A-B). The results were normalized to basal due to an unfasted experimental condition. However, no differences on basal blood glucose could be observed (Figure 30C).

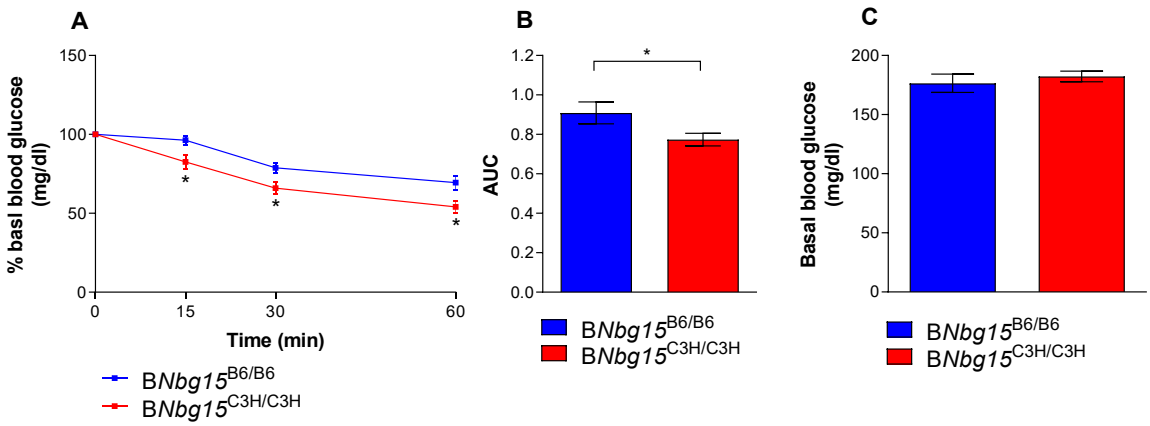


Figure 30. Intraperitoneal insulin tolerance test. Blood glucose (A), basal blood glucose levels (B) i.p.GTT area under the curve (C) in RCS.B6.C3H-*Nbg15* mice fed a 60 % kcal high-fat diet. Results are represented as MEAN ± SEM of B*Nbg15*$^{B6/B6}$ =13 mice; B*Nbg15*$^{C3H/C3H}$ = 19 mice. 2-way ANOVA followed by Sidak's multiple comparison post-hoc-test or unpaired ttest, *p<0.05, **p<0.01, ***p<0.001.

3.4.5 Final fasting measurements

RCS.B6.C3H-*Nbg15* mice were followed up for 21 weeks as B6 lines are known to have less risk to develop T2D. Thus, effects were expected to appear at later time points than for mice generated in a NZO background.

3.4.5.1 Final blood glucose and insulin

At 21 weeks of age, RCS.B6.C3H-*Nbg15* mice were fasted for 6 h for final blood glucose and insulin measurements. The results showed no differences in blood glucose levels when the B*Nbg15*$^{C3H/C3H}$ mice were compared to B*Nbg15*$^{B6/B6}$ (Figure 31A). Similarly, no significant differences were observed in final plasma insulin levels (Figure 31B).

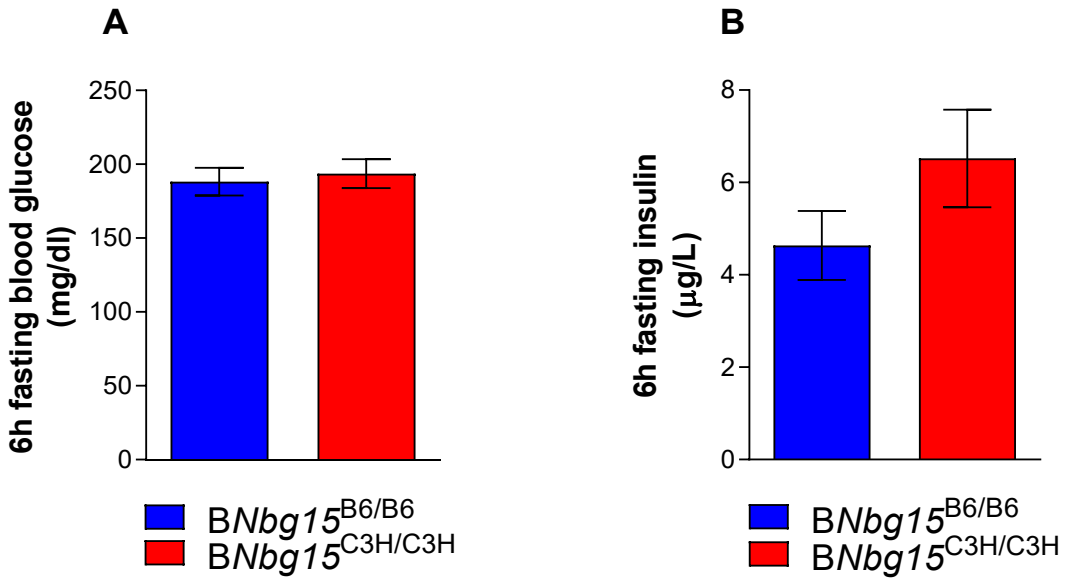


Figure 31. Final blood glucose and insulin levels in RCS.B6.C3H-*Nbg15* mice. 6 h fasted blood glucose levels measured at 21 weeks of age (A), final measurement of insulin (B). Results are represented as MEAN ± SEM of B*Nbg15*$^{B6/B6}$ =13 mice; B*Nbg15*$^{C3H/C3H}$ = 19 mice. Unpaired t-test, *p<0.05, **p<0.01, ***p<0.001.

3.4.5.2 Organs weight

Liver weight of the B*Nbg15*$^{C3H/C3H}$ group was not different compared to the B*Nbg15*$^{B6/B6}$ mice (Figure 32A). Additionally, no significant differences in gonadal fat were observed when the same groups were compared (Figure 32B). However, fat seems to accumulate more subcutaneously in the B*Nbg15*$^{B6/B6}$ than in B*Nbg15*$^{C3H/C3H}$ mice (B*Nbg15*$^{B6/B6}$ 2.05 ± 0.22 g, B*Nbg15*$^{C3H/C3H}$ 1.37 ± 0.15 g, p<0.001) (Figure 32C).

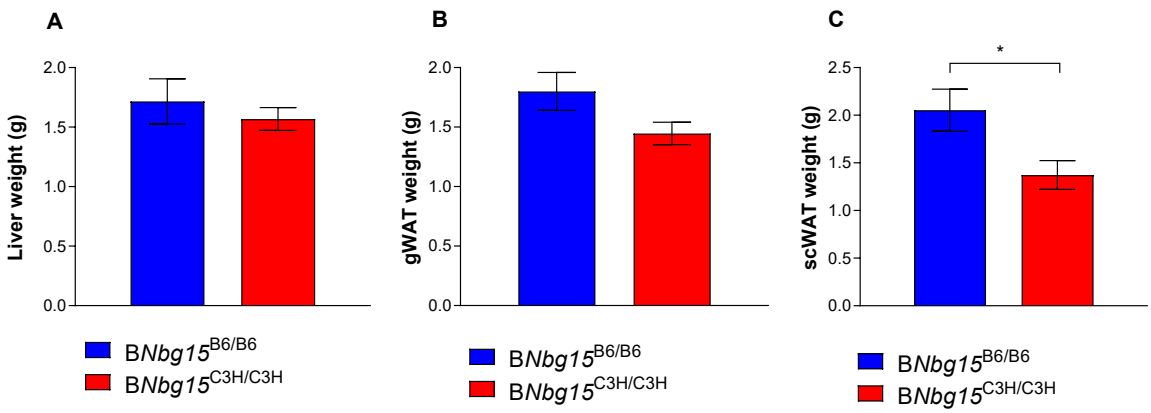


Figure 32. Final organs weight in RCS.B6.C3H-*Nbg15* mice. Liver weight measured at 21 weeks of age (A), gonadal white adipose tissue (gWAT) (B), subcutaneous adipose tissue (scWAT) (C). Results are represented as MEAN ± SEM of B*Nbg15*$^{B6/B6}$ =13 mice; B*Nbg15*$^{C3H/C3H}$ = 19 mice. Unpaired ttest, *p<0.05, **p<0.01, ***p<0.001.

3.5 Identification of causal gene variants for *Nbg15*

The results obtained in the RCS lines showed prove of evidence that *Nbg15* harbors at least one gene that contribute to pancreatic islets adaptations in a NZO background and improve glucose tolerance in a B6 background. However, as abovementioned, 503 genes are annotated within *Nbg15*. Therefore, some strategies were used to select genes with the highest likelihood to be link to the observed phenotypes.

For this purpose, a first strategy used was the gene expression profiling in the parental NZO and C3H mice. The gene expression in the parental mice was determined using microarray analysis and differentially expressed genes were further validated by Real time qPCR. Further, the expression of the validated genes was assessed in the corresponding tissue of the whole backcross (N_2(C3HXNZO)) population. Genes that presented a *cis*-eQTL were selected as stronger candidates for *Nbg15*. Furthermore, this data was overlaid with the haplotype and the variant effect predictor analysis, mainly with the aim to select genes associated to a pancreatic islets driven phenotype.

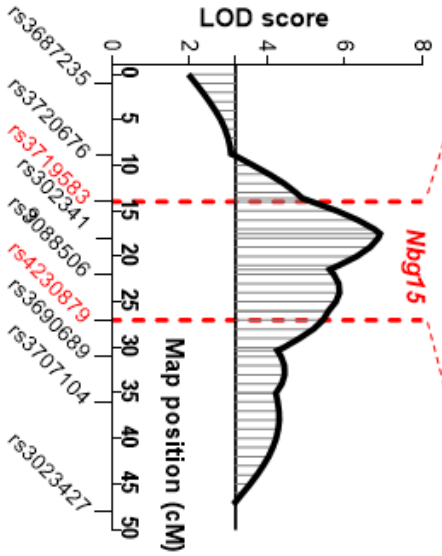
3.5.1 Gene expression profiling

Taking the assumption that genes harbored by the *Nbg15* locus are differentially regulated when parental mice are given a high fat diet, the parental NZO and C3H were fed a high-fat diet (45 % kcal) starting right after weaning (3 weeks of age). At 6 weeks of age, the mice were sacrificed and gonadal white adipose tissue

(gWAT), brown adipose tissue (BAT), quadriceps, liver and pancreatic islets were collected for expression profiling using microarray analysis. These would therefore reflect the early genetic adaptations to diet intervention. Moreover, the tissues were collected at an age of 21 weeks for later characterization of the genetic modulation.

Lastly, significantly differentially expressed genes pre-selected in the microarray analysis were further validated via qPCR. Thus, using this strategy, differentially expressed genes at 6 and 21 weeks were further selected as likely candidate genes for *Nbg15*.

The results of the microarray analysis revealed that a total of 54 genes annotated at the *Nbg15* locus were found to be differentially expressed within the parental C3H and NZO. In the liver, 8 out of 54 genes were differentially expressed, being 2 of them upregulated in the C3H strain and 6 in NZO. In the quadriceps, 5 genes were differentially expressed, 3 were highly expressed in the C3H mice and 2 in NZO. Moreover, most of the genes were found to be differentially regulated in the white and brown adipose tissue as well as in the pancreatic islets. In BAT, 9 out of 15 genes were overexpressed in the C3H strain and 6 in the NZO. In gWAT and pancreatic islets 22 and 23 genes were respectively found to be differentially expressed between the C3H and NZO mice. In gWAT 10 genes were upregulated in the NZO mice and 12 in the C3H while in pancreatic islets 16 genes were overexpressed in the NZO strain and 7 in C3H (Figure 33).



Gene ID	Mb position	Expression Ratio C3H/NZO Liver	Quadr	gWAT	BAT	Islets
Slc30a8	52.30	0.97	0.92	1.13	1.00	1.76
Ext1	53.34	0.91	1.15	1.05	1.41	0.95
Tnfrsf11b	54.25	0.90	0.87	1.21	0.94	0.61
Sntb1	55.64	1.00	1.03	1.13	1.50	0.91
Has2	56.67	0.94	1.10	1.13	1.02	0.69
Derl1	57.87	0.96	1.04	0.94	1.45	0.99
Tbc1d31	57.96	1.27	1.23	1.10	1.34	1.17
Klhl38	58.31	1.02	1.51	1.00	1.02	1.11
Anxa13	58.34	0.73	0.97	1.09	0.98	0.73
Myc	61.99	0.87	0.88	0.70	0.71	0.62
Lrrc6	66.38	0.97	0.96	1.70	0.94	1.58
Tmem71	66.53	0.87	0.85	0.75	1.03	0.58
Sla	66.78	0.85	0.93	0.64	0.90	1.00
Wisp1	66.89	0.89	0.97	1.00	1.03	0.70
Ndrg1	66.93	0.90	1.11	1.81	1.22	0.93
St3gal1	67.10	0.68	0.94	0.96	1.14	0.87
Gm25987	68.67	0.82	0.68	0.96	0.84	0.97
Ptp4a3	73.75	1.03	0.78	1.34	1.01	0.82
Ly6d	74.76	0.89	0.92	1.17	0.99	0.59
Ly6e	74.96	0.65	0.94	0.98	0.71	0.64
Ly6a	74.99	0.12	0.75	0.78	0.77	0.40
Ly6c1	75.05	0.25	0.91	0.91	0.86	0.78
9030619P08Rik	75.43	0.98	0.92	1.53	0.94	1.05
Ly6h	75.56	1.03	0.91	1.13	0.97	1.38
Gpihbp1	75.60	1.16	0.90	1.34	0.89	1.03
Mafa	75.75	1.14	1.08	0.98	1.07	1.43
Mapk15	76.00	1.08	0.96	1.52	0.95	1.08
Nrbp2	76.09	0.91	1.01	1.84	1.04	0.89
Adck5	76.59	1.06	1.07	1.06	1.21	1.59
Slc39a4	76.61	1.26	0.90	1.15	0.94	0.71
Gpt	76.70	1.04	0.93	1.34	0.99	1.03
Lrrc24	76.72	1.09	0.91	1.05	1.01	1.72
Mb	77.02	1.19	0.97	0.89	1.57	0.97
Apol6	77.05	0.96	1.39	1.01	0.74	1.04
Apol7a	77.39	1.43	1.00	1.07	0.94	1.04
Apol9a	77.40	7.74	1.06	1.03	1.04	1.01
Foxred2	77.94	1.06	0.96	1.43	1.01	0.99
Pvalb	78.19	1.11	0.99	0.52	2.33	0.99
Ncf4	78.24	0.90	1.00	0.55	1.02	0.84
Csf2rb2	78.28	0.30	0.50	0.20	0.40	0.18
Csf2rb	78.35	1.15	0.82	0.56	0.52	0.68
Tst	78.40	1.21	1.11	1.63	1.94	0.97
Rac2	78.56	0.93	0.89	0.50	0.90	0.97
Cyth4	78.60	0.97	0.98	0.51	0.89	0.85
Cdc42ep1	78.85	1.02	1.17	1.53	1.07	0.69
Lgals2	78.85	0.94	1.03	1.26	1.03	0.16
Tmem184b	79.36	1.17	1.04	1.18	1.43	0.95
Kdelr3	79.52	0.99	0.97	1.02	1.04	0.69
Apobec3	79.89	0.94	0.78	0.78	0.70	0.61
Syngr1	80.09	0.94	1.02	0.71	0.97	0.98
Mgat3	80.21	0.99	0.95	1.02	0.92	0.71
Fam83f	80.67	0.64	0.96	0.76	0.95	0.90
Mchr1	81.24	1.00	0.65	1.02	0.97	1.02
Slc25a17	81.32	1.17	1.39	1.42	1.86	2.00

Figure 33. Differentially expressed genes annotated in the *Nbg15* locus. Relative expression ratio (C3H/NZO) of genes located at *Nbg15* (50-80 Mb) in the liver, quadriceps, gWAT, BAT and pancreatic islets. Genes are assemble as Mb position (ENSEMBL). Significant differential gene expression is marked in red (overexpressed in C3H) and blue (overexpressed in NZO). Significant results were considered when $p<0.05$, one-sided Wilcoxon signed rank test. Quadr= quadriceps, gWAT= gonadal white adipose tissue, BAT= brown adipose tissue, Mb= mega base pairs.

3.5.1.1 Validation of *Nbg15* differential expressed genes via real time qPCR

Genes found to be significantly differentially expressed within the parental NZO and C3H in the microarray analysis were further selected for validation via qPCR. Thus, as mentioned above, gene expression was evaluated in the tissues collected from young (6 weeks of age) and adult (21 weeks of age) mice to assure that expression was not changed over time, with exception of pancreatic islets which were only collected from young mice. The validation of the gene expression was considered when the expression was also found to be significantly different between NZO and C3H via qPCR and followed the direction of regulation found in the microarray analysis.

The results revealed that out of 8 genes, 3 could be validated in the liver (*Apol7a*, *Csf2rb2*, *Fam83f*), 2 genes out of 5 were validated in quadriceps (*Apol6, Mchr1*). In the gWAT the genes *Csf2rb2* and *Slc25a17* could be validated from a total of 22 genes. In BAT, the genes *Csf2rb2*, *Tst* and *Tmem184b* were validated, representing 3 genes out of 15. Additionally, 8 genes out of 23 could be validated in Islets (*Slc30a8*, *Has2*, *Ly6d*, *Adck5*, *Lrrc24*, *Csf2rb2*, *Cdc42ep1*, *Kdelr3*). Thus, a total of 15 genes were selected as likely candidate genes for *Nbg15* based on the expression criteria. The majority of the selected genes were originated from the analysis in pancreatic islets. However, the lack of analysis in islets of adult mice might be the reason for the higher number of validated genes in this tissue. To add, less genes were differentially expressed at later stages of age compared to the young mice. This can be used as criteria to narrow down candidate genes as it represents a stable genetic regulation among the development of the phenotypic characteristics that protect the C3H mice from developing diabetes or elevate the T2D risk in the NZO mice (Table 13).

Table 13. Validated candidate genes of the *Nbg15* locus via qPCR in the young and adult parental C3H and NZO mice. Results are represented as relative gene expression ($2^{-\Delta\Delta Ct}$) ratio C3H/NZO using *Actb* as housekeeping gene for qPCR validation and as relative gene expression for microarray analysis. The genes are indexed according to the different analyzed tissues and Mb position on chromosome 15. Significant differentially expressed genes were considered when *p<0.05, **p<0.01, ***p<0.001 using two-tailed Unpaired t-test, n=6-8. QUADR= Quadriceps, BAT= brown adipose tissue, gWAT= gonadal white adipose tissue, Mb-Position= mega base pair position.

Age			6 Weeks		21 Weeks
Tissue	**Mb position**	**Gene**	**Ratio C3H/NZO Microarray**	**Ratio C3H/NZO qPCR_2(-ΔΔCt)**	**Ratio C3H/NZO qPCR_2(-ΔΔCt)**
LIVER	77.4	*Apol7a*	1.4300	6.2506 ***	6.9410 ***
	78.3	*Csf2rb2*	0.3020	0.0112 ***	0.0041 ***
	80.7	*Fam83f*	0.6350	0.6115 *	0.1255 **
QUADR	77.0	*Apol6*	1.3900	2.0440 *	1.5517 *
	81.2	*Mchr1*	0.6460	0.2224 **	0.2119 **
gWAT	78.3	*Csf2rb2*	0.2000	0.0002 **	0.0001 **
	81.3	*Slc25a17*	1.4200	2.3748 ***	2.7362 ***
BAT	78.3	*Csf2rb2*	0.4000	0.0116 **	0.0011 *
	78.4	*Tst*	1.9400	3.4113 **	3.2560 ***
	79.4	*Tmem184b*	1.4300	1.4467 p=0.09	2.8030 ***
ISLETS	52.2	*Slc30a8*	1.7600	2.8519 *	
	56.7	*Has2*	0.6850	0.1053*	
	74.8	*Ly6d*	0.5930	0.1132 *	
	76.6	*Adck5*	1.5900	2.5058 p=0.08	
	76.7	*Lrrc24*	1.7200	1.8617 p=0.08	
	78.3	*Csf2rb2*	0.1830	0.0002 p=0.05	
	78.8	*Cdc42ep1*	0.6890	0.2465 p=0.06	
	79.5	*Kdelr3*	0.6890	0.3916 *	

3.5.2 Investigation of expression Quantitative Trait Loci (eQTL) of the *Nbg15* candidate genes

A first selection of genes was based on the expression profiling conducted in the parental C3H and NZO mice. However, genetic regulatory networks are very complex and diverse between different mouse models. Therefore, an additional approach was used to investigate whether genomic regulatory regions match to the locus where the selected candidate genes were found. For this purpose, the mRNA transcript levels of the selected candidate genes were measured in different tissues of the male N2 backcross population (N_2(C3HXNZO)). Furthermore, the expression levels were used as quantitative traits and linked to the available N_2(C3HXNZO) genotypes for the calculation of the expression quantitative trait loci (eQTL). Using this criteria, the genes that show a *cis*-eQTL, meaning that if the locus where the gene is

located explains the fraction of its variance, were further selected as likely candidates for the *Nbg15* locus. Moreover, the genes exhibiting a *trans*-QTL or which associations to expression were not located within the locus where the gene is located were therefore considered as less likely candidate for *Nbg15*. The associations were represented by the LOD Score, being a high LOD Score an indication of greater association between genomic regions and gene expression or an indication of linkage.

In addition, the gene expression levels were correlated to phenotypic traits measured in the backcross population. Association with islets could not be performed as no pancreatic islets were collected from the backcross population.

3.5.2.1 Investigation of gene eQTL and correlation to phenotypic traits in the liver

The results of the gene expression profiling in the liver led to the validation of the genes *Apol7a*, *Csf2rb2* and *Fam83f*. Furthermore, the calculations of linkage showed a significant cis-eQTL for the gene *Apol7a* and *Cs2rb2* while for *Fam83f* only a *trans*-eQTL on chromosome 14 was detected (Figure 34A-C). The highest linkage was found for the *Apol7a* gene. A LOD score of 23.8 at 25 cM was detected for the *Apol7a* when significance threshold equaled to 2.4 LOD. In addition, the peak also matched to the gene-annotated region (77.4 mb) (Figure 34A). Similar results were obtained for the *Csf2rb2* gene which annotated peak region reached a LOD score of 5.5 located at 28 cM (Figure 34B). As genomic regulatory regions of the *Fam83f* gene were predicted to be located at another locus on chromosome 14, this gene was excluded as candidate (Figure 34C).

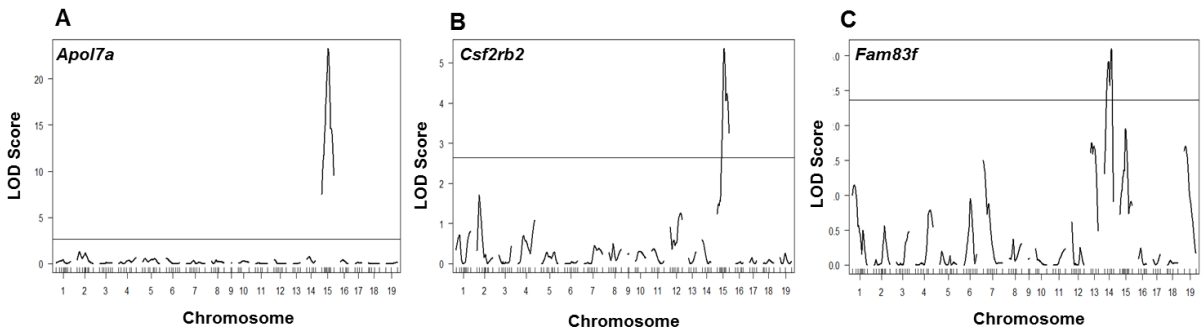


Figure 34. Genome-wide eQTL analysis of *Nbg15* candidate genes identified in the liver. The horizontal line in the graphs represent the significance threshold for the calculated eQTL in the liver. Significant results were determined by 100 permutation test using R/qtl software. LOD= Logarithm of the odds.

Both genes *Apol7a* and *Csf2rb2* are annotated at the region mapped to the distal peak of the *Nbg15* locus (Figure 35A). Furthermore, mRNA transcript levels were segregated according to the peak-nearest available genotyping marker (rs4230879 – 79 mb). Mice carrying the C3H allele at this position had significantly higher mRNA levels of the *Apol7a* gene compared to NZO allele carriers (Figure 35B). This result was in accordance to the expression of *Apol7a* obtained in the liver of the parental C3H and NZO mice (Table 13). The results obtained for the *Csf2rb2* gene also showed a similar pattern as observed in the parental mice with the *Csf2rb2* gene being upregulated in mice carrying the NZO alleles compared to the C3H allele carriers (Figure 35C).

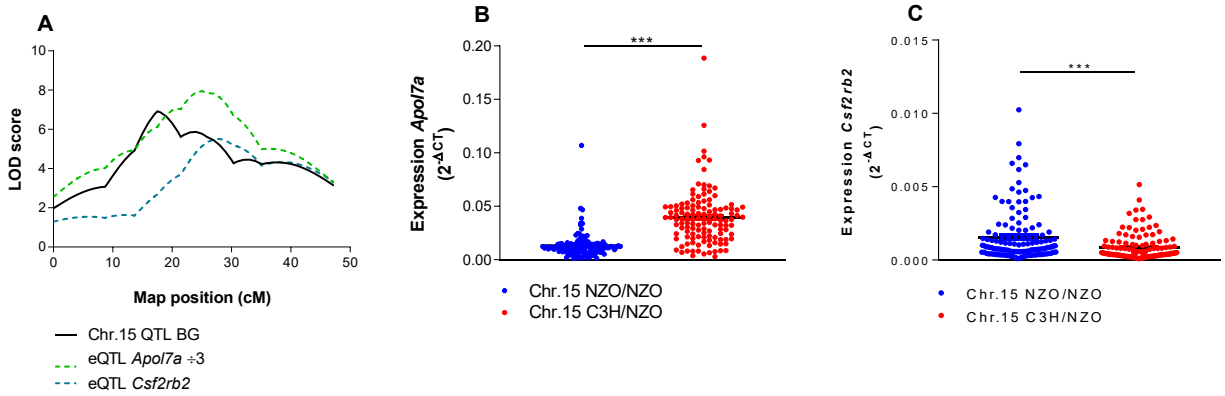


Figure 35. Chromosome 15 *cis*-eQTL in the liver and genotype dependent gene expression. Comparison of the detected chromosome15 *cis*-eQTL overlaid with the QTL on blood glucose (A). LOD score values of the eQTL for *Apol7a* was divided by three for better visualization (A). Allelic segregation of the *Apol7a* expression using the nearest genotype marker (*rs4230979*) of the eQTL peak (B). Allelic segregation of the *Csf2rb2* expression using the closest genotype marker (*rs4230979*) to the eQTL peak (C). The eQTL analysis was performed using R/qtl (100 permutation test). For the gene expression comparisons the two-tailed unpaired t-test was used. Statistically significant results were considered when *p<0.05, **p<0.01, ***p<0.001, (n= 140-186). QTL= Quantitative trait loci, LOD= Logarithm of the odds. Chr.15 QTL BG – Blood glucose QTL on chromosome 15, eQTL= Expression quantitative trait loci.

In the following, the mRNA transcripts levels of the *Apol7a* and *Csf2rb2* genes were correlated to related phenotypic traits measured in the backcross population. The results revealed no correlation between the *Apol7a* the metabolic traits of the N2 population. On the other hand, the *Csf2rb2* gene was negatively correlated with plasma free-fatty acid levels (Table 14).

Table 14. Correlations of the gene expression in the liver and metabolic traits of the backcross N_2(C3HxNZO) population. The correlations are indicated by R Square values (r^2) and calculated p-value. Statistical significance was calculated by linear regression. Significantly negative correlations are indicated in bold blue color and positive correlations in bold red color. 21W = 21 weeks of age; BMI=Body mass index, TG=triglycerides, FFA=Free fatty acids.

Correlation expression *vs* phenotype				
	Genes			
	Apol7a		*Csf2rb2*	
Traits	**R Square**	**P value**	**R Square**	**P value**
Blood glucose (21W)	0.00465	0.27680	0.00453	0.28440
Insulin (21W)	0.00001	0.95540	0.00013	0.86180
Body weight (21W)	0.00278	0.41240	0.00014	0.85410
BMI (21W)	0.00120	0.59050	0.00008	0.88980
Plasma cholesterol (21W)	0.00649	0.20080	0.00103	0.64000
Plasma TG (21W)	0.00445	0.29440	0.00139	0.55940
Plasma FFA (21W)	0.00762	0.16390	**0.01945**	**0.02600**
Liver TG (21W)	0.00119	0.58220	0.00430	0.29680
Liver weight (21W)	0.00069	0.67740	0.00482	0.27120
Liver glycogen (21W)	0.00000	0.99660	0.00118	0.60060

3.5.2.2 Investigation of gene eQTL and correlation to phenotypic traits in the quadriceps

In the gene expression profiling of the quadriceps the genes *Apol6* and *Mchr1* were selected as potential candidates. The further performed eQTL analysis revealed that both genes exhibited highly significant *cis*-eQTLs (Figure 36A). The *Apol6* gene exhibited a maximal LOD score of 15.7 at 26 cM and *Mchr1* eQTL revealed a LOD score of 6.9 at 25 cM. Both peaks matched to the genomic region were the genes are annotated and corresponded to the distal peak of the *Nbg15* locus (Figure 36B). Further allelic segregation of the gene expression demonstrated that *Apol6* expression is elevated in mice carrying the C3H alleles at the eQTL peak position in comparison to the NZO mice. On the other hand, the allelic segregation of the expression of the *Mchr1* gene showed that mice carrying the NZO alleles at the eQTL peak overexpressed *Mchr1* in comparison to the C3H mice (Figure 36C-D). For both genes the expression pattern observed in the parental mice could be reproduced in the allelic segregation of the backcross population.

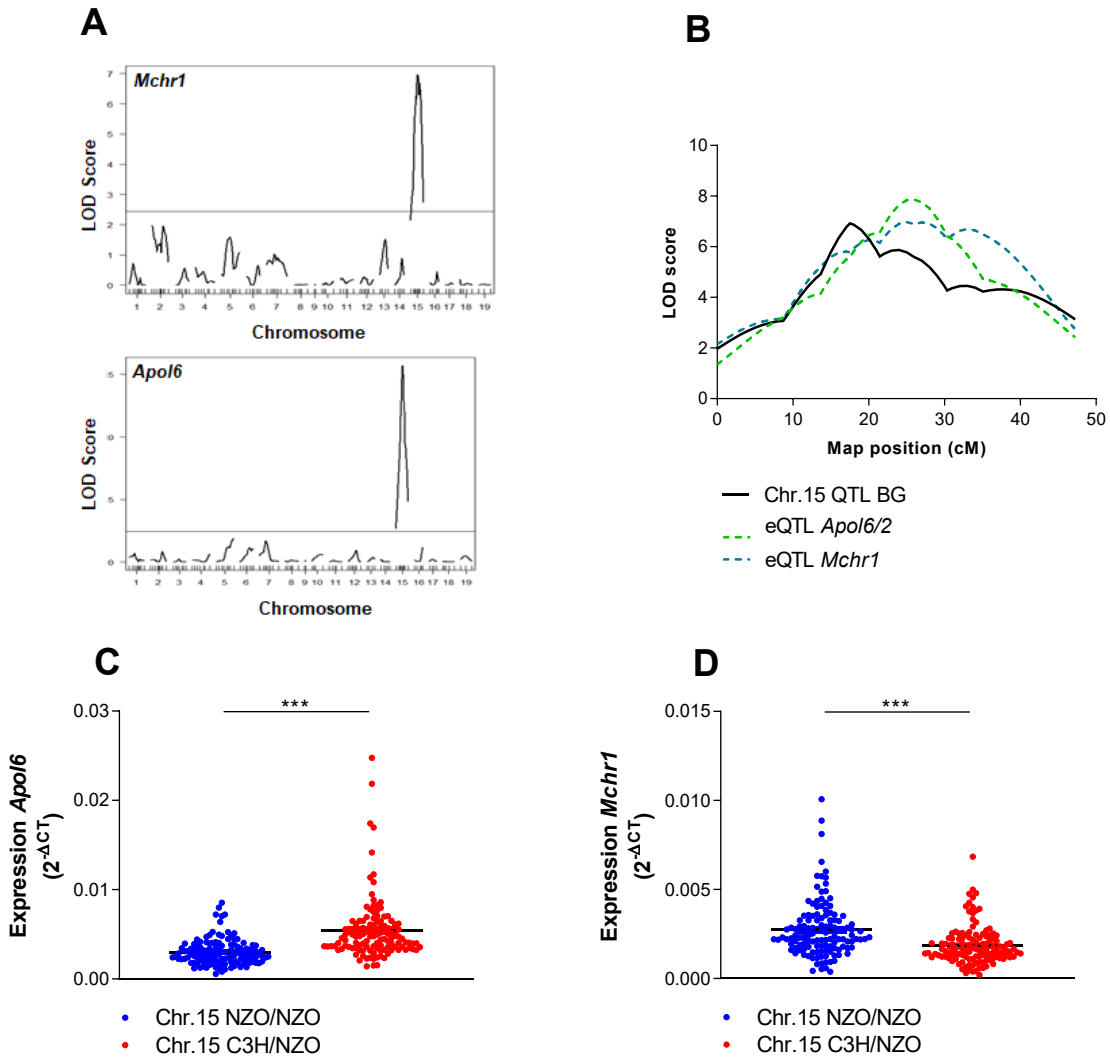


Figure 36. Chromosome 15 *cis*-eQTL in the quadriceps and genotype dependent gene expression. eQTL analysis of the candidate genes identified in the quadriceps of the N_2(C3HxNZO) population (A). Horizontal lines in the A graphs represent the significant threshold for the *cis*-eQTL analysis in quadriceps. The LOD score values of the *Apol7a* was divided by two for better visualization (B). Allelic segregation of the *Apol6* expression using the closest genotype marker (*rs4230879*) in the eQTL peak (B). Allelic segregation of the *Mchr1* expression using the nearest genotype marker (*rs4230879*) of the eQTL peak (C). The eQTL analysis was performed using R/qtl (100 permutation test). For the gene expression comparisons the two-tailed unpaired t-test was used. Statistically significant results were considered when *p<0.05, **p<0.01, ***p<0.001, (n= 140-186). QTL= Quantitative trait loci, LOD= Logarithm of the odds. Chr.15 QTL BG – Blood glucose QTL on chromosome 15, eQTL= Expression quantitative trait loci.

In the sequence, the correlations of the *Apol6* gene to the metabolic traits of the N_2(C3HxNZO) were performed and revealed that blood glucose levels, body weight and plasma triglycerides were negatively correlated to the expression of the *Apol6* gene. On the other hand, plasma insulin was positively correlated to *Apol6* gene expression. Similarly, *Mchr1* expression levels were positively correlated to plasma insulin levels (Table 15).

Table 15. Correlations of the gene expression in the quadriceps and metabolic traits of the backcross N_2(C3HxNZO) population. The correlations are indicated by R Square values (r^2) and calculated p-value. Statistical significance was calculated by linear regression. Significantly negative correlations are indicated by bold blue color and positive correlations in bold red color. 21W = 21 weeks of age; BMI=Body mass index, TG=triglycerides, FFA=Free fatty acids.

Correlation expression *vs* phenotype				
	Genes			
	Apol6		*Mchr1*	
Traits	**R Square**	**P value**	**R Square**	**P value**
Blood Glucose (21W)	**0.03774**	**0.0019**	0.004075	0.3128
Insulin (21W)	**0.02556**	**0.0135**	**0.01969**	**0.0304**
Body weight (21W)	**0.01525**	**0.0556**	0.002011	0.4884
BMI (21W)	0.01393	0.0674	0.0002675	0.8006
Plasma cholesterol (21W)	0.008973	0.1345	0.001731	0.5118
Plasma TG (21W)	**0.02641**	**0.0108**	0.001364	0.5651
Plasma FFA (21W)	0.008904	0.1352	0.0004056	0.7504
Lean mass (21W)	0.006457	0.2027	0.0001153	0.8651

3.5.2.3 Investigation of gene eQTL and correlation to phenotypic traits in the gWAT

In gWAT, *Csf2rb2* and *Slc25a17* were selected as likely candidates for *Nbg15* based on the expression profiling criteria. Further determination of their expression in the backcross N_2(C3HXNZO) males followed by association to the available genotypes revealed a significant *cis*-QTL on chromosome15 for both genes, highlighting *Csf2rb2,* which presented a maximal LOD score of 23 at 24 cM, being 2.6 LOD score the calculated significance threshold. A smaller link was found for *Slc25a17*, being 2.9 the maximal LOD score detected at 27 cM (Figure 37A) (Pelligra 2016);(Updated analysis of the Bachelor thesis from Angela Pelligra, 2016).

For both *Csf2rb2* and *Slc25a17* the eQTL peak matched to their coding region in the distal subloci of *Nbg15* (Figure 37B). Furthermore, mRNA transcript levels were segregated according to the peak-nearest genotyping marker (rs4230879 – 79 mb). As observed in the parental mice, *Csf2rb2* was downregulated in mice carrying the C3H alleles compared to the mice carrying the NZO alleles (Figure 37C). However, an opposite effect was observed for *Slc25a17*. The parental C3H mice exhibited elevated mRNA levels of *Slc25a17* compared to the NZO mice while in the backcross population NZO allele carriers mice had higher expression levels of *Slc25a17* compared to the mice carrying the C3H alleles (Figure 37D).

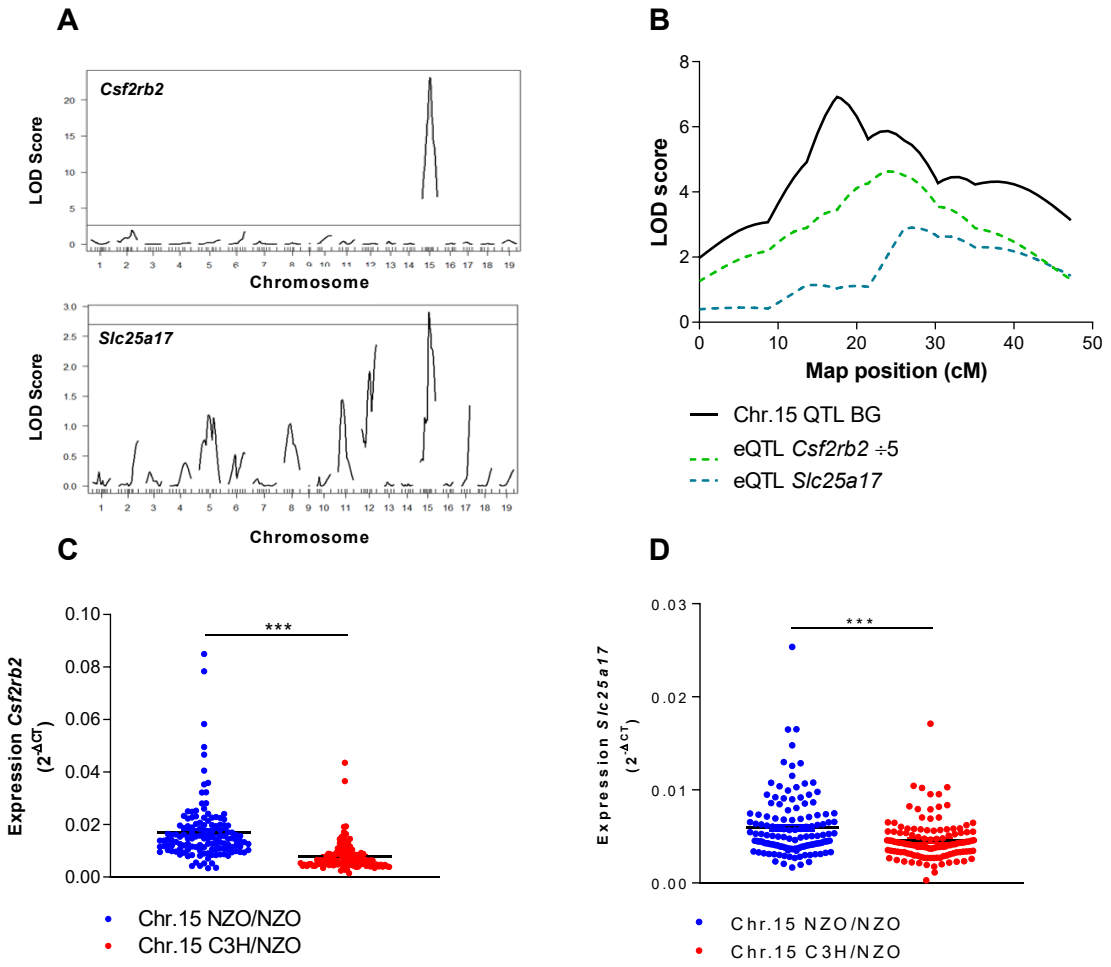


Figure 37. Genome-wide *cis*-eQTL analysis of *Nbg15* candidate genes identified in gWAT. Horizontal line in the graphs (A) represent the significant LOD threshold for the *cis*-eQTL in gWAT. To compare the LOD score distribution, all *cis*-eQTLs in gWAT were overlaid with the QTL for blood glucose on chromosome 15 (B). The LOD score values of the eQTL for the *Csf2rb2* gene was divided by five for better visualization (B). Allelic segregation of the *Csf2rb2* gene expression using the nearest genotype marker (*rs4230879*) in the eQTL peak (C). Allelic segregation of *Slc25a17* expression using the closest genotype marker (*rs4230879*) in the eQTL peak (D). The eQTL analysis was performed using R/qtl (100 permutation test). For the gene expression comparisons the two-tailed unpaired t-test was used. Statistically significant results were considered when *p<0.05, **p<0.01, ***p<0.001, (n= 140-186). QTL= Quantitative trait loci, LOD= Logarithm of the odds. Chr.15 QTL BG – Blood glucose QTL on chromosome 15, eQTL= Expression quantitative trait loci.

Correlations of mRNA transcript levels of the *Csf2rb2* and *Slc25a17* genes to phenotypic traits of the backcross population revealed a positive correlation of *Csf2rb2* to blood glucose and plasma triglycerides levels. In contrast, for the *Slc25a17* gene a negative correlation was found to plasma insulin levels, body weight, BMI and fat mass. Moreover, expression levels of the *Slc25a17* gene was positively correlated to blood glucose and plasma triglycerides (Table 16).

Table 16. Correlations of gene expression in gWAT and metabolic traits of the backcross N_2(C3HxNZO) population. Correlations are indicated by R Square values (r^2) and calculated p-value. Statistical significance was calculated by linear regression. Significantly negative correlations are indicated in bold blue color and positive correlations in bold red color. 21W = 21 weeks of age; BMI=Body mass index, TG=triglycerides, FFA=Free fatty acids.

Correlation expression *vs* phenotype				
	Genes			
	Csf2rb2		*Slc25a17*	
Traits	**R Square**	**P value**	**R Square**	**P value**
Blood glucose (21W)	**0.08500**	***<0.0001***	***0.03195***	***0.0040***
Insulin (21W)	*0.01454*	*0.0632*	***0.03477***	***0.0039***
Body weight (21W)	*0.00790*	*0.1672*	***0.06814***	***<0.0001***
BMI (21W)	*0.00703*	*0.1926*	***0.07137***	***<0.0001***
Plasma cholesterol (21W)	*0.00508*	*0.2567*	*0.01359*	*0.0630*
Liver TG (21W)	*0.00902*	*0.1303*	*0.00689*	*0.1866*
Plasma TG (21W)	***0.01956***	***0.0270***	***0.08622***	***<0.0001***
Plasma FFA (21W)	*0.00097*	*0.6200*	*0.00872*	*0.1354*
Fat mass (15W)	*0.00001*	*0.9682*	***0.06107***	***<0.0001***

3.5.2.4 Investigation of gene eQTL and correlation to phenotypic traits in the BAT

In brown adipose tissue the *Csf2rb2*, *Tst* and *Tmem184b* were selected as candidate genes. Expression QTL analysis conducted with these genes resulted in two significant *cis*-eQTL, for the *Csf2rb2* and *Tst* gene, while for the *Tmem184b* gene no significant eQTL was identified (Figure 38A-C). The eQTL of the *Csf2rb2* gene exhibited a peak located at 24 cM with a LOD score of 6.8, when the calculated significant threshold was set at 2.6 LOD (Figure 38A). However, the strongest linkage was found for the *Tst* gene. The eQTL peak for the *Tst* gene was identified at 26 cM and exhibited a LOD score of 9.2 being 2.7 LOD the significance threshold for this gene (Figure 38B). Moreover, the *Tmem184b* gene was excluded as candidate, since no link to expression regulation could be found on chromosome15 (Figure 38C).

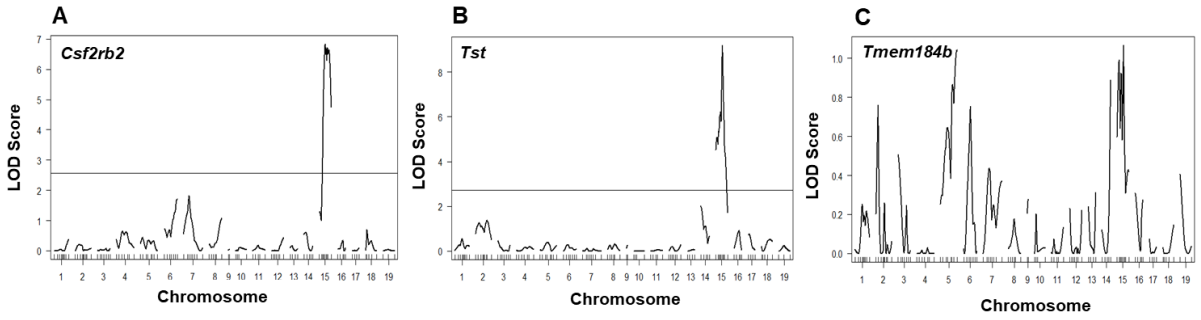


Figure 38. Genome-wide eQTL analysis of *Nbg15* candidate genes identified in BAT. Horizontal line in the graphs represent the significant LOD threshold for the identified *cis*-eQTL in BAT. Significant results were determined by 100 permutation test using R/qtl software.

For both *Csf2rb2* and *Tst* genes the expression should be regulated by the same genomic region were they are located, also matching to the distal subloci of *Nbg15*. However, the *Csf2rb2* gene exhibited a larger peak overlaying with further genomic regions on chromosome 15 (Figure 39A). The allelic segregation of the *Csf2rb2* expression revealed its downregulation in mice carrying the heterozygous C3H alleles on chromosome 15 compared to NZO allele carriers at the same position, results that corroborate with the findings in the parental mice (Figure 39B). The allelic segregation of the *Tst* gene demonstrated that the gene was overexpressed in mice carrying the heterozygous C3H alleles on chromosome 15 compared to the mice carrying the NZO alleles, validating the observation when *Tst* expression was measured in the parental C3H and NZO mice (Figure 39C).

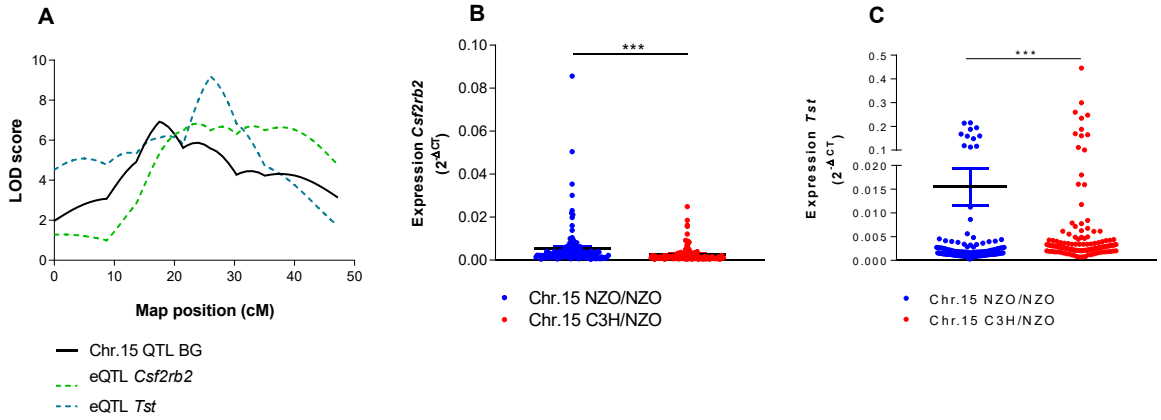


Figure 39. Genome-wide gene eQTL analysis in BAT and genotype dependent gene expression. Comparison of the detected *cis*-eQTL for the BAT candidate genes overlaid with the QTL on blood glucose (A). Allelic segregation of the *Csf2rb2* gene expression in BAT using the nearest genotype marker (*rs4230879*) in the eQTL peak (B). Allelic segregation of *Tst* expression based on the nearest EQTL genotype marker (*rs4230879*) (C). The eQTL analysis was performed using R/qtl (100 permutation test). For the gene expression comparisons the two-tailed unpaired t-test was used. Statistically significant results were considered when *p<0.05, **p<0.01, ***p<0.001, (n= 140-186). QTL= Quantitative trait toci, LOD= Logarithm of the odds. Chr.15 QTL BG – Blood glucose QTL on chromosome 15, eQTL= Expression quantitative trait loci.

Moreover, the expression of the *Csf2rb2* gene in BAT of the N2 backcross population was positively correlated to blood glucose levels and BMI and negatively correlated to plasma cholesterol and FFA. On the other hand, the *Tst* gene expression was positively correlated to blood glucose levels, while a negative correlation was found to plasma insulin levels (Table 17).

Table 17. Correlations of gene expression in BAT and metabolic traits of the backcross N_2(C3HxNZO) population. Correlations are indicated by R Square values (r^2) and calculated p-value. Statistical significance was calculated by linear regression. Significantly negative correlations are indicated in bold blue color and positive correlations in bold red color. 21W = 21 weeks of age; BMI=Body mass index, TG=triglycerides, FFA=Free fatty acids.

Correlation expression *vs* phenotype				
	Genes			
	Csf2rb2		*Tst*	
Traits	**R Square**	**P value**	**R Square**	**P value**
Blood glucose (21W)	**0.02015**	**0.02260**	**0.01499**	**0.04950**
Insulin (21W)	0.00609	0.21900	**0.03045**	**0.00570**
Body weight (21W)	0.00517	0.26350	0.00979	0.12330
BMI (21W)	**0.18710**	**<0.0001**	0.00018	0.83050
Plasma cholesterol (21W)	**0.06426**	**<0.0001**	0.00621	0.20890
Plasma TG (21W)	0.00000	0.98420	0.00020	0.82250
Plasma FFA (21W)	**0.02221**	**0.01660**	0.00006	0.90150

3.6 Identification of C3H and NZO polymorphic regions and predictions of gene transcriptional impact

Another criteria used to narrow down *Nbg15* candidate genes was the identification of polymorphisms between C3H and NZO parental mice. Considering that these inbreed strains descend from a common ancestor, the identification of genomic region containing specific C3H or NZO polymorphic regions were therefore assumed to have a higher likelihood to harbor the causal gene variants for the *Nbg15* locus. For this purpose, available data of single nucleotide polymorphisms (SNP) annotated in the *Nbg15* locus were collected from Wellcome Trust Sanger data base of SNPs. Furthermore, the frequency of SNPs that were different between NZO and C3H, defined as non-identical by descendent (non-IBD) SNPs, were quantified within intervals of 150 kbp (kilobase pairs). These were further compared to the total annotated SNPs counted in each interval when compared to the reference C57BL/6J (GRCm38.p6 assembly) strain.

The SNP analysis revealed that within the interval of 63 to 68 mb 100 % of the total annotated SNPs were different between C3H and NZO mice, corresponding to the region harboring the highest density of polymorphisms in the *Nbg15* locus. Beyond the proximal peak of *Nbg15* (50-71 mb) 71 % of the total SNPs were polymorphic. Similar frequency was found within the distal peak (71-80 mb), where 68 % of the SNPs were different between the C3H and NZO strain (Figure 40).

In addition, a total of 284 genes annotated in the *Nbg15* locus were identified carrying non-IBD SNPs. From these, 90 genes were annotated in the proximal

sublocus of *Nbg15* while the majority or a total of 194 genes were found to be located at the distal region of *Nbg15* locus (Figure 40).

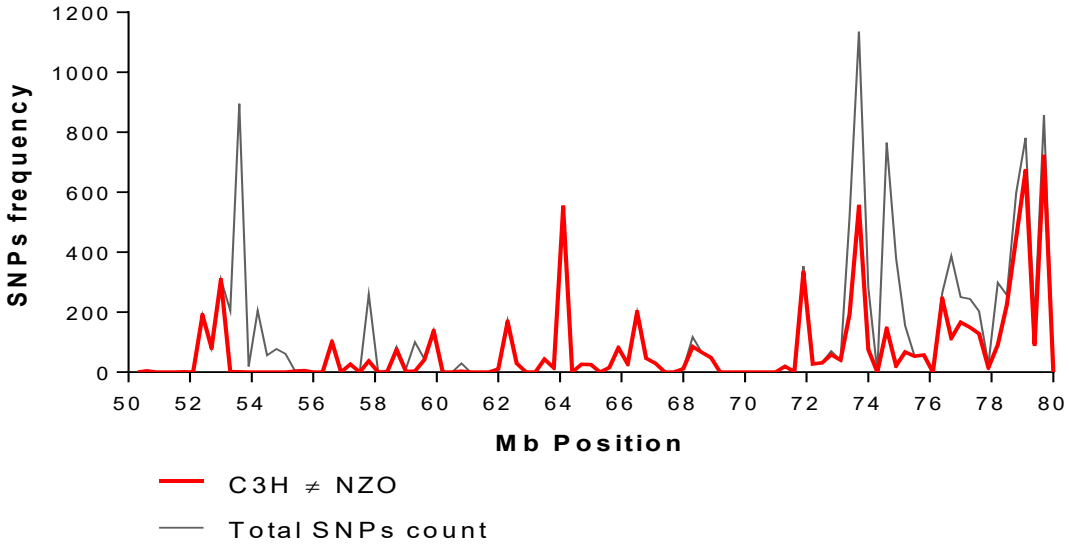


Figure 40. *Nbg15* polymorphic C3H and NZO regions. Frequency of polymorphic SNPs (C3H≠NZO) is indicated in red line considering an interval of 150 kb. Grey line indicates the total number of SNP annotated at the specific interval, considering C57BL/6J (GRCm38.p6 assembly) as reference strain. Mb= mega base pairs.

Additionally, the identified non-IBD SNPs were used for analysis of their functional impact in transcripts variants as well as in the corresponding translated protein. Thus, available dbSNP identifiers for the non-IBD SNPs were used as data input in the so called variant effect predictor analysis (VEP – Ensembl). VEP combines a variety of algorithms to depict SNPs driven amino acids substitution by comparing specific variants to the reference genome (*Mus musculus*-C57BL/6J). Using a based-ruled approach a set of transcripts consequences can be predicted. These are pre-defined consequences by the data base Sequence Ontology (SO) and are mainly based on the gene region which is usually affected.

In a second step this information is used to predict effects on protein function by considering its sequence homology as well as physical properties of the amino acids substitutions. The output consists of a subjective classification of impact, being “high” and “moderate”, more likely to change protein effectiveness, and “low” and “modifier” less likely to change protein function (Table 18). In addition, a so called SIFT (Sorting Intolerant from Tolerant) tool was used to indicate a score of impact in protein function when a qualitative moderate impact was predicted. Using SIFT, variants impact were

scored between 0 to1. A score closer to 0 was more likely to have a deleterious influence on protein function.

Table 18. Classification of SNP impact on gene and protein function. Using the SnpEff (Open Source tool) tool a subjective indication of protein impact was made: high (red), moderate (orange), modifier (blue), low (green). The table presents the description of every impact, which are defined by the tool available at - Wellcome Trust Sanger Institute (UK) via VEP (Variant Effect Predictor analysis).

Functional prediction of impact toolbox (SnpEff)	
High impact	The variant is assumed to have high (disruptive) impact in the protein, probably causing protein truncation, loss of function or triggering nonsense mediated decay.
Moderate impact	A non-disruptive variant that might change protein effectiveness.
Modifier consequence	Usually non-coding variants or variants affecting non-coding genes, where predictions are difficult or there is no evidence of impact.
Low	A variant that is assumed to be mostly harmless or unlikely to change protein behaviour.

Source**:** https://www.ensembl.org/Help/Glossary

The results of the VEP analysis revealed that a total the 284 genes carrying polymorphic SNPs were narrowed down to 65 genes carrying at least one SNP predicted to affect protein behavior. 12 genes out of the 65 were located at the proximal peak of *Nbg15* while the remaining 53 were found within the distal region. Additionally, in 62 genes predicted to affect protein function carried non-IBD SNPs located in exons. Moreover, 101 genes were identified carrying SNPs located in exons with a prediction of low impact on protein function (Figure 41A). The majority of the predicted consequences were classified as “modifier” due to the detection of SNPs between gene upstream or downstream genes (59.23 %). The classifications of “low” impact were mainly defined by synonymous amino acids exchanges (15.77 %). In the following, “moderate” impact was fully predicted by the detection of missense variants (11.73 %). Non-IBD SNPs with a high impact on protein function were so predicted due stop lost or gain codons (0.77 %) that could results in elongated or shortened transcripts, as well as by the detection of spliced variants (0.96 %) (Figure 41B).

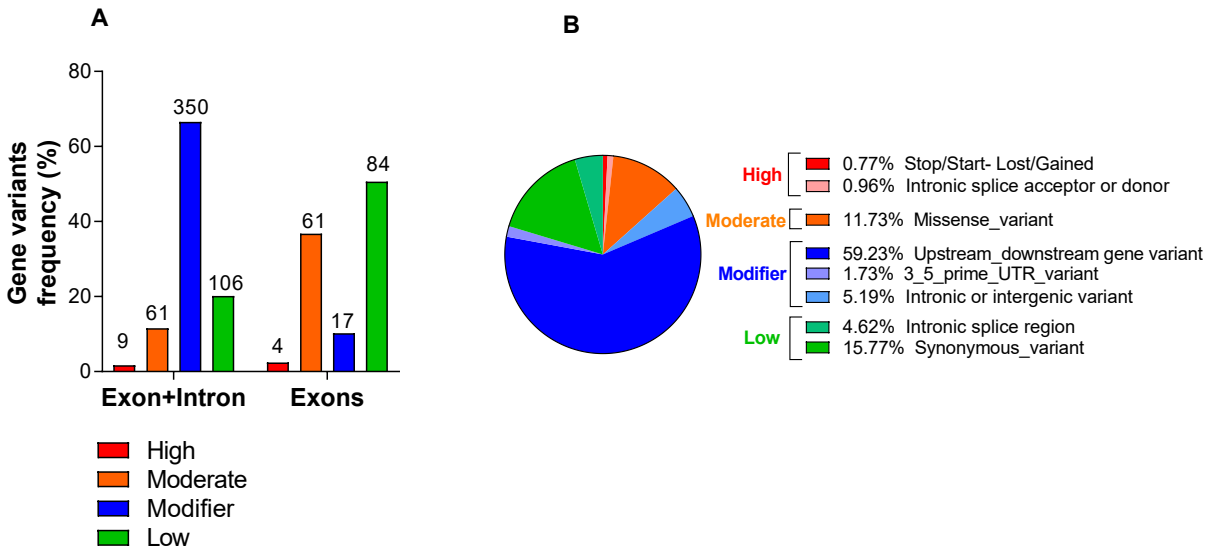


Figure 41. Prediction of functional impact of *Nbg15* non-IBD SNPs.
Frequency of gene variants carrying non-IBD SNPs in introns and exons predicted to affect protein function (A). Frequency of consequences detected for each rate of impact according to Sequence Ontology (SO) terms (B).

Subsequently, gene variants classified with a moderate impact were ranked according to the SIFT score of impact (Ensembl, VEP, 2019). Out of 62 gene variants carrying non-IBD SNPs in exons, 30 were classified with a higher likelihood to impact protein function, by considering SIFT <0.5 (Table 19). The four genes *Col22a1*, *Gsdmc2*, *Adgrb1* and *Csf2rb* were classified as likely deleterious, with a SIFT score of 0-0.01. The SIFT code applies for missense variants only. Therefore, for the genes *Exosc4*, *Fam227a*, *Apobec3* and *Cbx7* no SIFT score was calculated. These were initially classified as variants with "high" impact consequences.

Lastly, the expression profiling strategy was overlaid with the haplotype analysis strategy and revealed 3 genes as potential candidate genes for the *Nbg15* locus. *Apol6* was found to be overexpressed in quadriceps of the parental C3H mice compared to the NZO mice. *Kdelr3* was overexpressed in islets of the 6 weeks old NZO mice in comparison to the C3H mice, while *Apol7a* was identified as a liver candidate and overexpressed in the C3H parental mice when compared to the NZO strain. *Apol6*, *Kdelr3* and *Apol7a* contain polymorphic SNPs leading to a non-synonymous amino acid exchange predicted to affect protein function.

Table 19. *Nbg15* ranking of genes carrying disruptive polymorphisms. Genetic polymorphisms located in exons were ranked according to SIFT score of impact followed by subjective pre-classification of impact using SnpEff tool and prediction of consequences based on SO terms. Genes identified as *Nbg15* candidates by the expression profiling criteria are highlighted in bold red (overexpressed in C3H) and bold blue (overexpressed in NZO). No SIFT was calculated for variants already classified as highly disruptive by the SnpEff tool. SO= Sequence Ontology, SIFT= Sorting intolerant from tolerant

Chromosome15 Position (mb)	Gene	Consequence (SO)	IMPACT (SnpEff)	SIFT
71.8	Col22a1	missense_variant	Moderate	0.00
63.8	Gsdmc2	missense_variant	Moderate	0.01
74.5	Adgrb1	missense_variant	Moderate	0.01
78.3	Csf2rb	missense_variant	Moderate	0.01
76.7	Recql4	missense_variant	Moderate	0.02
77.7	Apol7e	missense_variant	Moderate	0.02
77.1	**Apol6**	**missense_variant**	**Moderate**	**0.04**
79.9	Cbx7	missense_variant	Moderate	0.05
52.0	Aard	missense_variant	Moderate	0.07
76.4	Mroh1	missense_variant	Moderate	0.08
78.9	Sh3bp1	missense_variant	Moderate	0.08
79.5	**Kdelr3**	**missense_variant**	**Moderate**	**0.08**
77.4	**Apol7a**	**missense_variant**	**Moderate**	**0.10**
79.0	Triobp	missense_variant	Moderate	0.10
78.9	Nol12	missense_variant	Moderate	0.11
71.5	Fam135b	missense_variant	Moderate	0.17
75.9	Eef1d	missense_variant	Moderate	0.18
57.9	Tbc1d31	missense_variant	Moderate	0.22
79.8	Cbx6	missense_variant	Moderate	0.22
76.9	Zfp7	missense_variant	Moderate	0.24
79.4	Maff	missense_variant	Moderate	0.25
76.6	Cpsf1	missense_variant	Moderate	0.26
57.3	Slc22a22	missense_variant	Moderate	0.27
79.9	Apobec3	missense_variant	Moderate	0.27
77.5	Apol10a	missense_variant	Moderate	0.34
66.5	Tmem71	missense_variant	Moderate	0.35
58.9	Trmt12	missense_variant	Moderate	0.36
76.9	Zfp251	missense_variant	Moderate	0.38
79.7	Sun2	missense_variant	Moderate	0.38
75.7	Rhpn1	missense_variant	Moderate	0.46
76.3	Exosc4	stop_lost,NMD_transcript_variant	High	–
79.6	Fam227a	start_lost	High	–
79.9	Apobec3	stop_gained	High	–
79.9	Cbx7	stop_gained	High	–

3.7 Association of candidate genes to diabetes protective phenotype of *Nbg15*

As observed in the backcross population, no correlations of *Apol7a* to phenotypic trait could be detected (Table 14). Additionally, although a decreased lean mass could be observed in the RCS.NZO.C3H-*Nbg15* mice, this effect was likely a result of diabetes-driven complications (Figure 17). On the other hand, RCS.NZO.C3H-*Nbg15* phenotyping characterization strongly pointed to an islet protective phenotype addressed by the C3H alleles (Figure 23, 25). Therefore, further results focused in the identification of the identified candidate gene in the pancreatic islets. The gene *Kdelr3*

was the only selected gene in islets that matched all the criteria applied to narrow down the candidates located in the *Nbg15* locus.

3.7.1 Investigation of *Kdelr3* as a candidate gene for *Nbg15*

As revealed by the gene expression profiling, *Kdelr3* was differentially expressed in islets of the parental C3H and NZO mice. This result was specifically observed in islets. *Kdelr3* was not differentially expressed in liver, quadriceps, BAT and gWAT. Additionally, differentially expression was also observed in B6 (C57BL/6J) mice, another known diet-induced diabetes resistant mouse strain (Figure 42A). The *Kdelr3* gene belongs to a gene family composed by another 2 genes, the *Kdelr1* and *Kdelr2* gene. The results obtained with the microarray analysis of the pancreatic islets revealed that among the *Kdel* gene family, the *Kdelr3* gene is the third most abundant gene. Nevertheless, significant differences in expression were only observed for *Kdelr3* while the other members of the family were not differentially expressed in the parental C3H and NZO islets (Figure 42B).

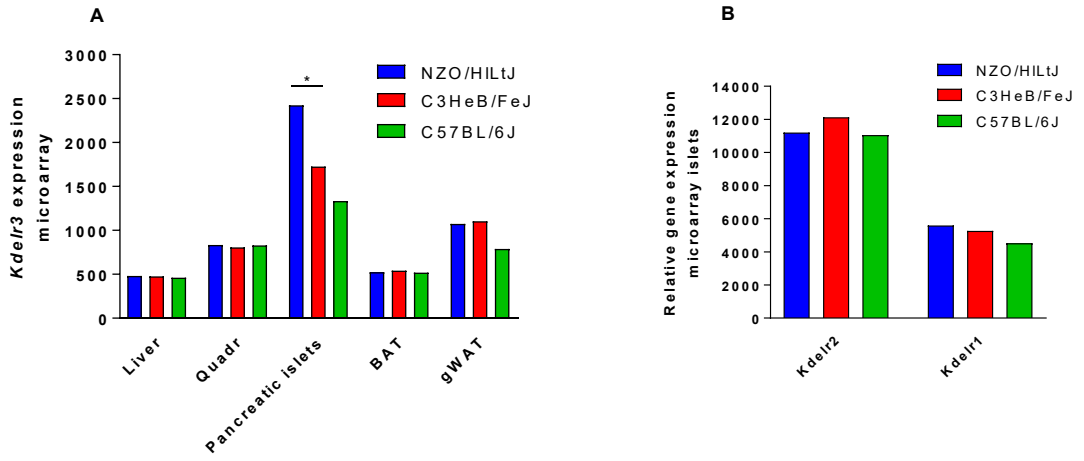


Figure 42. Expression profiling of *Kdelr3* gene and *Kdel* gene family members. Relative expression of *Kdelr3* among different tissues of the parental NZO (NZO/HlLtJ), C3H (C3HeB/FeJ) and B6 (C57BL/6J) mouse strains (A). Expression of *Kdel* gene family members in pancreatic islets of the parental mice (B). Significant results were considered when $p<0.05$, one-sided Wilcoxon signed rank test. Quadr= quadriceps, BAT=brown adipose tissue, gWAT= gonadal white adipose tissue.

The investigation of non-IBD regions between the parental C3H and NZO mice followed by the prediction of the impact in the targeted transcripts led to the identification of a coding missense mutation in *Kdelr3* gene of the NZO mouse strain. To confirm this result, the RNA sequence of the isolated pancreatic islets from a RCS line carrying the NZO alleles on chromosome15 was analyzed by RNASeq. The results

confirmed the presence of an exchange of the base pair adenine to thymine at position 79.52 bp (A79T). As a consequence, a substitution of the amino acid glutamic acid (E-Glu) to valine (V-Val) at position 96 of KDELR3 protein (96-G**A**G-G**T**G) is translated (Figure 43).

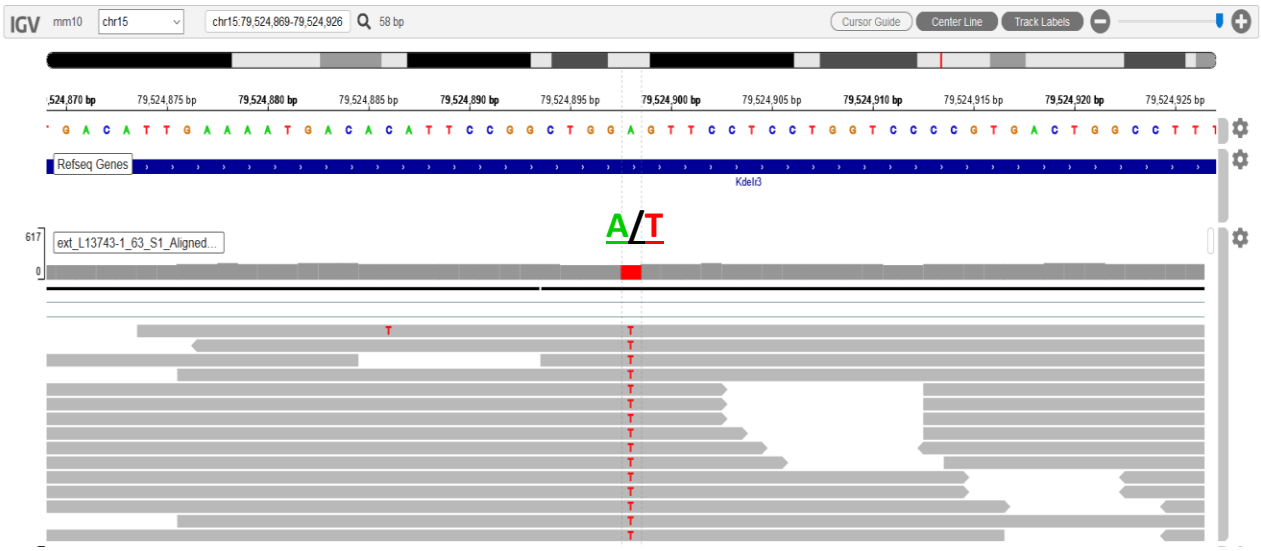


Figure 43. Pancreatic islets exome alignment of the *Kdelr3* gene. Screenshot of IGV - Integrative Genomic Viewer of RNA sequencing from pancreatic islets isolated from RCS.NZO.C3H-*Cdp7*con mice carrying the homozygous NZO alleles on chromosome 15. Each gray box represents a sequence read that is aligned to a reference genome (C57BL/6J) showed in the bottom up. A consistent mutation is detected when present in all reads as shown in red for the detected base pair substitution adenine (A) present in the reference strain to thymine (T) detected in the analyzed sample.

Subsequently, available protein sequence database was used for alignment of KDELR3 across different species. The results showed that the presence of a glutamic acid at position 96 in KDELR3 is highly conserved across several species, including mouse and human, suggestive of an important function of this domain for the protein (Figure 44).

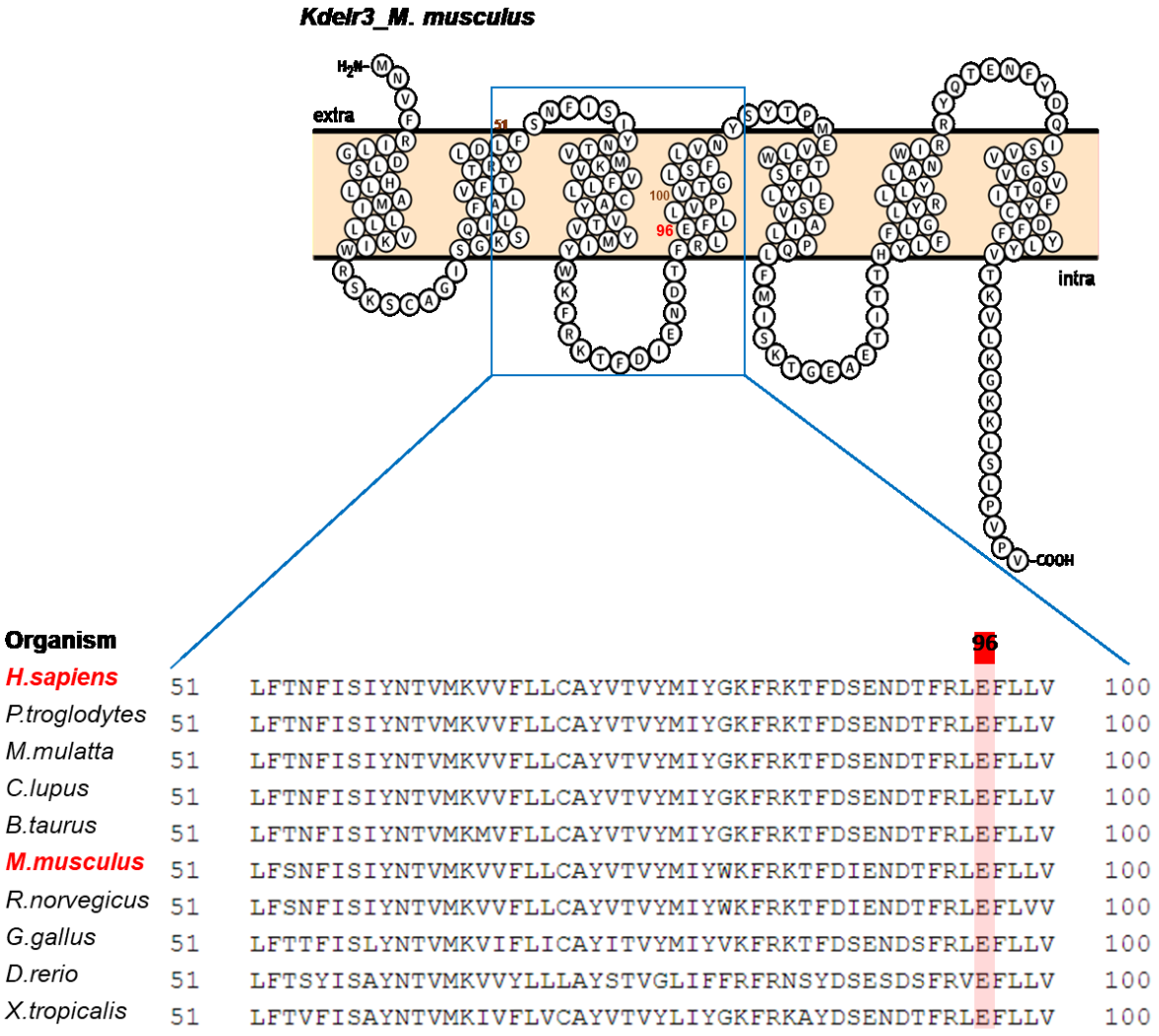


Figure 44. Topological structure and multiple sequence alignment of a KDELR3 domain. Multiple alignment of KDELR3 protein domain from amino acid position 51 to 100. Multiple alignment was performed using NCBI (National Center for Biotechnology Information) BLAST tool for homologous protein alignments. Position 96 highlighted in red represents the conserved amino acid sequence mutated in NZO where a substitution of glutamic acid (E) to valine (V) was identified. Source: NCBI alignment tool: blast.ncbi.nlm.nih.gov/ and Protter-visualize proteoforms: *(Omasits, et al. 2014).*

Furthermore, the N_2(C3HxNZO) backcross population was investigated for the presence of *Kdelr3*-A79T mutation. Homozygous mutants (*Kdelr3*-A79T$^{NZO/NZO}$) displayed significantly elevated blood glucose levels compare to heterozygous mutants (*Kdelr3*-A79T$^{NZO/NZO}$ 391.7 ± 13.05 mg/dl, *Kdelr3*-A79T$^{C3H/NZO}$ 308 ± 12.37 mg/dl, $p<0.001$) (Figure 45A). Moreover, 80 % of the homozygous mutants were diabetic at an age of 20 weeks while 59 % of the heterozygous carriers of *Kdelr3* mutation were diabetic at the same age (Figure 45B). Additionally, plasma insulin levels were significantly lower in the homozygous compared to heterozygous mutants (*Kdelr3*-A79T$^{C3H/NZO}$) (*Kdelr3*-A79T$^{NZO/NZO}$ 2.1 ± 0.16 µg/L, *Kdelr3*-A79T$^{C3H/NZO}$ 2.7 ± 0.17 µg/L, $p<0.05$) (Figure 45C).

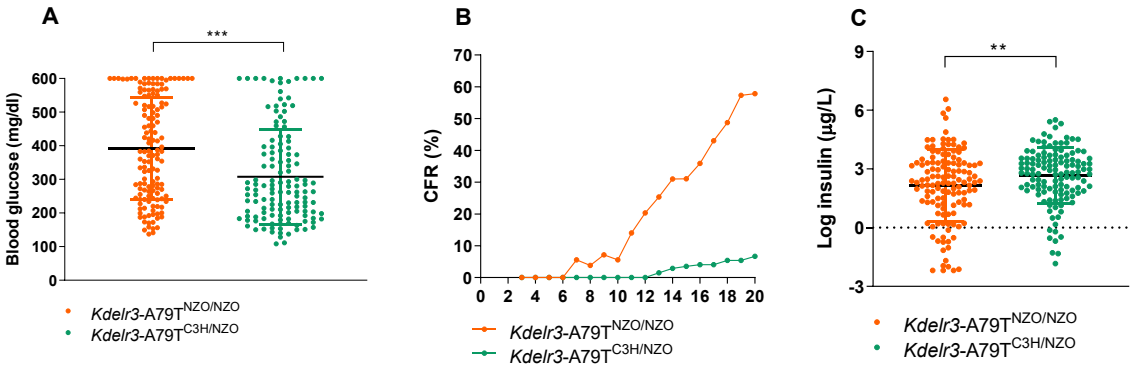


Figure 45. Metabolic features of the *Kdelr3*-A79T mutant. The backcross N_2(C3HxNZO) population was segregated based on the marker for the *Kdelr3*-A79T mutation. Blood glucose levels of *Kdelr3*-A79T mutant (A), Case fatality rate (CFR) (B) and plasma insulin levels (C). *Kdelr3*-A79T$^{NZO/NZO}$ = homozygous mutant, *Kdelr3*-A79T$^{C3H/NZO}$ = heterozygous mutant. The results are expressed as MEAN ± SEM. Statistically significant results were considered when *p<0.05, **p<0.01, ***p<0.001, (n= 140-186) the Student's ttest. Re-analysis of data collected by Tanja Schallschmidt (PhD thesis, 2018), Figure A and C represent the dot-plot representation of random blood glucose and plasma insulin levels measured in the N_2(C3HxNZO) population at 20 weeks of age displayed in Figure 9B and 10D respectively.

3.7.2 Association of *Kdelr3* gene expression in human islets

In collaboration with the research group Genomics Diabetes and Endocrinology leading by Prof. Dr. Ola Hansson at the Lund University (Sweden) *Kdelr3* expression could be assessed in isolated human pancreatic islets (personal communication). Higher expression of *Kdelr3* in human islets was associated to lower insulin levels. Along with this result, a positive correlation was also found with *Kdelr3* expression and the Insulin Amyloid Polypeptide (IAPP), indicating that *Kdelr3* overexpression could be related to increased type-2 diabetes susceptibility

4 DISCUSSION

The pathophysiology of type 2 diabetes (T2D) is strongly influenced by genetic factors. A combination of hundreds of genes is believed to determine individual differences in the susceptibility to develop T2D. However, many of these factors remain unknown or still just account for a small proportion of the estimated genetic heritability of the disease. Thus, this project had the aim to identify novel gene variants that account for the so-called “missing heritability” of T2D. For this purpose, animal models were used as these offer several experimental advantages, which include the genetic homogeneity of the progeny combined with the control of environmental factors, often addressing confounding aspects to the analysis. Thus, the obese and diabetes-prone NZO mouse strain was crossbred with the lean and diabetes-resistant C3H mouse. These mouse strains naturally diverge in their susceptibility towards T2D development and are therefore ideal for the study of the genetic composition of the disease. Moreover, diabetes development in NZO mice strongly resembles the human course of the disease, considered therefore a polygenic model of T2D (Joost and Schurmann 2014).

The generation and metabolic characterization of a N_2(NZOxC3H) backcross population revealed the *Nbg15* locus on chromosome 15 that manly contributes to blood glucose regulation (Schallschmidt 2018). In this study, this locus was further characterized by introgressing genomic fragments containing *Nbg15* from the C3H parental strain into both the NZO and the B6 background by natural breeding strategies, thus generating different recombinant congenic strains (RCS). In this study, the characterization of these RCS was combined with additional strategies including gene expression profiling and haplotype analysis for the identification of potential underlying causal genes for *Nbg15*.

4.1 *Nbg15* harbors two T2D associated subloci

In order to increase the resolution of the *Nbg15* blood glucose QTL, new additional genetic markers were chosen. This approach facilitates the definition of a QTL confidence interval, defined as the region which contains the potential QTL causal genes. This is achieved because the use of dense genotyping markers gives a better information on the recombination frequency, or block of genes that are inherited together. Although in backcrosses progeny the recombination frequency is generally

predicted to be low, genomic regions with high recombination activity are well known among mammalians and can occur among all chromosomes (Laurie, et al. 2005; Paigen and Petkov 2010; Pritchard and Przeworski 2001).

The refined linkage analysis on chromosome 15 revealed two different peaks among *Nbg15* locus. A region located at 63 Mb corresponded to the proximal peak (*Nbg15*-p) of the *Nbg15* locus, while the distal peak (*Nbg15*-d) was found at 74 Mb. Both subloci resulted in an average blood glucose effect size of 92.4 mg/dl. These results represented the largest phenotypic differences between the genotypes, observed at the age of 15 weeks.

Interestingly, the average blood glucose levels at 12 weeks of age reached the diabetes threshold for all genotypes. On the other hand, while the mice carrying the C3H alleles at the proximal and distal *Nbg15* sublocus maintained an average blood glucose level around 300 mg/dl, the NZO allele carriers of *Nbg15* had levels around 400 mg/dl. Moreover, at 12 weeks of age, the diabetes prevalence reached its peak among all genotypes. However, the prevalence of diabetes was strikingly lower among the proximal and distal C3H allele carriers of *Nbg15* in comparison to the NZO carriers of this locus. Interestingly though was the fact that although around 50 % of the mice carrying the C3H alleles at proximal and distal *Nbg15* sublocus were diabetic, only around 3 % of these mice died. Contrarily, among the 77 % diabetic mice in the NZO-*Nbg15* group, around 15 % of the mice died due to severe diabetes (results section 3.1).

These results lead to the revision of the QTL analysis for the insulin levels on chromosome 15. The QTL analysis for the insulin levels failed to reach statistical significance. On the other hand, some technical reasons could have led to this result. One reason is that the measurements of insulin were only obtained at 21 weeks of age. Certainly, the higher mortality rate among the NZO allele carriers of *Nbg15* substantially reduce the statistical power of the analysis as the highly diabetic mice presenting thereby the group with lower insulin levels were technically eliminated from the analysis. Moreover, additional limitations on the analysis itself might underlie the failure in detecting small and single effects that generally mitigate the complex interaction between the phenotypes. In this context, although the differences in insulin levels between the C3H and NZO allele carriers of *Nbg15* were significantly different, the effect size was not large enough at the time point of data collection and/or the variation within individual of the same genotype was too high to statistically link the

plasma insulin levels to the *Nbg15* locus. Therefore, the estimation of insulin levels for the missing samples was the best alternative to solve this research question. The results obtained in this analysis showed that the insulin QTL peak, located at 14 -26 cM, directly matched to the blood glucose peak of the *Nbg15* locus, which is expected given that the regulation of blood glucose is tightly coupled to insulin secretion. Thus, the original QTL for insulin levels, even if not reaching significance, was taken into consideration as a genomic region that contributes to the blood glucose effects.

4.1.1 *Nbg15* is linked to protection from pancreatic islet failure

Despite of the suggestive QTL (LOD<3) detected for plasma insulin levels, the results obtained in the linkage analysis of the backcross N_2(C3HxNZO) point towards a protection from islet dysfunction that is addressed by the C3H alleles on chromosome 15. One evidence for this hypothesis is the finding that the effects on blood glucose conferred by the *Nbg15* locus were dependent on the increased diabetes prevalence among the genotypes. This conclusion is supported by several studies that corroborate with the idea that beta cell demise precedes T2D onset (Roden and Shulman 2019). Therefore, these results are consistent with a sign of islets failure in compensating the insulin resistance in the NZO allele carriers in comparison to the C3H allele carriers of *Nbg15.* In addition, no differences in body weight development could be observed, what is usually expected to be a strong phenotype addressed to the insulin resistance pathogenesis. These results are assured by the fact that the development of obesity is known to be a strong metabolic feature of diabetes development in the NZO strain (Junger et al. 2002; Jurgens et al. 2006). Moreover, linkage studies performed with the obese NZO and the lean B6 and DBA mouse strains demonstrated a strong linkage of obesity-related QTL on chromosome 4, while no significant association were linked to chromosome 15 in any of these crossbreedings (Joost 2012; Junger et al. 2002; Vogel, et al. 2018). In addition, an improved survival rate in the C3H allele carriers on chromosome 15 corroborate with an improved capacity of the pancreatic islets to compensate for an increasing degree of insulin resistance presumably addressed by the genetic NZO background.

Interestingly, another linkage study performed in the Takeshita et al. (2006) laboratory using an outcross population generated between the non-obese type 1 diabetic Akita mice and the diabetes and obesity resistant A/J mouse strain presented a QTL on chromosome 15 for fasting blood glucose with a LOD score of 6.17. Elevated

blood glucose levels during a glucose tolerance test were attributed to an insulin resistance phenotype in the homozygous Akita mice. However, since plasma insulin levels did not change during the test, these results could still be a sign for a failure in insulin secretion as normal insulin resistance compensatory mechanisms would indeed lead to increase insulin secretion in response to elevated blood glucose levels (Hong, et al. 2007; Takeshita, et al. 2012; Takeshita, et al. 2006).

4.1.2 Homozygous C3H allele carriers of *Nbg15* validated the islet-driven phenotype associated to type 2 diabetes protection

4.1.2.1 Islets of the C3H allele carriers of *Nbg15* have improved hyperglycemia-induced metabolic compensation

The backcross N_2(C3HxNZO) population is composed by mice that represent a genetic mosaic of the parental C3H and NZO mice. Therefore, the identified QTL for blood glucose on chromosome 15 could also be influenced by QTLs in other chromosomes. Thus, the contribution of *Nbg15* to blood glucose regulation needed to be validated. This was accomplished with the generation of the recombinant congenic lines carrying the homozygous C3H alleles in the whole, the proximal or the distal *Nbg15* locus, onto the NZO genome. This was successfully achieved by repetitive intercrosses of N5-N6 backcross generation of the N_2(C3HxNZO) population. The phenotypic characterization of these mice demonstrated that at least in three different metabolic states, improved blood glucose levels in the C3H allele carriers of *Nbg15* locus was accompanied by elevated plasma insulin levels. Firstly, after refeeding, *Nbg15c*$^{C3H/C3H}$ mice had lower blood glucose levels attributed to elevated plasma insulin. Secondly, this result was reproduced in the *Nbg15c*$^{C3H/C3H}$ mice after 16 h fasting prior to i.p.GTT, effect that is attributed to the further improved glucose disposal displayed by this group after intraperitoneal injection of glucose. Moreover, although in these two occasions the *Nbg15d*$^{C3H/C3H}$ mice also had on average a similar phenotypic pattern as to the *Nbg15c*$^{C3H/C3H}$, the results of the comparison to the *Nbg15*$^{NZO/NZO}$ group failed to reach statistical significance. On the other hand, after 6 hour fasted prior to experimental termination, the observed *Nbg15c*$^{C3H/C3H}$ insulin associated-phenotype was phenocopied in the *Nbg15d*$^{C3H/C3H}$ mice when compared to *Nbg15*$^{NZO/NZO}$. Contrarily, mice carrying the C3H alleles at the proximal sublocus of *Nbg15* (*Nbg15p*$^{C3H/C3H}$) did not show a reduction in blood glucose levels, neither during fasting

nor the refeeding stages or in further experiments with shorter fasting periods (results section 3.3.3).

These results led to the conclusion that the distal sublocus of *Nbg15* is mainly required for the overall improved glucose homeostasis conferred by the C3H alleles to the NZO background, although a mild contribution or interaction to the proximal locus cannot be discarded, given that only mice carrying the full *Nbg15* locus from the C3H strain (consomic) showed stable improvements in random blood glucose levels.

Furthermore, no indication of improved insulin response during an i.p.ITT was observed among all C3H lines compared to the NZO controls. These results demonstrate that the C3H alleles on chromosome 15 do not generally protect NZO mice from developing peripheral insulin resistance, despite the fact that lower levels of plasma triglycerides detected in the *Nbg15*c$^{C3H/C3H}$ and *Nbg15*d$^{C3H/C3H}$ mice during the refeeding state could have beneficial effects on whole-body glycemia. On the other hand, although a significant reduction of plasma FFA occurred in all groups comparing fasting and refeeding state, among the groups, the levels of FFA were similar. Additionally, final measurements of liver weight showed no differences between the genotypes. This indicates that a variable degree of insulin resistance may occur among the genotypes, but does not represent the main mechanism by which the C3H alleles act on blood glucose lowering.

4.1.2.2 C3H allele carriers for *Nbg15* are protected from diabetes progression

In order to investigate if the diabetes status was a feature underlying the compensatory mechanisms associated to abnormal blood glucose levels among the genotypes, a diabetes severity threshold was set. The mice with random blood glucose over 500 mg/dl were considered severely diabetic, with possible impaired insulin secretion.

Diabetes severity was dramatically reduced in the mice carrying the distal and consomic C3H alleles of *Nbg15*. Contrary, despite the fact that the diabetes prevalence was lower for the mice carrying both C3H alleles in the proximal sublocus of *Nbg15*, all diabetic mice were severely hyperglycemic at the age of 15 weeks.

In addition, these results were independent of the known diabetes clustering phenomenon that correspond to the high phenotypic variability of these mouse strain. NZO mice are inbreed and therefore considered genetically identical. On the other hand, upon high-fat diet intervention, not 100 % of the mice develop severe obesity

and/or hyperglycemia. A certain proportion do not respond to diet intervention and remain healthy, although the majority display a severe phenotype. Among other studies, over 70 % of the NZO mice develop overt hyperglycemia when fed a high-fat diet (Jurgens, et al. 2007; Lubura, et al. 2012; Reifsnyder and Leiter 2002). In this study, if the number of mice per group would be adjusted to an equal number and these would all develop overt diabetes, they would still not reach the diabetes prevalence displayed by the *Nbg15*$^{NZO/NZO}$ group. For example, the *Nbg15c*$^{C3H/C3H}$ group was composed by 22 mice and 4 out of these were diabetic, representing 22.7 %. On the other hand, out of 31 *Nbg15*$^{NZO/NZO}$ mice, 20 were diabetic, which corresponds to 64.5 %. Thus, if 9 additional mice would be phenotyped in the *Nbg15c*$^{C3H/C3H}$ group and these would all develop diabetes, still 12 out of 31 mice would be diabetic, which results in 38.7 % of diabetes prevalence in this group.

Again, these results indicated a protective impact of *Nbg15* on islets of Langerhans. Thus, the observation in the N_2(C3HxNZO) population that led to the hypothesis that the progression of diabetes complications is retarded within the diabetic mice carrying the C3H alleles could be confirmed in the RCS carriers of the C3H alleles at distal and consomic *Nbg15* locus.

4.1.2.3 C3H allele carriers of the *Nbg15* locus improve glycaemia in a body weight independent manner

Observations in the body weight development among the RCS groups showed that statistical differences were only observed at later stages of life, exactly from the age (10 weeks of age) when diabetes prevalence reached its highest percentage among the groups. For this reason, it is difficult to distinguish if the effects on body weight precede the effects on blood glucose. However, considering the observation that body weight and gain of weight in the *Nbg15c*$^{C3H/C3H}$ and *Nbg15d*$^{C3H/C3H}$ mice were significantly higher than in the *Nbg15*$^{NZO/NZO}$, these observations support the hypothesis that these effects were secondarily caused by diabetes complications in the *Nbg15*$^{NZO/NZO}$ genotype. In addition, these results corroborate the finding of reduced body fat content in the *Nbg15p*$^{C3H/C3H}$ and *Nbg15*$^{NZO/NZO}$ groups, consistent with a physiological condition of elevated lipolysis as a consequence of severe hyperglycemia and lack of insulin action in these mice.

Furthermore, the measurements of fat disposal were in concordance to the abovementioned results showing that both gonadal and subcutaneous fat deposition were increased in the distal and consomic C3H allele carriers of the *Nbg15* locus.

4.1.2.3 Histological analysis of pancreatic islets demonstrated a larger number of small islets in homozygous C3H allele carriers of the *Nbg15* locus

To ultimately determine whether the C3H allele carriers of the *Nbg15* locus protect the pancreatic islets from metabolic stress induced by hyperglycemia, histological analysis of the pancreas was conducted.

The results of the histological analysis supported the findings that the C3H allele carriers protect NZO from diabetes by maintaining pancreatic islets function. Although the average total number of islets did not differ between the groups, among the clusters of small islets (0-5000 mm^2) the distal and consomic C3H allele carriers of *Nbg15* exhibited higher number of islets. These results could support the hypothesis that the islets secretory machinery in the carriers of the C3H alleles is improved leading to an elevated rate of neogenesis and/or proliferation. Moreover, several studies support the finding that the secretory factors and improvements in beta cell function, including beta cell expansion are part of the first compensatory mechanisms activated by the increase demand of insulin (Xu, et al. 2008). Along with this, it can be speculated on the fact that probably beta cell expansion occurred among all genotypes, however, beta cells that are functionally healthy were reduced in the *Nbg15*$^{NZO/NZO}$ group. The determination of the plasmatic levels of cytokines could give a hint towards a contribution of pro-inflammatory cytokines in this process since TNF-a, MCP-1 and IL-1b, all of them often associated with increased apoptotic signaling in islets, were diminished among all C3H allele carriers of the *Nbg15* locus compared to the carriers of the NZO alleles.

In summary, the results from the histological analysis clearly show that the C3H alleles protect NZO mice from developing diabetes via mechanisms that target the islets of Langerhans. However, the exact mechanism, whether it involves proliferation or just early ameliorated adaptations are still to be determined. Moreover, certainly, both *Nbg15* subloci are required to strengthen the anti-diabetic effects of the C3H alleles. However, the protective effects are mainly driven by the C3H distal sublocus of *Nbg15*. On the other hand, the existence of a diabetes cluster among the

proximal line, certainly still masks the resolution and interpretation of the effects that the proximal line may attribute to the diabetes protection effects.

4.2.3 Characterization of RCS.B6.C3H-*Nbg15* revealed improved glucose tolerance in C3H allele carriers

4.1.3.1 Mice carrying the full *Nbg15* locus from C3H show improved glucose tolerance

The introgression of the *Nbg15* locus into the B6 background had the purpose to confirm the protective effects of the C3H alleles without the diabetes clustering of the NZO strain.

Differently from the results observed in the RCS.NZO.C3H-*Nbg15,* the introgression of the C3H alleles in the B6 background mainly addressed effect on glucose disposal. The mice carrying the C3H alleles at the *Nbg15* locus presented ameliorated glucose tolerance during the i.p.GTT without evidence for differences in the plasmatic insulin levels during i.p.GTT. These data, together with improved insulin sensitivity, point towards a phenotype that targets exclusively the periphery action of insulin. Although B*Nbg15*$^{C3H/C3H}$ exhibited higher lean mass than B*Nbg15*$^{B6/B6}$ mice, it seems that the improved glucose tolerance in these mice resulted rather from the differences in fat disposal, given that the fat content as well as the plasmatic levels of triglycerides during refeeding were significantly reduced in the B*Nbg15*$^{C3H/C3H}$ group compared to the B*Nbg15*$^{B6/B6}$. Additionally, although the comparison between the levels of FFA during the fasted and the refed state did not show differences between the groups, the B*Nbg15*$^{C3H/C3H}$ mice had a higher rate of FFA reduction after refeeding than the B*Nbg15*$^{B6/B6}$ ones, indicating that the lipid metabolism could be a mechanism underlying the major insulin resistance phenotype observed among these lines (results section 3.4).

In summary, although it is reported that C3H mice show higher insulin secretion in comparison to B6, the C3H alleles contributed to protective effects to the B6 background only in terms of improvements of glucose disposal and insulin sensitivity (Freeman et al. 2006; Lee, et al. 1995). On the other hand, it has been reported that B6, when given a high-fat diet, rapidly develop hyperinsulinemia and can maintain the compensatory mechanism for a long time period, indicating an enhanced pancreatic islets function in comparison to NZO mice. The high degree of phenotypic similarity concerning protection of pancreatic islets from hyperglycemia in both the C3H

and the B6 mouse strain may be the cause why no differences could be observed when the C3H alleles were introgressed onto the B6 background (Roat, et al. 2014). Furthermore, probably other genes than those in the NZO background are underlying these effects, since the pathophysiological mechanisms behind diabetes development in NZO are quite peculiar and still not completely understood. Moreover, the pancreatic islet protective phenotype of the C3H alleles onto the NZO background was mainly attributed to the distal sublocus of *Nbg15*, while the full chromosome 15 from C3H was introgressed onto the B6 background (= consomic), reinforcing that the effects of the proximal sublocus of *Nbg15* could underlie the causal genes for the insulin resistant phenotype in the B6 background and therefore not being observed in the NZO background due to the early islets demise.

Furthermore, it is not surprising that the introgression of the C3H alleles into different genetic backgrounds would produce variable phenotypes. A multiplicity of studies throughout the last decades have reported dramatically differences between phenotypes dependent on the background strain (Enríquez 2019). The generation of diabetic mutants using the C57BL/KsJ and C57BL/6J substrains produced severe hyperglycemia and pancreatic islet damage in the C57BL/KsJ strain, while the same mutation lead to only transitory hyperglycemia and sustained elevated insulin levels that were combined with increased proliferation of beta cells in the C57BL/6J mice (Hummel, et al. 1972). Reported differences in insulin secretion were also observed within the C57BL/6J and C57BL/AJ mice during a high-fat diet intervention (Lee et al. 1995). Moreover, divergent levels of insulin have been observed in insulin receptor knockout mouse models generated in C57BL/6 [B6], 129Sv, and DBA mouse strains (Kulkarni, et al. 2003). A range of phenotypic diversity between the C57BL/6J and C57BL/6N mouse strains, including glucose tolerance, was also reported in another study (Simon, et al. 2013).

4.2 Identification of *Nbg15* candidate genes

The use of linkage studies to dissect monogenic diseases with a high population penetrance results in the identification of very resolved genomic regions that usually precisely match to the causal gene. On the other hand, the application of linkage studies to identify gene variants associated to polygenic diseases with a low penetrance, for instance type 2 diabetes often result in large genomic regions associated to a certain phenotype. The identification of the causal gene variants

therefore requires further approaches to narrow down the genomic region to smaller fragments containing about a handful of genes. Each criterion alone has disadvantages and can fail in the identification of causal-disease genes, especially when those are linked to complex traits. Thus, the combination of several criteria has been applied in this study to secure and strengthen the reliability of the filtering approaches in identifying the *Nbg15* associated genes.

4.2.1 Gene expression analysis lead to the validation of 16 candidate genes

In this study, several approaches were combined to identify possible causal gene variants for the *Nbg15* locus. Firstly, a gene expression profiling was conducted in the pancreatic islets, gonadal white adipose tissue, subcutaneous adipose tissue, brown adipose tissue, quadriceps and liver of the parental C3H and NZO. Using microarray expression analysis, 54 genes out of 503 annotated genes for the locus were identified as candidate genes. However, although microarray analyses enable a quick overview on transcript levels over the whole genome, the results can generate non-specific binding and background noise (Zhang, et al. 2005). Therefore, qPCR analysis is often used subsequent to the microarray analysis for validation purposes (Morey, et al. 2006). In this study, the validation of gene expression levels was performed in young and adult mice as the transcript levels often vary between developmental stages (Southworth, et al. 2009). The application of qPCR resulted in the validation of 18 candidate genes that were differentially expressed in both the young and old parental C3H and NZO mice among the aforementioned tissues.

Furthermore, the gene expression criteria were extended to the backcross population N_2(NZOxC3H). In this step, gene expression values were associated to the genotypes in the eQTL analysis, which represents a powerful tool to identify the genomic regions responsible for the regulation of gene expression (Lowe and Reddy 2015; Nica and Dermitzakis 2013; Schadt et al. 2008; Wray 2007). Thus, the identification of mRNA levels that were associated to a specific genomic region located far from where the gene is located (so-called *trans*-eQTL) were excluded as candidate genes. The application of these criteria led to the exclusion of the *Fam83f* gene as a candidate in the liver and *Tmem184b* candidacy in brown adipose tissue, therefore totalizing 16 candidate genes validated across the four analyzed tissues.

However, the eQTL analysis in the islets was not possible to be performed due to technical limitations in the isolation of islets from severe diabetic mice in the

N_2(NZOxC3H) backcross population. Therefore, all qPCR validated genes for this tissue were selected. Thus, 8 out 16 genes were validated in the pancreatic islets of Langerhans. Gene expression levels of the three genes *Slc30a8*, *Adck5* and *Lrrc24* were increased in the parental C3H islets compared to the NZO, whereas expression levels of *Has2*, *Ly6d*, *Csf2rb2*, *Cdc42ep1* and *Kdelr3* were decreased in the pancreatic islets of the parental NZO mice.

In the liver, the candidacy of *Apol7a* and *Csf2rb2* could be validated, that were respectively overexpressed in C3H and NZO parental mice. Likewise, *Apol6* and *Mchr1* genes were validated in the skeletal muscle. *Apol6* was found to be higher expressed in the parental C3H mice and *Mchr1* in the NZO strain. Moreover, *Csf2rb2* and *Slc25a17* were the validated candidate genes in the white adipose tissue, demonstrating increased expression levels in NZO and in C3H mice, respectively. In the brown adipose tissue, the candidacy of the two genes *Csf2rb2* and *Tst* was validated. Like in the other tissues, *Csf2rb2* was upregulated in the NZO mice, whereas gene expression of *Tst* was found to be increased in the parental C3H mice.

Interestingly, the gene *Csf2rb2* was the only gene which expression levels were validated among 4 of the analyzed tissues. In addition, in the liver, the mRNA levels of this gene significantly correlated with the plasmatic levels of FFA, and in the white adipose tissue correlations to TG levels and blood glucose were observed. Likewise, in the BAT, *Csf2rb2* transcript levels positively correlated to blood glucose and BMI and negatively correlated to plasma cholesterol and FFA in the N2 population. The results observed in the FFA and TG levels contradict the correlations of the *Cfs2rb2* expression with blood glucose and BMI. On the other hand, the observed correlations were rather weak when taking into account the R^2 values. The gene expression levels of *Csf2rb2* in BAT were attributed to a variation of 18.7 % in the BMI of the backcross N_2(C3HxNZO) population. However, in the other tissues, the *Csf2rb2* expression resulted in less than 10 % variation within the metabolic parameters analyzed (result section 3.5.1).

4.2.1 Haplotype analysis and *in silico* prediction of protein function revealed three genes as major candidates for the *Nbg15* locus

The mRNA levels are frequently used as criteria to narrow down candidate genes in genomic association studies. However, transcripts levels alone represent a weak filtering criterion as it is also largely reported that gene expression generally

results from the existence of different variants of the same gene. (Consortium, et al. 2008; Hasimu, et al. 2003; Wang, et al. 2004). Therefore, the combination of expression analysis with haplotype analysis represents stronger evidence for a gene candidacy. The results of the haplotype analysis revealed a total of 284 genes carrying C3H polymorphic SNPs. The haplotype analysis alone is also a weak criterion as the existence of variation in the DNA sequence are likely to occur among synonymous amino acid exchanges, therefore not implying the existence of functional impact. For this reason, the haplotype analysis was combined with the *in silico* predictions of gene-protein impact.

Using the variant effect predictor analysis (VEP), the genes *Exosc4*, *Fam227a*, *Apobec3* and *Cbx7* were identified carrying polymorphic SNPs that may have a high impact on protein function. In the *Exosc4* gene the polymorphisms were associated to a stop lost mutation, while in the *Fam227a* a start loss mutation was predicted. These are mutations in a start or stop codon respectively, leading to an abnormal protein elongation or in case of start loss, a reduction or even elimination of protein could be observed. For both *Apobec3* and *Cbx7* genes stop gained mutations were associated to the polymorphisms identified in these genes, which could result in an abnormal shortening of the protein due to a gain of a stop codon. Moreover, 61 genes were found to carry missense mutations in coding exons and therefore were classified as having a moderate impact on protein function. Using the SnpEff to qualitatively predict protein function, all identified non-synonymous amino acid substitutions were categorized as having similar or moderate impact. However, no predictions of the degree of pathogenicity could be applied using this tool. To segregate amino acid substitutions with equivalent physical properties or present in regions were the translated protein is not highly conserved or not highly contributing to regulatory function from those that may be disruptive polymorphisms, the SIFT tool was applied (Ng and Henikoff 2003). Using the SIFT, 30 genes were filtered as having potential disruptive polymorphisms. The three genes *Apol6*, *Apol7a* and *Kdelr3* were also found to be differentially expressed in the parental NZO and C3H tissues and therefore matched all filtering criteria, being selected as the most likely candidate genes for the *Nbg15* locus.

The *Apol6* gene was identified as a candidate gene in muscle. However, the general abundance of the gene was rather low in this tissue. In addition, databases present *Apol6* as a gene highly expressed in the adipose tissue (Biogps 2019; NCBIgene 2019). *Apol6* is one of the members of the ApoLs gene family and it encodes

for a protein that is associated to extracellular lipoproteins. However, it has been also demonstrated that *Apol6* is intracellularly expressed, where it participates in the binding of lipids to organelles (Tan et al. 2019). This might lead to the assumption that *Apol6* could be a tracer of intramuscular accumulation of fat in muscle and therefore a diabetes risk gene. Some studies support this hypothesis by having suggested the pro-apoptotic activity of *Apol6* (Liu, et al. 2005). On the other hand, another study demonstrated the involvement of *Apol6* in adipogenesis (Tan et al. 2019). Moreover, the gene *Apol2*, another member of the ApoLs gene family has also been associated to anti-apoptotic effects (Liao, et al. 2019). These last two studies corroborate the findings of *Apol6* expression in the muscles of the backcross N_2(C3HxNZO) population. In these, the higher expression levels of *Apol6* were associated to a diabetes protective phenotypes. Additionally, in the *Apol6* gene, three missense variants were detected. Two SNPs were leading to amino acid exchanges with similar properties, while one lead to an exchange of the negatively charged aspartate amino acid to the non-polar alanine (Tab.1 supplements). This last substitution occurs in amino acid 256 located at the transmembrane domain of the APOL6 protein (Tab.1 supplements, UniProt, 2019).

In addition, the gene apolipoprotein L 7a (*Apol7a*) was identified as a candidate gene in the liver. This is another member of the ApoLs family of genes. Although databases present this gene as being predominantly expressed in hepatocytes in the liver and this has been confirmed in the qPCR analysis performed in the liver of the N2 backcross population, no correlations of this gene with metabolic traits could be detected. Recent studies suggested an association of the *Apol7a* (also known as *Apol3*) with non-diabetic nephropathy (Skorecki, et al. 2018). Moreover, a knockdown of the *Apol7a* gene in endothelial cells was linked to the blockage of angiogenesis markers (Khalil, et al. 2018). Furthermore, the presence of two amino acid substitutions in a coding region of the gene were detected. Both substitutions lead to amino acid exchanges with different physical properties. At position 272, the polar amino acid threonine is exchanged to the non-polar isoleucine. On the other hand, at position 216 the opposite was observed, an isoleucine is exchanged by a threonine (Tab.1 supplements). Moreover, amino acid substitutions have also been reported in the human *Apol3* gene and are likely pathogenic. However, although *Apol7a* has been associated with nephropathy, in the liver, no evidence of the role of the gene has been

investigated so far and the pathogenicity of the missense variants remain to be validated (NCBI-dbSNP 2019) (results section 3.6).

4.2 Identification of *Kdelr3* as *Nbg15* islet candidate gene

Kdelr3 was the major candidate gene found in the pancreatic islets that matched all filtering criteria that aimed targeting the gene with the highest likelihood to be causatively associated to the phenotypes driven by the *Nbg15* locus. The KDELR3 protein is composed by seven transmembrane domains responsible for the recognition of a C-terminal tetra-peptide termed KDEL motif (Lys-Asp-Glu-Leu) in certain proteins or chaperones, thereby preventing them to be secreted to the Golgi complex by activating the retro transport of these proteins to the endoplasmic reticulum. In mouse, another two genes, *Kdelr1* and *Kdelr2* also participate in this process. On the one hand, apart from *KDELR1* and *KDELR2*, two isoforms of the *KDELR3* gene have been described in humans (Capitani and Sallese 2009).

The retrograde transport function of the *Kdel* genes was described more than 20 years ago. However, so far the regulation of these mechanisms are not completely understood. Moreover, emerging findings suggest that *Kdel* also participates in intracellular signaling processes that triggers the trafficking of proteins from the Golgi to the plasma membrane (Capitani and Sallese 2009; Saudek 2012). A more recent study showed evidence that *Kderl2* and *Kdelr3*, but not *Kdelr1* are part of the unfolded protein response (UPR) and therefore participate in the adaptation mechanisms coupling disease-associated ER stress like diabetes (Trychta, et al. 2018). These findings support the results obtained from the expression analysis of *Kdelr3* in the pancreatic islets of the parental NZO and C3H mouse strains. The mRNA levels of *Kdelr3* were less abundant in the pancreatic islets than *Kdelr1* and *Kdelr2*. On the other hand, among the tissues, the pancreatic islets exhibited the highest mRNA levels of the gene (Fig. 42). Moreover, islets of Langerhans from NZO mice presented higher mRNA levels of the *Kdelr3* gene in comparison to the C3H pancreatic islets, reinforcing the idea that elevated ER stress in NZO islets could lead to early activation of the UPR and consequently beta cells apoptosis.

Furthermore, the haplotype analysis revealed that NZO mice carry a coding missense mutation in this gene that is predicted to affect protein function. In addition, this mutation was associated to elevated diabetes associated deaths and pancreatic islet damage in the backcross N_2(NZOxC3H) population. Furthermore, RNA

sequencing of the parental NZO islets of Langerhans validated the existence of a missense mutation in the position 96 of the KDELR3 protein. This region is highly conserved among species, supporting the finding of a possible deleterious impact in the protein. This polymorphism has been reported in other mouse strains (Sanger_SNP_Database 2020). BTBR T Itpr3tf/J mice as well as NZB/B1NJ have been linked to obesity associated phenotypes, while WSB/EiJ was associated to abnormal insulin secretion in one study (Karunakaran, et al. 2013; Lee, et al. 2011; Wergedal, et al. 2007). On the other hand, BUB/BnJ, LEWES/EiJ, RF/J, ST/bJ and ZALENDE/EiJ mice have not been described in the context of obesity or T2D. The base pair exchange adenine to thymine in the DNA leads to an amino acid substitution of the acidic and negatively charged glutamic acid to the aliphatic amino acid valine. This alteration could significantly disturb the function of the proteins as both amino acid display totally different physical properties.

Mutational studies performed in the *KDELR1* gene (also known as *Edr2*) did not report alterations on the subcellular localization of this gene upon amino acid substitution at position 96 (Glu^{96}), although amino acid exchanges in the cytoplasmic loop of *Kdelr1*, same subcellular localization of the *Kdelr3* mutation, are believed to generally affect this property (Townsley, et al. 1993).

A more recent study performed extensive experiments on the structure of KDELR2 (Bräuer, et al. 2019). Accordingly, the KDELR receptors are organized in 7 transmembrane (TM) alpha helices and two internal triple helix bundles (THB), connected by the linker helix TM4. The acidic pH of the Golgi is required for ligand binding. Upon this event, conformational changes occur projecting the cytosolic bundles to form a negatively charged pocket. This is, in sum, essential for electrostatic interaction to other proteins. One of these proteins is the COPI and COPII vesicle trafficking complexes. The mutated residue Glu^{96} is located in the cytosolic TM4 helix linker. Thus, one hypothesis is that Glu^{96} contributes to the formation of the interaction pocket in the protein. Although this has not been reported by Bräuer and collaborators (2019), in other evolutionary related proteins like sugar transporters SWEET, TM4 has been related to be important for protein function (Becker, et al. 2016; Feng and Frommer 2016; Saudek 2012; Tao, et al. 2015; Xu, et al. 2014).

Additionally, the hypothesis of a structural function of TM4 cannot be discarded as the protein movement in flushing to the luminal membrane upon ligand binding followed by its cytosolic projection must require TM4 specialized function. Thus, we

can imply that the exchange of a charge to a non-charged amino acid at this domain could probably indirectly lead to variable affinities for ligand binding. This hypothesis is partially supported by another study that demonstrated that complete deletion of the C-terminal cytoplasmic domain of the KDEL receptor did not alter the subcellular localization of the receptor, but did influence ligand binding capacity (Capitani and Sallese 2009). Although extensive studies have been performed in regards of KDEL retrograde transport, it is still not completely understood how further signaling processes are regulated by this family of proteins, especially in disease. Regarding *Kdelr3*, although its participation in the UPR has been described, nothing is known about this gene in the context of type 2 diabetes. It is known however that the UPR response is activated in type 2 diabetes upon ER stress resulted from overload of misfolded proteins (Back and Kaufman 2012; Cnop, et al. 2012). Moreover, upon failure of this response, cell death signals arise, triggering beta cell demise and apoptosis in type 2 diabetes (Kataoka and Noguchi 2013). In this context, *KDELR3* mutations or variations could be directly linked to type 2 diabetes. Probably, the compensatory response to ER stress in the NZO is impaired, leading to the upregulation of *Kdelr3*. Alternatively, the *Kdelr3* NZO mutant could be more sensitive to ER stress signals. This hypothesis is supported by the findings of Yamamoto and co-workers who demonstrated that *Kdel* is associated to p38 mitogen-activated protein kinases (MAPKs) and c-Jun amino-terminal kinases (JNKs) in intracellular stress condition, which are also important effectors mediating beta cell dysfunction and insulin resistance (Back and Kaufman 2012; Herbert and Laybutt 2016; Kataoka and Noguchi 2013; Yamamoto, et al. 2003).

In addition, further mechanisms linking *Kdel* gene to metabolism could direct *Kdelr3* to the pathogenesis of type 2 diabetes. It has been postulated that *Kdel* interacts with the proto-oncogene tyrosine-protein kinase Src to trigger the signal that drive proteins from the Golgi to the plasma membrane (Pulvirenti, et al. 2008). Experimental evidence also suggests that the downregulation of Src alters insulin secretion dynamics in INS1 cells *in vitro* (Sato, et al. 2016). In addition, some studies have demonstrated that KDEL interacts with the protein kinase A (PKA) and protein kinase C (PKC), two important mediators in the regulation of insulin secretion (Cabrera, et al. 2003; Cancino, et al. 2013; Nesher, et al. 2002). Moreover, given that insulin biosynthesis requires the activation of several protein complexes of vesicles formation and trafficking from ER to Golgi and plasma membrane, including COPII, the

hypothesis that Kdelr3 takes part in this process is almost certain, but experimentally not investigated so far (Fang, et al. 2015).

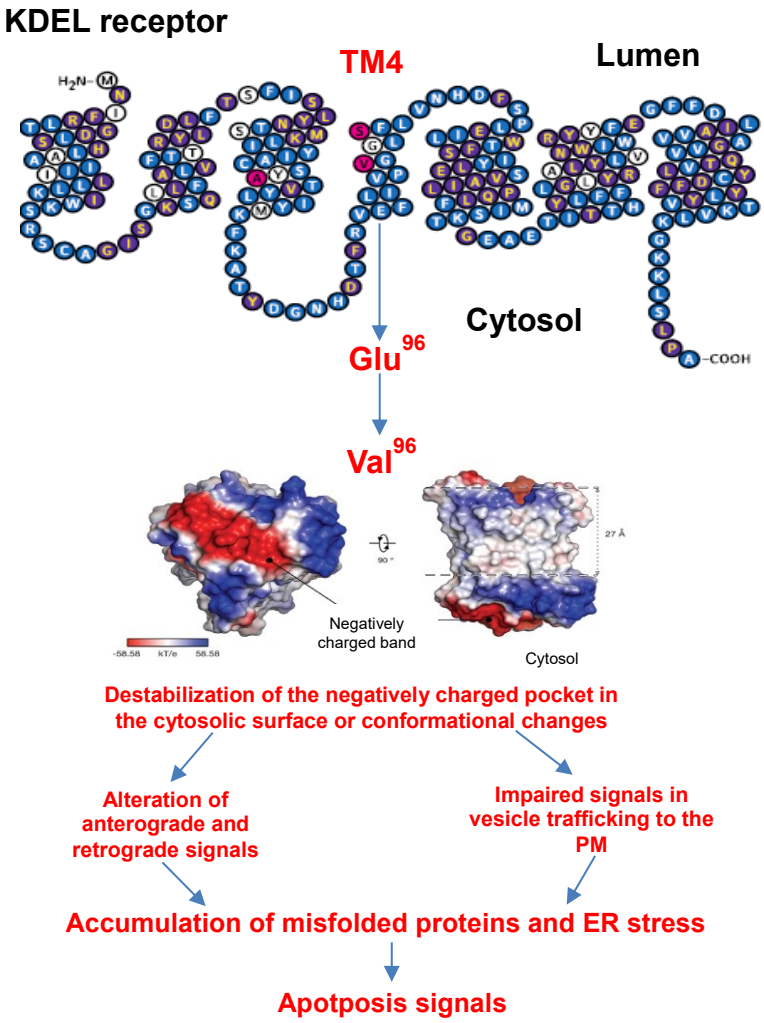


Figure 46. Hypothesis for *Kdelr3* mutant functional impact. Alterations in pancreatic islets stability in the NZO mice could be associated to the existence of a missense mutation in the *Kdelr3* gene. The negatively charged glutamic acid (Glu^{96}) is exchange to the non-charged amino acid Valine (Val^{96}) in the transmembrane 4 domain (TM4). The cytosolic location of the mutation could lead to alterations in electrostatic interactions to proteins or lead to abnormal conformation. These could lead to defective COPI and COPII signals for the trafficking of the receptor from the Golgi to ER and vice versa or disturb the anchorage of other proteins involved in vesicle trafficking to the plasma membrane (PM). Accumulation of misfolded proteins in the ER could be a result of these impaired signals that in a final step would trigger apoptosis in the beta cells and thus increase T2D risk. Source: (Brauer, et al. 2019).

4.3 Conclusion and future perspectives

It is known that a complex interaction between environmental factors and the genetic susceptibility of an individual play a crucial role in the development of metabolic and chronic diseases like type 2 diabetes (T2D). The estimated genetic heritability of T2D varies between 30-80 % among the human population. However, the genes identified so far account only for a small proportion of this and generally can explain

only 20 % or less of the genetic heritability associated to T2D. A combination of several human genomic association studies in the diabetes field have identified about 40 genes as causal variants for the disease (T2DM-Knowledge-Portal 2019). On the other hand, genomic association studies in human populations imply several disadvantages, especially because environmental factors are not possible to be controlled and often influence the outcome associations. Therefore mouse models have been largely used to dissect the gene variants influencing diabetes development (Broman 2009; Joost and Schurmann 2014).

In this study, a diabetes-associated locus on chromosome 15 (*Nbg15*) identified in a backcross population generated between the diabetes prone NZO with the diabetes resistant C3H mouse strain was metabolically characterized. The further generation of recombinant congenic strains (RCS) carrying the full, proximal and distal genomic fragments of the *Nbg15* locus, respectively, revealed that ameliorated glucose metabolic status addressed by the C3H alleles results from a pancreatic islets driven phenotype. Moreover, this phenotype was mainly attributed to the genomic region located distally on chromosome 15, whereas the proximal region may contribute to effects on the level of intracellular insulin signaling. However, further experimental approaches need to be performed to better understand the contribution of this specific genomic region to variations in the insulin sensitivity.

Investigations on the gene expression level identified *Kdelr3* as a causal candidate gene underlying the *Nbg15* locus. Furthermore, the identification of a disruptive mutation in the *Kdelr3* of the NZO strain is hypothesized to negatively influence adaptive responses activated upon metabolic stress induced by chronic energy imbalance in these mice. As a consequence, accumulation of misfolded proteins may occur intracellularly leading to the activation of cell death signals in the beta cells of the NZO. Interestingly, this gene was also found to be associated to alterations in insulin expression in human islets of Langerhans. However, further experiments need to be performed to investigate the underlying mechanism that result in the association of *Kdelr3* to diabetes risk and to beta-cells stability.

Furthermore, although the application of the filtering criteria was successful to identify a candidate gene in pancreatic islets, it is not surprising that every criterion applied corresponds to a biased and simplified method to select genes and might lead to a loss of important genes contributing to the phenotypes, therefore leading to the selection of genes that only contribute to small proportion of the phenotypic variations.

For this reason, further analysis could be performed including prediction of gene-gene interactions and predictions on the intergenic regulatory regions could also be addressed. The ultimately approach would involve the generation of a knockout mouse model to finally prove the influence of *Kdelr3* in the metabolic regulation associated to type 2 diabetes susceptibility.

6 REFERENCES

1. Ahlqvist E, Ahluwalia TS & Groop L 2011 Genetics of type 2 diabetes. *Clin Chem* **57** 241-254.
2. Akter R, Cao P, Noor H, Ridgway Z, Tu LH, Wang H, Wong AG, Zhang X, Abedini A, Schmidt AM, et al. 2016 Islet Amyloid Polypeptide: Structure, Function, and Pathophysiology. *J Diabetes Res* **2016** 2798269.
3. Ali O 2013 Genetics of type 2 diabetes. *World J Diabetes* **4** 114-123.
4. Appleby MW & Ramsdell F 2003 A forward-genetic approach for analysis of the immune system. *Nat Rev Immunol* **3** 463-471.
5. Asthana S, Mallick B, Alexandrescu AT & Jha S 2018 IAPP in type II diabetes: Basic research on structure, molecular interactions, and disease mechanisms suggests potential intervention strategies. *Biochim Biophys Acta Biomembr*.
6. Augustin R 2010 The protein family of glucose transport facilitators: It's not only about glucose after all. *IUBMB Life* **62** 315-333.
7. Back SH & Kaufman RJ 2012 Endoplasmic reticulum stress and type 2 diabetes. *Annu Rev Biochem* **81** 767-793.
8. BCBC 2019. *Beta Cell Biology Consortium (BCBC). Available at: https://www.betacell.org/ 10.12.2019*.
9. Becker B, Shaebani MR, Rammo D, Bubel T, Santen L & Schmitt MJ 2016 Cargo binding promotes KDEL receptor clustering at the mammalian cell surface. *Sci Rep* **6** 28940.
10. Biden TJ, Boslem E, Chu KY & Sue N 2014 Lipotoxic endoplasmic reticulum stress, beta cell failure, and type 2 diabetes mellitus. *Trends Endocrinol Metab* **25** 389-398.
11. Bielschowsky M & Goodall CM 1970 Origin of inbred NZ mouse strains. *Cancer Res* **30** 834-836.
12. Biogps 2019. *Available at: http://biogps.org/#goto=welcome 10.10.19*.
13. Bonnefond A & Froguel P 2015 Rare and common genetic events in type 2 diabetes: what should biologists know? *Cell Metab* **21** 357-368.
14. Brauer P, Parker JL, Gerondopoulos A, Zimmermann I, Seeger MA, Barr FA & Newstead S 2019 Structural basis for pH-dependent retrieval of ER proteins from the Golgi by the KDEL receptor. *Science* **363** 1103-1107.
15. Bräuer P, Parker JL, Gerondopoulos A, Zimmermann I, Seeger MA, Barr FA & Newstead S 2019 Structural basis for pH-dependent retrieval of ER proteins from the Golgi by the KDEL receptor. *Science* **363** 1103-1107.
16. Briant L, Salehi A, Vergari E, Zhang Q & Rorsman P 2016 Glucagon secretion from pancreatic alpha-cells. *Ups J Med Sci* **121** 113-119.
17. Brockmann GN, Christina. 2012 Positional Cloning of Diabetes Genes. Methods in molecular biology (Clifton, N.J.). *Methods in molecular biology (Clifton, N.J.). 933. 275-89. 10.1007/978-1-62703-068-7_18.*
18. .
19. Broman KW 2001 Review of statistical methods for QTL mapping in experimental crosses. *Lab Anim (NY)* **30** 44-52.
20. Broman KWS, S. 2009 A Guide to QTL Mapping with R/qtl *Statistics for Biology and Health Springer*.
21. Browning BL & Browning SR 2011 A fast, powerful method for detecting identity by descent. *Am J Hum Genet* **88** 173-182.
22. Bryant NJ, Govers R & James DE 2002 Regulated transport of the glucose transporter GLUT4. *Nat Rev Mol Cell Biol* **3** 267-277.

23. Cabrera M, Muniz M, Hidalgo J, Vega L, Martin ME & Velasco A 2003 The retrieval function of the KDEL receptor requires PKA phosphorylation of its C-terminus. *Mol Biol Cell* **14** 4114-4125.
24. Cancino J, Jung JE & Luini A 2013 Regulation of Golgi signaling and trafficking by the KDEL receptor. *Histochem Cell Biol* **140** 395-405.
25. Cantley J & Ashcroft FM 2015 Q&A: insulin secretion and type 2 diabetes: why do beta-cells fail? *BMC Biol* **13** 33.
26. Capitani M & Sallese M 2009 The KDEL receptor: new functions for an old protein. *FEBS Lett* **583** 3863-3871.
27. Castracane VD, Henson, Michael C. 2007 Leptin. *Springer*.
28. Cawthorn WP & Sethi JK 2008 TNF-alpha and adipocyte biology. *FEBS Lett* **582** 117-131.
29. Chadt A, Leicht K, Deshmukh A, Jiang LQ, Scherneck S, Bernhardt U, Dreja T, Vogel H, Schmolz K, Kluge R, et al. 2008 Tbc1d1 mutation in lean mouse strain confers leanness and protects from diet-induced obesity. *Nat Genet* **40** 1354-1359.
30. Chadt ASSJHA-H, H. 2018 Molecular links between Obesity and Diabetes: "Diabesity". *Available from: https://www.ncbi.nlm.nih.gov/books/NBK279051/ 10.12.2019*.
31. Champy MF, Selloum M, Zeitler V, Caradec C, Jung B, Rousseau S, Pouilly L, Sorg T & Auwerx J 2008 Genetic background determines metabolic phenotypes in the mouse. *Mamm Genome* **19** 318-331.
32. Charron NGaMJ 2003 What We Know about Facilitative Glucose Transporters. *BIOCHEMISTRY AND MOLECULAR BIOLOGY EDUCATION*
33. **31** 163–172
34. Clee SM & Attie AD 2007 The genetic landscape of type 2 diabetes in mice. *Endocr Rev* **28** 48-83.
35. Cnop M, Foufelle F & Velloso LA 2012 Endoplasmic reticulum stress, obesity and diabetes. *Trends Mol Med* **18** 59-68.
36. Cohen P & Frame S 2001 The renaissance of GSK3. *Nat Rev Mol Cell Biol* **2** 769-776.
37. Consortium S, Saar K, Beck A, Bihoreau MT, Birney E, Brocklebank D, Chen Y, Cuppen E, Demonchy S, Dopazo J, et al. 2008 SNP and haplotype mapping for genetic analysis in the rat. *Nat Genet* **40** 560-566.
38. Consortium STD, Williams AL, Jacobs SB, Moreno-Macias H, Huerta-Chagoya A, Churchhouse C, Marquez-Luna C, Garcia-Ortiz H, Gomez-Vazquez MJ, Burtt NP, et al. 2014 Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* **506** 97-101.
39. Cox RD & Church CD 2011 Mouse models and the interpretation of human GWAS in type 2 diabetes and obesity. *Dis Model Mech* **4** 155-164.
40. Das SK & Elbein SC 2006 The Genetic Basis of Type 2 Diabetes. *Cellscience* **2** 100-131.
41. Dean L MJ 2004 The Genetic Landscape of Diabetes [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2004. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1667/. *Available from: https://www.ncbi.nlm.nih.gov/books/NBK1667/*.
42. DeFronzo RA, Ferrannini E, Groop L, Henry RR, Herman WH, Holst JJ, Hu FB, Kahn CR, Raz I, Shulman GI, et al. 2015 Type 2 diabetes mellitus. *Nat Rev Dis Primers* **1** 15019.

43. Deng J, Lu PD, Zhang Y, Scheuner D, Kaufman RJ, Sonenberg N, Harding HP & Ron D 2004 Translational repression mediates activation of nuclear factor kappa B by phosphorylated translation initiation factor 2. *Mol Cell Biol* **24** 10161-10168.
44. Diamond J 2003 The double puzzle of diabetes. *Nature* **423** 599-602.
45. Donath MY 2014 Targeting inflammation in the treatment of type 2 diabetes: time to start. *Nat Rev Drug Discov* **13** 465-476.
46. Drews G, Krippeit-Drews P & Dufer M 2010 Oxidative stress and beta-cell dysfunction. *Pflugers Arch* **460** 703-718.
47. Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O'Connell P & Stern MP 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.
48. Duncan RE, Ahmadian M, Jaworski K, Sarkadi-Nagy E & Sul HS 2007 Regulation of lipolysis in adipocytes. *Annu Rev Nutr* **27** 79-101.
49. Enríquez JA 2019 Mind your mouse strain. *Nature Metabolism* **1** 5-7.
50. Fang J, Liu M, Zhang X, Sakamoto T, Taatjes DJ, Jena BP, Sun F, Woods J, Bryson T, Kowluru A, et al. 2015 COPII-Dependent ER Export: A Critical Component of Insulin Biogenesis and beta-Cell ER Homeostasis. *Mol Endocrinol* **29** 1156-1169.
51. Fawcett KA & Barroso I 2010 The genetics of obesity: FTO leads the way. *Trends Genet* **26** 266-274.
52. Feng L & Frommer WB 2016 Evolution of Transporters: The Relationship of SWEETs, PQ-loop, and PnuC Transporters. *Trends Biochem Sci* **41** 118-119.
53. Flamment M, Hajduch E, Ferre P & Foufelle F 2012 New insights into ER stress-induced insulin resistance. *Trends Endocrinol Metab* **23** 381-390.
54. Flint J & Eskin E 2012 Genome-wide association studies in mice. *Nat Rev Genet* **13** 807-817.
55. Fonseca SG, Gromada J & Urano F 2011 Endoplasmic reticulum stress and pancreatic beta-cell death. *Trends Endocrinol Metab* **22** 266-274.
56. Freeman HC, Hugill A, Dear NT, Ashcroft FM & Cox RD 2006 Deletion of nicotinamide nucleotide transhydrogenase: a new quantitive trait locus accounting for glucose intolerance in C57BL/6J mice. *Diabetes* **55** 2153-2156.
57. Fu Z, Gilbert ER & Liu D 2013 Regulation of insulin synthesis and secretion and pancreatic Beta-cell dysfunction in diabetes. *Curr Diabetes Rev* **9** 25-53.
58. Gaisano HY, Macdonald PE & Vranic M 2012 Glucagon secretion and signaling in the development of diabetes. *Front Physiol* **3** 349.
59. Galochkina T, Ng Fuk Chong M, Challali L, Abbar S & Etchebest C 2019 New insights into GluT1 mechanics during glucose transfer. *Sci Rep* **9** 998.
60. Gerber PA & Rutter GA 2017 The Role of Oxidative Stress and Hypoxia in Pancreatic Beta-Cell Dysfunction in Diabetes Mellitus. *Antioxid Redox Signal* **26** 501-518.
61. Ghaben AL & Scherer PE 2019 Adipogenesis and metabolic health. *Nat Rev Mol Cell Biol* **20** 242-258.
62. Gjesing AP, Kjems LL, Vestmar MA, Grarup N, Linneberg A, Deacon CF, Holst JJ, Pedersen O & Hansen T 2011 Carriers of the TCF7L2 rs7903146 TT genotype have elevated levels of plasma glucose, serum proinsulin and plasma gastric inhibitory polypeptide (GIP) during a meal test. *Diabetologia* **54** 103-110.
63. Glans F, Elgzyri T, Shaat N, Lindholm E, Apelqvist J & Groop L 2008 Immigrants from the Middle-East have a different form of Type 2 diabetes compared with Swedish patients. *Diabet Med* **25** 303-307.

64. Gromada J, Chabosseau P & Rutter GA 2018 The alpha-cell in diabetes mellitus. *Nat Rev Endocrinol* **14** 694-704.
65. Groop L, Forsblom C, Lehtovirta M, Tuomi T, Karanko S, Nissen M, Ehrnstrom BO, Forsen B, Isomaa B, Snickars B, et al. 1996 Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. *Diabetes* **45** 1585-1593.
66. Guenet JL 2005 The mouse genome. *Genome Res* **15** 1729-1740.
67. Habegger KM, Heppner KM, Geary N, Bartness TJ, DiMarchi R & Tschop MH 2010 The metabolic actions of glucagon revisited. *Nat Rev Endocrinol* **6** 689-697.
68. Haeusler RA, McGraw TE & Accili D 2018 Biochemical and cellular properties of insulin receptor signalling. *Nat Rev Mol Cell Biol* **19** 31-44.
69. Halban PA, Polonsky KS, Bowden DW, Hawkins MA, Ling C, Mather KJ, Powers AC, Rhodes CJ, Sussel L & Weir GC 2014 beta-cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *Diabetes Care* **37** 1751-1758.
70. Han HS, Kang G, Kim JS, Choi BH & Koo SH 2016 Regulation of glucose metabolism from a liver-centric perspective. *Exp Mol Med* **48** e218.
71. Hasimu B, Nakayama T, Mizutani Y, Izumi Y, Asai S, Soma M, Kokubun S & Ozawa Y 2003 Haplotype analysis of the human renin gene and essential hypertension. *Hypertension* **41** 308-312.
72. Haskell BD, Flurkey K, Duffy TM, Sargent EE & Leiter EH 2002 The diabetes-prone NZO/HlLt strain. I. Immunophenotypic comparison to the related NZB/BINJ and NZW/LacJ strains. *Lab Invest* **82** 833-842.
73. Herbert TP & Laybutt DR 2016 A Reevaluation of the Role of the Unfolded Protein Response in Islet Dysfunction: Maladaptation or a Failure to Adapt? *Diabetes* **65** 1472-1480.
74. Hess ME & Bruning JC 2014 The fat mass and obesity-associated (FTO) gene: Obesity and beyond? *Biochim Biophys Acta* **1842** 2039-2047.
75. Hong EG, Jung DY, Ko HJ, Zhang Z, Ma Z, Jun JY, Kim JH, Sumner AD, Vary TC, Gardner TW, et al. 2007 Nonobese, insulin-deficient Ins2Akita mice develop type 2 diabetes phenotypes including insulin resistance and cardiac remodeling. *Am J Physiol Endocrinol Metab* **293** E1687-1696.
76. Hou JC, Min L & Pessin JE 2009 Chapter 16 Insulin Granule Biogenesis, Trafficking and Exocytosis. In *Insulin and IGFs*, pp 473-506.
77. Hou JC & Pessin JE 2007 Ins (endocytosis) and outs (exocytosis) of GLUT4 trafficking. *Curr Opin Cell Biol* **19** 466-473.
78. Hsieh PSaC, Pei-Chi 2017 The Role of Adipocyte Hypertrophy and Hypoxia in the Development of Obesity-Associated Adipose Tissue Inflammation and Insulin Resistance.
79. Hu P, Han Z, Couvillon AD, Kaufman RJ & Exton JH 2006 Autocrine tumor necrosis factor alpha links endoplasmic reticulum stress to the membrane death receptor pathway through IRE1alpha-mediated NF-kappaB activation and down-regulation of TRAF2 expression. *Mol Cell Biol* **26** 3071-3084.
80. Hummel KP, Coleman DL & Lane PW 1972 The influence of genetic background on expression of mutations at the diabetes locus in the mouse. I. C57BL-KsJ and C57BL-6J strains. *Biochem Genet* **7** 1-13.
81. IDF 2019 IDF DIABETES ATLAS Ninth edition 2019 **Available at: https://www.diabetesatlas.org/upload/resources/2019/IDF_Atlas_9th_Edition_2019.pdf 10.12.19**.

82. Igel M, Becker W, Herberg L & Joost HG 1997 Hyperleptinemia, leptin resistance, and polymorphic leptin receptor in the New Zealand obese mouse. *Endocrinology* **138** 4234-4239.
83. Joost H-G, Al-Hasani, Hadi, Schürmann, Annette (Eds.) 2012 Animal Models in Diabetes Research. *Methods in molecular biology (Clifton, N.J.). 933. 275-89. 10.1007/978-1-62703-068-7_18.*
84. Joost HG & Schurmann A 2014 The genetic basis of obesity-associated type 2 diabetes (diabesity) in polygenic mouse models. *Mamm Genome* **25** 401-412.
85. Junger E, Herberg L, Jeruschke K & Leiter EH 2002 The diabetes-prone NZO/Hl strain. II. Pancreatic immunopathology. *Lab Invest* **82** 843-853.
86. Jurgens HS, Neschen S, Ortmann S, Scherneck S, Schmolz K, Schuler G, Schmidt S, Bluher M, Klaus S, Perez-Tilve D, et al. 2007 Development of diabetes in obese, insulin-resistant mice: essential role of dietary carbohydrate in beta cell destruction. *Diabetologia* **50** 1481-1489.
87. Jurgens HS, Schurmann A, Kluge R, Ortmann S, Klaus S, Joost HG & Tschop MH 2006 Hyperphagia, lower body temperature, and reduced running wheel activity precede development of morbid obesity in New Zealand obese mice. *Physiol Genomics* **25** 234-241.
88. Kahn SE, Hull RL & Utzschneider KM 2006 Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* **444** 840-846.
89. Kaku KM, J.; Province, M. and Permutt, M.A. 1989 A single major gene controls most of the difference in susceptibility to
90. streptozotocin-induced diabetes between C5 7BL/6J and C3H/HeJ mice. *Diabetologia* **32** 716-723.
91. Kalwat MA & Thurmond DC 2013 Signaling mechanisms of glucose-induced F-actin remodeling in pancreatic islet beta cells. *Exp Mol Med* **45** e37.
92. Kanatsuka A, Kou S & Makino H 2018 IAPP/amylin and beta-cell failure: implication of the risk factors of type 2 diabetes. *Diabetol Int* **9** 143-157.
93. Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, Stengard J & Kesaniemi YA 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.
94. Karunakaran S, Manji A, Yan CS, Wu ZJ & Clee SM 2013 Moo1 obesity quantitative trait locus in BTBR T+ Itpr3tf/J mice increases food intake. *Physiol Genomics* **45** 191-199.
95. Kasahara T & Kasahara M 1996 Expression of the rat GLUT1 glucose transporter in the yeast Saccharomyces cerevisiae. *Biochem J* **315 (Pt 1)** 177-182.
96. Kataoka HU & Noguchi H 2013 ER Stress and beta-Cell Pathogenesis of Type 1 and Type 2 Diabetes and Islet Transplantation. *Cell Med* **5** 53-57.
97. Khalil A, Poelvoorde P, Fayyad-Kazan M, Rousseau A, Nuyens V, Uzureau S, Biston P, El-Makhour Y, Badran B, Van Antwerpen P, et al. 2018 Apoliporotein L3 interferes with endothelial tube formation via regulation of ERK1/2, FAK and Akt signaling pathway. *Atherosclerosis* **279** 73-87.
98. Khan S & Wang CH 2014 ER stress in adipocytes and insulin resistance: mechanisms and significance (Review). *Mol Med Rep* **10** 2234-2240.
99. Kraakman MJ, Kammoun HL, Allen TL, Deswaerte V, Henstridge DC, Estevez E, Matthews VB, Neill B, White DA, Murphy AJ, et al. 2015 Blocking IL-6 trans-signaling prevents high-fat diet-induced adipose tissue macrophage recruitment but does not improve insulin resistance. *Cell Metab* **21** 403-416.
100. Kratz M, Coats BR, Hisert KB, Hagman D, Mutskov V, Peris E, Schoenfelt KQ, Kuzma JN, Larson I, Billing PS, et al. 2014 Metabolic dysfunction drives a

mechanistically distinct proinflammatory phenotype in adipose tissue macrophages. *Cell Metab* **20** 614-625.
101. Kulkarni RN, Almind K, Goren HJ, Winnay JN, Ueki K, Okada T & Kahn CR 2003 Impact of genetic background on development of hyperinsulinemia and diabetes in insulin receptor/insulin receptor substrate-1 double heterozygous mice. *Diabetes* **52** 1528-1534.
102. Laboratory TJ 2019. *Available at: https://www.jax.org/#*.
103. Lander E & Kruglyak L 1995 Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* **11** 241-247.
104. Lander ES & Botstein D 1989 Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121** 185-199.
105. Larance M, Ramm G, Stockli J, van Dam EM, Winata S, Wasinger V, Simpson F, Graham M, Junutula JR, Guilhaus M, et al. 2005 Characterization of the role of the Rab GTPase-activating protein AS160 in insulin-regulated GLUT4 trafficking. *J Biol Chem* **280** 37803-37813.
106. Laurie CC, Nickerson DA, Anderson AD, Weir B, Livingston RJ, Dean MD, Smith KL, Schadt EE & Nachman MW 2005 Linkage disequilibrium in wild mice. *PLoS Genetics* **preprint**.
107. Lee KT, Karunakaran S, Ho MM & Clee SM 2011 PWD/PhJ and WSB/EiJ mice are resistant to diet-induced obesity but have abnormal insulin secretion. *Endocrinology* **152** 3005-3017.
108. Lee SK, Opara EC, Surwit RS, Feinglos MN & Akwari OE 1995 Defective glucose-stimulated insulin release from perifused islets of C57BL/6J mice. *Pancreas* **11** 206-211.
109. Leiter EH & Reifsnyder PC 2004 Differential levels of diabetogenic stress in two new mouse models of obesity and type 2 diabetes. *Diabetes* **53 Suppl 1** S4-11.
110. Leiter EH, Reifsnyder PC, Flurkey K, Partke HJ, Junger E & Herberg L 1998 NIDDM genes in mice: deleterious synergism by both parental genomes contributes to diabetogenic thresholds. *Diabetes* **47** 1287-1295.
111. Lenzen S 2008 Oxidative stress: the vulnerable beta-cell. *Biochem Soc Trans* **36** 343-347.
112. Leto D & Saltiel AR 2012 Regulation of glucose transport by insulin: traffic control of GLUT4. *Nat Rev Mol Cell Biol* **13** 383-396.
113. Li X, Li Y, Song B, Guo S, Chu S, Jia N & Niu W 2012 Hematopoietically-expressed homeobox gene three widely-evaluated polymorphisms and risk for diabetes: a meta-analysis. *PLoS One* **7** e49917.
114. Liao Z, She C, Ma L, Sun Z, Li P, Zhang X, Wang P & Li W 2019 KDELR2 Promotes Glioblastoma Tumorigenesis Targeted by HIF1a via mTOR Signaling Pathway. *Cell Mol Neurobiol* **39** 1207-1215.
115. Liu Z, Lu H, Jiang Z, Pastuszyn A & Hu CA 2005 Apolipoprotein l6, a novel proapoptotic Bcl-2 homology 3-only protein, induces mitochondria-mediated apoptosis in cancer cells. *Mol Cancer Res* **3** 21-31.
116. Lowe WL, Jr. & Reddy TE 2015 Genomic approaches for understanding the genetics of complex disease. *Genome Res* **25** 1432-1441.
117. Lu CC, Chen YT, Chen SY, Hsu YM, Lin CC, Tsao JW, Juan YN, Yang JS & Tsai FJ 2018 Hematopoietically expressed homeobox gene is associated with type 2 diabetes in KK Cg-A(y)/J mice and a Taiwanese Han Chinese population. *Exp Ther Med* **16** 185-191.
118. Lu J, Zhao J, Meng H & Zhang X 2019 Adipose Tissue-Resident Immune Cells in Obesity and Type 2 Diabetes. *Front Immunol* **10** 1173.

119. Lubura M, Hesse D, Neumann N, Scherneck S, Wiedmer P & Schurmann A 2012 Non-invasive quantification of white and brown adipose tissues and liver fat content by computed tomography in mice. *PLoS One* **7** e37026.
120. Lyon MF, S. Rastan, and B. D.M. 1996 Genetic variants and strains of the laboratory mouse, 3rd ed.
121. . Oxford University Press, Oxford, England.
122. Mafakheri S, Florke RR, Kanngiesser S, Hartwig S, Espelage L, De Wendt C, Schonberger T, Hamker N, Lehr S, Chadt A, et al. 2018 AKT and AMP-activated protein kinase regulate TBC1D1 through phosphorylation and its interaction with the cytosolic tail of insulin-regulated aminopeptidase IRAP. *J Biol Chem* **293** 17853-17862.
123. Makki K, Froguel P & Wolowczuk I 2013 Adipose tissue in obesity-related inflammation and insulin resistance: cells, cytokines, and chemokines. *ISRN Inflamm* **2013** 139239.
124. MayoClinic 2019 Prediabetes: Fasting blood sugar test. *Available at: https://www.mayoclinic.org/diseases-conditions/prediabetes/diagnosis-treatment/drc-20355284 10.12.19*.
125. McBride MW, Charchar FJ, Graham D, Miller WH, Strahorn P, Carr FJ & Dominiczak AF 2004 Functional genomics in rodent models of hypertension. *J Physiol* **554** 56-63.
126. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P & Cunningham F 2016 The Ensembl Variant Effect Predictor. *Genome Biol* **17** 122.
127. McLaughlin T, Ackerman SE, Shen L & Engleman E 2017 Role of innate and adaptive immunity in obesity-associated metabolic disease. *J Clin Invest* **127** 5-13.
128. Mietlicki-Baase EG 2016 Amylin-mediated control of glycemia, energy balance, and cognition. *Physiol Behav* **162** 130-140.
129. Morey JS, Ryan JC & Van Dolah FM 2006 Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proced Online* **8** 175-193.
130. Mueckler M & Thorens B 2013 The SLC2 (GLUT) family of membrane transporters. *Mol Aspects Med* **34** 121-138.
131. Muller TD, Finan B, Clemmensen C, DiMarchi RD & Tschop MH 2017 The New Biology and Pharmacology of Glucagon. *Physiol Rev* **97** 721-766.
132. Nadeau JHS, J.B.; Matin, A.; Lander, E.S. 2000 Analysing complex genetic traits with chromosome substitution strains. *Nature Genetics* **24(3)** 221-225.
133. Naidoo N, Pawitan Y, Soong R, Cooper DN & Ku CS 2011 Human genetics and genomics a decade after the release of the draft sequence of the human genome. *Hum Genomics* **5** 577-622.
134. NCBI-dbSNP 2019. *Available at: https://www.ncbi.nlm.nih.gov/snp/ 10.10.19*.
135. NCBIgene 2019. *Available at: https://www.ncbi.nlm.nih.gov/gene 10.10.19*.
136. Nedumpully-Govindan P & Ding F 2015 Inhibition of IAPP aggregation by insulin depends on the insulin oligomeric state regulated by zinc ion concentration. *Sci Rep* **5** 8240.
137. Nesher R, Anteby E, Yedovizky M, Warwar N, Kaiser N & Cerasi E 2002 Beta-cell protein kinases and the dynamics of the insulin response to glucose. *Diabetes* **51 Suppl 1** S68-73.
138. Ng PC & Henikoff S 2003 SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31** 3812-3814.

139. Nica AC & Dermitzakis ET 2013 Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci* **368** 20120362.
140. Olokoba ABO, O. A.; Olokoba, L.B. 2012 Type 2 Diabetes Mellitus: A Review of Current Trends. *Oman Medical Journal* **27**.
141. Omasits U, Ahrens CH, Muller S & Wollscheid B 2014 Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **30** 884-886.
142. Ouchi N, Parker JL, Lugus JJ & Walsh K 2011 Adipokines in inflammation and metabolic disease. *Nat Rev Immunol* **11** 85-97.
143. Ozcan L, Ergin AS, Lu A, Chung J, Sarkar S, Nie D, Myers MG, Jr. & Ozcan U 2009 Endoplasmic reticulum stress plays a central role in development of leptin resistance. *Cell Metab* **9** 35-51.
144. Paigen K & Petkov P 2010 Mammalian recombination hot spots: properties, control and evolution. *Nat Rev Genet* **11** 221-233.
145. Peck GR, Ye S, Pham V, Fernando RN, Macaulay SL, Chai SY & Albiston AL 2006 Interaction of the Akt substrate, AS160, with the glucose transporter 4 vesicle marker protein, insulin-regulated aminopeptidase. *Mol Endocrinol* **20** 2576-2583.
146. Pelligra BtA 2016 Analysis of gene expression in gonadal white adipose tissue of a murine model
147. for the identification of novel risk genes for obesity and type-2-diabetes. *Heinrich Heine University Düsseldorf, Germany.*
148. Petkov PM, Ding Y, Cassell MA, Zhang W, Wagner G, Sargent EE, Asquith S, Crew V, Johnson KA, Robinson P, et al. 2004 An efficient SNP system for mouse genome scanning and elucidating strain relationships. *Genome Res* **14** 1806-1811.
149. Poitout V & Robertson RP 2002 Minireview: Secondary beta-cell failure in type 2 diabetes--a convergence of glucotoxicity and lipotoxicity. *Endocrinology* **143** 339-342.
150. Prasad RB & Groop L 2015 Genetics of type 2 diabetes-pitfalls and possibilities. *Genes (Basel)* **6** 87-123.
151. Prentki M & Nolan CJ 2006 Islet beta cell failure in type 2 diabetes. *J Clin Invest* **116** 1802-1812.
152. Pritchard JK & Przeworski M 2001 Linkage disequilibrium in humans: models and data. *Am J Hum Genet* **69** 1-14.
153. Pulvirenti T, Giannotta M, Capestrano M, Capitani M, Pisanu A, Polishchuk RS, San Pietro E, Beznoussenko GV, Mironov AA, Turacchio G, et al. 2008 A traffic-activated Golgi-based signalling circuit coordinates the secretory pathway. *Nat Cell Biol* **10** 912-922.
154. Qi L, Kraft P, Hunter DJ & Hu FB 2008 The common obesity variant near MC4R gene is associated with higher intakes of total energy and dietary fat, weight change and diabetes risk in women. *Hum Mol Genet* **17** 3502-3508.
155. Qiao Q, Wei G, Yao D & Song Z 2019 Formation of alpha-helical and beta-sheet structures in membrane-bound human IAPP monomer and the resulting membrane deformation. *Phys Chem Chem Phys* **21** 20239-20251.
156. RADHAKRISHNAN SS, S. E.; GORODKIN, J.G. 2015 Emerging tools for rna structure analysis in polymorphic data. *BioZoom*.
157. Reifsnyder PC & Leiter EH 2002 Deconstructing and reconstructing obesity-induced diabetes (diabesity) in mice. *Diabetes* **51** 825-832.

158. Roat R, Rao V, Doliba NM, Matschinsky FM, Tobias JW, Garcia E, Ahima RS & Imai Y 2014 Alterations of pancreatic islet structure, metabolism and gene expression in diet-induced obese C57BL/6J mice. *PLoS One* **9** e86815.
159. Roden M & Shulman GI 2019 The integrative biology of type 2 diabetes. *Nature* **576** 51-60.
160. Roder PV, Wu B, Liu Y & Han W 2016 Pancreatic regulation of glucose homeostasis. *Exp Mol Med* **48** e219.
161. Ronkainen J, Huusko TJ, Soininen R, Mondini E, Cinti F, Makela KA, Kovalainen M, Herzig KH, Jarvelin MR, Sebert S, et al. 2015 Fat mass- and obesity-associated gene Fto affects the dietary response in mouse white adipose tissue. *Sci Rep* **5** 9233.
162. Rorsman P & Braun M 2013 Regulation of insulin secretion in human pancreatic islets. *Annu Rev Physiol* **75** 155-179.
163. Rorsman P, Eliasson L, Renstrom E, Gromada J, Barg S & Gopel S 2000 The Cell Physiology of Biphasic Insulin Secretion. *News Physiol Sci* **15** 72-77.
164. Rorsman P & Renstrom E 2003 Insulin granule dynamics in pancreatic beta cells. *Diabetologia* **46** 1029-1045.
165. Saltiel AR & Kahn CR 2001 Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* **414** 799-806.
166. Samuel VT & Shulman GI 2012 Mechanisms for insulin resistance: common threads and missing links. *Cell* **148** 852-871.
167. Sanger_SNP_Database 2020. *Available at: https://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1505 15.03.20*.
168. Sato H, Nagashima K, Ogura M, Sato Y, Tahara Y, Ogura K, Yamano G, Sugizaki K, Fujita N, Tatsuoka H, et al. 2016 Src regulates insulin secretion and glucose metabolism by influencing subcellular localization of glucokinase in pancreatic beta-cells. *J Diabetes Investig* **7** 171-178.
169. Saudek V 2012 Cystinosin, MPDU1, SWEETs and KDELR belong to a well-defined protein family with putative function of cargo receptors involved in vesicle trafficking. *PLoS One* **7** e30876.
170. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, et al. 2008 Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* **6** e107.
171. Schafer SA, Tschritter O, Machicao F, Thamer C, Stefan N, Gallwitz B, Holst JJ, Dekker JM, t Hart LM, Nijpels G, et al. 2007 Impaired glucagon-like peptide-1-induced insulin secretion in carriers of transcription factor 7-like 2 (TCF7L2) gene polymorphisms. *Diabetologia* **50** 2443-2450.
172. Schallschmidt T 2018 Identification of novel susceptibility genes for diabetes-related traits in a backcross of obese NZO with lean C3H mice. *Heinrich Heine University Düsseldorf, Germany.*
173. Schallschmidt T, Lebek S, Altenhofen D, Damen M, Schulte Y, Knebel B, Herwig R, Rasche A, Stermann T, Kamitz A, et al. 2018 Two Novel Candidate Genes for Insulin Secretion Identified by Comparative Genomics of Multiple Backcross Mouse Populations. *Genetics* **210** 1527-1542.
174. Scheen AJ 2003 PATHOPHYSIOLOGY OF TYPE 2 DIABETES. *Acta Clinica Belgica* **58-6**.
175. Scherneck S, Nestler M, Vogel H, Bluher M, Block MD, Berriel Diaz M, Herzig S, Schulz N, Teichert M, Tischer S, et al. 2009 Positional cloning of zinc finger domain transcription factor Zfp69, a candidate gene for obesity-associated diabetes contributed by mouse locus Nidd/SJL. *PLoS Genet* **5** e1000541.

176. Schwenk RW, Vogel H & Schurmann A 2013 Genetic and epigenetic control of metabolic health. *Mol Metab* **2** 337-347.
177. Scott RV & Bloom SR 2018 Problem or solution: The strange story of glucagon. *Peptides* **100** 36-41.
178. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orru M, Usala G, et al. 2007 Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* **3** e115.
179. Seino S, Shibasaki T & Minami K 2011 Dynamics of insulin secretion and the clinical implications for obesity and diabetes. *J Clin Invest* **121** 2118-2125.
180. Senn JJ, Klover PJ, Nowak IA & Mooney RA 2002 Interleukin-6 induces cellular insulin resistance in hepatocytes. *Diabetes* **51** 3391-3399.
181. Simon MM, Greenaway S, White JK, Fuchs H, Gailus-Durner V, Wells S, Sorg T, Wong K, Bedu E, Cartwright EJ, et al. 2013 A comparative phenotypic and genomic analysis of C57BL/6J and C57BL/6N mouse strains. *Genome Biol* **14** R82.
182. Skorecki KL, Lee JH, Langefeld CD, Rosset S, Tzur S, Wasser WG, Shemer R, Hawkins GA, Divers J, Parekh RS, et al. 2018 A null variant in the apolipoprotein L3 gene is associated with non-diabetic nephropathy. *Nephrol Dial Transplant* **33** 323-330.
183. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, et al. 2007 A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445** 881-885.
184. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gomez-Marin C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, et al. 2014 Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507** 371-375.
185. Southworth LK, Owen AB & Kim SK 2009 Aging mice show a decreasing correlation of gene expression within genetic modules. *PLoS Genet* **5** e1000776.
186. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ & Cheung VG 2007 Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet* **39** 226-231.
187. Storey JD, Madeoy J, Strout JL, Wurfel M, Ronald J & Akey JM 2007 Gene-expression variation within and among human populations. *Am J Hum Genet* **80** 502-509.
188. Strawbridge RJ, Dupuis J, Prokopenko I, Barker A, Ahlqvist E, Rybin D, Petrie JR, Travers ME, Bouatia-Naji N, Dimas AS, et al. 2011 Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* **60** 2624-2634.
189. Su WL, Sieberts SK, Kleinhanz RR, Lux K, Millstein J, Molony C & Schadt EE 2010 Assessing the prospects of genome-wide association studies performed in inbred mice. *Mamm Genome* **21** 143-152.
190. Sundar Rajan S, Srinivasan V, Balasubramanyam M & Tatu U 2007 Endoplasmic reticulum (ER) stress & diabetes. *Indian J Med Res* **125** 411-424.
191. T2DM-Knowledge-Portal 2019. *Available at: http://www.type2diabetesgenetics.org/ 10.10.19*.
192. Takeshita S, Kitayama S, Suzuki T, Moritani M, Inoue H & Itakura M 2012 Diabetic modifier QTL, Dbm4, affecting elevated fasting blood glucose concentrations in congenic mice. *Genes Genet Syst* **87** 341-346.

193. Takeshita S, Moritani M, Kunika K, Inoue H & Itakura M 2006 Diabetic modifier QTLs identified in F2 intercrosses between Akita and A/J mice. *Mamm Genome* **17** 927-940.
194. Tam V, Patel N, Turcotte M, Bosse Y, Pare G & Meyre D 2019 Benefits and limitations of genome-wide association studies. *Nat Rev Genet* **20** 467-484.
195. Tan Y, Gan M, Fan Y, Li L, Zhong Z, Li X, Bai L, Zhao Y, Niu L, Shang Y, et al. 2019 miR-10b-5p regulates 3T3-L1 cells differentiation by targeting Apol6. *Gene* **687** 39-46.
196. Taniguchi CM, Emanuelli B & Kahn CR 2006 Critical nodes in signalling pathways: insights into insulin action. *Nat Rev Mol Cell Biol* **7** 85-96.
197. Tao Y, Cheung LS, Li S, Eom JS, Chen LQ, Xu Y, Perry K, Frommer WB & Feng L 2015 Structure of a eukaryotic SWEET transporter in a homotrimeric complex. *Nature* **527** 259-263.
198. Tokarz VL, MacDonald PE & Klip A 2018 The cell biology of systemic insulin function. *J Cell Biol* **217** 2273-2289.
199. Tornatore L, Thotakura AK, Bennett J, Moretti M & Franzoso G 2012 The nuclear factor kappa B signaling pathway: integrating metabolism with inflammation. *Trends Cell Biol* **22** 557-566.
200. Townsley FM, Wilson DW & Pelham HR 1993 Mutational analysis of the human KDEL receptor: distinct structural requirements for Golgi retention, ligand binding and retrograde transport. *EMBO J* **12** 2821-2829.
201. Trujillo ME, Sullivan S, Harten I, Schneider SH, Greenberg AS & Fried SK 2004 Interleukin-6 regulates human adipose tissue lipid metabolism and leptin production in vitro. *J Clin Endocrinol Metab* **89** 5577-5582.
202. Trychta KA, Back S, Henderson MJ & Harvey BK 2018 KDEL Receptors Are Differentially Regulated to Maintain the ER Proteome under Calcium Deficiency. *Cell Rep* **25** 1829-1840 e1826.
203. Tucker PK, Lee BK, Lundrigan BL & Eicher EM 1992 Geographic origin of the Y chromosomes in "old" inbred strains of mice. *Mamm Genome* **3** 254-261.
204. Urano F, Wang X, Bertolotti A, Zhang Y, Chung P, Harding HP & Ron D 2000 Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1. *Science* **287** 664-666.
205. Van Baak TE, Coarfa C, Dugue PA, Fiorito G, Laritsky E, Baker MS, Kessler NJ, Dong J, Duryea JD, Silver MJ, et al. 2018 Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol* **19** 2.
206. Vogel H, Kamitz A, Hallahan N, Lebek S, Schallschmidt T, Jonas W, Jahnert M, Gottmann P, Zellner L, Kanzleiter T, et al. 2018 A collective diabetes cross in combination with a computational framework to dissect the genetics of human obesity and Type 2 diabetes. *Hum Mol Genet* **27** 3099-3112.
207. Wade CM, Kulbokas EJ, 3rd, Kirby AW, Zody MC, Mullikin JC, Lander ES, Lindblad-Toh K & Daly MJ 2002 The mosaic structure of variation in the laboratory mouse genome. *Nature* **420** 574-578.
208. Wang X, Korstanje R, Higgins D & Paigen B 2004 Haplotype analysis in multiple crosses to identify a QTL gene. *Genome Res* **14** 1767-1772.
209. Weibel ER 1969 Stereological principles for morphometry in electron microscopic cytology. *Int Rev Cytol* **26** 235-302.
210. Weisberg SP, McCann D, Desai M, Rosenbaum M, Leibel RL & Ferrante AW, Jr. 2003 Obesity is associated with macrophage accumulation in adipose tissue. *J Clin Invest* **112** 1796-1808.
211. Wergedal JE, Ackert-Bicknell CL, Beamer WG, Mohan S, Baylink DJ & Srivastava AK 2007 Mapping genetic loci that regulate lipid levels in a

NZB/B1NJxRF/J intercross and a combined intercross involving NZB/B1NJ, RF/J, MRL/MpJ, and SJL/J mouse strains. *J Lipid Res* **48** 1724-1734.
212. Wheeler E & Barroso I 2011 Genome-wide association studies and type 2 diabetes. *Brief Funct Genomics* **10** 52-60.
213. Wray GA 2007 The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* **8** 206-216.
214. Wu C & Shea JE 2013 Structural similarities and differences between amyloidogenic and non-amyloidogenic islet amyloid polypeptide (IAPP) sequences and implications for the dual physiological and pathological activities of these peptides. *PLoS Comput Biol* **9** e1003211.
215. Xi XX, Sun J, Chen HC, Chen AD, Gao LP, Yin J & Jing YH 2019 High-Fat Diet Increases Amylin Accumulation in the Hippocampus and Accelerates Brain Aging in hIAPP Transgenic Mice. *Front Aging Neurosci* **11** 225.
216. Xu H, Barnes GT, Yang Q, Tan G, Yang D, Chou CJ, Sole J, Nichols A, Ross JS, Tartaglia LA, et al. 2003 Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. *Journal of Clinical Investigation* **112** 1821-1830.
217. Xu J, Wang J & Chen B 2012 SLC30A8 (ZnT8) variations and type 2 diabetes in the Chinese Han population. *Genet Mol Res* **11** 1592-1598.
218. Xu X, D'Hoker J, Stange G, Bonne S, De Leu N, Xiao X, Van de Casteele M, Mellitzer G, Ling Z, Pipeleers D, et al. 2008 Beta cells can be generated from endogenous progenitors in injured adult mouse pancreas. *Cell* **132** 197-207.
219. Xu Y, Tao Y, Cheung LS, Fan C, Chen LQ, Xu S, Perry K, Frommer WB & Feng L 2014 Structures of bacterial homologues of SWEET transporters in two distinct conformations. *Nature* **515** 448-452.
220. Yamamoto K, Hamada H, Shinkai H, Kohno Y, Koseki H & Aoe T 2003 The KDEL receptor modulates the endoplasmic reticulum stress response through mitogen-activated protein kinase signaling cascades. *J Biol Chem* **278** 34525-34532.
221. Yamazaki H, Hiramatsu N, Hayakawa K, Tagawa Y, Okamura M, Ogata R, Huang T, Nakajima S, Yao J, Paton AW, et al. 2009 Activation of the Akt-NF-kappaB pathway by subtilase cytotoxin through the ATF6 branch of the unfolded protein response. *J Immunol* **183** 1480-1487.
222. Yu JaB, F. C. 2016 Hypoglycemia and PDX 1 Targeted Therapy. *Available at> https://www.semanticscholar.org/ 10.12.2019*.
223. Zhang J, Finney RP, Clifford RJ, Derr LK & Buetow KH 2005 Detecting false expression signals in high-density oligonucleotide arrays by an in silico approach. *Genomics* **85** 297-308.
224. Zheng Y, Ley SH & Hu FB 2018 Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol* **14** 88-98.

7 SUPLEMENTS

7.1 List of publication and scientific contributions

Tanja Schallschmidt**,** Sandra Lebek, **Delsi Altenhofen**, Mareike Damen, Yvonne Schulte, Birgit Knebel, Ralf Herwig, Axel Rasche, Torben Stermann, Anne Kamitz, Nicole Hallahan, Markus Jähnert, Heike Vogel, Annette Schürmann, Alexandra Chadt and Hadi Al-Hasani (2018): "Two novel candidate genes for insulin secretion identified by comparative genomics of multiple backcross mouse populations". *GENETICS* 210:1527–1542.

Contribution: The results generated in this thesis originated from the first analysis performed in this paper. Here I contributed to the haplotype analysis (Fig.6) and to the correction and revision of the manuscript.

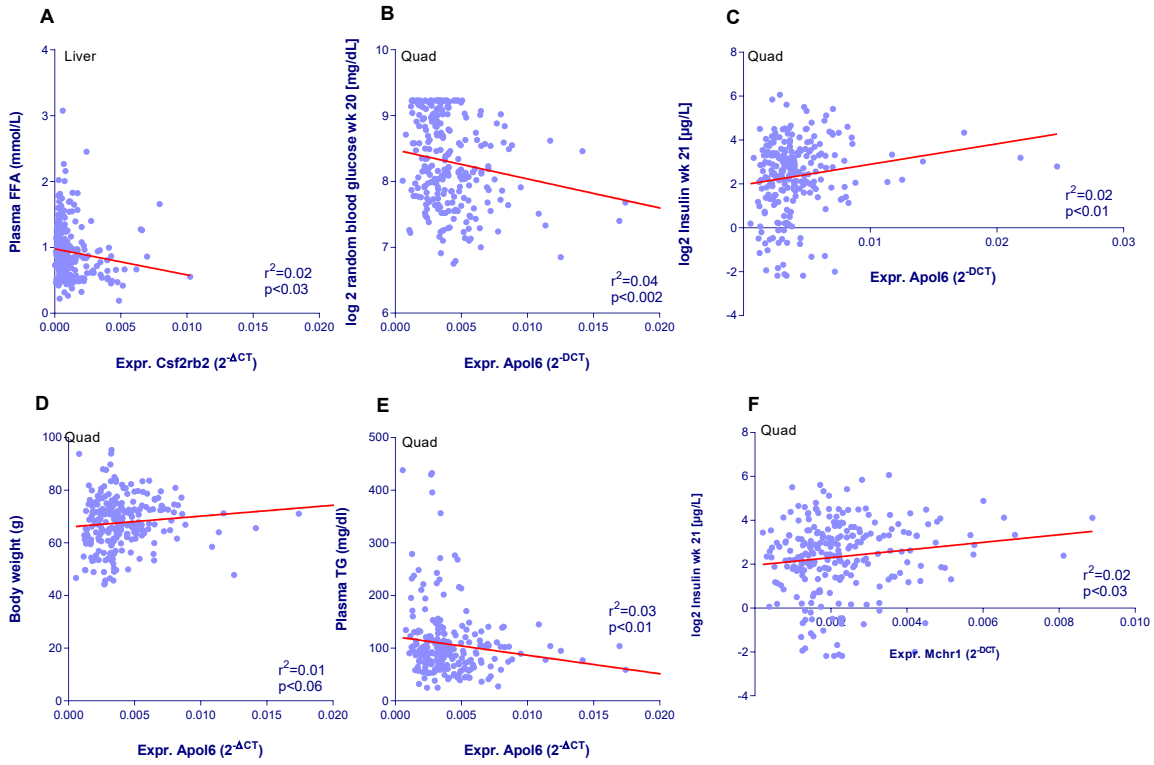
Heike Vogel, Anne Kamitz, Nicole Hallahan, Sandra Lebek, Tanja Schallschmidt, Wenke Jonas, Markus Jähnert, Pascal Gottmann, Lisa Zellner, Timo Kanzleiter, Mareike Damen, **Delsi Altenhofen**, Ralph Burkhardt, Simone Renner, Maik Dahlhoff, Eckhard Wolf, Timo D. Müller, Matthias Blüher, Hans-Georg Joost, Alexandra Chadt, Hadi Al-Hasani and Annette Schürmann (2018): "A collective diabetes cross in combination with a computational framework to dissect the genetics of human obesity and type 2 diabetes". *Human Molecular Genetics* 27: 3099-3112.

Contribution: This paper is part of the collective diabetes cross with the aim to identify novel diabetes susceptibility genes. In this paper I participated in the phenotyping of the parental mice which is summarized in Figures S1 and S2.

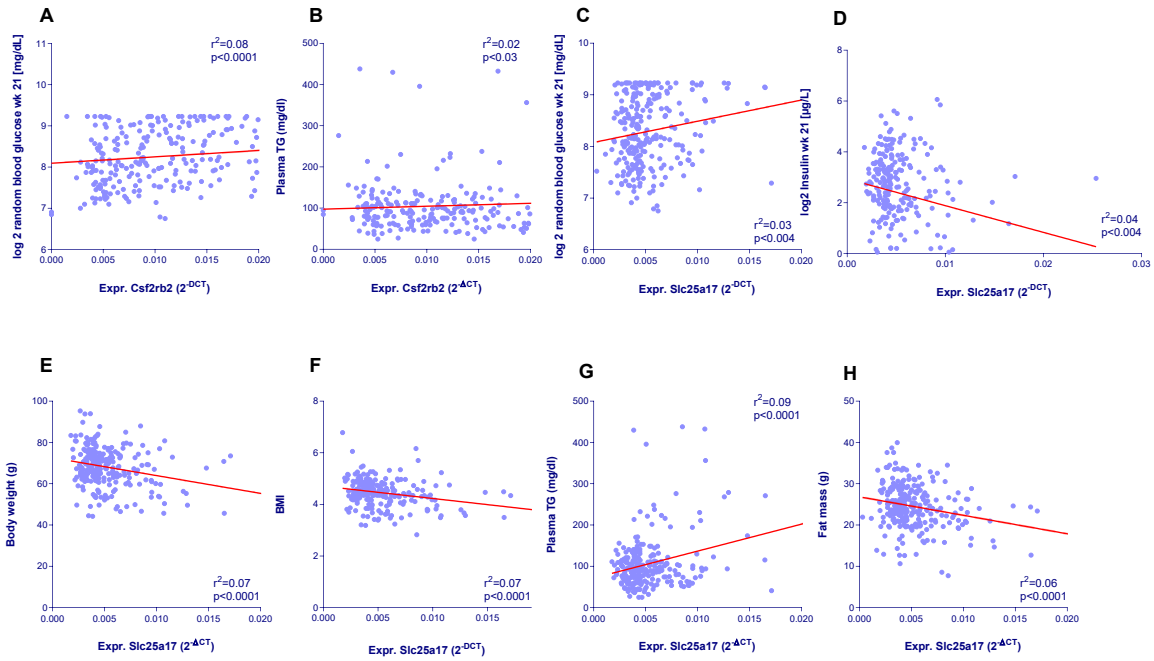
Torben Stermann, Franziska Menzel, Carmen Weidlich, Kay Jeruschke, Jürgen Weiss, **Delsi Altenhofen**, Tim Benninghoff, Anna Pujol, Fatima Bosch, Ingo Rustenbeck, D. Margriet Ouwens, G. Hege Thoresen, Christian de Wendt, Sandra Lebek, Tanja Schallschmidt, Martin Kragl, Eckhard Lammert,, Alexandra Chadt, and Hadi Al-Hasani (2018): Deletion of the RabGAP TBC1D1 Leads to Enhanced Insulin Secretion and Fatty Acid Oxidation in Islets From Male Mice. Endocrinology 159(4):1748–1761.

Contribution: In this paper I have performed the determination of pro-insulin content (Figure 3) and the pancreatic islet isolation for the palmitate uptake experiments (Figure 6).

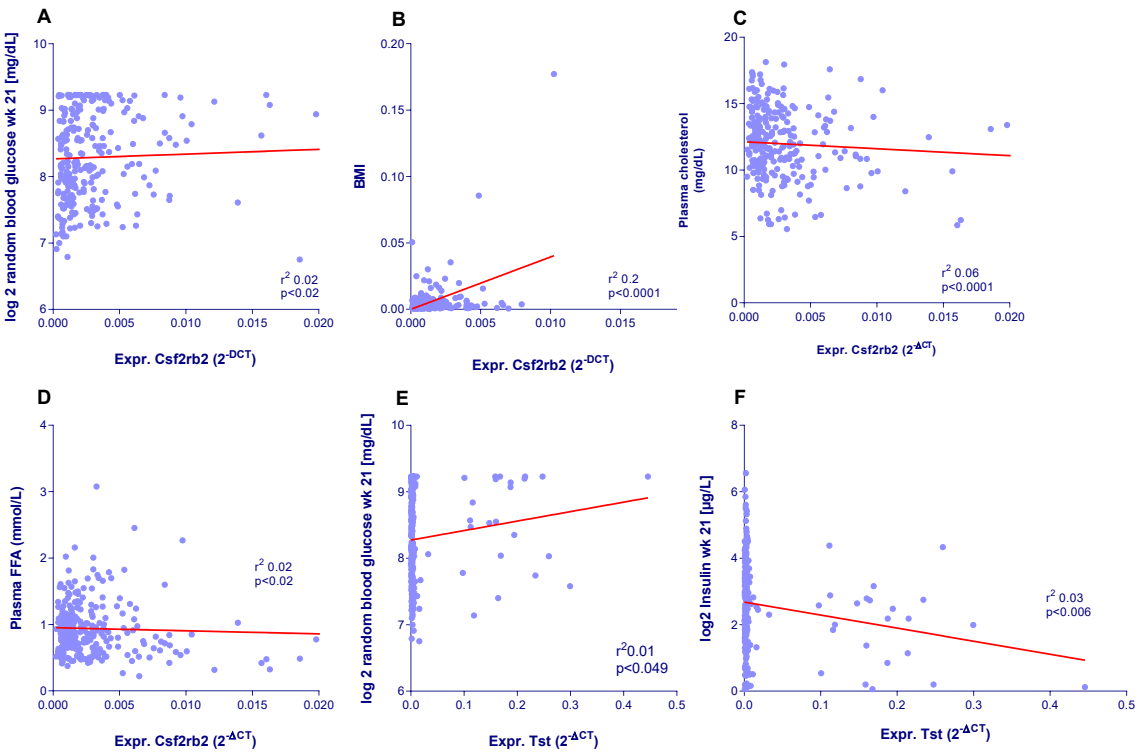
7.2 Supplement figures



Supplementary Figure 1. Significant correlations of gene expression and metabolic traits in the quadriceps and liver of the N_2(C3HxNZO) population. Expression of *Csf2rb2* in the liver significantly correlated to plasmatic levels of Free-fatty acids (A). Expression of *Apol6* in the quadriceps (Quad) correlated to blood glucose levels (B), plasma insulin (C), body weight (D) and plasma triglycerides (TG) (E). Expression of *Mchr1* in the quadriceps correlated to plasma insulin (F). Significant correlations were determined by using linear regression analysis indicated by the coefficient of determination r^2 (R squared); r^2=1 represents 100 % correlation or 100 % of the phenotypic variation is explained by the gene expression. Expr= expression.



Supplementary Figure 2. Significant correlations of gene expression and metabolic traits in the gonadal adipose tissue of the N_2(C3HxNZO) population. Expression of *Csf2rb2* in the gWAT significantly correlated to blood glucose levels (A) and plasma triglycerides (TG). *Slc25a17* correlated to blood glucose (C), plasma insulin (D), body weight (E), body mass index (BMI) (F), plasma TG (G) and fat mass (H). Significant correlations were determined by using linear regression analysis indicated by the coefficient of determination r^2 (R squared); r^2=1 represents 100 % correlation or 100 % of the phenotypic variation is explained by the gene expression. Expr= expression.



Supplementary Figure 3. Significant correlations of gene expression and metabolic traits in the gonadal adipose tissue of the N_2(C3HxNZO) population. Expression of *Csf2rb2* in the BAT significantly correlated to blood glucose levels (A), body mass index (BMI) (B), cholesterol levels (C) and plasma free fatty acids (FFA) (D). *Tst* correlated to blood glucose (E) and plasma insulin (F). Significant correlations were determined by using linear regression analysis indicated by the coefficient of determination r^2 (R squared); r^2=1 represents 100 % correlation or 100 % of the phenotypic variation is explained by the gene expression. Expr= expression.

7.3 Supplement Tables

Supplementary Table 1 – Identification of coding amino acid substitution in the *Nbg15* candidate genes

Gene	SNP-ID	Codon	Amino acid	Protein position	Domain
Apol6	rs261533897	GA**G** ⇒ GA**C**	E [Glu] ⇒ D [Asp]	122	Transmembrane
	rs234397075	**A**TG ⇒ **T**TG	M [Met] ⇒ L [Leu]	144	Transmembrane
	rs240031711	G**A**T ⇒ G**C**T	D [Asp] ⇒ A [Ala]	256	Transmembrane
Apol7a	rs229171902	A**C**C ⇒ A**T**C	T [Thr] ⇒ I [Ile]	272	Low complexity
	rs245423796	A**T**C ⇒ A**C**C	I [Ile] ⇒ T [Thr]	216	Low complexity
Kdelr3	rs256912165	G**A**G ⇒ G**T**G	E [Glu] ⇒ V [Val]	96	Transmembrane helical

Supplementary Table 2 – Summary of phenotyping of the RCS.NZO.C3H.*Nbg15*.

		Nbg15$^{NZO/NZO}$		*Nbg15*-c$^{C3H/C3H}$			*Nbg15*-p$^{C3H/C3H}$			*Nbg15*-d$^{C3H/C3H}$		
	Weeks of age	Mean	SEM	Mean	SEM	p value	Mean	SEM	p value	Mean	SEM	p value
Random blood glucose (mg/dl)	3	149.92	4.81	153.12	4.78	0.9991	141.64	10.35	0.9862	161.42	4.76	0.9662
	4	216.00	5.04	225.14	5.79	0.9854	214.15	7.50	0.9999	232.73	7.75	0.9431
	5	220.29	5.79	209.32	9.36	0.9754	205.05	5.34	0.9433	244.93	8.40	0.8454
	6	238.42	11.96	205.20	7.59	0.5585	205.56	8.13	0.5666	202.04	12.40	0.4973
	7	252.03	16.42	222.41	12.56	0.6938	221.40	18.14	0.6907	248.13	9.30	0.9991
	8	277.58	20.63	239.95	16.44	0.5234	255.40	24.59	0.8502	289.93	24.68	0.9756
	9	345.45	26.33	241.91	22.23	0.0048	300.10	37.09	0.3960	296.80	20.36	0.4140
	10	365.00	23.81	239.80	24.53	0.0001	324.24	30.14	0.3953	353.17	23.88	0.9633
	11	398.13	29.22	263.18	25.08	0.0001	316.35	39.61	0.0421	328.93	30.05	0.1538
	12	417.84	30.69	264.23	24.99	0.0001	312.65	36.10	0.0055	331.53	31.30	0.0533
	13	401.87	32.66	242.41	22.54	0.0001	338.65	44.91	0.1564	342.20	37.45	0.2571
	14	393.66	34.17	235.86	26.39	0.0001	341.30	43.03	0.2922	256.93	36.57	0.0008
	15	361.78	33.06	239.88	24.83	0.0003	307.88	41.08	0.2165	268.73	23.33	0.0125

		Nbg15$^{NZO/NZO}$		*Nbg15*-c$^{C3H/C3H}$			*Nbg15*-p$^{C3H/C3H}$			*Nbg15*-d$^{C3H/C3H}$		
	Weeks of age	**Mean**	**SEM**	**Mean**	**SEM**	**p value**	**Mean**	**SEM**	**p value**	**Mean**	**SEM**	**p value**
Random body weight (g)	3	12.72	0.37	14.05	0.39	0.4389	13.82	0.72	0.6030	14.05	0.67	0.4641
	4	22.27	0.63	20.80	0.59	0.4348	21.33	1.01	0.7673	20.09	0.89	0.2118
	5	31.41	0.58	30.35	0.49	0.6837	29.41	0.95	0.2089	30.29	0.77	0.7230
	6	38.04	0.47	36.85	0.49	0.5300	36.18	0.46	0.1838	37.83	0.80	0.9949
	7	42.82	0.52	42.20	0.70	0.9142	40.97	0.48	0.2677	43.27	0.56	0.9741
	8	47.36	0.41	45.86	0.66	0.4182	44.94	0.47	0.0965	49.09	0.44	0.3997
	9	49.87	0.39	49.27	0.78	0.9195	47.84	0.32	0.1963	51.78	0.48	0.3184
	10	52.27	0.54	52.09	0.75	0.9970	49.05	0.31	0.0055	55.23	0.59	0.0150
	11	53.41	0.73	54.41	1.07	0.7242	50.03	0.58	0.0097	57.97	0.65	0.0010
	12	54.42	0.82	57.54	1.15	0.0156	51.26	0.84	0.0173	60.17	0.66	<0.0001
	13	54.64	1.04	58.25	1.18	0.0040	51.50	1.06	0.0195	60.88	1.09	<0.0001
	14	53.63	1.30	59.99	1.24	<0.0001	52.57	1.31	0.7087	60.29	1.12	<0.0001
	15	54.61	1.17	61.30	1.35	<0.0001	53.03	1.27	0.3418	62.46	1.49	<0.0001
Fat mass (g)	3	2.12	0.15	2.10	0.12	0.9999	2.66	0.22	0.9142	1.92	0.14	0.9961
	6	9.92	0.27	9.40	0.27	0.8970	9.49	0.22	0.9388	9.91	0.59	0.9999
	10	18.02	0.59	18.60	0.49	0.8641	16.93	0.33	0.5027	19.51	0.90	0.2647
	15	19.16	1.15	26.26	0.93	0.0001	19.38	1.24	0.9927	25.51	1.40	0.0001
Lean mass (g)	3	10.46	0.29	11.74	0.34	0.0512	11.30	0.62	0.3086	11.65	0.52	0.1341
	6	26.95	0.32	26.67	0.30	0.8905	25.35	0.32	0.0043	26.87	0.45	0.9964
	10	32.41	0.46	31.80	0.38	0.4585	30.15	0.25	0.0001	33.25	0.29	0.2403
	15	32.96	0.31	33.09	0.49	0.9888	31.12	0.31	0.0016	34.36	0.34	0.0256

		Nbg15$^{NZO/NZO}$		*Nbg15*-c$^{C3H/C3H}$			*Nbg15*-p$^{C3H/C3H}$			*Nbg15*-d$^{C3H/C3H}$		
	Weeks of age	**Mean**	**SEM**	**Mean**	**SEM**	**p value**	**Mean**	**SEM**	**p value**	**Mean**	**SEM**	**p value**
16h fasted BG (mg/dl)	12	226.58	29.95	109.59	7.25	0.0166	202.15	40.71	0.9202	138.38	9.63	0.1612
2h refed BG(mg/dl)	12	390.06	33.64	213.68	25.89	0.0001	350.85	44.24	0.7399	269.00	27.55	0.0276
16h fasted Insulin (μg/L)	12	2.27	0.21	2.45	0.23	0.9925	1.70	0.18	0.9421	2.62	0.35	0.9759
2h refed Insulin (μg/L)	12	10.79	1.50	14.60	1.81	0.0497	10.37	2.11	0.9679	14.32	1.76	0.0930
16h fasted TG (mg/dl)	12	158.64	7.76	118.25	10.23	0.2806	165.85	17.86	0.9835	130.86	6.28	0.4239
2h refed TG (mg/dl)	12	303.39	21.22	222.84	16.54	0.0035	354.08	67.26	0.4219	176.87	7.26	<0.0001
16h fasted FFA (mmo/L)	12	0.93	0.06	1.05	0.08	0.4094	1.13	0.08	0.0619	0.90	0.03	0.9802
2h refed FFA (mmo/L)	12	0.74	0.04	0.77	0.03	0.9748	0.88	0.11	0.2219	0.58	0.03	0.1767
GTT_BG AUC	13	13.90	0.93	9.60	0.72	0.0011	13.20	1.11	0.9278	12.28	0.51	0.5217
GTT_Insulin AUC	13	0.07	0.01	0.09	0.01	0.4855	0.06	0.01	0.3091	0.08	0.01	0.6939
ITT AUC	14	5.50	0.198	5.37	0.398	0.9672	5.50	0.100	>0.9999	5.41	0.194	0.9866
FFBG (mg/dl)	15	358.32	32.16	190.50	23.38	0.0002	284.96	40.42	0.2229	173.62	13.02	0.0001
FFI (μg/L)	15	5.89	0.82	12.83	1.86	0.0021	4.11	0.53	0.7124	15.16	2.35	0.0001
Liver weight (g)	15	2.39	0.09	2.37	0.13	0.9981	2.14	0.10	0.2352	2.53	0.15	0.7435
gWAT weight (g)	15	2.40	0.17	2.96	0.21	0.1107	2.42	0.18	0.9996	3.14	0.21	0.0287
scWAT weight (g)	15	2.83	0.25	4.05	0.26	0.0028	2.80	0.28	0.9998	4.62	0.16	<0.0001
Average number of islets	15	32.80	5.42	48.71	8.54	0.3273	35.70	4.76	0.9875	44.22	9.33	0.5909
Average islets number (cluster 0-5000 mm^2)	15	12.21	1.94	18.85	3.62	<0.0001	13.75	0.64	0.1664	15.86	0.89	<0.0001
Average pancreas area (mm^2)	15	48.22	3.76	40.92	6.76	0.5927	33.14	3.45	0.0890	38.40	1.50	0.3560
Islet area/pancreas area (mm^2)x1000	15	0.37	0.05	0.59	0.05	0.1229	0.54	0.12	0.2663	0.61	0.07	0.0833

Supplementary Table 2 – Summary of phenotyping of the RCS.B6.C3H.*Nbg15*.

		B*Nbg15*$^{B6/B6}$		B*Nbg15*$^{C3H/C3H}$		
	Weeks of age	Mean	SEM	Mean	SEM	p value
Random blood glucose (mg/dl)	3	157.62	11.09	132.79	5.37	0.8387
	4	191.08	8.75	193.37	7.85	>0.9999
	5	191.31	10.93	182.84	5.72	>0.9999
	6	188.92	8.01	175.53	10.70	0.9998
	7	170.08	7.69	161.74	7.65	>0.9999
	8	172.46	8.83	162.11	5.65	>0.9999
	9	165.77	5.91	158.21	5.68	>0.9999
	10	170.54	6.92	157.63	4.09	0.9999
	11	161.77	5.76	158.37	5.04	>0.9999
	12	173.46	6.31	168.16	6.09	>0.9999
	13	181.92	6.61	166.68	6.02	0.9989
	14	167.23	5.39	167.11	4.63	>0.9999
	15	190.31	5.52	176.68	6.58	0.9998
	16	178.15	10.73	156.00	4.11	0.9322
	17	188.15	14.97	165.47	6.13	0.9177
	18	213.15	15.99	185.21	12.28	0.6764
	19	194.54	15.75	207.11	16.79	>0.9999
	20	181.77	12.62	204.63	17.11	0.9122
	21	198.85	9.41	198.42	24.54	>0.9999
Random body weight (g)	3	11.07	0.84	11.38	0.77	>0.9999
	4	17.66	0.94	19.08	1.07	>0.9999
	5	21.47	0.67	22.82	0.66	>0.9999
	6	23.97	0.74	25.45	0.51	0.9999
	7	26.48	0.91	27.63	0.46	>0.9999
	8	28.28	1.01	29.19	0.46	>0.9999
	9	29.94	1.10	31.24	0.56	>0.9999
	10	31.90	1.10	32.23	0.60	>0.9999
	11	32.48	1.11	32.57	0.56	>0.9999
	12	34.28	1.26	34.14	0.75	>0.9999
	13	35.78	1.39	35.92	0.81	>0.9999
	14	36.85	1.48	35.91	0.82	>0.9999
	15	36.44	2.03	37.74	1.01	>0.9999
	16	39.72	1.74	39.05	1.15	>0.9999
	17	42.00	1.73	41.22	1.31	>0.9999
	18	43.80	1.83	42.88	1.46	>0.9999
	19	44.66	1.91	44.49	1.52	>0.9999
	20	45.94	1.85	45.42	1.47	>0.9999
	21	46.12	1.76	45.71	1.54	>0.9999
Fat mass (g)	3	0.94	0.12	1.02	0.09	>0.9999
	6	2.32	0.15	2.48	0.15	0.9995
	10	6.66	0.73	4.70	0.54	0.0849
	15	11.09	1.23	7.86	0.80	0.0008
Lean mass (g)	3	10.07	0.68	10.37	0.64	0.9950
	6	20.95	0.63	22.60	0.44	0.2459
	10	24.74	0.77	27.39	0.39	0.0147
	15	26.56	1.11	29.08	0.43	0.0226

		B*Nbg15*$^{B6/B6}$		B*Nbg15*$^{C3H/C3H}$		
	Weeks of age	**Mean**	**SEM**	**Mean**	**SEM**	**p value**
16h fasted BG (mg/dl)	10	98.67	4.08	84.00	3.80	0.0308
2h refed BG(mg/dl)	10	156.20	4.02	164.89	4.47	0.2695
16h fasted Insulin (µg/L)	10	0.44	0.09	0.42	0.10	0.9534
2h refed Insulin (µg/L)	10	2.50	0.24	3.34	0.34	0.0147
16h fasted TG (mg/dl)	10	75.33	10.58	55.44	2.89	0.4623
2h refed TG (mg/dl)	10	208.07	20.57	142.55	12.45	0.0007
16h fasted FFA (mmo/L)	10	1.44	0.12	1.61	0.06	0.2452
2h refed FFA (mmo/L)	10	1.08	0.06	0.90	0.08	0.2254
Ratio F/R FFA(mmo/L)	10	1.39	0.14	2.04	0.18	0.0129
GTT_BG AUC	13	9.74	0.34	7.66	0.42	0.0020
GTT_Insulin AUC	13	0.06	0.01	0.05	0.01	0.8270
ITT AUC	15	0.91	0.06	0.77	0.03	0.0298
FFBG (mg/dl)	21	188.20	9.30	193.70	9.75	0.7158
FFI (µg/L)	21	4.64	0.75	6.52	1.06	0.2387
Liver weight (g)	21	1.72	0.19	1.57	0.10	0.4416
gWAT weight (g)	21	1.80	0.16	1.45	0.09	0.0521
scWAT weight (g)	21	2.05	0.22	1.37	0.15	0.0150

7.4 List of Figures

7.5 List of abbreviations

a.u.	Arbitrary units
Actb	*Beta actin*
ANOVA	Analysis of variance
AUC	Area under the curve
B6	C57BL/6J
BAT	Brown adipose tissue
BMI	Body mass index
BNbg15$^{B6/B6}$	Chromosome 15 homozygous B6 allele carriers
BNbg15$^{C3H/C3H}$	Chromosome 15 homozygous C3H allele carriers on B6 genetic background
Bp	Base pairs
BSA	Bovine serum albumin
C3H	C3HeB/FeJ mouse
cDNA	Complementary DNA
Chr.	Chromosome
cM	Centimorgan
con	Consomic
Ct	Cycle threshold
DDZ	Deutsches Diabetes Zentrum Düsseldorf (German Diabetes Cener)
DMEM	Dulbecco's Modified Eagles Medium
DNA	Desoxyribo nucleic acid
dNTP	Desoxyribo nucleotide triphosphate
EDTA	Ethylene diamine tetra acetate
ELISA	Enzyme-linked immuno sorbent assay
eQTL	Expression QTL
et al.	Latin: et alia - and others
F1	First filial generation
FFA	Free fatty acids
Fig.	Figure
GSIS	Glucose-stimulated insulin secretion
GWAs	Genome-wide association study
gWAT	white adipose tissue
h	Hour
HFD	High-fat diet
i.p.	Intraperitoneal
i.p.GTT	Intraperitoneal glucose tolerance test
i.p.ITT	Intraperitoneal insulin tolerance test
IBD	Identity by descent
Ins	Insulin
KASP	Kompetitive Allele Specific PCR
kcal	Kilo calories
KRH	Krebs-Ringer-HEPES
LOD	Logarithm of the odds
max.	Maximal
Mb	Mega base pairs
Min	Minute
MIN6	Mouse insulinoma cell line
MODY	Maturity-onset diabetes of the young

mRNA	Messenger RNA
n	Number
N2	Backcross
Nbg15	NZO blood glucose locus on chromosome 15
Nbg15c$^{C3H/C3H}$	Chromosome 15 homozygous C3H allele carriers on NZO genetic background
Nbg15d$^{C3H/C3H}$	Distal chromosome 15 homozygous C3H allele carriers
Nbg15$^{NZO/NZO}$	Chromosome 15 homozygous NZO allele carriers
Nbg15p$^{C3H/C3H}$	Proximal chromosome 15 homozygous C3H allele carriers
Nir4	NZO insulin resistance on distal chromosome 4
NMR	Nuclear magnetic resonance
ns	Not significant
NZB	New Zealand black
NZO	New Zealand obese
Pos	Position
qPCR	Quantitative polymerase chain reaction
QTL	Quantitative trait locus
Quadr.	Quadriceps
r2	Coefficient of determination
rcf	Relative centrifugal force
RCS	Recombinant congenic strain
Rev	Reverse
Rhpn2	Rhophilin, Rho GTPase binding protein 2
RNA	Ribonucleic acid
RT	Room temperature
SAT	Subcutanous adipose tissue
sec	Seconds
sec	Seconds
SEM	Standard error of the mean
siRNA	Small interfering RNA
SNP	single nucleotide polymorphism
Suppl.	Supplementary
T1D	Type 1 diabetes mellitus
T2D	Type 2 diabetes mellitus
Tab.	Table
TCA	Tricarboxylic acid cycle
TG	Triglycerides
TG	Triglycerides
tRNA	Transfer RNA
Vs.	Versus
WHO	World health organization
wk	Week

Acknowledgments

I have learned that time is the best gift you can get from someone, is something that never comes back. In the last years I was gifted by many people that provide me with their precious time and the best support one can have.

I am immensely thankful to my supervisors Prof. Dr. Hadi Al-Hasani and Dr. Alexandra Chadt. Thank you for accepting me in your lab and for all the opportunities that you provided to me. Thank you also for teaching me the very complicated but fascinating world of genetics. Thank you Alex for your great support and for being always available for discussions and planning the experiments. I am also very thankful to my co-supervisor Prof. Dr. Axel Gödecke for always having time for great discussions and for giving great ideas for this project.

I am very thankful to the QTL crew, especially to Dr. Tanja Schallschmidt and Dr. Sandra Lebek who initiated this project. Thank you Tanja for all the work you have done to make the running of the project on chromosome 15 possible. Thank you Sandra and Tanja, for the nice time in and outside the lab and for being very supportive colleagues during all the time. You made me feel very welcomed in the lab!

I also thank the youngest members of the QTL crew Sarah Görigk and Jenny Khoung. Thank you for your support and understanding, for the great discussions about the projects and great atmosphere in the lab. I also thank you Jenny for helping in the histology analysis during your master thesis.

A big thank to all the technicians who helped in experiments and for also kindly introducing me the techniques. I feel really lucky to be part of such a great team from whom I have also learned a lot. Thank you Angelika Horrighs, Anette Kurowski, Heidrun Podini, Antonia Osmers, Dagmar Grittner and Birgit Knobloch. Thank you to Lothar Bohne for his great work in the determination of the fat cell size. I also want to express my gratitude to Nicole Krause for helping me a lot to find accommodation and to arrange the documents for my matriculation and general bureaucratic requirements.

I am also very thankful to Dr. Birgit Knebel for performing the microarray analysis and for being very supportive to solve our questions. Thank you also to Axel Rasche for the statistical analysis of the microarray data. Many thanks also to Dr. Sonja Hartwig for helping in the measurement of the cytokines. Thank you to Dr. Jürgen Weiß and Kay Jeruschke for the support in the histology analysis.

Moreover, this project could not have been carry out without the great help of all animals caretakers. Therefore, I will be always very thankful to all of you. I specially thank Peter Herdt for the great work he has done to generate the mouse lines.

Thank you to the bachelor students Angela Pelligra, Maik Hoeffs and Janek Masuch for the help in the generation of data for this project. Thank you Angela for also helping in the genotyping and for the funny time we spent outside the lab as well.

I am also very thankful to all the colleagues at the DDZ. Thank you for the nice atmosphere and for the very funny moments during lunch time. I specially thank Dr.

Torben Stermann for also introducing me the pancreatic islets isolation and for giving me the opportunity to collaborate with his project. My special thanks to Christian Binsch for always offering his great computational knowledge to improve our analysis and solve related problems. Thank you to Dr. Pia Fahlbusch for introducing the hepatocytes isolation technique and for helping giving input in the miRNA analysis. Thank you to Dr. Dhiraj Kabra for the great suggestions and for introducing me the delicious Indian culinary. Thank you to Daebin Kim for the funny moments and taking the cooking rules very serious providing delicious Korean food to a very hungry group of students. Thank you to Christian de Wendt for the nice time in the lab and during the great time we were the representative students in the DDZ and in other events.

I am also very thankful to have been gifted with the best friend one can have. Thank Dr. Samaneh Eickelschulte for being the person I can always count on.

I am so thankful for being surrounded by special friends. Thank you to my friends from the other side of the Atlantic ocean who always motivated me to follow my goals. Thank you to the friends that live nearby for supporting and encouraging me. My special thanks to Heejin, Amanda, Barbara, Emília, Santos, Franscesco and Sebastian.

I am immensely grateful to my family for the endless support, for helping me to stand all difficulties and for the understanding. My special thanks to my husband Simão for your never-end support, for have embarked with me in this journey and for helping me to stand difficult moments providing me strength and sometimes making fun of everything. The last 14 years of my live would have been very boring without you. Love you all!

I would also like to express my gratitude to the German Academic Exchange Service (DAAD) and the Brazilian National Council for Scientific and Technological Development (CNPq) for the financial support and for giving me the opportunity to perform my full doctoral study in Germany.

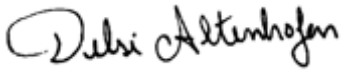
Thank you to the German Center for Diabetes Research (DZD) for providing the funding to execute this project.

Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Ferner versichere ich, dass ich bisher noch keinen Promotionsversuch unternommen habe. Die Arbeit wurde bisher an keiner anderen Hochschule eingereicht.

Düsseldorf, den 30.04.2020



(Delsi Altenhofen)