Technische Universität München

Fakultät für Medizin

Heterogeneity of Childhood Diabetes

Markus Hippich

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzender: Prof. Dr. Heiko Lickert

Prüfende/-r der Dissertation:

1. Prof. Dr. Anette-Gabriele Ziegler

2. Prof. Dr. Dr. Fabian Theis

Die Dissertation wurde am 20.03.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 13.08.2019 angenommen.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AAB	autoantibody
ADA	American Diabetes Association
BMI	body mass index
BMI SDS	body mass index standard deviation score
CART	classification and regression tree
CNRQ	Calibrated Normalized Relative Quantities
DCCT	Diabetes Control and Complications Trial
DIPP	Type 1 Diabetes Prediction and Prevention
FDR	false-discovery-rate
GAD	glutamic acid decarboxylase
GRS	genetic risk scores
GWAS	genome-wide association studies
HLA	human leukocyte antigen
HRs	Hazard ratios
IA2	insulinoma-associated autoantigen 2
IFN	interferon
IGRP	Islet-Specific Glucose-6-Phosphatase-Related Protein
IQR	interquartile range
LD	linkage disequilibrium
LIPS	Luciferase Immunoprecipitation System
LOD	limit of detection
MAF	minor allele frequencies
MODY	maturity-onset diabetes of the young
OR	odds ratio
PBMCs	peripheral blood mononuclear cells
qPCR	quantitative polymerase chain reaction
RBA	radiobinding assays
SNP	single nucleotide polymorphism
T1D	type 1 diabetes
T2D	type 2 diabetes
TEDDY	The Environmental Determinants of Diabetes in the Young
TPO	thyroid peroxidase
TSPAN7	Tetraspanin-7
VNTR	variable number tandem repeat
WTCCC	Wellcome Trust Case Control Consortium
ZnT8	zinc transporter 8

SUMMARY

The most prevalent type of diabetes in childhood and adolescence is type 1 diabetes, followed by type 2 diabetes and other types of diabetes. Differential diagnosis and classification of diabetes type, especially in islet autoantibody negative patients, is challenging and exacerbated due to high costs of genetic typing for monogenic forms of diabetes. In islet autoantibody positive patients, variation in onset of autoimmunity and subsequently progression to clinical manifestation additionally implies the presence of heterogenous endotypes of type 1 diabetes. Yet, a systematic analysis of endotypes of type 1 diabetes is missing, as is a description of the variability and associated factors of residual β-cell function at the time of diagnosis. This is clinically relevant as residual β-cell function correlates with quality of metabolic control and secondary diseases and, thus, indicates a possible benefit for the application of personalized therapies. The risk of developing type 1 diabetes is influenced by genetic susceptibility, family history of the disease, environmental factors and their interplay which additionally contributes to heterogeneity.

The aim of my study was to define the heterogeneity of diabetes in childhood and adolescence. I approached this task at three points: classification of endotypes in childhood diabetes, determination of genetic contribution to the excess risk in first-degree relatives of patients with type 1 diabetes, and the influence of two environmental factors, maternal type 1 diabetes and viral infections in early life, on the development of autoimmunity.

First, I conducted detailed phenotyping in over 1000 new onset patients in the DiMelli study and subsequently developed a classifier based on residual β-cell function and routine measurements which allowed the separation of mild and severe forms of type 1 diabetes and additionally indicated that other forms of diabetes mellitus are present. The best predictors for this classifier were age at diagnosis, HbA1c and body mass index standard deviation score (BMI SDS). Among islet autoantibody negative patients, I could distinguish three groups corresponding to those who had features of monogenic diabetes, type 1 diabetes, and type 2 diabetes, respectively. Among the islet autoantibody positive patients, I was able to identify five endotypes. They differed markedly in residual β-cell function, severity at onset, concentrations of inflammatory cytokines, BMI SDS and gender and may be relevant for disease prognosis.

Second, I found in an analysis of the TEDDY (The Environmental Determinants of Diabetes in the Young) study that genetic susceptibility contributes to the elevated risk of autoimmunity and type 1 diabetes in first degree relatives of patients with type 1 diabetes as compared to the general population. In children with matched high-risk HLA genotypes, different frequencies of HLA DR4 subtypes, type 1 diabetes susceptibility alleles and a

higher genetic risk score were observed in children with as compared to children without a first-degree relative with type 1 diabetes. Additionally, I discovered yet unknown allelic enrichment at SNPs in *BTNL2*, located in the HLA complex II, which was associated with development of type 1 diabetes in children with a first-degree family history of type 1 diabetes. Stratification based on these genetic elements identified a high risk stratum in which children from the general population and children who were first-degree relatives of patients with type 1 diabetes had the same risk, and showed that there was increasing divergence in risk of autoimmunity between the two groups as the genetic risk diminished. I, therefore, concluded that other factors, including environment, would act heterogeneously on risk in a manner that partially depended upon the genetic susceptibility load of an individual.

Third, I investigated environmental influences. I investigated how the well-known protective effect of maternal type 1 diabetes on developing islet autoantibodies interacts with genetic risk using the TEDDY study. I found that protection is dominant over genetic risk and can reduce the risk of developing islet autoantibodies even in children in the highest genetic susceptibility stratum. To further investigate environmental heterogeneity, I searched for viral sequences in peripheral blood mononuclear cells from children who developed islet autoantibodies. Despite the a priori selection of cells expressing a type 1 interferon signature, viral sequences were rare and my findings could not substantiate a viral cause in these type 1 interferon positive cases.

Altogether, in my thesis, I have translated a concept of heterogeneity into distinct, clinically relevant phenotypes of type 1 diabetes and demonstrated how different genetic and environmental factors may contribute to pathogenetic heterogeneity.

ZUSAMMENFASSUNG

Diabetes mellitus im Kindes- und Jugendalter umfasst ein heterogenes Krankheitsbild. Typ 1 Diabetes ist die häufigste Form von Diabetes, gefolgt von Typ 2 Diabetes und einer Reihe anderer Diabetestypen. Die klinische Differentialdiagnose und Klassifikation der unterschiedlichen Diabetesformen ist insbesondere bei Inselautoantikörper-negativen Personen anspruchsvoll und gewissermaßen erschwert durch die enormen Kosten einer Typisierung der monogenetischen Diabeteserkrankungen. Auch bei der autoimmunen (Inselautoantikörper-positiven) Form des Typ 1 Diabetes werden heterogene Verlaufsformen beschrieben, die sich im Hinblick auf Krankheitsentstehung und Krankheitsprogression unterscheiden. Bisher gibt es jedoch keine systematische Analyse von Subtypen oder Endotypen des Typ 1 Diabetes. Wenig bekannt ist auch, wie hoch die Variabilität der Betazellrestfunktion bei Patienten mit neumanifestem Typ 1 Diabetes ist, und welche Faktoren diese beeinflussen. Die Variabilität der verbliebenen Betazellrestfunktion hat insofern große klinische Relevanz, da sie mit der Güte der Stoffwechselkontrolle und dem Auftreten von diabetischen Folgeerkrankungen korreliert ist. Die Klassifizierung in heterogene Krankheitsformen könnte somit auch therapeutische Konsequenzen haben. Das Risiko, Typ 1 Diabetes zu entwickeln wird durch genetische Prädisposition, Familienanamnese und Umweltfaktoren beeinflusst, wobei deren Wechselwirkung zusätzlich zur Heterogenität beiträgt.

Ziel meiner Arbeit war es, die Heterogenität des Diabetes im Kindes- und Jugendalter zu untersuchen. Dabei ging ich das Thema auf drei unterschiedlichen Ebenen an: der Klassifizierung von Endotypen des neumanifesten Diabetes mellitus im Kinder- und Jugendalter anhand der Variabilität der körpereigenen Insulinsekretion; der Bestimmung des genetischen Einflusses auf das erhöhte Risiko von Typ 1 Diabetes bei Patienten mit einem erstgradigen Verwandten mit Typ 1 Diabetes; und dem Einfluss von Umweltfaktoren, nämlich maternaler Typ 1 Diabetes sowie Virusinfektionen im frühkindlichen Alter, auf die Entwicklung von Autoimmunität.

Zuerst führte ich eine detaillierte Phänotypisierung von Bioproben von mehr als 1000 pädiatrischen Patienten mit neumanifestem Diabetes der DiMelli-Studie durch. Mittels der erhobenen Daten entwickelte ich einen Algorithmus, der die Unterteilung von Endotypen mit unterschiedlicher Ausprägung des Typ 1 Diabetes wie auch anderer Diabetesformen ermöglichte. Die Parameter Manifestationsalter, HbA1c, und Standard Deviation Score des Body-Mass-Index (BMI SDS) stellten sich als die besten Prädiktoren für die ß-Zellfunktion heraus. In Inselzellautoantikörper-negativen Patienten konnten mit Hilfe dieser einfachen Parameter in einem classification and regression tree (CART) die Diabetesformen definiert werden, die die klare Charakteristika von monogenetischem Diabetes, von Typ 1 , bzw. Typ

2 Diabetes zeigten. Innerhalb der Inselzellautoantikörper-positiven Patienten mit Typ 1 Diabetes identifizierte ich fünf Endotypen. Diese zeichneten sich durch signifikante Unterschiede in ß-Zellrestfunktion, inflammatorischer Zytokinkonzentration, BMI SDS, und Geschlecht aus, und können möglicherweise für die Prognose des Krankheitsverlaufs von Bedeutung sein.

Weiterhin fand ich bei einer Analyse der Daten der TEDDY (The Environmental Determinants of Diabetes in the Young)-Studie heraus, dass das erhöhte Risiko für Autoimmunität und Typ 1 Diabetes bei Kindern mit einem erstgradigen Verwandten mit Typ 1 Diabetes zum Teil durch genetische Prädisposition erklärt werden kann. Dabei konnte ich zeigen, dass Kinder mit gleichen HLA-Hochrisikogenen ausgeprägte Unterschiede in DR4-Subtypenverteilungen, Verteilungen von non-HLA-Risikogenen, und einem Typ 1 Diabetes-spezifischen genetischen Risikoscore aufweisen, wenn sie einen bzw. keinen Verwandten mit Typ 1 Diabetes haben. Zusätzlich konnte ich ein neues Risikogen, *BTNL2*, dass sich im HLA-Komplex II befindet, bei Verwandten ersten Grades häufiger ist, und mit dem Auftreten von Typ 1 Diabetes assoziiert ist, identifizieren. Durch eine Stratifizierung anhand dieser Marker konnte ich Kinder aus der Allgemeinbevölkerung identifizieren, deren Risiko für Typ 1 Diabetes ebenso hoch ist wie das von Kindern mit einem erstgradigen Verwandten mit Typ 1 Diabetes. Diese Überlagerung war in der Gruppe mit dem höchsten genetischen Risiko zu beobachten und driftete in den Gruppen mit niedrigeren Risiken zusehends auseinander. Daher kam ich zum Schluss, dass sich andere Faktoren, Umweltfaktoren eingeschlossen, heterogen und zum Teil von der individuellen, genetischen Prädisposition abhängig, auf das Risiko für Autoimmunität und Typ 1 Diabetes auswirken. Zu guter Letzt untersuchte ich den Einfluss von Umweltfaktoren auf Typ 1 Diabetes. Ich überprüfte, abermals in der TEDDY-Studie, wie sich der protektive Effekt von maternalem Typ 1 Diabetes in unterschiedlichen genetischen Risikogruppen auswirkt. Ich fand heraus, dass der bereits bekannte protektive Effekt des maternalen Typ 1 Diabetes auf allen Stufen eines genetischen Risikos wirkt und damit das Risiko auf Autoimmunität auch bei Kindern mit dem höchsten genetischen Risiko reduzieren kann. Um die Auswirkung von Umweltfaktoren auf die Heterogenität des Typ 1 Diabetes weiter zu untersuchen, wandte ich mich der Analyse viraler Infektionen zu. Ich untersuchte periphere mononukleare Blutzellen von Kindern, die vor dem Auftreten der Autoimmunität gewonnen wurden, auf die Existenz viraler Sequenzen. Obwohl diese Patienten zudem eine antivirale Typ 1 Interferon-Signatur aufwiesen, konnten sehr wenige virale Sequenzen nachgewiesen werden und damit kein Rückschluss auf einen kausalen Zusammenhang zur Entstehung der Autoimmunität bei diesen Kindern gezogen werden.

Alles in allem, habe ich meiner Doktorarbeit das Konzept der Heterogenität in einem Model klinisch relevanter Phänotypen des Typ 1 Diabetes dargestellt und veranschaulicht wie verschiedene genetische Merkmale und Umweltfaktoren zu dieser Heterogenität beitragen können.

1 INTRODUCTION

Type 1 diabetes (T1D) is a chronic autoimmune disease that is characterized by the loss of insulin producing β-cells. The symptomatic onset of T1D is preceded by a susceptibility stage marked by genetic and environmental predisposition, and pathogenetic stages defined by the appearance of autoantibodies (AABs) against β-cell-antigens (stage 1) and subsequently metabolic alterations such as dysglycemia and impaired glucose tolerance (stage 2). This stage, which often starts with the autoimmunity between age 6 months and 3 years, has a markedly variable duration until the clinical manifestation of T1D with classical symptoms of metabolic decompensation of T1D (stage 3). The marked variation of progression between the susceptibility stage and the stage of clinical manifestation suggests heterogeneous paths to disease. Heterogeneity may have clinical relevance as it implies pathogenetic endotypes with different prognosis regarding disease progression and the possibility to personalize treatment.

The risk of developing islet autoimmunity and subsequently T1D is influenced by genetic susceptibility, family history of the disease and environmental factors. The weight and interplay of these factors are likely to associate to the heterogeneous paths to islet autoimmunity. Therapeutic relevant endotypes at onset of clinical T1D are likely to represent such pathogenetic paths. They are also likely to be reflected by residual β-cell function, as has been shown for monogenic forms of childhood diabetes. c-peptide is a measure of residual β-cell function and, therefore, disease severity. It is a prognostic marker for future complications of the disease, and, as such, could be used to identify potentially therapeutic endotypes of T1D.

I approached the clarification of T1D heterogeneity from three sides. First, I created a clinically relevant classification of new onset patients of the DiMelli study with regard to residual β-cell function based on routine measurements. Application of this classifier will help to separate severe and mild phenotypes of the disease and also points to fractions of patients that do not have a T1D phenotype but another form of diabetes mellitus. Second, I addressed the genetic arm of the disease. As incidence is increased within affected families compared to the general population, I asked the question how much of this increase is caused by genetics. I confirmed known and found new genetic elements that explain part of the increased risk in the TEDDY (The Environmental Determinants of Diabetes in the Young) study. After stratification of subjects based on their genetics, I could identify genetic backgrounds where the risk of children from the general population and affected families completely converges and increasingly diverges with lower genetic susceptibility. Third, I focused on environmental factors. On different genetic backgrounds, I investigated, again in

TEDDY, an environmental factor that is known to influence T1D risk, namely, maternal T1D, which has a relative protection to children compared to children with a father or a sibling with T1D. Unexpectedly, the genetic background did not influence this protective environmental effect. I, therefore, proposed additional environmental factors that interact with genetic susceptibility to create pathogenetic heterogeneity. To this end, I had access to a unique dataset of high-risk children (BABYDIET) that had antiviral type 1 interferon signatures preceding seroconversion. In an exploratory experiment I searched for viruses in peripheral mononuclear cells of these children as a potential cause of the autoimmunity in some cases of T1D.

My thesis successfully demonstrates and describes heterogeneity in childhood diabetes and provides a model to establish endotype-based diagnosis and investigation of childhood diabetes.

2 STATE OF THE ART – BACKGROUND

2.1 Current classification of childhood diabetes

Diabetes mellitus in childhood is classified into type 1 diabetes (T1D), type 2 diabetes (T2D), and other forms including monogenetic and neonatal diabetes (1). T1D is the most prevalent form (~ 90% depending on the country) (2, 3), and itself, is classified into an autoimmune (type 1A) and an idiopathic (type 1B) type.

In comparison, T2D is ~ 10-fold lower in frequency in children (also depending on the country). Other types of diabetes, including 14 forms of maturity-onset diabetes of the young (MODY) and neonatal diabetes, together account for 1 - 4% of childhood diabetes. Other forms comprise uncommonly observed types of diabetes mellitus related to genetic defects in insulin action, diseases of the exocrine pancreas, endocrinopathies, induction via chemicals or drugs, infections or other genetic syndromes (1, 4). Correct classification of diabetes mellitus is important for therapy, prognosis, as well as patient education but clinical diagnosis becomes more complicated as features of specific types, e.g. overweight (T2D) or ketoacidosis (T1D), tend to blur. In addition, genetic typing for diagnosis of monogenetic forms of diabetes is costly and therefore genetic typing and counselling needs to be restricted to patients who have features associated with these forms.

2.2 ß-cell autoimmunity as characteristic feature of type 1 diabetes

The majority of patients with T1D are positive for one or more islet cell AABs. Previous studies have demonstrated that insulin AABs are most prevalent in infants and young children, while AABs to glutamic acid decarboxylase (GAD), insulinoma-associated autoantigen 2 (IA2), and zinc transporter 8 (ZnT8) AABs are more frequently seen in adolescence (5-7). Previous studies have also shown that the presence of one or more islet AABs is a characteristic feature of T1D in childhood (8-12). Furthermore, a few studies described that islet AABs can disappear during the asymptomatic pre-diabetic stage of T1D (13). It is therefore expected that a proportion of children with T1D may be negative for islet AABs at disease diagnosis.

Islet AABs are also used to diagnose early asymptomatic stages of T1D, and serve as primary outcome in natural history studies. The presence of two or more islet cell AABs characterizes stage 1 T1D;

the additional presence of impaired glucose tolerance characterizes stage 2 T1D. The clinical onset of T1D is termed stage 3 (Figure 1, Table 1) (14).

Table 1. Stages of T1D (adapted from Insel, et al. (14)).

Stage 1	Multiple islet AABs, normal glucose tolerance, pre-symptomatic
Stage 2	Multiple islet AABs, impaired glucose tolerance, pre-symptomatic
Stage 3	Islet autoimmunity, abnormal (diabetic) glucose tolerance, symptomatic

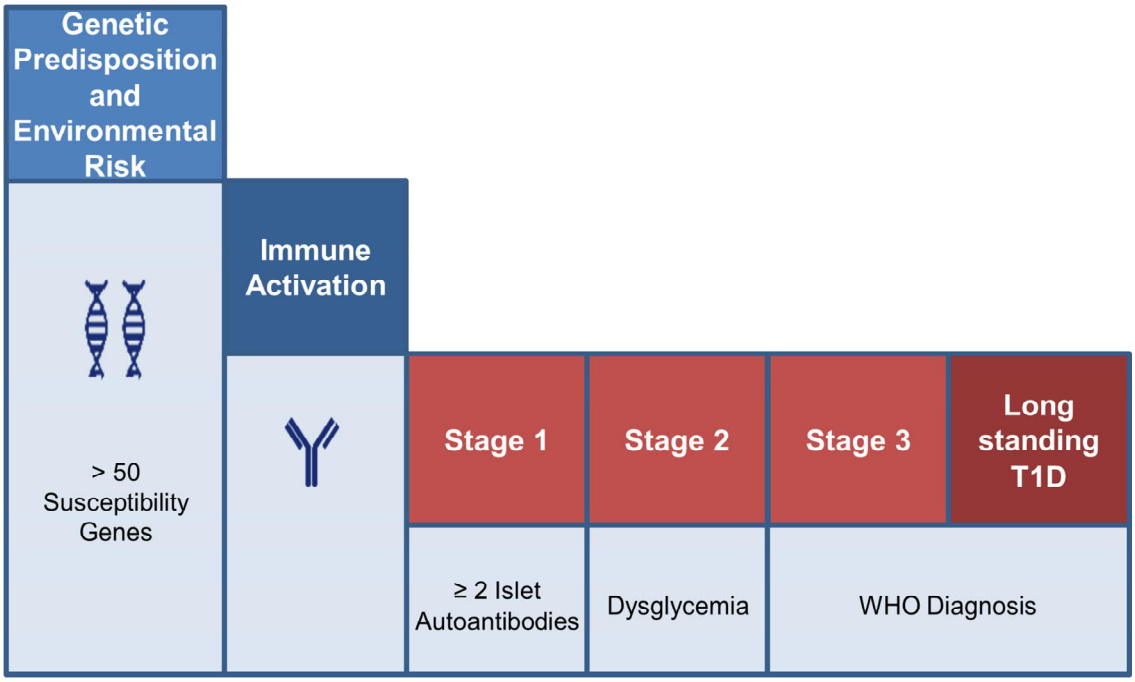


Figure 1. Stages of T1D.

Generally, islet AABs appear early in life. The first islet AABs are detected at an age of around 6 months. The peak incidence of islet AABs occurs around the age of two years (7). Incidence patterns of AABs are comparable between general population and in affected families (15, 16).

Besides AABs against the aforementioned 'classical' antigens there are other AABs (e.g. Tetraspanin-7 (TSPAN7), chromogranin A, Islet-Specific Glucose-6-Phosphatase-Related Protein (IGRP)) that are also found in T1D patients, but to a lesser extent (17-21). Furthermore, neoepitopes formed by posttranslational modifications like citrullination, transglutamination or peptide fusion have been reported (22).

2.3 Residual ß-cell function as marker of disease prognosis, and diabetes type

c-peptide or 'connecting peptide' is a peptide that connects the α and β chain of insulin in the proinsulin molecule. In the process of insulin synthesis, c-peptide and insulin molecules are secreted in equimolar concentrations. c-peptide is a marker of insulin secretion and also a surrogate for β-cell function itself (23-25). It is known, that patients with T1D have little or no residual c-peptide at diagnosis; in contrast patients with T2D or monogenetic forms of diabetes have relatively normal or even elevated levels of c-peptide (25, 26). With years of disease progression, functional β-cell mass can decline in any diabetes type, but most rapidly declines in patients with T1D (27). Among the contributing factors associated to the heterogeneous ß-cell function and decline are age, diabetes duration, and acute metabolic decompensation or infections (25, 28). Noteworthy, not all patients are diagnosed at the same point in the progression of the disease. (29, 30).

In islet AAB negative patients algorithms based on residual endogenous β-cell function have been suggested to aid the discrimination of autoimmune and non-autoimmune forms of diabetes (8, 10, 31). However, c-peptide levels overlap amongst patients with T1D and T2D at diagnosis, and other factors such as age and body mass index (BMI) are also relevant for distinguishing various subtypes of diabetes (30, 32, 33). Regarding maturity of diabetes in the young (MODY) patients, it is also reported that the persistence of c-peptide over many years is more characteristic and relevant more the measurement of c-peptide at the time of diagnosis (34).

The Diabetes Control and Complications Trial (DCCT) demonstrates lower HbA1c and reduced complications such as incidences of retinopathy, nephropathy and hypoglycemia in T1D patients with higher c-peptide levels (35). Furthermore, a reduction in microvascular complications with higher c-peptide can be found in T1D patients (30).

In addition, c-peptide as a measure of residual β-cell function is commonly used as outcome measure for immune intervention trials (25). Hence, c-peptide is a valuable measure at diabetes diagnosis with impact on acute therapy, long-term prognosis and the potential to classify various subtypes of diabetes in childhood and adolescence that represent clinically relevant phenotypes with mild and severe characteristics of T1D or even other forms of diabetes mellitus.

2.4 Diabetes as distinct phenotypes in the Young? Previous classification from the DiMelli study

Various factors drive heterogeneity of childhood diabetes. A previous analysis of the DiMelli study used the number of islet AABs to classify new onset diabetes patients into subgroups with multiple, one or no islet AAB and compared their clinical, metabolic and genetic parameters (8, 36). The most prominent differences between these groups were found in BMI percentiles, weight loss before diagnosis and fasting c-peptide, but also in age, ketonuria, insulin dependency and HLA (human leukocyte antigen) genotype. Patients without islet AABs had higher BMI, least weight loss, highest fasting c-peptide and were older with less ketonuria, less insulin dependency and less high risk HLA genotypes compared to people with islet AABs.

BMI percentiles were lower, the frequency of high risk HLA genotypes and ketonuria was higher in patients with multiple AABs as compared to patients with one AAB. Application of a classification strictly based on the number of islet AABs resulted in important characteristics of those groups but did not lead to a differentiation of the groups without overlap (8). This overlap of features demands for an improvement of the classifier via integration of other heterogeneity-associated features.

2.5 Monogenic forms of diabetes

Monogenic diabetes is an uncommon form of diabetes in children and adolescents (1-4%) and caused by one or more defects in a single gene. Heredity can be dominant, recessive or non-Mendelian. Spontaneous *de novo* mutations are also possible. Clinical scientists have described more than 40 genetic subtypes of monogenic diabetes, each one with an accompanying subphenotype. Subtypes include neonatal diabetes, autosomal dominant forms of mild, nonketotic diabetes (MODY), forms of diabetes in association with additional genetic syndromes, and diabetes caused by insulin resistance syndromes.

In neonatal diabetes, defined by disease onset before the age of 6 months, the most common causes are abnormalities in a region spanning the genes *PLAGL1* and *HYMAI* causing a transient form of diabetes, or mutations in *KCNJ11*, *ABCC8* or *INS*, causing a persistent form of neonatal diabetes. Further, rare forms, worth mentioning are caused by mutations in *EIF2AK3*, *GCK* and *FOXP3*. In total, more than 30 subtypes of neonatal diabetes have been described.

At least 14 different genes have been reported to cause MODY-like phenotypes, which differ in age of disease onset, disease severity and treatment (37). The majority of MODY cases are due to mutations in *GCK* (MODY2, 30-50% of MODY cases), *HNF1A* (MODY3, 30-65%

of MODY cases) and *HNF4A* (MODY1, 5-10% of MODY cases) (38). The genetic variant has direct implication to treatment approaches. MODY2 is characterized by a slightly elevated HbA1c, which is still lower than 7.5%, together with an intact insulin secretion and is treated solely with diet. MODY3 and MODY1 are very sensitive to treatment with sulfonylureas and usually do not require insulin therapy (39). Further genetic syndromes that occur in conjunction with diabetes are Wolfram syndrome, renal cysts and diabetes syndrome (MODY5), mitochondrial diabetes and diabetes secondary to monogenic diseases of the pancreas. Insulin resistance syndromes are caused by defects in primary insulin signaling, abnormalities in adipose tissue or other complex syndromes, including but not limited to Alström syndrome or Bardet-Biedl syndrome (39, 40).

2.6 Epidemiology of T1D

Worldwide, the incidence of T1D increases by a rate of ~3% per year (41). In Europe an annual increase of T1D incidence was observed over the last decades with highest rates in the youngest age group (42). This was confirmed in Germany, where the incidence of T1D in children younger than 15-years increased from 1999-2003 to 2004-2008 by 18% which corresponds to an annual increase of 3.4% (43). Predictive analysis anticipated a doubling of numbers in children <5 years of age and an increase by factor 1.7 in children <15 years of age (42). Such an increase in numbers cannot be attributed to genetics alone (24).

2.7 Family history

Children with a first-degree family history of T1D have a ~ 10 fold higher risk of developing islet AABs and T1D than children from the general population without a family history of the disease (15, 44, 45). The risk is higher for children who have a sibling or father with T1D or who have more than one relative with T1D than for children born to a mother with T1D (Table 2).

Having one first degree relative with T1D and a high risk HLA genotype elevates the risk to 10-20% while the risk for children with multiple relatives with T1D, independent of HLA genotype, is 20-25%. Children with identical twins with T1D or children with an affected sibling plus high HLA risk genotype both have a risk of 30-70%.

Table 2 Risk of developing T1D in dependence of family history related to the incidence in U.S.-born Caucasian children (adapted from (44))

first-degree family member with T1D	T1D risk (%)
none	0.4
one	5
mother	3
father	5
sibling	8
multiple	20 - 25

2.8 Genetics

T1D is a polygenic disease, i.e. multiple variants with different effect sizes contribute to the risk of developing the disease (41, 46-48). The largest effect, approximately 50 - 60% of the genetic contribution, is attributed to the HLA locus on the short arm of chromosome 6 (49).

2.8.1 HLA

The HLA locus encodes a variety of proteins involved in immune processes including cell surface proteins of class I (A, B, C) and class II (DR, DQ, DP) which are similar in structure and function as presenters of antigenic peptides to T cells (50, 51). This holds true for the presentation of self-antigens during thymic selection. Therefore, the association of the HLA locus and T1D as well as other autoimmune diseases is understandable.

It has been shown that in T1D the HLA-*DRB1-DQA1-DQB1* haplotypes are most informative (52, 53). With an odds ratio (OR) of 11.37 HLA *DRB1*0405-DQA1*0301-DQB1*0302* is the most susceptible haplotype in Caucasians with a frequency of 2.5% in patients with T1D and 0.2% in control subjects. Further risk conferring haplotypes are HLA *DRB1*0401-DQA1*0301-DQB1*0302* (OR 8.39, T1D patients: 28.1%, controls: 4.5%), HLA *DRB1*0301-DQA1*0501-DQB1*0201* (OR 3.64, T1D patients: 34.1%, controls:12.5%), HLA *DRB1*0402-DQA1*0301-DQB1*0302* (OR 3.63, T1D patients: 3.5%, controls: 1.0%), HLA *DRB1*0404-DQA1*0301-DQB1*0302* (OR 1.59, T1D patients: 5.0%, controls: 3.2%). The most protective haplotypes with OR ranging from 0.02 to 0.08 are HLA *DRB1*0701-DQA1*0201-DQB1*0303* (T1D patients: 0.1%, controls: 4.3%), HLA *DRB1*1401-DQA1*0101-DQB1*0503* (T1D patients: 0.0%, controls: 2.1%), HLA *DRB1*1501-DQA1*0102-DQB1*0602* (T1D patients: 0.4%, controls: 12.0%), HLA *DRB1*1104-DQA1*0501-DQB1*0301* (T1D patients: 0.2%, controls: 2.3%), HLA *DRB1*1303-DQA1*0501-DQB1*0301* (T1D patients: 0.1%, controls: 1.0%) (52).

Table 3. HLA susceptibility categorization and frequencies in the German BABYDIET study according to the HLA DR-DQ genotypes as previously described by Walter, et al. (54).

Risk category	HLA genotype	DRB1	DQB1	T1D patients (%) n = 538	Control subjects (%) n = 917
Very high risk	DR3/4-DQ08	03/04	02/0302	26.6	1.2
High risk	DR04-DQ08/ DR04-DQ08	04/04	0302/0302	9.3	0.7
Moderate DR4-DQ8	DR04-DQ08/ DRx-DQx	04/08	0302/0402	2.4	0.3
		04/13	0302/0604	4.1	0.8
		04/16	0302/0502	1.1	0.2
		04/01	0302/0501	8.0	2.4
Moderate DR3	DR03-DQ02/ DRx-DQx	03/16	02/0502	2.2	0.3
		03/13	02/0604	2.8	0.8
		03/03	02/02	5.8	1.9
Neutral	Neutral	rest		30.5	29.0
Protective	Protective	Z/11 or 12	z/0301	2.5	23.5
		Z/07	z/0303	0.6	6.8
		Z/13	z/0603	1.1	12.8
		Y/14	x/0503	0.4	5.5
		Y/Y	x/0602	2.6	26.4

*reference genotype; x = any DQB1 allele; Y = any DRB1 allele; Z-z = any DRB1-DQB1 haplotype except DRB1*04-DQB1*0302. #Subjects who have two protective haplotypes e.g. DR11-DQB1*0301 and DR7-DQB1*0303) are represented in each of the protective genotypes

HLA allele frequencies differ markedly by country (55, 56) (http://www.allelefrequencies.net). In Germany, the very high risk genotype HLA *DRB1*03-DQA1*0501-DQB1*0201/ DRB1*04-DQA1*0301-DQB1*0302* (*DR3/DR4-DQ8*) is present in 26.6% in T1D patients and 1.2% in controls (OR: 21.1) followed by the high risk HLA *DRB1*04-DQA1*0301-DQB1*0302/ DRB1*04-DQA1*0301-DQB1*0302* (*DR4-DQ8/DR4-DQ8*) with a frequency of 9.3% in T1D patients and 0.7% in control subjects (OR: 13.5). The moderate HLA DR4-DQ8/x and HLA DR3/x genotypes (detailed haplotype combinations in Table 3) are found in 15.6% and 10.8% of T1D patients and 3.7% and 3.0% in controls, respectively (OR range from 3.0 to 7.1, depending on specific haplotype). Protective HLA genotypes are present in 7.2% of T1D patients and 75.0% of controls (54).

The prevalence of the high risk HLA *DR3/4* is highest in patients with younger disease onset (10, 57). Protective HLA genotypes are rarely present (<1%) in patients with T1D but common in the general population (20%) (52).

2.8.2 Other susceptibility genes

Due to improved methodology and availability of single nucleotide polymorphism (SNP) genotyping, genome-wide association studies (GWAS) enabled the identification of SNPs associated with risk of developing multifactorial, complex disorders over time. A GWAS of the Wellcome Trust Case Control Consortium (WTCCC) found more than 40 susceptibility loci associated with T1D (58, 59), additionally to the HLA locus. In comparison to HLA the newly discovered loci show relative low contribution to T1D risk (Figure 2). In general, most of these genes are associated with immune response or expressed in human pancreatic islets. Indeed, they are grouped into three categories: I) immune function II) insulin expression III) β-cell-function (51). Most of the specific loci are in non-coding regions, but might express their effect through modulation of gene expression and/or identification of involved cell types by mapping to specific markers (60).

The loci (other than HLA) with the highest ORs in GWAS are located or associated to the following genes: The *INS* gene is located on chromosome 11p15 and encodes for preproinsulin. The precursor signal is cleaved first, and the resulting proinsulin is then processed to insulin and c-peptide in the same proportion. Within the gene region of INS polymorphisms in a variable number tandem repeat (VNTR) element confers risk to T1D with an OR of 2.2 (61). The protective allelic variant is associated with insulin mRNA in the thymus which supports the association of antigen level and negative selection of high-avidity autoreactive T cells (62, 63). Indeed, it was shown that low-avidity anti-proinsulin T cells are independent of *INS*, but that INS variants influence anti-proinsulin T cells with higher avidity. Thereby *INS* genotype is associated with an early checkpoint of autoimmunity regarding T cell frequency and repertoire (44, 64, 65).

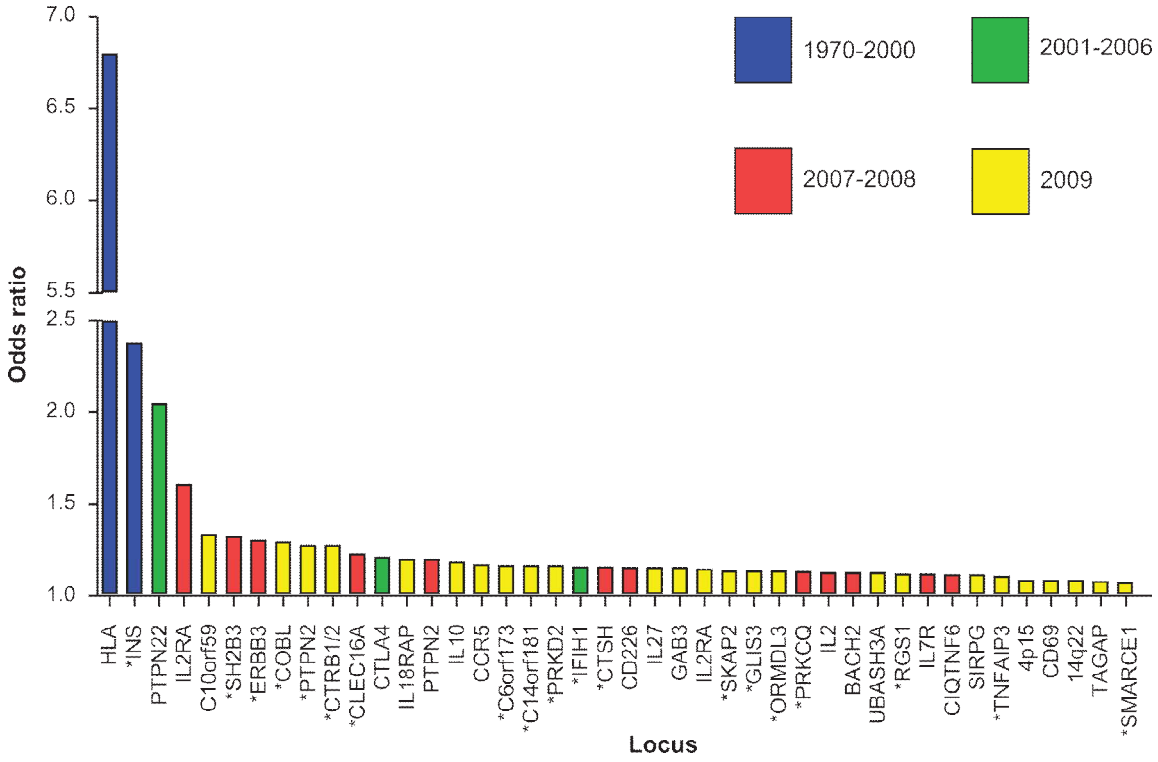


Figure 2. Susceptibility loci in association with T1D risk. Color-coding designates year of discovery of these candidate genes. The y-axis indicates the best estimate of the OR for risk alleles at each of the indicated loci on the basis of currently published data (58). For each genomic region where convincing association with T1D has been reported, the gene of interest or containing the most associated SNP is indicated on the x-axis. * marks genes expressed in the pancreas (http://www.t1dbase.org/) (66) (adapted from Pociot, et al. (67)).

The *PTPN22* gene on chromosome 1p13 encodes a Lymphoid-Specific Protein Tyrosine Phosphatase which acts as a negative regulator of T cell activation by dephosphorylation of the CBL protein. It was shown that the disease-associated variant results in down-regulation of both B cell receptor and T cell receptor signals upon stimulus which can lead to occurrence of polyreactive B cell potentially associated with autoimmunity (44). Besides T1D various other autoimmune diseases, including rheumatoid arthritis, systemic lupus erythematosus and Graves' disease are associated with mutations in PTPN22 (68).

On chromosome 10p15 *IL2RA* encodes the alpha chain of the interleukin 2 receptor. This receptor is involved in the control of the activity of regulatory T cells which in turn suppress autoreactive T cells. In T1D the susceptibility allele is associated with the concentration of the interleukin 2 receptor. Besides T1D, association to other autoimmune disease like multiple sclerosis, vitiligo and rheumatoid arthritis exist.

Summarized, GWAS resulted in numerous genotypic associations with T1D but the phenotypes underlying these variants remain largely unknown (69). The connection between

genetic risk and the impact of environmental factors as well as an enrichment of genetic susceptibility in affected families need to be elucidated.

2.8.3 Genetic Risk Scores

The odds ratio of each single locus on the risk of developing T1D is low in comparison to HLA, but their combination was utilized in order to improve risk prediction. Initially, each risk allele was counted in a small number of susceptibility loci in 1290 subjects of the BABYDIAB study (70). Subsequently, the number of loci was substantially increased, expanded for typing in the HLA region and incorporated with an assigned weight which was calculated using multivariate regression models in combination with Bayesian feature selection in the larger Type 1 Diabetes Genetics consortium (n = 5781) (71). Thus, genetic risk scores could further improve stratification in the population for the risk of developing islet autoimmunity or T1D in addition to HLA or family history. Others developed genetic risk scores to discriminate between T1D and T2D and predict insulin deficiency in young adults of the WTCCC (72). Bonifacio and colleagues combined the scores from Winkler, et al. (71) and Oram, et al. (72) to a total of 42 non HLA *DR-DQ* and 3 HLA *DR-DQ* SNPs (73). Application of these scores in high-risk HLA children of the TEDDY - study can identify children without a close relative with T1D with a 10% risk of developing multiple islet AABs by the age of 6 years as compared to a background risk of 0.3% (73). Furthermore, also in TEDDY, progression from islet autoimmunity to clinical T1D can also be stratified using genetic risk scores (74). Besides risk prediction, genetic variants are also associated with the decline of c-peptide and with effectivity of various treatments (14).

2.9 Environmental factors

It is generally assumed that environmental factors play an essential role in T1D and are responsible for the rise in disease incidence (75), and the divergence in risk of identical twins (44). A large list of environmental factors has been proposed in the past to impact on the risk to develop T1D including in utero environment (76-78), early nutrition (79-86), and infections (87-95). Two factors have been consistently shown to influence T1D risk across countries and across different study populations, namely maternal T1D as a protective condition and early life viral infections as risk condition.

2.9.1 Maternal type 1 diabetes

Maternal diabetes has repeatedly been shown to reduce the relative risk of T1D in affected families. Independent of HLA genotype, children with a mother with T1D have an almost 50 percent lower risk of developing T1D (3%) as compared to children with a father (5%) or

sibling (8%) with T1D (44, 96). Especially, the incidence of insulin AABs is reduced (76). Factors shown to contribute to the protection are moderately increased HbA1c during gestation (76) and the transfer of islet AABs from mother to child during pregnancy which was found to have a mild protective effect in children with HLA other than *DR3/4-DQ8* (97). Recent work postulates that increased (pro-) insulin production during pregnancy and the transplacental transfer of maternal islet AABs leads to an increase in immunological tolerance against T1D antigens in the fetus already *in utero*.

2.9.2 Viral infections

Another environmental factor that has consistently been shown to increase risk of T1D in families and the general population are respiratory and viral infections. The evidence comes from prospective birth cohort studies investigating the natural course of islet autoimmunity and T1D and from Bavarian claims data (98, 99). The German BABYDIET study reported that infections in the first 6 months of life increased the risk of islet autoimmunity (99). The association was strongest for viral respiratory infections; the more infections a child had in this period, the higher the risk was. A similar association of multiple viral respiratory infections and T1D risk was found in Bavarian claims data from nearly 300,000 children (98). The TEDDY study also confirmed the association and additionally reported an increased risk of viral infections preceding the onset of islet autoimmunity (100). Finally, the DIPP (Type 1 Diabetes Prediction and Prevention) study reported a close temporal association of enterovirus infections and islet autoimmunity onset (101). Also the TEDDY study confirmed that timing of viral infection may play a crucial role as respiratory infection episodes up to 9 months prior to seroconversion were associated with increase in risk of islet autoimmunity (100). Especially multiple infections in the period until age 6 months increases the risk of T1D with a hazard ratio of 1.42 (98). Furthermore, the DIPP study later identified enterovirus related with the progression from islet autoimmunity to T1D, indicating multiple working points of enteroviral infections (88). Also, a meta-analysis confirmed the association of enterovirus and islet autoimmunity (OR 3.7) and T1D (OR 9.8) (102).

Beside these epidemiological findings, direct evidence of enteroviruses infecting ß-cells comes from studies of human donor and autopsy material of patients with T1D, where enteroviruses were identified in ß-cell in the pancreas (103). Already in 1979, Coxsackie B4 was isolated from the pancreas of a child who died of diabetic ketoacidosis (104).

Moreover, weak and partly unconfirmed associations have been reported for Rotavirus (105), Cytomegalovirus (106), Parvovirus (107), Retrovirus (108, 109) and mumps virus (110).

There are different hypotheses how viral infections may act towards islet autoimmunity: one is by direct infection and killing of the ß-cell; another possible way is the activation of T cells through molecular mimicry, or the unspecific stimulation of immune responses through multiple infections and inflammation. Additionally, it cannot be excluded that virus infections are the result of a genetic predisposition leading to impaired immune responsiveness and clearing of the virus as suggested by work of Ashton and colleagues who found that children with insulin AABs have an incompetent immune response to the VP1 antigen of Coxsackie virus (111) (summarized in (112)).

In addition, several candidate genes for T1D are associated with antiviral responses in the immune system and in the β-cell itself and therefore support the theory of viral involvement in T1D development (112, 113). Consistent with this, the expression of an antiviral type 1 interferon (IFN) transcriptional signature is increased in peripheral blood mononuclear cells (PBMCs) of genetically predisposed infants before the development of islet AABs (114). The upregulation of IFN-inducible genes is transient and correlates with recent self-reported incidence of respiratory infections.

2.10 Precision medicine

Precision medicine is based upon the notion that within diseases, there are patient subgroups with well-defined and divergent pathogenetic causes of the disease. Therefore, the therapy that appropriately targets the mechanisms specific to the subgroup maximizes the treatment success for patients in this subgroup. The concept of precision medicine is furthermore based on the application and utilization of biomarkers. Biomarkers are defined as “any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease” (115). We currently lack biomarkers in relation to disease mechanism or therapy and subsequently clear definitions of patient subgroups.

3 OBJECTIVES

The overall objective of this thesis was to define heterogeneity of childhood diabetes. Thereby, I focused on three factors that overall contribute to this heterogeneity: **ß-cell function**, **genetic predisposition**, and **environmental exposure**. I had access to various human biomaterial and datasets and performed large scale ImmunoChip genotyping, unbiased high-throughput sequencing for viruses, electrochemoluminescence assays for inflammatory cytokines and serum Vitamin D3 and complex computational analyses to enrich these dataset to answer the following questions:

1) **How can childhood diabetes be classified into subtypes related to disease type and severity by using residual ß-cell reserve as the outcome?**

I applied a classification and regression tree algorithm to a cohort of 1194 patients with newly diagnosed diabetes aged < 20 years. I incorporated a range of metabolic, demographic, genetic, and immune marker variables in the analysis to define clinically relevant disease endotypes.

2) **How can genetic markers and family history of diabetes be used to define heterogeneous type 1 diabetes risk categories?**

Here, I analyzed data from over 3000 prospectively followed children in the TEDDY natural history study. In particular, I incorporated the analysis of HLA and HLA region loci, and over 100,000 genome wide SNPs to define risk categories and compare these between children with and without a first degree family history of type 1 diabetes.

3) **How do environmental factors interplay with genetic features with respect to the development of islet autoimmunity and T1D?**

Using data from the large TEDDY cohort, I examined the interaction between the protective factor maternal type 1 diabetes and genetic risk, and using clinical material from children who participated in the German BABYDIET study, I searched for viral genomes to assess if a type 1 interferon signature seen in some at risk children could be linked to specific infection.

4 POPULATIONS AND METHODS

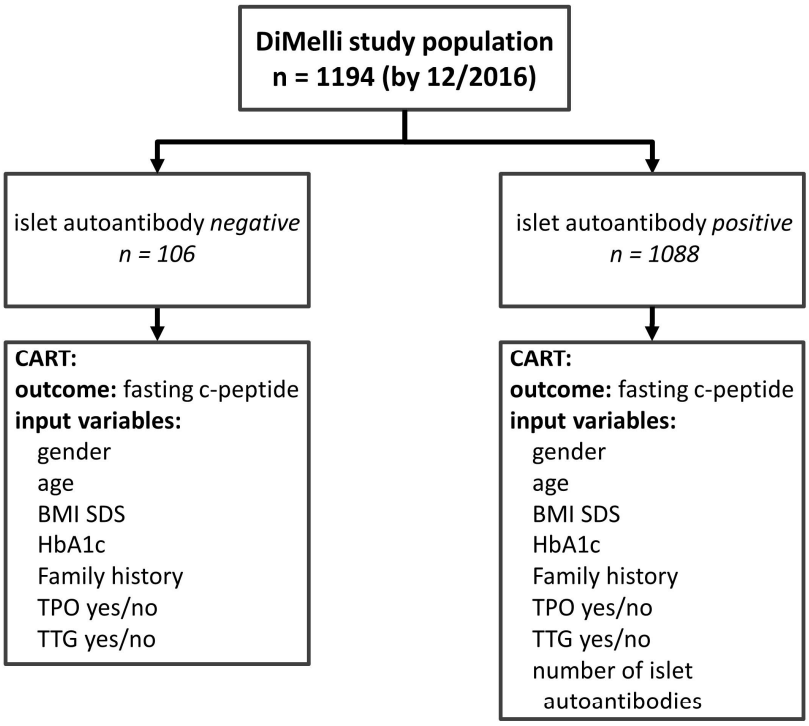
4.1 Populations

4.1.1 DiMelli

DiMelli is a population-based cohort study of diabetes incident cases in Bavaria, German. All patients were diagnosed with diabetes of any type according to to American Diabetes Association (ADA) criteria (116) before age 20 years, and enrolled within 6 months of diagnosis. DiMelli started in 2009 and ended in 2018. The study design has been described in detail previously (8, 36). At enrollment of each patient, a fasting blood sample was collected and a structured questionnaire was completed by the local physician at the hospital or the primary care center, and weight and height were assessed by trained staff (nurses or physicians) in accordance with the instructions given on the questionnaire. The blood sample was sent to the Institute of Diabetes Research, Helmholtz Zentrum München, by overnight express courier where all blood parameters were measured centrally. Each patient and/or parent provided written informed consent to participate in DiMelli. The DiMelli cohort study was approved by the medical ethics committee of Bavaria, Germany (Bayerische Landesaerztekammer, #08043) (8).

The overall objective of the DiMelli study was to understand disease heterogeneity and define subtypes of diabetes in childhood that are characterized by differences in pathogenesis, treatment, or prognosis. A total of 1378 patients were enrolled, 1194 of which had the full set of data required for analysis (Figure 3). Classification and regression Tree (CART) analysis was used to identify diabetes subtypes in patients stratified by the presence or absence of islet AABs.

Figure 3. Flowchart of the DiMelli study population with respect to CART analysis and validation.



4.1.2 TEDDY

The TEDDY study aims to decipher the impact of genetic predisposition and the role of environmental factors (including infectious agents, dietary factors, or stress) contributing to the initiation of islet autoimmunity, or progression to T1D mellitus (117). The TEDDY study screened 424788 newborns for T1D-associated HLA genotypes (Table 4) between 2004 and 2010, of which 8676 children with the respective high risk genotypes were enrolled and followed prospectively for islet AABs and T1D. The study was performed in six centers located in the USA, Finland, Germany, and Sweden. Detailed information on the study design, eligibility and methods has been published (46, 118, 119).

Table 4. HLA genotypes used to enroll newborns in the TEDDY study (adapted from Hagopian, et al. (117)).

a	*DR4- DQA1∗0301-DQB1∗0302@/ DR3- DQA1∗0501-DQB1∗0201*
b	*DR4- DQA1∗0301-DQB1∗0302 / DR4- DQA1∗0301-DQB1∗0302*
c	*DR4- DQA1∗0301-DQB1∗0302@/ DR8- DQA1∗0401-DQB1∗0402*
d	*DR3-DQA1∗0501-DQB1∗0201 / DR3-DQA1∗0501-DQB1∗0201*
e	*DR4- DQA1∗0301-DQB1∗0302@/ DR4- DQA1∗0301-DQB1∗0201*
f	*DR4- DQA1∗0301-DQB1∗0302@/ DR1#- DQA1∗0101-DQB1∗0501*
g	*DR4- DQA1∗0301-DQB1∗0302@/DR13-DQA1∗0102-DQB1∗0604*
h	*DR4- DQA1∗0301-DQB1∗0302 / DR4- DQA1∗0301-DQB1∗0304*
i	*DR4- DQA1∗0301-DQB1∗0302@/ DR9- DQA1∗0301-DQB1∗0303*
j	*DR3- DQA1∗0501-DQB1∗0201 / DR9- DQA1∗0301-DQB1∗0303*

Genotypes a–d are eligible from the general population and genotypes a–j for children with a first-degree relative with T1D.
@ Acceptable alleles in this haplotype include both *DQB1∗0302* and *∗0304*.
# In this *DQB1∗0501* haplotype, *DR10* must be excluded. Only *DR1* is eligible.

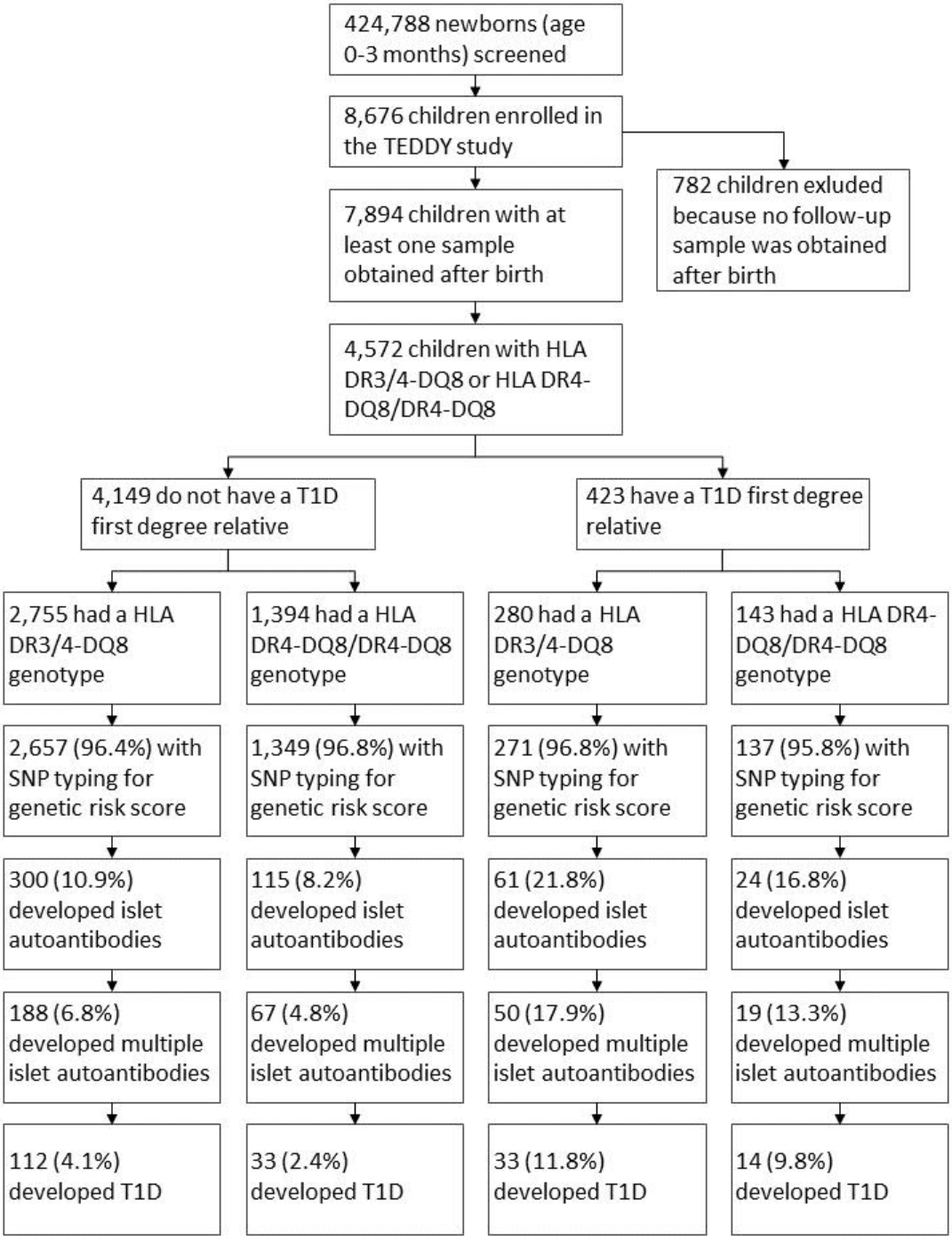
Written informed consents were obtained for all participants from a parent or primary caretaker, separately, for genetic screening and for participation in prospective follow-up. The study was approved by local Institutional Review Boards and is monitored by the External Evaluation Committee formed by the National Institutes of Health.

In children with high risk HLA genotypes who were enrolled in the prospective follow-up, blood samples were obtained quarterly until age 4 years and twice per year thereafter. Islet AABs (IAA, GADA, IA2A) were analyzed by radiobinding assays (RBA) as described previously (16, 120). The date of islet AABs seroconversion (time to first AAB) was defined as the date of drawing the first of two mandatory consecutive positive samples. The

presence of multiple islet AABs was defined as the presence of at least two islet AABs. Diabetes was diagnosed according ADA criteria.

For the analysis presented in this thesis, data as of 30 June 2017 was used from 4572 children with the *DR3-DQA1*0501-DQB1*0201/DR4-DQA1*030X-DQB1*0302* genotype (HLA *DR3/DR4-DQ8*), or the *DR4-DQA1*030X-DQB1*0302/DR4-DQA1*030X-DQB1*0302* genotype (*HLA DR4-DQ8/DR4-DQ8*), if at least one sample was obtained after birth as shown in Figure 4 (121).

Figure 4. Flowchart of TEDDY study population for the present analysis.

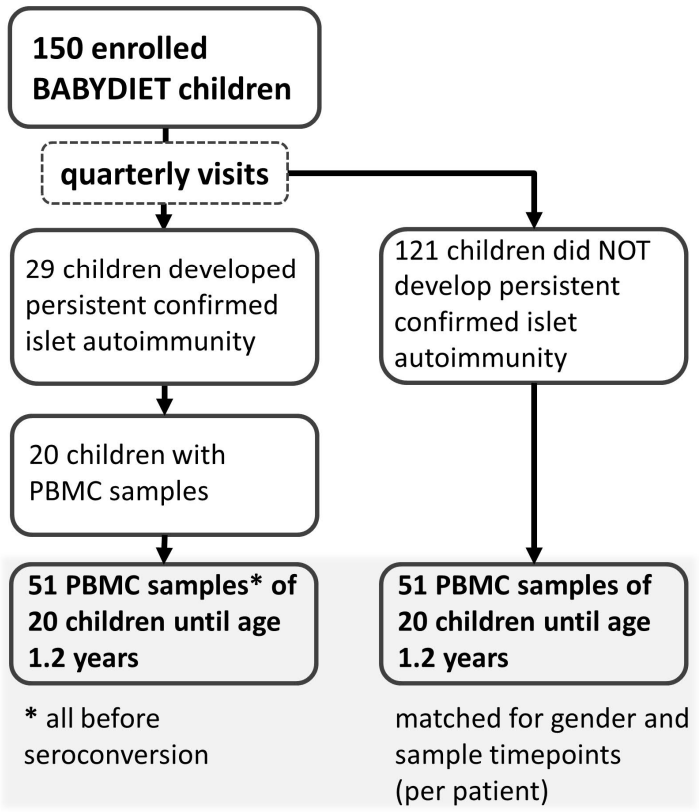


4.1.3 BABYDIET

The BABYDIET Study was a dietary intervention study in which children were randomized to gluten exposure at age 6 or 12 months (ClinicalTrials.gov NCT01115621) (79). In total, 150 children, selected for having both a first-degree relative with T1D and T1D risk HLA genotypes (*DRB1*03-DQA1*0501-DQB1*0201 / DRB1*04-DQA1*0301-DQB1*0302; DRB1*04-DQA1*0301-DQB1*0302 / DRB1*04-DQA1*0301-DQB1*0302 or DRB1*03-DQA1*0501-DQB1*0201 / DRB1*03-DQA1*0501-DQB1*0201*), participated in the BABYDIET Study (122). The BABYDIET study was approved by the ethical committee of Bavaria, Germany (Ludwig-Maximilian University No. 329/00, respectively), and all families gave written informed consent. Investigations were carried out in accordance with the principles of the Declaration of Helsinki, as revised in 2000 (7, 122). Children were recruited between 2000 and 2004 and followed from the age of 3 months with three-monthly blood samples until age 3 years, when the trial ended, and in the context of a natural follow-up study subsequently with 6-12 monthly blood samples.

Blood samples were utilized for the measurement of β-cell AABs (to insulin, GAD65, IA-2, and ZnT8) (122). Additionally, as part of an ancillary study, blood samples were used for the isolation of PBMCs by density gradient centrifugation using Lympholyte (Cedarlane) at each visit, and aliquots of 1×10^6 cells were stored at− 80 °C in TRIzol. The ancillary study aimed to identify environmental determinants of T1D prior to the initiation of islet autoimmunity. Subsequently, and as part of my thesis, RNA was isolated from individual aliquots using the phenol–chloroform method (114), and 102 of these samples from 40 children were subjected to VirCapSeq-VERT as shown in the flowchart in Figure 5 (123) with the purpose to search for potential viruses in blood associated with islet autoimmunity initiation. Furthermore, families were asked to document infections in their children daily up to the age of 3 years as previously described (122).

Figure 5. Flowchart of BABYDIET study participants regarding the sample selection for the search for viral infections using VirCapSeq-VERT.



4.2 Methods

4.2.1 Heterogeneity in β-cell reserve (DiMelli Study)

In order to define heterogeneity, I contributed to a comprehensive dataset by performing genetic typing using the ImmunoChip and measurement of inflammatory markers.

4.2.1.1 Questionnaire Data

The enrollment questionnaire included sex, the date of birth and the date of diagnosis of diabetes, first degree family history of T1D, T2D, or any other form of diabetes, and current anti-diabetic medications. Weight and height were assessed by trained staff (nurses or physicians) in accordance with the instructions given on the questionnaire.

4.2.1.2 Fasting c-peptide and HbA1c

Fasting c-peptide concentrations was measured in aprotinin-stabilized EDTA plasma samples using an automated immunoassay analyzer (AIA 360, Tosoh, San Francisco, CA, USA). The limit of detection (LOD) for c-peptide is 0.07 nmol/l. Values below the LOD were set to 0.07 nmol/l. HbA1c levels were measured in EDTA samples using a glycohemoglobin analyzer (TOSOH-723 G7; Tosoh).

4.2.1.3 Autoantibody measurement

AABs to islet autoantigens (insulin, glutamic acid decarboxylase (GAD65), insulinoma-associated protein 2, and zinc transporter 8) were measured by RBAs as previously described (11, 16, 120, 124). DiMelli samples negative in the RBAs were additionally measured by Luciferase Immunoprecipitation System (LIPS) assays for AABs to insulin, GAD65, IA-2, ZnT8, and TSPAN7, and by immunofluorescence for ICA (Gainesville, University of Florida, USA). The upper limit of normal for all RBAs (124, 125) and LIPS assays against ZnT8 and TSPAN7 corresponded to the 99th percentile (17), while upper limit of normal for LIPS assays against GAD65 and IA-2 correspond to the 99.9th percentile of control subjects (126). The number of positive islet AABs was categorized into none, 1, 2, 3, or 4 positive AABs, and as yes/no for each of the AABs IAA, GADA, IA2A, and ZnT8A. IAA were not included in the categorization of positive/negative if the sample was obtained >14 days after diagnosis, because antibodies to insulin can be induced by exogenous insulin therapy (8, 127). Transglutaminase and thyroid peroxidase (TPO) AABs were measured by RBA as described elsewhere (128).

4.2.1.4 Chemokine and Cytokine profiling

Serum concentrations of IFNγ, IL-10, IL-12p70, IL-1β, IL-2, IL-6, IL8 and TNFα were measured by electrochemiluminescence using a subset of the V-PLEX Proinflammatory Panel 1 Human Kit and MESO QuickPlex SQ120 instrument (Mesoscale Diagnostics) to assess the inflammatory status of the patients. This was done according to manufacturer's protocol in n = 666 DiMelli patients. Electrochemiluminescence signal intensity was converted to concentrations by back-fitting to a four-parameter logistic calibrator curve in DISCOVERY WORKBENCH 4.0 Software. For concentration below the LOD, values were imputed with the minimum concentration value for each cytokine. Samples were measured in duplicates and mean concentration is reported. The analytes were chosen with respect to the identification of clusters associated with hypo-inflammation, IL-1/IL-6 dominant inflammation, IL-8 associated inflammation, Th1/IFNγ associated inflammation, Th2 associated inflammation, regulation, and virus associate inflammation.

4.2.1.5 Triglycerides and Vitamin D3

Triglycerides as well as 25-OH-Vitamin D3 were analyzed using an enzymatic colorimetric test on a cobas 8000® analyzer with a c502 module (Roche Diagnostics, Basel, Switzerland) in cooperation with the Institute for Clinical Chemistry and Pathobiochemistry (MRI, TUM, for triglycerides), and the laboratory of Institute of Diabetes Research, Helmholtz Zentrum München (25-OH-Vitamin D) from 200µl of serum.

4.2.1.6 Genotyping

Genomic DNA from DiMelli participants was isolated from EDTA-blood and purified using the Genomic DNA Clean & Concentrator^{TM}-10 Kit (Zymo Research, Orange, CA, USA) according to manufacturer's protocol. Concentration was assessed using a NanoDrop 8000 (ThermoFisher) and quality by gel-electrophoresis.

HLA typing in the DiMelli study was performed by high-resolution sequencing-based typing of exons 2 and 3 of HLA-*DRB1* and HLA-*DQB1*, including heterozygous ambiguity resolution (Conexio Genomics, Fremantle, Western Australia). For additional genotyping of the DiMelli study, ImmunoChip analysis was performed following manufacturer's protocol, and by using the Infinium ImmunoArray-24 v2 BeadChip (Illumina), which covers >250,000 immune specific markers. Briefly, DNA was amplified for 24h, fragmented, precipitated with Isopropanol and resuspended for subsequent hybridization to the BeadChip. Following extension and staining the BeadChip was scanned by the Illumina iScan system. Clusters of SNPs were generated using GenomeStudio including Genotyping Module v1.0 (Illumina)

and exported as Cluster File. Quality control was conducted by application of following thresholds using PLINK software: missing genotype rate 5%, minor-allele frequency 5% missing individual rate 5% (129).

For the description of each individual's genetic risk besides HLA *DR-DQ* genotype the genetic risk score was calculated from 40 SNPs similar to the previously described merged genetic score (73) except the value of the HLA *DR-DQ* genotype (3.98 for *DR3/DR4-DQ8* or 3.15 for *DR4-DQ8/DR4-DQ8* in the TEDDY cohort) was not included in the score (Table 5). In the DiMelli dataset 5 SNPs were missing: rs1264813 (HLA A), rs9388489 (C6orf173), rs12722495 (IL2RA), rs17574546 (non-coding region), rs2290400 (non-coding region). The corresponding population median weight was added to reach comparability between datasets.

Table 5. Weights for SNPs used to calculate the genetic risk score from Bonifacio, et al. (73).

SNP	Gene, Allele, Genotype	Merged Score Weight
HLA class II		
rs17426593	*HLA DR4-DQ8/ DR4-DQ8*	3.15
rs2187668		
rs7454108	*HLA DR3/ DR4-DQ8*	3.98
Other SNPs		
rs1264813	*HLA A 24*	0.43
rs2395029	*HLA B 5701*	0.92
rs2476601	*PTPN22*	0.76
rs2816316	*RGS1*	0.16
rs3024505	*IL10*	0.22
rs1990760	*IFIH1*	0.16
rs3087243	*CTLA4*	0.16
rs10517086	*C4orf52*	0.19
rs2069763	*IL2*	0.11
rs6897932	*IL7RA*	0.19
rs3757247	*BACH2*	0.19
rs9388489	*C6orf173*	0.14
rs6920220	*TNFAIP3*	0.15
rs1738074	*TAGAP*	0.05
rs7804356	*SCAP2*	0.15
rs4948088	*COBL*	0.17
rs7020673	*GLIS3*	0.23
rs12722495	*IL2RA*	0.47
rs947474	*PRKCQ*	0.15
rs10509540	*RNLS/C10orf59*	0.25
rs1004446	*INS*	0.65
rs4763879	*CD69*	0.06
rs2292239	*ERBB3*	0.36
rs3184504	*SH2B3*	0.24
rs1465788	*ZFP36L1*	0.13

SNP	Gene, Allele, Genotype	Merged Score Weight
rs17574546	*RASGRP1*	0.13
rs3825932	*CTSH*	0.15
rs12708716	*CLEC16A*	0.15
rs4788084	*IL27*	0.20
rs7202877	*CTRB2*	0.19
rs2290400	*ORMDL3*	0.25
rs7221109	*CCR7*	0.15
rs45450798	*PTPN2*	0.09
rs763361	*CD226*	0.12
rs425105	*PRKD2*	0.21
rs2281808	*SIRPG*	0.07
rs3788013	*UBASH3a*	0.16
rs5753037	*RPS3AP51*	0.15
rs229541	*IL2B*	0.18
rs5979785	*TLR8*	0.09
rs2664170	*GAB3*	0.14
rs917997	*IL18RAP*	0.14*

* Not included in the genetic score calculations for the TEDDY cohort

4.2.1.7 Statistics and Data Analysis

All statistical analyses were performed using R version 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria; http://cran.r-project.org) and the packages childsds, dendextend, GWASTools, mclust, party, purr, SNPRelate and snpStats.

In general, group comparisons were performed using the Kruskal–Wallis test for continuous variables and the $\chi 2$ test or, if applicable, Fisher's exact test with false-discovery-rate correction for multiple testing for categorical variables. Associations between continuous variables were assessed using Pearson's correlation coefficient. For the DiMelli dataset P-values of <0.01 were considered statistically significant.

Within DiMelli patients, I applied a CART analysis (130) to investigate potential subphenotypes of pre-stratified groups of patients negative and positive for islet AABs. Fasting c-peptide as measure of endogenous insulin secretion was used as outcome variable. The maximum depth of the tree was set to 3. The following variables were considered as potential predicting factors for residual c-peptide concentrations: sex, age at diagnosis, BMI (SDS), HbA1c, first degree family history of T1D (yes/no) or any other form of diabetes (yes/no), number of positive islet AABs (1 to 4), TPO AABs (yes/no), and transglutaminase AABs (yes/no). Sex- and age-specific BMI SDS were calculated based on national reference values (131), and defined overweight as ±1SD, obesity as >+2SD and thinness as <-2SD (according to world health organization guideline). As a sensitivity analysis, we explored whether the CARTs changed when the genetic risk score or its

individual risk SNPs or the HLA genotype were added as input variables, or positivity of transglutaminase AAB and TPO AAB and number of islet AABs were left out.

Chemokine and cytokine concentrations were first correlated and then categorized into centiles (0-90%, 90-92.5%, 92.5%-95%, 95%-97.5%, 97.5%-100%) and scaled from 0 to 1 separately for each cytokine. These profiles were subjected to hierarchical clustering using Euclidean distance and ward.D2 clustering.

4.2.2 Heterogeneity in genetic predisposition (TEDDY study)

In order to assess the influence of genetic predisposition on the risk of developing islet autoimmunity and T1D, I used a dataset of the TEDDY study which combined complete genetic information and information of islet autoimmunity and progression to T1D. Inclusion criteria of the TEDDY study already included HLA genotyping which confirmed by the central HLA Laboratory at Roche Molecular Systems (Oakland, CA) for enrolled subjects (16). Additionally, data on SNP genotyping, using the Illumina ImmunoChip as described elsewhere (16), was available. Similar to the calculation described for the DiMelli study, a genetic risk score was applied. In the TEDDY dataset SNP rs2290400 was missing and the corresponding population median weight was added to reach comparability.

4.2.2.1 Functional siRNA-mediated gene silencing in monocyte-derived dendritic cells

To gain information about the newly discovered risk gene *BTNL2*, experiments on *in vitro* immune responses were performed in cooperation with my colleague Jan Knoop. Therefore, 6 healthy donors were recruited from the Munich Diabetes Bioresource with informed consent and ethical approved protocols (Ethical approval number 5049/11).

Briefly, activation of isolated $CD4^+$ $CD25^-$ T cells by allogeneic monocyte-derived dendritic cells transfected with non-targeting siRNA or *BTNL2* targeting siRNA in mixed lymphocyte cultures was measured. CD71 expression, previously reported as a marker of allo-reactive T cell activation, was used to measure the activation of the $CD4^+$ T cells. Dendritic cells were transfected as previously described (132).

4.2.2.2 qPCR experiments

For the validation of *BTNL2* knockdown experiments 200ng of RNA from siRNA transfected cells was subjected to cDNA synthesis using a mix of oligo-dT-primer and random primer with the iScript cDNA synthesis Kit (Bio-Rad) followed by pre-amplification with gene specific primers (Table 6) and the SsoAdvanced PreAmp Supermix (Bio-Rad). qbase+-software (Biogazelle) was used for analysis of qPCR experiments. *BTNL2* expression was normalized to reference genes *TELO2* and *TRMT61A*.

Table 6. Primers for qPCR experiments.

Assay	Sequence 5‘ – 3‘
BTNL2	
>BTNL2_fwd	GACAATGAAGCAGTCAGAAGACTT
>BTNL2_rev	TCTCAGTCACCTCCACTCCA
TRMT61A	
>TRMT61A_fwd_1	CTTCCTGGACATCCCATCAC
>TRMT61A_rev_2	TAGACCTGTGGCAGCACCT
TELO2	
>TELO2_fwd_1	GGCCGACTATCTGACCTCAC
>TELO2_rev_1	CCAGGCCTAGACAGCTCCT

4.2.2.3 Statistics and Data Analysis

All statistical analyses were performed using R version 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria; http://cran.r-project.org) and the packages GWASTools, haplo.stats, qqman and survminer. In the analysis of the TEDDY dataset groups were compared using the Kruskal–Wallis test for continuous variables and the $\chi 2$ test or, if applicable, Fisher’s exact test with Bonferroni’s correction for multiple testing for categorical variables. The risks of developing one or more islet AABs, multiple AABs, and diabetes in children with a first-degree relative with T1D and children from the general population were assessed using Kaplan–Meier analysis, and groups were compared using the log-rank test. Hazard ratios (HRs) were computed using Cox’s proportional hazards model. P-values <0.05, were considered statistically significant.

4.2.3 Heterogeneity in environmental exposure (TEDDY and BABYDIET studies)

To address the environmental contribution of T1D development, I used two approaches. First, in the TEDDY dataset described above, I investigated the characteristic of a protective environmental effect, namely maternal T1D, in affected families with respect to the genetic susceptibility. Second, in the BABYDIET dataset, I searched for a link of viral infection in early life and the development of islet autoimmunity and T1D. Based on the genetic stratification, I enquired if the influence of maternal T1D as a protective environmental factor in subjects with a first-degree relative with T1D is influenced by genetic stratification. In the TEDDY study the information on affected family members was available through questionnaires and categorized into mother, father or sibling with T1D. Combinations were allowed. In order to be able to associate viral infections in early life to the development of islet autoimmunity and subsequently T1D, I collected data of consecutive islet AAB measurements to define the outcomes. Viral infections were investigated by a high-throughput sequencing based method, VirCapSeq-VERT. To apply this technique, which was newly developed by Briese et al. at the Columbia University, New York (133) , I spent two months in the laboratory of Dr. Thomas Briese and performed the sequencing as well as the analysis under the supervision of Dr. Thomas Briese. The visit was supported by a stipend of the HELENA graduate school.

AABs to islet autoantigens (IA, GAD, IA2, and ZnT8) were measured by standardized RBA as previously described (11) for participants of BABYDIET. Similar to the TEDDY study, the date of islet AABs seroconversion was defined as the date of drawing the first positive samples. Diabetes was diagnosed according to clinical diagnosis.

4.2.3.1 VirCapSeq-VERT

The virome capture sequencing platform for vertebrate viruses (VirCapSeq-VERT) combines unbiased virus detection and increased sensitivity through positive selection of viral sequences followed by high-throughput sequencing. Viral Sequence selection is guaranteed via a biotinylated library of 1,993,200 oligonucleotide probes that cover the genomes of all 2017 virus taxa known to infect vertebrates, including humans, as of December 2014 (SeqCap EZ Choice; Roche/NimbleGen, Basel, Switzerland). Design of probes additionally allows the detection of novel virus genera (133). This method was applied to PBMC derived RNA from 102 samples of a total of 40 children from the BABYDIET study, whereof 50% progressed to islet autoimmunity.

4.2.3.1.1 Virome Capture Sequencing

Sequencing library generation follow mainly the standard KAPA protocol including the viral sequence capture library. Briefly, RNA was reverse transcribed using SuperScript III (ThermoFisher) with random hexamers. The cDNA was RNase H treated prior to second-strand synthesis with Klenow fragment (New England Biolabs). The resulting double-stranded cDNA was sheared to an average fragment size of 200 bp using the manufacturer's standard settings (Covaris E210 focused ultrasonicator). The sheared products were purified (AxyPrep), and libraries were constructed using KAPA library preparation kits (KAPA) with Roche/NimbleGen adapter kits. The quality and quantity of libraries were checked by Bioanalyzer (Agilent). The libraries were then mixed with a SeqCap HE universal oligonucleotide, SeqCap HE index oligonucleotides, and COT DNA (Roche) and vacuum dried at 60°C for approximately 40 min. Dried samples were mixed with 2x hybridization buffer and hybridization component A (Roche) prior to denaturation at 95°C for 10 min. The VirCap probe library (4.5 µl) was added and hybridized at 47°C for 12 h in a PCR thermocycler. SeqCap Pure capture beads (Roche) were washed twice, mixed with the hybridization mix, and kept at 47°C for 45 min with vortexing for 10 s every 10 to 15 min. The streptavidin capture beads complexed with biotinylated VirCapSeq-VERT probes were trapped (DynaMag-2 magnet; ThermoFisher) and washed once at 47°C and then twice more at room temperature with wash buffers (Roche). Finally, beads were suspended in 50 µl water and directly subjected to posthybridization PCR (SeqCap EZ accessory kit V2; Roche). The PCR products were purified (Agencourt Ampure DNA purification beads; Beckman Coulter) and quantitated by Bioanalyzer for sequencing on an Illumina HiSeq 4000 (133).

4.2.3.1.2 Bioinformatic Analysis

After sequencing, samples were demultiplexed using Illumina software, and FastQ files were generated. Demultiplexed and Q30-filtered FastQ files were preprocessed using PRINSEQ (v 0.20.2) (134) software, and quality filtered reads were aligned against a human host reference databases to remove the host background. The resulting reads were de novo assembled using MIRA (v 4.0) (135). Viral read numbers were generated from counting the number of reads mapping to contig sequences and unassembled singletons. Contigs and unique singletons were subjected to homology search using MegaBlast against the GenBank nucleotide database; sequences that showed poor or no homology at the nucleotide level were screened by BLASTX against the viral GenBank protein database. Viral sequences from BLASTX analysis were subjected to another round of BLASTX homology search against the entire GenBank protein database to correct for biased e-values and taxonomic misassignments. Based on the contigs and singletons for the various viruses or viral strains identified, the best matching GenBank sequences identified by the BLAST searches were downloaded and used as reference for mapping the whole data set to recover partial or complete genome sequences (Bowtie2 mapper 2.0.6 (http://bowtie-bio.sourceforge.net)). SAMtools (v 0.1.19) (136) were used to generate the consensus genomes and coverage statistics. Geneious (v10; www.geneious.com) or Tablet (137) were used to visualize and evaluate read mappings (133).

5 RESULTS

5.1 Heterogeneity in β-cell reserve (DiMelli study)

Diabetes in childhood and adolescence is most frequently classified as type 1a, which is characterized by the presence of islet AABs (4). There are additional monogenic forms of diabetes and T2D diagnosed in this age period. Algorithms based on residual endogenous β-cell function (8, 10, 31) or the use of genetic risk scores (72) have been suggested to aid the discrimination of true non-autoimmune forms of diabetes in islet AAB negative patients. Without extensive genetic analysis, however, there is diagnostic uncertainty for monogenic diabetes, and T2D is not always distinguishable from other forms of diabetes in young age (9).

Similar to what has been described for diabetes in adults (138), the notion of heterogeneity and endotypes has been gathering evidence also for type 1a diabetes and diabetes in childhood. There are differences in HLA genetic load between younger and older diabetes cases (139), as well as clinically relevant parameters such as functional β-cell reserve at diagnosis (140). More recently, two age-related endotypes of T1D have been proposed based on features of the T and B cell responses to autoantigens and the predominant cell type found infiltrating the islets of patients (141). There is also evidence for age-related differences in response to immune-therapy (142), further suggesting that there are pathogenetically and therapy responsive distinct endotypes of T1D.

The objective of this analysis was to define clinically relevant endotypes that relate to c-peptide at onset of diabetes that is diagnosed before 20 years of age. I developed these endotypes using the CART model in ~1200 patients with new onset diabetes under the age of 20 years identified in the DiMelli cohort study performed in Bavaria, Germany (8, 36). Separate searches were performed in islet AAB negative and islet AAB positive patients. Factors that are readily obtainable at diagnosis such as age, HbA1c, and BMI SDS were used as predictors in the search model so that endotype classification could be easily introduced into practice.

5.1.1 Comparison of islet Autoantibody positive and islet Autoantibody negative patients

The dataset used to build the CART contained 1194 patients (528 girls, 44.2%) with new onset diabetes who were enrolled at a median diabetes duration of 9 days (range 0-180 days; interquartile range [IQR], 6-13), and all tested for islet AABs (insulin, GAD65, IA2, and ZnT8R/ZnT8W). Of these, 106 patients (47 girls, 44.3%) were islet AAB negative, and 1088 patients (481 girls, 44.2%) were islet AAB positive (116 one, 253 two, 359 three, 360 four islet AABs).

Comparison of islet AAB positive patients and islet AAB negative patients showed that islet AAB negative patients had higher fasting c-peptide concentrations (0.28 [0.14; 0.52] vs. 0.13 [0.07; 0.23] nmol/l, p<0.001), were older at diabetes onset (13.50 [10.26; 15.53] vs. 10.21 [6.93; 13.22] years; p<0.001), had higher BMI SDS (0.37 [-0.64; 1.46] vs. -0.58 [-1.38;0.22]; p<0.001) and blood triglyceride concentrations (111.00 [74.50;153.00] vs 78.00 [60.00;104.00] mg/dl; p<0.001), and lower HbA1c (10.00 [8.12; 12.05] vs. 10.80 [9.30;12.50] %; p=0.004), had more family history of diabetes that was not T1D (n=20 (18.8%) vs n=25 (2.3%); p<0.001), were less frequently treated with insulin (n=86 (81.1%) vs n=1065 (97.9%); p<0.001), and had lower frequency of ketonuria (n=60 (67.4%) vs n=903 (88.9%); p<0.001). The islet AAB negative patients also had less genetic features associated with T1D, with a predominance of protective HLA genotypes (n=35 (37.6%) vs n=62 (6.6%); p<0.001), and lower non-HLA genetic risk scores (10.20 [9.50; 11.08] vs 10.65 [9.95; 11.25]; p=0.009) (Table 7).

5.1.2 CART analysis of islet autoantibody negative patients

Islet AAB negative patients had a wide range of fasting c-peptide values from undetectable (<0.07 nmol/l) to >3.18 nmol/l. Since c-peptide has prognostic and diagnostic value in T1D (29, 30), I searched for heterogeneity within the islet AAB negative patients using CART analysis to predict fasting c-peptide concentration. As possible predictor variables I chose sex, age at diagnosis, HbA1c, BMI (SDS), and first degree family history of type 1 or any other form of diabetes as these are readily obtainable soon after diagnosis.

Application of CART algorithm to 106 islet AAB negative patients indicated that BMI SDS was the best predictor of fasting c-peptide concentration and prediction was further stratified by HbA1c in patients who were not in the high BMI SDS category (Figure 6A). Among the three groups (A1, A2, A3), fasting c-peptide was lowest in the 65 patients with BMI SDS of <1.6 and HbA1c >7.8 (A2;0.17 (0.10; 0.26] nmol/l), and was highest in the 24 patients with a BMI SDS >1.6 (A3: 0.86 (0.55; 1.35] nmol/l), and in the 17 patients with a BMI SDS of <1.6 and HbA1c <7.8 (A1: 0.50 (0.43; 0.66] nmol/l). The characteristics of the three groups differed substantially (Table 8).

5.1.3 CART analysis of islet autoantibody positive patients

The ability of the CART analysis to classify islet AAB negative patients led me to apply a similar analysis in the islet AAB positive patients (Figure 6B). In 1088 islet AAB positive patients, the CART analysis identified 7 subgroups with median c-peptide concentrations ranging from 0.10 nmol/l (group B3/B4) to 0.38 nmol/l (group B5; Figure 6B, Table 9). Age at diagnosis was the best predictor for residual c-peptide concentrations at disease onset in these patients. Prediction was improved by further partitions with HbA1c and BMI SDS.

5.1.4 Characteristics of diabetes subgroups

5.1.4.1 c-peptide

Groups with low median fasting c-peptide were identified in both, the islet AAB negative patients (A2) and the islet AAB positive patients (B1, B3, B4, B6). The groups with the highest median fasting c-peptide were the groups A1 and A3 in the islet AAB negatives, but also the majority of islet AAB positive patients in groups B2, B5 and B7 had a fasting c-peptide concentration >0.2 nmol/l (Figure 6). High c-peptide groups A1, B2, and B5 were characterized by HbA1c values that were modestly elevated (highest HbA1c: 8.3%), and groups A3 and B7 by a high BMI SDS (1.71 [1.28; 2.29] vs -0.62 [-1.40; 0.14] in remainder; p<0.001).

5.1.4.2 Sex, family history, autoantibodies

A female bias was observed in the islet AAB negative A1 group (64.7% females, p=0.11 vs rest). Male bias was observed in islet AAB positive groups B5 (64.3% males, p=0.21), B6 (60.6% males, p=0.018), and B7 (75.0% males, p=0.007) (Table 8 and Table 9). Group A1 also differed by an excess prevalence of cases with a family history of diabetes, and in particular diabetes that was neither classified as type 1 nor type 2 diabetes (11.8%, p=0.030). Group A3 had a large excess of T2D family history (41.7%, p<0.001) and group B5 was slightly increased in T2D family history (12.5%, p<0.001). A first-degree family history of T1D ranged from 4.2% in group A3 to 17.6% in group A1.

The additional presence of thyroid autoimmunity (TPO autoantibodies) was seen in a minority of cases in each of the patient groups except A1 and B2; the presence was highest in groups B6 (13.9%) and B7 12.5(%). The prevalence of celiac disease-associated transglutaminase autoantibodies was highest in groups B1 12.8(%), B3 (10.4%) and A2 (10.8%). No differences in the number of islet AABs was observed between the islet AAB positive patient groups. Group B1, representing the younger cases had the highest prevalence of IAA (88.1%, p=0.009) (Table 9).

5.1.4.3 Treatment

The islet AAB negative patient group A1 was noticeable by the paucity of insulin treatment given at the onset of disease (Table 8). Twelve (70.6%, p<0.001) of 17 patients in group A1 were started with diet or no treatment. A substantial proportion (29.2%) of patients in group A3 were given metformin and no insulin treatment. Among the islet AAB positive groups, there was reduced prevalence of insulin treatment given to patients at clinical onset in groups B1 (91.5%, p=0.015), B2 (86.4%, p=0.010) and B5 (89.3%, p<0.001) (Table 9). In all other groups, including group A2, the frequency of patients receiving insulin treatment was >95%. Insulin dose varied between groups (Figure 7A).

5.1.4.4 Vitamin D3 status

Vitamin D3 is often low in patients with T1D (143, 144) or T2D (145). Consistent with this, all groups except group A1 had low median vitamin D3 concentrations (Table 8 and Table 9; Figure 7B). Group A3 patients had the lowest vitamin D3 concentrations. The frequency of patients with vitamin D3 below 30 ng/ml ranged from 82.6% in group A3 to 31.2% in group A1.

5.1.4.5 Inflammatory markers

The cytokines IFNγ, IL-10, IL-12p70, IL-1β, IL-2, IL-6, IL-8, TNFα were measured in 666 patients. Cytokine values in patients were correlated (Figure 7C), suggesting that there were patients in whom multiple cytokines were systematically raised. A heatmap of the overall data confirmed this and identified patients in whom there was substantial increase in the majority of the cytokines measured, including IL-1β, IL-6, and/or TNFα (Figure 7D). An additional small group of patients had raised IFNγ without an accompanying increase in IL-1β, IL-6 or TNFα, and some patients had raised IL-2 only. Patients with any of these signatures were relatively abundant in the young B1 islet AAB positive group (39.1% vs. 14.3%; range 0.0% – 23.4% in the remainder, p=0.009). Occasional patients in other groups had very high concentrations of individual or multiple cytokines. Of note, all cytokine concentrations in the patients in group A3 were relatively low in comparison to other groups (Table 8 and Table 9).

5.1.4.6 Genetics

HLA genotype distribution was consistent with groups A1 and A3 defining non-T1D forms. None of the typed patients in these groups had high T1D risk genotypes *DR3/DR4-DQ8* or *DR4-DQ8/DR4-DQ8* and >90% of patients in each of these groups had either neutral or protective HLA genotypes (Table 8 and Table 9; Figure 7E). In contrast, protective or neutral genotypes were found in <60% of cases in all other groups. There was, however, substantial variation in the relative frequencies between these groups. Noticeably, group A2, B5 and B7 had more patients with protective or neutral genotypes. We also examined non-HLA genetic risk using a genetic risk score that did not include HLA *DR-DQ* risk (Figure 6F). Within the islet AAB negative and positive patients, the genetic risk score was inversely correlated with the HLA genotype risk category (R2=-0.11; p=0.005). The median genetic risk score was lowest in Groups A1 (10.15) and A3 (9.74) and highest in groups B1 (11.06) and B7 (11.21). Patient genetic risk score varied substantially within each group and there was evidence of a bimodal distribution in Group B2 (Figure 7F).

5.1.4.7 Relationships between variables

In addition to the strong correlations between cytokine concentrations, vitamin D3 concentration was correlated with IL-8 concentration and also weakly with IL-12p70 concentration. Vitamin D3 was also inversely correlated with the age of onset, HbA1c, insulindose and triglycerides. c-peptide was strongly positively correlated with BMI SDS, triglycerides, the age of diabetes onset and family history of T2D and negatively correlated

with insulin therapy, islet AABs, ketonuria and HLA risk genotypes (Figure 8A).

5.1.5 Endotypes

We examined the CART subgroups for evidence of endotype (Table 10). Patients in Group A1 had few features of T1D, and contained the highest proportion of monogenic diabetes cases. Group A3 patients were obese, older and had a number of features associated with T2D. Group A2 had clear features of T1D. Group B1 was characterized by young patients and was the only group that exhibited a substantial proportion of cases with an inflammatory cytokine profile. This group also had a relatively low HbA1c and only 55% of patients presented with ketonuria. It may, therefore, be considered a more acute form of T1D with a strong inflammatory component. Group B2 appeared as a milder and less inflammatory endotype of Group B1 with a slightly older age of onset. Groups B3, B4, and B6 had similar genetic, autoimmune, and typical clinical characteristics of T1D and differed with respect to age (B6 older) and BMI (B4 higher), suggesting that they may be continuums within a similar endotype. Group B5 and B7 had a teenage onset form and similar characteristics except for BMI (higher in B7) and HbA1c (lower in B5). B5 appears to be a milder or earlier diagnosed form of the same endotype.

Table 7. Study characteristics of 1194 DiMelli patients, 106 with no and 1088 with islet AABs.

	Islet autoantibody negative		Islet autoantibody positive		
variable	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	***P*-value**
outcome					
fasting c-peptide [nmol/l]	106	0.28 [0.14;0.52]	1088	0.13 [0.07;0.23]	<0.001
diabetes duration [days]	105	10 [6; 15]	1086	9 [6; 12]	0.36
predictor variables					
sex (girls)		47 (44.3%)		481 (44.2%)	1.00
age [years]	106	13.50 [10.26;15.53]	1088	10.21 [6.93;13.22]	<0.001
BMI SDS	106	0.37 [-0.64;1.46]	1088	-0.58 [-1.38;0.22]	<0.001
HbA1c [%]	106	10.00 [8.12;12.05]	1088	10.80 [9.30;12.50]	0.004
Family History					
T1D		12 (11.3%)		78 (7.2%)	0.18
T2D		17 (16.0%)		25 (2.3%)	<0.001
other		3 (2.8%)		0 (0.0%)	<0.001
number of islet autoantibodies			1088		<0.001
0		106 (100.0%)		0 (0.0%)	
1		0 (0.0%)		116 (10.7%)	
2		0 (0.0%)		253 (23.3%)	
3		0 (0.0%)		359 (33.0%)	
4		0 (0.0%)		360 (33.1%)	
TPO positive		4 (3.8%)		110 (10.1%)	0.052
TTG positive		7 (6.6%)		84 (7.7%)	0.82

	Islet autoantibody negative		Islet autoantibody positive		
variable	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	***P*-value**
other variables not included in prediction					
treatment					<0.001
Insulin	106	86 (81.1%) *	1088	1065 (97.9%)	
Metformin		7 (6.6%) *		0 (0.0%)	
diet or no treatment started		15 (14.2%)		23 (2.1%)	
Insulindose [U/day/kg]	85	0.67	947	0.89	<0.001
		[0.28;1.09]		[0.61;1.18]	
Ketonuria		60 (67.4%)		903 (88.9%)	<0.001
genetic risk score †	63	10.20	596	10.65	0.009
		[9.50;11.08]		[9.95;11.25]	
HLA genotypes					<0.001
DR3/4-DQ08		10 (10.8%)		193 (20.6%)	
DR04-DQ08/DR04-DQ08		1 (1.1%)		83 (8.9%)	
DR04-DQ08/DRx-DQx		10 (10.8%)		165 (17.6%)	
DR03-DQ02/DRx-DQx		10 (10.8%)		80 (8.5%)	
neutral		27 (29.0%)		353 (37.7%)	
protective		35 (37.6%)		62 (6.6%)	
triglycerides [mg/dl] ‡	95	111.00	847	78.00	<0.001
		[74.50;153.00]		[60.00;104.00]	

* two patients are treated with metformin plus insulin
† without HLA *DR/DQ*
‡ not enough blood to assess all parameters in each child
A total of 1194 patients fulfilled data completeness for CART analysis, i.e. full data on the outcome variable fasting c-peptide and the predictor variables age, sex, BMI SDS, HbA1c, family history, number of islet autoantibodies and antibodies against thyreoid peroxidase or tissue transglutaminase. Italics mark categorization of variables in relation to the CART analysis.

Table 8. Study characteristics of 106 islet AAB negative patients included in the CART analysis.

	A1		A2		A3		
variable	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	***P*-value**
outcome							
fasting c-peptide [nmol/l]	17	0.50 [0.43;0.66]	65	0.17 [0.10;0.26]	24	0.86 [0.55;1.35]	<0.001
below LOD	0		8		0		
predictor variables							
sex (girls)		11 (64.7%)		24 (36.9%)		12 (50.0%)	0.099
age [years]	17	13.49 [10.24;15.61]	65	12.55 [8.11;15.32]	24	14.29 [13.43;15.63]	0.094
BMI SDS	17	0.34 [-0.50;0.75]	65	-0.29 [-1.28;0.50]	24	2.29 [1.79;2.84]	<0.001
HbA1c [%]	17	6.50 [5.80;6.90]	65	10.80 [9.30;13.00]	24	10.05 [7.90;10.90]	<0.001
Family History							
T1D		3 (17.6%)		8 (12.3%)		1 (4.2%)	0.37
T2D		1 (5.9%)		6 (9.2%)		10 (41.7%)	<0.001
other		2 (11.8%)		0 (0.0%)		1 (4.2%)	0.030
TPO positive		0 (0.0%)		3 (4.6%)		1 (4.2%)	0.67
TTG positive		0 (0.0%)		7 (10.8%)		0 (0.0%)	0.094
other variables not included in prediction							
treatment							<0.001
Insulin		5 (29.4%)		64 (98.5%)		17 (70.8%) *	
Metformin		0 (0.0%)		0 (0.0%)		7 (29.2%) *	
diet or no treatment started		12 (70.6%)		1 (1.5%)		2 (0.8%)	
Insulindose [U/day/kg]	12	0.00 [0.00;0.08]	57	0.83 [0.52;1.10]	16	0.80 [0.19;1.17]	<0.001
Ketonuria		3 (23.1%)		47 (82.5%)		10 (52.6%)	<0.001
genetic risk score†	14	10.15 [9.76;10.64]	37	10.38 [9.47;11.19]	12	9.74 [9.46;10.80]	0.79
HLA genotypes							0.004
DR3/4-DQ08		0 (0.0%)		10 (17.2%)		0 (0.0%)	
DR04-DQ08/DR04-DQ08		0 (0.0%)		1 (1.7%)		0 (0.0%)	
DR04-DQ08/DRx-DQx		0 (0.0%)		9 (15.5%)		1 (4.8%)	
DR03-DQ02/DRx-DQx		1 (7.1%)		9 (15.5%)		0 (0.0%)	
neutral		4 (28.6%)		17 (29.3%)		6 (28.6%)	
protective		9 (64.3%)		12 (20.7%)		14 (66.7%)	
triglycerides [mg/dl] ‡	14	106.00 [61.25;145.75]	58	92.50 [70.25;129.00]	23	156.00 [106.00;277.50]	<0.001
calcidiol [ng/ml] ‡	16	41.10 [27.62;48.81]	63	22.53 [15.90;35.31]	23	10.90 [7.69;25.81]	<0.001

	A1		A2		A3		
variable	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	***P*-value**
IFN-γ[pg/ml] ‡	13	4.07 [3.44;7.97]	37	6.10 [3.76;10.15]	13	4.14 [3.03;5.03]	0.045
IL-10 [pg/ml] ‡	13	0.45 [0.41;0.55]	37	0.86 [0.46;1.59]	13	0.44 [0.31;0.57]	0.017
IL-12p70 [pg/ml] ‡	13	0.23 [0.10;0.82]	37	0.38 [0.19;0.96]	13	0.30 [0.21;0.38]	0.37
IL-1β [pg/ml] ‡	13	0.44 [0.28;3.62]	37	1.07 [0.35;6.30]	13	0.62 [0.25;1.01]	0.23
IL-2 [pg/ml] ‡	13	0.02 [0.00;0.45]	37	0.33 [0.12;1.11]	13	0.00 [0.00;0.08]	<0.001
IL-6 [pg/ml] ‡	13	0.93 [0.48;1.54]	37	3.02 [0.68;10.31]	13	1.02 [0.89;1.66]	0.42
IL-8 [pg/ml] ‡	13	94.89 [12.37;258.85]	37	374.22 [19.60;3871.30]	13	17.33 [8.16;438.32]	0.086
TNF-α [pg/ml] ‡	13	3.23 [2.42;4.55]	37	3.98 [2.98;7.83]	13	2.95 [2.21;4.12]	0.085

* two patients are treated with metformin plus insulin
† without HLA *DR/DQ*
‡ not enough blood to assess all parameters in each child

Table 9. Study characteristics of 1088 islet AAB positive patients included in the CART analysis.

	B1		B2		B3		B4		B5		B6		B7		
variable	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	***P*-value**
outcome															
fasting c-peptide [nmol/l]	47	0.13 [0.00;0.20]	22	0.33 [0.21;0.46]	357	0.10 [0.00;0.13]	170	0.10 [0.07;0.20]	56	0.38 [0.26;0.54]	388	0.17 [0.10;0.23]	48	0.28 [0.20;0.40]	<0.001
below LOD	15		3		168		55		1		73		2		
predictor variables															
sex (girls)		21 (44.7%)		9 (40.9%)		185 (51.8%)		81 (47.6%)		20 (35.7%)		153 (39.4%)		12 (25.0%)	0.001
age [years]	47	5.02 [3.05;6.08]	22	9.13 [8.14;10.18]	357	7.21 [4.57;9.06]	170	7.79 [4.88;9.20]	56	14.09 [12.62;15.95]	388	13.53 [12.06;14.85]	48	13.14 [11.85;15.71]	<0.001
BMI SDS	47	-0.66 [-1.35;0.13]	22	-0.49 [-0.79;0.53]	357	-1.17 [-1.81;-0.62]	170	0.62 [0.21;1.06]	56	0.01 [-0.54;1.23]	388	-0.74 [-1.49;-0.10]	48	1.42 [1.15;1.87]	<0.001
HbA1c [%]	47	7.20 [6.40;7.65]	22	7.10 [6.45;7.60]	357	10.80 [9.70;12.10]	170	10.80 [9.43;12.07]	56	7.30 [6.30;7.93]	388	11.85 [10.50;13.50]	48	10.85 [10.00;12.10]	<0.001
Family History															
T1D		6 (12.8%)		3 (13.6%)		22 (6.2%)		10 (5.9%)		7 (12.5%)		26 (6.7%)		4 (8.3%)	0.30
T2D		0 (0.0%)		0 (0.0%)		1 (0.3%)		4 (2.4%)		7 (12.5%)		12 (3.1%)		1 (2.1%)	<0.001
other		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)	1
number of islet autoantibodies															0.34
1		6 (12.8%)		4 (18.2%)		42 (11.8%)		16 (9.4%)		7 (12.5%)		34 (8.8%)		7 (14.6%)	
2		9 (19.1%)		4 (18.2%)		90 (25.2%)		34 (20.0%)		12 (21.4%)		97 (25.0%)		7 (14.6%)	
3		12 (25.5%)		5 (22.7%)		108 (30.3%)		60 (35.3%)		26 (46.4%)		133 (34.3%)		15 (31.2%)	
4		20 (42.6%)		9 (40.9%)		117 (32.8%)		60 (35.3%)		11 (19.6%)		124 (32.0%)		19 (39.6%)	
TPO positive		3 (6.4%)		0 (0.0%)		26 (7.3%)		16 (9.4%)		5 (8.9%)		54 (13.9%)		6 (12.5%)	0.043
TTG positive		6 (12.8%)		2 (9.1%)		37 (10.4%)		11 (6.5%)		3 (5.4%)		23 (5.9%)		2 (4.2%)	0.19

	B1		B2		B3		B4		B5		B6		B7		
variable	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	***P*-value**
other variables not included in prediction															
IAA positive		37 (88.1%)		12 (66.7%)		268 (76.4%)		122 (72.6%)		36 (76.6%)		231 (60.3%)		32 (69.6%)	<0.001
GADA positive		36 (76.6%)		14 (63.6%)		220 (61.6%)		120 (70.6%)		37 (66.1%)		284 (73.2%)		34 (70.8%)	0.029
IA2A positive		36 (76.6%)		17 (77.3%)		267 (74.8%)		134 (78.8%)		40 (71.4%)		311 (80.2%)		37 (77.1%)	0.62
ZnT8A positive		31 (66.0%)		20 (90.9%)		259 (72.5%)		128 (75.3%)		40 (71.4%)		297 (76.5%)		39 (81.2%)	0.24
treatment															<0.001
Insulin		43 (91.5%)		19 (86.4%)		354 (99.2%)		169 (99.4%)		50 (89.3%)		384 (99.0%)		46 (95.8%)	
Metformin		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)		0 (0.0%)	
diet or no treatment started		4 (8.5%)		3 (13.6%)		3 (0.8%)		1 (0.6%)		6 (10.7%)		4 (1.0%)		2 (4.2%)	
Insulindose [U/day/kg]	37	0.43 [0.10;0.75]	16	0.40 [0.18;0.53]	313	1.00 [0.71;1.30]	154	0.79 [0.58;1.02]	48	0.40 [0.30;0.64]	340	0.97 [0.74;1.26]	39	0.86 [0.58;1.06]	<0.001
Ketonuria		22 (55.0%)		12 (66.7%)		313 (91.5%)		147 (91.3%)		31 (68.9%)		336 (92.3%)		42 (91.3%)	<0.001
genetic risk score *	23	11.06 [10.16;11.65]	15	10.40 [9.32;11.53]	204	10.69 [9.93;11.21]	95	10.67 [10.20;11.18]	24	10.48 [9.80;10.77]	215	10.55 [9.93;11.21]	20	11.21 [10.18;11.41]	0.28
HLA genotypes															0.47
DR3/4-DQ08		9 (25.0%)		6 (33.3%)		71 (23.6%)		34 (22.5%)		7 (15.6%)		64 (18.5%)		2 (5.1%)	
DR04-DQ08/ DR04-DQ08		3 (8.3%)		1 (5.6%)		19 (6.3%)		15 (9.9%)		5 (11.1%)		35 (10.1%)		5 (12.8%)	
DR04-DQ08/ DRx-DQx		8 (22.2%)		0 (0.0%)		55 (18.3%)		30 (19.9%)		5 (11.1%)		59 (17.1%)		8 (20.5%)	
DR03-DQ02/ DRx-DQx		4 (11.1%)		4 (22.2%)		23 (7.6%)		13 (8.6%)		4 (8.9%)		29 (8.4%)		3 (7.7%)	
neutral		9 (25.0%)		6 (33.3%)		111 (36.9%)		53 (35.1%)		21 (46.7%)		134 (38.7%)		19 (48.7%)	
protective		3 (8.3%)		1 (5.6%)		22 (7.3%)		6 (4.0%)		3 (6.7%)		25 (7.2%)		2 (5.1%)	
triglycerides [mg/dl] †	41	64.00 [53.00;81.00]	19	62.00 [55.50;78.50]	280	73.00 [57.00;91.00]	134	75.00 [57.25;108.00]	45	74.00 [57.00;106.00]	290	85.50 [67.25;108.75]	38	118.50 [82.25;155.00]	<0.001

	B1		B2		B3		B4		B5		B6		B7		
variable	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	**n**	**median [IQR] or n(%)**	***P*-value**
calcidiol [ng/ml] †	42	27.56 [18.22;34.97]	21	28.76 [20.62;37.31]	329	26.40 [17.12;34.76]	161	28.19 [18.24;37.42]	51	20.28 [13.03;32.07]	356	24.15 [15.10;32.61]	46	22.32 [14.43;29.90]	0.017
IFN-γ [pg/ml] †	23	10.33 [6.46;21.20]	15	6.30 [4.68;10.24]	209	7.03 [4.55;13.90]	95	5.99 [4.08;9.64]	24	4.56 [3.25;8.73]	216	4.74 [3.45;7.35]	21	4.91 [3.35;6.16]	<0.001
IL-10 [pg/ml] †	23	1.02 [0.61;1.45]	15	0.83 [0.48;1.27]	209	0.95 [0.62;2.30]	95	0.73 [0.55;1.60]	24	0.83 [0.52;1.33]	216	0.65 [0.48;1.24]	21	0.60 [0.37;1.04]	<0.001
IL-12p70 [pg/ml] †	23	0.34 [0.15;0.58]	15	0.33 [0.16;0.67]	209	0.34 [0.18;0.83]	95	0.28 [0.13;0.63]	24	0.32 [0.13;0.44]	216	0.26 [0.12;0.57]	21	0.26 [0.10;0.51]	0.21
IL-1β [pg/ml] †	23	1.05 [0.29;2.29]	15	0.35 [0.18;5.09]	209	0.87 [0.35;3.90]	95	0.78 [0.34;3.24]	24	0.64 [0.26;2.61]	216	0.83 [0.26;2.90]	21	0.78 [0.31;2.01]	0.80
IL-2 [pg/ml] †	23	0.27 [0.15;0.94]	15	0.43 [0.08;0.76]	209	0.36 [0.13;1.49]	95	0.18 [0.04;0.52]	24	0.12 [0.02;0.81]	216	0.24 [0.08;0.71]	21	0.19 [0.01;0.55]	0.010
IL-6 [pg/ml] †	23	1.90 [0.70;3.53]	15	1.75 [0.54;9.31]	209	2.04 [0.67;7.33]	95	1.66 [0.73;5.85]	24	1.32 [0.65;8.49]	216	1.32 [0.65;4.88]	21	1.41 [0.76;3.68]	0.87
IL-8 [pg/ml] †	23	70.32 [15.89;1393.30]	15	98.08 [16.77;1973.58]	209	286.30 [42.55;2628.54]	95	234.97 [55.48;2072.16]	24	146.49 [63.66;925.24]	216	245.04 [56.48;1666.04]	21	470.27 [112.50;1235.35]	0.83
TNF-α [pg/ml] †	23	4.60 [2.93;8.00]	15	3.89 [2.01;6.90]	209	3.79 [2.12;7.42]	95	3.81 [2.39;6.88]	24	3.15 [1.87;5.82]	216	3.04 [1.69;5.41]	21	3.42 [1.85;4.63]	0.045

* without HLA *DR/DQ*

† not enough blood to assess all parameters in each child

Table 10. Endotypes of diabetes in childhood.

E1 (CART A1)	Islet AAB negative, relatively abundant c-peptide, normal BMI, moderate HbA1c, normal Vitamin D, no HLA risk, protective T1D genes, monogenetic family history
E2 (CART A2)	Islet AAB negative, low c-peptide, low or normal BMI, ketones, T1D genetic susceptibility, celiac autoimmunity, elevated IL-10 and IL-2 – T1D like
E3 (CART A3)	Islet AAB negative, high BMI, high first degree family history of T2D, no HLA risk – T2D like
E4 (CART B1)	Islet AAB positive, young age, moderate HbA1c, low c-peptide, inflammatory (high cytokines), low ketones, high genetic risk, IAA positive
E5 (CART B2)	Islet AAB positive, young age, moderate HbA1c, *positive* c-peptide, non-inflammatory,
E6 (CART B3/B4/B6)	Islet AAB positive, typical T1D with variation in age at onset
E7 (CART B5)	Islet autoantibody positive, teenage, moderate HbA1c, c-peptide positive, T2D family history, less insulin treatment, less ketones
E8 (CART B7)	Islet AAB positive, teenage, c-peptide positive, overweight, male, high triglycerides

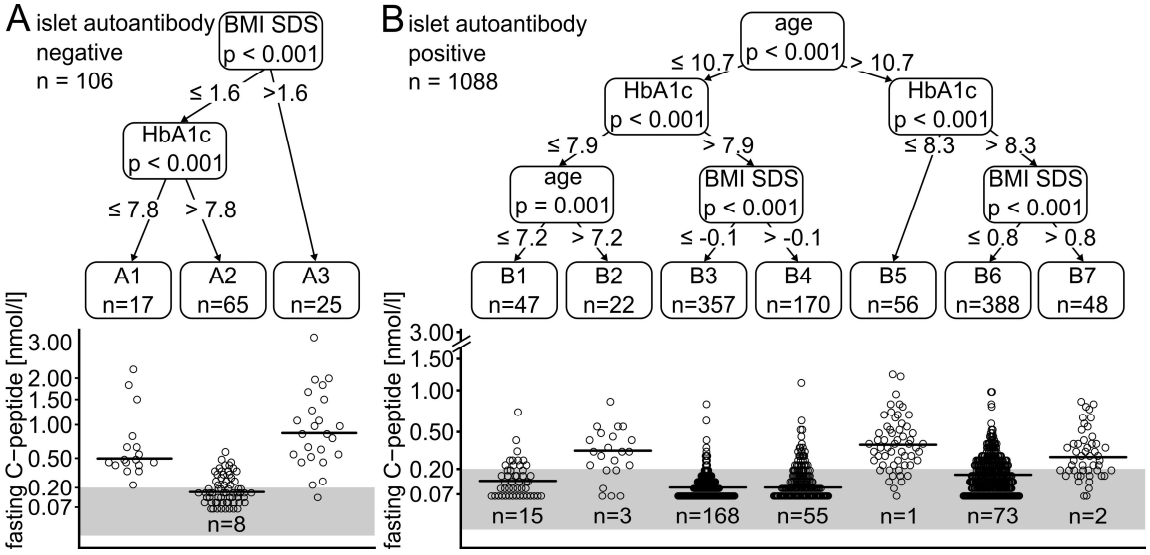


Figure 6. CART analysis of (A) islet autoantibody negative and (B) islet autoantibody positive patients. The limit of detection (LOD) for c-peptide is 0.07 nmol/l. The number in the shaded area indicates the number of samples at/below the LOD.

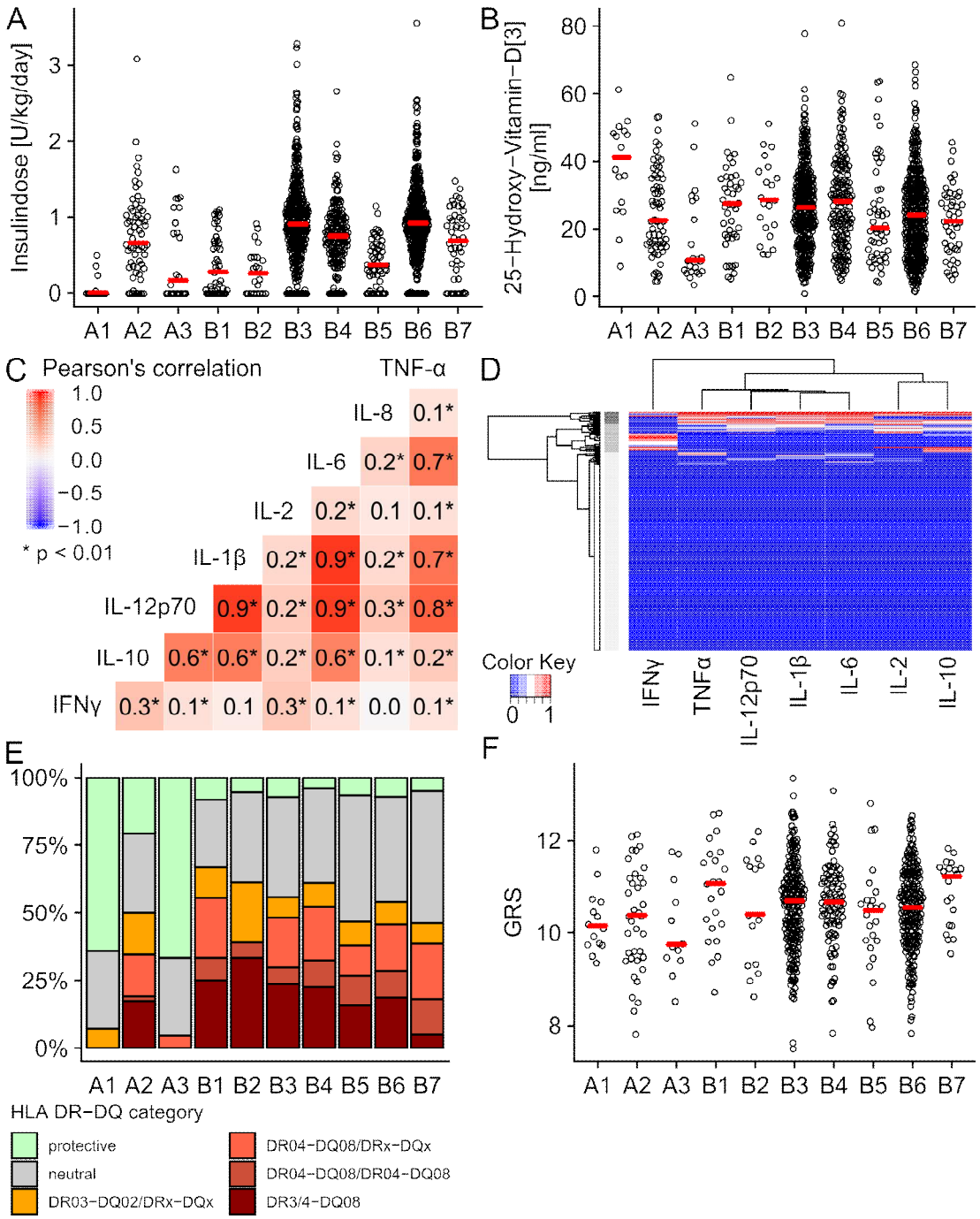


Figure 7. Distribution of features in DiMelli patients grouped according to CART analysis of islet AAB negative (A1-3) and positive (B1-7) subjects. (A) Insulindose and **(B)** 25-Hydroxy-Vitamin D3 and **(C)** Pearson's correlation of serum concentrations of inflammatory cytokines among 666 DiMelli patients with the full set of measurements. **(D)** Heatmap on cytokine concentrations categorized into centiles (0-90%, 90-92.5%, 92.5%-95%, 95%-97.5%, 97.5%-100%) and scaled from 0 to 1 separately for each cytokine. Hierarchical clustering on the samples revealed three clusters indicated by the grey-scaled column left of the heatmap. Description of genetic features of nodes from CART analysis assessed by **(E)** distribution of HLA *DR-DQ* genotypes and **(F)** genetic risk score (GRS) of non-HLA loci.

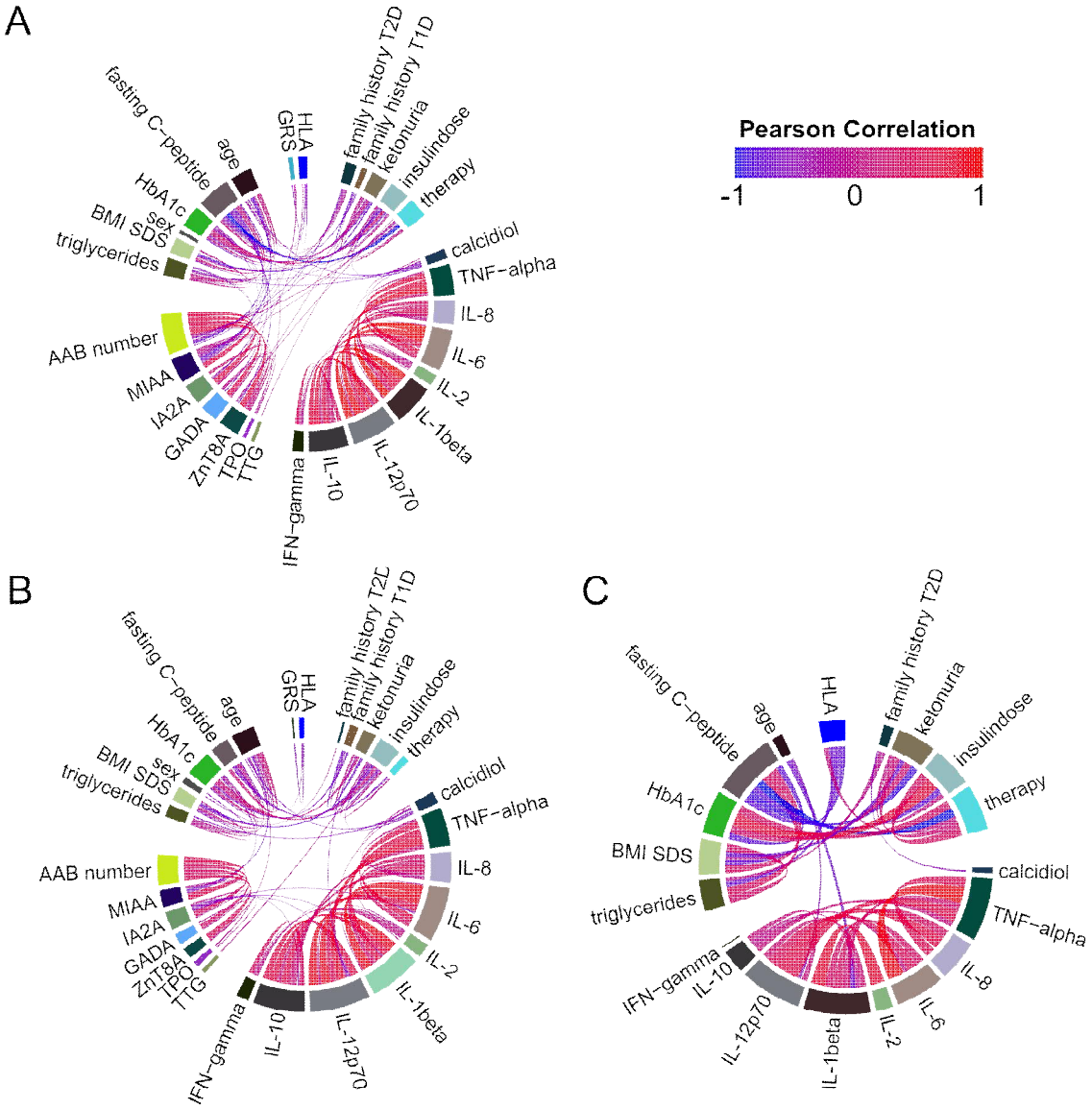


Figure 8. Relationships between assessed variables by Pearson Correlation in (A) all patients or (B) islet autoantibody positive and (C) negative patients. Only correlations with a p<0.01 and R^2>0.1 are displayed.

5.2 Heterogeneity in genetic predisposition (TEDDY study)

Children with a first-degree family history of T1D have a risk of developing islet AABs and diabetes that is significantly higher than the risk in children from the general population without a family history of the disease (15, 44, 45). Enrichment of T1D-susceptibility genotypes of HLA and other genes is likely to contribute to the inflated risk in children with a first-degree family history of T1D. Understanding the genetic differences and their contributions to the divergent risks between these groups could provide paradigms to identify novel genetic and environmental factors that modify risk, and to identify children from the general population whose *a priori* risk of developing T1D is similar to that of children with a first-degree family history of T1D. The TEDDY cohort includes both, children with a first-degree family history of T1D and children from the general population, providing a rare opportunity to examine the excess risk of developing islet AABs and diabetes in affected families. Extensive genotype data is available (48). I calculated genetic risk scores representing cumulative genetic susceptibility (71-73), and investigated which genes, beyond the known susceptibility regions, may contribute to risk.

Using this approach, I detected enrichment of genetic susceptibility for multiple known risk genes and a novel risk gene. Matching children with a first-degree family history of T1D and children from the general population for genetic risk completely abrogated the excess risk of children with a first-degree family history of T1D in the highest genetic risk stratum, but not in the lower genetic risk strata. These findings provide evidence that additional factors preferentially contribute to T1D risk in children without a full complement of genetic susceptibility.

The results of this analysis have been published as:
Hippich, M., A. Beyerlein, W. A. Hagopian, J. P. Krischer, K. Vehik, J. Knoop, C. Winker, J. Toppari, A. Lernmark, M. J. Rewers, A. K. Steck, J. X. She, B. Akolkar, C. C. Robertson, S. Onengut-Gumuscu, S. S. Rich, E. Bonifacio and A. G. Ziegler (2019). "Genetic Contribution to the Divergence in Type 1 Diabetes Risk Between Children From the General Population and Children From Affected Families." Diabetes. 2019 Jan 17. pii: db180882. doi: 10.2337/db18-0882. [Epub ahead of print]

5.2.1 Risk of developing islet autoantibodies and diabetes according to HLA genotype and first-degree relative status

Of the TEDDY children with at least one follow-up sample (n=7894), 3035 (38.4%) had the HLA *DR3/DR4-DQ8* genotype, and 1537 (19.5%) had the HLA *DR4-DQ8/DR4-DQ8* genotype. Of these 4572 children, 423 (9.3%) had a first-degree relative with T1D and 4149 were from the general population (Table 11, Figure 9). One or more islet AABs developed in 500 (10.9%) children, of which 324 (7.1%) had multiple islet AABs. Diabetes was diagnosed in 192 (4.2%) children.

Matching for HLA *DR-DQ* provided an estimate of the excess risk of developing islet AABs or diabetes in children with a first-degree relative with T1D that was due to factors other than enrichment for these genotypes (Figure 8). Matching for HLA genotypes was sufficient to reduce the >10-fold excess risk usually observed in children from affected families to below 3-fold. The cumulative risk (95%CI) by 6 years of age in HLA *DR3/DR4-DQ8* children with a first-degree relative with T1D was 20.5% (15.4–25.4%) for one or more islet AABs, 17.0% (12.1–21.5%) for multiple islet AABs, and 6.8% (3.8–9.7%) for diabetes compared with 10.0% (8.8–12.2%; p<0.001), 6.4% (5.4–7.4%; p<0.001), and 2.7% (2.0–3.3%; p<0.001), respectively, in children from the general population with these genotypes (Figure 9A,C,E). Similar differences were observed in the HLA *DR4-DQ8/DR4-DQ8* children (Figure 9B,D,F). A first-degree family history of T1D was associated with an increased incidence of islet autoimmunity in the first 3 years of life in children with the high-risk HLA genotypes (Figure 10A). This was similar if the outcome was defined as the detection of IAA before other AABs or GADA as the first islet AAB (Figure 10B).

5.2.2 DRB1*04 allele subtype enrichment in children from affected families

The risk for T1D is influenced by the *HLA-DRB1*04* allele (52, 146). I, therefore, searched for enrichment of *DRB1*04* subtypes in children with a first-degree relative with T1D (Table 12). The high risk *DRB1*04:01* allele was more frequent in the children with a first-degree relative with T1D than in children from the general population (p<0.001) for children with either the *DR3/4-DQ8* (p=0.007) or *DR4-DQ8/DR4-DQ8* (p<0.001) genotypes. In contrast, the lower risk *DRB1*04:04* (*P*<0.001) and *DRB1*04:07* (p=0.035) alleles were less frequent in the children with a first-degree relative with T1D than in children from the general population. There were no differences in the frequencies of the *DRB1*04:02* or the

*DRB1*04:05* alleles between the children with a first-degree relative with T1D and children from the general population. The remaining *DRB1*04* alleles were infrequent in the study population and were not considered.

5.2.3 Genetic risk scores in children with a first-degree relative with T1D and children from the general population

The additional risk conferred by the non-HLA *DR-DQ* genes was expressed as a genetic risk score from 40 of the non-HLA *DR-DQ* genes (73). Genetic risk scores were higher in children with a first-degree relative with T1D (median, interquartile range: 10.3, 9.7–11.0) than in children from the general population (10.1, 9.4–10.7; $p<0.001$; Figure 11A). Enrichment of risk genotypes reached significance for six of the 40 SNPs included in the score (Table 13). Similar genetic risk scores were observed in the multiple islet AAB positive children with a first-degree relative with T1D (median, interquartile range: 10.4, 10.0–11.3) and in children from the general population (10.5, 10.0–11.1; $p=0.97$; Figure 11B).

5.2.4 Additional genes with allele enrichment in children with a first-degree relative with T1D

HLA *DR-DQ* susceptible genotypes have additional variants in susceptibility genes that are in linkage disequilibrium (LD) with HLA *DR-DQ* (147). I reasoned that the frequencies of susceptibility genotypes of such genes may be increased in children with a first-degree relative with T1D and could account for some of the excess risk in these children. With the rare opportunity to examine a large number of children from the general population and children with a first-degree relative with T1D matched for HLA *DR-DQ* genotype in the TEDDY study, I examined 6097 SNPs with minor allele frequencies >5% (MAF > 0.05) on the short arm of chromosome 6, containing the HLA genes (6p21.32). Two SNPs, rs3763305 ($p=7.27\times10^{-7}$) and rs3817964 ($p=8.26\times10^{-7}$) were enriched in children with a first-degree relative with T1D (Figure 12A). Both these SNPs are intronic variants of *BTNL2* and are in complete LD ($r^2=1$). The *BTNL2* gene is located in the human MHC, in a cluster of over 100 genes that include the HLA class I and HLA class II genes close to the HLA-*DRA*/HLA-*DRB5*/HLA-*DRB1* cluster and HLA-*DQB1*. Extension of the analysis to all 111,069 ImmunoChip SNPs that passed QC filters identified rs7735139 (intronic SNP in *ITGA1*, Integrin Subunit Alpha 1, 5q11.2) with allelic enrichment in children with a first-degree relative with T1D ($p=4.34\times10^{-8}$; Figure 12B).

5.2.5 Genetic contribution to the additional risk for islet autoantibodies and diabetes in children with a first-degree relative with T1D

Cox proportional hazards models were used to assess the development of one or more islet AABs, multiple islet AABs, and diabetes (Table 14). Model 1 examined the excess risk conferred by a first-degree family history of T1D adjusted for HLA genotype (*DR3/4-DQ8* vs *DR4-DQ8/DR4-DQ8*), sex, and country of origin. Hazard ratios (HRs) in children with a first-degree relative with T1D were 2.12 (95% CI, 1.65–2.72) for one or more islet AABs, 2.77 (95% CI, 2.09–3.68) for multiple islet AABs, and 3.69 (95% CI, 2.60–5.23) for diabetes.

To determine whether the genetic factors that were enriched in children with a first-degree relative with T1D contributed to the excess risk in children with a first-degree relative with T1D, Model 2 additionally included the non-HLA *DR-DQ* genetic risk score, *DRB1*04* subtype *(enriched: DR3/DRB1*04:01* and *DRB1*04:01/DR*04:xx* where **04:xx* is neither **04:04* nor **04:07* vs non-enriched: other genotypes), the *BTNL2* SNP rs3763305 (enriched: GG vs non-enriched: GA/AA genotypes), and the *ITGA1* SNP rs7735139 (enriched: GG vs non-enriched: GA/AA genotypes). HRs associated with a first-degree family history of T1D were reduced to 1.82 (95% CI, 1.42–2.35) for one or more islet AABs, 2.26 (95% CI, 1.70–3.02) for multiple islet AABs, and 2.92 (95% CI, 2.05–4.16) for diabetes. The enriched *DRB*04:01* subtype (HR, 1.48; 95% CI, 1.08–2.01; p=0.014), and the non-HLA *DR-DQ* genetic risk score (HR, 1.66; 95% CI, 1.47–1.88; p<0.001) contributed to the risk of multiple islet AABs. Similar HRs for these variables were also observed for the risk of one or more islet AABs and diabetes, some of which reached significance. The *BTNL2* SNP rs3763305 GG genotype conferred additional risk for diabetes (HR, 1.80; 95% CI, 1.11–2.93; p=0.017; Table 14), and this additional risk was also observed when the analysis was restricted to children with a HLA *DR3/4-DQ8* genotype (HR, 1.92; 95% CI, 1.11–3.35; p=0.021; Table 15). The *ITGA1* SNP was not associated with the risk of islet AABs or diabetes.

5.2.6 Association between DRB1*04 subtypes and islet autoantibodies or diabetes

To further assess whether the *DRB1*04:01* allele was associated with increased risk compared with the other *DRB1*04* alleles in the TEDDY children, I removed the confounder of family history, and performed a Kaplan-Meier analysis in children from the general population. Among the 1,876 children with the *DRB1*04:01* allele, the cumulative risks (95% CI) at 6 years old were 11.4% (9.8–12.9%) for one or more islet AABs, 7.7% (6.3–9.0%) for

multiple islet AABs, and 2.9% (2.1–3.6%) for diabetes compared with 6.8% (5.7–8.0%; p<0.001), 4.0% (3.1–4.9%; p<0.001), and 1.4% (0.9–1.9%; *P*<0.001), respectively, among 2,176 children without *DRB1*04:01* (Figure 13A,C,E). These differences remained when the analysis was limited to children from the general population with the HLA *DR3/DR4-DQ8* genotype (Figure 13B,D,F).

5.2.7 Association between BTNL2 genotypes and islet autoantibodies or diabetes

Variants in the *BTNL2* gene have not been implicated as independent genetic risk factors for T1D previously, likely due to the extensive LD in the MHC region and inadequate sample size. The *BTNL2* rs3763305 GG genotype distribution was increased in children who developed one or more islet AABs (p<0.001), multiple islet AABs (p<0.001), or diabetes (p<0.001), compared with children who remained islet AAB negative. These associations were observed separately for children with HLA *DR3/4-DQ8* or *DR4-DQ8/DR4-DQ8* (Table 16). The association between the *BTNL2* rs3763305 GG genotype and T1D was also validated in a separate case-control study after stratification for HLA *DR3/4-DQ8* (Table 17).
The ImmunoChip contained 88 SNPs within the *BTNL2* gene, including 34 that passed quality criteria. R package haplo.stats was used to generate haplotypes and subsequently genotypes (Table 18). The risk associated with the four most frequent genotypes among children with *DR3/DR4-DQ8* and the four most frequent genotypes among children with *DR4-DQ8/DR4-DQ8* was stratified by the presence of haplotype 28, which uniquely contained an A allele at *BTNL2* rs3763305, and a T allele at *BTNL2* rs3817964 (Figure 14).
BTNL2 lies close to the HLA-*DRB5*/HLA-*DRB6*/HLA-*DRB1* protein-coding genes in a region of high LD. HLA DR3 was in nearly complete LD with the rs3763305 G allele: the *BTNL2* rs3763305 GG genotype was identified in 1,608 (99.3%) of 1,619 children who had the HLA *DR3/DR3* genotype (Table 19). I then examined the second *BTNL2* rs3763305 allele in children with *DR3/DR4-DQ8*. The *BTNL2* rs3763305 G allele was in nearly complete LD with *DRB1*04:01* (allele frequency, 99.5%), *DRB1*04:02* (99.4%), and *DRB1*04:05* (100.0%), whereas the *BTNL2* rs3763305 A allele was associated with *DRB1*0404* (39.2%) and *DRB1*0407* (34.4%) (p<0.001). These associations were confirmed in a separate cohort of 149 children with T1D and the *DR3/DR4-DQ8* genotype from Bavaria, Germany (Table 20). The *BTNL2* rs3763305 A allele was also observed together with the protective HLA *DRB1*04:03* allele in five (56%) of nine informative genotypes, and with the protective HLA *DRB1*13:01* allele in 10 (25%) of 40 informative genotypes, but not with other HLA *DRB1* alleles in the German cohort.

The *DRB1*04:04* allele was relatively frequent in the TEDDY children and was found with either the *BTNL2* rs3763305 G or the *BTNL2* rs3763305 A allele and, therefore, provided an opportunity to determine whether the *BTNL2* gene conferred an independent risk to the HLA-*DR4* subtype. Risks associated with *BTNL2* genotypes were examined in children with the *DR3/DRB1*04:04-DQ8* or *DRB1*04:04-DQ8/DRB1*04:04-DQ8* genotypes. The cumulative risks (95% CI) at 6 years old in children with the *BTNL2* rs3763305 GG genotype were 9.8% (5.6–13.8%) for one or more islet AABs, 6.3% (2.9–9.6%) for multiple islet AABs, and 3.7% (1.3–6.0%) for diabetes, compared with 8.0% (6.2–9.8%; p=0.46), 4.7% (3.3–6.2%; p=0.096), and 1.6% (0.8–6.0%; P=0.005) in the children with the GA or AA genotypes (Figure 15). Cox proportional hazards models adjusted for the potential contribution of *DR3/DRB1*04:04-DQ8* to risk as compared with the *DRB1*04:04-DQ8/ DRB1*04:04-DQ8* genotype replicated the additional risk for diabetes conferred by the *BTNL2* rs3763305 GG genotype (p=0.009, Table 21). The additional risk conferred by the *BTNL2* rs3763305 GG genotype may be due to a specific association with specific *DRB1*04:04* subtypes. We, therefore, examined the relationship between the *BTNL2* rs3763305 alleles and *DRB1*04:04* subtypes in the German cohort of patients who had been HLA genotyped by sequencing of HLA *DRB1* exon 2, which harbors variations in all 12 subtypes of *DRB1*04:04*. All subjects with *DRB1*04:04* had the *DRB1*04:04:01* allele regardless of whether the *BTNL2* rs3763305 was A or G, indicating that the *BTNL2* rs3763305 A allele does not appear to mark a subtype of *DRB1*04:04*.

Finally, we examined the effect of BTNL2 knockdown on *in vitro* immune responses (Figure 16). Compared with nontargeting siRNA control treated dendritic cells, *BTNL2*-targeted siRNA-treated dendritic cells had increased naïve alloreactive CD4+ T-cell activation (p=0.031) but not memory antigen-specific CD4+ T-cell activation (p=0.43).

5.2.8 Risk excess in children with a first-degree relative with T1D after stratification by genetic risk

I asked whether the observed enrichment of T1D genetic susceptibility in children with a first-degree relative with T1D could account for their excess risk and defined four risk strata by HLA *DRB1*04* subtype and genetic risk score. These strata were able to discriminate the risk of developing islet AABs and T1D in children from the general population (Figure 17A,C,E). A similar stratification in children from the general population was observed if the strata were defined by the children's *BTNL2* genotype and genetic risk score (Figure 18). In contrast to children from the general population, a discrimination of risk in children with a first-degree relative with T1D was only achieved in the lowest risk stratum (Figure 17B,D,F). Comparing children with a first-degree relative with T1D and children from the general

population showed complete convergence of the risks of developing islet AABs and diabetes in the highest-risk stratum, and divergence of risk in the lower-risk strata (Table 22). The fold difference in risk for multiple islet AABs between children with a first-degree relative with T1D and children from the general population was 1.1 in the highest-risk stratum (14.3% vs. 12.7%), 1.9 in the second risk stratum (17% vs. 9%), 3.3 in the third risk stratum (14.8% vs. 4.5%), and 5.8 in the lowest-risk stratum (9.2% vs. 1.6%).

Table 11. Study characteristics by first-degree relative status.

Variable	children with a first-degree relative with T1D (n=423)	children from the general population (n=4149)
Males	200 (47.3%)	2082 (50.2%)
HLA genotype		
DR3/4-DQ8	280 (66.2%)	2755 (66.4%)
DR4-DQ8/DR4-DQ8	143 (33.8%)	1394 (33.6%)
Country		
US	194 (45.9%)	1750 (42.2%)
Finland	51 (12.1%)	792 (19.1%)
Germany	92 (21.7%)	209 (5.0%)
Sweden	86 (20.3%)	1398 (33.7%)
First-degree relative with T1D		
None	0 (0.0%)	4149 (100.0%)
Mother	146 (34.5%)	0 (0.0%)
Father	180 (42.6%)	0 (0.0%)
Sibling	79 (18.7%)	0 (0.0%)
Multiplex	18 (4.3%)	0 (0.0%)
Outcome events		
One or more islet autoantibodies	85 (20.1%)	415 (10.0%)
Multiple islet autoantibodies	69 (16.3%)	255 (6.1%)
First-appearing IAA	51 (12.1%)	227 (5.5%)
First-appearing GADA	46 (10.9%)	250 (6.0%)
Diabetes	47 (11.1%)	145 (3.5%)
Genetic risk score available	408 (96.5%)	4006 (96.6%)

Table 12. Allelic enrichment of *DRB1*04* subtypes in children with a first-degree relative with T1D.

DR4 subtype	children with a first-degree relative with T1D (alleles, n=535)	children from the general population (alleles, n=5,415)	*P*-value*
*DRB1*04:01*	329 (60.15%)	2788 (51.49%)	<0.001
*DRB1*04:02*	33 (6.03%)	291 (5.37%)	0.42
*DRB1*04:04*	137 (25.05%)	1928 (35.60%)	<0.001
*DRB1*04:05*	23 (4.20%)	226 (4.17%)	0.91
*DRB1*04:06*	1 (0.18%)	1 (0.02%)	0.17
*DRB1*04:07*	7 (1.28%)	153 (2.83%)	0.035
*DRB1*04:08*	4 (0.73%)	25 (0.46%)	0.33
*DRB1*04:10*	0 (0.00%)	1 (0.02%)	1
*DRB1*04:11*	1 (0.18%)	1 (0.02%)	0.17
*DRB1*04:13*	0 (0.00%)	1 (0.02%)	1

Each *DRB*04* allele was counted separately (once for children with the *DR3/DR4-DQ8* genotype and twice for children with the *DR4-DQ8/DR4-DQ8* genotype). *DRB*04* subtype information was missing in 77 *DR3/DR4-DQ8* and 35 *DR4-DQ8/DR4-DQ8* children. Children with the non-risk *DRB1*04:03* allele were excluded *a priori* from the TEDDY study unless they had a first-degree relative with T1D, and the 12 occurrences of this allele in children with a first-degree relative with T1D were therefore not considered. * *P*-values were calculated using Fisher's test.

Table 13 Genotype frequencies for SNPs used in the genetic risk score in children with a first-degree relative with T1D and children from the general population.

		Genotype frequency [%] *						
		children with a first-degree relative with T1D			Children from the general population			
SNP	**Gene**	**PP**	**SP**	**SS**	**PP**	**SP**	**SS**	***P*-value**
rs6897932	*IL7R*	11.8	34.5	53.7	7.7	41.0	51.3	0.002
rs1004446	*INS*	9.0	44.7	46.3	14.4	45.6	40.0	0.002
rs3825932	*CTSH*	10.9	40.0	49.2	13.9	44.1	42.1	0.014
rs3024505	*IL10*	2.4	20.9	76.8	2.2	27.0	70.8	0.025
rs3184504	*SH2B3*	27.4	46.8	25.8	30.9	48.8	20.3	0.027
rs2292239	*ERBB3*	42.8	44.0	13.2	46.4	44.1	9.6	0.043
rs1990760	*IFIH1*	13.5	50.1	36.3	18.3	47.4	34.3	0.053
rs10517086	Unknown	46.8	46.6	6.6	51.4	40.7	7.9	0.058
rs229541	*IL2B*	29.8	49.4	20.8	34.5	48.3	17.2	0.064
rs7804356	*SCAP2*	3.3	34.0	62.6	5.7	35.2	59.1	0.079
rs2069763	*IL2*	35.5	51.1	13.5	39.4	45.5	15.1	0.088
rs3757247	*BACH2*	28.9	48.6	22.5	33.2	47.7	19.1	0.10
rs4948088	*COBL*	0.2	5.9	93.9	0.3	8.8	90.9	0.12
rs7020673	*GLIS3*	21.7	47.3	31.0	24.4	49.0	26.5	0.12
rs2476601	*PTPN22*	76.1	22.5	1.4	80.3	18.6	1.1	0.12
rs2816316	*RGS1*	5.0	29.6	65.5	3.3	30.4	66.3	0.20
rs9388489	*C6ORF*	26.7	51.8	21.5	29.7	47.6	22.7	0.25
rs425105	*PRKD2*	1.4	28.6	70.0	2.7	27.1	70.2	0.26
rs45450798	*PTPN2*	66.7	29.3	4.0	69.6	27.5	2.9	0.26
rs5753037	Unknown	42.8	43.3	13.9	40.6	47.2	12.2	0.27
rs3087243	*CTLA4*	14.7	47.8	37.6	17.5	47.0	35.5	0.31

		Genotype frequency [%] *						
		children with a first-degree relative with T1D			Children from the general population			
SNP	**Gene**	**PP**	**SP**	**SS**	**PP**	**SP**	**SS**	***P*-value**
rs2395029	*HLA_B_5701*	0.0	1.2	98.8	0.0	2.2	97.7	0.32
rs17574546	Unknown	61.6	33.4	5.0	64.3	31.8	3.9	0.41
rs2281808	*SIRPG*	9.7	46.0	44.3	11.8	44.5	43.7	0.43
rs1465788	*ZFP36L1*	7.1	44.4	48.5	7.9	41.2	50.9	0.43
rs3788013	*UBASH3A*	35.9	50.1	13.9	35.0	48.6	16.4	0.44
rs4788084	*IL27*	17.5	47.8	34.8	19.4	48.3	32.3	0.50
rs763361	*CD226*	26.0	48.7	25.3	27.1	49.8	23.0	0.57
rs6920220	*TNFAIP3*	64.5	30.5	5.0	63.6	32.0	4.4	0.73
rs1738074	*TAGAP*	19.9	46.3	33.8	18.6	48.0	33.4	0.76
rs12708716	*CLEC16A*	12.6	42.2	45.3	12.2	44.0	43.8	0.76
rs7221109	Unknown	13.9	44.2	41.8	13.6	45.9	40.4	0.79
rs2664170	*GAB3*	56.5	22.7	20.8	57.6	21.4	20.9	0.83
rs4763879	*CD69*	37.7	49.1	13.3	38.9	47.7	13.4	0.86
rs947474	*PRKCQ*	3.5	30.5	66.0	3.2	30.1	66.8	0.89
rs5979785	*TLR8*	17.5	18.4	64.1	18.2	18.5	63.3	0.92
rs7202877	Unknown	79.7	19.1	1.2	79.1	19.5	1.3	0.95
rs1264813	*HLA_A_24*	82.7	16.6	0.7	82.3	17.0	0.7	0.98
rs10509540	*C10orf59*	7.1	38.3	54.6	7.3	38.3	54.4	0.98
rs12722495	*IL2R*	0.7	16.5	82.7	0.7	16.6	82.7	1.00

* P refers to the protective allele and S refers to the susceptible allele

Table 14. Cox proportional hazards models for developing islet autoantibodies and diabetes in children with a first-degree relative with T1D compared with children from the general population (reference).

	One or more islet autoantibodies				Multiple islet autoantibodies				Diabetes			
	Model 1 *		Model 2 *		Model 1 *		Model 2 *		Model 1 *		Model 2 *	
Variable	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**
first-degree relative with T1D	2.12 (1.65–2.72)	<0.001	1.82 (1.42–2.35)	<0.001	2.77 (2.09–3.68)	<0.001	2.26 (1.70–3.02)	<0.001	3.69 (2.60–5.23)	<0.001	2.92 (2.05–4.16)	<0.001
*DRB1*0401/x* †			1.42 (1.10–1.83)	0.007			1.48 (1.08–2.01)	0.014			1.38 (0.93–2.04)	0.10
Genetic risk score ‡			1.48 (1.34–1.64)	<0.001			1.66 (1.47–1.88)	<0.001			1.67 (1.42–1.96)	<0.001
BTNL2 rs3763305 GG §			1.05 (0.79–1.40)	0.73			1.39 (0.96–2.00)	0.080			1.80 (1.11–2.93)	0.017
ITGA1 rs7735139 GG §			1.18 (0.86–1.63)	0.30			1.06 (0.70–1.60)	0.77			1.11 (0.67–1.87)	0.68

* Model 1 and 2 are adjusted for sex, country (reference: US) and HLA genotype (reference: *DR4-DQ8/DR4-DQ8*)
† reference: *DRB1* without *0401* or *0401/0404* and *0401/0407*; ‡ per unit increase; § reference: GA/AA genotype

Table 15. Cox proportional hazards models for the development of islet autoantibodies and diabetes in children with a first-degree relative with T1D compared with children from the general population in children with the HLA *DR3/4-DQ8* genotype.

	One or more islet autoantibodies				Multiple islet autoantibodies				Diabetes			
	Model 1 *		Model 2 *		Model 1 *		Model 2 *		Model 1 *		Model 2 *	
Variable	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**
first-degree relative with T1D	2.24 (1.67–3.01)	<0.001	1.95 (1.45–2.63)	<0.001	2.92 (2.09–4.07)	<0.0001	2.47 (1.76–3.46)	<0.001	3.48 (2.30–5.28)	<0.0001	2.86 (1.88–4.36)	<0.001
*DRB1*0401/x* †			1.26 (0.95–1.68)	0.11			1.24 (0.88–1.74)	0.22			1.17 (0.76–1.79)	0.47
Genetic risk score ‡			1.50 (1.33–1.68)	<0.001			1.63 (1.41–1.88)	<0.001			1.56 (1.29–1.87)	<0.001
BTNL2 rs3763305 GG §			1.09 (0.79–1.51)	0.60			1.43 (0.95–2.17)	0.090			1.92 (1.11–3.35)	0.021
ITGA1 rs7735139 GG §			1.26 (0.88–1.82)	0.20			1.14 (0.73–1.80)	0.56			1.25 (0.72–2.18)	0.43

* Model 1 and 2 are adjusted for sex, country (reference: US);
† reference: *DRB1* without *0401* or *0401/0404* and *0401/0407*; ‡ per unit increase; § reference: GA/AA genotype

Table 16. BTNL2 SNP genotype frequencies in relation to the development of islet autoantibodies and diabetes among TEDDY children with *HLA DR3/DR4-DQ8* or *DR4-DQ8/DR4-DQ8* and available genotype information.

	Islet autoantibody negative	One or more islet autoantibodies	Multiple islet autoantibodies	Diabetes
HLA *DR3/4-DQ8* (n=3,024)				
BTNL2 SNP rs3763305				
GG	1,839 (69.1%)	272 (75.3%)	190 (79.8%)	120 (82.8%)
GA	823 (30.9%)	89 (24.7%)	48 (20.2%)	25 (17.2%)
AA	1 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
P-value *		0.048	0.002	0.002
HLA *DR4-DQ8/DR4-DQ8* (n=1,532)				
BTNL2 SNP rs3763305				
GG	692 (49.7%)	89 (64.0%)	63 (73.3%)	36 (76.6%)
GA	571 (41.0%)	44 (31.7%)	21 (24.4%)	10 (21.3%)
AA	130 (9.3%)	6 (4.3%)	2 (2.3%)	1 (2.1%)
P-value *		0.003	<0.001	0.001

* Children who remained islet autoantibody negative were compared with children who developed one or more islet autoantibodies, multiple islet autoantibodies and diabetes using Fisher's test.

Table 17. Genotype frequencies of *BTNL2* SNP rs3765503 in validation cohort.

	BTNL2 rs3765503 genotype			
	GG	**AG**	**AA**	***P*-value***
Control	487 (71%)	200 (29%)	0 (0%)	<0.001
Diabetes	3496 (81%)	842 (19%)	0 (0%)	

* *P*-value was calculated using Fisher's exact test.

The validation cohort consists of 5,025 Caucasian subjects with European decent and HLA *DR3/4-DQ8* genotype according to the algorithm defined by Barker et al. (148), which is based on the tag SNPs rs7454108 and rs2040410.

Samples were genotyped on the Illumina ImmunoChip array and imputed to the TOPMed Reference Panel. rs3763305 was directly genotyped on the ImmunoChip. rs2040410, and rs7454108 were imputed with high confidence (R^2>0.99). Principal components were generated by calculating PC axes in unrelated controls using a set of 83,458 LD-pruned variants and projecting the remaining samples onto this PC space. A set of 33,249 European ancestry unrelated case-control subjects were identified for analysis by comparing PCs to 1000 Genomes Phase 3 subjects (149). We ensured samples were unrelated (less than second degree relationship) using KING version 2.13 (http://people.virginia.edu/~wc9c/KING/). 5,025 of these samples have the HLA *DR3/4-DQ8* genotype.

Table 18. Haplotypes of 34 SNPs in *BTNL2* and their frequencies in HLA *DR3/4-DQ8* and *DR4-DQ8/DR4-DQ8*. SNPs in bold are those that were identified as enriched within children with a first-degree relative with T1D. Haplotype ID 'X' indicates that they are not among the 30 most frequent haplotypes.

frequency [%] in																																				
DR3/4-DQ8	***DR4-DQ8/ DR4-DQ8***	**Haplotype ID**	rs3817969	rs3129953	rs28362678	rs2076530	rs9268480	rs2076529	rs9268481	rs3129954	rs3129955	rs4248166	rs2294884	rs2294883	rs2294882	rs2294881	rs2294880	rs9268482	rs2294878	rs3817966	rs3817964	rs3817963	rs3817962	rs3763305	rs2076525	rs2076523	rs2076522	rs3793127	rs10947260	rs10947261	rs10947262	rs3806155	rs3806156	rs3806157	rs2395158	rs3763307
42.82	0.26	9	G	A	G	A	G	A	A	G	G	A	A	T	A	A	A	A	C	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	G	T
18.12	18.70	25	G	G	G	G	A	G	G	G	G	A	A	T	A	A	G	T	A	G	A	G	A	G	G	G	G	A	A	C	G	A	A	C	A	A
15.14	14.62	28	G	G	G	G	A	G	G	G	G	A	A	T	A	A	G	T	A	G	T	G	A	A	G	G	G	A	A	C	G	T	A	C	A	A
12.48	12.29	26	G	G	G	G	A	G	G	G	G	A	A	T	A	A	G	T	A	G	A	G	A	G	G	G	G	A	A	C	G	T	A	C	A	A
9.86	3.56	12	G	G	G	A	G	A	A	A	A	A	A	T	A	A	A	A	C	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	A	T
0.60	0.09	27	G	G	G	G	A	G	G	G	G	A	A	T	A	A	G	T	A	G	A	G	A	G	G	G	G	G	A	C	G	T	A	C	A	A
0.31	0.27	16	G	G	G	A	G	A	A	G	G	A	A	T	A	A	A	A	C	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	A	T
0.26	0.07	5	A	G	A	G	G	G	A	G	G	G	C	A	G	G	A	A	A	A	A	A	C	G	A	G	C	G	G	A	A	T	A	C	A	T
0.15	0.10	21	G	G	G	A	G	A	A	G	G	G	C	A	G	G	A	A	A	A	A	A	C	G	A	G	C	G	G	A	A	T	A	C	A	T
0.09	0.03	4	A	G	A	G	G	G	A	G	G	G	C	A	G	G	A	A	A	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	A	T
0.05	0.03	23	G	G	G	G	A	G	G	G	G	A	A	T	A	A	A	T	A	G	A	G	A	G	G	G	G	A	A	C	G	T	A	C	A	A
0.00	0.07	3	A	G	A	G	G	G	A	G	G	A	A	A	G	G	G	A	C	G	A	G	C	G	A	G	C	G	A	A	A	T	A	A	A	T
0.03	0.00	8	G	A	G	A	G	A	A	G	G	A	A	T	A	A	A	A	C	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	A	T
0.03	0.00	13	G	G	G	A	G	A	A	A	A	A	A	T	G	G	A	A	A	A	A	A	C	G	A	G	C	G	G	A	A	T	A	C	A	T
0.02	0.00	17	G	G	G	A	G	A	A	G	G	A	A	T	A	A	A	A	C	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	G	T
0.00	0.00	1	A	A	A	G	G	G	A	G	G	G	C	A	G	G	A	A	A	A	A	A	C	G	A	G	C	G	G	A	A	T	A	C	A	T

frequency [%] in																																						
DR3/4-DQ8	***DR4-DQ8/ DR4-DQ8***	**Haplotype ID**	rs3817969	rs3129953	rs28362678	rs2076530	rs9268480	rs2076529	rs9268481	rs3129954	rs3129955	rs4248166	rs2294884	rs2294883	rs2294882	rs2294881	rs2294880	rs9268482	rs2294878	rs3817966	rs3817964	rs3817963	rs3817962	rs3763305	rs2076525	rs2076523	rs2076522	rs3793127	rs10947260	rs10947261	rs10947262	rs3806155	rs3806156	rs3806157	rs2395158	rs3763307
0.02	0.00	6	A	G	G	G	G	G	A	G	G	G	A	A	G	G	G	A	C	G	A	A	A	G	A	G	C	G	G	C	A	T	A	A	A	T
0.00	0.02	10	G	A	G	A	G	A	A	G	G	A	A	T	G	G	A	A	A	A	T	A	C	A	A	G	C	G	G	A	A	T	A	C	A	T
0.02	0.00	11	G	A	G	G	A	G	G	G	G	A	A	T	A	A	G	T	A	G	T	G	A	A	G	G	G	A	A	C	G	T	A	C	A	A
0.00	0.00	14	G	G	G	A	G	A	A	A	G	A	A	T	A	A	A	A	C	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	G	T
0.00	0.00	15	G	G	G	A	G	A	A	G	A	A	A	T	A	A	A	A	C	A	A	A	C	G	A	A	C	G	A	C	G	T	C	A	A	T
0.00	0.02	24	G	G	G	G	A	G	G	G	G	A	A	T	A	A	G	T	A	G	A	G	A	G	G	G	G	A	A	C	A	T	A	C	A	A
0.00	0.00	x	G	A	G	A	G	A	A	G	G	A	A	T	G	G	A	A	A	A	A	A	C	G	A	G	C	G	G	A	A	T	A	C	A	T
0.00	0.00	x	G	G	G	A	G	A	A	G	G	G	C	A	G	G	A	A	A	A	T	A	C	A	A	G	C	A	G	A	A	T	A	C	A	T

Table 19. Relationship between *BTNL2* SNP rs3765503 genotypes and HLA *DR3* and *DRB1*04* subtypes in children with the *HLA DR3/DR4-DQ8* genotype.

HLA *DR* genotype	*BTNL2* rs3763305 genotype			Allele frequency (%)	
	AA	GA	GG	A	G
DR3/DR3	2 (0.1%)	9 (0.6%)	1608 (99.3%)	0.4	99.6
*DR3/DRB1*04:01*	0 (0.0%)	15 (1.0%)	1496 (99.0%)	0.5	99.5
*DR3/DRB1*04:02*	0 (0.0%)	2 (1.2%)	165 (98.8%)	0.6	99.4
*DR3/DRB1*04:04*	1 (0.1%)	837 (78.1%)	233 (21.8%)	39.2	60.8
*DR3/DRB1*04:05*	0 (0.0%)	0 (0.0%)	134 (100.0%)	0.0	100.0
*DR3/DRB1*04:06*	0 (0.0%)	1 (100.0%)	0 (0.0%)	50.0	50.0
*DR3/DRB1*04:07*	0 (0.0%)	33 (68.8%)	15 (31.2%)	34.4	65.6
*DR3/DRB1*04:08*	0 (0.0%)	1 (4.8%)	20 (95.2%)	2.4	97.6
*DR3/DRB1*04:10*	0 (0.0%)	0 (0.0%)	1 (100.0%)	0.0	100.0

Table 20. Validation of the association of the *BTNL2* rs3763305 genotype with HLA *DRB1*04* subtype alleles in children from DiMelli.

*DRB1*04* subtype	*BTNL2* rs3763305 genotype		
	AA	AG	GG
*DR3/DRB1*04:01*	0 (0.0%)	0 (0.0%)	89 (100.0%)
*DR3/DRB1*04:02*	0 (0.0%)	0 (0.0%)	21 (100.0%)
*DR3/DRB1*04:04*	0 (0.0%)	18 (75.0%)	6 (25.0%)
*DR3/DRB1*04:05*	0 (0.0%)	0 (0.0%)	7 (100.0%)

Table 21. Cox proportional hazards models for developing one or more islet autoantibodies, multiple islet autoantibodies and diabetes according to *BTNL2* rs3763305 in children with the HLA *DR3/DRB1*0404-DQ8* or HLA *DRB1*0404-DQ8/DRB1*0404-DQ8* genotypes adjusted for the HLA *DR3/DR4-DQ8* genotype.

Variable	One or more islet autoantibodies		Multiple islet autoantibodies		Diabetes	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
HLA *DR3/DR4-DQ8* *	1.48 (0.81–2.70)	0.20	1.64 (0.70–3.82)	0.25	1.26 (0.44–3.59)	0.66
BTNL2 rs3763305 GG †	1.13 (0.74–1.74)	0.57	1.49 (0.88–2.54)	0.14	2.38 (1.25–4.55)	0.009

* reference: *DR4-DQ8/DR4-DQ8*; † reference: GA/AA genotype

Table 22. Cumulative risk of developing islet autoantibodies or diabetes by 6 years old in children stratified by genetic risk score and HLA DRB1*04 subtype.

	One or more islet autoantibodies		Multiple islet autoantibodies		Diabetes	
Risk strata	**children from the general population**	**children with a first-degree relative with T1D**	**children from the general population**	**children with a first-degree relative with T1D**	**children from the general population**	**children with a first-degree relative with T1D**
1: High-risk *DR4* *AND 1st quartile GRS	17.4% (13.6–21.2%)	16.7% (7.9–24.6%)	12.7% (9.3–16.0%)	14.3% (6.1–21.9%)	4.1% (2.2–5.9%)	4.8% (0.1–9.2%)
2: High-risk *DR4* AND 2nd quartile GRS OR low-risk *DR4* * AND 1st quartile GRS	11.9% (9.6–14.1%)	23.4% (14.7–31.2%)	9.0% (7.0–10.9%)	17.0% (9.3–24.1%)	4.1% (2.8–5.3%)	9.1% (3.5–14.3%)
3: High-risk *DR4* AND GRS <50th centile OR low-risk *DR4* AND 2nd quartile GRS	8.2% (6.7–9.7%)	19.3% (12.2–25.9%)	4.5% (3.3–5.7%)	14.8% (8.4–20.7%)	1.5% (0.9–2.2%)	7.1% (2.8–11.2%)
4: Low-risk *DR4* AND GRS <50th centile	4.3% (2.9–5.5%)	11.1% (2.3–19.1%)	1.6% (0.8–2.4%)	9.2% (1.2–16.6%)	0.5% (0.1–0.9%)	1.6% (0.0–4.8%)
P-value †	<0.001	0.45	<0.001	0.60	<0.001	0.18

* High-risk *DR4* was defined as *DR3/DRB1*0401* or *DRB1*0401-DQ8/DR4* without *0404 or *0407, and low risk were all other genotypes;
† *P*-values were calculated as log-rank tests per column over the four strata. GRS, genetic risk score

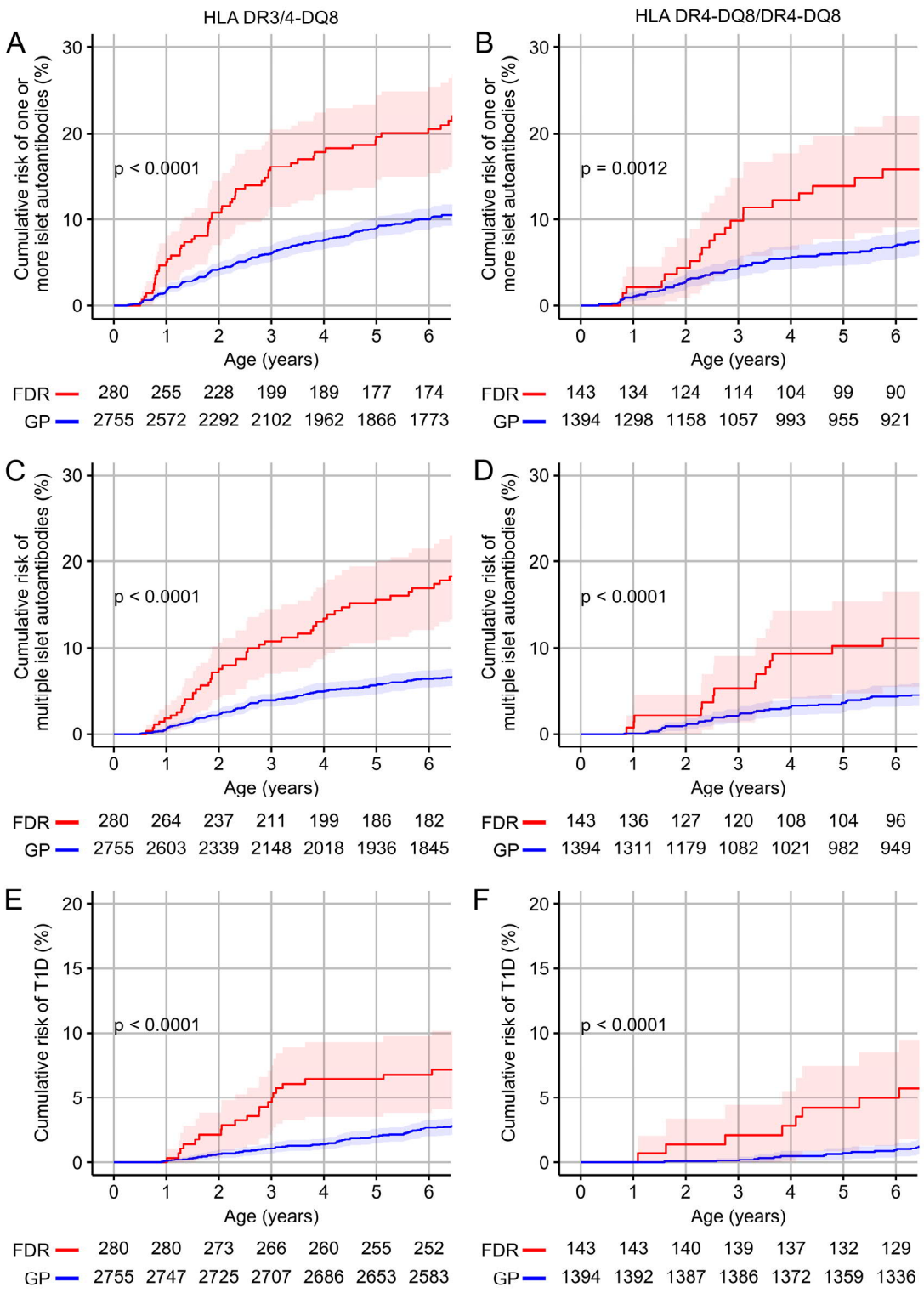


Figure 9. Cumulative risks of islet autoantibodies and diabetes in TEDDY children with HLA *DR3/DR4-DQ8* or HLA *DR4-DQ8/DR4-DQ8* genotypes. Kaplan–Meier curves for the risk of one or more islet autoantibodies **(A, B)**, multiple islet autoantibodies **(C, D)**, and diabetes **(E, F)** in children with a first-degree relative with T1D (red) and in children from the general population (blue), stratified into children with the HLA *DR3/DR4-DQ8* **(A, C, E)** or HLA *DR4-DQ8/DR4-DQ8* **(B, D, F)** genotypes. Shaded areas represent the 95% CI. Numbers represent children at risk. *P*-values were calculated using log-rank tests.

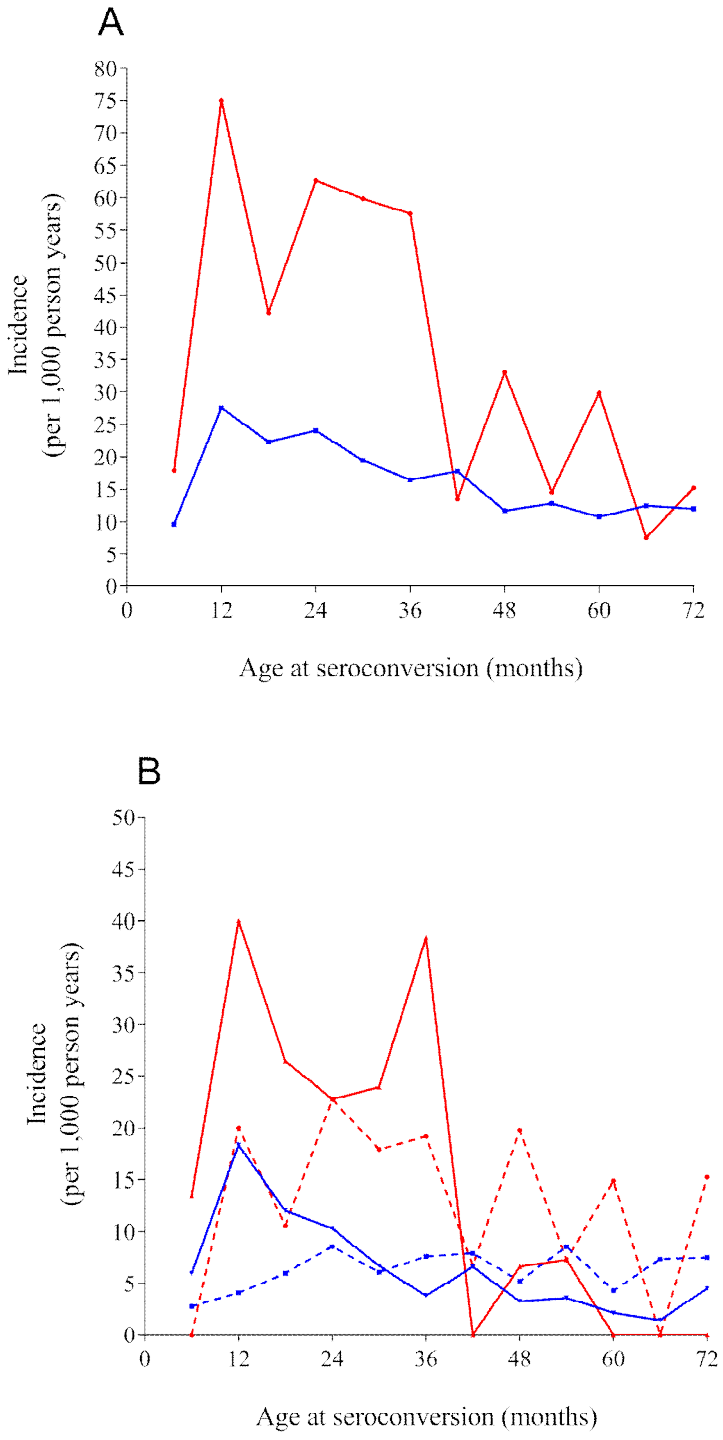


Figure 10. Incidence of one or more islet autoantibodies. Incidences **(A)** among *DR3/4-DQ8* and *DR4-DQ8/DR4-DQ8* children with a first-degree relative with T1D (red) compared with children from the general population (blue) by the age of seroconversion. **(B)** Incidence of first-appearing IAA (solid lines) and first-appearing GADA (broken lines) at seroconversion in *DR3/4-DQ8* and *DR4-DQ8/DR4-DQ8* children with a first-degree relative with T1D (red) compared with children from the general population (blue) by the age of seroconversion.

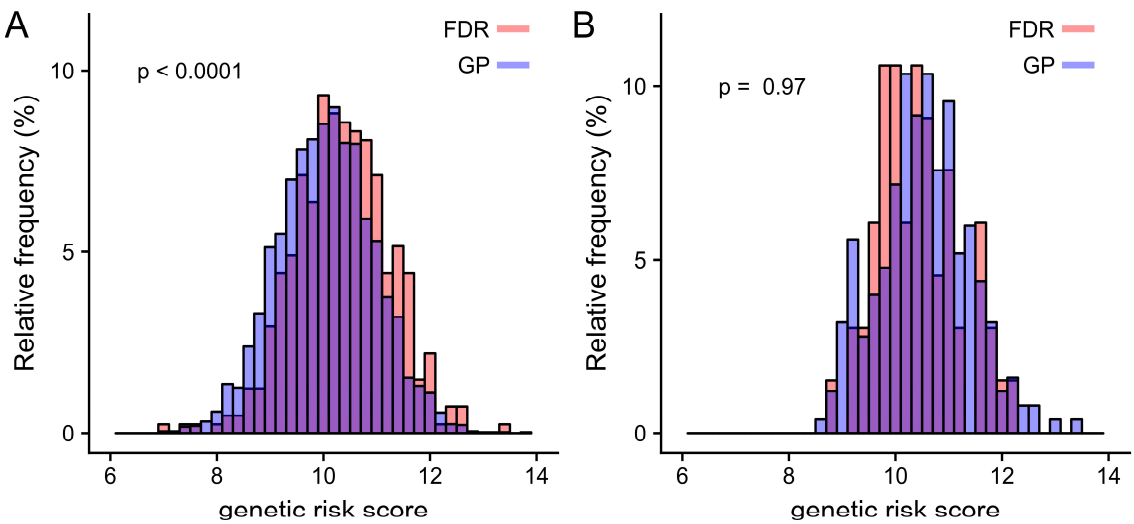


Figure 11. Distribution of non-HLA *DR-DQ* genetic risk scores. In **(A)** all 4,414 *DR3/DR4-DQ8* or *DR4-DQ8/DR4-DQ8* children stratified into children with a first-degree relative with T1D (red) and children from the general population (blue). **(B)** Distribution of non-HLA *DR-DQ* genetic risk scores in 317 *DR3/DR4-DQ8* or *DR4-DQ8/DR4* children who developed multiple islet autoantibodies (children with a first-degree relative with T1D, red; children from the general population, blue). *P*-values were calculated using the two-sided Mann–Whitney U test.

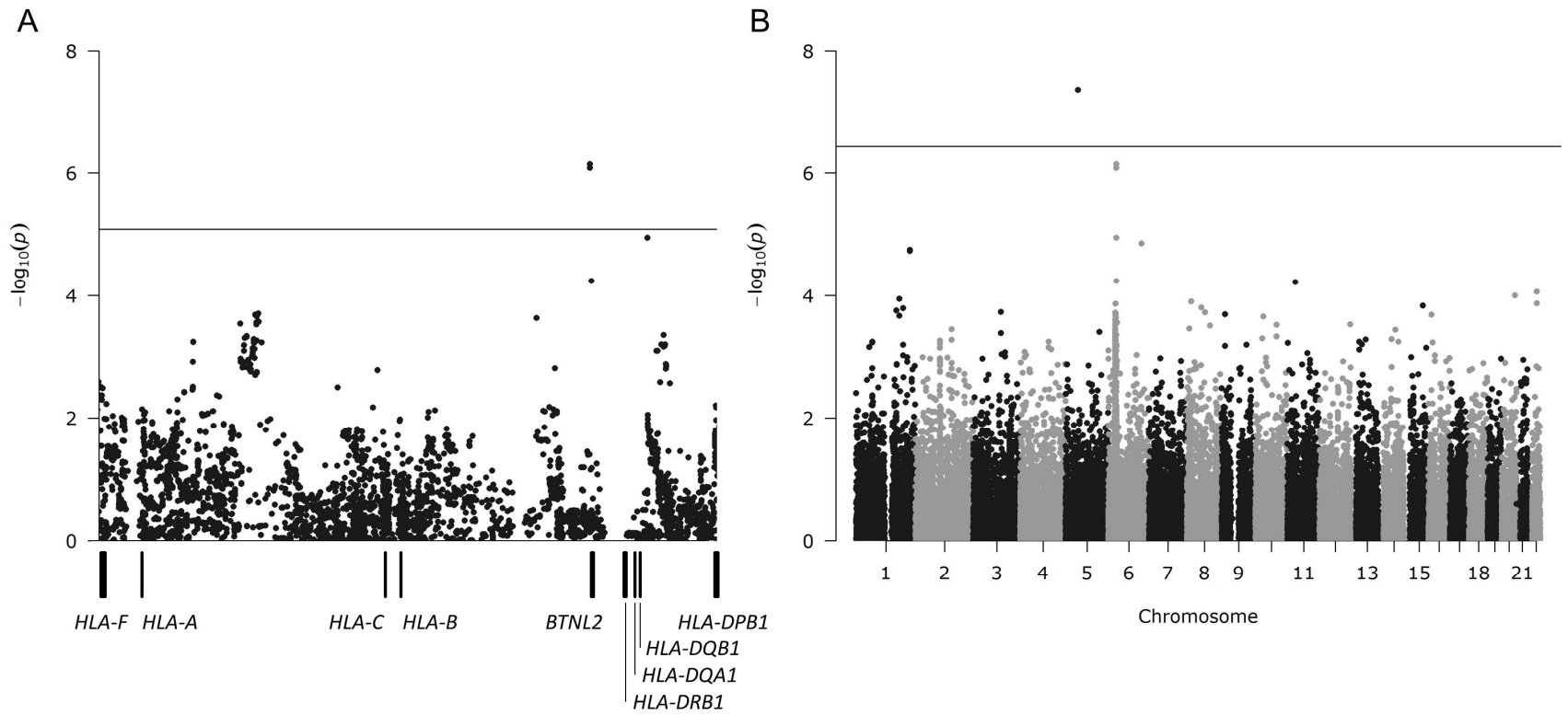


Figure 12. Manhattan plot of allele enrichment in children with a first-degree relative with T1D. SNPs were analyzed across **(A)** the HLA region on chromosome 6 and across **(B)** all ImmunoChip data.

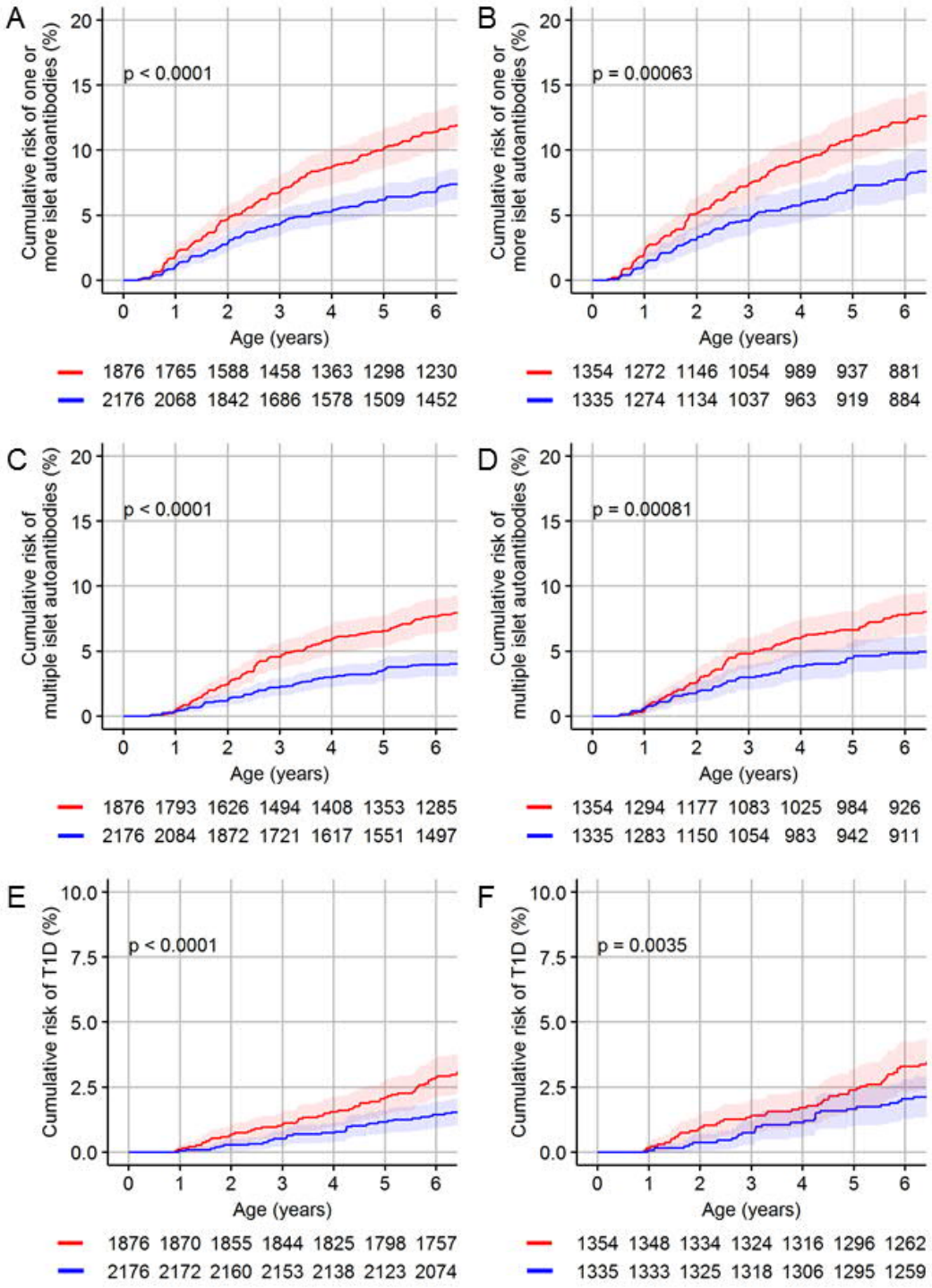


Figure 13. Risk of developing one or more islet autoantibodies, multiple islet autoantibodies and diabetes with respect to *DRB1*04:01*. Risk of developing one or more islet autoantibodies **(A, B)**, multiple islet autoantibodies **(C, D)** and diabetes **(E, F)** in children from the general population with HLA *DR3/DRB1*04:01-DQ8* or *DRB1*04:01-DQ8/DRB1*04:xx-DQ8*, where *04:xx was any allele other than DRB1*04:04 or *DRB1*04:07* (red) vs children without *DRB1*04:01* (blue). The risks are also shown separately for children from the general population with HLA *DR3/DR4*04:01-DQ8* **(B, D, F)**. *P*-values were calculated using log-rank tests.

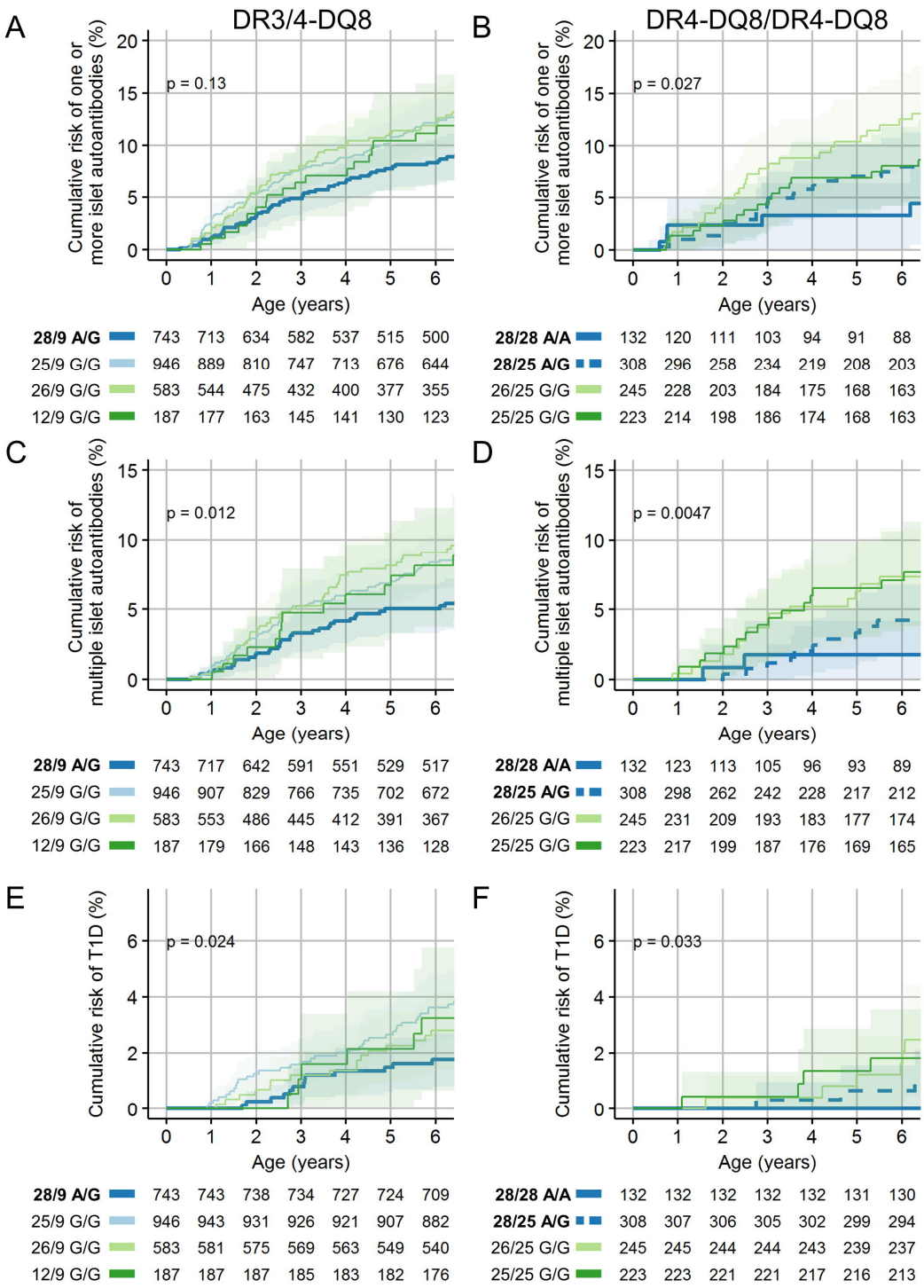


Figure 14. Kaplan–Meier curves for the risk of one or more islet autoantibodies, multiple islet autoantibodies, and diabetes according to *BTNL2* haplotypes. Kaplan–Meier curves for the risk of one or more islet autoantibodies **(A, B)**, multiple islet autoantibodies **(C, D)**, and diabetes **(E, F)** in children from the general population stratified into children with the HLA *DR3/DR4-DQ8* **(A, C, E)** or HLA *DR4-DQ8/DR4-DQ8* **(B, D, F)** genotypes and according to *BTNL2* haplotypes. For both HLA genotypes, the 4 major *BTNL2* genotypes are shown. The genotypes that include haplotype 28, which is the only *BTNL2* haplotype that has the SNP rs3763305 A allele, are indicated as thick blue lines. Shaded areas represent the 95% CI. Numbers represent children at risk. *P*-values were calculated using log-rank tests.

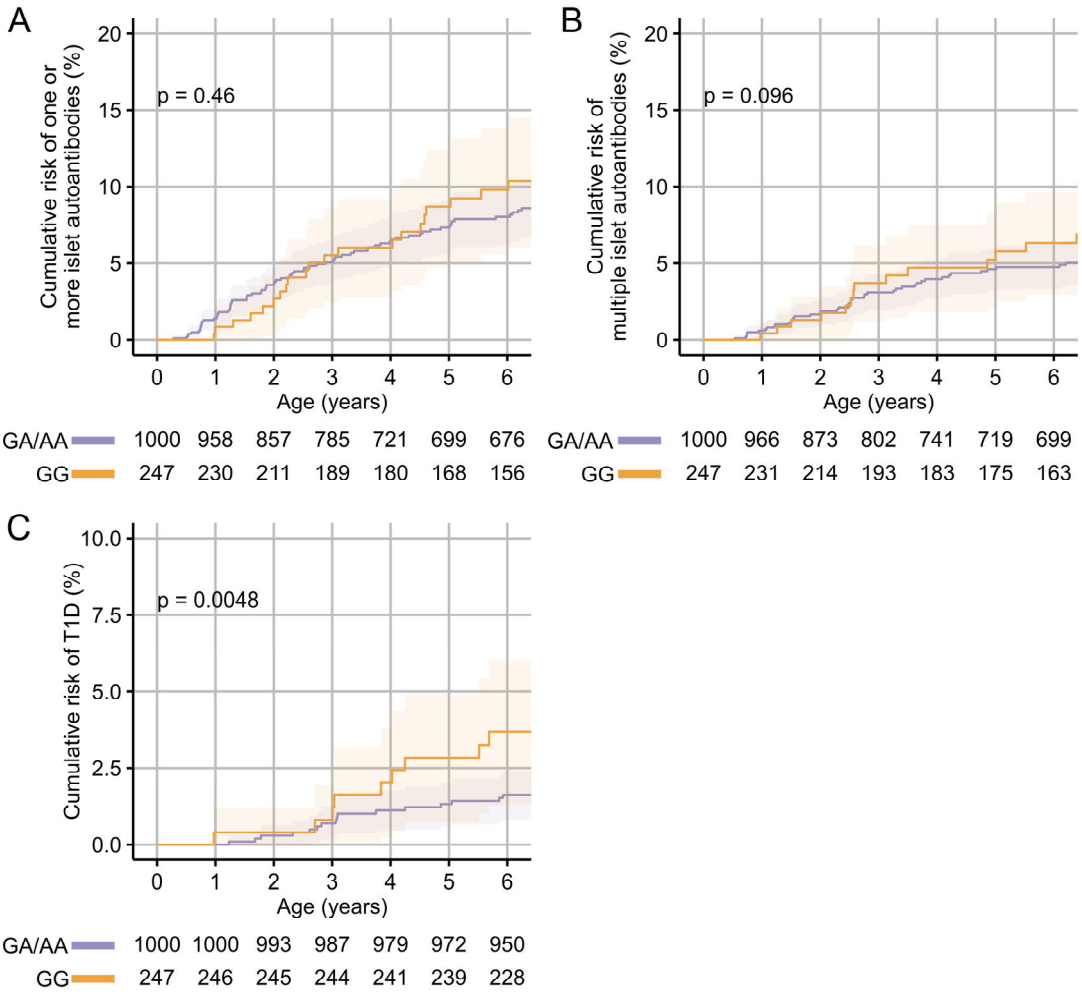


Figure 15. Modification of risk by the *BTNL2* SNP rs3763305 on the development of islet autoantibodies and diabetes.The modification of risk by the *BTNL2* SNP rs3763305 on the development of one or more islet autoantibodies **(A)**, multiple islet autoantibodies **(B)** and diabetes **(C)** in children with the *DR3/DRB1*04:04-DQ8* or *DRB1*04:04-DQ8/DRB1*04:04-DQ8* genotypes.Risks are shown for the GG genotype (orange-brown) versus the GA or AA genotypes (purple) at rs3763305. *P*-values were calculated using log-rank tests.

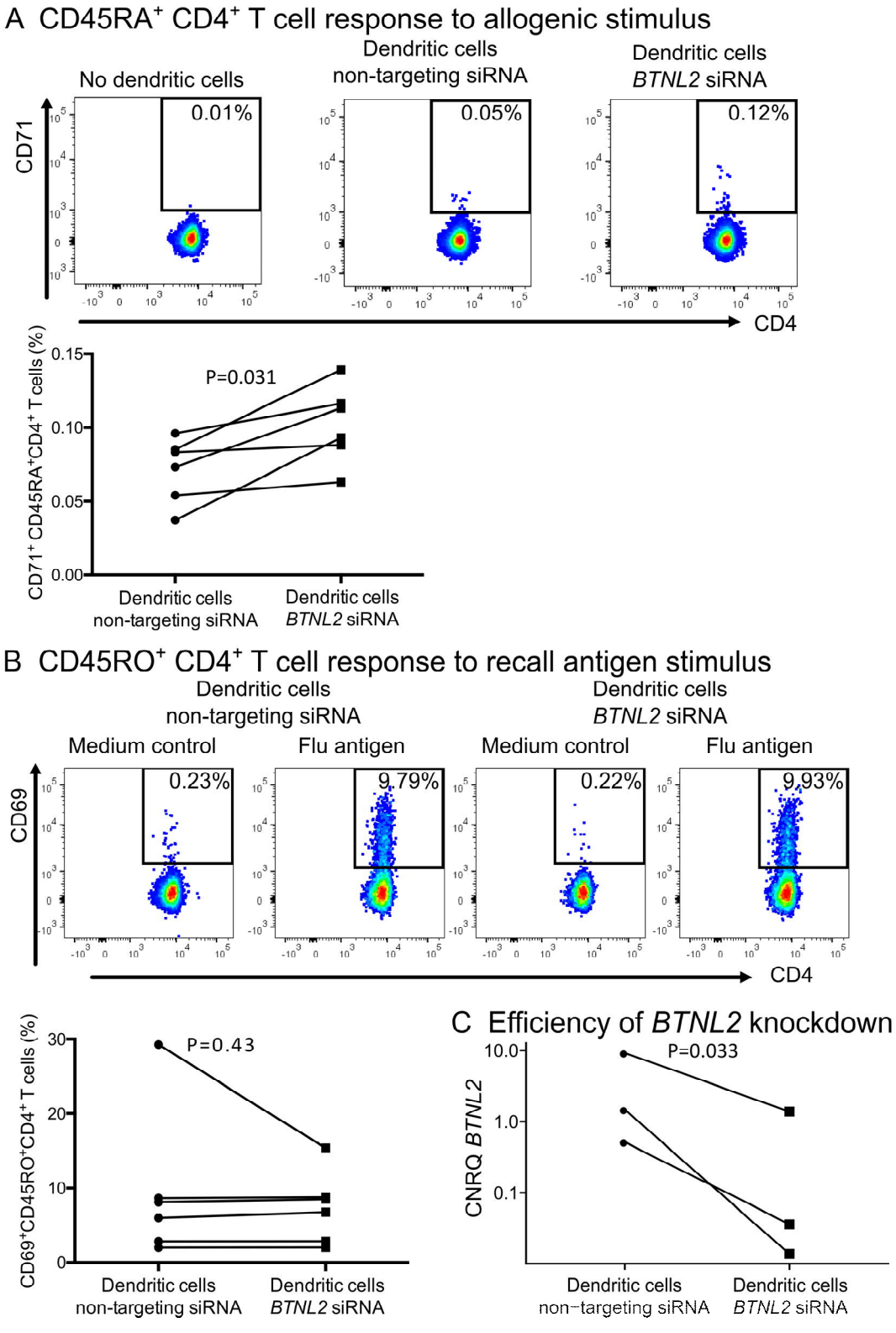

A CD45RA+ CD4+ T cell response to allogenic stimulus
No dendritic cells
Dendritic cells non-targeting siRNA
Dendritic cells BTNL2 siRNA
CD71
0.01%
0.05%
0.12%
CD4
CD71+ CD45RA+CD4+ T cells (%)
P=0.031
Dendritic cells non-targeting siRNA
Dendritic cells BTNL2 siRNA
B CD45RO+ CD4+ T cell response to recall antigen stimulus
Dendritic cells non-targeting siRNA
Dendritic cells BTNL2 siRNA
Medium control
Flu antigen
Medium control
Flu antigen
CD69
0.23%
9.79%
0.22%
9.93%
CD4
CD69+CD45RO+CD4+ T cells (%)
P=0.43
Dendritic cells non-targeting siRNA
Dendritic cells BTNL2 siRNA
C Efficiency of BTNL2 knockdown
CNRQ BTNL2
P=0.033
Dendritic cells non-targeting siRNA
Dendritic cells BTNL2 siRNA

Figure 16. Effect of *BTNL2* knockdown in monocyte-derived dendritic cells on $CD4^+$ T cell activation. (A) Activation of isolated $CD4^+$ $CD25^-$ T cells by allogeneic monocyte-derived dendritic cells transfected with non-targeting siRNA or *BTNL2* targeting siRNA in mixed lymphocyte cultures. CD71 expression, previously reported as a marker of allo-reactive T cell activation, was used to measure the activation of the $CD4^+$ T cells. Dendritic cells were transfected as previously described (132). The upper panels are exemplary FACS plots for 42 hour cultures of $CD4^+$ T cells only and in the presence of allo-reactive transfected dendritic cells. The lower graph indicates the mean frequency $CD71^+$ $CD45RA^+$ $CD4^+$ T cells at the end of the 42 hour culture (4 replicates) after subtraction of the mean frequency of $CD71^+$ $CD45RA^+$ $CD4^+$ T cells in quadruplicate cultures without dendritic cells. Each of three dendritic cell samples was tested against two different allogeneic $CD4^+$ T cell preparation yielding 6 data sets. Activation was increased when $CD4^+$ T cells were activated with dendritic cells transfected with *BTNL2*-targeting siRNA as compared to non-targeting siRNA (p=0.031, Wilcoxon matched pair sign test). **(B)** Activation of isolated $CD4^+$ $CD25^-$ T cells by autologous monocyte-derived dendritic cells transfected with non-targeting siRNA or *BTNL2*-targeting siRNA in the presence of flu or tetanus toxoid antigen. CD69 expression was used to measure the activation of the $CD4^+$ T cells. The upper panels are exemplary FACS plots for 42 hour cultures of $CD4^+$ T cells plus dendritic cells in the presence and absence of flu antigen. The lower graph indicates the mean frequency $CD69^+$ $CD4^+$ T cells at the end of the 42 hour culture (triplicates) after subtraction of the mean frequency of $CD69^+$ $CD4^+$ T cells in triplicates cultures without antigen. Each of three dendritic cell samples was tested against flu and tetanus toxoid yielding 6 data sets. Activation was not different when $CD4^+$ T cells were activated with dendritic cells transfected with *BTNL2*-targeting siRNA as compared to non-targeting siRNA (p=0.43, Wilcoxon matched pair sign test). **(C)** Efficiency of knockdown with *BTNL2* vs non-targeting siRNA in dendritic cells used for **(A)** and **(B)** *BTNL2* gene expression was normalized to reference genes *TELO2* and *TRMT61A* and the Calibrated Normalized Relative Quantities (CNRQ) relative to the treatment with non-targeting siRNA is shown (p=0.033).

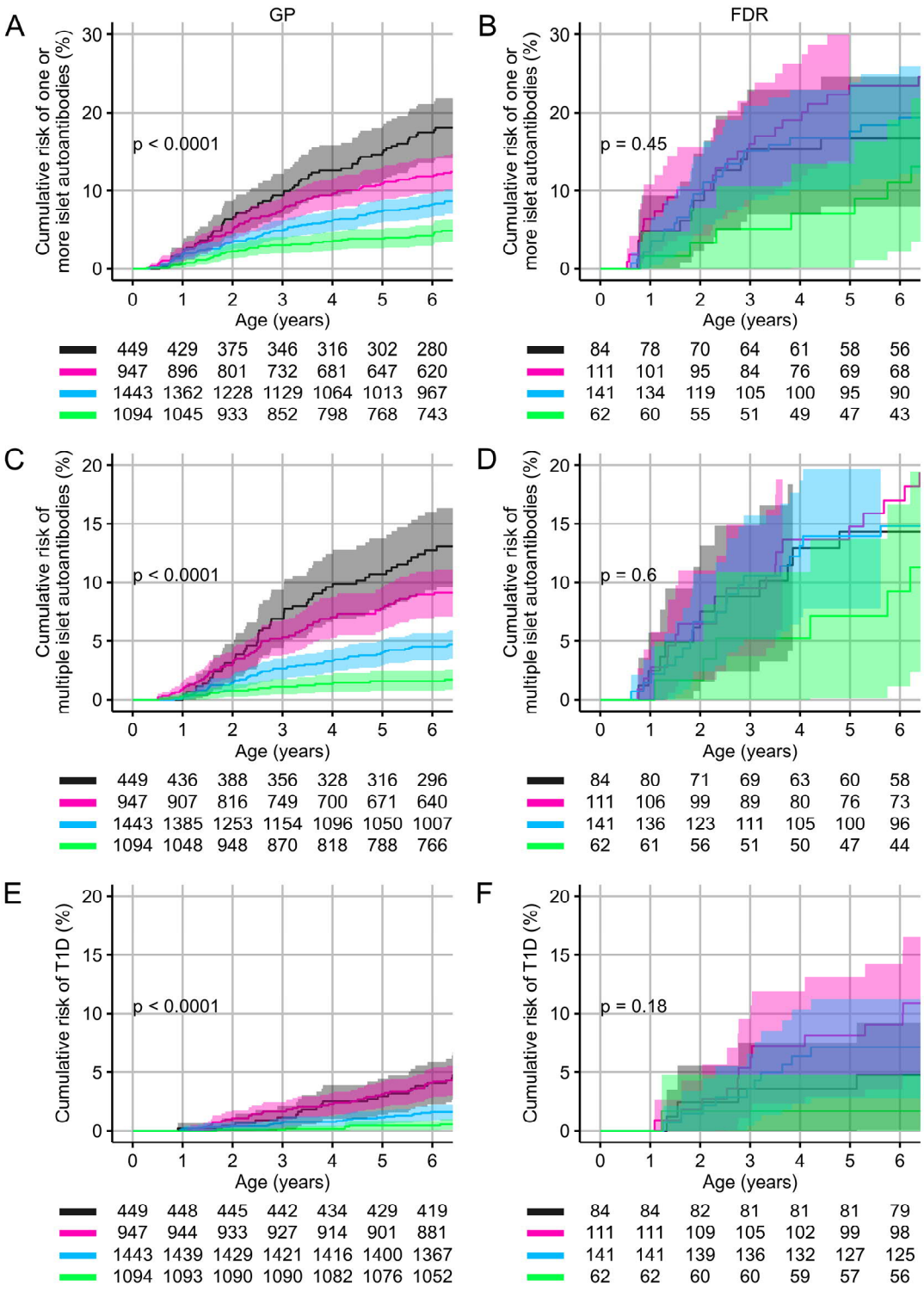

A
GP
Cumulative risk of one or more islet autoantibodies (%)
p < 0.0001
Age (years)
449 429 375 346 316 302 280
947 896 801 732 681 647 620
1443 1362 1228 1129 1064 1013 967
1094 1045 933 852 798 768 743
B
FDR
Cumulative risk of one or more islet autoantibodies (%)
p = 0.45
Age (years)
84 78 70 64 61 58 56
111 101 95 84 76 69 68
141 134 119 105 100 95 90
62 60 55 51 49 47 43
C
Cumulative risk of multiple islet autoantibodies (%)
p < 0.0001
Age (years)
449 436 388 356 328 316 296
947 907 816 749 700 671 640
1443 1385 1253 1154 1096 1050 1007
1094 1048 948 870 818 788 766
D
Cumulative risk of multiple islet autoantibodies (%)
p = 0.6
Age (years)
84 80 71 69 63 60 58
111 106 99 89 80 76 73
141 136 123 111 105 100 96
62 61 56 51 50 47 44
E
Cumulative risk of T1D (%)
p < 0.0001
Age (years)
449 448 445 442 434 429 419
947 944 933 927 914 901 881
1443 1439 1429 1421 1416 1400 1367
1094 1093 1090 1090 1082 1076 1052
F
Cumulative risk of T1D (%)
p = 0.18
Age (years)
84 84 82 81 81 81 79
111 111 109 105 102 99 98
141 141 139 136 132 127 125
62 62 60 60 59 57 56

Figure 17. Risk of developing islet autoantibodies and diabetes according to genetic susceptibility strata based on HLA *DRB1*04* subtype and genetic risk score. Risk of developing islet autoantibodies and diabetes in children with a first-degree relative with T1D **(B, D, F)** and in children from the general population **(A, B, C)** according to genetic susceptibility strata based on HLA *DRB1*04* subtype and genetic risk score (GRS). Risks are shown for the development of one or more islet autoantibodies **(A, B)**, multiple islet autoantibodies **(C, D)**, and diabetes **(E, F)**. All of the children had the *DR3/DR4-DQ8* or *DR4-DQ8/DR4-DQ8* genotype. Genetic susceptibility strata were defined as follows: **1.** high-risk *DRB1*04* subtype (*DR3/DRB1*0401* or *DRB1*0401-DQ8/DR4* without 0404 or 0407) AND GRS in the upper quartile (grey); **2.** high-risk *DRB1*04* subtype AND GRS in the second quartile, OR lower-risk *DRB1*04* subtype AND GRS in the upper quartile (pink); **3.** high-risk *DRB1*04* subtype AND GRS in the lower 50th centile OR lower-risk *DRB1*04* subtype AND GRS in the second quartile (light blue); and **4.** lower-risk *DRB1*04* subtype AND GRS in the lower 50th centile (green). The strata appear in this order from top to bottom in the risk tables. P-values were calculated across all strata using log-rank tests.

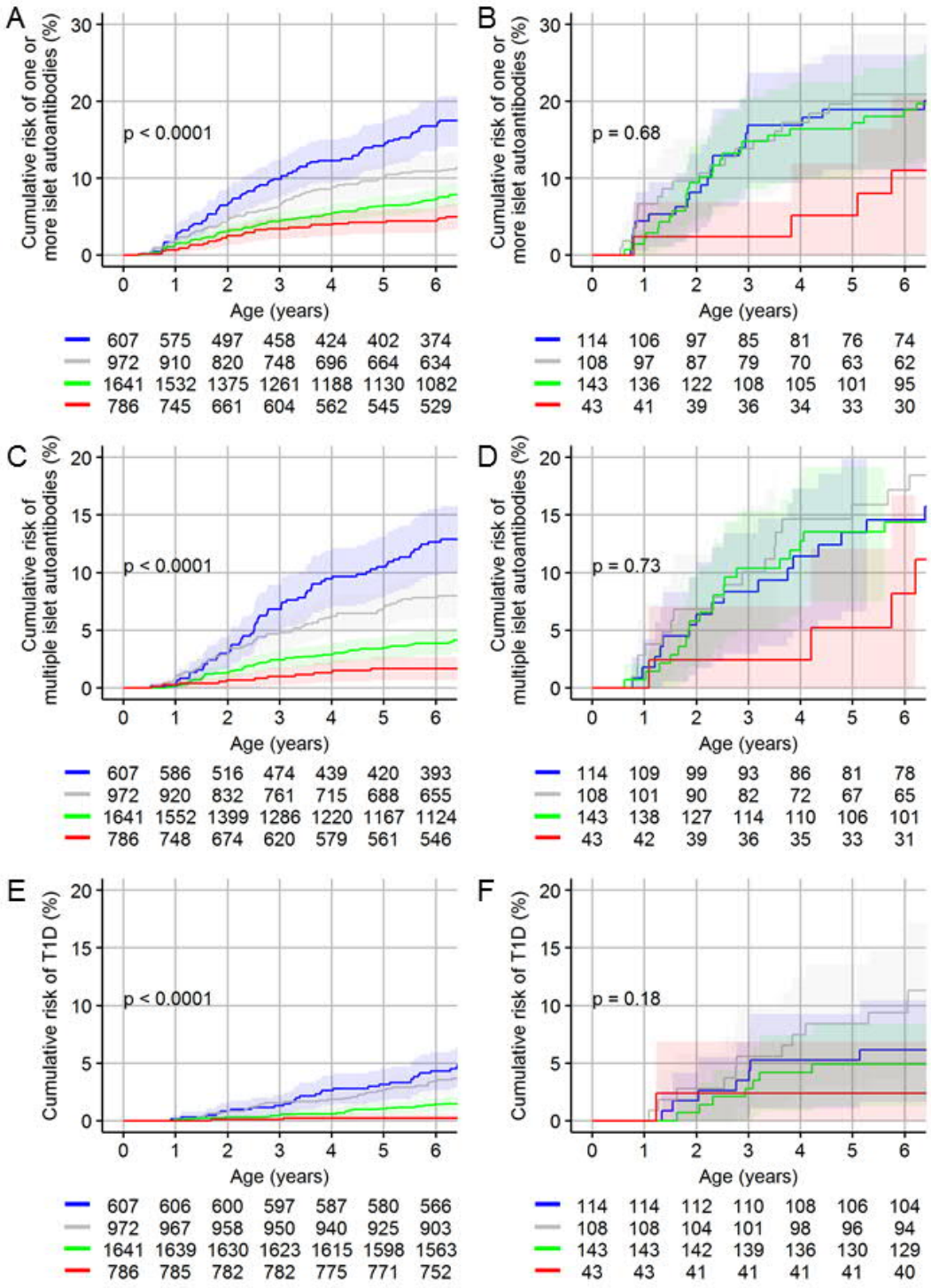

A
Cumulative risk of one or more islet autoantibodies (%)
p < 0.0001
Age (years)
607 575 497 458 424 402 374
972 910 820 748 696 664 634
1641 1532 1375 1261 1188 1130 1082
786 745 661 604 562 545 529
B
Cumulative risk of one or more islet autoantibodies (%)
p = 0.68
Age (years)
114 106 97 85 81 76 74
108 97 87 79 70 63 62
143 136 122 108 105 101 95
43 41 39 36 34 33 30
C
Cumulative risk of multiple islet autoantibodies (%)
p < 0.0001
Age (years)
607 586 516 474 439 420 393
972 920 832 761 715 688 655
1641 1552 1399 1286 1220 1167 1124
786 748 674 620 579 561 546
D
Cumulative risk of multiple islet autoantibodies (%)
p = 0.73
Age (years)
114 109 99 93 86 81 78
108 101 90 82 72 67 65
143 138 127 114 110 106 101
43 42 39 36 35 33 31
E
Cumulative risk of T1D (%)
p < 0.0001
Age (years)
607 606 600 597 587 580 566
972 967 958 950 940 925 903
1641 1639 1630 1623 1615 1598 1563
786 785 782 782 775 771 752
F
Cumulative risk of T1D (%)
p = 0.18
Age (years)
114 114 112 110 108 106 104
108 108 104 101 98 96 94
143 143 142 139 136 130 129
43 43 41 41 41 41 40

Figure 18. Risk of developing islet autoantibodies and diabetes according to genetic susceptibility strata based on *BTNL2* SNP rs3763305 subtype and genetic risk score. Risk of developing islet autoantibodies and diabetes in children with a first-degree relative with T1D **(B, D, F)** and in children from the general population **(A, B, C)** according to genetic susceptibility strata based on *BTNL2* SNP rs3763305 and genetic risk score (GRS). Risks are shown for the development of one or more islet autoantibodies **(A, B)**, multiple islet autoantibodies **(C, D)**, and diabetes **(E, F)**. All of the children had the *DR3/DR4-DQ8* or *DR4-DQ8/DR4-DQ8* genotype. Genetic susceptibility strata were defined as follows: **1.** rs3763305 GG AND GRS in the upper quartile (blue); **2.** rs3763305 GG AND GRS in the second quartile, OR rs3763305 GA or AA AND GRS in the upper quartile (grey); **3.** rs3763305 GG AND GRS in the lower 50th centile OR rs3763305 GA or AA AND GRS in the second quartile (green); and **4.** rs3763305 GA or AA AND GRS in the lower 50th centile (red). *P*-values were calculated across all strata using log-rank tests.

5.3 Heterogeneity in environmental exposure (TEDDY and BABYDIET studies)

Various environmental factors play an essential role in T1D and are jointly responsible for the rise in disease incidence and contribute to the heterogeneity of the disease (75). Across ethnicities and across study populations, two factors, namely maternal T1D as a protective condition and early life viral infections as risk condition, have repeatedly been shown to influence the incidence of T1D.

Based upon an *a priori* genetic susceptibility to the development of islet autoimmunity and T1D, defined by risk strata, I investigated the influence of the effect of relative protection of maternal T1D compared to paternal or sibling T1D.

Furthermore, I looked into a potential causal link of viral infections in early life by unbiased deep sequencing of viral sequences in longitudinal PBMC samples of at-risk children expressing an antiviral type 1 interferon signature who afterwards developed islet AABs and those who did not.

5.3.1 Risk in children with a first-degree relative with type 1 diabetes is modified by maternal, paternal, and sibling type 1 diabetes (TEDDY)

The analysis of the TEDDY data resulted in risk divergence between children with a first-degree relative with T1D and children from the general population at lower genetic susceptibility strata (Figure 17 and Figure 18). If the first-degree relative index case was a father the HRs (95% CI) for one or more islet AABs (2.37; 1.71–3.28; p<0.001), multiple islet AABs (2.89; 2.00–4.17; p<0.001), and diabetes (3.06; 1.89–4.94; p<0.001), were increased as compared to children from the general population. The HRs were also increased if the first-degree relative was a sibling. By contrast, if the first-degree relative index case was the mother, the HR (95% CI) was not increased for one or more islet AABs (0.85; 0.50–1.44; p=0.54), multiple islet AABs (0.97; 0.52–1.79; p=0.91), and diabetes (1.39; 0.67–2.87; p=0.38; Table 23). This relative protection conferred by maternal T1D versus paternal or sibling T1D was observed across various risk strata (Figure 19) indicating that this protective effect is independent of genetic susceptibility.

The results of this analysis have been published as:

Hippich, M., A. Beyerlein, W. A. Hagopian, J. P. Krischer, K. Vehik, J. Knoop, C. Winker, J. Toppari, A. Lernmark, M. J. Rewers, A. K. Steck, J. X. She, B. Akolkar, C. C. Robertson, S. Onengut-Gumuscu, S. S. Rich, E. Bonifacio and A. G. Ziegler (2019). "Genetic Contribution to the Divergence in Type 1 Diabetes Risk Between Children From the General Population and Children From Affected Families." Diabetes. 2019 Jan 17. pii: db180882. doi: 10.2337/db18-0882. [Epub ahead of print]

Table 23. Hazard ratios (HRs) and 95% CIs for developing islet autoantibodies and diabetes with respect to affected family member. Hazard ratios (HRs) and 95% CIs for developing one or more islet autoantibodies, multiple islet autoantibodies and diabetes in children with a first-degree relative with T1D (first-degree relative with T1D mother vs father vs sibling vs multiplex), adjusted for sex, genetic factors (reference: *DR4-DQ8*), and country (reference US).

	One or more islet autoantibodies		Multiple islet autoantibodies		Diabetes	
	HR (95% CI)	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**
first-degree relative with T1D multiplex	1.94 (0.62–6.12)	0.26	3.37 (1.06–10.66)	0.039	5.90 (1.84–18.94)	0.003
first-degree relative with T1D mother	0.85 (0.50–1.44)	0.54	0.97 (0.52–1.79)	0.91	1.39 (0.67–2.87)	0.38
first-degree relative with T1D father	2.37 (1.71–3.28)	<0.001	2.89 (2.00–4.17)	<0.001	3.06 (1.89–4.94)	<0.001
first-degree relative with T1D sibling	2.68 (1.70–4.22)	<0.001	3.41 (2.07–5.62)	<0.001	5.15 (2.94–9.03)	<0.001
*DRB1*04:01/x* *	1.44 (1.11–1.86)	0.006	1.49 (1.09–2.04)	0.012	1.40 (0.94–2.07)	0.096
Genetic risk score †	1.49 (1.35–1.65)	<0.001	1.68 (1.48–1.90)	<0.001	1.68 (1.43–1.98)	<0.001
BTNL2 rs3763305 GG ‡	1.02 (0.77–1.36)	0.88	1.34 (0.93–1.94)	0.12	1.74 (1.07–2.84)	0.026
ITGA1 rs7735139 "GG" ‡	1.25 (0.91–1.73)	0.17	1.14 (0.76–1.72)	0.52	1.21 (0.72–2.03)	0.47

* Reference: *DRB1* without *0401* or *0401/0404* and *0401/0407*; † per unit increase; ‡ reference: GA/AA genotype

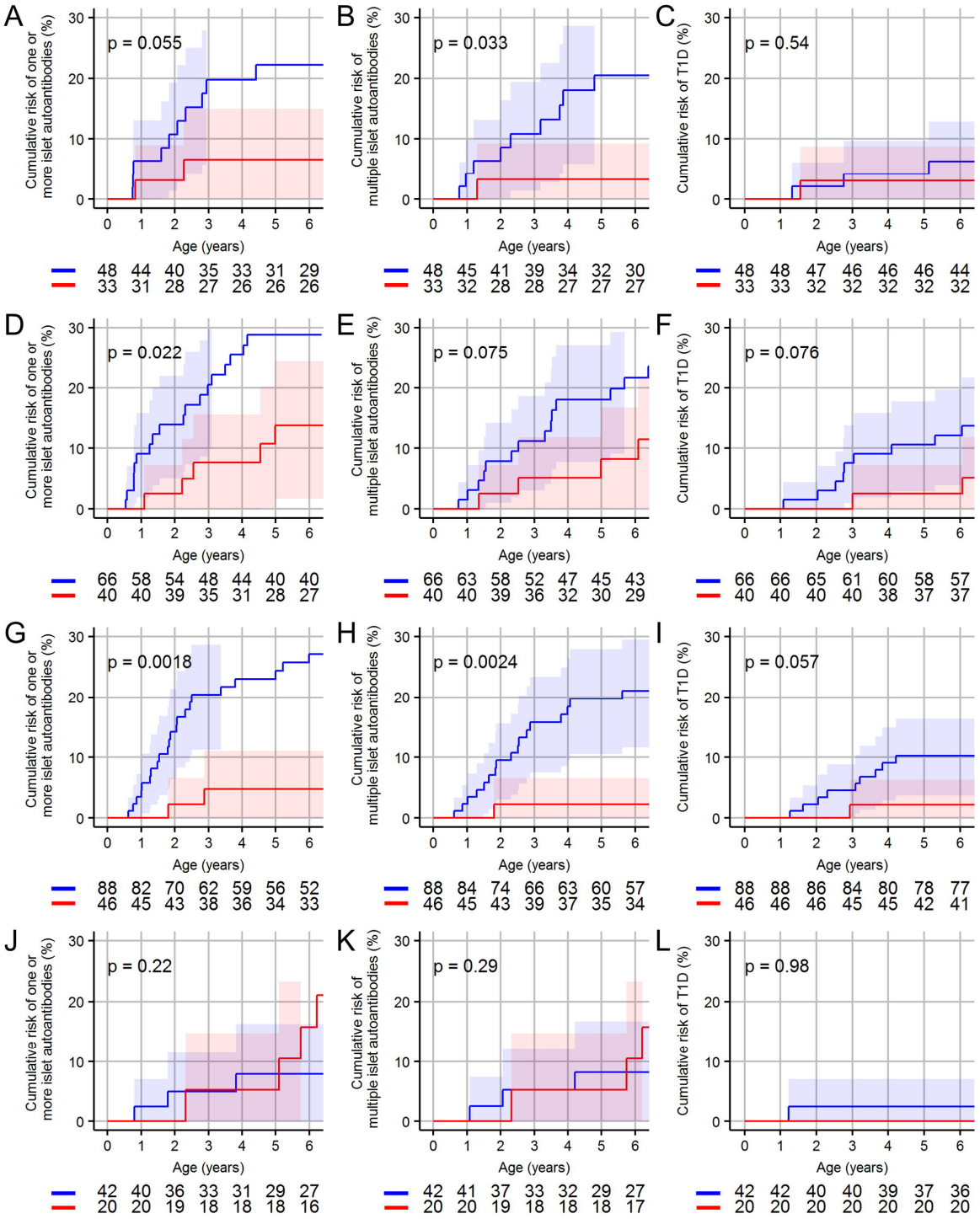


Figure 19. Risk of developing islet autoantibodies and diabetes in maternal and paternal type 1 diabetes. Risk of developing one or more islet autoantibodies **(A, D, G, J)**, multiple islet autoantibodies **(B, E, H, K)** and diabetes **(C, F, I, L)** in children with a mother with T1D (red) compared with children with a father or sibling with T1D (blue). Children have been stratified by genetic risk score and HLA *DRB1*04* subtype into four risk strata from highest genetic susceptibility **(A, B, C)**, to the lowest genetic susceptibility **(J, K, L)**. *P*-values were calculated using log-rank tests.

5.3.2 Detection of viral sequences in peripheral blood mononuclear cells of children with viral respiratory tract infections preceding islet autoimmunity (BABYDIET)

To further investigate the interplay of genetic and environment and its contribution to heterogeneity in the risk of developing islet autoimmunity, I searched for viral sequences in the BABYDIET cohort, which included high-risk children that expressed antiviral type 1 interferon transcript signatures prior to seroconversion. I aimed to detect viral infections in PBMCs to find a potential causal relation to the development of islet autoimmunity.

After a median follow-up of 4.1 years (IQR 1.4–7.9 years), 29 children developed islet AABs and 16 progressed to T1D (7). In 20 of the 29 children with islet AABs, PBMC samples were obtained in the first year; the remaining 9 children had no PBMC samples collected (7). For comparison, PBMC samples of 20 randomly selected BABYDIET children who remained islet AAB negative were also studied (Table 24) (7, 123).

In one of 102 PBMC samples a virus was detected by application of VirCapSeq-VERT. This positive PBMC sample contained rotavirus sequence, identified at age 1.05 years in a child who converted to islet AABs at the age of 2.1 years (Table 24 and Table 25). Rotavirus is the leading cause of diarrhea hospitalization among children worldwide and was the most frequently reported disease in Germany between 2001 and 2008 when this child was studied (https://survstat.rki.de and Koch and Wiese-Posselt (150)). For the rotavirus-positive child, a respiratory infection (ICD-10 J20.9) as well as an infectious gastroenteritis (ICD-10 A09) were reported during the sampling period.

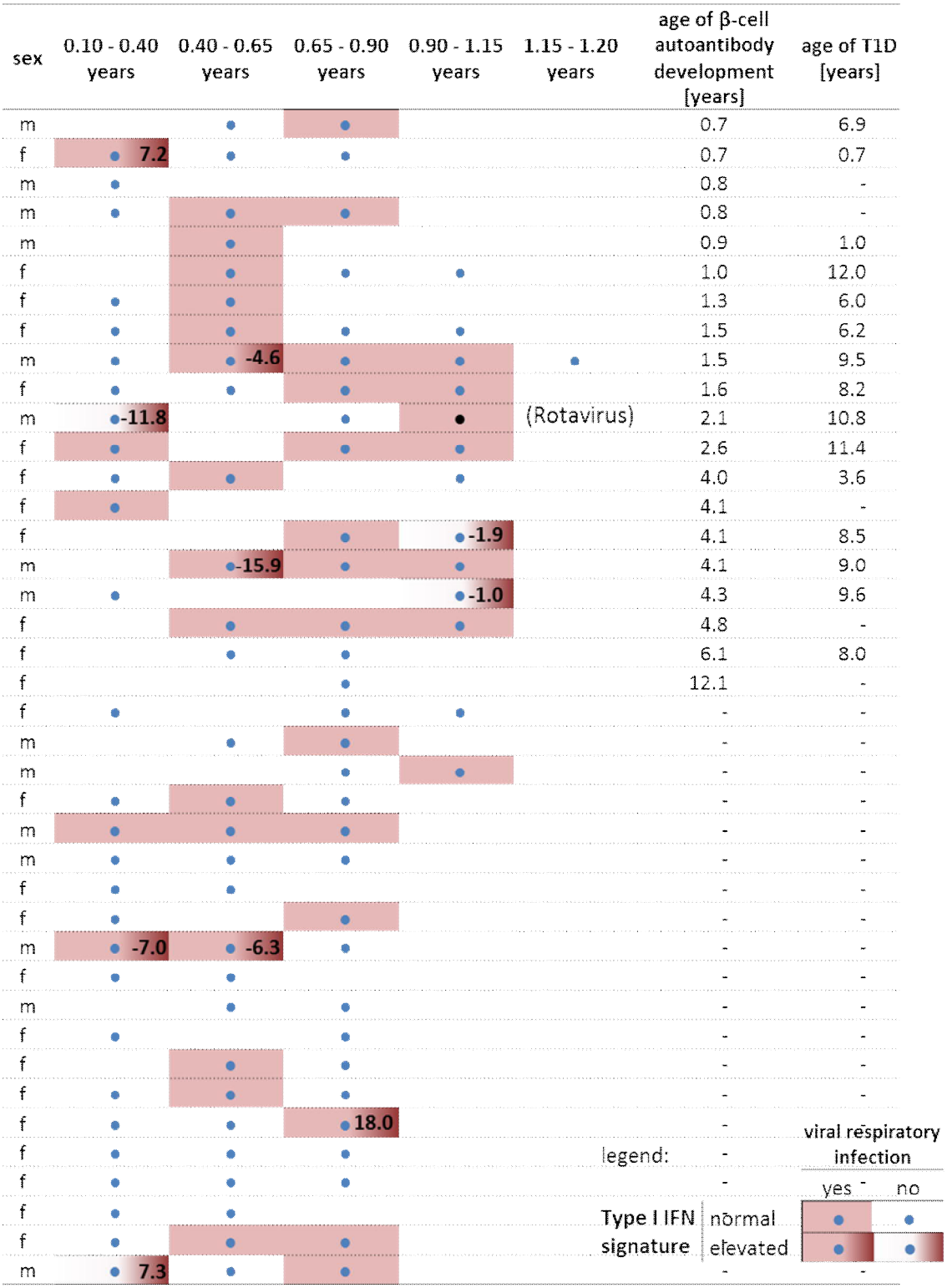
All other PBMC samples of children with and without progression to islet autoimmunity were negative of viral nucleic acids (Table 25).

Table 24. Description of BABYDIET sample demographics.

	controls	patients
sex (m/f)	7/13	8/12
median time to seroconversion, years (range)	-	1.8 (0.7 - 12.1)
developed T1D (n)	0	15
median time to T1D, years (range)	-	8.2 (0.7 - 12.0)
HLA *DR3* (%, frequency of subjects)	70	75
HLA *DR4* (%, frequency of subjects)	70	95

The results of this analysis have been published as:

Hippich, M., A. Oleynik, K. Jain, C. Winkler, R. C. Ferreira, E. Bonifacio, A. G. Ziegler and T. Briese (2018). "Searching peripheral blood mononuclear cells of children with viral respiratory tract infections preceding islet autoimmunity for viruses by high-throughput sequencing." Acta Diabetol 55(8): 881-884.



sex	0.10 - 0.40 years	0.40 - 0.65 years	0.65 - 0.90 years	0.90 - 1.15 years	1.15 - 1.20 years	age of β-cell autoantibody development [years]	age of T1D [years]
m		●	●			0.7	6.9
f	● 7.2	●	●			0.7	0.7
m	●					0.8	-
m	●	●	●			0.8	-
m		●				0.9	1.0
f		●	●	●		1.0	12.0
f	●	●				1.3	6.0
f	●	●	●	●		1.5	6.2
m	●	● -4.6	●	●	●	1.5	9.5
f	●	●	●	●		1.6	8.2
m	● -11.8		●	●	(Rotavirus)	2.1	10.8
f	●		●	●		2.6	11.4
f	●	●		●		4.0	3.6
f	●					4.1	-
f			●	● -1.9		4.1	8.5
m		● -15.9	●	●		4.1	9.0
m	●			● -1.0		4.3	9.6
f		●	●	●		4.8	-
f		●	●			6.1	8.0
f			●			12.1	-
f	●		●	●		-	-
m		●	●			-	-
m			●	●		-	-
f	●	●	●			-	-
m	●	●	●			-	-
m	●	●	●			-	-
f	●	●				-	-
f	●		●			-	-
m	● -7.0	● -6.3	●			-	-
f	●	●				-	-
m		●	●			-	-
f	●		●			-	-
f		●	●			-	-
f	●	●	●			-	-
f	●	●	● 18.0			-	-
f	●	●	●			-	-
f	●	●	●			-	-
f	●	●				-	-
f	●	●				-	-
m	● 7.3	●	●			-	-

Table 25. Overview of BABYDIET patient and control samples processed for sequencing. Description of abbreviations and symbols in the table: f female; m male; ● shows analyzed samples and ● denotes the positive rotavirus detection; Background color integrates Type I IFN signature and viral respiratory infection history around the time of blood draw (30 days before to 7 days after sampling date). Elevated Type I IFN signature is displayed as quantified by Principal Component 1 (114).

6 DISCUSSION

In my thesis, I addressed the definition of heterogeneity of diabetes in childhood and adolescence based on the contributions of β-cell function, genetic predisposition and environmental factors. I developed a clinically relevant classifier of endotypes in new onset diabetes patients which was based on residual β-cell function and routine measurements. This decision algorithm allowed the separation of mild and severe forms of T1D and additionally indicated the presence of other forms of diabetes mellitus. Also, I defined the genetic contribution of the excess risk of developing T1D associated autoimmunity in children with a first-degree relative with T1D as compared to children from the general population, and discovered a new potential risk gene *BTNL2*. Stratification based on these genetic elements identified a high risk stratum in which children from the general population and children who were first-degree relatives of patients with T1D had the same risk, and showed that there was increasing divergence in risk of autoimmunity between the two groups as the genetic risk diminished. Finally, I showed that the protective effect of maternal T1D is dominant over genetic susceptibility and can reduce the risk of developing islet autoantibodies even in children in the, previously defined, highest genetic susceptibility stratum. Furthermore, the search for viral sequences was unable to identify a specific virus preceding T1D associated autoimmunity in type 1 interferon positive cases.

6.1 Heterogeneity in β-cell reserve (DiMelli study)

I used a surrogate for residual β-cell function, the measure of fasting c-peptide, to classify childhood diabetes into endotypes with relevance to disease pathology, severity and possibly therapy. C-peptide at the time of diagnosis was used as the outcome variable (and discriminator) for the application of a CART algorithm in a dataset of 1194 new onset diabetes patients <20 years of age. The introduction of demographic, genetic and immune marker variables and metabolic features helped define 8 clinically relevant endotypes. Among islet AAB negative patients, I distinguished three groups corresponding to those who had features of monogenic diabetes, T1D, and T2D, respectively. Among the islet AAB positive patients, I was able to identify five endotypes. They differed markedly in residual β-cell function, severity at onset, concentrations of inflammatory cytokines, BMI SDS and gender and may be relevant for disease prognosis.

In previous studies, heterogeneity in childhood diabetes was rather classified by using a single marker, or one or two markers, but a more complex approach was still missing.

Ludvigsson et al. has studied a cohort of 2734 children and adolescents with new onset diabetes from Sweden, and found that c-peptide is highest in patients with T2D, followed by MODY, and T1D. The authors further conclude that children with a c-peptide ≥1.0 nmol/l have a 46% chance of having T2D or MODY and that children with a c-peptide <0.2 nmol/l have a 99.9% chance of having T1D (25). Earlier, Törn et al describes in a smaller study of around 900 patients that c-peptide distinguishes between patients with and without islet AABs (32). The findings of this large study described by Ludvigsson et al. are in line with my findings with regard to elevated levels of fasting c-peptide in (non-T1D) endotypes E1 and E3 that correspond to higher random c-peptide levels in MODY or T2D children. However, my analysis expands on a further classification, especially of T1D, beyond the classical clinical diagnosis applied by Ludvigsson and colleagues.

Dabelea et al. have used islet AABs against GAD and IA-2 in combination with insulin sensitivity to describe four subgroups in 2291 diabetes patients from the SEARCH study diagnosed at <30 years of age. The four groups represent either classical T1D (autoimmune, insulin sensitive), classical T2D (non-autoimmune, insulin resistant), T1D combined with obesity (autoimmune, insulin resistant), and a group that is thought to include patients who have T1D with undetected autoimmunity but also monogenic diabetes (non-autoimmune, insulin sensitive) (151). The endotypes defined in my thesis by the CART approach have a higher resolution than those of Dabelea and colleagues but some of them align with their categorization: my endotype E3 corresponds to the classical T2D, E1 and E2 to undetected autoimmunity/MODY, E7 and E8 to T1D with insulin resistance, and E5 and E6 to classical T1D.

In the same study, Dabelea et al. were able to look at follow-up measurements of c-peptide at a median time of 8 months after diagnosis. In islet AAB positive patients, the authors describe a progressive decline of fasting c-peptide. In contrast, in the islet AAB negative patients, they report heterogeneity and speculate again about a subgroup with undetected autoimmunity and similar decline as islet AAB positive patients and a subgroup with limited decline and presumably insulin-resistant diabetes (10). The inclusion of follow-up samples appears to have some added value to the classification. It is missing in the DiMelli study.

Verkauskiene et al. studied 1209 patients with diabetes of any diabetes duration from Lithuania, and reported that islet AABs negative patients have higher frequencies of family history of diabetes and retinopathy, and a lower frequency of ketosis at presentation, but similar age at onset, HbA1c, frequency of nephropathy and neuropathy compared to islet AAB positive patients (152). While I cannot conclude on retinopathy, nephropathy and neuropathy in the DiMelli study as DiMelli only included patients with new onset disease, I

cannot confirm the age of diagnosis, as the age in DiMelli was higher in patients with T2D or endotypes of T1D and obesity.

Warncke et al. have previously analyzed 630 patients from the DiMelli study and categorized them into patients without, with one or multiple islet AABs. They show that the groups differ in BMI percentile, weight loss before diagnosis, fasting c-peptide and insulin sensitivity sore. Yet, they find considerable overlap between phenotypes and suggest that a refined classification may require immune as well as metabolic phenotyping (8). My analysis of the same, but extended, study population confirms differences between islet AAB negative and positive patients but increases resolution of phenotypes as the CARTs result in 3 nodes (3 endotypes) in the islet AAB negatives and 7 nodes in the islet AAB positives (5 endotypes), respectively.

Patel et al. have taken the classification of diabetes in childhood one step further; they investigate a population of 1963 patients with T1D, 805 patients with MODY and 242 patients with neonatal diabetes, and show that a genetic risk score, initially developed for T1D, can discriminate between patients with T1D and MODY (153).

The CART analysis of the DiMelli study was also able to identify one endotype E1 with characteristic features of MODY. Patients with this endotype had no high-risk HLA genotypes, a low average genetic risk score, and normal values of vitamin D which was atypical for patients with either T1D or T2D in my analysis. An important difference to Patel et al., was that my CART analysis demonstrated that a discrimination of potential MODY, T1D, and T2D in islet AAB negative patients can be achieved by a very simple and cost-effective approach; namely the combination of BMI SDS and HbA1c at diagnosis.

Finally, there are some studies addressing heterogeneity of diabetes in adult onset diabetes. In a study of 1180 adults, Thunander et al. report that fasting c-peptide is a better discriminator between islet AAB negative and positive patients compared to age and BMI and that fasting c-peptide increases with age and BMI (33). Ahlqvist et al. performs k-means and hierarchical clustering on scaled and centered data of GAD-antibodies, BMI, HbA1c, HOMA2-B and HOMA2-IR from a total of 14755 adult patient from 3 cohorts. Five clusters are identified predicting disease progression and development of diabetic complications. This is a clear novel step forward for the classification of adult diabetes (138) with improvement compared to simpler classification based on c-peptide, BMI and age (33). It remains to be seen whether cluster – targeted treatment approaches will be able to prevent complications in the future in those patient groups with high risk. CART analysis has been successfully applied in cancer research and other autoimmune diseases such as systemic lupus erythematosus but has not previously be attempted in childhood diabetes. Jiang et al. used CART analysis to successfully define intratumoral heterogeneity in relation to patient

survival in 160 (40 per stage) clear cell renal cell carcinomas based on markers assessed by immunohistochemistry (154). Also, Barnholtz-Sloan et al. successfully applied CART to model the risk of prostate cancer in two different genetic populations using androgen pathway genes in combination with age and family history of the disease (155). Banchereau et al. have used linear mixed models of longitudinal blood transcriptome profiles, demographics, disease activity and nephritis class of 159 systemic lupus erythematosus patients to identify seven patient subgroups with differences in IFN and plasmablast signatures (156). This comprehensive approach demonstrates the benefit of the definition of disease endotypes by providing explanations for failure of previous clinical trials and suggestions for new trial design and tailored therapies.

6.1.1 Strength and limitations

In order to develop a classifier of diabetes in childhood and adolescence with clinical relevance, the model was required to handle prediction and explanation at the same time. On one side, so called ‘black box’ systems, i.e. support vector machines are very good predictors but lack explanation (157). On the other side, regression models, such as Cox models, do explain the relevance of single covariates (158). The CART algorithm is a regression model selecting the covariate representing the best split with regard to the outcome variable. It has been advanced over the years to solve overfitting and selection bias problems. Another important aspect is the possibility to handle nominal, ordinal, discrete and continuous variables simultaneously (130). Finally, the result of the algorithm is simple, easy to interpret and, despite the amount of input variables, only requires readily available clinical variables to stratify patients into groups with potentially different pathogenesis as well as prognosis. “Therefore, information from CART models can be used to develop individualized interventions and/or treatments, while information from regression models applies to the average member of a population only” (155).

The analysis of the DiMelli study presented here has some limitations. First of all, the cohort was recruited from a locally limited area, Bavaria in Germany and can, therefore, not draw any conclusion of the developed decision tree in other ethnicities. Because the stability of clusters was not validated so far, the model cannot claim that this classification is the best possible. Also, a validation and further genetic testing, especially in islet AAB negative patients, is missing to date. This is planned for the future and is thought to strengthen this model and reveal different pathologies among endotypes. Another limitation is the lack of longitudinal follow-up samples. Ahlquist et al. demonstrated how this further improves the definition of endotypes (138). Especially in patients with suspected monogenic diabetes, the course of c-peptide has been shown to differentiate from T1D and T2D patients (34).

Therefore, a second samples obtained at ≥6 months after diagnosis would help to further refine such an algorithm in a future approach.

The clear strength of the multivariable CART approach was based on a combination of stratification for islet AABs and the inclusion of previously associated input variables; by using this approach clear boundaries could be defined that assigned patients into different nodes. To my knowledge, this novel representation of endotypes in childhood diabetes is the most comprehensive one and goes beyond current practice. I also demonstrated that inclusion of genetics did not improve classification at the time of diagnosis. To this end, the developed algorithm represented a cheap and simple way for the identification of candidates for genetic testing for monogenic diabetes, for children, although islet AAB negative, who have T1D and may require insulin and for the classification of autoimmune endotypes. Furthermore, the inflammatory endotype E4 (CART B1) might be of relevance with respect to immunotherapy.

6.1.2 Outlook

Further refinement of endotypes could advance the resolution and lead to hypotheses on mechanistic differences. For example, the titer of islet AABs could be implemented into the model as it is associated with risk of developing T1D (125, 159). It would also be interesting, especially in endotypes E4, E5 and E6, to assess the response to different stimuli including, but not limited, to IL-2 and its downstream target STAT5 for investigation on regulatory T cell development and homeostasis (160), IL-7 for investigations on adaptive immunity and thymopoiesis (161), and LPS (162) or IFN-α (163) for investigation on infectious events in combination with examination of different cell populations (monocytes, B lymphocytes, memory and naïve CD4+ T cells, CD8+ T cells and Treg cells) for immune cell profiling in these particular patient subgroups.

6.2 Heterogeneity in genetic predisposition (TEDDY study)

In my thesis, I aimed to define the genetic contribution to the excess risk of developing islet autoimmunity and T1D in children with an affected first-degree family member. I investigated children from the TEDDY study with two genotypes, HLA *DR3/4-DQ8* and *DR4-DQ8/DR4-DQ8*, and studied their genetic profile by comparing children with a first degree family history with T1D to children from the general population. In terms of the genetic profile, I investigated the frequency of *DRB1*04*-subtypes, the distribution of T1D susceptibility genes other than HLA, the genetic risk score of these susceptibility genes, and x additional SNPs from the ImmunoChip.

In the analysis of genetic contribution to T1D in the TEDDY study I found that the excess risk for islet AABs and diabetes in children with a first-degree relative with T1D as compared to children from the general population could be abrogated by accounting for specific *DRB1*04* subtypes and increased load of T1D susceptibility alleles at multiple loci, including a novel susceptibility region marked by SNPs within the *BTLN2* gene. The risk of developing islet autoimmunity and T1D converged between children with a first-degree relative with T1D and children from the general population when they were matched at the highest genetic susceptibility and became increasingly divergent as genetic susceptibility was attenuated. This is of practical relevance as it identifies children from the general population whose risk for islet AABs and T1D is as high as that in the highest-risk children with a first-degree family history of T1D.

The study was performed in a large number of children with and without a first-degree relative with T1D of mainly European descent who were matched for the two highest-risk HLA class II genotypes. This unique cohort allowed assessing the contributions of other genetic factors. After selection by HLA genotype, the excess risk for islet AABs and diabetes was around 2- to 3-fold higher in children with a first-degree relative with T1D, which is markedly less than the >10-fold excess observed without HLA selection. Enrichment of genetic susceptibility was observed for HLA *DRB1*04* subtypes and by an increased genetic risk score for non-HLA loci. The addition of these genetic markers further reduced the excess risk in children with a first-degree relative with T1D, but the adjusted HRs remained above 2 for the development of multiple islet AABs or diabetes. Remarkably, this excess risk was heterogeneous, and depended on the *a priori* genetic susceptibility.

The excess risk that remained unaccounted for by susceptibility genes in families is likely due to further genetic enrichment, including rare variants that may be more frequent in

familial cases, or other factors, such as a shared environment. The study provided the opportunity to search for additional genetic factors that may contribute to risk by exploring genes with allelic enrichment in children with a first-degree family history of T1D. A limitation of this approach is that, despite the size of the TEDDY study, there was relatively little power to find these genes across the whole genome, particularly for genes with low minor allele frequencies.

I was successful in finding an enrichment of alleles for two additional genes. One of the genes with allelic enrichment in the children with a first-degree family history of T1D, *BTNL2*, lies within the HLA class II region. SNPs within *BNTL2* were previously shown to be associated with other HLA DR-linked diseases, but in almost all cases, including T1D, the risk was attributed to LD with HLA *DR* (164). The study, which included over 3000 children with the HLA *DR3/DR4-DQ8* genotype and over 1500 with the *DR4-DQ8/DR4-DQ8* genotype, had sufficient power to adequately test the independent contribution of *BTNL2*. The G allele of the SNP rs3763305 increased the risk for T1D with a HR of around 1.7 in these HLA-selected children. Although the analyses also controlled for the HLA *DRB1*04* subtype, the possibility that the *BTNL2* SNP marks HLA *DR4* extended haplotypes cannot be excluded. However, there was an association between the non-susceptible *BTNL2* allele and *DRB1*04* subtypes that are protective or confer relatively low risk. It is, therefore, equally possible that some of the associations between *DRB1*04* subtype and T1D risk are due to variation in *BTNL2* rather than or in addition to HLA *DR*.

BTNL2 is a negative regulator of immunity that is expressed on antigen-presenting cells and affects the generation, proliferation, and function of regulatory T cells (165-167). It was demonstrated that *BTNL2* SNPs confer risk for sarcoidosis (168), a T cell-related inflammatory disease, independently of HLA *DR* (169) and influence antibody responses to dietary antigens (170). A relationship between the minor allele of the *BTNL2* rs3763305 genotype and *BTNL2* transcriptomic expression has been reported (171). Further studies are required to determine whether there are functional differences between *BTNL2* genotypes that may be relevant to T1D susceptibility

The remaining increased risk for islet autoimmunity and T1D in children with a first-degree relative with T1D, after accounting for genetic load implies that other factors, which are shared or enriched within affected families, contribute to the child's risk. It is known that a family history of T1D is associated with changes in parental practices in an effort to reduce the risk in their unaffected children (172). It is likely that such practices are more frequent in the children of affected families, and it seems possible that some of these practices may be associated with increased risk. It is also possible that family members more often share infections or diet that increase the risk for islet autoimmunity.

6.2.1 Strength and limitations

A limitation of the study is that I could not examine children with other HLA genotypes and, therefore, cannot assess whether the divergence continues in children with HLA genotypes associated with moderate or low risk. Although TEDDY is a unique study with unprecedented numbers of children with and without a first-degree family history of T1D for comparisons, the findings require further validation, especially in different ethnical populations as it has been shown that a genetic risk score cannot be applied uniformly on different ethnic populations (173, 174).

6.2.2 Outlook

In summary, we have shown that the increased risk of developing islet autoimmunity in children with a first-degree relative with T1D is largely due to an excess load of genetic susceptibility, we identified a potential novel gene that confers risk for islet autoimmunity, and we have shown that accounting for the excess genetic susceptibility leads to convergence in high-risk strata and divergence in lower-risk strata for the risk of developing islet autoantibodies and diabetes between children with a first-degree relative with T1D and children from the general population. These findings stress that environmental risk factors of disease will likely exert different effects in a gene-dependent manner, and that searching for these factors may require genetic stratification. In the future, to improve risk stratification on a genetic basis the predictive value of genetic risk scores might benefit from the addition of further, low-frequency SNPs and the addition of informative haplotypes.

6.3 Heterogeneity in environmental exposure (TEDDY and BABYDIET studies)

To elucidate the environmental side of the heterogeneity in childhood diabetes, I investigated two environmental factors that have been repeatedly shown to influence the risk of T1D; on one side the relative protection of maternal T1D in children with a first-degree family history of T1D (76) and on the other side viral infections in early life prior to the onset of islet autoimmunity (98, 99).

The Cox proportional hazards model of the TEDDY study confirmed that unlike children whose father or sibling had T1D there was no excess risk in children whose mother had T1D as compared to children from the general population (76). The relative protection conferred by maternal as compared to paternal T1D was pronounced in the higher genetic susceptibility strata, suggesting that maternal T1D harbors a dominantly protective environment in the presence of enriched genetic susceptibility. In contrast, in the lowest risk stratum no difference between maternal T1D and paternal T1D can be detected.

This may suggest that maternal protection requires a certain threshold of genetic susceptibility for T1D or certain affected pathways to the disease to be effective. In addition, higher insulin levels in the fetal and neonatal period have been reported in offspring of mothers with T1D which could lead to improved peripheral tolerance to insulin (76). The transfer of islet AABs from mother to child during pregnancy (97) is also well known and may affect the development of the child *in utero*. Of particular relevance, these data also suggest that the shared environment of siblings and fathers with T1D may be a source from which to identify environmental risk factors.

With regard to the environmental factor of viral infections, I hypothesized that we may be able to identify a pathogenic virus from PBMCs of infants who experienced respiratory tract infections before they developed islet AABs. I considered that the identification of a virus in the blood of predisposed children may provide important insights into the association of respiratory infections and T1D and reveal a potential causal link. To search for viruses in PBMCs, I used a novel highly sensitive method for viral enrichment in high-throughput sequencing and investigated samples collected from infants within the first year of life in which transcriptomics had been performed. I had assumed that children with an infection and a positive transcriptomic type I IFN signature would provide an optimal scenario for detecting potentially associated viruses in PBMCs, because the PBMCs revealed signs of antiviral reactivity. However, the search in PBMCs before the conversion to islet autoimmunity failed to consistently detect viral sequences. VirCapSeq-VERT sequencing

identified viruses in only one of 102 PBMC samples. The positive PBMC sample contained rotavirus sequence, identified at the age of 1.05 years in a child who presented with beta cell autoantibodies at the age of 2.1 years. All other PBMC samples of children with and without progression to islet autoimmunity were negative.

Rotavirus is the leading cause of diarrhea hospitalization among children worldwide and was the most frequently reported disease in Germany between 2001 and 2008 when this child was studied (150). Rotavirus RNA has been previously reported in 70% of PBMC samples during the acute phase (<72h) and in 4-8% during the convalescent phase (3 weeks) of an infection (175). Rotaviruses have been previously associated with T1D (109).

Overall, these findings are in line with a previous publication using high throughput sequencing to search for viruses in plasma samples, where viral RNA in plasma was not more frequently detected in children with rapid-onset T1D than in controls (176).

A limitation of the study is that it could not investigate virus in plasma, nasal swabs, or stool samples taken concurrently. The analysis of the gut microbiome by Kim and colleagues, using the exact same method, revealed higher abundance of enterovirus A species in 45 children with islet autoimmunity as compared to 48 matched controls. The same study reports that detection of viral sequences in gut do not overlap viral sequences detected in plasma. This phenomenon is attributed to the short duration of viremia (177).

In summary, my analysis showed that even with highly sensitive sequencing methods it is challenging to identify possible causal agents of beta cell autoimmunity in blood and that a variety of sample materials and tighter sampling schedule might be necessary to draw conclusions on the interaction of viral infections and the development of T1D.

REFERENCES

1. Mayer-Davis EJ, et al. ISPAD Clinical Practice Consensus Guidelines 2018: Definition, epidemiology, and classification of diabetes in children and adolescents. *Pediatric Diabetes.* 2018;19 Suppl 27:7-19.
2. Gale EA. Type 1 diabetes in the young: the harvest of sorrow goes on. *Diabetologia.* 2005;48(8):1435-8.
3. of Torpy JM, Campbell A, and Glass RM. Chronic diseases of children. *JAMA.* 2010;303(7):682-.
4. Craig ME, Jefferies C, Dabelea D, Balde N, Seth A, and Donaghue KC. ISPAD Clinical Practice Consensus Guidelines 2014. Definition, epidemiology, and classification of diabetes in children and adolescents. *Pediatric Diabetes.* 2014;15 Suppl 20:4-17.
5. Vardi P, MD, et al. Concentration of Insulin Autoantibodies at Onset of Type I Diabetes: Inverse Log-Linear Correlation With Age. *Diabetes Care.* 1988;11(9):736-9.
6. Bonifacio E, Lampasona V, Genovese S, Ferrari M, and Bosi E. Identification of Protein Tyrosine Phosphatase-Like iA2 (Islet Cell Antigen 51 2) as the Insulin-Dependent Diabetes-Related 37/40K Autoantigen and a Target of Islet-Cell Antibodies. *The Journal of Immunolgy.* 1995;155(11):5419-26.
7. Ziegler AG, Bonifacio E, and BABYDIAB-BABYDIET Study Group. Age-related islet autoantibody incidence in offspring of patients with type 1 diabetes. *Diabetologia.* 2012;55(7):1937-43.
8. Warncke K, Krasmann M, Puff R, Dunstheimer D, Ziegler AG, and Beyerlein A. Does diabetes appear in distinct phenotypes in young people? Results of the diabetes mellitus incidence Cohort Registry (DiMelli). *PloS one.* 2013;8(9):e74339.
9. Hamman RF, et al. The SEARCH for Diabetes in Youth study: rationale, findings, and future directions. *Diabetes Care.* 2014;37(12):3336-44.
10. Dabelea D, et al. Clinical evolution of beta cell function in youth with diabetes: the SEARCH for Diabetes in Youth study. *Diabetologia.* 2012;55(12):3359-68.
11. Ziegler AG, Hummel M, Schenker M, and Bonifacio E. Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. *Diabetes.* 1999;48(3):460-8.
12. Bonifacio E, et al. Islet autoantibody markers in IDDM: risk assessment strategies yielding high sensitivity. *Diabetolgia.* 1995;38:816-22.
13. Dopico XC, et al. Widespread seasonal gene expression reveals annual differences in human immunity and physiology. *Nature Communications.* 2015;6:7000.
14. Insel RA, et al. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care.* 2015;38(10):1964-74.
15. Krischer JP, et al. Genetic and Environmental Interactions Modify the Risk of Diabetes-Related Autoimmunity by 6 Years of Age: The TEDDY Study. *Diabetes Care.* 2017;40(9):1194-202.

16. Krischer JP, et al. The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study. *Diabetologia.* 2015;58(5):980-7.
17. Walther D, et al. Tetraspanin 7 autoantibodies in type 1 diabetes. *Diabetologia.* 2016;59(9):1973-6.
18. McLaughlin KA, et al. Identification of Tetraspanin-7 as a Target of Autoantibodies in Type 1 Diabetes. *Diabetes.* 2016;65(6):1690-8.
19. Gottlieb PA, et al. Chromogranin A is a T cell antigen in human type 1 diabetes. *J Autoimmun.* 2014;50:38-41.
20. Mukherjee R, Wagar D, Stephens TA, Lee-Chan E, and Singh B. Identification of CD4+ T cell-specific epitopes of islet-specific glucose-6-phosphatase catalytic subunit-related protein: a novel beta cell autoantigen in type 1 diabetes. *J Immunol.* 2005;174(9):5306-15.
21. Ouyang Q, et al. Recognition of HLA class I-restricted beta-cell epitopes in type 1 diabetes. *Diabetes.* 2006;55(11):3068-74.
22. Pugliese A. Autoreactive T cells in type 1 diabetes. *The Journal of Clinical Investigation.* 2017;127(8):2881-91.
23. VanBuecken DE, and Greenbaum CJ. Residual C-peptide in type 1 diabetes: what do we really know? *Pediatric Diabetes.* 2014;15(2):84-90.
24. Soltesz G, Patterson C, and Dahlquist G. Worldwide childhood type 1 diabetes incidence – what can we learn from epidemiology? *Pediatric Diabetes.* 2007;8(s6):6-14.
25. Ludvigsson J, et al. C-peptide in the classification of diabetes in children and adolescents. *Pediatric Diabetes.* 2012;13(1):45-50.
26. Velho G, and Froguel P. Genetic, metabolic and clinical characteristics of maturity onset diabetes of the young. *European journal of endocrinology.* 1998;138(3):233-9.
27. Chen C, Cohrs CM, Stertmann J, Bozsak R, and Speier S. Human beta cell mass and function in diabetes: Recent advances in knowledge and technologies to understand disease pathogenesis. *Molecular Metabolism.* 2017;6(9):943-57.
28. Ludvigsson J, and Heding LG. C-peptide in children with juvenile diabetes. A preliminary report. *Diabetologia.* 1976;12(6):627-30.
29. Leighton E, Sainsbury CA, and Jones GC. A Practical Review of C-Peptide Testing in Diabetes. *Diabetes Ther.* 2017;8(3):475-87.
30. Jones AG, and Hattersley AT. The clinical utility of C-peptide measurement in the care of patients with diabetes. *Diabetic Medicine.* 2013;30(7):803-17.
31. Shields BM, et al. C-Peptide Decline in Type 1 Diabetes Has Two Phases: An Initial Exponential Fall and a Subsequent Stable Phase. *Diabetes Care.* 2018;41(7):1486-92.
32. Törn C, Landin-Olsson M, and Scherstén B. Predictability of C-peptide for autoimmune diabetes in young adult diabetic patients. *Practical Diabetes International.* 2001;18(3):83-8.
33. Thunander M, Torn C, Petersson C, Ossiansson B, Fornander J, and Landin-Olsson M. Levels of C-peptide, body mass index and age, and their usefulness in classification of diabetes in relation to autoimmunity, in adults with newly diagnosed diabetes in Kronoberg, Sweden. *European journal of endocrinology.* 2012;166(6):1021-9.

34. Majidi S, et al. Can Biomarkers Help Target Maturity-Onset Diabetes of the Young Genetic Testing in Antibody-Negative Diabetes? *Diabetes Technol Ther.* 2018;20(2):106-12.
35. Steffes MW, Sibley S, Jackson M, and Thomas W. β-Cell Function and the Development of Diabetes-Related Complications in the Diabetes Control and Complications Trial. *Diabetes Care.* 2003;26(3):832.
36. Thumer L, et al. German new onset diabetes in the young incident cohort study: DiMelli study design and first-year results. *The Review of Diabetic Studies.* 2010;7(3):202-8.
37. Murphy R, Ellard S, and Hattersley AT. Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nature clinical practice Endocrinology & metabolism.* 2008;4(4):200-13.
38. Naylor R, Knight Johnson A, and del Gaudio D. Maturity-Onset Diabetes of the Young Overview. https://www.ncbi.nlm.nih.gov/books/NBK500456/. Accessed 06.02., 2019.
39. Hattersley AT, et al. ISPAD Clinical Practice Consensus Guidelines 2018: The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatric Diabetes.* 2018;19 Suppl 27:47-63.
40. Rubio-Cabezas O, and Ellard S. Diabetes mellitus in neonates and infants: genetic heterogeneity, clinical approach to diagnosis, and therapeutic options. *Hormone Research in Paediatrics.* 2013;80(3):137-46.
41. Skyler JS, et al. Differentiation of Diabetes by Pathophysiology, Natural History, and Prognosis. *Diabetes.* 2017;66(2):241-55.
42. Patterson CC, Dahlquist GG, Gyürüs E, Green A, and Soltész G. Incidence trends for childhood type 1 diabetes in Europe during 1989–2003 and predicted new cases 2005–20: a multicentre prospective registration study. *The Lancet.* 2009;373(9680):2027-33.
43. Bendas A, et al. Trends in Incidence Rates during 1999-2008 and Prevalence in 2008 of Childhood Type 1 Diabetes Mellitus in GERMANY – Model-Based National Estimates. *PloS one.* 2015;10(7):e0132716.
44. Ziegler AG, and Nepom GT. Prediction and pathogenesis in type 1 diabetes. *Immunity.* 2010;32(4):468-78.
45. Raab J, et al. Capillary blood islet autoantibody screening for identifying pre-type 1 diabetes in the general population: design and initial results of the Fr1da study. *BMJ Open.* 2016;6(5).
46. Hagopian WA, et al. The Environmental Determinants of Diabetes in the Young (TEDDY): genetic criteria and international diabetes risk screening of 421 000 infants. *Pediatr Diabetes.* 2011;12(8):733-43.
47. Schenker M, et al. Early expression and high prevalence of islet autoantibodies for DR3/4 heterozygous and DR4/4 homozygous offspring of parents with Type I diabetes: the German BABYDIAB study. *Diabetologia.* 1999;42(6):671-7.
48. Torn C, et al. Role of Type 1 Diabetes-Associated SNPs on Risk of Autoantibody Positivity in the TEDDY Study. *Diabetes.* 2015;64(5):1818-29.
49. Noble JA, et al. HLA class I and genetic susceptibility to type 1 diabetes: results from the Type 1 Diabetes Genetics Consortium. *Diabetes.* 2010.

50. Hu X, et al. Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nat Genet.* 2015;47(8):898.
51. Noble JA, and Erlich HA. Genetics of Type 1 Diabetes. *Cold Spring Harb Perspect Med.* 2012;2(1).
52. Erlich H, et al. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. *Diabetes.* 2008;57(4):1084-92.
53. Noble JA. Immunogenetics of type 1 diabetes: A comprehensive review. *J Autoimmun.* 2015;64:101-12.
54. Walter M, et al. IDDM2/insulin VNTR modifies risk conferred by IDDM1/HLA for development of Type 1 diabetes and associated autoimmunity. *Diabetologia.* 2003;46(5):712-20.
55. Mack SJ, et al. HLA-A, -B, -C, and -DRB1 allele and haplotype frequencies distinguish Eastern European Americans from the general European American population. *Tissue Antigens.* 2009;73(1):17-32.
56. Kempson F, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res.* 2014;43(D1):D784-D8.
57. Caillat-Zucman S, et al. Age-dependent HLA Genetic Heterogeneity of Type 1 Insulin-dependent Diabetes Mellitus. *Journal of Clinical Investigation.* 1992;90:2242-50.
58. Barrett JC, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet.* 2009;41(6):703-7.
59. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447(7145):661-78.
60. Cho JH, and Feldman M. Heterogeneity of autoimmune diseases: pathophysiologic insights from genetics and implications for new therapies. *Nat Med.* 2015;21(7):730-8.
61. Bennett ST, et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat Genet.* 1995;9(3):284-92.
62. Pugliese A, et al. The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. *Nat Genet.* 1997;15(3):293-7.
63. Vafiadis P, et al. Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. *Nat Genet.* 1997;15(3):289-92.
64. Yang Z, et al. The CD4(+) regulatory T-cells is decreased in adults with latent autoimmune diabetes. *Diabetes Res Clin Pract.* 2007;76(1):126-31.
65. Durinovic-Bello I, Wu RP, Gersuk VH, Sanda S, Shilling HG, and Nepom GT. Insulin gene VNTR genotype associates with frequency and phenotype of the autoimmune response to proinsulin. *Genes Immun.* 2010;11(2):188-93.
66. Kutlu B, et al. Detailed transcriptome atlas of the pancreatic beta cell. *BMC Med Genomics.* 2009;2(1):3.
67. Pociot F, et al. Genetics of type 1 diabetes: what's next? *Diabetes.* 2010;59(7):1561-71.
68. Steck AK, and Rewers MJ. Genetics of type 1 diabetes. *Clin Chem.* 2011;57(2):176-85.

69. Atkinson MA, Eisenbarth GS, and Michels AW. Type 1 diabetes. *The Lancet.* 2014;383(9911):69-82.
70. Winkler C, et al. A strategy for combining minor genetic susceptibility genes to improve prediction of disease in type 1 diabetes. *Genes Immun.* 2012;13(7):549-55.
71. Winkler C, et al. Feature ranking of type 1 diabetes susceptibility genes improves prediction of type 1 diabetes. *Diabetologia.* 2014;57(12):2521-9.
72. Oram RA, et al. A Type 1 Diabetes Genetic Risk Score Can Aid Discrimination Between Type 1 and Type 2 Diabetes in Young Adults. *Diabetes Care.* 2016;39(3):337-44.
73. Bonifacio E, et al. Genetic scores to stratify risk of developing multiple islet autoantibodies and type 1 diabetes: A prospective study in children. *PLoS Medicine.* 2018;15(4):e1002548.
74. Beyerlein A, et al. Progression from islet autoimmunity to clinical type 1 diabetes is influenced by genetic factors: results from the prospective TEDDY study. *J Med Genet.* 2018.
75. Knip M, and Simell O. Environmental triggers of type 1 diabetes. *Cold Spring Harb Perspect Med.* 2012;2(7):a007690-a.
76. Bonifacio E, Pfluger M, Marienfeld S, Winkler C, Hummel M, and Ziegler AG. Maternal type 1 diabetes reduces the risk of islet autoantibodies: relationships with birthweight and maternal HbA(1c). *Diabetologia.* 2008;51(7):1245-52.
77. Sorensen IM, Joner G, Jenum PA, Eskild A, Torjesen PA, and Stene LC. Maternal serum levels of 25-hydroxy-vitamin D during pregnancy and risk of type 1 diabetes in the offspring. *Diabetes.* 2012;61(1):175-8.
78. Miettinen ME, et al. Serum 25-hydroxyvitamin D level during early pregnancy and type 1 diabetes risk in the offspring. *Diabetologia.* 2012;55(5):1291-4.
79. Hummel S, Bonifacio E, Pflueger M, Ziegler A-G, and Hummel M. Primary Dietary Intervention Study to Reduce the Risk of Islet Autoimmunity in Children at Increased Risk for Type 1 Diabetes - The BABYDIET study. *Diabetes Care.* 2011;34:1301–5.
80. Norris JM, et al. Timing of initial cereal exposure in infancy and risk of islet autoimmunity. *JAMA.* 2003;290(13):1713-20.
81. Atkinson MA. Timing of initial cereal exposure in infancy and risk of islet autoimmunity. *The Journal of pediatrics.* 2004;144(5):684-5.
82. Ziegler AG, Schmid S, Huber D, Hummel M, and Bonifacio E. Early infant feeding and risk of developing type 1 diabetes-associated autoantibodies. *JAMA.* 2003;290(13):1721-8.
83. Virtanen SM, et al. Early introduction of root vegetables in infancy associated with advanced ss-cell autoimmunity in young children with human leukocyte antigen-conferred susceptibility to Type 1 diabetes. *Diabetic Medicine.* 2011;28(8):965-71.
84. Frederiksen B, et al. Infant exposures and development of type 1 diabetes mellitus: The Diabetes Autoimmunity Study in the Young (DAISY). *JAMA Pediatrics.* 2013;167(9):808-15.
85. Knip M, Virtanen SM, and Akerblom HK. Infant feeding and the risk of type 1 diabetes. *The American journal of clinical nutrition.* 2010;91(5):1506s-13s.
86. Chmiel R, Beyerlein A, Knopff A, Hummel S, Ziegler AG, and Winkler C. Early infant feeding and risk of developing islet autoimmunity and type 1 diabetes. *Acta Diabetol.* 2015;52(3):621-4.

87. Stene LC, and Rewers M. Immunology in the clinic review series; focus on type 1 diabetes and viruses: the enterovirus link to type 1 diabetes: critical review of human studies. *Clin Exp Immunol.* 2012;168(1):12-23.
88. Stene LC, et al. Enterovirus infection and progression from islet autoimmunity to type 1 diabetes: the Diabetes and Autoimmunity Study in the Young (DAISY). *Diabetes.* 2010;59(12):3174-80.
89. Richardson SJ, Willcox A, Bone AJ, Foulis AK, and Morgan NG. The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes. *Diabetologia.* 2009;52(6):1143-51.
90. Krogvold L, et al. Detection of a low-grade enteroviral infection in the islets of langerhans of living patients newly diagnosed with type 1 diabetes. *Diabetes.* 2015;64(5):1682-7.
91. Resic Lindehammer S, et al. Seroconversion to islet autoantibodies after enterovirus infection in early pregnancy. *Viral Immunol.* 2012;25(4):254-61.
92. Dahlquist G, Frisk G, Ivarsson SA, Svanberg L, Forsgren M, and Diderholm H. Indications that maternal coxsackie B virus infection during pregnancy is a risk factor for childhood-onset IDDM. *Diabetologia.* 1995;38(11):1371-3.
93. Dahlquist GG, Ivarsson S, Lindberg B, and Forsgren M. Maternal enteroviral infection during pregnancy as a risk factor for childhood IDDM. A population-based case-control study. *Diabetes.* 1995;44(4):408-13.
94. Viskari H, et al. Maternal enterovirus infection as a risk factor for type 1 diabetes in the exposed offspring. *Diabetes Care.* 2012;35(6):1328-32.
95. Rewers M, and Ludvigsson J. Environmental risk factors for type 1 diabetes. *The Lancet.* 2016;387(10035):2340-8.
96. Warram JH, Krolewski AS, Gottlieb MS, and Kahn CR. Differences in risk of insulin-dependent diabetes in offspring of diabetic mothers and diabetic fathers. *The New England journal of medicine.* 1984;311(3):149-52.
97. Koczwara K, Bonifacio E, and Ziegler AG. Transmission of maternal islet antibodies and risk of autoimmune diabetes in offspring of mothers with type 1 diabetes. *Diabetes.* 2004;53(1):1-4.
98. Beyerlein A, Donnachie E, Jergens S, and Ziegler A. Infections in early life and development of type 1 diabetes. *JAMA.* 2016;315(17):1899-901.
99. Beyerlein A, Wehweck F, Ziegler AG, and Pflueger M. Respiratory infections in early life and the development of islet autoimmunity in children at increased type 1 diabetes risk: evidence from the BABYDIET study. *JAMA Pediatrics.* 2013;167(9):800-7.
100. Lönnrot M, et al. Respiratory infections are temporally associated with initiation of type 1 diabetes autoimmunity: the TEDDY study. *Diabetologia.* 2017.
101. Oikarinen S, et al. Enterovirus RNA in Blood Is Linked to the Development of Type 1 Diabetes. *Diabetes.* 2011;60(1):276-9.
102. Yeung WC, Rawlinson WD, and Craig ME. Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies. *BMJ.* 2011;342:d35.

103. Dotta F, et al. Coxsackie B4 virus infection of β cells and natural killer cell insulitis in recent-onset type 1 diabetic patients. *Proceedings of the National Academy of Sciences.* 2007;104(12):5115.
104. Yoon JW, Austin M, Onodera T, and Notkins AL. Isolation of a virus from the pancreas of a child with diabetic ketoacidosis. *The New England journal of medicine.* 1979;300(21):1173-9.
105. Menser MA, Forrest JM, Honeyman MC, and Burgess JA. Letter: Diabetes, HL-A antigens, and congenital rubella. *The Lancet.* 1974;2(7895):1508-9.
106. Pak CY, Eun HM, McArthur RG, and Yoon JW. Association of cytomegalovirus infection with autoimmune type 1 diabetes. *The Lancet.* 1988;2(8601):1-4.
107. Kasuga A, Harada R, and Saruta T. Insulin-dependent diabetes mellitus associated with parvovirus B19 infection. *Ann Intern Med.* 1996;125(8):700-1.
108. Conrad B, Weissmahr RN, Boni J, Arcari R, Schupbach J, and Mach B. A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. *Cell.* 1997;90(2):303-13.
109. Honeyman MC, et al. Association Between Rotavirus Infection and Pancreatic Ilset Autoimmunity in Children at Risk of Developing Type 1 Diabetes. *Diabetes.* 2000;49:1319-24.
110. Goto A, et al. A case of fulminant type 1 diabetes associated with significant elevation of mumps titers. *Endocr J.* 2008;55(3):561-4.
111. Ashton MP, et al. Incomplete immune response to coxsackie B viruses associates with early autoimmunity against insulin. *Sci Rep.* 2016;6:32899.
112. Op de Beeck A, and Eizirik DL. Viral infections in type 1 diabetes mellitus - why the beta cells? *Nature Reviews Endocrinology.* 2016;12(5):263-73.
113. Giltiay NV, Chappell CP, and Clark EA. B-cell selection and the development of autoantibodies. *Arthritis Research and Therapy.* 2012;14:S1.
114. Ferreira RC, et al. A Type I Interferon transcriptional Signature precedes Autoimmunity in Children Genetically at Risk for Type 1 Diabetes. *Diabetes.* 2014;63(63):2538–50.
115. Strimbu K, and Tavel JA. What are biomarkers? *Current Opinion in HIV and AIDS.* 2010;5(6):463-6.
116. Barker JM, et al. Pre-type 1 diabetes dysmetabolism: maximal sensitivity achieved with both oral and intravenous glucose tolerance testing. *The Journal of pediatrics.* 2007;150(1):31-6.e6.
117. Hagopian WA, et al. TEDDY--The Environmental Determinants of Diabetes in the Young: an observational clinical trial. *Ann N Y Acad Sci.* 2006;1079:320-6.
118. Teddy Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY) study: study design. *Pediatr Diabetes.* 2007;8(5):286-98.
119. Teddy Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY) Study. *Ann N Y Acad Sci.* 2008;1150:1-13.
120. Bonifacio E, et al. Harmonization of glutamic acid decarboxylase and islet antigen-2 autoantibody assays for national institute of diabetes and digestive and kidney diseases consortia. *J Clin Endocrinol Metab.* 2010;95(7):3360-7.
121. Hippich M, et al. Genetic Contribution to the Divergence in Type 1 Diabetes Risk Between Children From the General Population and Children From Affected Families. *Diabetes.* 2019.

122. Schmid S, Buuck D, Knopff A, Bonifacio E, and Ziegler AG. BABYDIET, a feasibility study to prevent the appearance of islet autoantibodies in relatives of patients with Type 1 diabetes by delaying exposure to gluten. *Diabetologia.* 2004;47(6):1130-1.
123. Hippich M, et al. Searching peripheral blood mononuclear cells of children with viral respiratory tract infections preceding islet autoimmunity for viruses by high-throughput sequencing. *Acta Diabetol.* 2018;55(8):881-4.
124. Achenbach P, et al. Autoantibodies to zinc transporter 8 and SLC30A8 genotype stratify type 1 diabetes risk. *Diabetologia.* 2009;52(9):1881-8.
125. Achenbach P, et al. Stratification of Type 1 Diabetes Risk on the Basis of Islet Autoantibody Characteristics. *Diabetes.* 2004;53(2):384.
126. Burbelo P, et al. *Extrapancreatic Autoantibody Profiles in Type I Diabetes.* 2012.
127. Greenbaum CJ, et al. Preservation of beta-cell function in autoantibody-positive youth with diabetes. *Diabetes Care.* 2009;32(10):1839-44.
128. Bonifacio E, Mayr A, Knopff A, and Ziegler AG. Endocrine autoimmunity in families with type 1 diabetes: frequent appearance of thyroid autoimmunity during late childhood and adolescence. *Diabetologia.* 2009;52(2):185-92.
129. Purcell S, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet.* 2007;81(3):559-75.
130. Hothorn T, Hornik K, and Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics.* 2006;15(3):651-74.
131. Kromeyer-Hauschild K, et al. Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschrift Kinderheilkunde.* 2001;149(8):807-18.
132. Troegeler A, et al. An efficient siRNA-mediated gene silencing in primary human monocytes, dendritic cells and macrophages. *Immunol Cell Biol.* 2014;92(8):699-708.
133. Briese T, et al. Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. *MBio.* 2015;6(5):e01491-15.
134. Schmieder R, and Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics.* 2011;27(6):863-4.
135. Chevreux B, Wetter T, and Suhai S. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. *Computer Science and Biology Proceedings of the German Conference on Bioinformatics, GCB'99.* 1999;99:45–56.
136. Li H, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-9.
137. Milne I, et al. Tablet--next generation sequence assembly visualization. *Bioinformatics.* 2010;26(3):401-2.
138. Ahlqvist E, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The Lancet Diabetes and Endocrinology.* 2018;6(5):361-9.
139. Todd JA. Etiology of type 1 diabetes. *Immunity.* 2010;32(4):457-67.

140. Hao W, Woodwyk A, Beam C, Bahnson HT, Palmer JP, and Greenbaum CJ. Assessment of beta Cell Mass and Function by AIRmax and Intravenous Glucose in High-Risk Subjects for Type 1 Diabetes. *The Journal of clinical endocrinology and metabolism.* 2017;102(12):4428-34.
141. Morgan NG, Leete P, Foulis AK, and Richardson SJ. Islet inflammation in human type 1 diabetes mellitus. *IUBMB Life.* 2014;66(11):723-34.
142. Herold KC, et al. Teplizumab treatment may improve C-peptide responses in participants with type 1 diabetes after the new-onset period: a randomised controlled trial. *Diabetologia.* 2013;56(2):391-400.
143. Raab J, et al. Prevalence of vitamin D deficiency in pre-type 1 diabetes and its association with disease progression. *Diabetologia.* 2014;57(5):902-8.
144. Chakhtoura M, and Azar ST. The Role of Vitamin D Deficiency in the Incidence, Progression, and Complications of Type 1 Diabetes Mellitus. *Int J Endocrinol.* 2013;2013:10.
145. Martin T, and Campbell RK. Vitamin D and Diabetes. *Diabetes Spectr.* 2011;24(2):113.
146. Reijonen H, et al. HLA-DR4 subtype and -B alleles in DQB1*0302-positive haplotypes associated with IDDM. *Eur J Immunogenet.* 1997;24(5):357-63.
147. Nejentsev S, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature.* 2007;450(7171):887-92.
148. Barker JM, et al. Two single nucleotide polymorphisms identify the highest-risk diabetes HLA genotype: potential for rapid screening. *Diabetes.* 2008;57(11):3152-5.
149. The Genomes Project C, et al. A global reference for human genetic variation. *Nature.* 2015;526:68.
150. Koch J, and Wiese-Posselt M. Epidemiology of rotavirus infections in children less than 5 years of age: Germany, 2001-2008. *The Pediatric Infectious Disease Journal.* 2011;30(2):112-7.
151. Dabelea D, et al. Etiological approach to characterization of diabetes type: the SEARCH for Diabetes in Youth Study. *Diabetes Care.* 2011;34(7):1628-33.
152. Verkauskiene R, et al. The course of diabetes in children, adolescents and young adults: does the autoimmunity status matter? *BMC Endocr Disord.* 2016;16(1):61.
153. Patel KA, et al. Type 1 Diabetes Genetic Risk Score: A Novel Tool to Discriminate Monogenic and Type 1 Diabetes. *Diabetes.* 2016;65(7):2094-9.
154. Jiang W, et al. Intratumoral heterogeneity analysis reveals hidden associations between protein expression losses and patient survival in clear cell renal cell carcinoma. *Oncotarget.* 2017;8(23):37423-34.
155. Barnholtz-Sloan JS, Guan X, Zeigler-Johnson C, Meropol NJ, and Rebbeck TR. Decision tree-based modeling of androgen pathway genes and prostate cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* 2011;20(6):1146-55.
156. Banchereau R, et al. Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. *Cell.* 2016;165(3):551-65.
157. Barakat N, and Bradley AP. Rule extraction from support vector machines: A review. *Neurocomputing.* 2010;74(1):178-90.

158. Mittlböck M, and Schemper M. Explained Variation for Logisitic Regression. *Stat Med.* 1996;15(19):1987-97.
159. Orban T, et al. Pancreatic islet autoantibodies as predictors of type 1 diabetes in the Diabetes Prevention Trial–Type 1. *Diabetes care.* 2009;32(12):2269-74.
160. Mahmud SA, Manlove LS, and Farrar MA. Interleukin-2 and STAT5 in regulatory T cell development and function. *JAK-STAT.* 2013;2(1):e23154-e.
161. Gao J, Zhao L, Wan YY, and Zhu B. Mechanism of Action of IL-7 and Its Potential Applications and Limitations in Cancer Immunotherapy. *Int J Mol Sci.* 2015;16(5):10267-80.
162. Straub T, Freudenberg MA, Schleicher U, Bogdan C, Gasteiger G, and Pircher H. Bacterial coinfection restrains antiviral CD8 T-cell response via LPS-induced inhibitory NK cells. *Nature communications.* 2018;9(1):4117-.
163. Medrano RFV, Hunger A, Mendonça SA, Barbuto JAM, and Strauss BE. Immunomodulatory and antitumor effects of type I interferons and their application in cancer therapy. *Oncotarget.* 2017;8(41):71249-84.
164. Orozco G, et al. Analysis of a functional BTNL2 polymorphism in type 1 diabetes, rheumatoid arthritis, and systemic lupus erythematosus. *Human Immunol.* 2005;66(12):1235-41.
165. Arnett HA, Escobar SS, and Viney JL. Regulation of costimulation in the era of butyrophilins. *Cytokine.* 2009;46(3):370-5.
166. Swanson RM, et al. Butyrophilin-like 2 modulates B7 costimulation to induce Foxp3 expression and regulatory T cell development in mature T cells. *J Immunol.* 2013;190(5):2027-35.
167. Sinisalo J, et al. Novel 6p21.3 Risk Haplotype Predisposes to Acute Coronary Syndrome. *Circ Cardiovasc Genet.* 2016;9(1):55-63.
168. Valentonyte R, et al. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat Genet.* 2005;37:357.
169. Wolin A, et al. SNP Variants in Major Histocompatibility Complex Are Associated with Sarcoidosis Susceptibility-A Joint Analysis in Four European Populations. *Front Immunol.* 2017;8:422.
170. Rubicz R, et al. Genome-wide genetic and transcriptomic investigation of variation in antibody response to dietary antigens. *Genet Epidemiol.* 2014;38(5):439-46.
171. Fehrmann RS, et al. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet.* 2011;7(8):e1002197.
172. Baughcum AE, Johnson SB, Carmichael SK, Lewin AB, She J-X, and Schatz DA. Maternal Efforts to Prevent Type 1 Diabetes in At-Risk Children. *Diabetes Care.* 2005;28(4):916.
173. Perry DJ, et al. Application of a Genetic Risk Score to Racially Diverse Type 1 Diabetes Populations Demonstrates the Need for Diversity in Risk-Modeling. *Scientific reports.* 2018;8(1):4529.
174. Tosur M, and Redondo MJ. Heterogeneity of Type 1 Diabetes: The Effect of Ethnicity. *Curr Diabetes Rev.* 2018;14(3):266-72.

175. Moon S, Wang Y, Dennehy P, Simonsen KA, Zhang J, and Jiang B. Antigenemia, RNAemia, and innate immunity in children with acute rotavirus diarrhea. *FEMS Immunol Med Microbiol.* 2012;64(3):382-91.
176. Lee HS, et al. Next-generation sequencing for viruses in children with rapid-onset type 1 diabetes. *Diabetologia.* 2013;56(8):1705-11.
177. Kim KW, et al. Higher abundance of enterovirus A species in the gut of children with islet autoimmunity. *Scientific reports.* 2019;9(1):1749.

DANKSAGUNG

Zu allererst bedanke ich mich bei Prof. Dr. Anette-Gabriele Ziegler für die Möglichkeit meine Doktorarbeit an ihrem Institut durchzuführen. Vielen Dank für die Bereitstellung meines Promotionsthemas, die wissenschaftlichen Freiheiten, die vielen hilfreichen Diskussionen, die ermöglichten Teilnahmen an internationalen Kongressen und Laboraustauschen und natürlich für den Zugang zu einzigartigem Probenmaterial.

Ich bedanke mich ganz herzlich auch bei Prof. Dr. Dr. Fabian Theis für die Betreuung meiner Doktorarbeit als Zweitbetreuer, für das Feedback, die interessanten Diskussionen und die methodischen Denkanstöße in den Thesis Committee Meetings.

Außerdem gilt mein Dank meinem Mentor Prof. Dr. Ezio Bonifacio. Vielen Dank für das Interesse an meiner Arbeit, die geteilten Ansichten, die hilfreichen Diskussionen und die motivierende Art.

Vielen Dank an Dr. Thomas Briese, der mir den Forschungsaufenthalt in seinem Labor an der Columbia University in New York ermöglicht hat, für die Betreuung vor Ort und die Wissensvermittlung über die Grenzen meines Projekts hinaus.

Ich danke dem gesamten Team des Institutes für Diabetesforschung und den Mitwirkenden der DiMelli und BABYDIET Studien für die hervorragende Arbeitsatmosphäre. Besonderes danke ich Rainer Fürst und Andreas Beyerlein wie auch Manja, Nana, Christiane, Florian, Marlon, Lisa, Lorena und Cigdem für die wertvolle Unterstützung und die angenehme Zeit. Ein spezieller Dank gilt an dieser Stelle ebenfalls allen Studienbetreuern/innen, Ärzten/innen und natürlich den Studienteilnehmern/innen, die den Grundstein unserer Arbeit bilden. Außerdem bedanke ich mich bei der TEDDY Study Group für die Möglichkeit an diesem einmaligen Datensatz zu arbeiten.

Ein ganz besonderes Dankschön gilt an dieser Stelle meinen Homies von der Doktorandeninsel Lisa, Martin und Jan! Es war und ist mir eine Ehre mit euch zu arbeiten.

Jan,
ein spezieller Dank noch für dein Engagement für die Revision kurz vor Weihnachten!

Herzlichst bedanke ich mich bei meinen Eltern, meiner Schwester Laura und meiner Freundin Kate für die unendliche Unterstützung und den Freiraum, den ihr mir gebt. Danke, dass ihr immer für mich da seid und ich mich immer auf euch verlassen kann!

SCIENTIFIC PUBLICATIONS

Hippich, M., A. Oleynik, K. Jain, C. Winkler, R. C. Ferreira, E. Bonifacio, A. G. Ziegler and T. Briese (2018). "Searching peripheral blood mononuclear cells of children with viral respiratory tract infections preceding islet autoimmunity for viruses by high-throughput sequencing." Acta Diabetol 55(8): 881-884.

Hippich, M., A. Beyerlein, W. A. Hagopian, J. P. Krischer, K. Vehik, J. Knoop, C. Winker, J. Toppari, A. Lernmark, M. J. Rewers, A. K. Steck, J. X. She, B. Akolkar, C. C. Robertson, S. Onengut-Gumuscu, S. S. Rich, E. Bonifacio and A. G. Ziegler (2019). "Genetic Contribution to the Divergence in Type 1 Diabetes Risk Between Children From the General Population and Children From Affected Families." Diabetes. 2019 Jan 17. pii: db180882. doi: 10.2337/db18-0882. [Epub ahead of print]

Beyerlein A, Bonifacio E, Vehik K, **Hippich M**, Winkler C, Frohnert BI, Steck AK, Hagopian WA, Krischer JP, Lernmark Å, Rewers MJ, She JX, Toppari J, Akolkar B, Rich SS, Ziegler AG; TEDDY Study Group. “Progression from islet autoimmunity to clinical type 1 diabetes is influenced by genetic factors: results from the prospective TEDDY study.” J Med Genet. 2018 Oct 4. pii: jmedgenet-2018-105532. doi: 10.1136/jmedgenet-2018-105532. [Epub ahead of print]

Bonifacio E, Beyerlein A, **Hippich M**, Winkler C, Vehik K, Weedon MN, Laimighofer M, Hattersley AT, Krumsiek J, Frohnert BI, Steck AK, Hagopian WA, Krischer JP, Lernmark Å, Rewers MJ, She JX, Toppari J, Akolkar B, Oram RA, Rich SS, Ziegler AG; TEDDY Study Group. “Genetic scores to stratify risk of developing multiple islet autoantibodies and type 1 diabetes: A prospective study in children.” PLoS Med. 2018 Apr 3;15(4):e1002548. doi: 10.1371/journal.pmed.1002548. eCollection 2018 Apr.

Kindt ASD, Fuerst RW, Knoop J, Laimighofer M, Telieps T, **Hippich M**, Woerheide MA, Wahl S, Wilson R, Sedlmeier EM, Hommel A, Todd JA, Krumsiek J, Ziegler AG, Bonifacio E. “Allele-specific methylation of type 1 diabetes susceptibility genes.” J Autoimmun. 2018 May;89:63-74. doi: 10.1016/j.jaut.2017.11.008. Epub 2017 Dec 8.

Serr I, Scherm MG, Zahm AM, Schug J, Flynn VK, **Hippich M**, Kälin S, Becker M, Achenbach P, Nikolaev A, Gerlach K, Liebsch N, Loretz B, Lehr CM, Kirchner B, Spornraft M, Haase B, Segars J, Küper C, Palmisano R, Waisman A, Willis RA, Kim WU, Weigmann B, Kaestner KH, Ziegler AG, Daniel C. “A miRNA181a/NFAT5 axis links impaired T cell tolerance induction with autoimmune type 1 diabetes.” Sci Transl Med. 2018 Jan 3;10(422). pii: eaag1782. doi: 10.1126/scitranslmed.aag1782. Erratum in: Sci Transl Med. 2019 Feb 20;11(480):.

Knoop J, Gavrisan A, Kuehn D, Reinhardt J, Heinrich M, **Hippich M**, Eugster A, Ockert C, Ziegler AG, Bonifacio E. „GM-CSF producing autoreactive CD4+ T cells in type 1 diabetes.” Clin Immunol. 2018 Mar;188:23-30. doi: 10.1016/j.clim.2017.12.002. Epub 2017 Dec 8.

Ungethüm K, Jolink M, **Hippich M**, Lachmann L, Haupt F, Winkler C, Hummel S, Pitchika A, Kordonouri O, Ziegler AG, Beyerlein A. “Physical activity is associated with lower insulin and C-peptide during glucose challenge in children and adolescents with family background of diabetes.” Diabet Med. 2019 Mar;36(3):366-375. doi: 10.1111/dme.13819. Epub 2018 Oct 26.