AGE-RELATED ETIOLOGIC HETEROGENEITY IN THE ONSET OF ISLET AUTOIMMUNITY

by

BRITTNI NOEL FREDERIKSEN
B.A., Boston University, 2008
M.P.H., Tulane University School of Public Health and Tropical Medicine, 2010

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Epidemiology Program
2014
This thesis for the Doctor of Philosophy degree by

Brittni Noel Frederiksen

has been approved for the

Epidemiology Program

by

Marci Sontag, Chair

Jill M. Norris, Advisor

Anna Barón

Tessa Crume

Marian Rewers

Date 5/1/2014
ABSTRACT

Type 1 diabetes (T1D) is preceded by a pre-clinical phase of islet autoimmunity (IA) where the insulin-producing β-cells of the pancreas are destroyed and circulating autoantibodies can be detected. Autoantibodies can begin appearing in late infancy, with incidence peaks in infancy and adolescence. This thesis investigated whether early onset IA is associated with different risk factors than late onset IA by assessing age-related heterogeneity in a prospective cohort of children at increased genetic risk for T1D. Three methods were used to test and model non-proportional hazards, or age-related heterogeneity: the supremum test, weighted Schoenfeld residuals, and restricted cubic splines. Well-studied, but inconsistently associated determinants in the etiology of T1D were investigated for age-related heterogeneity in the development of IA: height and weight growth, timing of infant diet exposures, and vitamin D. Height growth velocity remained consistently associated with increased IA risk across childhood with slightly decreasing risk with increasing age, while weight growth velocity had no effect. Vitamin D intake did not exhibit an age-related effect. However, plasma 25(OH)D levels were associated with increased IA risk before age 3 years and then decreased risk after age 7.75 years. A non-HLA T1D candidate single nucleotide polymorphism (SNP), rs10517086, demonstrated an age-related effect on IA risk, with increased risk before age 2 years, but not older ages. Non-Hispanic white ethnicity and maternal age also demonstrated a significant age-related effect on IA risk, with increased risk early on that became protective with increasing age. While not exhibiting age-related heterogeneity, vitamin D metabolism SNPs, DHC7/NADSYN1 rs12785878 and CYP27B1 rs4646536, were associated with IA, while VDR rs1544410 interacted with PTPN2 rs1893217 in the risk of progression to T1D. With regard to T1D, both early and late first exposure to any solid food predicted development of T1D. Specifically, early exposure to fruit
and late exposure to rice/oat predicted T1D, while breastfeeding at introduction to wheat/barley
conferred protection. However, these infant diet exposures did not exhibit age-related
heterogeneity. Assessing heterogeneity in disease etiology is encouraged as it enables researchers
to identify associations that may lead to better understanding of complex chronic diseases, and
thus improve prevention efforts.

The form and content of this abstract are approved. I recommend its publication.

Approved: Jill M. Norris
To my mom for always valuing education
ACKNOWLEDGEMENTS

I would like to thank Jill Norris for her inspiration, dedication, and guidance. I am so grateful for the time we have spent together. I am thankful for the continual challenges and endless opportunities she has given me. Her mentorship has had a significant impact on my life. I would also like to thank Marci Sontag for being a constant source of inspiration, smiles, and comfort. I hope I can effortlessly share my light with others one day as she does every day. Thank you to Anna Barón for her statistical guidance and friendship. It has been an honor to have the opportunity to work with and learn from her. Thank you to Tessa Crume for sharing her knowledge and friendship with me. I hope we have the opportunity to work together in the future. Finally, I would like to thank Marian Rewers for his insight and guidance. I am continually impressed by his brilliance.

Last, but certainly not least, I would like to thank my family for their endless support, encouragement, and love. This would not have been possible without them.
# TABLE OF CONTENTS

## CHAPTER

I. **INTRODUCTION**

1.1 Background 1

1.2 Specific Aims and Hypotheses 2

1.3 Significance 3

II. **LITERATURE REVIEW**

2.1 Trends in Islet Autoimmunity and Type 1 Diabetes 6

2.2 Height and Weight Growth and Islet Autoimmunity 6

2.3 Timing of Infant Diet Exposures and Islet Autoimmunity 8

2.4 Vitamin D and Islet Autoimmunity 10

III. **METHODS**

3.1 Study Population 12

3.2 Definition of Outcome 12

3.3 Data Collection 13

3.3.1 Height and Weight Growth 13

3.3.2 Timing of Infant Diet Exposures 13

3.3.3 Vitamin D 14

3.3.3.1 Measurement of plasma 25(OH)D levels 14

3.3.3.2 Measurement of dietary intake of vitamin D 14

3.3.4 SNP Genotyping 15

3.4 Statistical Analysis 16

3.4.1 Assessment of the Proportional Hazards Assumption 16

3.4.2 Fixed Covariates 17

3.4.3 Time-varying Covariates 17

3.4.4 Cox Proportional Hazards Regression 17
II. ASSOCIATION BETWEEN VITAMIN D METABOLISM GENE POLYMORPHISMS AND RISK OF ISLET AUTOIMMUNITY AND PROGRESSION TO TYPE 1 DIABETES: THE DIABETES AUTOIMMUNITY STUDY IN THE YOUNG (DAISY)

2.1 Abstract

2.2 Introduction

2.3 Materials and Methods

2.3.1 Subjects

2.3.2 Measurement of Autoantibodies

2.3.3 Vitamin D SNP Genotyping

2.3.4 Measurement of Plasma 25(OH)D Levels

2.3.5 Analysis Population

2.3.6 Statistical Analyses

2.3.6.1 Analysis of Vitamin D SNPs and Risk of IA and T1D in IA Positive Children

2.3.6.2 Longitudinal Analysis of Plasma 25(OH)D Levels and Vitamin D SNPs

2.3.7 Results

2.3.7.1 Development of Islet Autoimmunity

2.3.7.2 Progression to T1D in Children with IA

2.3.7.3 Longitudinal Analysis of Plasma 25(OH)D Levels and Vitamin D SNPs

2.3.8 Discussion

III. EVIDENCE OF STAGE- AND AGE-RELATED HETEROGENEITY OF NON-HLA SNPS AND RISK OF ISLET AUTOIMMUNITY AND TYPE 1 DIABETES: THE DIABETES AUTOIMMUNITY STUDY IN THE YOUNG (DAISY)

3.1 Abstract
1.1 Introduction 67

1.2 Materials and Methods 69
   1.2.1 Subjects 69
   1.2.2 Measurement of Autoantibodies 69
   1.2.3 Non-HLA SNP Genotyping 70
   1.2.4 Analysis Population 70
   1.2.5 Statistical Analyses 71

1.3 Results 72
   1.3.1 Development of IA 72
   1.3.2 Progression to T1D in Children with IA 72
   1.3.3 Development of T1D 73

1.4 Discussion 77

1.5 Conclusions 82

II. ASSESSING AGE-RELATED ETIOLOGIC HETEROGENEITY IN THE ONSET OF ISLET AUTOIMMUNITY 83

2.1 Abstract 83

2.2 Introduction 83
   2.2.1 Supremum Test 85
   2.2.2 Weighted Schoenfeld Residuals 85
   2.2.3 Restricted Cubic Splines 85
   2.2.4 Diabetes Autoimmunity Study in the Young (DAISY) 86

2.3 Methods 86
   2.3.1 Study Population 86
   2.3.2 Height Growth Velocity 87
   2.3.3 Measurement of Dietary Vitamin D Intake 87
   2.3.4 Measurement of Plasma 25(OH)D Levels 88
LIST OF TABLES

TABLE

1 Demographic and Perinatal Characteristics of DAISY Birth Cohort by T1DM Status 27
2 Infant Dietary Exposure Characteristics of DAISY Birth Cohort by T1DM Status 28
3 Investigation of Independent Effects of Infant Diet Exposures on Risk of T1DM in DAISY Birth Cohort 29
4 Characteristics of Timing of Rice/Oat Introduction 32
5 Characteristics of DAISY Children With and Without Complete Data for the VDR and PTPN2 SNPs 40
6 Characteristics of IA Positive Children With and Without Complete Data for the VDR and PTPN2 SNPs 40
7 Genotype Frequencies of the Analysis Population for Development of Islet Autoimmunity (IA) 42
8 Genotype Frequencies of Islet Autoantibody (IA) Positive Subjects by Type 1 Diabetes (T1D) Status 42
9 Characteristics of the Analysis Population by Islet Autoimmunity (IA) Status 43
10 Association Between Variants in VDR and PTPN2 and Risk of Islet Autoimmunity 44
11 Characteristics of Islet Autoantibody (IA) Positive Subjects by Type 1 Diabetes (T1D) Status 45
12 Association Between Variants in VDR and PTPN2 and Risk of Progression to Type 1 Diabetes in IA Positive Children 47
13 Characteristics of the Analysis Sample by IA Status and Association of Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population 58
14 Observed Number of Cases with Endpoint and Number of Person-Years for Each Genotype of Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population 59
15 Characteristics of IA-Positive Subjects by T1D Status and Association of Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population 60
16 Plasma Vitamin D Levels Predicted by Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population 61
17 Mean 25(OH)D Levels for Each Genotype of Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population

18 Demographic Characteristics of DAISY Non-Hispanic White Population

19 Association Between Non-HLA T1D Candidate SNPs and Development of IA, Progression from IA to T1D, and Development of T1D Unadjusted for HLA-DR3/4, DQB1*0302 and First Degree Relative with T1D

20 Association Between Non-HLA T1D Candidate SNPs and Development of IA, Progression from IA to T1D, and Development of T1D Adjusted for HLA-DR3/4, DQB1*0302 and First Degree Relative with T1D

21 Non-HLA T1D Candidate SNPs Associated with Development of IA, Progression from IA to T1D, and/or Development of T1D in DAISY

22 Characteristics of DAISY Cohort by IA Status

23 Assessment of the Proportional Hazards Assumption in DAISY Cohort
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Analysis Populations Based on Stages of the Disease Process</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Study Participants</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>First Solid Food Introduced to 1835 Diabetes Autoimmunity Study in the Young Infants</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Flow Chart Illustrating Formation of Analysis Cohorts from the DAISY Study Population</td>
<td>41</td>
</tr>
<tr>
<td>5</td>
<td>Association Between VDR rs1544410 and Risk of Progression to T1D in IA Positive Children by PTPN2 rs1893217 Genotype</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>Association Between SNP rs10517086 and Development of IA Modeled Using a Restricted Cubic Spline</td>
<td>77</td>
</tr>
<tr>
<td>7</td>
<td>Weighted Schoenfeld Residual Plot for the Model with HLA-DR3/4, DQB1*0302 Genotype and First-Degree Relative (FDR) with Type 1 Diabetes (T1D) Status</td>
<td>95</td>
</tr>
<tr>
<td>8</td>
<td>Weighted Schoenfeld Residual Plots for Non-Hispanic White Ethnicity (NHW) and a 5 Year Difference in Maternal Age</td>
<td>96</td>
</tr>
<tr>
<td>9</td>
<td>Restricted Cubic Splines for the High Risk HLA-DR3/4, DQB1*0302 Genotype and First-Degree Relative with Type 1 Diabetes Status</td>
<td>97</td>
</tr>
<tr>
<td>10</td>
<td>Restricted Cubic Splines for Non-Hispanic White Ethnicity and a 5 Year Difference in Maternal Age</td>
<td>98</td>
</tr>
<tr>
<td>11</td>
<td>Restricted Cubic Splines for Height Growth Velocity, Vitamin D Intake from Food and Supplements, and Plasma 25(OH)D</td>
<td>99</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

CI: CONFIDENCE INTERVAL
C1QTNF6: C1Q AND TUMOR NECROSIS FACTOR RELATED PROTEIN 6
C6ORF173: CHROMOSOME 6 OPEN READING FRAME 173
C14ORF181: CHROMOSOME 14 OPEN READING FRAME 181
CTLA4: CYTOTOXIC T-LYMPHOCYTE-ASSOCIATED PROTEIN 4
CYP27B1: CYTOCHROME P450 FAMILY 27 SUBFAMILY B PEPTIDE 1
DAISY: DIABETES AUTOIMMUNITY STUDY IN THE YOUNG
DHCR7/NADSYN1: DEHYDROCHOLESTEROL REDUCTASE/NICOTINAMIDE-ADENINE DINUCLEOTIDE SYNTHETASE 1
HLA: HUMAN LEUKOCYTE ANTIGEN
HR: HAZARD RATIO
IA: ISLET AUTOIMMUNITY
IFIH1: INTERFERON INDUCED WITH HELICASE C DOMAIN 1
IL2: INTERLEUKIN 2
IL2RA: INTERLEUKIN 2 RECEPTOR, ALPHA
IL7R: INTERLEUKIN 7 RECEPTOR
INS: INSULIN
PRKCQ: PROTEIN KINASE C, THETA
PTPN2: PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 2 GENE
PTPN22: PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 22
SKAP2: SRC KINASE ASSOCIATED PHOSPHOPROTEIN
SMARCE1: SWI/SNF RELATED, MATRIX ASSOCIATED, ACTIN DEPENDENT REGULATOR OF CHROMATIN, SUBFAMILY E, MEMBER 1
SNP: SINGLE NUCLEOTIDE POLYMORPHISM
T1D: TYPE 1 DIABETES
TLR8: TOLL-LIKE RECEPTOR 8
UBASH3A: UBIQUITIN ASSOCIATED AND SH3 DOMAIN CONTAINING A
VDR: VITAMIN D RECEPTOR GENE
CHAPTER I
INTRODUCTION

Background

Type 1 diabetes (T1D) is a chronic autoimmune disease in which the insulin-producing beta cells of the pancreas are destroyed. There is typically a preclinical phase of circulating autoantibodies, called islet autoimmunity (IA) that precedes the clinical diagnosis of T1D. T1D is widely believed to be caused by an environmental factor on a susceptible genetic background. A number of genetic markers, as well as environmental exposures, have been implicated in the etiology of T1D. However, the cause of T1D remains unknown. In order to better understand when these determinants play a role in the etiology of T1D, age-related heterogeneity, or identifying determinants that influence the disease process at certain ages, can be explored.

Islet autoimmunity appears to exhibit two incidence peaks, the first around age 1-2 years with a second possible peak in adolescence. Distinct characteristics of the autoantibodies at each incidence peak have been reported (1). There are well-documented differences in autoantibody frequency and levels among patients manifesting T1D at various ages (2–4), suggesting heterogeneity in the pathogenesis of T1D. Age of diagnosis of T1D is strongly correlated with age of appearance of first autoantibody and autoantibody levels (4). HLA genes associated with disease susceptibility are more frequent in patients diagnosed during childhood than in those diagnosed during adulthood (5–11). Children developing T1D between the ages of 1 and 2 years have also been found to be significantly more likely to have a parental history of T1D than older children (9). Interestingly, different serum metabolite profiles were found to be associated with the age of IA development, including a reduction of methionine concentration from around seroconversion in children with early but not late IA (12). This study investigates age-related heterogeneity of three well-studied, but inconsistently associated determinants of T1D in order to more fully understand their mechanism in the etiology of T1D: height and weight growth, timing of infant diet exposures, and vitamin D.
Specific Aims and Hypotheses

The overall aim of the study was to more fully understand the mechanism of three well-studied, but inconsistently associated determinants in the etiology of type 1 diabetes by investigating their age-related heterogeneity in the development of islet autoimmunity (IA): height and weight growth, infant diet exposures, and vitamin D.

This research had three specific aims:

1. To evaluate whether height and weight growth are associated with increased risk of IA in late onset IA, but not in early onset IA.

   **Hypothesis 1:** Height and weight growth are associated with increased risk of IA in late onset IA, but not in early onset IA.

2. To evaluate whether timing of infant diet exposures is associated with an increased risk of IA in early onset IA, but not in late onset IA.

   **Hypothesis 2:** Early (< 4 months) and late (≥ 6 months) first exposures to selected foods in the infant diet is associated with an increased risk of IA in early onset IA, but not late onset IA.

3. To evaluate whether vitamin D, including reported dietary intake of vitamin D, plasma 25(OH)D levels, and related single nucleotide polymorphisms (SNPs), is more protective in early onset IA compared to late onset IA.

   **Hypothesis 3.1:** Dietary intake of vitamin D and plasma 25(OH)D levels are inversely associated with IA in early onset IA, but not late onset IA.

   **Hypothesis 3.2:** Variation in vitamin D pathway genes is associated with risk of IA at all ages.
Significance

This study enters an area of research in the etiology of T1D, age-related heterogeneity, which has not been prospectively examined up to this point. The first prospective cohort study of the etiology of T1D only began in 1989 in Germany, making the oldest subject at most 23 years old. Puberty, or 12-18 years of age, is the time of peak T1D incidence. Since 1993, DAISY has studied two large cohorts of children at high risk for T1D: 1) 1,424 general population newborns with high-risk HLA-DR, DQ alleles with a current median age of 15.3 years; and 2) 1,123 young non-diabetic relatives of patients with T1D with a current median age of 12.2 years. The current median age of the DAISY cohort should bring rise to a number of new T1D cases as they enter puberty, allowing this study to compare risk factors for early- vs. late-onset IA and T1D, or age-related etiologic heterogeneity of T1D. This thesis is one of the first studies to investigate age-related heterogeneity of prospectively collected determinants in the etiology of T1D.

There are very few studies that have explored age-related heterogeneity in the development of the pre-clinical phase of T1D, islet autoimmunity. To our knowledge, none of the major T1D prospective cohort studies have explored age-related heterogeneity of height and weight growth and vitamin D and their association with IA. Additionally, there currently appears to be only one T1D prospective cohort study that has published evidence of age-related heterogeneity of infant diet exposures. Therefore, this thesis has immense potential to provide valid results pertaining to age-related heterogeneity of the three proposed exposures that the existing studies have not investigated.

Not only was this study able to examine age-related heterogeneity of disease determinants across childhood and the natural history of T1D, but we were also able to combine prospective dietary data with prospectively collected dietary biomarkers and non-HLA gene variants. As it becomes increasingly clear that T1D is a complex chronic disease, with multiple genetic and environmental risk factors playing roles at different critical time-points in the disease’s natural history, it will be necessary to combine prospective dietary data with
prospectively collected dietary biomarkers and non-HLA gene variants in order to more fully understand metabolic pathways that cannot be explained by dietary data, biomarkers, or non-HLA gene variants alone. This research took well-studied determinants in the etiology of T1D and examined them on a new level of age-related etiology heterogeneity.

**Figure 1: Analysis Populations Based on Stages of the Disease Process**

DAISY was the first study that made the leap from studying the pathogenesis of T1D in relatives of affected patients to a large-scale general population cohort. It is only one of a few studies that have been following a genetically at-risk cohort from birth for the development of T1D and its preclinical asymptomatic phase of IA. While children are normally seen in clinic annually, children positive for any autoantibody are followed in 3-6 month intervals. Due to the change in frequency of clinic visits and the suspected behavior changes that occur after a child is informed of islet autoantibody positivity, we have separated our cohort into three separate analysis populations based on stages of the disease process displayed in **Figure 1**: 1) Birth to IA, 2) IA to T1D, and 3) Birth to T1D. This thesis will focus on the preclinical phase of T1D, development of IA, with the goal of finding interventions to slow the overall progression of T1D. These methods will allow us to look at age-related etiologic heterogeneity of T1D across childhood, which will be novel in T1D research. This thesis will contribute to DAISY’s role as a pivotal vanguard project to The Environmental Determinants of Diabetes in the Young study (TEDDY) by assisting with the prioritization of data collection protocols, use of stored samples, and particularly analysis strategies. Due to the young age of the cohort, TEDDY will not be able
to address the questions posed in this thesis for another 10 years. The findings of this study may have a profound effect on design of future screening and prevention programs.
CHAPTER II

LITERATURE REVIEW

Trends in Islet Autoimmunity and Type 1 Diabetes

The incidence of T1D is increasing worldwide with an average relative increase in incidence of 3% to 4% per year (13–18). This increase in incidence appears to be greatest in children < 5 years of age (13,14,16,18). The incidence of T1D varies widely among countries with the highest rates in Finland (> 60 per 100,000/year) (13,14,19). Using data from the Colorado Insulin-Dependent Diabetes Mellitus (IDDM) Study Registry (1978-1988) and SEARCH for Diabetes in Youth (2002-2004), Vehik et al. (2007) found the incidence of T1D to be 14.8 per 100,000/year in 1978-1988 and 23.9 per 100,000/year in 2002-2004 among youth aged 0-17 in the state of Colorado (20). From 1978-2004, the incidence of T1D increased by 2.3% (1.6-3.1) per year (20). This increase in incidence was significant for both non-Hispanic white and Hispanic youth (20). The cause of this increase in incidence remains unknown, however, it appears to be occurring in the very young and those with moderate genetic susceptibility, which may implicate an early life or in utero environmental exposure (21).

The preclinical phase of T1D, islet autoimmunity, can begin very early in life, with those who develop IA early, subsequently developing T1D in early childhood (22). Children who develop autoantibodies within the first 2 years of life are more likely to develop multiple islet autoantibodies and progress to T1D during childhood (22). It has also been shown that children who develop autoantibodies later develop multiple antibodies and T1D more slowly (22). The German BABYDIAB cohort has observed two peaks in autoimmunity incidence around 1 to 2 years of age and around puberty that appear to exhibit different characteristics (1), which may imply age-related heterogeneity of disease determinants in the etiology of T1D.

Height and Weight Growth and Islet Autoimmunity

Growth has long been of interest in the development of IA and progression to T1D. The ‘overload hypothesis’ and the ‘accelerator hypothesis’ are two common hypotheses that suggest
the current childhood obesity epidemic is driving the increasing incidence and earlier age of T1D onset. The ‘overload hypothesis’ proposes that the high insulin demand on the beta cell that results from the overfeeding and resultant accelerated growth of children and adolescents makes the beta cells vulnerable to autoimmune attack and apoptosis (23). The ‘accelerator hypothesis’ suggests that insulin resistance caused by excess weight gain may accelerate beta cell apoptosis in individuals at genetic risk (24,25). The increasing incidence of T1D and the shift toward a younger age of onset suggest accelerated disease progression (23). Risk factors that may accelerate beta cell destruction, such as child growth, weight, and birth weight, have shown a steady increase in the population (23).

Descriptive epidemiological studies from around the world have shown the incidence of T1D to peak during puberty in both boys and girls (23). It has been suggested that this incidence peak in puberty could be due to the high growth rate during puberty causing a relative increase in insulin needs as a result of the effect of growth hormone upon insulin resistance (23,26). Prospective growth data from a population-based study showed that children who developed T1D were consistently taller than age-matched controls for a number of years before the onset of T1D (23). A study of Colorado young people recently showed that increasing height over time accounted for 15% of the decreasing age at diagnosis with T1D (27). Both puberty and high linear growth at any age has been associated with an increased risk of childhood T1D (23). Several prospective studies have more recently confirmed the association with height and have also shown height, weight, and BMI to be associated with increased risk (28,29).

Obesity has also been proposed as a predictor of T1D in addition to increased linear growth (23). Knowing that the rise in incidence of T1D, especially in the younger ages, is most likely due to an environmental factor accelerating the disease process, the increasing obesity epidemic seems a plausible culprit. Children are experiencing obesity and insulin resistance early in childhood, which may be programmed as early as the intrauterine environment (30).
Growth depends on genetic, nutritional, and hormonal factors that should all be considered when looking at age-related heterogeneity of growth and its association with the development of IA and T1D. DAISY has the unique ability to evaluate the genetic, nutritional, and hormonal factors that may contribute to an association between growth and development of IA and T1D.

DAISY has previously looked at height, weight, BMI and velocities of growth in height, weight and BMI, for association with development of IA and T1D (31). Greater height growth velocity was associated with IA development and progression to T1D in IA positive children (31). However, these analyses were limited to children < 11.5 years of age (31). There is a need to further explore this association in pubertal children. This is an opportune time to begin to explore risk factors that may be important in the development of IA or progression to T1D during puberty as the median age of the DAISY population is 12.8 years. With the cohort getting older, it will make it easier to examine age-related heterogeneity of risk factors, particularly growth.

**Timing of Infant Diet Exposures and Islet Autoimmunity**

Given that T1D, and its preclinical IA, appear early in life, infant and childhood diet have been implicated as potential initiating exposures in the etiology of the disease. However, findings have been contradictory and inconsistent. Breastfeeding has been of particular interest. Of four cohort studies examining IA as the outcome, three showed no effect of the duration of exclusive breastfeeding (32,33) or age at exposure to cow’s milk (33,34) on the risk of IA, and one showed that short-term exclusive breastfeeding and early exposure to cow’s milk increased risk for IA (35). These discrepant findings may be explained by the fact that exposure to cow’s milk may be correlated with the actual diabetogenic exposure in some populations but not in others. Studies of other foods in the infant diet also have been contradictory. With regard to introduction of solid foods, two studies found T1D cases had been exposed to solid foods earlier than controls (36,37), two studies found no association (38,39), and one study reported that T1D cases had been exposed to solid foods later than controls (40). A cross-sectional study of children with a family
history of T1D found no association between exposure to cereals before age 6 months and IA (33).

In 2003, however, Norris et al. found early (0-3 months) and late (≥ 7 months) introduction of cereals when compared with first introduction in the 4th to 6th month of life to be associated with development of IA (41). Age at exposure of 4 to 6 months was chosen as the reference group based on the general recommendation by US pediatricians to introduce solid foods, especially cereals, between the ages of 4 and 6 months. The bimodal nature of this association would make it easy to miss with conventional analyses that compared mean age at exposure or use just one age cutoff for exposure (41). Ziegler et al. (2003) found a similar increased risk of IA with early (0-3 months) introduction of gluten-containing foods (42). Recently, Virtanen et al. (2011) found early introduction to root vegetables (≤ 4 months) associated with increased risk of IA in a prospective population-based cohort in Finland (43). While different first solid food exposures have been found to be associated with increased risk of IA, it may not be the type of infant food exposure, but the timing of first solid food exposures. Cereals happen to be the first food most frequently introduced in the United States and root vegetables are often the first food introduced in Finland.

With evidence that the timing of certain infant diet exposures is associated with development of IA and T1D, it is important to explore age-related heterogeneity of these exposures. To our knowledge, there is only one T1D prospective cohort study that has published evidence of age-related heterogeneity of infant diet exposures. Virtanen et al. found that early introduction of gluten and egg (< 4 months of age) was associated with an increased risk of IA, but only during the first 3 years of life (43). In DAISY, we found similar results where early introduction of cereal (< 3 months of age) was associated with an increased risk of IA in the first three years, but not after three years (44). Based on these two studies, it appears that early introduction of solid foods may only be important for increased IA risk in the first three years of life. However, all infant diet exposures need to be investigated for age-related heterogeneity to
see if this trend is consistent across diet exposures. This study allowed us to explore age-related heterogeneity of infant diet exposures.

**Vitamin D and Islet Autoimmunity**

Vitamin D has been another environmental factor of interest in preventing T1D because of its immunosuppressive or immunomodulating effects. After 1,25-dihydroxyvitamin D was found to prevent autoimmune diabetes in nonobese diabetic mice, a model for human T1D, two epidemiologic studies showed associations between the use of vitamin D supplements in the first year of life and a lower risk of T1D (45,46).

The EURODIAB substudy 2 assessed whether the protective effect of vitamin D supplementation was different for T1D onset in different age groups and Mantel-Haenszel pooled odds ratios were calculated for age at onset before 5 years (Odds Ratio (OR): 0.83, 95% CI: 0.50-1.40), between 5 and 9 years (OR: 0.81, 95% CI: 0.55, 1.20) and between 10 and 14 years (OR: 0.47, 95% CI: 0.33, 0.68), with the effect most obvious in the oldest onset patients (45). This case-control study hinted at a heterogeneous age-related etiologic effect of vitamin D, which was explored in DAISY through this thesis.

DAISY investigated the association between vitamin D intake and 25(OH)D levels and the risk of IA and subsequent T1D in children at increased risk of T1D and found that intake of vitamin D was not associated with the risk of IA (adjusted HR: 1.13, 95% CI: 0.95, 1.35, P value = 0.18) nor progression to T1D in IA positive children (adjusted HR: 1.30, 95% CI: 0.91, 1.86, P value = 0.15) (47). Moreover, 25(OH)D levels were not associated with the risk of IA (adjusted HR: 1.12, 95% CI: 0.88, 1.43, P value = 0.36), nor progression to T1D in IA positive children (adjusted HR: 0.91, 95% CI: 0.68, 1.22, P value = 0.54) (47). This thesis explored age-related heterogeneity of vitamin D intake and 25(OH)D levels and the risk of IA in the DAISY cohort. With a cohort of both early and late converters to IA and T1D, age-related heterogeneity of vitamin D was explored across childhood.
In 2010, a genome-wide association study (GWAS) in approximately 30,000 individuals of European descent identified variants at four loci that were associated with 25(OH)D levels: GC rs2282679, DHCR7 rs12785878, CYP2R1 rs10741657, and CYP24A1 rs6013897 (48). A second GWAS of 25(OH)D levels confirmed the findings with GC, DHCR7, and CYP2R1 (49). These variants are located within or near genes involved in vitamin D transport (GC), cholesterol synthesis (DHCR7), and hydroxylation (CYP2R1 and CYP24A1) (48). Cooper et al. (50) more recently tested genetic variants influencing 25(OH)D metabolism for an association with both circulating 25(OH)D concentrations and T1D. They replicated the associations found in the aforementioned GWAS of the four vitamin D metabolism genes (GC, DHCR7, CYP2R1, and CYP24A1) with 25(OH)D in controls subjects (48,50) and found that CYP27B1, DHCR7, and CYP2R1 were associated with T1D (2). CYP27B1 had previously been associated with T1D in 2004 (51).

The DAISY cohort was genotyped for 12 SNPs in genes associated with metabolism of vitamin D, including DHCR7/NADSYN1, GC, CYP2R1, CYP27B1, CYP24A1, and VDR. These SNPs in vitamin D pathway genes were investigated for association with risk of IA. The vitamin D receptor gene (VDR) has been an inconsistently associated candidate gene for T1D. Protein tyrosine phosphatase, non-receptor type 2 (PTPN2), also associated with T1D, contains a novel intronic binding site for VDR. We explored the association between VDR and PTPN2 variants, as well as their interaction with both risk of IA and progression to T1D. This targeted investigation of genes in the vitamin D pathway suggests that the picture is more complex than simply 25(OH)D levels, and we explored the vitamin D pathway in terms of IA risk. Having multiple SNPs in the vitamin D metabolism pathway, as well as prospectively collected measures of 25(OH)D, aids in putting together the pieces of the puzzle regarding the vitamin D metabolism pathway in the development of T1D.
CHAPTER III

METHODS

Study Population

DAISY is a prospective study composed of two groups of children at increased risk for T1D who were recruited between 1993 and 2004 and are being followed prospectively for the development of IA and T1D. One group is made up of first degree relatives of patients with T1D, recruited between birth and eight years of age. The second group consists of infants born at St. Joseph’s Hospital in Denver, Colorado, whose umbilical cord blood was screened for diabetes-susceptibility genotypes in the HLA region. Based on their HLA genotype, newborns were categorized into three risk groups determined by the odds of developing T1D by the age of 20 years: high, odds of developing T1D by the age of 20 years, 1:16; moderate, 1:75 in non-Hispanic whites or 1:230 in Hispanics; or low, less than 1:300. All newborns found to be at high risk and a sample of those found to be at moderate risk were asked to participate in the follow-up. The St. Joseph’s Hospital newborn population is representative of the general population of the Denver metropolitan area. Details of the newborn screening, sibling and offspring recruitment, and follow-up of both cohorts have been published previously (52,53). Cord blood or the first available blood sample (depending on enrollment group) was sent to Roche Molecular Systems, Inc., Alameda, CA for PCR-based HLA class II typing. All study protocols were approved by the Colorado Multiple Institutional Review Board, and informed consent was given by parents of all participating children.

Definition of Outcome

Autoantibodies were tested at 9, 15, and 24 months, and annually thereafter, or at their first visit and annually thereafter if the child enrolled after birth. Radio-immunoassays were used to measure serum autoantibodies to insulin, GAD-65, and IA-2 (BDC512), as previously described (54–57), with rigorous confirmation of all positive and a subset of negative results.
The cut-off for positivity was established as the 99\textsuperscript{th} percentile of healthy controls. Children who tested autoantibody positive were put on an accelerated testing schedule of every 3-6 months.

Cases of IA were defined as those children positive for at least one islet autoantibody (IAA, GAD-65, IA-2) on two or more consecutive visits. Type 1 diabetes was diagnosed by a physician and defined according to the criteria for the diagnosis of diabetes by the American Diabetes Association: HbA\textsubscript{1c} (A1C) \(\geq 6.5\%\) or fasting plasma glucose (FPG) \(\geq 126\, \text{mg/dL (7.0 mmol/L)}\) or 2-hour plasma glucose \(\geq 200\, \text{mg/dL (11.1 mmol/L)}\) during an oral glucose tolerance test (OGTT) or in a patient with classic symptoms of hyperglycemia or hyperglycemic crisis, a random plasma glucose \(\geq 200\, \text{mg/dL (11.1 mmol/L)}\) (58).

**Data Collection**

**Height and Weight Growth**

Weight was measured at every clinic visit on a scale with a precision of \(\pm 0.1\, \text{kg}\). Height was first measured when the child was able to stand cooperatively, about 2 years old, and at every clinic visit thereafter, using a stadiometer with a precision of \(\pm 1\, \text{mm}\). Length before age 2 was inconsistently collected due to the difficulty of collecting accurate length measurements in infants and therefore, was not analyzed. BMI was calculated as weight (kg)/height (m)\(^2\) for all clinic visit where the child was at least 2 years old.

**Timing of Infant Diet Exposures**

Data for infant diet were collected during telephone or face-to-face interviews at 3, 6, 9, 12, and 15 months of age. At each interview, mothers were asked to report the date of introduction and frequency of exposure (i.e., number of servings per day) of all milks, formulas, and foods the infants consumed during the previous 3 months. Exclusive breastfeeding duration was determined by the reported age at which the infant was exposed to any foods or liquids other than breast milk or water.
**Vitamin D**

*Measurement of plasma 25(OH)D levels*

Blood was drawn and kept from light at all times during processing. Plasma was separated immediately, snap-frozen in liquid nitrogen, and stored at -70°C until it was sent to the University of Colorado Pediatric Clinical Translational Research Center Core Laboratory, which has a certificate of proficiency from the Vitamin D External Quality Assessment Scheme. 25-hydroxyvitamin D3 was measured by radioimmunoassay (DiaSorin, Stillwater, MN, USA), with a CV of 7.5%. Quality control was assessed via assay of blinded duplicate samples, and excellent agreement (intraclass correlation coefficient of 0.92 in 229 pairs) was observed. 25(OH)D levels are reported in ng/mL.

*Measurement of dietary intake of vitamin D*

Current dietary intake of vitamin D was assessed using one of two instruments. For children aged 2-9 years of age, parents filled out a food frequency questionnaire (FFQ) (Nutrition Questionnaire Service Center, Boston, MA, USA) on the child’s behalf annually. Starting between the ages of 10 and 12 years, children were asked to recall their own diets and complete the youth/adolescent questionnaire (YAQ), a food frequency questionnaire geared toward adolescents that is based on the FFQ. Both surveys collected the average food consumption over the course of the previous years, using structured responses regarding how often a child consumes commonly sized portions (example: one slice of bread), with the response options ranging from ‘never’ to ‘6+ servings per day.’ Fluid cow’s milk is a major source of dietary vitamin D in US children, because it is routinely fortified to a level of 400 IU per 0.95 l. We conducted an instrument-comparison study in our DAISY population, in which we determined that data from these two instruments may be combined when an instrument indicator variable (i.e. type of food frequency questionnaire [FFQ vs. YAQ] was included in the model. The FFQ has been validated against multiple 24 h recalls and has been found to correlate with micronutrient and fatty acid biomarkers in our population. Dietary data were linked to an autoantibody measurement if the 1-
year time period of the questionnaire encompassed the time directly preceding the clinic visit at which the autoantibody was measured.

**SNP Genotyping**

DAISY children were genotyped for VDR rs1544410 (BsmI), VDR rs2228570 (FokI), VDR rs11568820 (Cdx2), PTPN2 rs1893217, PTPN2 rs478582, DHCR7/NADSYN1 rs12785878, CYP2R1 rs10741657, CYP2R1 rs12794714, CYP24A1 rs6013897, GC rs4588, GC rs7041, CYP27B1 rs4646536, C1QTNF6 rs229541, C6orf173 rs9388489, C14orf181 rs1465788, IL2 rs2069762, IL2 rs4505848, IL2RA rs12722563, IL2RA rs2104286, IL7R rs6897932, PRKCCQ rs947474, SKAP2 rs7804356, SMARCE1 rs7221109, TLR8 rs5979785, UBAH3A rs3788013, and SNP rs10517086. VDR rs1544410 was genotyped using a linear array method at Roche Molecular Systems, Inc., as described in Mirel et al. (2002). PTPN2 rs1893217 was genotyped as part of a second project using Illumina 48-plex VeraCode technology following the manufacturer’s protocol. Genotyping data analysis and clustering was performed in Illumina’s GenomeStudio. Clustering clouds were manually investigated and adjusted if necessary. All plates included one duplicate sample and one positive control. VDR rs2228570, VDR rs11568820, VDR rs478582, DHCR7/NADSYN1 rs12785878, CYP2R1 rs10741657, CYP2R1 rs12794714, CYP24A1 rs6013897, GC rs4588, GC rs7041, CYP27B1 rs4646536, C1QTNF6 rs229541, C6orf173 rs9388489, C14orf181 rs1465788, IL2 rs2069762, IL2 rs4505848, IL2RA rs12722563, IL2RA rs2104286, IL7R rs6897932, PRKCCQ rs947474, SKAP2 rs7804356, SMARCE1 rs7221109, TLR8 rs5979785, UBAH3A rs3788013, and SNP rs10517086 were typed utilizing the Taqman SNP genotype-based OpenArray platform [Applied Biosystems, CA, USA]. Custom designed arrays were loaded using the OpenArray AccuFill system and cycling was performed on a GeneAmp 9700 PCR system [Applied Biosystems, CA, USA], all according to manufacturer’s protocol. Alleles were analyzed using the OpenArray SNP genotyping analysis software v.1.0.3 and Taqman Genotyper Software 2.0 [Applied Biosystems, CA, USA]. All 26 SNPs had a 95% call rate or higher and were in Hardy-Weinberg equilibrium.
Statistical Analysis

Assessment of the Proportional Hazards Assumption

As a preliminary step preceding assessment of PH, continuous variables were examined for correct functional form using martingale residuals with all important covariates and confounders included in the model. We assessed violation of the PH assumption using three different methods: the supremum test, weighted Schoenfeld residuals, and restricted cubic splines. The supremum test was performed in SAS version 9.3 (SAS Institute, Cary, NC) using the ASSESS statement in the PHREG procedure. Weighted Schoenfeld residuals were plotted using the cox.zph function in R 2.15.2 (59). A significant supremum test ($P < 0.30$) or a non-zero slope for the loess smoothed curve of the weighted Schoenfeld residuals indicated a violation of PH (60). Using the weighted Schoenfeld residuals, a global test of PH was assessed first; if the global test $P$ value was not large ($< 0.30$), the individual covariate tests of PH were used to identify the source(s) of the non-PH. Restricted cubic splines were modeled using the RCS macro in SAS (61). Due to a limited number of events, the number of knots for the RCS was selected to be 3, placed at the 5th, 50th, and 95th percentiles of age of the IA cases; this minimizes the number of coefficients to fit the RCS models (61). The RCS macro provides the three statistical tests described above which should be performed hierarchically. The first test has 3 df for a 3-knot spline model and tests whether the covariate of interest is associated with the event. If the null hypothesis is rejected ($P < 0.05$), the second statistical test with 2 df for a 3-knot spline model can be performed to determine whether the association is non-constant with time ($P < 0.05$ indicating a violation of PH). Finally, if the null hypothesis is rejected for both the first and second statistical tests, the third statistical test with 1 df for a 3-knot spline model can be performed to determine whether the relationship between the HR and time is linear ($P < 0.05$ indicates non-linearity) (61).
Fixed Covariates

All statistical analyses were conducted using SAS software, Version 9.3 of the SAS System for Windows. Copyright © 2002-2010 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA. The only exception was the weighted Schoenfeld residuals, which were generated and plotted in R (62). The three methods (supremum test, weighted Schoenfeld residuals, and RCS) were evaluated for the fixed covariates: HLA-DR3/4, DQB1*0302 genotype, FDR status, NHW ethnicity, and maternal age.

Time-varying Covariates

Currently available tools are limited with regard to examination of PH with time-varying covariates. The supremum test for violation of the PH assumption can theoretically accommodate time-varying covariates, but would require higher dimensional plots for time-varying covariates. The cox.zph function used to plot the weighted Schoenfeld residuals is valid for time-varying covariates; however, the software has been written to assume the variance of the time-varying covariate is constant over time (60). If this assumption is violated, results from the weighted Schoenfeld residual tests are not reliable. Motivated by these limitations in assessing PH with time-varying covariates, Kroehl et al. (unpublished) recently adapted RCS for use with time-varying covariates and evaluated their performance in identifying and modeling a non-constant HR. Using this approach, non-PH were investigated for height growth velocity and vitamin D intake from food and supplements.

Cox Proportional Hazards Regression

If a variable exhibited proportional hazards, hazard ratios and 95% confidence intervals were estimated using Cox proportional hazards regression to account for right-censored data. Clustered time-to-event analyses were performed treating siblings from the same family as a cluster, and robust sandwich variance estimates (63) were used for statistical inference. The
significance threshold was defined \textit{a priori} as \( \alpha < 0.05 \). Because our analyses were based on \textit{a priori} hypotheses, \( P \) values were not corrected for multiple testing.
CHAPTER IV

INFANT EXPOSURES AND DEVELOPMENT OF TYPE 1 DIABETES MELLITUS:
THE DIABETES AUTOIMMUNITY STUDY IN THE YOUNG (DAISY)1

Abstract

The prevalence of type 1 diabetes mellitus (T1DM) is increasing worldwide, with the most rapid increase among children younger than 5 years of age. The objective of this study was to examine the associations between perinatal and infant exposures, especially early infant diet, and the development of T1DM. The Diabetes Autoimmunity Study in the Young (DAISY) is a longitudinal, observational study. Newborn screening for human leukocyte antigen (HLA) was done at St. Joseph’s Hospital in Denver, Colorado. First-degree relatives of individuals with T1DM were recruited from the Denver metropolitan area. A total of 1835 children at increased genetic risk for T1DM followed up from birth with complete prospective assessment of infant diet. Fifty-three children developed T1DM. Early (<4 months of age) and late (≥6 months of age) first exposure to solid foods was compared with first exposures at 4 to 5 months of age (referent). Risk for T1DM was diagnosed by a physician. Both early and late first exposure to any solid food predicted development of T1DM (hazard ratio [HR], 1.91; 95% CI: 1.04-3.51 and HR, 3.02; 95% CI: 1.26-7.24, respectively), adjusting for the HLA-DR genotype, first-degree relative with T1DM, maternal education, and delivery type. Specifically, early exposure to fruit and late exposure to rice/oat predicted T1DM (HR: 2.23; 95% CI: 1.14-4.39, and HR, 2.88; 95% CI, 1.36-6.11, respectively), while breastfeeding at the time of introduction to wheat/barley conferred protection (HR, 0.47; 95% CI, 0.26-0.86). Complicated vaginal delivery was also a predictor of T1DM (HR, 1.93; 95% CI, 1.03-3.61).

These results suggest the safest age to introduce solid foods in children at increased genetic risk for T1DM is between 4 and 5 months of age. Breastfeeding while introducing new foods may reduce T1DM risk.

**Introduction**

Type 1 diabetes (T1DM) is increasing worldwide with the most rapid increase among children younger than 5 years of age (14,15,18,20,64,65). Perinatal factors, such as birth delivery type, birth weight, infant growth, neonatal jaundice, and maternal age have been associated with T1DM and islet autoimmunity (IA), the preclinical stage of T1DM (66–74).

Exposures in the infant diet have been of particular interest in the etiology of T1D (75,76). Of the retrospective studies, 2 reports found T1DM cases had been exposed to solid foods earlier than control subjects (36,37), while 2 reports found no association (39,38) and 1 report showed T1DM cases had been exposed to solid foods later than control subjects (40). Prospective studies examining the development of IA have been more consistent. We found early (0-3 months of life) and late (≥7 months of life) introduction of cereals, compared with the introduction in the 4th to 6th month of life, to predict IA (41). Ziegler et al. found a similar increased risk for IA with early (<3 months of age) introduction of gluten-containing foods (42). Recently, Virtanen et al. found early introduction to root vegetables (≤4 months of age) associated with an increased risk for IA (43).

The purpose of this prospective study was to examine infant exposures, with a particular focus on infant diet, and their association with development of T1DM in a birth cohort of children at increased genetic risk for T1DM. While many of these exposures have been previously examined for association with IA in this cohort, to our knowledge, this is the first time these infant exposures have been examined prospectively for the risk for T1DM.
Methods

Study Population

The Diabetes Autoimmunity Study in the Young (DAISY) is a prospective study following up 2 groups of children at increased genetic risk for developing T1DM. One group consists of babies born at St. Joseph’s Hospital in Denver, Colorado, and screened by umbilical cord blood for diabetes-susceptibility alleles in the human leukocyte antigen (HLA) region (77,78). The details of the newborn screening have been published elsewhere (52).

The second group consists of unaffected children with a first-degree relative (either a mother, father, or sibling) with T1DM identified and recruited between birth and 8 years of age. Only children followed up from birth are included in these analyses. DAISY subjects enrolled at birth completed clinic visits at 9, 15, and 24 months, and annually thereafter. The Colorado Multiple Institutional Review Board approved all study protocols. Informed consent was obtained from the parents/legal guardians of all children. Assent was obtained from children age 7 years and older.

We examined the following descriptive and perinatal variables: HLA genotype (HLA-DR3/4, DQB1*0302 vs. other), first-degree relative with T1DM (mother vs. father or sibling vs. none), maternal age at child’s birth, maternal education (> 12 years vs. ≤ 12 years), sex (female vs. male), race/ethnicity (non-Hispanic white vs. other race/ethnicity), birth weight, birth season (September-February vs. March-August), type of delivery (cesarean delivery vs. complicated vaginal delivery vs. uncomplicated vaginal delivery), and maternal smoking during pregnancy (yes vs. no).

Weight Measurement

Weight in kilograms was measured at study visits at 9, 15, and 24 months, and annually thereafter on a scale with a precision of ± 0.1 kg. The infant growth rate was calculated for 0 to 9 months and 0 to 24 months as (weight at 9 months – birth weight)/(age at 9-month clinic visit)
and \((\text{weight at 24 months} - \text{birth weight})/(\text{age at 24-month clinic visit})\) to represent how rapidly the child grew \((\text{kg/y})\) in the first year and 2 years of life, respectively.

The first measurement of height was not taken until the child was able to stand cooperatively, at around 2 years of age; therefore, height was not taken into consideration in these analyses.

**Infant Diet Measurement**

Data for infant diet were collected during telephone or face-to-face interviews at 3, 6, 9, 12, and 15 months of age. At each interview, mothers were asked to report the date of introduction and frequency of exposure (i.e., number of servings per day) of all milks, formulas, and foods the infants consumed during the previous 3 months. Exclusive breastfeeding duration was determined by the reported age at which the infant was exposed to any foods or liquids other than breast milk or water. Breast-milk months is a novel breastfeeding variable we created to examine whether relative amounts of breast milk, rather than simply timing, are important in T1DM risk. Breast-milk months is a relative quantity of breast milk based on the proportion of breast milk to formula over the first 9 months of life. For example, for infants exclusively breastfed for the first 9 months, the proportion of breast milk to formula for each month was 1.0 and the number of breast-milk months summed to 9.0 breast-milk months. For infants who received both breast milk and formula, the total number of servings of breast milk for each month was divided by the total number of servings of formula and breast milk for that month, and these were summed over the first 9 months to arrive at the number of breast-milk months. Based on previous work showing a protective effect of breastfeeding when introducing cereals (for IA) (41) or gluten (for celiac disease) (79), we created 3 additional breastfeeding variables to represent whether the child was breastfed at the time of the introduction to any solid foods, cereals, and food containing wheat/barley.

We created an overall variable of age first exposure to any solid foods, as well as variables that were components of this solid-foods variable, such as age exposure to cereal
(wheat/barley/oats/rice), wheat/barley, rice/oat, fruit (not including fruit juice), vegetables, and
meat. There were no reports of introducing rye in the infant diet in DAISY children. Juices were
not included in the fruit variable because we were interested in solid-food introductions. The
study was an observational study; therefore, no dietary advice was given to the participating
families.

**Diagnosis of Type 1 Diabetes**

Type 1 diabetes was diagnosed by a physician and defined as typical symptoms of
polyuria and/or polydipsia and a random glucose level of 200 mg/dL or greater or an oral glucose
tolerance test with a fasting plasma glucose of 126 mg/dL or greater or a 2-hour glucose of 200
mg/dL or greater.

**Analysis Population**

Of the 2547 children followed by DAISY, 1886 (74.0%) had a clinic visit before the age
of 1.35 years, and of these, 1835 children (97.3%) had complete prospectively-collected solid
food exposure data. Our analysis population comprised these 1835 children, 53 of whom
developed T1DM during follow-up, as shown in Figure 2.

**Statistical Analysis**

The SAS version 9.3 (SAS Institute Inc) statistical software package was used for all
statistical analyses. Hazard ratios and 95% confidence intervals were estimated using Cox
proportional hazards regression to account for right-censored data. A clustered time-to-event
analysis was performed treating siblings from the same family as clusters, and robust sandwich
variance estimates (80) were used for statistical inference. Calculations of follow-up time began
at birth. When modeling the association between T1DM and perinatal and infant dietary
exposures, we adjusted for HLA, first-degree relative with T1DM, and delivery type because
these were associated with T1DM. We also evaluated maternal age, maternal education, sex, and
race/ethnicity as potential confounders.
Figure 2: Study Participants. Flowchart illustrating formation of analysis cohort from the Diabetes Autoimmunity Study in the Young (DAISY) population. T1DM indicates type 1 diabetes mellitus.
Only the inclusion of maternal education changed the hazard ratio of the perinatal or infant diet variable by more than 10%, and therefore was included in the final models. The significance threshold was defined as $\alpha < 0.05$. Because our analyses were based on a priori hypotheses and because we were interested in analyzing a complete diet in which the timing of introductions is highly correlated (i.e., not independent) among foods, $P$ values were not corrected for multiple testing.

**Results**

Children who developed T1DM were more likely to have the HLA-DR3/4, DQB1*0302 genotype and to have a father or sibling with T1DM (Table 1). Children who experienced a complicated vaginal delivery (i.e., breech or use of forceps or vacuum) were more likely to develop T1DM than those with an uncomplicated vaginal delivery, adjusting for HLA, first-degree relative with T1DM, and maternal education. The risk for T1DM was not predicted by birth weight, weight growth at 9 months or 2 years, exposure to smoking in utero, or season of birth.

Based on American Academy of Pediatrics guidelines that recommend introduction of solid foods between 4 and 6 months of age (81), we categorized age at first exposure to foods as younger than 4 months (early, introduced prior to the 4-month birthday), and 6 months of age or older (late, introduced on the 6-month birthday or later) and compared them to 4 to 5 months of age (referent) (Table 2). Adjusting for HLA, first-degree relative with T1DM, maternal education, and delivery type, both early ($<$4 months) and late ($\geq$6 months) exposure to any solid food predicted T1DM. We then examined individual components of the solid-foods variable to determine whether certain foods were driving this association. We first tested whether the age at first exposure to any cereals predicted T1DM because this variable has previously been found to predict IA in our cohort (41). In the current analysis, early ($<$4 months) and late ($\geq$6 months) exposure to any cereal (wheat/barley/oats/rice) increased the risk for T1DM. We then divided this variable into foods containing wheat/barley (i.e., gluten containing) and foods containing...
Late (≥6 months) first exposure to rice/oat increased the risk for T1DM, while early (<4 months) first exposure did not predict T1DM. First exposure to wheat/barley did not predict the development of T1DM. Early, but not late, exposure to fruits increased the risk for T1DM. The timing of introduction of vegetables or meat did not predict T1DM.

Children who were still breastfed when wheat/barley were introduced had a significantly lower risk for T1DM than children who were not breastfed at the time when wheat/barley were introduced, adjusting for HLA, first-degree relative with T1DM, maternal education, and delivery type (Table 2). More breast-milk months marginally decreased the risk for T1DM. Partial and exclusive breastfeeding duration and age at first exposure to cow’s milk did not predict T1DM.

As many food introductions happen together and therefore may not be independent, we placed all food-introduction and breastfeeding variables found to either marginally or significantly predict T1DM in the same model to evaluate the independent effects of each (Table 3). Adjusting for HLA, first-degree relative with T1DM, maternal education, and delivery type, the association between T1DM and early (<4 months) exposure to fruit was attenuated with the inclusion of the other food variables, while breastfeeding at the time of first exposure to foods containing wheat/barley and late (≥6 months) exposure to rice/oat remained significantly predictive of T1DM.

As rice/oat is the most common first solid food introduced in DAISY children (Figure 3), it was important to understand the differences, demographically and dietwise, between mothers who waited until after 6 months of age to introduce rice/oat compared with mothers who introduce it earlier (Table 4). Mothers who introduced rice/oat late (≥6 months) were more likely to be older, more educated, and have T1DM than mothers who introduced rice/oat earlier. Mothers who introduced rice/oat late (≥6 months) were also more likely to breastfeed (both partially and exclusively) for a longer duration and introduce all other solid foods later than mothers who introduced rice/oat earlier.
Table 1: Demographic and Perinatal Characteristics of DAISY Birth Cohort By T1DM Status

<table>
<thead>
<tr>
<th>Demographic factors</th>
<th>Developed T1DM (n = 53)</th>
<th>Did Not Develop T1DM (n = 1,782)</th>
<th>HR</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DR3/4, DQB1*0302</td>
<td>30 (56.6%)</td>
<td>421 (23.6%)</td>
<td>3.69</td>
<td>2.12, 6.40</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>First-degree relative with T1DM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>22 (41.5%)</td>
<td>1193 (67.0%)</td>
<td>1.0</td>
<td>[Reference]</td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>3 (5.7%)</td>
<td>182 (10.2%)</td>
<td>0.89</td>
<td>0.27, 2.96</td>
<td>0.85</td>
</tr>
<tr>
<td>Father or Sibling</td>
<td>28 (52.8%)</td>
<td>407 (22.8%)</td>
<td>3.32</td>
<td>1.85, 5.97</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Maternal age</td>
<td>30.9 ± 6.0</td>
<td>30.0 ± 5.7</td>
<td>1.01</td>
<td>0.95, 1.06</td>
<td>0.79</td>
</tr>
<tr>
<td>Maternal education (&gt; 12 years)</td>
<td></td>
<td></td>
<td>1.02</td>
<td>0.58, 1.80</td>
<td>0.94</td>
</tr>
<tr>
<td>Female</td>
<td>26 (49.1%)</td>
<td>867 (48.7%)</td>
<td>1.02</td>
<td>0.58, 1.80</td>
<td>0.94</td>
</tr>
<tr>
<td>Race/ethnicity, non-Hispanic white</td>
<td></td>
<td></td>
<td>1.80</td>
<td>0.78, 4.15</td>
<td>0.17</td>
</tr>
<tr>
<td>Perinatal factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight (kg)</td>
<td>3.3 ± 0.6</td>
<td>3.3 ± 0.6</td>
<td>0.90</td>
<td>0.53, 1.55</td>
<td>0.70</td>
</tr>
<tr>
<td>9-mo weight growth rate (kg/yr)</td>
<td></td>
<td></td>
<td>7.0 ± 1.4</td>
<td>7.0 ± 1.3</td>
<td>1.02</td>
</tr>
<tr>
<td>2-y weight growth rate (kg/yr)</td>
<td></td>
<td></td>
<td>4.4 ± 0.7</td>
<td>4.4 ± 0.7</td>
<td>0.97</td>
</tr>
<tr>
<td>Birth season, September – February</td>
<td></td>
<td></td>
<td>20 (37.7%)</td>
<td>838 (47.0%)</td>
<td>0.71</td>
</tr>
<tr>
<td>Delivery type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncomplicated vaginal delivery</td>
<td></td>
<td></td>
<td>32 (60.4%)</td>
<td>1133 (66.5%)</td>
<td>1.0</td>
</tr>
<tr>
<td>Complicated vaginal delivery</td>
<td></td>
<td></td>
<td>12 (22.6%)</td>
<td>192 (11.3%)</td>
<td>1.93</td>
</tr>
<tr>
<td>Cesarean delivery</td>
<td>9 (17.0%)</td>
<td>380 (22.3%)</td>
<td>0.83</td>
<td>0.37, 1.89</td>
<td>0.66</td>
</tr>
<tr>
<td>Exposure to maternal cigarette smoke in utero</td>
<td></td>
<td></td>
<td>5 (9.4%)</td>
<td>178 (10.5%)</td>
<td>1.19</td>
</tr>
</tbody>
</table>

Abbreviations: T1D, type 1 diabetes; HR, hazard ratio

aHRs are adjusted for HLA genotype, first-degree relative with T1D, maternal education, and delivery type

bHR is adjusted for HLA genotype, first-degree relative with T1D, and maternal education
Table 2: Infant Dietary Exposure Characteristics of DAISY Birth Cohort by T1DM Status

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Age at first exposure</th>
<th>Developed T1DM (n = 53)</th>
<th>Did Not Develop T1D&lt; (n = 1782)</th>
<th>Adjusted HR&lt;sup&gt;a&lt;/sup&gt;</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclusive breastfeeding duration (in months)</td>
<td>-</td>
<td>1.4 ± 2.0</td>
<td>1.3 ± 1.7</td>
<td>0.97</td>
<td>0.83, 1.14</td>
<td>0.73</td>
</tr>
<tr>
<td>Breastfeeding duration (in months)</td>
<td>-</td>
<td>5.8 ± 7.0</td>
<td>6.4 ± 6.9</td>
<td>0.97</td>
<td>0.92, 1.01</td>
<td>0.17</td>
</tr>
<tr>
<td>Breastfeeding-milk months (in months)</td>
<td>-</td>
<td>3.7 ± 3.3</td>
<td>4.2 ± 3.4</td>
<td>0.92</td>
<td>0.84, 1.01</td>
<td>0.08</td>
</tr>
<tr>
<td>Breastfeeding at introduction of solid foods, yes vs no</td>
<td>-</td>
<td>28 (52.8%)</td>
<td>980 (55.0%)</td>
<td>0.70</td>
<td>0.38, 1.28</td>
<td>0.25</td>
</tr>
<tr>
<td>Breastfeeding at introduction of cereals (wheat/barley/oats/rice), yes vs no</td>
<td>-</td>
<td>27 (50.9%)</td>
<td>976 (54.8%)</td>
<td>0.66</td>
<td>0.36, 1.21</td>
<td>0.18</td>
</tr>
<tr>
<td>Breastfeeding at introduction of wheat/barley, yes vs no</td>
<td>-</td>
<td>17 (32.1%)</td>
<td>765 (42.9%)</td>
<td>0.47</td>
<td>0.26, 0.86</td>
<td>0.01</td>
</tr>
<tr>
<td>Age first exposure to cow’s milk (in months)</td>
<td>&lt; 4 months old</td>
<td>4.4 ± 4.0</td>
<td>3.5 ± 3.3</td>
<td>1.01</td>
<td>0.93, 1.10</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at first exposure to any solid food&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>28 (52.8%)</td>
<td>763 (42.8%)</td>
<td>1.91</td>
<td>1.04, 3.51</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>7 (13.2%)</td>
<td>115 (6.5%)</td>
<td>3.02</td>
<td>1.26, 7.24</td>
<td>0.01</td>
</tr>
<tr>
<td>Age at first exposure to any cereal&lt;sup&gt;b&lt;/sup&gt; (wheat/barley/oats/rice)</td>
<td>&lt; 4 months old</td>
<td>25 (47.2%)</td>
<td>715 (40.1%)</td>
<td>1.72</td>
<td>0.95, 3.12</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>9 (17.0%)</td>
<td>137 (7.7%)</td>
<td>3.33</td>
<td>1.54, 7.18</td>
<td>0.002</td>
</tr>
<tr>
<td>Age at first exposures to foods containing wheat/barley&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>6 (11.3%)</td>
<td>124 (7.0%)</td>
<td>2.08</td>
<td>0.76, 5.68</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>34 (64.2%)</td>
<td>1049 (58.9%)</td>
<td>1.26</td>
<td>0.67, 2.38</td>
<td>0.48</td>
</tr>
<tr>
<td>Age at first exposure to foods containing rice/oat&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>24 (45.3%)</td>
<td>696 (39.1%)</td>
<td>1.62</td>
<td>0.90, 2.92</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>9 (17.0%)</td>
<td>159 (8.9%)</td>
<td>2.88</td>
<td>1.36, 6.11</td>
<td>0.01</td>
</tr>
<tr>
<td>Age at first exposure to fruit, excluding fruit juice&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>15 (28.3%)</td>
<td>265 (14.9%)</td>
<td>2.23</td>
<td>1.14, 4.39</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>14 (26.4%)</td>
<td>533 (29.9%)</td>
<td>1.03</td>
<td>0.51, 2.09</td>
<td>0.94</td>
</tr>
<tr>
<td>Age at first exposure to vegetables&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>7 (13.2%)</td>
<td>189 (10.6%)</td>
<td>1.19</td>
<td>0.49, 2.89</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>16 (30.2%)</td>
<td>544 (30.5%)</td>
<td>1.06</td>
<td>0.55, 2.01</td>
<td>0.87</td>
</tr>
<tr>
<td>Age at first exposure to meat&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>2 (3.8%)</td>
<td>13 (0.7%)</td>
<td>2.52</td>
<td>0.44, 14.49</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>42 (79.3%)</td>
<td>1527 (85.7%)</td>
<td>0.63</td>
<td>0.31, 1.28</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Abbreviations: DAISY, Diabetes Autoimmunity Study in the Young; HR, hazard ratio; T1DM, type 1 diabetes mellitus.

<sup>a</sup>Adjusted HRs are adjusted for human leukocyte antigen genotype, first-degree relative with T1DM, maternal education, and delivery type.

<sup>b</sup>Reference = 4 or 5 months old.
Table 3: Investigation of Independent Effects of Infant Diet Exposures on Risk of T1DM in DAISY Birth Cohort

<table>
<thead>
<tr>
<th>Variable</th>
<th>Age at first exposure</th>
<th>Adjusted HR</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DR3/4,DQB1*0302</td>
<td>-</td>
<td>5.59</td>
<td>3.07, 10.19</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>First-degree relative with T1DM</td>
<td>None</td>
<td>1.0</td>
<td>[Reference]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mother</td>
<td>1.75</td>
<td>0.53, 5.85</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Father or sibling</td>
<td>5.86</td>
<td>3.15, 10.90</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Maternal education, &gt; 12 y</td>
<td>-</td>
<td>0.89</td>
<td>0.42, 1.87</td>
<td>0.75</td>
</tr>
<tr>
<td>Delivery type</td>
<td>Uncomplicated vaginal delivery</td>
<td>1.0</td>
<td>[Reference]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Complicated vaginal delivery</td>
<td>1.89</td>
<td>0.98, 3.65</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Cesarean section delivery</td>
<td>0.82</td>
<td>0.34, 1.96</td>
<td>0.66</td>
</tr>
<tr>
<td>Breastfeeding at introduction of wheat/barley, yes vs no</td>
<td>-</td>
<td>0.46</td>
<td>0.26, 0.80</td>
<td>0.01</td>
</tr>
<tr>
<td>Age first exposed to foods containing wheat/barley&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>2.03</td>
<td>0.71, 5.80</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>1.17</td>
<td>0.61, 2.22</td>
<td>0.64</td>
</tr>
<tr>
<td>Age first exposed to foods containing rice/oat&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>0.92</td>
<td>0.44, 1.93</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>3.32</td>
<td>1.46, 7.51</td>
<td>0.004</td>
</tr>
<tr>
<td>Age first exposed to fruit, excluding fruit juice&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt; 4 months old</td>
<td>1.99</td>
<td>0.88, 4.51</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>≥ 6 months old</td>
<td>0.67</td>
<td>0.31, 1.46</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Abbreviations: DAISY, Diabetes Autoimmunity Study in the Young; HLA, human leukocyte antigen; HR, hazard ratio; T1DM, type 1 diabetes mellitus.

<sup>a</sup>Reference = 4 or 5 months old.
We explored the distribution of ages at first exposure to rice/oat in the late (≥6 months) category and found 71 children (42.3%) were introduced to rice/oat on their 6-month birthday (i.e., the first day of the late category), perhaps reflecting the propensity to use benchmarks such as birthdays in decisions about diet. To explore the sensitivity of the age cutoff for this late category, we placed the 71 children introduced to rice/oat on their birthday into the previous category of 4 to 5 months (a difference of one day) and reexamined the risk for T1DM. The adjusted hazard ratio for late introduction of rice/oat when placing those children introduced to rice/oat on the 6-month birthday in the reference category (4-5 months) increased to 4.28 (95% CI, 1.90-9.66) from 2.88 (95% CI, 1.36-6.11) (Table 3), suggesting T1DM predicted by late introduction to rice/oat was most likely driven by even later introduction of rice/oat than on the 6-month birthday.

Figure 3: First Solid Food Introduced to 1835 Diabetes Autoimmunity Study in the Young Infants
Discussion

DAISY now has a sufficient number of T1DM cases to investigate risk factors for development of T1DM. The findings reported here are consistent with our previous report, in which we had found a window of time for introduction of cereals, 4 to 6 months, outside of which risk for development of IA increased.(41) Our current data suggest the increased risk for T1DM found with early (<4 months) and late (≥6 months) introduction of solid foods appears to be driven by late first exposure to rice/oat and early first exposure to wheat/barley and fruit.

Our previous report defined the categories of time for introduction in terms of months of life rather than months of age, with first exposure to foods in the 0 to 3 months of life (i.e., prior to 3 months of age) and 7 months or more of life (i.e., at or after 6 months of age) with the reference being the 4 to 6 months of life (41). Given the recommendation by the American Academy of Pediatrics to introduce solid foods between 4 and 6 months of age, we chose on or after the 4-month birthday and before the 6-month birthday as the reference group for the current analysis, which more accurately reflects the way mothers and pediatricians would interpret this recommendation. Moreover, instead of referring to the categories in terms of months of life, in the present study we refer to these in terms of months of age, which is how we believe mothers tend to report the age of their child when describing milestones such as introducing new foods, first sitting, and first crawling.

The relation between late (≥6 months) introduction of rice/oat and development of T1DM is of particular interest. Rice/oat is the most common first solid food introduced in DAISY children, reflecting US practices. While there were many differences between children who were exposed to cereals after 6 months compared with before, an increased T1DM risk remained after adjustment for these variables. It is possible late (≥6 months) introduction of rice/oat represents an unmeasured set of variables/behaviors that increases T1DM risk.
Table 4: Characteristics of Timing of Rice/Oat Introduction

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>≤4 mo</th>
<th>4-5 mo</th>
<th>≥6 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DR3/4, DQB1*0302</td>
<td>25.4%</td>
<td>24.6%</td>
<td>20.8%</td>
</tr>
<tr>
<td>First-degree relative with T1DM&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>70.0%</td>
<td>64.2%</td>
<td>61.3%</td>
</tr>
<tr>
<td>Mother</td>
<td>8.5%</td>
<td>10.7%</td>
<td>13.7%</td>
</tr>
<tr>
<td>Father or sibling</td>
<td>21.5%</td>
<td>25.1%</td>
<td>25.0%</td>
</tr>
<tr>
<td>Mean maternal age, yr&lt;sup&gt;a&lt;/sup&gt;</td>
<td>28.5 ± 5.6</td>
<td>30.9 ± 5.5</td>
<td>31.8 ± 5.4</td>
</tr>
<tr>
<td>Maternal education, &gt; 12 y&lt;sup&gt;a&lt;/sup&gt;</td>
<td>68.8%</td>
<td>79.4%</td>
<td>85.0%</td>
</tr>
<tr>
<td>Race/ethnicity, non-Hispanic white</td>
<td>67.7%</td>
<td>71.9%</td>
<td>71.9%</td>
</tr>
<tr>
<td>Mean exclusive breastfeeding, mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.7 ± 1.2</td>
<td>1.4 ± 1.8</td>
<td>2.5 ± 2.6</td>
</tr>
<tr>
<td>Mean breastfeeding duration, mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.3 ± 5.9</td>
<td>7.3 ± 6.8</td>
<td>10.4 ± 8.9</td>
</tr>
<tr>
<td>Mean breast-milk months&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.9 ± 3.0</td>
<td>4.9 ± 3.4</td>
<td>6.0 ± 3.3</td>
</tr>
<tr>
<td>Mean age at first exposure to dairy, mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.0 ± 3.0</td>
<td>3.7 ± 3.3</td>
<td>5.4 ± 4.4</td>
</tr>
<tr>
<td>Mean age at first exposure to wheat/barley, mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.5 ± 2.0</td>
<td>7.2 ± 1.6</td>
<td>8.2 ± 1.8</td>
</tr>
<tr>
<td>Mean age at first exposure to fruit, mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.1 ± 1.4</td>
<td>6.1 ± 1.2</td>
<td>7.3 ± 1.6</td>
</tr>
<tr>
<td>Mean age at first exposure to vegetable, mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.4 ± 1.4</td>
<td>6.2 ± 1.2</td>
<td>7.4 ± 1.6</td>
</tr>
<tr>
<td>Mean age at first exposure to meat, mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8.7 ± 3.1</td>
<td>9.3 ± 2.7</td>
<td>10.3 ± 2.8</td>
</tr>
</tbody>
</table>

Abbreviations: HLA, human leukocyte antigen; T1DM, type 1 diabetes mellitus.
<sup>a</sup>P < .05 for comparisons across groups.

Our findings align with the American Academy of Pediatrics recommendation to introduce solid foods between 4 and 6 months of age (81); although recently, the American Academy of Pediatrics Section on Breastfeeding reaffirmed its recommendation of exclusive breastfeeding for about 6 months (82). These apparently conflicting recommendations can result in confusion over the best practice. Our sensitivity analysis showing T1DM predicted by late introduction to rice/oat was mostly driven by later introduction of rice/oat than 6 months suggests there is some leeway in the timing of the introduction of cereals, such that women who choose to exclusively breastfeed their child for 6 months may not have to worry they are increasing their child’s risk for T1DM by waiting until this time to introduce cereals.

The risk predicted by early exposure to solid foods might suggest a mechanism involving an abnormal immune response to solid food antigens in an immature gut immune system in susceptible individuals. As the increased risk is not limited to a specific food, it is possible many solids, including cereals and fruits, contain a common component that triggers an immature response. The increased risk predicted by late exposure to solid foods may be related to the larger
amounts given at initial exposure to older children (41,79,83). Also, if solid foods are introduced too late, when breast milk alone no longer meets the infant’s energy and nutrient needs, nutrient deficiencies may occur (84), which may play a role in increasing T1DM risk. Additionally, the increased risk predicted by late exposure to solid foods may be related to the cessation of breastfeeding before solid foods are introduced, resulting in a loss of the protective effects of breast milk at the introduction of foreign food antigens.

Breastfeeding is thought to reduce T1DM risk by protecting against infections through secretory immunoglobulin A antibodies and enhancement of the infant’s immune response via increased β-cell proliferation (85), as well as delayed exposure to foreign milk antigens. However, to our knowledge, prospective studies in genetically at-risk children have not found an association (41–43) between breastfeeding duration or timing of exposure to cow’s milk and development of IA. The most strongly associated breastfeeding variable in our analyses was breastfeeding at wheat/barley introduction, suggesting that breast milk may protect against an abnormal immune response to new antigens in an immature gut. Previously, we found a significant reduction in IA risk if cereals were introduced while the child was still breastfeeding, independent of the age at first exposure to cereals (41). Similarly, Ivarsson et al. reported a reduced celiac disease risk in children who had been breastfed when gluten was introduced (79).

Our findings suggesting an increased risk for T1DM predicted by complicated vaginal delivery but not cesarean delivery are similar to what we found previously with the outcome of IA (66). A recent report, which suggested cesarean delivery may predict a faster progression to T1DM after the appearance of IA, did not differentiate between complicated and uncomplicated vaginal delivery (86). Perhaps the stress of a complicated vaginal delivery affecting the fetal immune system, or other unknown factors complicating the birth or leading to a decision to not have a cesarean delivery, may be related to T1DM risk.

In this prospective study, we examined infant exposures as risk factors for the development of T1DM in children at increased genetic risk. While much of the focus of infant
diet and T1DM research has been on the timing of the introduction of a single antigen (i.e., milk or gluten), our data suggest multiple foods/antigens play a role, and that there is a complex relationship between the timing and type of infant food exposures and T1DM risk. In summary, there appears to be a safe window in which to introduce solid foods between 4 and 5 months of age; solid foods should be introduced while continuing to breastfeed to minimize T1DM risk in genetically susceptible children. These findings should be replicated in a larger cohort for confirmation.
CHAPTER V

INVESTIGATION OF THE VITAMIN D RECEPTOR GENE (VDR) AND ITS INTERACTION WITH PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 2 GENE (PTPN2) ON RISK OF ISLET AUTOIMMUNITY AND TYPE 1 DIABETES: THE DIABETES AUTOIMMUNITY STUDY IN THE YOUNG (DAISY)²

Abstract

The present study investigated the association between variants in the vitamin D receptor gene (VDR) and protein tyrosine phosphatase, non-receptor type 2 gene (PTPN2), as well as an interaction between VDR and PTPN2 and the risk of islet autoimmunity (IA) and progression to type 1 diabetes (T1D). The Diabetes Autoimmunity Study in the Young (DAISY) has followed children at increased risk of T1D since 1993. Of the 1692 DAISY children genotyped for VDR rs1544410, VDR rs2228570, VDR rs11568820, PTPN2 rs1893217, and PTPN2 rs478582, 111 developed IA, defined as positivity for GAD, insulin or IA-2 autoantibodies on 2 or more consecutive visits, and 38 IA positive children progressed to T1D. Proportional hazards regression analyses were conducted.

There was no association between IA development and any of the gene variants, nor was there evidence of a VDR*PTPN2 interaction. Progression to T1D in IA positive children was associated with the VDR rs2228570 GG genotype (HR: 0.49, 95% CI: 0.26-0.92) and there was an interaction between VDR rs1544410 and PTPN2 rs1893217 (pinteraction = 0.02). In children with the PTPN2 rs1893217 AA genotype, the VDR rs1544410 AA/AG genotype was associated with a decreased risk of T1D (HR: 0.24, 95% CI: 0.11-0.53, p=0.0004), while in children with the

²This chapter was reprinted with permission from Frederiksen B, Liu E, Romanos J, Steck AK, Yin X, Kroehl M, Fingerlin TE, Erlich H, Eisenbarth GS, Rewers M, Norris JM. (2013). Investigation of the vitamin D receptor gene (VDR) and its interaction with protein tyrosine phosphatase, non-receptor type 2 gene (PTPN2) on risk of islet autoimmunity and type 1 diabetes: the Diabetes Autoimmunity Study in the Young (DAISY). The Journal of Steroid Biochemistry and Molecular Biology 133, 51-57.
PTPN2 rs1893217 GG/GA genotype, the VDR rs1544410 AA/AG genotype was not associated with T1D (HR: 1.32, 95% CI: 0.43-4.06, p = 0.62). These findings should be replicated in larger cohorts for confirmation. The interaction between VDR and PTPN2 polymorphisms in the risk of progression to T1D offers insight concerning the role of vitamin D in the etiology of T1D.

Introduction

Type 1 diabetes (T1D) is an autoimmune disease in which the insulin-producing beta cells of the pancreas are destroyed. There is typically a preclinical phase of circulating autoantibodies, called islet autoimmunity (IA) that precedes the clinical diagnosis of T1D.

Vitamin D deficiency has been associated with a number of diseases, including multiple sclerosis (87), rheumatoid arthritis, and T1D (88), although not consistently. The mechanism by which vitamin D may exert its effects on these diseases is still not understood completely, particularly with regard to underlying genetic risk. The gene for the vitamin D receptor (VDR), through which vitamin D acts, has long been a candidate gene for T1D. Initial small studies found VDR polymorphisms to be associated with T1D (89–95). However, Nejentsev et al. (2004) analyzed association of the 98 VDR single nucleotide polymorphisms (SNPs) in up to 3763 type 1 diabetic families and found no evidence of association with T1D in the populations tested (96). Moreover, in a meta-analysis, Guo et al. (2006) found no evidence of an association between VDR gene polymorphisms (FokI, BsmI, ApaI, and TaqI) and T1D risk (97). Finally, in a recent analysis of 19 genes for association with T1D in the Type 1 Diabetes Genetics Consortium families, none of the forty SNPs genotyped in the VDR region were associated with T1D (98). A recent study (99) showed that the vitamin D receptor binds to a number of genomic positions across the genome, including a novel intronic binding site in the protein tyrosine phosphatase, non-receptor type 2 gene (PTPN2), which has also been associated with T1D through a genome-wide association scan in 2007 (100,101). This suggests the possibility of a more complex relationship in which variation in both VDR and PTPN2 is necessary to have an effect on diabetes risk, which may explain why previous findings regarding VDR have been inconsistent.
The Diabetes Autoimmunity Study in the Young (DAISY) has been prospectively following children at increased T1D risk for the development of IA and progression to T1D since 1993. The purpose of this study was to examine the associations between 5 particular VDR and PTPN2 SNPs and the development of IA and progression to T1D in the prospective DAISY cohort. We also aimed to explore a potential gene-gene interaction between VDR and PTPN2 polymorphisms and the risk of IA and progression to T1D.

Materials and Methods

Subjects

DAISY is a prospective study composed of two groups of children at increased risk for T1D who were recruited between 1993 and 2004 and are being followed prospectively for the development of IA and T1D. One group is made up of first degree relatives of patients with type 1 diabetes mellitus, recruited between birth and eight years of age. The second group consists of infants born at St. Joseph’s Hospital in Denver, Colorado, whose umbilical cord blood was screened for diabetes-susceptibility genotypes in the HLA region. The St. Joseph’s Hospital newborn population is representative of the general population of the Denver metropolitan area. Details of the newborn screening, sibling and offspring recruitment, and follow-up of both cohorts have been published previously (52,53). Cord blood or the first available blood sample (depending on enrollment group) was sent to Roche Molecular Systems, Inc., Alameda, CA, for PCR-based HLA class II typing. All study protocols were approved by the Colorado Multiple Institutional Review Board, and informed consent was given by parents of all participating children.

Measurement of Autoantibodies

Autoantibodies were tested at 9, 15, and 24 months, and annually thereafter, or at their first visit and annually thereafter if the child enrolled after birth. Radio-immunoassays were used to measure serum autoantibodies to insulin, GAD-65, and IA-2 (BDC512), as previously described (54–57), with rigorous confirmation of all positive and a subset of negative results.
The cut-off for positivity was established as the 99\textsuperscript{th} percentile of healthy controls. Children who tested autoantibody positive were put on an accelerated testing schedule of every 3-6 months.

Cases of persistent IA were defined as those children positive for at least one islet autoantibody (IAA, GAD-65, IA-2) on two or more consecutive visits. Type 1 diabetes was diagnosed by a physician and defined as random blood glucose $>$200 mg/dL and/or HbA1c (A1C) $>$6.2\% with clinical symptoms of diabetes.

**VDR and PTPN2 Genotyping**

DAISY children were genotyped for VDR rs1544410 (BsmI), VDR rs2228570 (FokI), VDR rs11568820 (Cdx2), PTPN2 rs1893217, and PTPN2 rs478582. The three VDR SNPs were chosen based on previous associations with T1D (89,90,94,95,102) as well as function (96,103–105). The VDR rs1544410 SNP is located at the 3’ end of VDR and is not known to alter the structure or function of VDR (102,106,107). The VDR rs2228570 polymorphism is a T-C transition at the translation initiation codon of VDR that results in a shorter protein with increased biological activity (103–105). VDR rs11568820 is in the 5’ promoter region of VDR, results in reduced transcriptional activity of the promoter, and affects calcium absorption in the intestine (107–110). The PTPN2 SNPs, both of which are intronic, were chosen for their association with both T1D and celiac disease (101). Linkage disequilibrium (LD) between the three VDR SNPs and between the two PTPN2 SNPs was tested in our population using Haploview version 4.2 and none of the SNPs were found to be in LD as measured by $r^2$, with $r^2 \leq 0.005$ for the VDR SNPs, and an $r^2 = 0.124$ for the PTPN2 SNPs.

VDR rs1544410 was genotyped using a linear array method at Roche Molecular Systems, Inc., as described in Mirel et al. (2002) (111). PTPN2 rs1893217 was genotyped as part of a second project using Illumina 48-plex VeraCode technology following the manufacturer’s protocol. Genotyping data analysis and clustering was performed in Illumina’s GenomeStudio. Clustering clouds were manually investigated and adjusted if necessary. All plates included one duplicate sample and one positive control. VDR rs2228570, VDR rs11568820, and VDR
rs478582 were typed utilizing the Taqman SNP genotype-based OpenArray platform [Applied Biosystems, CA, USA]. Custom designed arrays were loaded using the OpenArray AccuFill system and cycling was performed on a GeneAmp 9700 PCR system [Applied Biosystems, CA, USA], all according to manufacturer’s protocol. Alleles were analyzed using the OpenArray SNP genotyping analysis software v.1.0.3 and Taqman Genotyper Software 2.0 [Applied Biosystems, CA, USA].

All five SNPs had a 95% call rate or higher and were in Hardy-Weinberg equilibrium. Each SNP was tested for consistency with Hardy-Weinberg proportions using a 1-degree of freedom χ² goodness-of-fit test with a p-value of 0.001 considered as evidence of a departure from Hardy-Weinberg equilibrium.

**Analysis Population**

We obtained genetic data for all five SNPs on 1692 children in the DAISY cohort. A flow-chart showing selection of subjects for the analyses is shown in Figure 3. Comparisons of the children who have complete genetic data with those who do not are presented in Table 5 for the cohort examining risk of IA, and Table 6 for the cohort examining progression to T1D in IA positive children. There were small differences in HLA, first-degree relative status and sex between children with and without complete genetic data in the cohort examining risk of IA, and no differences in the cohort examining progression to T1D in IA positive children. The 1692 children with complete genetic data included 122 children who developed persistent IA, of whom 38 went on to develop T1D. However, 11 IA cases were positive on their first clinic visit, thus considered left censored, and were removed from the development of IA analysis cohort.
Table 5: Characteristics of DAISY Children With and Without Complete Data for the VDR and PTPN2 SNPs

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Children with complete genetic data</th>
<th>Children without complete genetic data</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age(^a) (years)</td>
<td>8.9 (5.5)</td>
<td>9.0 (5.8)</td>
<td>0.57</td>
</tr>
<tr>
<td>HLA-DR3/4, DQB1*0302</td>
<td>308 (18.3%)</td>
<td>189 (22.3%)</td>
<td>0.02</td>
</tr>
<tr>
<td>First degree relative with T1D</td>
<td>719 (42.8%)</td>
<td>400 (47.2%)</td>
<td>0.04</td>
</tr>
<tr>
<td>Ethnicity (non-Hispanic white)</td>
<td>1210 (72.0%)</td>
<td>614 (74.2%)</td>
<td>0.25</td>
</tr>
<tr>
<td>Sex (Female)</td>
<td>783 (46.6%)</td>
<td>431 (50.8%)</td>
<td>0.04</td>
</tr>
</tbody>
</table>

\(^a\)age at first IA positive visit in autoantibody positive children or age at last follow-up in autoantibody negative children

Table 6: Characteristics of IA Positive Children With and Without Complete Data for the VDR and PTPN2 SNPs

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>IA Positive children with complete genetic data</th>
<th>IA Positive children without complete genetic data</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age(^a) (years)</td>
<td>8.2 (4.0)</td>
<td>8.0 (4.0)</td>
<td>0.84</td>
</tr>
<tr>
<td>Mean age at first IA positive visit (years)</td>
<td>3.6 (2.9)</td>
<td>3.9 (3.0)</td>
<td>0.75</td>
</tr>
<tr>
<td>HLA-DR3/4, DQB1*0302</td>
<td>37 (30.3%)</td>
<td>26 (41.9%)</td>
<td>0.12</td>
</tr>
<tr>
<td>First degree relative with T1D</td>
<td>78 (63.9%)</td>
<td>36 (58.1%)</td>
<td>0.44</td>
</tr>
<tr>
<td>Ethnicity (non-Hispanic white)</td>
<td>101 (82.8%)</td>
<td>47 (75.8%)</td>
<td>0.26</td>
</tr>
<tr>
<td>Sex (Female)</td>
<td>57 (46.7%)</td>
<td>35 (56.5%)</td>
<td>0.21</td>
</tr>
</tbody>
</table>

\(^a\)age at T1D diagnosis in diabetic children or age at last follow-up in non-diabetic children

Statistical Analyses

SAS version 9.3 (SAS Institute Inc., Cary, NC, USA) statistical software package was used for all statistical analyses. SNPs were analyzed for their association both with development of IA and with progression from IA to T1D. For each model described below, hazard ratios (HR) and 95% confidence intervals (CI) were estimated using Cox proportional hazards regression. A clustered time to event analysis was performed treating siblings from the same family as clusters, and robust sandwich variance estimates (80) were used for statistical inference. Analyses of time to development of IA were adjusted for the HLA-DR genotype (HLA-DR3/4, DQB1*0302 vs. other genotypes), presence of a first degree relative with T1D, and self-reported ethnicity (non-Hispanic white vs. other). Analyses of time to progression to T1D were adjusted, in addition, for age at first positive autoantibody. The significance threshold was defined a priori as \( \alpha < 0.05 \).
Because our analyses were based on an *a priori* hypothesis, p-values were not corrected for multiple testing.

Figure 4. Flow Chart Illustrating Formation of Analysis Cohorts from the DAISY Study Population. Flow chart illustrating formation of analysis cohorts from the DAISY study population. IA, islet autoimmunity; T1D, type 1 diabetes.

Genotype frequencies of the five SNPs are presented in Table 7 for the cohort examining risk of IA, and in Table 8 for the cohort examining progression to T1D in IA positive children. To avoid sparse cell counts, the genotypes of each SNP were dichotomized in the following manner: *VDR* rs1544410 genotypes were dichotomized as AA/AG vs. GG, *VDR* rs2228570 genotypes as AA/AG vs. GG, *VDR* rs11568820 genotypes as TT/TC vs. CC, *PTPN2* rs1893217 genotypes as GG/GA vs. AA, and the *PTPN2* rs478582 genotypes as CC/CT vs. TT for all analyses.

We analyzed *VDR* and *PTPN2* in separate models before testing the interaction. For the *VDR* gene analysis, all three *VDR* SNPs were included in the same model, adjusting for HLA-DR3/4, DQB1*0302 genotype, first degree relative with type 1 diabetes, and non-Hispanic white
ethnicity. Similarly, for the PTPN2 gene analysis, the two PTPN2 SNPs were included in the same model, adjusting for HLA-DR3/4, DQB1*0302 genotype, first degree relative with type 1 diabetes, and non-Hispanic white ethnicity. These analyses were performed for both the development of IA as well as for progression to T1D in IA positive children. Then, based on the finding by Ramagopalan et al. (2010), all possible VDR*PTPN2 SNP interactions were examined for association with development of IA and progression to T1D (99), with separate models for each interaction tested. For significant interaction terms, we computed HRs and 95% CIs from the coefficient and standard error estimates of the main effect and interaction terms to describe the interaction.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDR rs1544410</td>
<td>A/A</td>
<td>241</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>815</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>676</td>
</tr>
<tr>
<td></td>
<td>A/A</td>
<td>229</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>823</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>680</td>
</tr>
<tr>
<td></td>
<td>C/C</td>
<td>1927</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>589</td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>116</td>
</tr>
<tr>
<td>PTPN2 rs1893217</td>
<td>A/A</td>
<td>1154</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>430</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>35</td>
</tr>
<tr>
<td>PTPN2 rs478582</td>
<td>C/C</td>
<td>315</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>847</td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>526</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDR rs1544410</td>
<td>A/A</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>41</td>
</tr>
<tr>
<td>VDR rs2228570</td>
<td>A/A</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>48</td>
</tr>
<tr>
<td>VDR rs11568820</td>
<td>C/C</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>4</td>
</tr>
<tr>
<td>PTPN2 rs1893217</td>
<td>A/A</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>3</td>
</tr>
<tr>
<td>PTPN2 rs478582</td>
<td>C/C</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>37</td>
</tr>
</tbody>
</table>

### Results

**Development of Persistent Islet Autoimmunity**

We first examined whether variants in VDR were associated with development of persistent islet autoimmunity, and whether these interacted with variants in PTPN2. Table 9 describes the children by whether or not they developed persistent IA. Of a total of 122 IA positive children in DAISY, 11 had to be excluded from the subsequent proportional hazards
analyses of IA because they tested autoantibody positive on their first study visits (i.e., they were left-censored). The mean age at first IA positive visit was 5.5 years, and the mean age at last follow-up visit in children who did not develop IA was 9.1 years. IA positive children, were more likely to have the HLA-DR3/4, DQB1*0302 genotype or a first degree relative with T1D compared to children who did not develop IA.

Table 10 shows the association between the SNPs and the development of IA for the model containing all three VDR SNPs and the model containing both of the PTPN2 SNPs, adjusted for HLA DR3/4 status, having a first degree relative with T1D, and ethnicity. There was no evidence of significant association between the three VDR SNPs together and development of IA. The two PTPN2 variants (rs1893217 and rs478582) were also not associated with risk of IA.

We then explored whether variants in VDR and PTPN2 interacted to affect risk of IA. With three VDR SNPs and two PTPN2 SNPs, we tested six potential VDR*PTPN2 interactions for their associations with risk of IA and found no evidence of VDR*PTPN2 interactions, adjusting for HLA DR3/4 status, having a first degree relative with T1D, and ethnicity (data not shown).

Table 9: Characteristics of the Analysis Population by Islet Autoimmunity (IA) Status

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Children positive for IA n = 111</th>
<th>Children negative for IA n = 1570</th>
<th>Univariate HR and 95% CI</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age(^\text{a}) (years)</td>
<td>5.5 (3.8)</td>
<td>9.1 (5.6)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>HLA-DR3/4, DQB1*0302</td>
<td>34 (30.6%)</td>
<td>274 (17.5%)</td>
<td>2.18 (1.45, 3.26)</td>
<td>0.0002</td>
</tr>
<tr>
<td>First degree relative with T1D</td>
<td>67 (60.4%)</td>
<td>652 (41.5%)</td>
<td>1.66 (1.13, 2.43)</td>
<td>0.01</td>
</tr>
<tr>
<td>Ethnicity (non-Hispanic white)</td>
<td>90 (81.1%)</td>
<td>1120 (71.3%)</td>
<td>1.38 (0.85, 2.25)</td>
<td>0.20</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>53 (47.8%)</td>
<td>730 (46.5%)</td>
<td>1.07 (0.75, 1.54)</td>
<td>0.71</td>
</tr>
<tr>
<td>VDR rs1544410 (AA/AG)</td>
<td>76 (68.5%)</td>
<td>959 (61.1%)</td>
<td>1.32 (0.90, 1.95)</td>
<td>0.16</td>
</tr>
<tr>
<td>VDR rs2228570 (AA/AG)</td>
<td>69 (62.2%)</td>
<td>955 (60.8%)</td>
<td>0.99 (0.68, 1.45)</td>
<td>0.98</td>
</tr>
<tr>
<td>VDR rs11568820 (TT/TC)</td>
<td>36 (32.4%)</td>
<td>630 (40.1%)</td>
<td>0.72 (0.49, 1.07)</td>
<td>0.11</td>
</tr>
<tr>
<td>PTPN2 rs1893217 (GG/GA)</td>
<td>39 (35.1%)</td>
<td>438 (27.9%)</td>
<td>1.34 (0.90, 1.99)</td>
<td>0.15</td>
</tr>
<tr>
<td>PTPN2 rs478582 (CC/CT)</td>
<td>78 (70.3%)</td>
<td>1077 (68.6%)</td>
<td>1.08 (0.72, 1.62)</td>
<td>0.72</td>
</tr>
</tbody>
</table>

\(^{a}\)Age at first IA positive visit in autoantibody positive children or age at last follow-up in autoantibody negative children.
Table 10: Association Between Variants in VDR and PTPN2 and Risk of Islet Autoimmunity (n = 111 Affecteds and 1570 Unaffecteds)

<table>
<thead>
<tr>
<th>Variable</th>
<th>HR</th>
<th>95% CI</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model of VDR Variants</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VDR rs1544410 (AA/AG vs. GG)</td>
<td>1.39</td>
<td>0.94, 2.07</td>
<td>0.10</td>
</tr>
<tr>
<td>VDR rs2228570 (AA/AG vs. GG)</td>
<td>0.99</td>
<td>0.68, 1.45</td>
<td>0.96</td>
</tr>
<tr>
<td>VDR rs11568820 (TT/TC vs. CC)</td>
<td>0.73</td>
<td>0.49, 1.10</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Model of PTPN2 Variants</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPN2 rs1893217 (GG/GA)</td>
<td>1.36</td>
<td>0.90, 2.04</td>
<td>0.15</td>
</tr>
<tr>
<td>PTPN2 rs478582 (CC/CT)</td>
<td>1.23</td>
<td>0.80, 1.87</td>
<td>0.35</td>
</tr>
</tbody>
</table>

In the Model of VDR variants, the three VDR SNPs were included in a single model, adjusting for HLA-DR3/4, DQB1*0302 genotype, first degree relative with type 1 diabetes, and non-Hispanic white ethnicity. Likewise, in the Model of PTPN2 variants, the two PTPN2 SNPs were included in a single model, adjusting for HLA-DR3/4, DQB1*0302 genotype, first degree relative with type 1 diabetes, and non-Hispanic white ethnicity.

**Progression to T1D in Children with IA**

We examined whether variants in VDR were associated with progression to T1D in IA positive children, and whether these variants interacted with variants in PTPN2. The study population of IA positive children is described in Table 11. Of the 122 IA positive children in DAISY, 38 developed T1D; the mean age at T1D diagnosis was 8.2 years. The mean age at last follow-up visit in non-diabetic children with IA was 12.8 years. Children who develop T1D were younger when they first tested positive for an autoantibody than IA positive children who have not progressed to T1D, 3.6 and 6.1 years, respectively. They were also more likely to have the HLA-DR3/4, DQB1*0302 genotype compared to children who did not progress to T1D.

Table 12 displays the three VDR SNPs modeled together in one model and their association with the progression to T1D in IA positive children, adjusting for HLA DR3/4 status, having a first degree relative with T1D, ethnicity, and age at first IA positive visit. One of the two functional VDR SNPs, VDR rs2228570, was significantly associated with development of T1D when the three VDR SNPs were modeled together. The two PTPN2 SNPs were analyzed together similarly in a separate model to more completely describe the variation in the PTPN2 gene, but there was no evidence of significant associations with progression to T1D in IA positive children. We then explored whether variants in VDR and PTPN2 interacted to affect progression to T1D in IA positive children. We found a significant interaction between VDR 1544410 and
PTPN2 rs1893217 on progression to T1D, adjusting for HLA DR3/4 status, having a first degree relative with T1D, ethnicity and age at first IA positive visit (Table 12 and Figure 5).

**Table 11: Characteristics of Islet Autoantibody (IA) Positive Subjects by Type 1 Diabetes (T1D) Status**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>IA positive children who progressed to T1D</th>
<th>IA positive children who have not progressed to T1D</th>
<th>Univariate HR and 95% CI</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (years)</td>
<td>8.2 (4.0)</td>
<td>12.8 (4.6)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean age at first IA positive visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit (years)</td>
<td>3.6 (2.9)</td>
<td>6.1 (3.8)</td>
<td>0.88 (0.78, 0.98)</td>
<td>0.03</td>
</tr>
<tr>
<td>HLA-DR3/4, DQB1*0302</td>
<td>23 (60.5%)</td>
<td>14 (16.7%)</td>
<td>3.98 (2.13, 7.41)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>First degree relative with T1D</td>
<td>26 (68.4%)</td>
<td>52 (61.9%)</td>
<td>1.11 (0.57, 2.17)</td>
<td>0.76</td>
</tr>
<tr>
<td>Ethnicity (non-Hispanic white)</td>
<td>35 (92.1%)</td>
<td>66 (78.6%)</td>
<td>1.95 (0.57, 6.62)</td>
<td>0.29</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>20 (52.6%)</td>
<td>37 (44.1%)</td>
<td>1.47 (0.76, 2.85)</td>
<td>0.26</td>
</tr>
<tr>
<td>VDR rs1544410 (AA/AG)</td>
<td>24 (63.2%)</td>
<td>57 (67.9%)</td>
<td>0.67 (0.34, 1.29)</td>
<td>0.23</td>
</tr>
<tr>
<td>VDR rs2228570 (AA/AG)</td>
<td>18 (47.4%)</td>
<td>56 (66.7%)</td>
<td>0.63 (0.34, 1.17)</td>
<td>0.14</td>
</tr>
<tr>
<td>VDR rs11568820 (TT/TC)</td>
<td>14 (36.8%)</td>
<td>30 (35.7%)</td>
<td>1.37 (0.70, 2.71)</td>
<td>0.36</td>
</tr>
<tr>
<td>PTPN2 rs1893217 (GG/GA)</td>
<td>12 (31.6%)</td>
<td>31 (36.9%)</td>
<td>0.77 (0.40, 1.48)</td>
<td>0.43</td>
</tr>
<tr>
<td>PTPN2 rs478582 (CC/CT)</td>
<td>28 (73.7%)</td>
<td>57 (67.9%)</td>
<td>1.23 (0.61, 2.45)</td>
<td>0.57</td>
</tr>
</tbody>
</table>

*Age at T1D diagnosis in diabetic children or age at last follow-up in non-diabetic children.

The association between VDR rs1544410 and the progression to T1D in IA positive children differs by PTPN2 rs1893217 genotype, as shown in Figure 5. In children with the PTPN2 rs1893217 AA genotype, the VDR rs1544410 AA/AG genotype is associated with a significantly lower risk of progressing to T1D compared to children with the VDR rs1544410 GG genotype (HR: 0.24, 95% CI: 0.11-0.53, p-value=0.0004). In IA positive children with the PTPN2 rs1893217 GG/GA genotype, the VDR rs1544410 AA/AG genotype was not associated with progression to T1D. A subanalysis on non-Hispanic whites only produced similar results to those reported herein for development of IA and for progression to T1D in IA positive children (data not shown).

**Discussion**

In the DAISY population, the three VDR SNPs, VDR rs1544410, VDR rs2228570, and VDR rs11568820 we tested were not significantly associated with the appearance of IA. However, VDR rs2228570 was found to be associated with progression to T1D in IA positive
children, and VDR rs1544410 significantly interacted with PTPN2 rs1893217 on risk of progression to T1D, suggesting that the role of VDR in risk of T1D is complex. Besides the aforementioned interaction between PTPN2 rs1893217 and VDR rs1544410, the two PTPN2 SNPs we tested were not significantly associated with the appearance of IA or progression to T1D as main effects. This is in slight contrast to our recent report in which we found PTPN2 rs1893217 to be weakly associated with development of islet autoimmunity (112). This discrepancy can be attributed to the use of different IA case definitions. DAISY uses two definitions of IA, one that defines IA as the presence of at least one islet autoantibody on two consecutive visits (which is the definition used in the present study); and the other that further requires that the children still be autoantibody positive or diabetic on their most recent visit (which is the definition used in the previous study).

The studies that have examined VDR variants and the outcome of T1D have been inconsistent, and it is possible that the effects of VDR are only important for a faster development of T1D in the presence of IA, but not for the overall risk of T1D among genetically (HLA-DR,DQ) susceptible children. We found the functional VDR SNP, VDR rs2228570, to be a significant predictor in the progression to development of T1D. It also appears that VDR rs1544410, presumed to be non-functional because of its location in an intron, becomes mechanismically important for the progression to T1D in IA positive children when in combination with PTPN2 rs1893217. This, coupled with the finding that the vitamin D receptor has an intronic binding site in the PTPN2 gene (99), may provide insight into one of the ways in which the vitamin D receptor exerts its effects on T1D, which currently are still widely unknown. VDR rs1544410 has been shown to be in LD with a poly(A) microsatellite located in the 3’-untranslated region of the VDR gene, which has been discussed to influence VDR mRNA stability (113,114). Ramagopalan et al. (2010) identified two VDR binding intervals in the PTPN2 gene; and the SNP involved in the interaction (PTPN2 rs1893217) is 5 kb and 58 kb away from these
two intervals (99). Thus, it is possible that this SNP is in LD with a variant in one of the two identified VDR binding intervals. This unique finding should be confirmed in other populations.

**Table 12: Association Between Variants in VDR and PTPN2 and Risk of Progression to Type 1 Diabetes in IA Positive Children (n = 38 Affecteds and 84 Unaffecteds)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>HR</th>
<th>95% CI</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model of VDR variants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VDR rs1544410 (AA/AG)</td>
<td>0.61</td>
<td>0.31, 1.21</td>
<td>0.16</td>
</tr>
<tr>
<td>VDR rs2228570 (AA/AG)</td>
<td>0.50</td>
<td>0.26, 0.95</td>
<td>0.03</td>
</tr>
<tr>
<td>VDR rs11568820 (TT/TC)</td>
<td>1.87</td>
<td>0.93, 3.74</td>
<td>0.08</td>
</tr>
<tr>
<td>Model of PTPN2 variants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPN2 rs1893217 (GG/GA)</td>
<td>0.65</td>
<td>0.27, 1.60</td>
<td>0.35</td>
</tr>
<tr>
<td>PTPN2 rs478582 (CC/CT)</td>
<td>1.07</td>
<td>0.43, 2.67</td>
<td>0.89</td>
</tr>
<tr>
<td>VDR model with interaction term</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VDR rs1544410*PTPN2 rs1893217 interaction</td>
<td>*</td>
<td>*</td>
<td>0.02</td>
</tr>
<tr>
<td>PTPN2 rs1893217</td>
<td>*</td>
<td>*</td>
<td>0.007</td>
</tr>
<tr>
<td>VDR rs1544410 (AA/AG)</td>
<td>*</td>
<td>*</td>
<td>0.0004</td>
</tr>
<tr>
<td>VDR rs2228570 (AA/AG)</td>
<td>0.49</td>
<td>0.26, 0.92</td>
<td>0.03</td>
</tr>
<tr>
<td>VDR rs11568820 (TT/TC)</td>
<td>1.82</td>
<td>0.90, 3.66</td>
<td>0.10</td>
</tr>
</tbody>
</table>

In the model of VDR variants, the three VDR SNPs were included in a single model, adjusting for HLA-DR3/4, DQB1*0302 genotype, first degree relative with type 1 diabetes, non-Hispanic white ethnicity, and age at first IA positive visit. Likewise, in the Model of PTPN2 variants, the two PTPN2 SNPs were included in a single model, adjusting for HLA-DR3/4, DQB1*0302 genotype, first degree relative with type 1 diabetes, non-Hispanic white ethnicity, and age at first IA positive visit. In the VDR model with interaction term, the variables included were VDR rs1544410, VDR rs2228570, VDR rs11568820, PTPN2 rs1893217, the VDR rs1544410*PTPN2 rs1893217 interaction term, HLA-DR3/4, DQB1*0302 genotype, first degree relative with type 1 diabetes, non-Hispanic white ethnicity, and age at first IA positive visit.

Ramagopalan et al. found increased VDR binding in intronic and intergenic regions compared with the basal state upon stimulation with calcitriol, the hormonally active form of vitamin D (99). We were unable to examine what influence calcitriol had on the observed associations between VDR and progression to T1D because we do not have measures of 1,25-dihydroxyvitamin D on our children. Circulating 1,25-dihydroxyvitamin D has a short half-life of 15 hours, is closely regulated by parathyroid hormone, calcium, and phosphate (115), and measurement requires a large amount of plasma for our pediatric subjects. Therefore, DAISY measures annually only plasma 25[OH]D as an estimate of vitamin D status. Only in situations of severe vitamin D deficiency do levels of 1,25-dihydroxyvitamin D decrease (88,116).
Simpson et al. (2011) recently reported no association between 25(OH)D levels and the risk of IA, nor progression to diabetes in IA positive children in DAISY (47).

**Figure 5: Association Between VDR rs1544410 and Risk of Progression to T1D in IA Positive Children by PTPN2 rs1893217 Genotype.** The association between VDR rs1544410 and the risk of progression to T1D in IA positive children differs by PTPN2 rs1893217 genotype. Black dots represent the hazard ratio for the VDR rs1544410 AA/AG genotype (with GG genotype as the referent). The solid lines represent the 95% confidence intervals for the hazard ratios. In children with the PTPN2 rs1893217 AA genotype (n = 79), the VDR rs1544410 AA/AG genotype (n = 54) was associated with a significantly lower risk of progressing to T1D compared to children with the VDR rs1544410 GG genotype (n = 25) (HR: 0.24, 95% CI: 0.11-0.53, p-value = 0.0004). In IA positive children with the PTPN2 rs1893217 GG/GA genotype (n = 43), the VDR rs1544410 AA/AG genotype (n = 27) was not associated with progression to T1D (HR: 1.32, 95% CI: 0.43, 4.06, p-value = 0.62).

The observation that VDR rs2228570 is associated with progression to T1D but not development of IA suggests that the role of VDR may be in the acceleration or deceleration of progression to T1D in autoimmune children. VDR 2228570 has been inconsistently associated with T1D (96,102,106,117–123). VDR rs2228570 is a coding non-synonymous SNP located in the translational initiation codon that determines the formation of two protein variants: a longer version of the VDR protein (427-amino acids) that corresponds to the A allele and a form shortened by three amino acids corresponding to the G allele (424-amino acids) (103,124,125). Functional studies have shown that the shorter version of the protein has greater transcriptional activity and is more effective in transactivation of the vitamin D signal (103,124–127).
Therefore, the VDR rs2228570 minor allele is associated with the production of a longer VDR protein that is less transcriptionally active (124,128,129). In this particular study, we found the VDR rs2228570 minor allele (A) to be significantly protective for the development of T1D in IA positive children, which may imply that once autoimmune, less active transcription is beneficial. It is also interesting that VDR rs2228570 only becomes significant when adjusting for VDR 1544410 and VDR rs11568820, which may be an indication of the complexity of the way in which these polymorphisms act together.

The major strength of this study is the prospective long-term follow-up from birth of children at an increased risk for T1D, which allowed us to differentiate between genetic risk factors for the appearance of autoimmunity and the subsequent progression to T1D. However, the cost of assembling and following such a unique cohort has limited the number of IA positive children and children who progress to T1D that we could include in our analysis. While our study had adequate power to detect the novel interaction between variants in VDR and PTPN2 that we present herein, it is possible that we have missed other interactions of smaller magnitude due to limited power. This is the first study to observe this gene-gene interaction in the progression phase to T1D, and it is important that this be investigated in other populations, to rule out type 1 error and lend further evidence to this association.

In conclusion, a functional variant in VDR, VDR rs2228570 was found to be significantly associated with progression to T1D in IA positive children of the DAISY population. A significant interaction between a non-coding VDR variant, VDR rs1544410, and a variant in PTPN2, PTPN2 rs1893217, was also found to be associated with progression to T1D in IA positive children. Interestingly, these two VDR variants, VDR 2228570 and VDR 1544410, were not found to be associated with the appearance of IA. These findings should be explored in other prospective cohorts following children with IA for progression to T1D, as they may offer insight concerning the complex role of vitamin D in the etiology of T1D, as well as other autoimmune diseases associated with PTPN2, e.g., Crohn’s disease and celiac disease.
CHAPTER VI
ASSOCIATION BETWEEN VITAMIN D METABOLISM GENE POLYMORPHISMS
AND RISK OF ISLET AUTOIMMUNITY AND PROGRESSION TO TYPE 1
DIABETES: THE DIABETES AUTOIMMUNITY STUDY IN THE YOUNG (DAISY)³

Abstract

Vitamin D metabolism genes have been associated with type 1 diabetes (T1D) risk; however, these genes have not been investigated for association with the pre-clinical phase of T1D, islet autoimmunity (IA). Studies of vitamin D metabolism genes may elucidate the role of vitamin D in complex diseases. The objective of the study was to explore the association between seven vitamin D metabolism gene single-nucleotide polymorphisms (SNPs) and the risk of IA and progression to T1D. The Diabetes Autoimmunity Study in the Young is a longitudinal, observational study. Newborn screening for human leukocyte antigen, sibling and offspring recruitment, and follow-up took place in Denver, Colorado. A total of 1708 children at increased genetic risk of T1D participated in the study: 148 developed IA and 62 IA-positive children progressed to T1D. IA, defined as positivity for glutamic acid decarboxylase, insulin, or IA-2 autoantibodies on two or more consecutive visits, and T1D, diagnosed by a physician, were the main outcome measures. The risk of IA was associated with DHCR7/NADSYN1 rs12785878 and CYP27B1 rs4646536 (hazard ratio 1.36, 95% confidence interval 1.08-1.73 (for each additional minor allele) and hazard ratio 0.59, 95% confidence interval 0.39-0.89 (for A/G compared with the A/A genotype), respectively]. None of the vitamin D SNPs typed was associated with progression to T1D in IA-positive children. Six of the seven SNPs were significantly associated with 25-hydroxyvitamin D levels.

DHCR7/NADSYN1 rs12785878 and CYP27B1 rs4646536 may play an important role in islet autoimmunity, the preclinical phase of T1D. These findings should be replicated in larger cohorts for confirmation.

**Introduction**

Type 1 diabetes (T1D) is an autoimmune disease in which the insulin-producing β-cells of the pancreas are destroyed. There is typically a preclinical phase of circulating autoantibodies, called islet autoimmunity (IA), that precedes the clinical diagnosis of T1D.

Vitamin D insufficiency [25-hydroxyvitamin D [25(OH)D] <20 ng/mL or <50 nmol/L] has been associated with a number of extraskeletal chronic disorders (48,50), including autoimmune diseases, such as T1D, multiple sclerosis, Crohn’s disease, and rheumatoid arthritis (50,130–133). The mechanism by which vitamin D exerts its effects on these diseases is not understood completely, particularly with regard to underlying genetic risk. Determinants of circulating 25(OH)D, the inactive circulating form of vitamin D and an established marker of vitamin D status, include sun exposure and dietary intake. However, reported heritabilities of 25(OH)D of 28.8% (134) to 43% (135) suggest that genetic determinants also play a role (48,50).

In 2010, a genome-wide association study (GWAS) in approximately 30,000 individuals of European descent identified variants at four loci that were associated with 25(OH)D levels: GC rs2282679, DHCR7 rs12785878, CYP2R1 rs10741657 and CYP24A1 rs6013897 (48). A second GWAS of 25(OH)D levels confirmed the findings with GC, DHCR7, and CYP2R1 (49). These variants are located within or near genes involved in vitamin D transport (GC), cholesterol synthesis (DHCR7), and hydroxylation (CYP2R1 and CYP24A1) (48).

Cooper et al (50) more recently tested genetic variants influencing 25(OH)D metabolism for an association with both circulating 25(OH)D concentrations and T1D. They replicated the associations found in the aforementioned GWAS of the four vitamin D metabolism genes (GC, DHCR7, CYP2R1, and CYP24A1) with 25(OH)D in control subjects (48,50) and found that
*CYP27B1, DHCR7,* and *CYP2R1* were associated with type 1 diabetes (50). *CYP27B1* had previously been associated with type 1 diabetes in 2004 (51).

The Diabetes Autoimmunity Study in the Young (DAISY) has been prospectively following children at increased T1D risk for the development of IA and progression to T1D since 1993. The purpose of this study was to explore associations between seven previously studied vitamin D single nucleotide polymorphisms (SNPs) and the development of IA and progression to T1D in the prospective DAISY cohort.

**Materials and Methods**

**Subjects**

DAISY is a prospective study composed of two groups of children at increased risk for T1D who were recruited between 1993 and 2004 and are being followed up prospectively for the development of IA and T1D. One group is made up of first-degree relatives of patients with T1D, recruited between birth and 8 years of age (n = 850, 49.8%). The second group consists of infants born at St Joseph’s Hospital in Denver, Colorado, whose umbilical cord blood was screened for diabetes-susceptibility genotypes in the human leukocyte antigen (HLA) region (n = 858, 50.2%). Based on their HLA genotype, newborns were categorized into three risk groups determined by the odds of developing T1D by the age of 20 years: high, odds of developing T1D by the age of 20 years, 1:16; moderate, 1:75 in non-Hispanic whites or 1:230 in Hispanics; or low, less than 1:300. All newborns found to be at high risk and a sample of those found to be at moderate risk were asked to participate in the follow-up. The St Joseph’s Hospital newborn population is representative of the general population of the Denver metropolitan area. Details of the newborn screening, sibling and offspring recruitment, and follow-up of both cohorts have been published previously (52,53). Cord blood or the first available blood sample (depending on enrollment group) was sent to Roche Molecular Systems, Inc, for PCR-based HLA class II typing. All study protocols were approved by the Colorado Multiple Institutional Review Board, and informed consent was given by parents of all participating children.
Measurement of Autoantibodies

Autoantibodies were tested at 9, 15, and 24 months, and annually thereafter, or at their first visit and annually thereafter if the child enrolled after birth. RIAs were used to measure serum autoantibodies to insulin, glutamic acid decarboxylase-65, and IA-2 (BDC512), as previously described (54–57), with rigorous confirmation of all positive and a subset of negative results. The cutoff for positivity was established as the 99th percentile of healthy controls. Children who tested autoantibody positive were put on an accelerated testing schedule of every 3-6 months.

Cases of persistent IA were defined as those children positive for at least one islet autoantibody (IAA, glutamic acid decarboxylase-65, IA-2) on two or more consecutive visits. T1D was diagnosed by a physician and defined according to the criteria for the diagnosis of diabetes by the American Diabetes Association: hemoglobin A1c of 6.5% or greater or fasting plasma glucose of 126 mg/dL (7.0 mmol/L) or 2-hour plasma glucose of 200 mg/dL or greater (11.1 mmol/L) during an oral glucose tolerance test or in a patient with classic symptoms of hyperglycemia or hyperglycemic crisis, a random plasma glucose of 200 mg/dL or greater (11.1 mmol/L) (58).

Vitamin D SNP Genotyping

DAISY children were genotyped for seven SNPs in genes involved in components of vitamin D metabolism. DHCR7/NADSYN1 rs12785878, CYP2R1 rs10741657, CYP2R1 rs12794714, and CYP24A1 rs6013897 were genotyped based on their genome-wide significance with 25(OH)D concentrations (48). GC rs4588 and GC rs7041 were chosen because they are located in the GC gene that encodes vitamin D binding protein, which is a 52- to 59-kDA protein synthesized in the liver that binds and transports vitamin D and its metabolites [including 25(OH)D and 1,25-dihydroxyvitamin D (1,25[OH]2D)], and these two SNPs are the most widely studied GC variants (48). CYP27B1 rs4646536 was chosen based on its association with type 1 diabetes (50,136). Linkage disequilibrium was tested in our population using Haploview version
4.2 as measured by \( r^2 \), with \( r^2 = 0.508 \) for the two \( GC \) SNPs and \( r^2 = 0.487 \) for the two \( CYP2R1 \) SNPs.

The SNPs were genotyped utilizing the Taqman SNP genotype-based OpenArray platform (Applied Biosystems). Custom-designed, 48-sample arrays and normalized genomic DNA were loaded using the OpenArray AccuFill system, and cycling was performed on a GeneAmp 9700 PCR system (Applied Biosystems), all according to manufacturer protocol. Alleles were analyzed using the OpenArray SNP genotyping analysis software version 1.0.3 and Taqman Genotyper Software 2.0 (Applied Biosystems). All seven SNPs had a 95% call rate or higher.

Each SNP was tested for consistency with Hardy-Weinberg proportions using a 1-degree of freedom \( \chi^2 \) goodness-of-fit test with a value of \( P = .05 \) considered as evidence of a departure from Hardy-Weinberg equilibrium; all seven SNPs were in Hardy-Weinberg equilibrium.

**Measurement of Plasma 25(OH)D Levels**

Blood was drawn and kept from light at all times during processing. Plasma was separated immediately, snap-frozen in liquid nitrogen, and stored at -70°C until it was sent to the University of Colorado Pediatric Clinical Translational Research Center Core Laboratory, which has a certificate of proficiency from the Vitamin D External Quality Assessment Scheme. 25-Hydroxyvitamin D2 and D3 was measured by RIA (DiaSorin), with a coefficient of variation of 7.5%. Quality control was assessed via an assay of blinded duplicate samples, and excellent agreement (intraclass correlation coefficient of 0.92 in 229 pairs) was observed. The 25(OH)D levels are reported in nanograms per milliliter.

**Analysis Population**

We obtained genetic data on at least one of the seven vitamin D SNPs on 1708 non-Hispanic white children in the DAISY cohort. This included 148 children who developed persistent IA, of whom 62 went on to develop T1D. However, 17 IA cases were positive for autoantibodies on their first clinic visit; these left-censored cases were removed from the
development of IA analysis cohort but retained in the progression from IA to T1D cohort. All statistical analyses were limited to non-Hispanic whites in the DAISY cohort.

**Statistical Analyses**

SAS version 9.3 (SAS Institute Inc., Cary, NC, USA) statistical software package was used for all statistical analyses.

**Analysis of Vitamin D SNPs and Risk of IA and T1D in IA Positive Children**

SNPs were analyzed for their association both with development of IA and with progression from IA to T1D. For each model, hazard ratios (HRs) and 95% confidence intervals (CIs) were estimated using Cox proportional hazards regression. A clustered time-to-event analysis was performed treating siblings from the same family as clusters, and robust sandwich variance estimates (63) were used for statistical inference. There were 272 families with more than one child enrolled in the study and included in these analyses. Analyses of time to development of IA were adjusted for the HLA-DR genotype (HLA-DR3/4,DQB1*0302 vs other genotypes) and the presence of a first-degree relative with T1D. Analyses of time to progression to T1D were adjusted, in addition, for age at first positive autoantibody visit. The significance threshold was defined as an α of less than .05. Because our analyses were based on a priori hypotheses with SNPs previously found to be associated with 25(OH)D levels and T1D, P values were not corrected for multiple testing.

A priori, we evaluated the linearity assumption for each SNP in the Cox proportional hazards using a contrast of the coefficients from the additive model. CYP27B1 rs4646536 did not meet the linearity assumption and was therefore analyzed as a three-level categorical variable (i.e., genotypes). Because the global test for this categorical variable was statistically significant (P = .03), we did two pairwise tests using homozygosity for the major allele as the reference group. The remaining SNPs met the linearity assumption and were therefore analyzed treating the number of minor alleles as a continuous variable with the HR representing an increase in risk for each minor allele. We analyzed each vitamin D SNP in separate adjusted models.
We analyzed the relationship between plasma 25(OH)D levels and individual vitamin D SNPs using a linear mixed modeling approach in a subcohort of DAISY participants. The subcohort is a representative sample of children selected from an eligible pool in DAISY via stratified random sampling based on HLA-DR genotype and family history of type 1 diabetes. Of the 279 non-Hispanic white children selected for the subcohort, 262 children had at least one plasma 25(OH)D measurement and were genotyped for at least one of the vitamin D SNPs and therefore were included in the analyses. The linear mixed models distinguish variability between participants and the variability between repeated measurements over time within participants. Vitamin D intake and caloric intake were measured by annual food frequency questionnaires, as described previously (47). We tested each vitamin D SNP for association with 25(OH)D levels in separate models adjusted for HLA-DR3/4,DQB1*0302, first-degree relative with T1D, season of blood draw, age of study subjects, sex, average daily vitamin D intake from food and supplements, average daily intake of calories, and type of food frequency questionnaire completed [Food Frequency Questionnaire vs Youth Adolescent Questionnaire (17)]. Linear mixed models with a random intercept only, a random slope for age only, and both together were tested for the best fit based on the lowest Akaike information criteria. Linear mixed models with both a random intercept and a random slope for age represented the best fit to the data. The mixed model provides a regression coefficient, a SE, and a \( P \) value for each variable to indicate its contribution toward explaining variation in 25(OH)D.

**Results**

**Development of Islet Autoimmunity**

We first examined whether vitamin D variants were associated with the development of persistent islet autoimmunity (Table 13). Of a total of 148 IA-positive children in DAISY, 17 had to be excluded from the proportional hazards analyses of IA because they tested autoantibody positive on their first study visits (i.e., they were left censored). The mean age at first IA-positive
visit was 5.9 years, and the mean age at last follow-up visit in children who did not develop IA was 9.9 years. IA-positive children were more likely to have the HLA-DR3/4,DQB1*0302 genotype compared with children who did not develop IA.

The observed number of IA cases and the number of person-years for each genotype of the vitamin D metabolism SNPs in the DAISY non-Hispanic white population are shown in Table 14, published on The Endocrine Society’s Journal Online web site at http://jcem.endojournals.org. Adjusting for HLA-DR3/4,DQB1*0302 and first-degree relative with T1D, development of IA was associated with \textit{DHCR7/NADSYN1 rs12785878} and \textit{CYP27B1 rs4646536} \([HR 1.36, 95\% \text{ CI } 1.08, 1.73 \text{ (for each additional minor allele) and HR 0.59, 95\% \text{ CI } 0.39, 0.89 \text{ (for A/G compared with A/A genotype), respectively}]\). When both \textit{DHCR7/NADSYN1 rs12785878} and \textit{CYP27B1 rs4646536} were included in the model, both SNPs remained significantly associated with development of IA, adjusting for HLA-DR3/4,DQB1*0302 and a first-degree relative with T1D \([DHCR7/NADSYN1 rs12785878, HR 1.39, 95\% \text{ CI } 1.07, 1.80 \text{ (for each additional minor allele) and CYP27B1 rs4646536 HR 0.59, 95\% \text{ CI } 0.39, 0.89 \text{ (for A/G compared with A/A genotype)}\]).

We also created an allelic score accounting for the combined effects of \textit{DHCR7/NADSYN1 rs12785878} and \textit{CYP27B1 rs4646536}. \textit{DHCR7/NADSYN1 rs12785878} was treated continuously with an increase in the number of protective major alleles (the A allele), and \textit{CYP27B1 rs4646536} was treated dichotomously, with 0 representing individuals with no or two minor alleles and 1 representing individuals with the protective heterozygosity. The score ranged from 0 to 3, with 0 representing individuals with zero \textit{DHCR7} rs12785878 A alleles and zero or two \textit{CYP27B1} rs4646536 minor alleles and 3 representing individuals with two \textit{DHCR7} rs12785878 A alleles and heterozygosity at \textit{CYP27B1} rs4646536. The allelic score treated continuously was associated with decreased risk of islet autoimmunity (HR 0.67, 95\% CI 0.54, 0.84).
Table 13: Characteristics of the Analysis Sample by IA Status and Association of Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population (n = 1691)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Minor Allele</th>
<th>Children Positive for IA (n = 131)</th>
<th>Children Negative for IA (n = 1560)</th>
<th>Adjusted HR and 95% CI</th>
<th>P Value</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age, y&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.9 ± 4.2</td>
<td>9.9 ± 5.7</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>HLA-DR3/4,DQB1*0302</td>
<td>48 (36.6%)</td>
<td>261 (16.7%)</td>
<td>2.86 (1.96, 4.16)</td>
<td>&lt;.0001</td>
<td>1691</td>
<td></td>
</tr>
<tr>
<td>First-degree relative with T1D</td>
<td>84 (64.1%)</td>
<td>762 (48.9%)</td>
<td>1.44 (1.00, 2.07)</td>
<td>.05</td>
<td>1691</td>
<td></td>
</tr>
<tr>
<td>Sex (female)</td>
<td>68 (51.9%)</td>
<td>735 (44.4%)</td>
<td>1.17 (0.84, 1.65)</td>
<td>.35</td>
<td>1691</td>
<td></td>
</tr>
<tr>
<td>DHCR7/NADSYN1 rs12785878&lt;sup&gt;cd&lt;/sup&gt;</td>
<td>G 0.26</td>
<td>73 (55.7%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>687 (44.4%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.36 (1.08, 1.73)</td>
<td>.01</td>
<td>1677</td>
</tr>
<tr>
<td>GC rs4588&lt;sup&gt;cd&lt;/sup&gt;</td>
<td>T 0.28</td>
<td>62 (53.5%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>715 (48.2%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.11 (0.86, 1.44)</td>
<td>.43</td>
<td>1601</td>
</tr>
<tr>
<td>GC rs7041&lt;sup&gt;cd&lt;/sup&gt;</td>
<td>A 0.43</td>
<td>78 (67.2%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1027 (69.1%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.88 (0.68, 1.13)</td>
<td>.30</td>
<td>1603</td>
</tr>
<tr>
<td>CYP2RI rs10741657&lt;sup&gt;cd&lt;/sup&gt;</td>
<td>A 0.40</td>
<td>75 (65.2%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>940 (64.0%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.99 (0.76, 1.30)</td>
<td>.95</td>
<td>1583</td>
</tr>
<tr>
<td>CYP2RI rs12794714&lt;sup&gt;cd&lt;/sup&gt;</td>
<td>A 0.43</td>
<td>74 (64.9%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>987 (67.8%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.95 (0.72, 1.27)</td>
<td>.74</td>
<td>1570</td>
</tr>
<tr>
<td>CYP24A1 rs6013897&lt;sup&gt;cd&lt;/sup&gt;</td>
<td>A 0.22</td>
<td>45 (38.5%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>577 (38.8%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.10 (0.78, 1.54)</td>
<td>.60</td>
<td>1604</td>
</tr>
<tr>
<td>CYP27B1 rs4646536&lt;sup&gt;e&lt;/sup&gt;</td>
<td>G 0.32</td>
<td>62 (55.9%)</td>
<td>672 (46.0%)</td>
<td>Referent</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A/A</td>
<td>36 (32.4%)</td>
<td>638 (43.6%)</td>
<td>0.59 (0.39, 0.89)</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>13 (11.7%)</td>
<td>152 (10.4%)</td>
<td>1.03 (0.57, 1.86)</td>
<td>.92</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: MAF, minor allele frequency.

*MAF calculated for children negative for IA.

Age at first IA-positive visit in autoantibody-positive children or age at last follow-up in autoantibody-negative children.

Adjusted for HLA-DR3/4,DQB1*0302 and first-degree relative with T1D.

SNP analyzed treating the number of minor alleles as a continuous variable, with HR representing increase in risk for each additional minor allele.

*n and percentage for individuals with at least one minor allele.
Table 14: Observed Number of Cases with Endpoint and Number of Person-Years for Each Genotype of Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population

<table>
<thead>
<tr>
<th>SNP</th>
<th>Observed number of IA cases (n = 131)</th>
<th>Person-years of follow-up for Development of IA Cohort (n = 1691)</th>
<th>Observed number of T1D cases (n = 62)</th>
<th>Person-years of follow-up for Progression to T1D Cohort (n = 148)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DHCR7/NADSYN1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12785878</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/T</td>
<td>58 (44.3%)</td>
<td>9119.3</td>
<td>27 (43.6%)</td>
<td>424.9</td>
</tr>
<tr>
<td>G/T</td>
<td>60 (45.8%)</td>
<td>5717.8</td>
<td>28 (45.2%)</td>
<td>347.9</td>
</tr>
<tr>
<td>G/G</td>
<td>13 (9.9%)</td>
<td>1265.7</td>
<td>7 (11.3%)</td>
<td>113.8</td>
</tr>
<tr>
<td><strong>GC rs4588</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>54 (46.6%)</td>
<td>8080.7</td>
<td>27 (56.3%)</td>
<td>355.8</td>
</tr>
<tr>
<td>G/T</td>
<td>54 (46.6%)</td>
<td>6034.2</td>
<td>19 (39.6%)</td>
<td>380.1</td>
</tr>
<tr>
<td>T/T</td>
<td>8 (6.9%)</td>
<td>1301.0</td>
<td>2 (4.2%)</td>
<td>59.1</td>
</tr>
<tr>
<td><strong>GC rs7041</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/C</td>
<td>38 (32.8%)</td>
<td>4898.3</td>
<td>20 (40.8%)</td>
<td>258.5</td>
</tr>
<tr>
<td>A/C</td>
<td>65 (56.0%)</td>
<td>7968.5</td>
<td>24 (49.0%)</td>
<td>424.0</td>
</tr>
<tr>
<td>A/A</td>
<td>13 (11.2%)</td>
<td>2625.1</td>
<td>5 (10.2%)</td>
<td>116.9</td>
</tr>
<tr>
<td><strong>CYP2R1 rs10741657</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>40 (34.8%)</td>
<td>5335.3</td>
<td>19 (38.0%)</td>
<td>321.8</td>
</tr>
<tr>
<td>A/G</td>
<td>58 (46.5%)</td>
<td>7692.3</td>
<td>25 (50.0%)</td>
<td>337.6</td>
</tr>
<tr>
<td>A/A</td>
<td>17 (14.8%)</td>
<td>2319.3</td>
<td>6 (12.0%)</td>
<td>131.4</td>
</tr>
<tr>
<td><strong>CYP2R1 rs12794714</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>40 (35.1%)</td>
<td>4876.2</td>
<td>19 (39.6%)</td>
<td>291.6</td>
</tr>
<tr>
<td>A/G</td>
<td>53 (46.5%)</td>
<td>7664.5</td>
<td>20 (41.7%)</td>
<td>338.6</td>
</tr>
<tr>
<td>A/A</td>
<td>17 (14.8%)</td>
<td>2614.6</td>
<td>9 (18.8%)</td>
<td>165.8</td>
</tr>
<tr>
<td><strong>CYP24A1 rs6013897</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/T</td>
<td>72 (61.5%)</td>
<td>9535.5</td>
<td>30 (61.2%)</td>
<td>500.3</td>
</tr>
<tr>
<td>A/T</td>
<td>38 (32.5%)</td>
<td>5266.7</td>
<td>16 (32.7%)</td>
<td>259.3</td>
</tr>
<tr>
<td>A/A</td>
<td>7 (6.0%)</td>
<td>665.1</td>
<td>3 (6.1%)</td>
<td>36.1</td>
</tr>
<tr>
<td><strong>CYP27B1 rs4646536</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>62 (55.9%)</td>
<td>7216.4</td>
<td>29 (59.2%)</td>
<td>455.0</td>
</tr>
<tr>
<td>A/G</td>
<td>36 (32.4%)</td>
<td>6522.4</td>
<td>16 (32.7%)</td>
<td>227.9</td>
</tr>
<tr>
<td>G/G</td>
<td>13 (11.7%)</td>
<td>1461.1</td>
<td>4 (8.2%)</td>
<td>85.2</td>
</tr>
</tbody>
</table>

Abbreviations: SNP, single nucleotide polymorphism; IA, islet autoimmunity; T1D, type 1 diabetes.

Progression to T1D in Children with IA

We then examined whether vitamin D variants were associated with progression to T1D in IA-positive children (Table 15). Of the 148 IA-positive children in DAISY, 62 developed T1D; the mean age at T1D diagnosis was 8.6 years. The mean age at last follow-up visit in non-diabetic children with IA was 14.0 years.
Table 15: Characteristics of IA-Positive Subjects by T1D Status and Association of Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population (n = 148)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>IA-Positive Children Who Progressed to T1D (n = 62)</th>
<th>IA-Positive Children Who Have Not Progressed to T1D (n = 86)</th>
<th>Adjusted HR and 95% CI</th>
<th>P Value</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age, y(^a)</td>
<td>8.6 ± 3.9</td>
<td>14.0 ± 4.2</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Mean age at first IA-positive visit, y</td>
<td>3.9 ± 3.0</td>
<td>7.1 ± 4.2</td>
<td>0.87 (0.79, 0.96)</td>
<td>.01</td>
<td>148</td>
</tr>
<tr>
<td>HLA-DR3/4,DQB1*0302</td>
<td>32 (51.6%)</td>
<td>20 (23.3%)</td>
<td>2.39 (1.47, 3.90)</td>
<td>.001</td>
<td>148</td>
</tr>
<tr>
<td>First-degree relative with T1D</td>
<td>45 (72.6%)</td>
<td>56 (65.1%)</td>
<td>1.09 (0.63, 1.89)</td>
<td>.76</td>
<td>148</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>29 (46.8%)</td>
<td>46 (53.5%)</td>
<td>1.05 (0.62, 1.76)</td>
<td>.86</td>
<td>148</td>
</tr>
<tr>
<td>DHCR7/NADSYN1 rs12785878(^b,c)</td>
<td>35 (56.5%)(^d)</td>
<td>44 (51.2%)(^d)</td>
<td>0.98 (0.69, 1.40)</td>
<td>.92</td>
<td>148</td>
</tr>
<tr>
<td>GC rs4588(^b,c)</td>
<td>21 (43.8%)(^d)</td>
<td>47 (57.3%)(^d)</td>
<td>0.68 (0.42, 1.11)</td>
<td>.13</td>
<td>130</td>
</tr>
<tr>
<td>GC rs7041(^b,c)</td>
<td>29 (59.2%)(^d)</td>
<td>59 (72.0%)(^d)</td>
<td>0.76 (0.50, 1.18)</td>
<td>.22</td>
<td>131</td>
</tr>
<tr>
<td>CYP2R1 rs10741657(^b,c)</td>
<td>31 (62.0%)(^d)</td>
<td>51 (63.8%)(^d)</td>
<td>0.96 (0.68, 1.36)</td>
<td>.80</td>
<td>130</td>
</tr>
<tr>
<td>CYP2R1 rs12794714(^b,c)</td>
<td>29 (60.4%)(^d)</td>
<td>56 (69.1%)(^d)</td>
<td>0.92 (0.63, 1.35)</td>
<td>.68</td>
<td>129</td>
</tr>
<tr>
<td>CYP24A1 rs6013897(^b,c)</td>
<td>19 (38.8%)(^d)</td>
<td>31 (37.8%)(^d)</td>
<td>1.06 (0.68, 1.63)</td>
<td>.81</td>
<td>131</td>
</tr>
<tr>
<td>CYP27B1 rs4646536(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>126</td>
</tr>
<tr>
<td>A/A</td>
<td>29 (59.2%)</td>
<td>43 (55.8%)</td>
<td>Referent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/G</td>
<td>16 (32.7%)</td>
<td>25 (32.5%)</td>
<td>0.88 (0.46, 1.70)</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>4 (8.2%)</td>
<td>9 (11.7%)</td>
<td>1.21 (0.40, 3.70)</td>
<td>.73</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Age at T1D diagnosis in diabetic children or age at last follow-up in non-diabetic children.

\(^b\) SNP analyzed treating the number of minor alleles as a continuous variable, with HR representing an increase in risk for each additional minor allele.

\(^c\) Adjusted for HLA-DR3/4,DQB1*0302, first-degree relative with T1D, and age at first IA-positive visit.

\(^d\) n and percentage for individuals with at least one minor allele.
Table 16: Plasma Vitamin D Levels Predicted by Vitamin D Metabolism SNPs in DAISY Non-Hispanic White Population (n = 262)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Unadjusted</th>
<th>Adjusted&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Subjects</th>
<th>Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>95% CI</td>
<td>P Value</td>
<td>Estimate</td>
</tr>
<tr>
<td></td>
<td>(SE)</td>
<td>(SE)</td>
<td></td>
<td>(SE)</td>
</tr>
<tr>
<td>DHCR7/NADSYN1 rs12785878&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-1.15 (0.54)</td>
<td>-2.20, -0.09</td>
<td>.03</td>
<td>-1.13 (0.48)</td>
</tr>
<tr>
<td>GC rs4588&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-3.45 (0.54)</td>
<td>-4.52, -2.38</td>
<td>&lt;.0001</td>
<td>-3.11 (0.49)</td>
</tr>
<tr>
<td>GC rs7041&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-2.25 (0.53)</td>
<td>-3.29, -1.22</td>
<td>&lt;.0001</td>
<td>-2.10 (0.47)</td>
</tr>
<tr>
<td>CYP2R1 rs10741657&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.97 (0.52)</td>
<td>0.95, 3.00</td>
<td>.0002</td>
<td>2.21 (0.46)</td>
</tr>
<tr>
<td>CYP2R1 rs12794714&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-2.95 (0.51)</td>
<td>-3.95, -1.95</td>
<td>&lt;.0001</td>
<td>-2.82 (0.44)</td>
</tr>
<tr>
<td>CYP24A1 rs6013897&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-1.93 (0.69)</td>
<td>-3.28, -0.58</td>
<td>.01</td>
<td>-1.59 (0.62)</td>
</tr>
<tr>
<td>CYP27B1 rs4646536&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
</tr>
<tr>
<td></td>
<td>A/A</td>
<td>Referent</td>
<td></td>
<td>A/A</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>1.01 (0.80)</td>
<td>-0.57, 2.59</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>-0.43 (1.40)</td>
<td>-3.17, 2.31</td>
<td>.76</td>
</tr>
</tbody>
</table>

<sup>a</sup> Adjusted for HLA-DR3/4,DQB1*0302, first-degree relative with T1D, season of blood draw, age of study subjects, sex, vitamin D intake, caloric intake, and type of food frequency questionnaire completed.

<sup>b</sup> SNP analyzed treating the number of minor alleles as a continuous variable with estimate representing mean difference in 25(OH)D for each additional minor allele.
Table 17: Mean 25(OH)D Levels for Each Genotype of Vitamin D Metabolism SNPs in DAISY non-Hispanic White Population

<table>
<thead>
<tr>
<th>SNP</th>
<th>Mean Unadjusted 25(OH)D Estimate (SE)</th>
<th>Mean Adjusted 25(OH)D Estimate (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHCR7/NADSYN1 rs12785878</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/T</td>
<td>30.09 (0.47)</td>
<td>29.96 (0.42)</td>
</tr>
<tr>
<td>G/T</td>
<td>28.86 (0.62)</td>
<td>28.88 (0.56)</td>
</tr>
<tr>
<td>G/G</td>
<td>27.90 (1.17)</td>
<td>27.64 (1.04)</td>
</tr>
<tr>
<td>GC rs4588 G/G</td>
<td>31.35 (0.47)</td>
<td>31.11 (0.42)</td>
</tr>
<tr>
<td>G/T</td>
<td>28.06 (0.56)</td>
<td>28.08 (0.51)</td>
</tr>
<tr>
<td>T/T</td>
<td>24.16 (1.26)</td>
<td>24.74 (1.15)</td>
</tr>
<tr>
<td>GC rs7041 C/C</td>
<td>31.22 (0.64)</td>
<td>30.87 (0.57)</td>
</tr>
<tr>
<td>A/C</td>
<td>29.52 (0.50)</td>
<td>29.65 (0.45)</td>
</tr>
<tr>
<td>A/A</td>
<td>26.49 (0.87)</td>
<td>26.31 (0.79)</td>
</tr>
<tr>
<td>CYP2R1 rs10741657 G/G</td>
<td>28.24 (0.61)</td>
<td>27.96 (0.53)</td>
</tr>
<tr>
<td>A/G</td>
<td>29.55 (0.53)</td>
<td>29.49 (0.47)</td>
</tr>
<tr>
<td>A/A</td>
<td>32.50 (0.90)</td>
<td>32.71 (0.79)</td>
</tr>
<tr>
<td>CYP2R1 rs12794714 G/G</td>
<td>32.23 (0.63)</td>
<td>32.02 (0.55)</td>
</tr>
<tr>
<td>A/G</td>
<td>29.18 (0.50)</td>
<td>29.07 (0.44)</td>
</tr>
<tr>
<td>A/A</td>
<td>26.37 (0.83)</td>
<td>26.42 (0.72)</td>
</tr>
<tr>
<td>CYP24A1 rs6013897 T/T</td>
<td>30.33 (0.48)</td>
<td>30.15 (0.43)</td>
</tr>
<tr>
<td>A/T</td>
<td>28.32 (0.60)</td>
<td>28.51 (0.54)</td>
</tr>
<tr>
<td>A/A</td>
<td>27.01 (2.51)</td>
<td>27.29 (2.27)</td>
</tr>
<tr>
<td>CYP27B1 rs4646536 A/A</td>
<td>29.26 (0.53)</td>
<td>29.43 (0.48)</td>
</tr>
<tr>
<td>A/G</td>
<td>30.27 (0.61)</td>
<td>29.93 (0.55)</td>
</tr>
<tr>
<td>G/G</td>
<td>28.83 (1.29)</td>
<td>28.40 (1.17)</td>
</tr>
</tbody>
</table>

Abbreviations: SNP, single nucleotide polymorphism; SE, standard error.

*Adjusted for HLA-DR3/4,DQB1*0302, first-degree relative with type 1 diabetes, season of blood draw, age of study subjects, gender, vitamin D intake, caloric intake, and type of food frequency questionnaire completed.

Children who developed T1D were significantly younger when they first tested positive for an autoantibody than IA-positive children who have not progressed to T1D, 3.9 and 7.1 years, respectively. Children who developed T1D were more likely to have the HLA-DR3/4,DQB1*0302 genotype compared with children who did not progress to T1D.

The observed number of T1D cases and number of person-years for each genotype of the vitamin D metabolism SNPs in the DAISY non-Hispanic white population are shown in Table...
There was no evidence of association between any of the vitamin D SNP variants and progression to T1D in IA-positive children, adjusting for HLA-DR3/4,DQB1*0302, first-degree relative with T1D, and age at first autoantibody visit.

**Longitudinal Analysis of Plasma 25(OH)D Levels and Vitamin D SNPs**

To explore the mechanism behind these genetic associations, we examined whether individual vitamin D variants were associated with plasma 25(OH)D levels in a subcohort of DAISY. Plasma 25(OH)D levels of all visits ranged from 6.7 to 72.0 ng/mL, with a mean of 29.5 ng/mL. A total of 238 of the 2176 subcohort visits, or 10.9%, had inadequate 25(OH)D levels (≤ 20 ng/mL). Using separate linear mixed models for each vitamin D variant adjusting for HLA-DR3/4,DQB1*0302, first-degree relative with type 1 diabetes, season of blood draw, age of study subjects, gender, vitamin D intake, caloric intake, and type of food frequency questionnaire completed, six of the seven SNPs were associated with plasma 25(OH)D levels (Table 16). The one SNP not associated with plasma 25(OH)D levels was CYP27B1 rs4646536. The mean unadjusted and adjusted 25(OH)D levels for each genotype of the vitamin D metabolism SNPs in the DAISY non-Hispanic white population are shown in Table 17.

**Discussion**

In exploring the associations between seven previously studied vitamin D SNPs and the development of IA and progression to T1D in the prospective DAISY cohort, DHCR7/NADSYN1 rs12785878 and CYP27B1 rs4646536 were associated with development of IA but not associated with progression to T1D in IA-positive children, suggesting these genes play a role early in the development of T1D related to the initial appearance of autoimmunity. Up to this point, these genetic variants had been investigated only in case-control studies using T1D as case status and not in prospective studies with IA as an outcome.

DHCR7/NADSYN1 rs12785878 was recently discovered as a novel marker for association with T1D in a study of seven genes in the vitamin D pathway (50) and is a novel locus for association with vitamin D status (48). DHCR7/NADSYN1 rs12785878 is located in the second
intron of the *NADSYN1* gene 8 kb 5' from the transcription initiation site of *DHCR7* on chromosome 11q12 and is not predicted to exert functional effects such as transcription factor binding site or splicing modification (137). *DHCR7* encodes the rate-limiting enzyme 7-dehydrocholesterol reductase, which converts 7-dehydrocholesterol (7-DHC) to cholesterol in skin, removing the substrate 7-DHC from the synthetic pathway of the production of vitamin D$_3$, a precursor of 25(OH)D (48,137–140). Vitamin D$_3$ is made in the skin when 7-DHC reacts with UV light (137). The G allele of *DHCR7/NADSYN1* rs12785878 has been associated with lower levels of 25(OH)D and risk for T1D (50,137,139,140), which we also found with plasma 25(OH)D levels and risk for IA. Our findings that *DHCR7* is associated with both vitamin D levels and IA risk suggest the hypothesis that the effect of *DHCR7* is mediated by 25(OH)D levels. However, the lack of association between 25(OH)D and IA risk in DAISY that we have reported previously (47), and confirmed in our current investigation (data not shown) suggests that the effect of this *DHCR7* variant is not mediated through 25(OH)D levels, and that this enzyme may influence diabetes risk via other mechanisms.

*CYP27B1* rs4646536 was first found to be associated with T1D in a study of *CYP27B1* and *CYP24A1* gene polymorphisms (136). rs4646536 is located in intron 6 of the *CYP27B1* gene on chromosome 12q13.1-q13.3, which encodes 1α-hydroxylase, the enzyme that converts 25(OH)D into 1,25(OH)$_2$D. Although *CYP27B1* action is downstream of circulating 25(OH)D, causal variants could possibly alter their role in metabolic feedback loops or affect the speed at which 25(OH)D is metabolized. However, the observation that *CYP27B1* rs4646536 is not associated with 25(OH)D levels in our cohort, which is similar to what has been seen previously (48,50), suggests that the action of *CYP27B1* in determining risk of IA is not mediated through 25(OH)D levels.

Vitamin D is of interest in the etiology of T1D because of observations that vitamin D supplementation during infancy was inversely associated with T1D (130). In DAISY, however, we recently reported that neither childhood vitamin D intake nor 25(OH)D levels were associated
with the risk of IA or progression to T1D in IA-positive children (47). The low prevalence of vitamin D insufficiency [25(OH)D < 20 ng/mL or < 50 nmol/L] in the DAISY population (10.9%) may have been one reason we did not observe an association between 25(OH)D status and T1D and that these associations may be easier to detect in populations with a higher prevalence of vitamin D insufficiency, as was recently seen in a population of young adults (141). Results from the present study may elucidate the importance of vitamin D metabolism genes that affect the risk of IA, not necessarily through 25(OH)D levels. It is possible that other aspects of the metabolic pathway of vitamin D are more important than 25(OH)D levels in predicting the risk of IA and T1D. Other vitamin D biomarkers, such as 1,25(OH)\(_2\)D and vitamin D binding protein, should be investigated. Recently we found an interaction between rs1544410 in the vitamin D receptor gene and rs1893217 in the protein tyrosine phosphatase, non-receptor type 2 gene associated with progression from IA to T1D but not risk of IA (142). These findings combined with the current study suggest that vitamin D is acting at both phases in the natural history of T1D.

The major strength of this study is the prospective long-term follow-up from birth of children at an increased risk for T1D, which allowed us to differentiate between genetic risk factors for the appearance of autoimmunity and the subsequent progression to T1D. However, the cost of assembling and following such a unique cohort has limited the number of IA-positive children and children who progress to T1D that we could include in our analysis. Although we would have liked to explore specific, biologically relevant gene environment interactions between 25(OH)D levels (or vitamin D insufficiency) and vitamin D SNPs for association with the development of IA and the progression to T1D in children with IA, we did not have adequate power. This is the first study to observe an association between \textit{DHCR7/NADSYN1} rs12785878 and \textit{CYP27B1} rs4646536 and the development of IA. Therefore, it is important that these associations be investigated in other prospective cohorts to rule out type 1 error and lend further evidence to this association.
In conclusion, two novel intronic variants for the association with T1D, DHCR7/NADSYN1 rs12785878 and CYP27B1 rs4646536, were found to be significantly associated with the appearance of IA. Interestingly, these two variants were not found to be associated with progression to T1D in IA-positive children. These findings may offer insight concerning the complex role of vitamin D in the etiology of T1D.
CHAPTER VII

EVIDENCE OF STAGE- AND AGE-RELATED HETEROGENEITY OF NON-HLA SNPS AND RISK OF ISLET AUTOIMMUNITY AND TYPE 1 DIABETES: THE DIABETES AUTOIMMUNITY STUDY IN THE YOUNG (DAISY)

Abstract

Previously, we examined 20 non-HLA SNPs for association with islet autoimmunity (IA) and/or progression to type 1 diabetes (T1D). Our objective was to investigate fourteen additional non-HLA T1D candidate SNPs for stage- and age-related heterogeneity in the etiology of T1D. Of 1634 non-Hispanic white DAISY children genotyped, 132 developed IA (positive for GAD, insulin, or IA-2 autoantibodies at two or more consecutive visits); 50 IA positive children progressed to T1D. Cox regression was used to analyze risk of IA and progression to T1D in IA positive children. Restricted cubic splines were used to model SNPs when there was evidence that risk was not constant with age. $C1QTNF6$ (rs229541) predicted increased IA risk (HR: 1.57, CI: 1.20-2.05) but not progression to T1D (HR: 1.13, CI: 0.75-1.71). SNP (rs10517086) appears to exhibit an age-related effect on risk of IA, with increased risk before age 2 years (age 2 HR: 1.67, CI: 1.08-2.56) but not older ages (age 4 HR: 0.84, CI: 0.43-1.62). $C1QTNF6$ (rs229541), SNP (rs10517086), and $UBASH3A$ (rs3788013) were associated with development of T1D. This prospective investigation of non-HLA T1D candidate loci shows that some SNPs may exhibit stage- and age-related heterogeneity in the etiology of T1D.

Introduction

Type 1 diabetes (T1D) is a chronic autoimmune disease in which the insulin-producing beta cells of the pancreas are destroyed. There is typically a preclinical phase of circulating...
autoantibodies, called islet autoimmunity (IA), which precedes the clinical diagnosis of T1D.

T1D is widely believed to be caused by an environmental factor on a susceptible genetic background. The major susceptibility locus for T1D maps to the HLA class II genes at chromosome 6p21. These HLA class II alleles account for 30-50% of the familial clustering of T1D (143).

More than 50 non-HLA T1D susceptibility gene markers have been confirmed. The major non-HLA loci include \textit{INS} (144), \textit{CTLA4} (145), \textit{PTPN22} (146), \textit{IL2RA} (147), and \textit{IFIH1} (148). The DAISY study has previously investigated 20 non-HLA SNPs and found SNPs in \textit{PTPN22}, \textit{UBASH3A}, \textit{INS}, and \textit{IFIH1} associated with IA and/or progression to T1D (112,149–152).

Prospective birth cohorts have the unique ability to study two stages in the natural history of T1D: development of IA and progression to T1D in IA positive children. Different exposures have been associated with one or both stages. For instance, DAISY recently identified an association between a gene-gene interaction involving the vitamin D receptor gene (\textit{VDR}) and protein tyrosine phosphatase, non-receptor type 2 gene (\textit{PTPN2}) with progression to T1D in IA positive children, but not with development of IA (153). This would be an example of stage-related heterogeneity in the natural history of T1D. There is also evidence of age-related heterogeneity in the etiology of T1D when a gene or exposure is associated with the disease at certain ages, but not others. One example is a recent study that found differences in metabolite profiles relative to age, in which there was an association between lower methionine levels and presence of diabetes autoantibodies in younger onset (≤ 2 years), but not older onset (≥ 8 years) autoimmunity (12). The purpose of this analysis was to investigate stage- and age-related heterogeneity of fourteen non-HLA T1D candidate SNPs for their association with development of IA and progression to T1D in a prospective birth cohort of non-Hispanic white (NHW) children at increased genetic risk of T1D. Additionally, we investigated whether the fourteen
T1D candidate SNPs that were originally detected by GWAS using a case-control study design would be detected in time-to-event analyses of T1D risk in a prospective birth cohort.

**Materials and Methods**

**Subjects**

The Diabetes Autoimmunity Study in the Young (DAISY) is a prospective study composed of two groups of children at increased risk for T1D who were recruited between 1993 and 2004 and are being followed prospectively for the development of IA and T1D. One group is made up of first degree relatives of patients with T1D, identified and recruited between birth and eight years of age, mainly through the Barbara Davis Center for Childhood Diabetes \( n = 815 \). The second group consists of infants born at St. Joseph’s Hospital in Denver, CO, whose umbilical cord blood was screened for diabetes-susceptibility HLA-DR, DQ genotypes \( n = 819 \). Details of the newborn screening, sibling and offspring recruitment, and followup of both cohorts have been published previously (52,53). Cord blood or the first available blood sample (depending on enrollment group) was sent to Roche Molecular Systems, Inc., Alameda, CA, for PCR-based HLA-DR, DQ typing. All study protocols were approved by the Colorado Multiple Institutional Review Board, and informed consent was given by parents of all participating children.

**Measurement of Autoantibodies**

Autoantibodies were tested at 9, 15, and 24 months, and annually thereafter, or at their first visit and annually thereafter if the child enrolled after birth. Radioimmunoassays were used to measure serum autoantibodies to insulin, GAD-65, and IA-2 (BDC512), as previously described (54–57), with rigorous confirmation of all positive and a subset of negative results. The cut-off for positivity was established as the 99\({\text{th}}\) percentile of healthy controls. Children who tested autoantibody positive were put on an accelerated testing schedule of every 3-6 months.

Cases of IA were defined as those children positive for at least one islet autoantibody (IAA, GAD-65, IA-2) on two or more consecutive visits. T1D was diagnosed by a physician and
defined as random blood glucose > 200 mg/dL and/or HbA1c (A1C) > 6.5% with clinical symptoms of T1D.

Non-HLA SNP Genotyping

DAISY children were genotyped for fourteen non-HLA T1D candidate SNPs: *C1QTNF6* (rs229541), *C6orf173* (rs9388489), *C14orf181* (rs1465788), *IL2* (rs2069762), *IL2* (rs4505848), *IL2RA* (rs12722563), *IL2RA* (rs2104286), *IL7R* (rs6897932), *PRKCQ* (rs947474), *SKAP2* (rs7804356), *SMARCE1* (rs7221109), *TLR8* (rs5979785), *UBASH3A* (rs3788013), and SNP (rs10517086). Thirteen of the fourteen SNPs were chosen from three GWAS meta-analyses (154–156). *UBASH3A* (rs3788013) was chosen based on its strong LD with *UBASH3A* (rs876498), which was discovered for its association with T1D by Concannon *et al.* (157).

The SNPs were genotyped utilizing the Taqman SNP genotype based OpenArray platform (Applied Biosystems, CA, USA). Custom designed 48-sample arrays and normalized genomic DNA were loaded using the OpenArray AccuFill system and cycling was performed on a GeneAmp 9700 PCR system (Applied Biosystems, CA, USA), all according to manufacturer protocol. Alleles were analyzed using the OpenArray SNP genotyping analysis software v.1.0.3 and Taqman Genotyper Software 2.0 (Applied Biosystems, CA, USA). All fourteen SNPs had a 95% call rate or higher.

Each SNP was tested for consistency with Hardy-Weinberg proportions using a 1-degree of freedom $\chi^2$ goodness-of-fit test with a $P$ value of 0.05 considered as evidence of a departure from Hardy-Weinberg equilibrium; all fourteen SNPs were in Hardy-Weinberg equilibrium. Linkage disequilibrium (LD) was tested in our population using Haploview version 4.2 as measured by $r^2$ and $D'$, with $r^2 = 0.257$ and $D' = 0.862$ for the two *IL2RA* SNPs and $r^2 = 0.222$ and $D' = 1.0$ for the two *IL2* SNPs.

Analysis Population

We obtained genetic data on at least one of the fourteen non-HLA T1D candidate SNPs for 1634 non-Hispanic white children in the DAISY cohort. This included 132 children who
developed IA, of whom 50 went on to develop T1D. Fifteen IA cases were positive for autoantibodies on their first clinic visits; these left-censored cases were removed from the development of IA analysis cohort but retained in the progression from IA to T1D cohort. The same 50 IA positive children who went on to develop T1D are the same 50 T1D cases in the development of T1D analyses. However, not all IA positive children went on to develop T1D. All statistical analyses were limited to non-Hispanic whites in the DAISY cohort.

**Statistical Analyses**

SAS version 9.3 (SAS Institute Inc., Cary, NC, USA) statistical software package was used for all statistical analyses. SNPs were tested for violation of the proportional hazards assumption using a supremum test, with a $P$ value < 0.20 indicating possible departure from proportional hazards (158). If a SNP appeared not to meet this assumption, restricted cubic splines were used to evaluate the nature and extent of the violation. SNPs were analyzed for their association both with development of IA and with progression from IA to T1D. For each model, hazard ratios (HR) and 95% confidence intervals (CI) were estimated using Cox regression analyses. There were 302 sibling pairs in the analysis cohort, so a clustered time to event analysis was performed treating siblings from the same family as a cluster, and robust sandwich variance estimates were used for statistical inference (80). Analyses of time to development of IA were adjusted for the HLA-DR genotype (HLA-DR3/4,DQB1*0302 versus other genotypes) and presence of a first degree relative with T1D. Analyses of time to progression to T1D were adjusted, in addition, for age at first positive autoantibody visit. TLR8 (rs5979785) was additionally adjusted for sex because it is on the X chromosome. The significance threshold was defined as $\alpha = 0.05$. Because our analyses were based on a priori hypotheses with SNPs previously found to be associated with T1D, $P$ values were not corrected for multiple testing. We analyzed each non-HLA SNP in separate, covariate adjusted models. In the analyses examining development of IA, all of the SNPs were treated additively, except IL2RA (rs12722563) and PRKCQ (rs947474), which were dichotomized on the minor allele due to small sample sizes.
Additionally, in the progression to T1D analyses, SNPs were treated additively, except
*C14orf181* (rs1465788), *IL2RA* (rs12722563), *IL2RA* (rs2104286), *IL7R* (rs6897932), *PRKCQ*
(rs947474), *SKAP2* (rs7804356), and SNP (rs10517086), which were dichotomized on the minor
allele due to small sample sizes.

**Results**

**Development of IA**

We first examined whether non-HLA variants were associated with development of IA. The mean age at first IA positive visit was 6.2 years, and the mean age at last follow-up visit in children who did not develop IA was 9.9 years (*Table 18*). IA positive children were more likely to have the HLA-DR3/4,DQB1*0302 genotype compared to DAISY children who did not develop IA (HR: 2.97, 95% CI: 2.01, 4.39).

Unadjusted SNP association analyses are presented in *Table 19* (see Supplementary Material available online at http://dx.doi.org/12.1155/2013/417657).

Adjusting for HLA-DR3/4,DQB1*0302 and first degree relative with T1D, *C1QTNF6*
(rs229541) was associated with development of IA (HR: 1.57, 95% CI: 1.20, 2.05 (for each additional minor allele)) (*Table 20*). SNP (rs10517086) did not appear to meet the assumptions of proportional hazards in the development of IA analysis and therefore was modeled using a restricted cubic spline. The restricted cubic spline shows that SNP (rs10517086) is associated with an increased risk of developing IA before the age of two or in younger ages but is not associated with developing IA in older ages (*Figure 6*). Children with a minor allele developed IA significantly earlier than children with no minor alleles (mean age at onset of IA: 7.2, 5.2, and 3.2 for 0, 1, and 2 minor alleles, resp., *P* = 0.003).

**Progression to T1D in Children with IA**

We then examined whether non-HLA variants were associated with progression to T1D in IA positive children. Of the 132 IA positive children in DAISY, 50 developed T1D; the mean age at T1D diagnosis was 8.7 years (*Table 18*). The mean age at last followup visit in non-
diabetic children with IA was 14.1 years. Children who developed T1D were younger when they first tested positive for an autoantibody than IA positive children who have not progressed to T1D, 3.9 and 7.2 years, respectively ($P = 0.003$). Children with IA who developed T1D were more likely to have the HLA-DR3/4,DQB1*0302 genotype compared to children with IA who did not progress to T1D (HR: 2.79, 95% CI: 1.65, 4.72). Adjusting for HLA-DR3/4,DQB1*0302, first degree relative with T1D, and age at first IA positive visit, none of the fourteen non-HLA T1D candidate SNPs was associated with progression to T1D in IA positive children (Table 20). SNP association analyses adjusted only for age at first IA positive visit are presented in Table 19.

**Development of T1D**

In order to evaluate the same outcome as previous GWAS to see if similar associations could be seen using a time-to-event analysis in a prospective birth cohort and to better understand the role these SNPs play in the natural history of T1D, we examined whether these non-HLA variants were associated with T1D in our population of 1619 children. All 50 children who developed T1D had developed IA previously, so the same 50 T1D cases were included in both the progression from IA to T1D (presented in Progression to T1D in Children with IA) and these development of T1D analyses. However, not all IA positive children went on to develop T1D during followup. Unadjusted SNP association analyses are presented in Table 19. Adjusting for HLA-DR3/4,DQB1*0302 and first degree relative with T1D, three of the fourteen non-HLA T1D candidate SNPs were associated with development of T1D (Table 20). Two of the SNPs associated with development of T1D, *C1QTNF6* (rs229541) and SNP (rs10517086), were also associated with development of IA, but not with progression to T1D in IA positive children. The other SNP associated with development of T1D, *UBASH3A* (rs3788013), was not associated with development of IA nor progression to T1D in IA positive children.
Table 18: Demographic Characteristics of DAISY Non-Hispanic White Population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Development of islet autoimmunity (IA) ( (n = 1619) )</th>
<th>Progression from IA to type 1 diabetes (T1D) ( (n = 132) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Children positive for IA ( (n = 117) )</td>
<td>Children negative for IA ( (n = 1502) )</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td>6.2 ± 4.2(^a)</td>
<td>9.9 ± 5.7(^b)</td>
</tr>
<tr>
<td>Mean age at first IA positive visit (years)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>HLA-DR3/4,DQB1*0302</td>
<td>43 (36.8%)</td>
<td>248 (16.5%)</td>
</tr>
<tr>
<td>First degree relative with T1D</td>
<td>75 (64.1%)</td>
<td>737 (49.1%)</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>62 (53.0%)</td>
<td>712 (47.4%)</td>
</tr>
</tbody>
</table>

CI: confidence interval; DAISY: Diabetes Autoimmunity Study in the Young; HLA: human leukocyte antigen; HR: hazard ratio; IA: islet autoimmunity; T1D: type 1 diabetes.

\(^a\)Age at first IA positive visit.

\(^b\)Age at last followup.

\(^c\)Age at T1D diagnosis.
<table>
<thead>
<tr>
<th>SNP</th>
<th>Development of IA (n = 1619)</th>
<th>Progression from IA to T1D (n = 132)</th>
<th>Development of T1D (n = 1619)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR and 95% CI</td>
<td>P value</td>
<td>HR and 95% CI</td>
</tr>
<tr>
<td>CLQTNF6 (rs229541)</td>
<td><strong>1.55 (1.18, 2.02)</strong>^b</td>
<td>0.001</td>
<td><strong>1.94 (1.33, 2.84)</strong>^b</td>
</tr>
<tr>
<td>C6orf173 (rs9388489)</td>
<td>1.16 (0.88, 1.52)^b</td>
<td>0.29</td>
<td>1.05 (0.70, 1.56)^b</td>
</tr>
<tr>
<td>C14orf181 (rs1465788)</td>
<td>1.00 (0.77, 1.31)^b</td>
<td>0.98</td>
<td>0.79 (0.46, 1.39)^c</td>
</tr>
<tr>
<td>IL2 (rs2069762)</td>
<td>1.06 (0.80, 1.40)^b</td>
<td>0.69</td>
<td>1.19 (0.78, 1.81)^b</td>
</tr>
<tr>
<td>IL2 (rs4505848)</td>
<td>1.02 (0.79, 1.32)^b</td>
<td>0.88</td>
<td>1.16 (0.80, 1.68)^b</td>
</tr>
<tr>
<td>IL2RA (rs12722563)</td>
<td>0.83 (0.53, 1.30)^c</td>
<td>0.41</td>
<td>0.37 (0.15, 0.92)^c</td>
</tr>
<tr>
<td>IL2RA (rs2104286)</td>
<td>0.82 (0.60, 1.10)^b</td>
<td>0.19</td>
<td>0.66 (0.37, 1.19)^c</td>
</tr>
<tr>
<td>IL7R (rs6897932)</td>
<td>0.93 (0.68, 1.29)^b</td>
<td>0.67</td>
<td>0.84 (0.48, 1.49)^c</td>
</tr>
<tr>
<td>PRKSCQ (rs947474)</td>
<td>1.11 (0.74, 1.65)^c</td>
<td>0.62</td>
<td>0.75 (0.40, 1.42)^c</td>
</tr>
<tr>
<td>SKAP2 (rs7804356)</td>
<td>0.90 (0.66, 1.22)^b</td>
<td>0.49</td>
<td>0.96 (0.54, 1.69)^c</td>
</tr>
<tr>
<td>SMARCE1 (rs7221109)</td>
<td>0.94 (0.71, 1.23)^b</td>
<td>0.63</td>
<td>0.69 (0.41, 1.16)^b</td>
</tr>
<tr>
<td>TLR8 (rs5979785)^d</td>
<td>0.84 (0.64, 1.09)^b</td>
<td>0.18</td>
<td>0.92 (0.63, 1.34)^d</td>
</tr>
<tr>
<td>UBASH3A (rs3788013)</td>
<td>1.22 (0.92, 1.63)^b</td>
<td>0.17</td>
<td>1.76 (1.15, 2.69)^b</td>
</tr>
<tr>
<td>rs10517086</td>
<td>*</td>
<td>*</td>
<td>1.90 (1.28, 2.83)^b</td>
</tr>
</tbody>
</table>

CI: confidence interval; DAISY: Diabetes Autoimmunity Study in the Young; HLA: human leukocyte antigen; HR: hazard ratio; IA: islet autoimmunity; MAF: minor allele frequency; T1D: type 1 diabetes.

*SNP rs10517086 did not meet the assumptions of proportional hazards in the development of IA analysis.

**SNP analyzed additively with HR representing increase in risk for each additional minor allele.

^SNP analyzed dichotomously with HR representing increase in risk for at least one minor allele.

^TLR8 (rs5979785) is adjusted for sex because it is on the X chromosome.
Table 20: Association Between Non-HLA T1D Candidate SNPs and Development of IA, Progression from IA to T1D, and Development of T1D Adjusted for HLA-DR3/4, DQB1*0302 Genotype and First Degree Relative with T1D

<table>
<thead>
<tr>
<th>SNP</th>
<th>Minor allele</th>
<th>MAF&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Adjusted HR&lt;sup&gt;b&lt;/sup&gt; and 95% CI</th>
<th>(P) value</th>
<th>Adjusted HR&lt;sup&gt;c&lt;/sup&gt; and 95% CI</th>
<th>(P) value</th>
<th>Adjusted HR&lt;sup&gt;d&lt;/sup&gt; and 95% CI</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIQTNF6 (rs229541)</td>
<td>A</td>
<td>0.44</td>
<td>1.57 (1.20, 2.05)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.001</td>
<td>1.13 (0.75, 1.71)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.56</td>
<td>1.95 (1.33, 2.87)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.001</td>
</tr>
<tr>
<td>C6orf173 (rs9388489)</td>
<td>G</td>
<td>0.46</td>
<td>1.15 (0.88, 1.51)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.31</td>
<td>0.87 (0.61, 1.24)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.44</td>
<td>1.03 (0.69, 1.54)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.87</td>
</tr>
<tr>
<td>CI4orf181 (rs1465788)</td>
<td>T</td>
<td>0.28</td>
<td>0.97 (0.73, 1.27)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.80</td>
<td>1.09 (0.62, 1.92)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.77</td>
<td>0.70 (0.40, 1.23)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.21</td>
</tr>
<tr>
<td>IL2 (rs2069762)</td>
<td>C</td>
<td>0.29</td>
<td>1.08 (0.82, 1.43)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.58</td>
<td>1.22 (0.79, 1.87)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.38</td>
<td>1.20 (0.79, 1.83)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.40</td>
</tr>
<tr>
<td>IL2 (rs4505848)</td>
<td>G</td>
<td>0.35</td>
<td>1.04 (0.80, 1.34)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.78</td>
<td>0.93 (0.67, 1.28)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.65</td>
<td>1.18 (0.82, 1.70)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.37</td>
</tr>
<tr>
<td>IL2RA (rs12722563)</td>
<td>A</td>
<td>0.12</td>
<td>0.90 (0.58, 1.41)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.65</td>
<td>0.52 (0.17, 1.57)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.25</td>
<td>0.43 (0.17, 1.08)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.07</td>
</tr>
<tr>
<td>IL2RA (rs2104286)</td>
<td>C</td>
<td>0.27</td>
<td>0.85 (0.63, 1.15)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.28</td>
<td>1.26 (0.70, 2.27)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.44</td>
<td>0.77 (0.43, 1.38)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.38</td>
</tr>
<tr>
<td>IL7R (rs6897932)</td>
<td>T</td>
<td>0.28</td>
<td>0.93 (0.68, 1.28)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.66</td>
<td>0.86 (0.49, 1.53)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.61</td>
<td>0.88 (0.50, 1.55)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.65</td>
</tr>
<tr>
<td>PRKCQ (rs947474)</td>
<td>G</td>
<td>0.17</td>
<td>1.19 (0.80, 1.77)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.40</td>
<td>0.76 (0.41, 1.42)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.39</td>
<td>0.87 (0.47, 1.64)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.67</td>
</tr>
<tr>
<td>SKAP2 (rs7804356)</td>
<td>C</td>
<td>0.25</td>
<td>0.90 (0.66, 1.22)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.49</td>
<td>1.54 (0.83, 2.85)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.17</td>
<td>1.01 (0.57, 1.78)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.99</td>
</tr>
<tr>
<td>SMARCE1 (rs7221109)</td>
<td>T</td>
<td>0.35</td>
<td>0.94 (0.72, 1.22)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.63</td>
<td>0.94 (0.57, 1.57)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.82</td>
<td>0.73 (0.45, 1.20)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.21</td>
</tr>
<tr>
<td>TLR8 (rs5979785)</td>
<td>C</td>
<td>0.20</td>
<td>0.82 (0.64, 1.05)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.11</td>
<td>0.84 (0.61, 1.16)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.29</td>
<td>0.86 (0.60, 1.24)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.43</td>
</tr>
<tr>
<td>UBASH3A (rs3788013)</td>
<td>A</td>
<td>0.44</td>
<td>1.19 (0.89, 1.59)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.25</td>
<td>1.03 (0.72, 1.48)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.86</td>
<td>1.63 (1.04, 2.54)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.03</td>
</tr>
<tr>
<td>rs10517086</td>
<td>A</td>
<td>0.30</td>
<td>*</td>
<td>*</td>
<td>1.36 (0.77, 2.41)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.30</td>
<td>2.03 (1.35, 3.03)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.001</td>
</tr>
</tbody>
</table>

CI: confidence interval; DAISY: Diabetes Autoimmunity Study in the Young; HLA: human leukocyte antigen; HR: hazard ratio; IA: islet autoimmunity; MAF: minor allele frequency; T1D: type 1 diabetes.

<sup>a</sup>Minor allele frequency (MAF) calculated for children negative for IA.

<sup>b</sup>Adjusted for HLA-DR3/4, DQB1*0302 genotype and first degree relative with T1D. TLR8 (rs5979785) is additionally adjusted for sex because it is on the X chromosome.

<sup>c</sup>Adjusted for HLA-DR3/4, DQB1*0302 genotype, first degree relative with T1D, and age at first antibody positive visit. TLR8 (rs5979785) is additionally adjusted for sex because it is on the X chromosome.

<sup>d</sup>SNP analyzed additively with HR representing increase in risk for each additional minor allele.

<sup>e</sup>SNP analyzed dichotomously with HR representing increase in risk for at least one minor allele.

<sup>*SNP rs10517086 did not meet the assumptions of proportional hazards in the development of IA analysis and therefore was modeled using a restricted cubic spline (Figure 1).
In exploring associations between fourteen previously discovered non-HLA T1D candidate SNPs and the development of IA and progression to T1D in the prospective DAISY cohort, we found that \( C1QTNF6 \) (rs229541) predicts IA but not progression to T1D, demonstrating stage-related heterogeneity. Moreover, SNP (rs10517086) demonstrates age-related heterogeneity with predicting IA only in the youngest ages. These two SNPs were also associated with development of T1D in our cohort, as well as \( UBASH3A \) (rs3788013). It is possible that the observed associations between both \( C1QTNF6 \) (rs229541) and SNP (rs10517086) and development of T1D were driven by their association with development of IA.
Given that all of our T1D cases developed IA prior to clinical diagnosis, it was not possible to determine whether a gene was associated with T1D via a pathway other than IA.

*CIQTNF6* (rs229541), which was first identified through a meta-analysis of data from three genome-wide association studies (GWAS) (Combined *P* value = 1.98 x 10⁻⁸) (154), is an intronic SNP located on chromosome 22q13 between two genes, *CIQTNF6* (C1q and tumor necrosis factor related protein 6) and *SSTR3* (somatostatin receptor 3). We found that *CIQTNF6* (rs229541) is associated with development of IA but not associated with progression to T1D in IA positive children, suggesting this gene may play a role early in the development of T1D related to the initial appearance of autoimmunity.

SNP (rs10517086) exhibits an age-related effect with development of IA, with an increased risk of developing IA before the age of two or in younger ages, and a null effect in older ages. Children carrying a risk allele for SNP (rs10517086) developed IA significantly earlier than children without a risk allele. Based on these epidemiologic analyses, future studies should investigate mechanisms as to how this SNP influences risk of early autoimmunity. SNP (rs10517086), which was first discovered through another GWAS and meta-analysis of T1D, is located within a gene desert on chromosome 4. Loci in or near genes without a known function or in regions not containing annotated genes may indicate involvement of long-range gene expression regulatory elements and/or nonprotein-coding RNA genes (159).

*UBASH3A* (rs3788013) was associated with development of T1D, but not with either of the stages preceding this (development of IA and progression from IA to T1D). *UBASH3A* (rs3788013) is an intronic SNP located on chromosome 21q22 in the *UBASH3A* (ubiquitin associated and SH3 domain containing A) gene. *UBASH3A* is expressed predominantly in T-cells suppressing T-cell receptor signaling (160). *UBASH3A* has also been associated with other autoimmune diseases, such as celiac disease and rheumatoid arthritis (161). Another *UBASH3A* SNP, *UBASH3A* (rs11203203), was found to be associated with both development of IA and development of T1D in a previous DAISY analysis (112). The LD between *UBASH3A*
(rs3788013) and *UBASH3A* (rs11203203) is $r^2 = 0.491$ and $D' = 0.801$. DAISY uses two definitions of IA, one that defines IA as the presence of at least one islet autoantibody on two consecutive visits (which is the definition used in the present analysis) and the other that further requires that the children still be autoantibody positive or diabetic on their most recent visit (which is the definition used in the previous analysis). The definition used in the previous study is closer to T1D, which makes our *UBASH3A* (rs3788013) association with T1D consistent with what was previously found with *UBASH3A* (rs11203203).

In combination with those presented in this paper, DAISY has now investigated 34 non-HLA T1D candidate SNPs for association with development of IA, progression from IA to T1D, and/or development of T1D with multiple examples of stage-related heterogeneity, which are presented in Table 3. Investigating 20 non-HLA SNPs for development of IA, progression from IA to T1D, and/or development of T1D, Steck *et al.* found *PTPN22* (rs2476601) associated with development of IA, but not progression from IA to T1D, and *CTLA4* (rs231775) associated with progression from IA to T1D, but not development of IA (151). *PTPN2* (rs1893217) was only associated with development of IA, while *UBASH3A* (rs11203203) was associated with development of IA and development of T1D (112). *INS* (rs689) was not associated with development of IA nor progression from IA to T1D but was associated with development of T1D (112,151). *PTPN22* (rs2476601) was also associated with development of T1D (112). Here we investigated fourteen additional non-HLA T1D candidate SNPs and found *C1QTNF6* (rs229541) associated with development of IA and development of T1D, but not progression from IA to T1D. *UBASH3A* (rs3788013) was associated with development of T1D, but not with either of the stages preceding this (development of IA and progression from IA to T1D) and SNP (rs10517086) was associated with development of T1D, while exhibiting an age-related effect with IA risk but was not associated with progression from IA to T1D. The SNPs investigated in these three studies are summarized in Table 21. The distinction between the risk factors for islet autoimmunity versus progression to type 1 diabetes in IA positive children is important because it
may allow us to explore potentially different mechanisms of triggering islet autoimmunity versus epitope spreading and progressive loss of beta-cell mass leading to overt diabetes.

GWAS are important for identifying new candidate regions associated with a clinical outcome, such as T1D, in a large number of cases and controls. Prospective birth cohort studies, like DAISY, are then able to take these newly identified candidate regions and look for associations with different stages in the disease process and at different ages. As a prospective birth cohort study following children at increased risk for developing T1D from birth, we are able to capture the preclinical phase of T1D from birth, we are able to study two separate stages in the natural history of T1D: development of IA and progression to T1D in IA positive children. We are also able to study whether certain exposures are important at one age, but not another.

Due to the cost associated with following a large group of children from birth into adulthood, our sample sizes are much smaller than those obtained for GWAS. Our lack of association for many of these SNPs is not evidence against their association with T1D but may result from limited power, especially in the progression from IA to T1D stage. We also have a very unique population comprised of children with high risk HLA genotypes and it is possible that the effect of these SNPs differs based on one’s HLA risk status. The risk for non-HLA loci appears to be lower in individuals carrying high-risk HLA genotypes, as has been seen with PTPN22 (rs2476601) (155,162) and TCF7 (rs5742913) (163,164).

We believe that taking these GWAS identified candidate regions and studying them in the context of the natural history of T1D are central to better understanding the disease process and where in the disease process genetic loci may be important. This will allow us to create more accurate risk prediction models for both stages in the natural history of the disease, as well as inform the design of targeted interventions to prevent or slow the progression of IA and subsequent development of T1D.
Table 21: Non-HLA T1D Candidate SNPs Associated with Development of IA, Progression from IA to T1D, and/or Development of T1D in DAISY

<table>
<thead>
<tr>
<th>SNP</th>
<th>Development of IA</th>
<th>Progression from IA to T1D</th>
<th>Development of T1D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted HR and 95% CI</td>
<td>$P$ value</td>
<td>Adjusted HR and 95% CI</td>
</tr>
<tr>
<td><strong>C1QTNF6 (rs2295451)</strong></td>
<td>1.57 (1.20, 2.05)$^abc$</td>
<td>0.001</td>
<td>1.13 (0.75, 1.71)$^ab$</td>
</tr>
<tr>
<td><strong>CTLA4 (rs231775)</strong></td>
<td>1.12 (0.86, 1.46)$^{aef}$</td>
<td>0.42</td>
<td>**0.54 (0.33, 0.88)$^{aeg}$</td>
</tr>
<tr>
<td><strong>INS (rs689)</strong></td>
<td>1.39 (0.99, 1.95)$^{aef}$</td>
<td>0.05</td>
<td>1.34 (0.72, 2.52)$^{aeg}$</td>
</tr>
<tr>
<td><strong>PTPN2 (rs1893217)</strong></td>
<td>**1.42 (1.02, 1.99)$^{ach}$</td>
<td><strong>0.04</strong></td>
<td>0.65 (0.27, 1.60)$^{ijk}$</td>
</tr>
<tr>
<td><strong>PTPN22 (rs2476601)</strong></td>
<td>**1.83 (1.27, 2.63)$^{aef}$</td>
<td><strong>0.001</strong></td>
<td>0.98 (0.50, 1.93)$^{aeg}$</td>
</tr>
<tr>
<td><strong>UBASH3A (rs11203203)</strong></td>
<td>**1.46 (1.11, 1.91)$^{ach}$</td>
<td><strong>0.01</strong></td>
<td>**1.83 (1.28, 2.64)$^{ach}$</td>
</tr>
<tr>
<td><strong>UBASH3A (rs3788013)</strong></td>
<td>1.19 (0.89, 1.59)$^{abc}$</td>
<td>0.25</td>
<td>1.03 (0.72, 1.48)$^{abd}$</td>
</tr>
<tr>
<td>rs10517086</td>
<td>*</td>
<td>*</td>
<td>1.36 (0.77, 2.41)$^{bdj}$</td>
</tr>
</tbody>
</table>

DAISY: Diabetes Autoimmunity Study in the Young; HLA: human leukocyte antigen; HR: hazard ratio; CI: confidence interval; IA: islet autoimmunity; T1D: type 1 diabetes.

*a* SNP analyzed additively with HR representing increase in risk for each additional minor allele.

*b* SNP analyzed dichotomously with HR representing increase in risk for at least one minor allele.

*c* SNP rs10517086 did not meet the assumptions of proportional hazards in the development of IA analysis and therefore was modeled using a restricted cubic spline (Figure 1 of the current paper).
Conclusions

The effect of a SNP may act nonlinearly, with an effect at early ages but not later ages or vice versa. Our results provide evidence that SNP (rs10517086) is acting on early risk of IA, with the age at onset of IA occurring significantly earlier in children with a minor allele compared to children with no minor alleles. By ignoring heterogeneity in the etiology of disease, valuable associations may be missed that could aid in better understanding complex diseases, such as T1D.
CHAPTER VIII
ASSESSING AGE-RELATED ETIOLOGIC HETEROGENEITY IN THE ONSET OF ISLET AUTOIMMUNITY

Abstract

Type 1 diabetes (T1D), a chronic autoimmune disease, is often preceded by a pre-clinical phase of islet autoimmunity (IA) where the insulin-producing beta cells of the pancreas are destroyed and circulating autoantibodies can be detected. The goal of this study was to demonstrate methods for identifying exposures that differentially influence the disease process at certain ages by assessing age-related heterogeneity. The Diabetes Autoimmunity Study in the Young (DAISY) has followed 2,632 children at increased genetic risk for T1D from birth, 210 of whom developed IA. Using the DAISY population, we evaluated putative determinants of IA, including: HLA-DR3/4,DQB1*0302, first-degree relative with T1D, non-Hispanic white (NHW) ethnicity, maternal age, height growth velocity, vitamin D intake, and plasma 25-hydroxyvitamin D [25(OH)D] levels for age-related heterogeneity. The supremum test, weighted Schoenfeld residuals, and restricted cubic splines were used to assess non-proportional hazards, i.e. an age-related effect of the exposure on IA risk. NHW ethnicity, maternal age, and plasma 25(OH)D levels demonstrated a significant age-related effect on IA risk. Assessing heterogeneity in disease etiology enables researchers to identify associations that may lead to better understanding of complex chronic diseases, and thus improve prevention efforts.

Introduction

Type 1 diabetes (T1D) results from destruction of insulin-producing pancreatic beta cells. The incidence of T1D is increasing at an annual rate of about 3% worldwide (64). The most rapid increase has been in children younger than 5 (14,15,18,20,64,65).

T1D is preceded by a pre-clinical phase of islet autoimmunity (IA) where the body produces antibodies (IAA, GAD65, or IA-2) against the insulin-producing beta cells of the pancreas, which can be detected as early as 6 months of age (1). There appears to be two peaks in
IA incidence at approximately 1 to 2 years of age and in adolescence, with distinct characteristics at each peak (1).

IA and T1D development may be subject to age-related etiologic heterogeneity, where exposures influence the disease process more strongly at certain ages. T1D development is more likely to occur earlier in life for those with disease-associated HLA genotypes and a parental history of T1D (5–11). A recent study found differences in serum metabolite profiles relative to age; an association between lower methionine levels and presence of diabetes autoantibodies in younger onset (≤ 2 years), but not older onset (≥ 8 years) autoimmunity was described (12). Additionally, Virtanen et al. found early introduction of wheat, rye, oats and/or barley cereals, and egg, were associated with increased IA risk, but only during the first 3 years of life, suggesting an age-related effect (43).

Assessment of age-related heterogeneity allows understanding of if and when exposures play a role in the disease process. Valuable associations may be overlooked if their effects are averaged over time and not evaluated for heterogeneity. Knowing when exposures play a role in the disease process can guide treatment and prevention efforts by creating more accurate risk prediction models, and informing the design of targeted interventions.

Prospective cohorts of children at increased T1D risk are often followed from birth to IA and T1D development. Time-to-event analyses, frequently based on Cox proportional hazards (PH) regression, are utilized to identify risk factors. A Cox PH model assumes the hazard ratio (HR) is constant over time, meaning the effect of a covariate is the same at all time points. If age-related heterogeneity is present for a given variable, the effect of that variable changes over time (i.e., age) and the PH assumption is not valid. Therefore, age-related heterogeneity can be assessed by evaluating the PH assumption. We demonstrate the use of three methods for testing and modeling non-PH: the supremum test, weighted Schoenfeld residuals, and restricted cubic splines.
**Supremum Test**

The supremum test, a regression diagnostic for PH models, plots the path of the observed cumulative sum of martingale residuals for a covariate against time (158). Rather than a test statistic, it produces a $P$-value which represents the percentage of 1000 simulated paths embodying the PH assumption whose supremum (or largest) values exceed the supremum of the observed path for the covariate of interest (158). Higher $P$-values (ideally much greater than 0.05) are better, suggesting the supremum of the observed path is substantially smaller than a large proportion of the suprema of the 1000 simulated paths that actually follow the PH assumption for that covariate (158). The test is implemented in SAS PROC PHREG.

**Weighted Schoenfeld Residuals**

Weighted Schoenfeld residuals can be plotted as another PH regression diagnostic as described by Grambsch and Therneau (1994) (60). In the R package using the cox.zph function of the `survival` library, these residuals produced separately for each covariate for each individual are visualized through scatterplot smoothing. This effectively shows how the regression coefficient, $\beta(t)$, varies with time (60). If the assumption of PH is satisfied, the residuals will be independent of time; thus, a non-zero slope indicates a violation of the PH assumption.

**Restricted Cubic Splines**

Restricted cubic splines (RCS), piecewise polynomials smoothly joined at $k$ knot values, can also be used to identify and model non-PH (61). RCS provide a visual assessment of the HR as a function of time and allow for flexible modeling of the HR without a specific functional form, e.g. linear or quadratic. The number of knots selected for the splines is chosen based on Akaike information criterion (AIC), where a lower value indicates better fit. The SAS RCS macro, designed to assess PH for fixed covariates, first tests whether the covariate of interest is associated with the event. If the covariate is associated with the event, it can then test whether the association is non-constant with time (indicating a violation of the PH assumption), and, if so, whether the relationship between the HR and time is linear or not (61).
Diabetes Autoimmunity Study in the Young (DAISY)

DAISY has prospectively collected 20 years of data from birth on children at increased genetic risk for T1D. DAISY data can be used to study prospective exposures across childhood, and assess whether the risk associated with determinants of IA and T1D differs by the age at which the child develops IA and/or T1D. T1D has a complex etiology with numerous identified factors that either increase or decrease disease risk. The goal of this study was to demonstrate three methods of identifying exposures that influence the disease process at varying ages by assessing age-related heterogeneity (or lack of PH) of putative IA determinants, including: HLA-DR3/4,DQB1*0302 genotype, first-degree relative with T1D (FDR) status, non-Hispanic white (NHW) ethnicity, maternal age, height growth velocity, vitamin D intake from food and supplements, and plasma 25-hydroxyvitamin D [25(OH)D] levels. We were interested in assessing age-related heterogeneity of IA because if risk factors for IA are shown to have different effects by age, then age-appropriate interventions can be designed to prevent or slow the development of IA and ultimately, T1D.

Methods

Study Population

DAISY recruited two groups of children between 1993 and 2004, who are at increased risk for T1D and followed prospectively for IA and T1D development. One group is first-degree relatives of patients with T1D, identified and recruited between birth and age 8, mainly through the Barbara Davis Center for Childhood Diabetes. The second group is infants born at St. Joseph’s Hospital in Denver, CO, whose umbilical cord blood was screened for diabetes-susceptibility HLA-DR,DQ genotypes and recruited if they had these genotypes (77,78). Details of the newborn screening, sibling and offspring recruitment, and follow-up of both cohorts have been published previously (53). Cord blood or the first available blood sample (depending on enrollment group) was sent to Roche Molecular Systems, Inc., Alameda, CA, for PCR-based
HLA-DR,DQ typing. All study protocols were approved by the Colorado Multiple Institutional Review Board, and informed consent was given by parents of all participating children.

The DAISY cohort is composed of 2,632 children, of whom 210 have developed IA. Thirty-five IA cases were positive for autoantibodies on their first clinic visits; these left-censored cases were removed from the analysis. Plasma 25(OH)D levels were investigated in a case-cohort design, which is a representative sample of 380 children (i.e. subcohort) selected from the DAISY cohort via stratified random sampling based on HLA-DR genotypes and family history of type 1 diabetes. During follow-up, 24 children with 25(OH)D measurements developed IA within the subcohort. We supplemented these with 69 children who developed IA outside the subcohort. Therefore, 93 children with IA and 352 IA negative children were included in the analysis of plasma 25(OH)D levels.

We examined the following fixed covariates for age-related heterogeneity: HLA-DR3/4,DQB1*0302 genotype, FDR status, NHW ethnicity, and maternal age at birth. The time-varying covariates height growth velocity, vitamin D intake from food and supplements, and plasma 25(OH)D levels were also assessed for age-related heterogeneity.

**Height Growth Velocity**

Height was first measured when the child was able to stand cooperatively around 2 years of age and at subsequent visits, using a stadiometer with a precision of ±1 mm. The method for calculating instantaneous growth velocity for height using best linear unbiased predictors from a linear mixed model is described in Lamb et al. (2009) (31).

**Measurement of Dietary Vitamin D Intake**

Dietary vitamin D intake was evaluated using one of two instruments depending on the child’s age. For children age 2-9, parents completed a food frequency questionnaire (FFQ) (Nutrition Questionnaire Service Center, Boston, MA, USA) annually on their child’s behalf. Starting between age 10 and 12, children completed the Youth/Adolescent Questionnaire (YAQ), a food frequency questionnaire based on the FFQ, designed for adolescents. Both questionnaires
collected average food consumption over the previous year, using structured responses regarding how often a child consumed commonly sized portions (example: 8 oz. glass of milk), with response options ranging from ‘never/less than once per month’ to ‘6+ servings per day’. An instrument-comparison study conducted using the DAISY population determined data from these two instruments may be combined when an instrument indicator variable (i.e. type of food frequency questionnaire [FFQ vs. YAQ]) is included in the model (165). The FFQ has been validated against multiple 24 hour recalls (166) and correlates with micronutrient (167) and fatty acid (168) biomarkers in the DAISY population. Dietary data were linked to an autoantibody measurement if the 1-year time period of the questionnaire encompassed the time directly preceding the visit at which the autoantibody was measured.

**Measurement of Plasma 25(OH)D Levels**

Blood was drawn and kept from light at all times during processing. Plasma was separated immediately, snap frozen in liquid nitrogen, and stored at -70°C until it was sent to the University of Colorado Pediatric Clinical Translational Research Center Core Laboratory, which has a certificate of proficiency from the Vitamin D External Quality Assessment Scheme. 25-hydroxyvitamin D2 and D3 was measured by RIA (DiaSorin), with a coefficient of variation of 7.5%. Quality control was assessed via an assay of blinded duplicate samples, and excellent agreement (intraclass correlation coefficient of 0.92 in 229 pairs) was observed. The 25(OH)D levels are reported in nanograms per milliliter.

**Measurement of Autoantibodies**

Autoantibodies were tested at 9, 15, and 24 months, or at their first visit, and annually thereafter, depending on enrollment. Radioimmunoassays measured serum autoantibodies to insulin, glutamic acid decarboxylase (GAD)65, and insulinoma antigen (IA-2) (also known as BDC512), as previously described (55–57), with confirmation of all positive and a subset of negative results. Cut-off for positivity was established as the 99th percentile of healthy controls. Children testing autoantibody positive were put on an accelerated testing schedule of every 3–6
months. IA cases were defined as children positive for at least one autoantibody (IAA, GAD_{65}, IA-2) on ≥ 2 consecutive visits within 6 months.

**Statistical Analyses**

**Assessment of the PH Assumption**

We assessed violation of the PH assumption using three different methods: the supremum test, weighted Schoenfeld residuals, and restricted cubic splines. The supremum test was performed in SAS version 9.3 (SAS Institute, Cary, NC) using the ASSESS statement in the PHREG procedure. Weighted Schoenfeld residuals were plotted using the cox.zph function in R 2.15.2 (59). A significant supremum test ($P<0.30$) or a non-zero slope for the loess smoothed curve of the weighted Schoenfeld residuals indicated a violation of PH (60). Using the weighted Schoenfeld residuals, a global test of PH was assessed first; if the global test $P$-value was not large ($<0.30$), the individual covariate tests of PH were used to identify the source(s) of the non-PH. Restricted cubic splines were modeled using the RCS macro in SAS (61). Due to a limited number of events, the number of knots for the RCS was selected to be 3, placed at the 5th, 50th, and 95th percentiles of age of the IA cases; this minimizes the number of coefficients to fit the RCS models (61). The RCS macro provides the three statistical tests described above which should be performed hierarchically. The first test has 3 df for a 3-knot spline model and tests whether the covariate of interest is associated with the event. If the null hypothesis is rejected ($P<0.05$), the second statistical test with 2 df for a 3-knot spline model can be performed to determine whether the association is non-constant with time ($P<0.05$ indicating a violation of PH). Finally, if the null hypothesis is rejected for both the first and second statistical tests, the third statistical test with 1 df for a 3-knot spline model can be performed to determine whether the relationship between the HR and time is linear ($P<0.05$ indicates non-linearity) (61).

**Fixed Covariates**

All statistical analyses were conducted using SAS software, Version 9.3 of the SAS System for Windows. Copyright © 2002-2010 SAS Institute Inc. SAS and all other SAS Institute
Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA. The only exception was the weighted Schoenfeld residuals, which were generated and plotted in R (62). The three methods (supremum test, weighted Schoenfeld residuals, and RCS) were evaluated for the fixed covariates: HLA-DR3/4,DQB1*0302 genotype, FDR status, NHW ethnicity, and maternal age.

**Time-Varying Covariates**

Existing tools are limited with regard to examination of PH with time-varying covariates. The supremum test for violation of the PH assumption can theoretically accommodate time-varying covariates, but requires higher dimensional plots for time-varying covariates. The `cox.zph` function used to plot the weighted Schoenfeld residuals is valid for time-varying covariates; however, the software assumes the variance of the time-varying covariate is constant over time (60). If this assumption is violated, results from the weighted Schoenfeld residual tests are not reliable. Motivated by these limitations in assessing PH with time-varying covariates, Kroehl et al. (unpublished) recently adapted RCS for use with time-varying covariates and evaluated their performance in identifying and modeling a non-constant HR. Using this approach, non-PH were investigated for height growth velocity and vitamin D intake from food and supplements.

**Results**

**Table 22** describes the DAISY children by IA status. Of 210 IA positive children in DAISY, 35 were excluded from analyses of IA because they tested autoantibody positive on their first study visit (i.e., left-censored). Mean age at first IA positive visit was 6.4 years, and 9.0 years at last follow-up visit in those without IA. Children who developed IA were more likely to have the HLA-DR3/4,DQB1*0302 genotype and be a FDR compared to children who did not develop IA (**Table 22**).
Table 22: Characteristics of DAISY Cohort by IA Status

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Developed IA (n = 175)</th>
<th>Did Not Develop IA (n = 2422)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>6.4 (4.3)(^a)</td>
<td>9.0 (5.8)(^b)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>HLA-DR3/4,DQB1*0302 genotype</td>
<td>68 (38.9%)</td>
<td>472 (19.5%)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>First-degree relative with T1D</td>
<td>102 (58.3%)</td>
<td>1029 (42.5%)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Female</td>
<td>90 (51.4%)</td>
<td>1156 (47.7%)</td>
<td>0.34</td>
</tr>
<tr>
<td>Race/ethnicity, non-Hispanic white</td>
<td>137 (78.3%)</td>
<td>1727 (72.3%)</td>
<td>0.08</td>
</tr>
<tr>
<td>Maternal age, mean (SD), y</td>
<td>30.3 (5.8)</td>
<td>29.9 (5.7)</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Abbreviations: DAISY, Diabetes Autoimmunity Study in the Young; HLA, human leukocyte antigen; IA, islet autoimmunity; T1D, type 1 diabetes.
\(^a\)Age at first IA positive visit.
\(^b\)Age at last follow-up.

**Fixed Covariates**

We first assessed age-related heterogeneity of variables on which the DAISY cohort was sampled: HLA-DR3/4,DQB1*0302 and FDR status. As a reminder, higher P-values (>0.30) are an indication that PH are met. The supremum test P-value for the HLA-DR3/4,DQB1*0302 genotype adjusting for FDR status was 0.26, indicating a possible violation of PH (Table 23).

The supremum test P-value for FDR status adjusting for the HLA-DR3/4,DQB1*0302 genotype was 0.01, also indicating a violation of PH (Table 23). The weighted Schoenfeld residuals global test P-value for the model including the HLA-DR3/4,DQB1*0302 genotype and FDR status was 0.31, not necessarily indicating a violation of PH (Figure 7). Therefore, the individual covariate PH tests were not assessed. Modeling the RCS for the HLA-DR3/4,DQB1*0302 genotype adjusted for FDR status showed a significant overall effect with risk of IA (Effect P<0.0001) (Figure 9A). Since the null hypothesis was rejected, we also tested whether the effect was non-constant and the P-value was 0.17, indicating PH (Table 23). The non-linear effect was not tested because the null hypothesis for the non-constant effect was not rejected.

The RCS for FDR status adjusted for the HLA-DR3/4,DQB1*3020 genotype indicated a significant effect overall with IA risk (Effect P=0.001) (Figure 9B). Therefore, the non-constant effect was tested, which did not indicate a violation of PH (Non-constant P=0.31). The RCS for FDR status showed a significant increased IA risk in early childhood (age 2 HR: 2.49, 95%
confidence interval (CI): 1.51-4.12) that diminished while still remaining elevated, but not
significant, with increasing age (age 11 HR: 1.47, 95% CI: 0.93-2.32) (Figure 9B).

Two other demographic variables were assessed for age-related heterogeneity: NHW
ethnicity and maternal age. The supremum test $P$-value was 0.01 for NHW ethnicity adjusting
for HLA-DR3/4,DQB1*0302 genotype and FDR status, indicating a violation of PH (Table 23).
The weighted Schoenfeld residuals had a global PH test $P=0.04$ and an individual PH test
$P=0.01$, indicating a violation of PH (Figure 8A). We modeled the RCS to evaluate the HR as a
function of time. NHW ethnicity had an overall significant effect with risk of IA adjusting for
HLA-DR3/4,DQB1*0302 genotype and FDR status (Effect $P=0.046$) (Figure 10A). Based on
rejection of the null hypothesis, the non-constant effect was tested, producing a $P=0.02$,
indicating non-PH (Table 23). Finally, a non-linear effect was tested based on rejection of the
null hypothesis for the non-constant effect, which was not significant (Non-linear $P=0.62$),
indicating a linear decrease in IA risk associated with NHW ethnicity over time (Table 23). The
restricted cubic spline demonstrated an elevated risk in early childhood (age 2 HR: 1.68, 95% CI:
0.88-3.22) diminishing with increasing age (age 11 HR: 0.67, 95% CI: 0.41-1.11).

The supremum test $P$-value for maternal age was 0.002 adjusting for HLA-
DR3/4,DQB1*0302 genotype and FDR status, indicating a violation of PH (Table 23). The
weighted Schoenfeld residuals for maternal age also had a significant global PH test $P=0.01$ and
an individual PH test $P=0.003$, another indication of a PH violation (Figure 8B). The modeled
RCS resulted in a significant overall effect of maternal age with IA risk adjusting for HLA-
DR3/4,DQB1*0302 genotype and FDR status (Effect $P=0.001$) (Figure 10B). Based on rejection
of this null hypothesis, the non-constant effect of maternal age was tested with a resulting
$P=0.0004$, demonstrating age-related heterogeneity (Table 23). Finally, based on rejection of the
null hypothesis for the non-constant effect test, a non-linear effect was tested. In contrast with the
restricted cubic spline result for NHW ethnicity, the non-constant effect of maternal age for a
five-year difference was also non-linear (Non-linear $P=0.01$) with greater maternal age associated
with increased risk in early childhood (age 2 HR: 1.03, 95% CI: 0.98-1.07), which became protective later in adolescence (age 11 HR: 0.97, 95% CI: 0.93-1.01).

**Time-varying Covariates**

We assessed age-related heterogeneity for three time-varying covariates: height growth velocity, vitamin D intake from food and supplements, and plasma 25(OH)D levels. Restricted cubic spline models were fit for each of the time-varying covariates separately (Figure 11). Height growth velocity had an overall significant effect with IA risk (Effect $P=0.01$) adjusting for HLA-DR3/4,DQB1*0302 genotype, FDR status, sex, and NHW ethnicity, previously established in DAISY by Lamb et al. (2009) in children < 11.5 years of age (31). With additional years of follow-up into adolescence, the effect of height growth velocity on risk of IA did not exhibit a non-constant effect (Non-constant $P=0.67$) (Figure 11A).

Vitamin D intake from food and supplements did not have an overall effect on IA risk (Effect $P=0.94$), previously reported in DAISY by Simpson et al. (2011) (47). The restricted cubic spline displayed a straight line parallel to zero, indicating PH (age 2 HR: 1.08, 95% CI: 0.76-1.53 and age 11 HR: 1.05, 95% CI: 0.82-1.35) (Figure 11B).

Plasma 25(OH)D levels exhibited age-related heterogeneity with a significant overall effect on IA risk (Effect $P=0.001$), a significant non-constant effect (Non-constant $P=0.0003$), and a significant non-linear effect (Non-linear $P=0.047$) adjusting for HLA-DR3/4,DQB1*0302 genotype, FDR status, and NHW ethnicity. The restricted cubic spline demonstrated an increased IA risk early with increasing plasma 25(OH)D levels, which became protective with increasing age (age 2 HR: 1.73, 95% CI: 1.20-2.50 and age 11 HR: 0.62, 95% CI: 0.43-0.89) (Figure 11C).
### Table 23: Assessment of the Proportional Hazards Assumption in DAISY Cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Supremum test $P$</th>
<th>Schoenfeld residuals $P$</th>
<th>Schoenfeld residuals individual covariate $P$</th>
<th>Restricted cubic splines effect $P$</th>
<th>Restricted cubic splines non-constant $P$</th>
<th>Restricted cubic splines non-linear $P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DR3/4,DQB1*0302 genotype$^a$</td>
<td>0.26</td>
<td>0.31$^b$</td>
<td>*&lt;.0001$^c$</td>
<td>0.17$^c$</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>First-degree relative with T1D$^d$</td>
<td>0.01</td>
<td>0.31</td>
<td>*0.001$^e$</td>
<td>0.31$^e$</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white ethnicity$^f$</td>
<td>0.01</td>
<td>0.04</td>
<td>0.01$^g$</td>
<td>0.046$^h$</td>
<td>0.02$^h$</td>
<td>0.62$^h$</td>
</tr>
<tr>
<td>Maternal age$^i$</td>
<td>0.002</td>
<td>0.01</td>
<td>0.003$^j$</td>
<td>0.001$^j$</td>
<td>0.0004$^j$</td>
<td>0.01$^j$</td>
</tr>
<tr>
<td>Height growth velocity$^k$</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.01$^l$</td>
<td>0.67$^l$</td>
<td>N/A</td>
</tr>
<tr>
<td>Vitamin D intake from food and supplements$^m$</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.94$^n$</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Plasma 25-hydroxyvitamin D levels$^o$</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.001$^p$</td>
<td>0.0003$^p$</td>
<td>0.047$^p$</td>
</tr>
</tbody>
</table>

Abbreviations: DAISY, Diabetes Autoimmunity Study in the Young; HLA, human leukocyte antigen; T1D, type 1 diabetes.

$^a$Schoenfeld residual individual covariate $P$ value not calculated because the Schoenfeld residual global $P$-value was not <0.30.

N/A: Statistical tests were not performed based on the inability to reject the null hypothesis of the prior test in the hierarchical structure.

$^b$Adjusted for first-degree relative with T1D.

$^c$Figure 1.

$^d$Adjusted for HLA-DR3/4,DQB1*0302 genotype.

$^e$Figure 3A.

$^f$Adjusted for HLA-DR3/4,DQB1*0302 genotype and first-degree relative with T1D.

$^g$Figure 3B.

$^h$Adjusted for HLA-DR3/4,DQB1*0302 genotype and first-degree relative with T1D.

$^i$Figure 2A.

$^j$Figure 4A.

$^k$Adjusted for HLA-DR3/4,DQB1*0302 genotype, first-degree relative with T1D, sex, and non-Hispanic white ethnicity.

$^l$Figure 2B.

$^m$Adjusted for HLA-DR3/4,DQB1*0302 genotype, first-degree relative with T1D, reported energy intake, and the type of survey completed.

$^n$Figure 4B.

$^o$Adjusted for HLA-DR3/4,DQB1*0302 genotype, first-degree relative with T1D, and non-Hispanic white ethnicity.

$^p$Figure 5A.
The weighted Schoenfeld residual plot for the model with HLA-DR3/4, DQB1*0302 genotype and first-degree relative (FDR) with type 1 diabetes status produces a global PH test $P$-value of 0.31, not indicating a violation of the PH assumption.

Figure 7: Weighted Schoenfeld Residual Plot for the Model with HLA-DR3/4, DQB1*0302 Genotype and First-Degree Relative (FDR) with Type 1 Diabetes (T1D) Status. The weighted Schoenfeld residual plot for the model with HLA-DR3/4, DQB1*0302 genotype and first-degree relative (FDR) with type 1 diabetes status produces a global PH test $P$-value of 0.31, not indicating a violation of the PH assumption.
Figure 8: Weighted Schoenfeld Residual Plots for Non-Hispanic White Ethnicity (NHW) and a 5 Year Difference in Maternal Age. The weighted Schoenfeld residual plots are displayed for non-Hispanic white ethnicity (NHW) (8A) and a 5 year difference in maternal age (8B) in the prospective DAISY cohort. The global PH test $P$-values based on the Schoenfeld residuals are 0.04 and 0.01 for NHW ethnicity and a 5 year difference in maternal age, respectively, also indicating a violation of the PH assumption.
Figure 9: Restricted Cubic Splines for the High Risk HLA-DR3/4, DQB1*0302 Genotype and First-Degree Relative with Type 1 Diabetes Status. The restricted cubic spline for the high risk HLA-DR3/4,DQB1*0302 genotype (9A) in the prospective DAISY cohort demonstrates a significantly increased risk of islet autoimmunity (IA) associated with the high risk HLA-DR3/4,DQB1*0302 genotype across childhood with higher risk at the peaks of IA incidence: early infancy and puberty. The restricted cubic spline for first-degree relative with type 1 diabetes status and IA development (9B) demonstrates an elevated IA risk across childhood that is higher in the younger ages.
Figure 10: Restricted Cubic Splines for Non-Hispanic White Ethnicity and a 5 Year Difference in Maternal Age. The restricted cubic spline for non-Hispanic white ethnicity exhibits an increased risk of islet autoimmunity (IA) early on that then becomes protective in the older ages (10A). The restricted cubic spline for a 5 year difference in maternal age exhibits a slightly elevated risk of IA in early childhood that becomes protective in adolescence (10B).
Figure 11: Restricted Cubic Splines for Height Growth Velocity, Vitamin D Intake from Food and Supplements, and Plasma 25(OH)D
Figure 11: Restricted Cubic Splines for Height Growth Velocity, Vitamin D Intake from Food and Supplements, and Plasma 25(OH)D. The restricted cubic spline for height growth velocity and development of islet autoimmunity in the prospective DAISY cohort displays an increased risk across childhood that remains relatively constant or proportional (11A). The restricted cubic spline for vitamin D intake from food and supplements demonstrates proportional hazards with a constant ratio across childhood (11B). The restricted cubic spline for plasma 25-hydroxyvitamin D exhibits increased IA risk until age 3 years, which then becomes significantly protective between ages 7.75 and 12.5 years.
Discussion

Two peaks in IA incidence during infancy and adolescence motivated the assessment of age-related heterogeneity to determine whether risk factors differed for these two periods. RCS aid in visualizing risk over time and provide statistical tests about whether risk is non-constant with respect to time or age.

The assumption of PH is central to Cox PH analysis, and if violated, inferences made from an analysis may be incorrect. For example, the effect of an age-sensitive risk factor could be missed if the effect, averaged over time, is determined to be non-significant. Three of the seven variables evaluated in our analysis were shown to violate the PH assumption, indicating it may be somewhat common for risk factors to have time-varying effects. NHW ethnicity, maternal age, and plasma 25(OH)D levels that demonstrated significant age-related heterogeneity with statistically significant P-values for the non-constant effect have not been previously shown to demonstrate etiologic heterogeneity in T1D. It is important to recognize even if a covariate does not have a statistically significant P-value for the non-constant effect when assessing non-PH with RCS, marginal associations may be clinically interesting or meaningful. RCS also provide HRs and 95% CIs for any time point or age of interest. Attributable risks can then be calculated to determine whether the number of cases attributed to the disease at a certain age is meaningfully different from other ages (169–171).

We were interested in assessing age-related heterogeneity of IA, because if IA risk factors can be identified, preventive efforts may be designed to slow or even prevent progression to T1D. Also, for purposes of assessing age-related heterogeneity using RCS, the number of T1D cases is quite small in our prospective study, as it is for other prospective studies following children at increased risk of T1D. The larger number of IA cases allows for greater power for assessing age-related heterogeneity using RCS. The case definition we used for IA requires one to be positive for at least one autoantibody (insulin, GAD-65, and IA-2) on two consecutive visits. However, subjects who develop only one autoantibody may lose this autoantibody and are
less likely to develop T1D (1,172–174). We also tested a stricter definition of IA in which the subject had to be positive for at least one autoantibody on two consecutive visits and remain autoantibody positive, or develop diabetes by their last follow-up visit, which excluded subjects who subsequently became negative for their autoantibody. Even though this stricter definition decreased the number of IA outcomes, it may have increased our power by decreasing misclassification of the outcome, resulting in larger effect sizes (data not shown). However, our analyses of the covariates using the stricter IA definition led to similar conclusions; therefore, we present the original IA case definition in this paper to maximize our number of IA outcomes.

Assessing age-related heterogeneity is a valuable step in understanding etiologic heterogeneity of complex diseases and ensuring important associations are not overlooked. We encourage researchers using Cox PH regression analyses to assess covariates for violation of the PH assumption when using them in a Cox PH model. As a preliminary step preceding assessment of PH, continuous covariates should be examined for correct functional form with all important covariates and confounders included in the model (175). If the variable is not linearly associated with the HR, a transformation may be appropriate (such as a log transform) prior to assessing for PH. Martingale residuals are a common diagnostic tool for evaluating functional form. It should also be recognized that missing covariates can also erroneously induce non-proportionality (175).

Regarding the three methods for testing the PH assumption presented in this paper, we suggest the supremum test or the weighted Schoenfeld residuals for diagnostic testing of the PH assumption to determine if a fixed covariate can be used in a Cox PH model. However, if one is interested in examining the nature of age-related heterogeneity, we recommend modeling RCS and performing the hierarchical testing of the PH assumption by first determining whether the covariate of interest is associated with the event. If the null hypothesis is rejected and the covariate is associated with the event, a statistical test can be performed to determine whether the association is non-constant with time, violating the assumption of PH. Using this approach,
DAISY recently detected a single nucleotide polymorphism, rs10517086, that demonstrated age-related heterogeneity with IA risk, with increased risk before 2 years of age (age 2 HR: 1.67, 95% CI: 1.08-2.56), but not older ages (age 4 HR: 0.84, 95% CI: 0.43-1.62) (176). RCS are a powerful way of visualizing the true form of a variable with time with the ability to test whether this effect exhibits a non-constant or non-linear effect. They have the added advantage of being applicable to time-varying covariates. This method can be used by anyone interested in etiologic heterogeneity by age or time.
CHAPTER IX
DISCUSSION

The goal of this research was to better understand well studied, but inconsistently associated determinants in the etiology of T1D. One reason exposures may be inconsistently associated across study cohorts is because they play a role in the disease process at certain ages, but not others. This concept is called age-related heterogeneity. If an exposure is studied in a younger cohort and found to be associated with the outcome and then studied in an older cohort and is not found to be associated with the outcome, this inconsistency across cohorts may be an example of an exposure that exhibits age-related heterogeneity. If there is age-related heterogeneity, the effect of an exposure is averaged over time when using standard analytic approaches, when in fact it plays a role in the disease process at certain ages, but not others. Failure to adequately assess age-related heterogeneity could lead to false negative associations with important risk factors. A number of genetic and environmental factors are being elucidated in T1D and are believed to be important at different ages in the natural history of T1D. By studying factors at different ages in the etiology of T1D, we hoped to gain a better understanding of if and when these exposures play a role in the disease process.

The overarching aim of this thesis was to determine whether early onset IA is associated with different risk factors than late onset IA (i.e., age-related heterogeneity). To address our overarching aim, we used data from a prospective cohort study of children at increased genetic risk for T1D to more fully understand the mechanism of three well-studied, but inconsistently associated determinants in the etiology of T1D by investigating their age-related heterogeneity in the development of IA: height and weight growth, timing of infant diet exposures, and vitamin D. These three exposures were evaluated using three methods to test and model non-proportional hazards (i.e., age-related heterogeneity): the supremum test, weighted Schoenfeld residuals, and restricted cubic splines.
Our first interest was to assess age-related heterogeneity of height and weight growth based on previous research in DAISY by Lamb et al. (2009), which found an increased risk of IA associated with increased height growth velocity, but only in children < 11.5 years of age (31). We hypothesized that periods of increased growth, such as infancy or adolescence, may be associated with an even greater risk of IA. Specifically, we were interested in the adolescent period that had not been previously investigated in DAISY. By including children after 11.5 years of age in the analysis, we saw similar associations to those previously found by Lamb et al. (2009) (31). When we modeled the restricted cubic splines for height and weight growth velocity neither exhibited age-related heterogeneity. It may be important to explore the effect height and weight growth velocity by sex, as males and females experience increased height and weight growth at different periods, especially in adolescence. It is possible females would experience an increased risk of IA associated with height growth velocity earlier since they often experience a growth spurt earlier in puberty than males, whereas height growth velocity may be associated with an increased risk of IA later in males. In this cohort, stratifying by sex would most likely make it difficult to assess age-related heterogeneity using the restricted cubic splines due to small samples in each group. Based on our current findings, height growth velocity appears to be a risk factor for IA at all ages across childhood and is not associated with increased risk only in periods of proposed increased growth, like infancy and adolescence. Weight growth velocity does not appear to be a risk factor for IA in this cohort even when potential age-related heterogeneity is taken into account.

Our second interest was to assess age-related heterogeneity related to the timing of infant diet exposures. DAISY previously identified both early and late first exposure to cereals to be associated with risk of IA (41). However, when we revisited this analysis in DAISY with an older cohort, we no longer observed significant associations between early and late first exposure to cereal, or any other solid food exposure, and risk of IA. An unpublished analysis in DAISY using B-spline functions also suggested that the increased risk associated with timing of cereal
introduction occurs in earlier, but not later onset IA (44). In addition, the Finnish DIPP study found early introduction of wheat, rye, oats, and/or barley cereals ($P$ value = 0.013) and egg ($P$ value = 0.031) associated with an increased risk of IA, but only during the first 3 years of life (43), suggesting age-related heterogeneity related to infant diet exposures.

This previous work led us to hypothesize that early (< 4 months) and late (≥ 6 months) first exposures to selected foods in the infant diet are associated with an increased risk of IA in early onset IA, but not late onset IA. However, using the approaches described herein, we identified no significant age-related heterogeneity of early (< 4 months) or late (≥ 6 months) first exposure to any solid food in the infant diet.

The previous DAISY analysis that identified a potential age-related effect used a stricter IA case definition that required one to be positive for at least one autoantibody on two consecutive visits and remain autoantibody positive, or develop diabetes by their last follow-up visit. Our analyses used a case definition for IA that only required one to be positive for at least one autoantibody on two consecutive visits. This less strict IA case definition could have decreased our power to see an effect due to possible misclassification of the outcome, as within this less strict IA case definition there are subjects who develop only one autoantibody who may lose this autoantibody and are less likely to develop T1D (1,172–174). However, it was still difficult to model the restricted cubic splines with a smaller number of strict IA cases in the birth cohort, as Kroehl et al. (unpublished), through simulations, found on average 100 events were necessary to fit the restricted cubic splines. Revisiting the use of the B-splines is a possibility; however, B-splines rely on a visual assessment of non-proportional hazards, where the restricted cubic splines provide a statistical test to assess whether the effect exhibits non-proportional hazards.

While we were no longer able to see an association between timing of infant diet exposures and risk of IA, we had the opportunity to examine these infant exposures prospectively for development of T1D, which had not previously been investigated. Both early and late first
exposure to any solid food predicted development of T1D. Specifically, early exposure to fruit and late exposure to rice/oat predicted T1D, while breastfeeding at the time of introduction to wheat/barley conferred protection. These findings suggest the safest age to introduce solid foods in children at increased genetic risk for T1D is between 4 and 5 months of age and breastfeeding while introducing new foods may reduce T1D risk. These results also substantiate the importance of the timing and type of infant exposure previously identified with risk of IA by establishing an association with the clinical outcome of T1D.

This work brought to light the potential confusion that could occur based on the current complementary food and breastfeeding recommendations by the American Academy of Pediatrics. Regarding timing of infant diet exposures, we encourage the American Academy of Pediatrics to issue clearer and more specific recommendations regarding introduction of complementary foods, similar to the position paper by the European Society for Pediatric Gastroenterology, Hepatology and Nutrition (177). Based on the findings of this thesis, as well previous findings, it appears that early introduction of solid foods is an important risk factor in the etiology of T1D. DAISY findings also suggest an increased risk of IA and T1D with late introduction of cereals, particularly rice and oat. The biological mechanisms behind these associations have only been speculated and future research should be dedicated to identifying these mechanisms.

Our third aim evaluated whether vitamin D, including reported dietary intake of vitamin D, plasma 25(OH)D levels, and related SNPs, is more protective in early onset IA compared to late onset IA. This hypothesis was based off previous research that found diabetic children were less likely to have received vitamin D supplements during infancy than non-diabetic children, suggesting a protective role for vitamin early in life (45,46). Interestingly, previous work in DAISY found no association between vitamin D intake or 25(OH)D levels and risk of IA and progression to T1D (47). For this thesis, we hypothesized that our inability to detect a significant association could be due to an age-related effect of vitamin D that was averaged over time,
resulting in a null association. Therefore, we wanted to further investigate these vitamin D exposures for age-related heterogeneity. The restricted cubic spline for vitamin D intake from food and supplements displayed a straight line parallel to zero, indicating no age-related heterogeneity and a lack of effect across all ages (age 2 HR: 1.08, 95% CI: 0.76, 1.53 and age 11 HR: 1.05, 95% CI: 0.82, 1.35). It is possible that vitamin D intake should be evaluated stratified by HLA-DR genotype, as Lamb et al. (2014) recently found greater cow’s milk protein intake was associated with increased IA risk in children with low/moderate risk HLA-DR genotypes, but not in children with high risk HLA-DR genotypes (178). Cow’s milk is a primary source of dietary vitamin D in the United States. Therefore, it is possible future analyses with a greater number of IA cases may identify age-related heterogeneity of vitamin D intake stratifying by HLA-DR genotype.

We recently investigated age-related heterogeneity of plasma 25(OH)D levels. This analysis proved to be slightly more complex because a case cohort study design was utilized. However, our initial investigation suggests plasma 25(OH)D levels may exhibit an age-related effect with increased IA risk associated with increased plasma 25(OH)D early on, which then becomes protective later after adjusting for HLA-DR3/4, DQB1*0302 genotype, first-degree relative with T1D, and non-Hispanic white ethnicity. These results suggest the opposite of the protective association between vitamin D and IA early on that we hypothesized, which is intriguing. This age-related effect may be why plasma 25(OH)D levels were not previously found to be associated with IA in DAISY because the effect was averaged over time.

Regarding vitamin D SNPS, we hypothesized that vitamin D SNPs would play a role at all ages of development. This hypothesis was primarily based off of a lack of previous evidence indicating an age-related effect of vitamin D SNPs. It is also difficult to hypothesize when these SNPs play a role in the disease process without completely understanding the role of these SNPs in the vitamin D pathway. In our investigation, none of the vitamin D SNPs exhibited age-related heterogeneity. It is likely that the effect of many of these SNPs depends on plasma vitamin D
levels, including 1,25-dihydroxyvitamin D, 25(OH)D, and vitamin D binding protein and
interactions between vitamin D metabolism SNPs and plasma levels should be investigated in
larger cohorts, like the TEDDY study.

A number of novel associations between vitamin D SNPs and development of IA, as well
as progression from IA to T1D were identified. *DHCR7/NADSYN1* rs12785878 and *CYP27B1*
rs4646536, previously identified through GWAS for association with T1D, were, for the first
time, found to be associated with the pre-clinical phase of T1D, islet autoimmunity. Additionally,
an interaction between *VDR* rs1544410 and *PTPN2* rs1893217 was associated with progression
from IA to T1D. These novel SNP associations may offer insight into the role of vitamin D in the
etiology of T1D.

In our investigation of age-related heterogeneity, a non-HLA SNP previously identified
through GWAS, rs10517086, was found to exhibit an age-related effect with development of IA
with an increased risk of IA before age 2 years that diminished with age. In an investigation of
well-established determinants of IA, non-Hispanic white ethnicity and maternal age also
exhibited age-related heterogeneity. While these determinants have not previously been shown to
demonstrate etiologic heterogeneity, these findings may prove interesting for hypothesis
generation.

This thesis provides an impetus for investigating exposures for age-related heterogeneity
in the future. The majority of the T1D literature does not address determinants in disease
development that may exhibit an age-related effect, which is what makes this research innovative.
One of the strengths of this research is its ability to shed light on the importance of evaluating the
assumption of proportional hazards before including covariates in a Cox proportional hazards
model. It also may encourage researchers to develop hypotheses that specifically explore age-
related heterogeneity of exposures of interest, as well as revisit exposures that were previously
associated or not associated with the outcome, as their effect may have changed with increasing
follow-up time. The novel method for evaluating age-related heterogeneity of time-covariates
using restricted cubic splines developed for this research will allow investigators to explore an age-related effect of both fixed and time-varying covariates with the ability to statistically test a non-proportional effect.

The small number of exposures we identified to exhibit age-related heterogeneity could be due to a number of reasons. It is possible that not many exposures exhibit an age-related effect and are associated with the development of the disease at all ages. Exposures that can only occur early in life, like timing of infant diet exposures, may have a programming effect on the immune system that allows that exposure to be associated with risk across childhood. The DAISY population is also a unique population that is at increased genetic risk for developing T1D because of a high-risk HLA genotype or first-degree relative with T1D. This may make it difficult to identify environmental exposures that play a role in the disease process, particularly at different ages, because the influence of the genetic component is so great.

It is also possible that the relatively small number of IA cases limited our ability to see a significant age-related effect, particularly in certain cases where the restricted cubic splines visually demonstrate a non-constant hazard, but the non-constant effect is not statistically significant. When fitting the restricted cubic spline models, small samples limit the amount of information available to accurately predict the effect of a number of variables, such as the coefficients related to the splines. DAISY data also has a greater number of events at the beginning of follow-up time and fewer events toward the end of follow-up time making it difficult for time-dependent predictors related to the hazard ratio to accurately estimate the later ages. The large number of IA cases developing at the beginning of follow-up may also have a certain set of risk factors that once removed from the risk set may make it difficult to see an effect with the outcome later on.

Each of the exposures evaluated in this study is currently being evaluated in the TEDDY study in which 8,668 children at increased genetic risk for developing T1D are enrolled. Being the largest prospective birth cohort of T1D, the TEDDY study will be able to utilize these
methods for assessing age-related heterogeneity from the beginning without having to revisit exposures with null associations that could have been due to an averaged effect over time. TEDDY may be able to elucidate more of the mechanisms behind these well-studied, but inconsistently associated determinants of T1D with larger sample sizes, multi-country subjects, and advanced methodologies using metabolomics and epigenetics.

Assessment of age-related heterogeneity using survival analysis is an area of research that has significant potential for future development and application. One of the ultimate goals in DAISY is to be able to incorporate each of the determinants that have been implicated in T1D in one model to determine where in the natural history of the disease or at what ages each exposure plays a role. This could also be used to identify groups of exposures that together contribute to peaks in incidence of IA and T1D. Like many complex chronic diseases with multiple exposures, studying the etiology of T1D will require advanced methods like an assessment of age-related heterogeneity in order to understand when and where in the disease process these exposures are important. Assessing age-related heterogeneity is valuable in evaluating determinants of numerous disease processes in epidemiologic research, as well as other disciplines where certain exposures may be influential at certain ages. This information will help guide treatment and prevention efforts by creating more accurate risk prediction models, as well as inform the design of targeted interventions.
REFERENCES


