GENETIC INSIGHTS INTO LATENT AUTOIMMUNE DIABETES IN ADULTS

Rajashree Mishra

A DISSERTATION

in

Genomics & Computational Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

_____

Struan F.A Grant, Ph.D., Professor of Pediatrics

Graduate Group Chairperson

_____

Ben Voight, Ph.D., Associate Professor of Systems Pharmacology and Translational Therapeutics and Associate Professor of Genetics

Dissertation Committee

Marcella Devoto, Ph.D., Professor of Pediatrics

Ben Voight, Ph.D., Associate Professor of Systems Pharmacology and Translational Therapeutics and Associate Professor of Genetics

Andrew Wells, Ph.D., Associate Professor of Pathology and Laboratory Medicine

Stephen S. Rich, Ph.D., Professor of Public Health Sciences, University of Virginia

GENETIC INSIGHTS INTO LATENT AUTOIMMUNE DIABETES IN ADULTS

This work is dedicated to my loving parents, brother, and sister.

ACKNOWLEDGMENT

I extend my whole-hearted gratitude to everyone who has supported me in my academic endeavors.

This work would not have been made possible without the mentorship and guidance of my advisor Struan Grant. Struan, thank you for your constant support and push towards excellence. I will always be grateful for the opportunities you gave me and your dedication to my development as a scientist. You have taught me independence, a strong work-ethic, and most importantly how to communicate my work concisely and eloquently. It has truly been an honor and privilege to be your first PhD graduate student.

To my thesis committee, Marcella Devoto, Ben Voight, Steve Rich, and Andrew Wells, thank you for your helpful advice and guidance throughout my candidacy. I truly appreciate your time and dedication to my development. It's been an honor to learn from each of your expertise.

To the 'LADA team', 'Bone group', and co-authors, this work could not have been done without you. Special thanks to Diana Cousminer, David Leslie, Kenyaita Hodge, and Alessandra Chesi, and special thanks to Jonathan Mitchell for providing a quiet and remote space to write.

To the students, postdocs, faculty, and staff of the Grant Lab and Genomics and Computational Biology program, thank you for inspiring me every day and providing me with an amazing work and learning environment here at Penn/CHOP. Special thanks to Li-San, Maureen, and Hannah for their helpful coordination and support throughout my time at Penn.

To my close friends, thank you for being my support system throughout graduate school. Your encouragement and companionship continue to motivate me and keep me in high-spirits. Most importantly, thank you for believing in me even when I didn't believe in myself.

Finally, words cannot describe my appreciation and gratitude towards my family. To my parents, Nrusingha and Banita, my brother, Satish, and my sister, Biju, thank you for your endless love and support. Your inspiration gives me strength to work hard and strive to succeed every day.

This work was supported by National Institutes of Health R01 DK085212 grant.

ABSTRACT

GENETIC INSIGHTS INTO LATENT AUTOIMMUNE DIABETES IN ADULTS

Rajashree Mishra

Struan F.A. Grant

‘Latent autoimmune diabetes in adults’ (LADA) is a controversial subtype of diabetes characterized by initial insulin independency and the presence of diabetes associated autoantibodies. As a result, LADA is often misclassified and can represent 5-10% of apparent type 2 diabetes (T2D) cases and is potentially more prevalent than childhood-onset type 1 diabetes (T1D). Despite LADA sharing features with the two better characterized classic diabetes subtypes, the genetic etiology of LADA remains largely unknown. Once there is a more accurate definition of LADA, there will be an improvement in diabetes classification and consequently better treatment and therapeutic interventions. The objective of this thesis is to understand the genetic basis of LADA in order to bring clarity to the current definition of LADA by being the first to leverage genome-wide genotype data from a LADA cohort and the subsequent application of statistical genetics approaches. These investigations can be divided into three parts: 1) the role of T1D and T2D loci in LADA 2) the

first genome-wide association study (GWAS) of LADA, and 3) searching for genetic discrepancies between LADA and childhood-onset T1D in the human leukocyte antigen (HLA) region. Four out of the five strongest associations from the candidate locus study were known T1D loci (*HLA, PTPN22, INS* and *SH2B3*) and reached genome-wide significance in the GWAS meta-analysis. However, a novel independent signal at a known T1D locus was also observed to be genome-wide significant, near the *PFKFB3* gene, which had not been implicated in previous T1D or T2D GWAS. Additionally, major T1D-susceptibility HLA haplotypes were observed to be less frequent in LADA. Furthermore, contrary to observations in childhood-onset T1D studies, *HLA-B* and *HLA-A*, were not significantly associated with LADA, independent of *HLA-DQB1* and *HLA-DRB1* haplotypes. Overall, the genetics of LADA point to a strong T1D component, but a positive genetic correlation between LADA and T2D is also evident, strongly suggesting LADA has both a T1D and T2D component. However, it remains unresolved whether LADA is at the genetic intersection of T1D and T2D or simply a mixture of relatively poorly phenotyped individuals who have either T1D or T2D.

TABLE OF CONTENTS

ACKNOWLEDGMENT ..... *iii*

ABSTRACT ..... *vi*

TABLE OF CONTENTS ..... *viii*

LIST OF TABLES ..... *xi*

LIST OF ILLUSTRATIONS ..... *xii*

CHAPTER 1. INTRODUCTION *1*

1.1 Background **1**

1.2 Motivation and Objective **2**

1.3 Organization of Dissertation **2**

1.4 Autoantibodies and clinical heterogeneity **3**

1.5 Clinical implications of misdiagnosing LADA cases **5**

1.6 Previous genetic studies of LADA **7**

CHAPTER 2. A GLOBAL PERSPECTIVE OF LADA *11*

2.1 Global prevalence of diabetes **11**

2.2 Global prevalence of LADA **13**

2.2.1 Prevalence of LADA in populations of European Ancestry 13

2.2.2 Prevalence of LADA in populations of African Ancestry 15

2.2.3 Prevalence of LADA in African-American populations 16

Prevalence of LADA in Asian and Middle Eastern populations 18

2.3 Importance of autoantibody screening **22**

2.4 Concluding remarks **23**

CHAPTER 3. RELATIVE CONTRIBUTION OF T1D AND T2D LOCI TO THE GENETIC ETIOLOGY OF LADA *25*

3.1 Introduction **25**

3.2 Methods **27**

3.3 Results **33**

3.4 Discussion **41**

3.5 Conclusion **46**

CHAPTER 4. HERITABILITY OF LADA *48*

4.1 Introduction **48**

4.2 Materials & Methods **51**

4.3 Results **55**

4.4 Discussion **58**

4.5 Conclusion **60**

CHAPTER 5. THE FIRST GENOME-WIDE ASSOCIATION STUDY OF LADA 61

5.1 Introduction **61**

5.2 Methods **63**

5.3 Results **67**

5.4 Discussion **74**

5.5 Conclusion **78**

CHAPTER 6. GENETIC DISCRIMINATION BETWEEN LADA AND CHILDHOOD-ONSET TYPE 1 DIABETES WITHIN THE MHC..... 79

6.1 Introduction **79**

6.2 Materials & Methods **81**

6.3 Results **87**

6.4 Discussion **90**

6.5 Conclusion **94**

CHAPTER 7. FUTURE DIRECTIONS 103

7.1 Brief Summary **103**

7.2 Limitations **104**

7.3 Is LADA simply a mixture of T1D and T2D? **105**

7.4 Follow-up on genetic markers involved in autoimmune diabetes **109**

7.5 Towards precision medicine in diabetes **110**

7.6 Conclusion **111**

APPENDIX..... 113

Chapter 2: A Global Perspective of LADA Supplemental Table **113**

Chapter 3: RELATIVE CONTRIBUTION OF T1D AND T2D LOCI TO THE GENETIC ETIOLOGY OF LADA **114**

Supplemental Results 114

Tables 116

Supplemental Figures 128

Chapter 5: THE FIRST GENOME-WIDE ASSOCIATION STUDY IN LADA – SUPPLEMENTARY MATERIAL **132**

Supplemental Tables 132

Supplemental Figures 157
Supplemental Note: Cohort Information 160

Chapter 6. Genetic Discrimination between LADA and childhood-onset type 1 diabetes within the MHC 172
Supplemental Tables 172

BIBLIOGRAPHY 179

LIST OF TABLES

CHAPTER 2. A GLOBAL PERSPECTIVE OF LADA

Table 2-1 Prevalence of autoantibody-positive individuals among adults with diabetes across different population groups. 21

CHAPTER 3. RELATIVE CONTRIBUTION OF T1D AND T2D LOCI TO THE GENETIC ETIOLOGY OF LADA

Table 3-1 Association of established T1D loci with LADA 34
Table 3-2 Association of established T1D loci in LADA subjects positive for both GADA and IA2A 35
Table 3-3 Association of established T2D loci with LADA.. 36
Table 3-4 Association between established T2D loci in LADA cases positive for GAD and IA2 autoantibodies 37

CHAPTER 4. HERITABILITY OF LADA

Table 4-1 Heritability estimates for T1D, LADA, and T2D derived from the LDSC method with and without the MHC.. 56
Table 4-2 Heritability estimates for T1D, LADA, and T2D across two tools: GCTA and LDAK.. 57
Table 4-3 Power calculation for GCTA 57

CHAPTER 5. THE FIRST GENOME-WIDE ASSOCIATION STUDY OF LADA

Table 5-1 Genome-wide significant signals associated with LADA.. 68
Table 6-1 Independent association signals from the conditional analysis in T1D cases vs WTCCC controls.. 95

CHAPTER 6. GENETIC DISCRIMINATION BETWEEN LADA AND CHILDHOOD-ONSET TYPE 1 DIABETES WITHIN THE MHC

Table 6-2 Independent association signals from the conditional analysis in LADA cases vs controls 96
Table 6-3 Independent association signals from the conditional analysis in T1D vs controls in the Swedish replication cohort. 97
Table 6-4 Independent association signals from the conditional analysis in LADA vs controls in the Swedish replication cohort. 98
Table 6-5 Independent association signals from the conditional analysis in the downsampled cohort of T1D cases vs controls. 99
Table 6-6 Independent association signals from the conditional analysis in a sample of randomly selected T1D and T2D cases vs controls. 100

LIST OF ILLUSTRATIONS

CHAPTER 1. INTRODUCTION

Figure 1-1 The clinical heterogeneity of diabetes. 5

CHAPTER 3. RELATIVE CONTRIBUTION OF T1D AND T2D LOCI TO THE GENETIC ETIOLOGY OF LADA

Figure 3-1 T1D and T2D genetic risk scores tested in LADA cases and controls. 39
Figure 3-2 GRS distributions 41

CHAPTER 4. HERITABILITY OF LADA

Figure 4-1 Comparison of heritability estimates 56

CHAPTER 5. THE FIRST GENOME-WIDE ASSOCIATION STUDY OF LADA

Figure 5-1 LocusZoom plots for the *PFKFB3* locus. 70

CHAPTER 6. GENETIC DISCRIMINATION BETWEEN LADA AND CHILDHOOD-ONSET TYPE 1 DIABETES WITHIN THE MHC

Figure 6-1 Flow Chart of (A) datasets and (B) overall analysis 86
Figure 6-2 Conditional analysis in 1,985 T1D cases and 2,219 WTCCC controls. 101
Figure 6-3 Conditional analysis in 1,428 LADA cases and 2,979 WTCCC controls.. 102

CHAPTER 7. FUTURE DIRECTIONS

Figure 7-1 GRS distribution of varying proportions of T1D and T2D cases 108
Figure 7-2 T2D GRS plotted against T1D GRS 109

CHAPTER 1. INTRODUCTION

1.1 Background

Genetic studies of childhood-onset type 1 diabetes (T1D) and type 2 diabetes (T2D) have yielded dozens of loci associated with diabetes risk. However, the genetic etiology of latent autoimmune diabetes in adults (LADA) remains substantially less well characterized compared to the two classic forms of diabetes, thus highlighting a major undeveloped area in the field, including the debate on the clinical definition of LADA and its phenotypic heterogeneity. Despite being defined as a slowly progressive form of T1D by the World Health Organization (American Diabetes Association/ADA, 2018), LADA exhibits clinical features of both T1D and T2D, earning the nickname “type 1.5 diabetes.” Additionally, existing genetic studies have portrayed LADA as an intersection of the two diseases (Tuomi et al., 1999).

Patients diagnosed with LADA are typically defined by insulin independence for at least the first six months after diagnosis plus the presence of circulating diabetes-associated pancreatic islet autoantibodies. Diabetes associated autoantibodies highly predictive of LADA include glutamic acid decarboxylase (GAD) autoantibodies, islet cell autoantibodies (ICA), IA-2 autoantibodies and insulin autoantibodies (IAA). In certain populations, notably the Chinese in whom the frequency of GAD and IA-2 autoantibody positivity is low, zinc transporter (ZnT8) autoantibodies serve as an additional marker that has been shown to improve diagnostic sensitivity (Huang et al., 2013).

1.2 Motivation and Objective

Given the initial independence of insulin in patients with LADA, there is a high misdiagnosis rate among those with T2D (5-10%), and LADA continues to be overlooked in clinical practice with no clear diagnostic guidelines. In fact, the majority of patients with adult-onset autoimmune diabetes do not require insulin treatment for at least 6 months after diagnosis (Buzzetti, Zampetti, & Maddaloni, 2017). Despite recent studies that have shown that T1D is as prevalent in adults as in children (Thomas et al., 2017), these studies do not directly address LADA. The objective of this dissertation is to investigate the genetics of LADA for the first time in order to aid in the characterization of this puzzling clinically-defined subtype of diabetes.

1.3 Organization of Dissertation

The remainder of this chapter will cover the background of LADA, including the clinical definition and implications of misdiagnosing individuals, and previous genetic studies relevant to LADA. Chapter 2 covers the global prevalence and incidence of LADA and highlights the importance of autoantibody screening. The remaining chapters highlight the three specific aims of the dissertation. The first aim is to investigate the role of T1D and T2D loci in LADA in order to understand how LADA compares to these classic forms of diabetes (Chapter 3). The second aim is to take a genome-wide approach for understanding the genetics of LADA by investigating the heritability (Chapter 4) and performing the first discovery genome-wide association (GWAS) of LADA to attempt to identify putatively novel signals that are unique to LADA and could aid in the distinguishing of LADA from T1D and T2D (Chapter 5). The third aim is to investigate the genetic

discrepancies between childhood-onset T1D and LADA within the major histocompatibility complex (Chapter 6). Differences within this region between the two groups offers great promise to help us understand the progression of autoimmune diabetes. Finally, Chapter 7 will summarize the findings from aims 1-3, provide follow-up analyses and highlight future directions for studying the genetics of LADA.

1.4 Autoantibodies and clinical heterogeneity

The heterogeneity of LADA is evident not only in cases in which autoantibodies are present but also in the varying levels of these autoantibodies. The distribution of GAD titers across LADA patients has been shown to be bimodal in some populations, resulting in two subgroups with distinct characteristics (Buzzetti et al., 2017; Maioli et al., 2010; Zhou et al., 2013). However, more detailed analyses of bimodality, specifically on untransformed assay signals, should be performed. Typically, those with high GAD titer are younger (Turner et al., 1997) with lower BMI, lower prevalence of metabolic syndrome (Zinman et al., 2004), higher levels of HbA1c, lower C-peptide levels and a higher frequency of other diabetes-related autoantibodies (**Figure 1-1**). As a result, this particular subgroup resembles T1D more closely than individuals with lower GAD titer; alternatively, those with low GAD titer have metabolic presentation and loss of β-cell function that resembles a phenotype more closely aligned with T2D (Liu et al., 2015). This observation resonates with patterns seen in a cohort of individuals recently diagnosed with childhood-onset T1D, where those subjects who were positive for a single autoantibody (among those >12 years old) showed an association with

the T2D-associated *TCF7L2* locus, and had higher C-peptide and lower glucose levels in an oral glucose tolerance test (Redondo et al., 2017; Redondo et al., 2018). Notably, IA-2 autoantibodies against a specific epitope are associated with increased body mass index (BMI) in patients with T2D, and patients positive for IA-2 autoantibodies to this epitope have a delayed progression to insulin requirement than patients who were solely positive for the GAD autoantibody (Buzzetti et al., 2017). These observations have led to two hypothesized mechanisms for the onset of LADA, that diabetes can result from either chronic inflammatory responses in individuals with higher BMI or through classic T1D autoimmune response in leaner individuals (Buzzetti et al., 2017). However, an alternative explanation could be that autoantibodies are not necessarily pathogenic unless they recognize certain epitopes of an antigen.

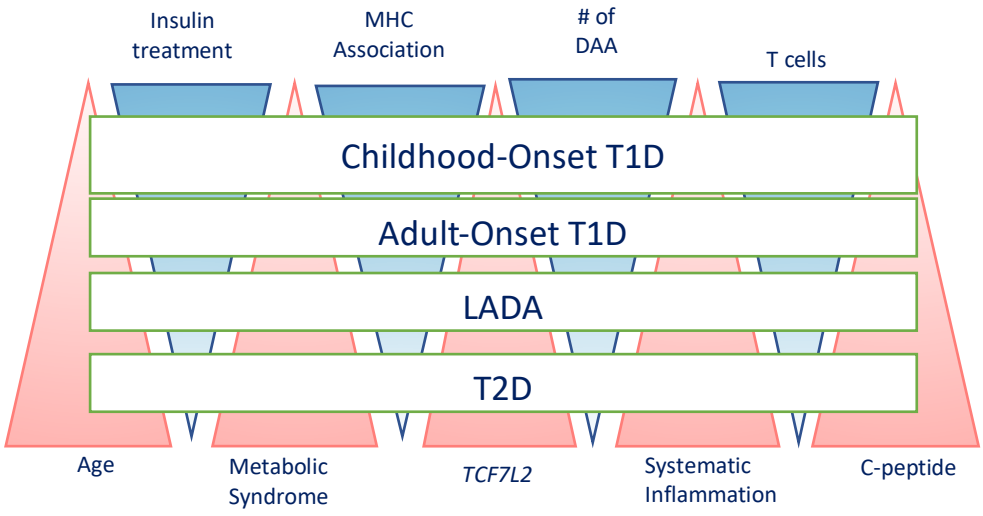


Figure 1-1 The clinical heterogeneity of diabetes, with LADA situated between T2D and late-onset T1D. This image is derived from (Leslie et al., 2008) and current knowledge of LADA (Buzzetti et al., 2017; Hawa et al., 2013). Abbreviations: DAA, diabetes associated, type 1 diabetes; T2D, type 2 diabetes.

1.5 Clinical implications of misdiagnosing LADA cases

The role of autoantibodies in the pathogenesis of LADA remains crucial for identifying patients who should be monitored more closely, especially if they require insulin therapy at an early stage. Insulin producing pancreatic β-cells in those with LADA tend to be compromised, presumably by the adverse immune response. Consequentially, the risk of progression to insulin therapy in patients with LADA is higher and earlier than in cases with T2D. If inadequately treated, ketoacidosis can result, representing a substantial complication, as it is the most common cause of death among pediatric diabetic cases (Tasker & Acerini, 2014). While the destruction of the β-cell leads to hyperglycemia in T1D, this complication is also common in T2D due to insulin insensitivity and pancreatic β-cell dysfunction or decreased β-cell mass (Parker et al., 2017).

Thus, LADA and T2D are presumed to have two distinct mechanisms for disease progression, and therefore should be treated appropriately immediately following diagnosis. Studies have indicated that patients with LADA should not be treated with sulfonylureas, which is a common treatment for T2D, as these drugs further exhaust β-cell reserves, although one caveat is that these studies have had limited power (Kobayashi et al., 2002). Agents which augment the incretin pathway will be uninformative if C-peptide levels are very low (Dennis et al., 2018). Therefore, accurate diagnoses are imperative, as the more rapid

progression to insulin dependence in misdiagnosed LADA cases could be life-threatening.

Several factors are important in making an accurate diagnosis. First, inadequate treatment is reflected in higher HbA1c levels in LADA cases than in T2D cases (Maddaloni et al., 2015; Zaharieva et al., 2017). A recent study found that the rate of deterioration of HbA1c in LADA was double that of patients with antibody-negative T2D (Donnelly et al., 2018). Furthermore, individuals with LADA have higher HbA1c independent of insulin usage (Andersen et al., 2013). Second, co-morbidities are different, where there is a higher risk of metabolic syndrome and cardiovascular disease among T2D cases, although large prospective studies are still ongoing in this context (Wod et al., 2018). Recent studies have suggested that macrovascular disease is as common in LADA as in T2D. However, in LADA there is a higher risk of thyroid and parietal cell autoimmunity, leading to a need for either thyroid replacement or vitamin B12 supplementation. Finally, LADA patients who are obese and scheduled for bariatric surgery should consider avoiding this operation as they are actually at risk of progression to insulin therapy independent of their insulin resistance status (De Luca et al., 2016).

In particular, the underuse of autoantibody screening has led to the underdiagnosis of LADA, hampering the ability to estimate accurate prevalence rates and to assess the genetic architecture of the disease. Strikingly, the prevalence estimates based on current reports for T1D, T2D and LADA differ widely across populations, and this is influenced by genetic, environmental and socioeconomic differences, with notable differences between clinic-based or population-based studies. There is a need for consistent and thorough, well-

powered genetic studies in populations of diverse ethnicities to aid in the full characterization of the pathogenesis of autoimmune diabetes; however, these studies cannot be performed if autoantibody screening is not performed immediately after the diagnosis of diabetes.

1.6 Previous genetic studies of LADA

The lack of awareness of LADA and the relatively unclear diagnostic criteria has hindered progress in understanding the genetics of this common disease. However, there have been many genetic studies regarding T1D and T2D, both of which have a clear heritable component. The risk for developing T2D is three times higher in first-degree relatives of patients with T2D than individuals without a family history, while the risk for developing T1D is 15 times more likely in first-degree relatives (Florez, Hirschhorn, & Altshuler, 2003). A stronger genetic similarity between LADA and T1D than between LADA and T2D was supported by a recent study based on 378 LADA cases which showed that having family history of T1D was associated with a 6-fold increased risk, whereas family history of T2D was associated with a 2-fold increased risk of LADA (Hjort et al., 2017). This study suggests that the T1D genetic contribution to LADA etiology is greater than that of T2D.

T1D and T2D appear to be due to distinct biological mechanisms, as shown by the paucity of overlapping significant loci identified by GWAS (Aylward et al., 2018; Bradfield et al., 2011). However, some overlapping signals from major histocompatibility complex (MHC) loci are beginning to emerge from the most recent GWAS of T2D (Mahajan et al., 2018; Ng et al., 2014; Scott et al., 2017), potentially due to the presence of unaccounted individuals with autoimmune diabetes. The MHC accounts for approximately 50% of the genetic risk for T1D

and harbors human leukocyte antigen (*HLA*) genes which encode highly polymorphic antigen-presenting proteins (Todd, 2010). Risk MHC class II haplotypes in combination with an MHC class I chain-related A5 gene polymorphism, located in the class III region, has been shown to be a strong genetic marker for T1D risk (Gambelunghe et al., 2000). The MHC class I genes *HLA-A* and *HLA-B* have also been pinpointed through conditional analyses as increasing T1D risk(Nejentsev et al., 2007). However, a signal from the *HLA-B* locus has emerged in a GWAS meta-analysis of T2D in African Americans along with a signal from the *INS-IGF2* locus, which has been strongly implicated in LADA (Cervin et al., 2008), further suggesting potential misclassification among those diagnosed with T2D.

Association of T1D and T2D genetic signals in LADA

Multiple loci are robustly associated with T1D and T2D (Bradfield et al., 2011; Ng et al., 2014; Onengut-Gumuscu et al., 2015; Scott et al., 2017). Established GWAS-implicated loci for these diseases have been investigated in cohorts of European cases diagnosed with LADA (Cervin et al., 2008; Howson et al., 2011). Several studies have shown that T2D genetic risk, particularly at *TCF7L2*, is also associated with LADA (Cervin et al., 2008; Howson et al., 2011). There is an emerging picture for a role for the *TCF7L2* locus in the pathogenesis of autoimmune diabetes, where it has been implicated in T1D heterogeneity (Redondo et al., 2014). In particular, T1D carriers of the known T2D *TCF7L2* risk allele tend to have less severe immunological and metabolic phenotypes with an age–dependent effect (>12 years at diagnosis being the cut-off above which the association was noted) (Redondo et al., 2017). Therefore, it is unclear to what extent T2D loci contribute to the onset of LADA, but these recent studies suggest

a T2D-like pathophysiological mechanism for a subset of affected individuals. A study of diabetes from Hungary (Lukacs et al., 2012) suggested that BMI influences the genetic effect on LADA risk, such that the lower the BMI, the higher the *TCF7L2* genetic effect, as also predicted in a smaller American study (Redondo et al., 2014).

Specific *HLA* haplotypes have been shown to be significantly associated with T1D. In particular, the *HLA-DR3* and *HLA-DR4* haplotypes play a consistent role across multiple ethnicities (Desai et al., 2007; Luo et al., 2016; Manan et al., 2010; Mbanya et al., 2010). Moreover, there is a greater increase in risk for patients with DR3/DR4 heterozygosity than those with only the DR3 haplotype. In a major Chinese study, protective T1D haplotypes were enriched in LADA compared to T1D, while, conversely, susceptibility T1D haplotypes were diminished in LADA when compared to T1D (Luo et al., 2016). Perhaps not surprisingly, another study in the same Chinese population comparing LADA and T2D found that these T1D susceptibility haplotypes had a significantly higher frequency in LADA cases (63.9%) than that in either T2D cases (47.1%) or controls (43.2%); indeed, the converse was also observed for the frequency of protective haplotypes (LADA: 22.8%; T2D: 33.3%; controls: 32.7%) (Zhou et al., 2013).

Taken together, the genetics studies of LADA thus far have highlighted the *HLA* region and *TCF7L2* contribution to LADA. This dissertation further expands these studies by investigating not only candidate loci, but also the genome-wide genetics of LADA in the largest cohort to date.

Majority of the content in this chapter is published in Mishra R, Hodge KM*, Cousminer DL., Leslie RD., Grant SFA. "A Global Perspective of Adult Autoimmune Diabetes." Trends in Endocrinology and Metabolism (2018). PMID: 30041834. *equal contribution. RM's contribution to the manuscript includes literature review, drafting, critical revision, and writing the final version of the manuscript.*

CHAPTER 2. A GLOBAL PERSPECTIVE OF LADA

This chapter highlights the prevalence rates reported across the globe and addresses the apparent challenges of performing genetic studies for LADA. The content of this chapter is published in Mishra R, Hodge KM, Cousminer DL., Leslie RD., Grant SFA. "A Global Perspective of Adult Autoimmune Diabetes." Trends in Endocrinology and Metabolism (2018). PMID: 30041834. RM's contribution to the manuscript includes literature review, drafting, critical revision, and writing the final version of the manuscript.

2.1 Global prevalence of diabetes

Diabetes is dramatically on the rise across the globe. Based on the studies available, the overall rate of diabetes is approaching epidemic levels, with the global prevalence projected to reach 10.4% by 2040, and with 5 million deaths reported as of 2015. T2D is by far the most prevalent form of diabetes (Alberti & Zimmet, 1998), accounting for approximately 90-95% of total diabetic cases world-wide, totaling approximately 400 million adults (Ogurtsova et al., 2017). Additionally, an increased rate of obesity and rise in socioeconomic status has influenced the rate of T2D (Martinell et al., 2017). In contrast, T1D makes up only a small percentage (4-10%) of the total burden of diabetes but has a stronger genetic etiology (Eisenbarth et al., 2008; Poulsen et al., 1999) and clear variation in prevalence around the world (American Diabetes Association, 2014; Maahs et al., 2010). In 2013, worldwide prevalence estimates indicated that there were almost 500,000 children under the age of 15 years old with T1D, with the largest

populations being in Europe and North America (Maahs et al., 2010). The recent increase in T1D incidence in certain areas could be explained by an environmental effect, such as the proposed hygiene hypothesis (Bach, 2002). However, there is a paucity of well-powered studies which have systematically addressed the prevalence of T1D in the under-developed regions of the world. Despite the relatively limited studies available, the epidemiological studies performed thus far have shown an overall increase in the number of new cases per year (incidence rate) especially in young children. However, trends reported on T1D are likely only available from countries with well-established public health and diabetes research programs. Therefore, there is a need for comparable research on T1D in developing countries. Additionally, many of the T1D studies represent trends in children, even though age-of-onset can extend to adulthood (Maahs et al., 2010; Mayer et al., 2017). Consequently, the prevalence rates reported for T1D should be carefully interpreted.

Using genetics to identify T1D cases amongst adults with diabetes

A recent study (Thomas et al., 2017) demonstrated that T1D is almost as common in adulthood as in childhood using genetic risk scores (GRS) in order to identify T1D cases in a cohort with adult-onset diabetes. The GRS in this particular study was based on data from a cohort of childhood-onset T1D, the majority of whom were diagnosed under the age of 16 years. However, substantial evidence points towards age-dependent differences in the genetic risk to develop autoimmune diabetes. Specifically, the genetic risk to develop autoimmune diabetes declines with increasing age-at-diagnosis without a clear cut-off, rather in a graded fashion (Frohnert et al., 2017; Pugliese et al., 2016;

Steck et al., 2017). Additionally, LADA was not addressed in a recent cohort study (Thomas et al., 2017), yet a proportion of diabetes cases genetically defined as T1D using a childhood-onset T1D derived GRS did not need insulin at diagnosis. Those defined as T1D represented only 22% of the whole cohort on insulin treatment and only 66% of that developing ketoacidosis. Therefore, it currently remains unclear how to discriminate T1D and LADA. At diagnosis, C-peptide levels or lack of insulin dependence is commonly used to discriminate LADA from T1D; however, there is scope for improvement on how to distinguish the two biologically, and that is the motivation for investigating the genetics of T1D and LADA.

2.2 Global prevalence of LADA

In contrast to T1D and T2D, there have been many fewer studies conducted on LADA, especially in ethnicities beyond European ancestral populations, with existing studies being somewhat underpowered. Additionally, the validity of autoantibody measurements and the application of diagnostic criteria may differ significantly across regions, which in turn are likely to influence prevalence estimates, leading to overly conservative reports and conclusions. Here, we report global trends in LADA and highlight how they compare to the two more classic forms of diabetes across populations of European, African, Asian and Middle Eastern ancestry.

2.2.1 Prevalence of LADA in populations of European Ancestry

Approximately 5-12% of cases in European populations with apparent T2D are in fact misdiagnosed LADA cases (Barinas-Mitchell et al., 2004; Bosi et al., 1999; Buzzetti et al., 2007; Hawa et al., 2013; Maioli et al., 2010; Radtke et al.,

2009; Tuomi et al., 1999; Turner et al., 1997; Wod et al., 2018; Zaharieva et al., 2017; Zinman et al., 2004) (**Table 2-1**). In most of these studies, LADA is defined by age-of-onset >30 years, with insulin independence for at least 6 months, and positivity for at least one autoantibody. However, trends in LADA diagnosis across European populations differ even between regions within a given country, especially between population-based and hospital clinic-based cohorts. For instance, there was a dramatically lower prevalence of LADA in a population-based study in Northern Italy (0.19%) (Bosi et al., 1999) when compared to a much larger study of subjects recruited throughout Italy from hospital clinics (6.6%) (Genovese et al., 2006). The reason for this discrepancy is likely due to the greater metabolic severity of LADA, such that hospital clinics are more likely to be enriched for these cases. However, these observations could also be explained by regional variations similar to those reported across Spain, where the south of the country has a higher prevalence rate of LADA compared to the north (10.9-14.7% vs 5.6-7.2%, respectively) (Soriguer-Escofet et al., 2002). These discrepancies could be attributed in part to sample size; however, genetic and environmental factors could influence LADA prevalence, an idea that is supported by the wide range in diabetes prevalence observed across the globe. An unexplored area is the proportion of cases with transient autoantibodies, in whom the presentation is autoantibody negative but who otherwise have most of the features of LADA. This possibility was highlighted in a recent large (n=14,775 cases) population-based study in Scandinavia of adult onset diabetes, of which 6.4% had LADA, but 17.5% had the same features including low C-peptide levels but were autoantibody negative (Ahlqvist et al., 2018).

Since the vast majority of adult-onset diabetes cases have T2D and the estimated prevalence of T2D in populations of European ancestry is approximately 10% (Ng et al., 2014), the prevalence of LADA in European populations is considerably higher than the previously reported European prevalence of T1D (~0.3%) (Dabelea et al., 2014; Maahs et al., 2010). T1D, however, is by far the most common type of diabetes in children and adolescents and is more frequent in populations of European ancestry (Maahs et al., 2010) (**Appendix: Supp. Table 2-1**).

2.2.2 Prevalence of LADA in populations of African Ancestry

The majority of well-powered diabetes studies are predominantly of European ancestry, despite the African region expecting to have the world's largest proportional increase of adults with diabetes (Ogurtsova et al., 2017). The estimated prevalence of diabetes in adults across Africa ranges widely, but is overall lower than in other parts of the world (Ogurtsova et al., 2017)(**Appendix: Supp. Table 2-1**). The majority of people in Africa with T2D live in cities, although the overall population is predominantly rural, and while those in rural areas may have less access to health centers, environmental factors like urbanization may also play a major role in susceptibility. In fact, the reported prevalence of T2D can vary largely between rural and urban areas (Mbanya et al., 2010; Motala et al., 2003). There is also a lack of corresponding studies on T1D in Africa, presumably leading to inaccurate incidence/prevalence estimates. Additionally, poor healthcare and high mortality rates leave many cases unaccounted for in such estimates (Mbanya et al., 2010).

Taking T1D/T2D epidemiology in African populations into account, there are only a handful of LADA studies performed in that part of the world. It has been reported that 10-14% of West Africans with apparent T2D are misdiagnosed LADA cases, higher than their European counterparts (8-10%) (Adeleye et al., 2012; Agyei-Frempong et al., 2008; Ipadeola et al., 2015; Muazu et al., 2016) (**Table 2-1**). However, there may be a lower LADA prevalence in East Africa, most obviously in Kenya (5.7%) (Otieno et al, 2008). These findings should be interpreted carefully given differences in study design, selection criteria, sample size and potential variation in access to healthcare. Additionally, current studies have been largely concentrated in West Africa, and therefore further studies expanding to regions throughout Africa are still required. Larger population-based studies will undoubtedly lead to more accurate comparisons of LADA prevalence between populations of European and African ancestry.

2.2.3 Prevalence of LADA in African-American populations

Limited research exists on the prevalence of LADA in African Americans, as the relative scarcity of sample collections continues to be a challenge. There is currently only one study of LADA in African Americans, where diabetes-associated autoantibodies were measured in 295 non-Hispanic black adults, with 4.7% of subjects being autoantibody positive (Barinas-Mitchell et al., 2004). This prevalence estimate was significantly higher than autoantibody positive nondiabetic controls (1.3%) and significantly lower than that of autoantibody positive non-Hispanic whites (8.6%) in the same study. In this particular analysis, there was a higher prevalence of autoantibody positive non-insulin requiring

diabetics in non-Hispanic white adults than in non-Hispanic black adults **(Table 2-1)**.

With respect to T2D and the difference in prevalence between those with European and African ancestry, racial admixture should also be considered. One of the highest prevalence rates of T2D has been seen in African Americans (18.7%) compared to 10.2% observed in European Americans (Cooke et al., 2012). This African American prevalence estimate is likely due to socioeconomic and behavioral risk factors that have resulted in relatively unhealthy lifestyles and higher rates of obesity (Cowie et al., 2006). However, in support of a genetic influence, African-European mixed-ancestry individuals have a higher risk of T2D compared to those of European ancestry alone (Ng et al., 2014). Indeed, one American study, which considered racial admixture, found a significant association between African Americans with a higher admixture of African ancestry and T2D (Cheng et al., 2012). Although T2D is more prevalent in African Americans than Africans (Tull & Roseman, 1995), this increased prevalence has been observed in African migrants (Mbanya et al., 2010), suggesting a strong impact of environmental factors contributing to the progression of T2D.

Similar to LADA, the T1D prevalence in African Americans is also relatively less than in European Americans (approximately 1.62 per 1000 individuals as of 2009 versus 2.55 per 1000 individuals, respectively) (Dabelea et al., 2014). Furthermore, one study observed a higher incidence of T1D among African American children living in Allegheny, Pennsylvania (11.8 per 100,000) than in Jefferson County, Alabama (4.4 per 100,000) (Tull & Roseman, 1995), where the genetic admixture is 21.2% and 17.0%, respectively. These results likely reflect

European Americans being at higher risk for childhood-onset T1D than African Americans. The trend in LADA diagnoses in racially admixed individuals remains unclear, and more studies with larger samples are needed to fully understand the epidemiology of LADA in African and African American populations.

Prevalence of LADA in Asian and Middle Eastern populations

LADA studies in other parts of the world, specifically non-European populations, will also help determine to what extent discrepancies in prevalence are due to genetics. Despite the high prevalence of T2D in populations of Asian ancestry, LADA prevalence has been reported to be generally lower in China, Japan, Korea and India (2.5-9.2%) than in Europe (**Table 2-1**) (Hwangbo et al., 2012; Katulanda et al., 2008; Park et al., 2011; Qi et al., 2011; Sachan et al., 2015; Takeda et al., 2002; Zhou et al., 2013). In a study of Indian, Malay and Chinese populations from Singapore (Ong et al., 2017), GAD autoantibodies were more frequent in a European cohort (13.9% vs. 11.4%, 6.0%, 5.8%, respectively), while IA-2 antibody positivity was higher in the Asian ethnic groups, which has also been shown in an East Indian populations (Kanungo & Sanjeevi, 2003). A high prevalence of LADA was observed in a small hospital-based study in Kerala, India (Unnikrishnan et al., 2004), where 25% of 83 lean patients with apparent T2D were positive for GAD autoantibody. Studies based on hospital patients, such as this study, are more likely to show a greater frequency of autoantibody positive patients given that they have more severe disease, as noted above. Overall, countries across Asia have collectively reported a wide range of prevalence estimates of LADA compared to European and African ancestral groups. Despite the high prevalence of T2D in populations of Asian

ancestry, the T1D prevalence and incidence rates are reported to be low (Dabelea et al., 2014). Collectively, however, both T1D and LADA have a lower prevalence in Asia, with T2D being by far the most prominent form of diabetes.

Strikingly**,** the prevalence of LADA in Middle Eastern countries may be as high as 14% (**Table 2-1**) (Hossein et al., 2015; Maddaloni et al., 2015). However, and in contrast, a large cross-sectional population-based study in the United Arab Emirates only reported LADA and classic T1D in 2.6% and 0.2%, respectively, of 17,072 subjects with adult-onset diabetes (>30 years) (Maddaloni et al., 2015). Prevalence rates of adult-onset diabetes therefore vary substantially across Northern Africa and the Middle-East. According to the International Diabetes Federation, the total number of adults with diabetes in the Middle East is also predicted to see a large increase in the prevalence of diabetes over time. A high T1D occurrence has recently been reported in the Middle-East (**Appendix: Supp. Table 2-1**), with the incidence rates in Kuwait amongst the highest in the world (Ogurtsova et al., 2017; Shaltout et al., 2017). Given these strikingly high prevalence estimates in Middle Eastern countries for both T1D and T2D, but also potentially for LADA cases, this region would particularly benefit from improvements in classification of diabetes subtypes.

As a brief overview of this section, the majority of autoimmune diabetes studies have, thus far, been in populations of European ancestry, with only limited research that is almost invariably underpowered conducted in other ethnicities. Childhood-onset T1D is known to be most prevalent in populations of European ancestry, but strikingly the converse can be true for adult-onset autoimmune diabetes including LADA. Indeed, the prevalence of LADA varies across Africa,

the Middle East and Asia. Unfortunately, many populations affected by this adult-onset form of autoimmune diabetes frequently have the least access to autoantibody screening, hampering efforts to distinguish LADA from T2D. Only with proper diagnoses and ascertainment of individuals afflicted with LADA from large populations of diverse ancestries can genetic studies be implemented.

Ethnicity	Author/ Year	Location	Type of study[1]	Study Size	Inclusion	Age of onset[2]	Measured AA	# Cases positive for AA	Frequency of AA positivity (%)
African	Muazu *et al.* (2016)	Northern Nigeria	C	200	Non-insulin requiring T2D patients	>30	GAD	21	10.5
	Ipadeola et al. (2014)	Southwest Nigeria	C	160	Individuals with T2D	>30	GAD	19	11.9
	Adeleye *et al.* (2012)	Southwest Nigeria	C	235	Individuals with T2D	>30	GAD	33	14
	Agyei *et al.* (2008)	Ghana	C	120	Recently diagnosed (<1 year) patients with diabetes	>35	GAD	14	11.7
	Otieno *et al.* (2008)	Kenya	C	124	Individuals with T2D	>40	GAD	7	5.7
Asian	Takeda *et al.* (2002)	Japan	C	4,980	Patients with diabetes	>20	GAD	188	3.8
	Qi *et al.* (2010)	China	P	8,109	Three step randomized sampling procedure from population, T2D were then identified	≥15	GAD	46	9.2
	Zhou, *et al.* (2013)	China	C	4,880	Individuals with diabetes duration of 1 year and no ketoacidosis or insulin dependence in the first 6 months after diagnosis.	≥30	GAD	287	5.9
	Hwangbo *et al.* (2012)	Korea	C	462	Individuals with T2D, within 5 years from the time of study	>20	GAD	20	4.3
	Park *et al.* (2011)	Korea	C	884	Patients with diabetes for > 2 months and < 5 years, and diabetes controlled by diet or by oral anti-diabetic agents for at least (≥) 2 months after diagnosis	35-70	GAD, IA2, or ZnT8A	39	4.4
	Katlunda et al.	India	C	992	Individuals with diabetes aged ≤45 years.	16-40	GAD	54	5.4
	Sachan, *et al.* (2014)	India	C/ P	618	Individuals with T2D	>30	GAD and/or IA2	9	2.6
European	Turner *et al.* (1997)	UK	P	3,672	Individuals recently diagnosed with non-insulin requiring diabetes	25–65	GAD and/or ICA	430	12
	Tuomi, *et al.*(1999)	Finland	P	1,122	Individuals diagnosed with non-insulin requiring diabetes	28–83	GAD and/or IA-2	104	9.3
	Zinman, *et al.*(2004)	USA, Europe	C	4,357	Individuals recently diagnosed with non-insulin requiring diabetes	30–75	GAD and/or IA-2	174	4.2
	Buzzetti, *et al.*(2007)	Italy	P	4,250	Participants in NIRAD Study with T2D	30–75	GAD and/or IA-2	193	4.5
	Radtke, *et al.* (2009)	Norway	P	1,049	Participants in HUNT study with diabetes	≥20	GAD	106	10
	Maioli *et al.* (2010)	Sardinia	C	5,568	Individuals with T2D	35–70	GAD	276	4.9
	Hawa, et al. (2013)	Europe	C	6,810	Individuals with diabetes	30–70	GAD and/or IA-2, ZnT8	598	9.7
	Bosi *et al.*(1999)	Italy (Northen)	P	2,076	Participants in Cremona Study, then individuals with diabetes were identified	>40	GAD	4	2.8
Middle Eastern	Maddaloni *et al.* (2015)	United Arab Emirates	C	17,072	Adult onset diabetes, with sufficent data available	30–70	GAD and/or IA-2	437	2.6
	Hossein *et al.*(2015)	Iran	P	500	Individuals with T2D with insulin independency at least first 6 months	> 35	GAD	71	14.2

Table 2-1 Prevalence of autoantibody-positive individuals among adults with diabetes across different population groups.

2.3 Importance of autoantibody screening

In order to perform well-powered population-based epidemiological and genetic studies, more extensive autoantibody screening is crucial for early detection of autoimmunity in adult-onset diabetes. The presence of a single autoantibody alone, however, cannot be relied on for a categorical diagnosis of LADA given the potential for false positive and false negative assay results, as well as the variation in observations across racial/ethnic groups (Xiang et al., 2015). Notably, in a cross-sectional study (Siraj et al., 2016) in Ethiopia measuring islet-cell associated antibodies in T1D, T2D and non-diabetic controls, IA-2 autoantibodies were absent in all groups, suggesting that the clinical utility of IA-2 autoantibodies may be limited in some populations. Furthermore, individuals with LADA in China and India also have lower average autoantibody titers, especially for GAD and IA-2 autoantibodies, than their European counterparts, making LADA much more challenging to diagnose in these populations (Sachan et al., 2015; Zhou et al., 2013). In particular, 92.7% of autoantibody positive cases in the Action LADA study (Hawa et al., 2013) had GAD autoantibody positivity while the remainder of cases were positive for IA-2 autoantibody or ZnT8A autoantibodies. However, in the LADA China study, GAD frequency was 67%, and non-GAD autoantibodies were more prevalent in this Chinese cohort (Zhou et al., 2013). Therefore, given that some autoantibodies are less prevalent in certain populations, it is crucial to measure more than one marker to capture the true frequency of autoimmune diabetes.

Despite the high frequency of misdiagnosed cases and the availability of clinical screening tools for identifying LADA risk, antibody screening is not common in practice and is usually limited to only those who are considered to be

at very high risk (invariably, lean young adults with high HbA1c (American Diabetes Association, 2009)). However, this approach will miss a substantial number of cases, and it is challenging to clinically identify patients with LADA, because some may be obese, with metabolic syndrome and mild metabolic dysfunction. As a consequence of stringent criteria, current standards for high risk patients still dismiss true LADA cases (Zinman et al., 2004). Thus, identifying at-risk patients through routine autoantibody screening for adult onset autoimmune diabetes may substantially decrease the number of misdiagnosed cases and could potenitally reduce health care costs. In addition, the development of oral therapy used to treat T1D, for example SGLT2 inhibitors (sotogliflozin) (Garg et al., 2017), and immune therapy to limit the immune effect could be potentially useful in treating LADA.

2.4 Concluding remarks

Despite the limited number of LADA studies in the ethnicities highlighted in this chapter, and the challenges in defining the trait, this diabetes subtype appears to be more prevalent than childhood-onset T1D, particularly among populations of African ancestry. Additionally, a proportion of cases with T2D are misdiagnosed LADA cases irrespective of ancestry, and this misdiagnosis rate is perhaps highest in cases of West African ancestry; however, it is unclear whether diagnostic practices are significantly different across regions, impacting prevalence estimates. Furthermore, the prevalence of autoantibody positivity in diabetes cases with non-European ancestry is lower than in European populations with misdiagnosis likely due to relatively low levels of GAD autoantibodies. Screening programs designed for populations at risk should

enable earlier intervention with improved metabolic outcomes, appropriate therapy and improved identification of comorbidities (Martinell et al., 2017).

Finally, genetic studies are particularly needed in African and Asian populations to identify ancestry-specific genetic susceptibility loci for autoimmune diabetes, most specifically LADA. Of the studies published, sample sizes have been relatively small and thus statistically underpowered to fully dissect the issues raised in this chapter. In addition to having more power, the collection of samples needs to be more representative of the population to gain more clarity and accurate prevalence estimates. The polymorphic nature of the MHC region means that the frequencies of susceptibility and protective alleles vary widely between and within ethnicities, further highlighting the need for population-specific studies. Future studies should focus on comparing genetic features among different subtypes across ethnic groups affected with diabetes in order to further improve prognosis, diagnosis and treatment of diabetes, ultimately allowing us to move beyond a dependence solely on autoantibodies to make such diagnoses.

CHAPTER 3. RELATIVE CONTRIBUTION OF T1D AND T2D LOCI TO THE GENETIC ETIOLOGY OF LADA

The content of this chapter is published in Mishra R., Chesi A.*, Cousminer DL. *, Hawa MI., Bradfield JP., Hodge KM., Guy VC., Hakonarson H.,BMDCS, Mauricio D., Schloot NC, Yderstræde KB., Voight BF., Schwartz S., Boehm BO., Leslie RD. and Grant SFA. "Relative Contribution of Type 1 Diabetes and Type 2 Diabetes Loci to the Genetic Etiology of Adult-Onset Non-Insulin Requiring Autoimmune Diabetes". BMC Medicine (2017); 15(1):88. PMID: 28438156. *equal contribution. RM's contribution to the manuscript includes study concept and design, analysis and interpretation of data, drafting, critical revision, and writing the final version of the manuscript. RM had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy for the data analysis.*

3.1 Introduction

Diabetes is a heterogeneous group of diseases resulting in hyperglycaemia due to insulin secretory dysfunction as well as insulin resistance. A substantial proportion of type 1 diabetes (T1D) cases present in adulthood, and despite the presence of diabetes-associated autoantibodies, the majority of these patients do not initially require insulin (Hawa et al., 2013; Maddaloni et al., 2015). The manifestation of this 'latent autoimmune diabetes in adulthood' (LADA) is clinically defined by (i) an adult age of onset, (ii) at least one diabetes-associated

autoantibody, and (iii) the lack of requisite insulin treatment for at least 6 months after diagnosis. This definition overall represents ~5-10% of all cases of adult-onset diabetes, potentially the most frequent form of autoimmune diabetes (Tuomi et al., 1999; Zhou et al., 2013).

However, classifying adult-onset autoimmune T1D, including LADA, remains challenging. The need for insulin treatment is a clinical decision, while diabetes-associated autoantibodies are neither pathogenic nor categorical features of LADA. Decisions are further confounded by false positives when large numbers of patients are screened (Tuomi et al., 1993). Since LADA has intermediate features between T1D and type 2 diabetes (T2D), there are limits to the current classification of diabetes. New paradigms are needed to distinguish LADA and ensure appropriate disease treatment and management.

Recently, several studies have used genetic information derived from diabetes-associated risk variants across the genome to reclassify diabetes (Oram et al., 2016). To date, comprehensive genetic studies of T1D and T2D have uncovered dozens of distinct susceptibility loci for each of these two diseases (Bradfield et al., 2011; DIAGRAM et al., 2014; Wellcome Trust Case Control Consortium, 2007). Initial analyses of T1D loci in relatively small LADA cohorts have consistently shown an association with the T1D locus *HLA-DQB1,* which resides in the major histocompatibility complex (MHC) (Desai et al., 2007a; Horton et al., 1999; Tuomi et al., 1999), as well as at *PTPN22* and *INS* (Cervin et al., 2008; Desai et al., 2006). Similar analyses of T2D loci have suggested an association in LADA with the strongest T2D locus harboring *TCF7L2* (Cervin et al., 2008; Lukacs et al., 2012; Zampetti et al., 2010) and the *ZMIZ1* locus (M K Andersen et al., 2014). A significant challenge of these studies has been the lack

of statistical power due to the small number of LADA patients included. Thus, the genetic etiology of LADA remains largely unresolved.

To quantify the genetic liability to LADA contributed by genetic risk factors for T1D and T2D, we amassed the largest LADA cohort to date. By assessing the association of these variants in LADA, our objective was to place LADA along the etiological diabetes spectrum and reshape our understanding of the relationship between LADA and classic diabetes phenotypes.

3.2 Methods

Study populations and antibody testing

We ascertained 978 LADA cases from two studies, a European Union-funded multicenter study (Action LADA) and a German Research Council study (DFG: SFB 518, A1), each aimed to identify features of adult-onset autoimmune diabetes. A description of the participants and study design has been published elsewhere (Hawa et al., 2013). Briefly, for this study, all participants were diagnosed with LADA if: aged 30-70 years old, positive for diabetes-associated Glutamic Acid Decarboxylase (GAD) autoantibodies, without insulin treatment for at least 6 months after diagnosis. These cross-sectional studies included adult-onset diabetic patients recruited between 2003 and 2007 from Barcelona (Spain), Düsseldorf (Germany), London (United Kingdom), Odense (Denmark) and Ulm (Germany); two centers recruited patients from a community or primary care setting (Düsseldorf and Odense), and the remaining three centers (Barcelona, London and Ulm) recruited patients in a hospital setting. The ethics committees of all participating centers approved the protocol, and all participants provided

written informed consent. Patients were diagnosed with diabetes according to standard criteria, with at least two recorded fasting blood glucose measurements ≥7mmol/L, or 2 hour post-oral glucose blood glucose >10mM. Patients were excluded if their data was incomplete, they were pregnant, had renal disease (raised creatinine or proteinuria) or an acute illness at the time of testing. An attending physician recorded the medication data and risk factors. Serum and plasma samples were collected according to standard procedures and stored at -20°C.

Samples were tested in two central laboratories (London, UK and Ulm, Germany) for serum autoantibodies to glutamic acid decarboxylase (GADA) and insulinoma associated antigen-2 (IA2A), using established radioimmunoprecipitation assays. Positive results were replicated, reducing false positives to <0·2%. Using the Diabetes Antibody Standardization Program (Torn et al., 2008) (and unpublished data), we determined the sensitivity and specificity of the GADA assays (90%, 93% in London and 86%, 95% in Ulm, respectively) and the sensitivity and specificity of IA2A assays (68%, 95% in London and 73%, 99% in Ulm, respectively).

Each assay included serially diluted sera from antibody positive individuals. The cut-off for positivity was selected arbitrarily based on the end point of the standard curve and further confirmed with Quantile-Quantile probability plots for London (UK) and the 99th centile for Ulm (Germany).

The population-based control cohort comprised non-diabetic children of European ancestry, aged 5-20 years old, enrolled in the Bone Mineral Density in Childhood Study (BMDCS). Subjects were randomly recruited

from five different centers in the US. As previously reported (Kalkwarf et al., 2007), enrollment criteria included healthy, normally developing children. Each participating center received approval of the study by their respective institutional review boards. Participants 18 years old and older provided written informed consent. Written informed consent for participants younger than 18 years of age was obtained from the parent or guardian and assent was obtained from the participants.

The population-based control cohort comprised 1,057 non-diabetic children of European ancestry, aged 5-20 years, enrolled in the Bone Mineral Density in Childhood Study (BMDCS). Subjects were randomly recruited from five different centers in the US. As previously reported, enrollment criteria included healthy, normally developing children. Each participating center received approval of the study by their respective institutional review boards.

Since BMDCS consists of European-descent children ascertained from the United States, while the LADA cases were adults ascertained from the UK and Germany, we also leveraged 2,820 healthy adult British Birth Cohort controls from the Wellcome Trust Case Control Consortium (WTCCC) (Wellcome Trust Case Control Consortium, 2007) to act as an extra set of controls to verify our observations. Principal component analysis (PCA) showed that BMDCS controls were well-matched with cases despite ascertainment in the United States, while the WTCCC controls were stratified (**Appendix: Supp. Figure 3-1**), principally due to

differences in the genotyping arrays used. Thus, BMDCS was used in the primary analyses, with verification in the WTCCC cohort. Our study also utilized publicly available childhood-onset T1D (n=2,000) and adult-onset T2D (n=1,999) from the Wellcome Trust Case Control Consortium; these individuals were recruited within England, Scotland and Wales (Wellcome Trust Case Control Consortium, 2007). Individual data from WTCCC is available through the Consortium's Data access committee (http://www.wtccc.org.uk).

Genotyping

LADA samples and BMDCS controls were genotyped on the Illumina Infinium II OMNI Express plus Exome BeadChip array (Illuminia, San Diego, CA, USA) at the Children's Hospital of Philadelphia Center for Applied Genomics (Philadelphia, PA, USA). WTCCC T1D and T2D cases were genotyped on the Affymetrix 500K genotype array. WTCCC control samples consisted of 3,000 1958 British Birth Cohort control samples genotyped on the Illumina 660K genotyping array. The genomic inflation factor for the pruned genome-wide SNPs is 0.966 and the QQ-plot can be found in **Appendix: Supp. Figure 3-2.**

Quality Control and Imputation

PLINK (Purcell et al., 2007) was used to exclude individuals with incorrect gender assignments or whose gender could not be determined by genotype, duplicate individuals, and individuals with missing genotype

rate >5%. Principal component analysis was used to exclude ethnic outliers. Single nucleotide polymorphisms (SNPs) with a call rate <95%, minor allele frequency <0.5%, Hardy-Weinberg equilibrium $P<10^{-5}$ and with A/T and G/C alleles were removed. After quality control, genotypes were imputed to the 1000 Genome Phase I Integrated Release Version 3 reference panel. A two-step imputation process was performed using SHAPEIT (O'Connell et al., 2014) for haplotype phasing and IMPUTE2 (Howie et al., 2012) for imputation. All SNPs in this study had imputation quality scores >0.40.

Individual candidate SNP association tests

To investigate the role of previously discovered T1D and T2D variants in LADA, we tested 67 T1D SNPs (from Immunobase; http://www.immunobase.org and 71 T2D SNPs (from the T2D study led by the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium (DIAGRAM et al., 2014)). Association between each SNP and case/control status was assessed using a univariate linear mixed model within GEMMA (Zhou & Stephens, 2012). This model accounts for population stratification and relatedness using the Wald test and the restricted maximum likelihood estimate of ß. We tested each SNP in LADA cases versus BMDCS controls and in LADA cases versus T1D or T2D cases. Significant associations were called after Bonferroni correction for multiple testing. Analysis was performed for all LADA cases (n=978),

LADA cases positive for GADA only (n=669), and LADA cases positive for both GADA and IA2A (n=309). Approximated odds ratios were calculated using μ (intercept) and β (effect size) estimates from the linear mixed model, with the formula: OR $=\frac{\beta}{e^{\mu(1-\mu)}}$ (Zhou & Stephens, 2012).

Genetic risk scores (GRS)

We calculated two GRS using 69 T1D SNPs and 71 T2D SNPs for T1D cases (n=1,990), T2D cases (n=1,960), LADA cases (n=978), LADA cases positive for GADA only (n=669), LADA cases restricted on GADA+ IA2A+ status (n=309), and BMDCS controls (n=1,057). Weights utilized for the scores were derived from published odds ratios (ORs) from T1Dbase (t1dbase.org) or a previous publication (Morris et al., 2012), respectively. Two SNPs, rs2187668 and rs7454108, were used to infer *HLA DR3/DR4/DQ8* haplotypes, and additional HLA SNPs tagging *HLA-A*, *HLA-B*, and *DRB1* haplotypes were included (Oram et al., 2016; Winkler et al., 2014). rs7111341 and rs11171710 did not have publicly available ORs, and rs7202877, implicated in both T1D and T2D (**Appendix: Supp. Table 3-1**), so these were excluded. Each GRS was calculated using PLINK by multiplying the number of risk-increasing alleles by the natural log of the OR at each locus and summing across risk loci for each individual. Logistic regression and receiver operating characteristic (ROC) curve analyses evaluated how well these GRS distinguished LADA cases from BMDCS controls (using the PredictABEL package (Kundu et al.,

2011)). We repeated the GRS calculation for GADA+ and IA2+ LADA cases and for GADA+, IA2A- LADA cases. Additionally, we combined the T1D and T2D SNPs (139 SNPs) and classified LADA and controls for both LADA groups. The distributions of the T1D and T2D GRS of the five groups were compared using the Wilcoxon rank sum test accounting for multiple comparisons (using a Bonferroni correction). Control samples were obtained from the WTCCC study, as described above.

3.3 Results

T1D loci:

Four T1D SNPs were significantly associated with LADA and survived multiple testing correction (P=0.05/67 loci tested =7.46x10^{-4}; **Table 3-1 and Appendix: Supp. Table 3-2**). The strongest association was in the MHC region (OR=1.46 P=9.64x10^{-11}). Strong association was also observed for variants at *PTPN22* (OR=1.47; P=6.38x10^{-6}), *SH2B3* (OR=1.28; P=1.10x10^{-5}), and *INS* (OR=1.27; P=2.39x10^{-4}). The association signal within the MHC region was significantly different between LADA and T1D cases ($P_{difference}$=1.26x10^{-17}), with the T1D risk allele of rs9272346 (A) less common in LADA than in T1D, but still at a higher frequency than in controls. The signals at *INS* and *SMARCE1* also yielded significant differences between LADA and T1D ($P_{difference}$=3.88x10^{-4} and 6.54x10^{-4}, respectively). The *INS* signal was more common in LADA than in either T1D or controls, while the frequency of the *SMARCE1* signal was lower in LADA than in T1D but similar to controls.

Locus	SNP	T1D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T1D *P*-value
			LADA	T1D	Control			
MHC	rs9272346	A/G	0.686	0.818	0.579	1.455 [1.427-1.483]	$9.6x10^{-11}$	$1.26x10^{-17}$
PTPN22	rs6679677	A/C	0.143	0.17	0.093	1.469 [1.427-1.510]	$6.38x10^{-6}$	$2.61x10^{-2}$
SH2B3	rs17696736	G/A	0.515	0.503	0.44	1.277 [1.250-1.304]	$1.10x10^{-5}$	0.542
INS	rs689	T/A	0.796	0.741	0.73	1.265 [1.234-1.296]	$2.39x10^{-4}$	$3.88x10^{-4}$
SMARCE1	rs7221109	C/T	0.621	0.687	0.632	0.954 [0.925-0.983]	0.423	$6.54x10^{-4}$

Table 3-1 Association of established T1D loci with LADA. Only T1D variants significantly associated with LADA are shown (LADA Association P-value), as well as signals significantly different between LADA and T1D (LADA vs. T1D P-value), with a significance threshold of $p = 7.46 \times 10^{-4}$. The locus reported is the closest gene of interest to the signal (a full list of genes is provided in **Appendix: Supp. Table 3-2**). The risk and other alleles reported refer to the alleles in T1D, and the following allele frequencies refer to the frequency of the risk allele reported in T1D for LADA, T1D and BMDCS control group. Odds ratios of the risk allele reported are derived from the BMDCS control data set (n=1,057), the WTCCC T1D (n=1,990), and the LADA cases (n=978).

To further understand the influence of antibody status on the clinical classification of LADA, the same analyses were carried out for 669 GADA+ LADA subjects (**Appendix: Supp. Table 3-3).** The MHC region was the only signal surviving correction for multiple comparison for cases against controls, as well as cases versus T1D (OR=1.30; P=$6.84x10^{-5}$, $P_{difference}$=$1.99x10^{-24}$).

In the restricted subset of GADA+ IA2A+ LADA cases (n=309), four loci were associated (**Table 3-2 and Appendix: Supp. Table 3-4**). The MHC (OR=1.98; P=$1.20x10^{-15}$), *PTPN22* (OR=1.86; P=$2.19x10^{-6}$), *SH2B3* (OR=1.48; P=$5.93x10^{-6}$), and *INS* (rs689; OR=1.44; P=$1.90x10^{-4}$) signals remained strongly associated and had stronger ORs in this constrained setting. However, the risk increasing allele at the MHC locus remained

significantly less than that in T1D cases. Two partially independent signals near *INS* (r^2=0.278) yielded a significant difference between T1D and GADA+ IA2A+ LADA in this restricted dataset, rs689 ($P_{difference}$=1.68x10^{-6}) and rs7111341 ($P_{difference}$=2.39x10^{-4}).

Locus	SNP	T1D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T1D *P*-value
			LADA	T1D	Control			
MHC	rs9272346	A/G	0.763	0.818	0.579	1.983 [1.954-2.012]	1.20x10^{-15}	4.01x10^{-3}
PTPN22	rs6679677	A/C	0.17	0.17	0.093	1.864 [1.819-1.909]	2.19x10^{-6}	0.603
SH2B3	rs17696736	G/A	0.542	0.503	0.44	1.481 [1.452-1.511]	5.93x10-6	0.180
*INS**	rs689	T/A	0.824	0.741	0.73	1.440 [1.407-1.474]	1.90x10^{-4}	1.68x10^{-6}
*INS**	rs7111341	C/T	0.812	0.75	0.73	1.360 [1.327-1.394]	1.82x10^{-3}	2.39x10^{-4}

Table 3-2 Association of established T1D loci in LADA subjects positive for both GADA and IA2A. Only T1D variants significantly associated with LADA are shown (LADA Association P-value), as well as signals significantly different between LADA and T1D (LADA vs. T1D P-value). Significance threshold is 7.46x10-4 after correcting for multiple comparison. The locus reported is the closest, well-known gene of interest to the signal (a full list of genes is provided in Appendix: Supp. Table 3-3). The risk and other alleles reported refer to the alleles in T1D, and the following allele frequencies refer to the frequency of the risk allele reported in T1D for LADA, T1D and BMDCS control group. Odds ratios of the risk allele reported are derived from the BMDCS control data set (n=1,057), the WTCCC T1D (n=1,990), and the constrained N LADA cases (n=309). *Independent signals (INS signals have a r^2=0.278)

T2D loci:

Only one T2D signal survived correction for multiple comparisons (*P*=0.05/71 loci =7.04x10^{-4}) in LADA cases, the *HNF1A* locus (OR=1.291; *P*=3.42x10^{-4}; **Table 3-3 and Appendix: Supp. Table 3-5**). Contrary to

previous reports (Cervin et al., 2008; Lukacs et al., 2012; Zampetti et al., 2010), the T2D risk allele (rs7903146-T) at *TCF7L2* was not enriched among LADA cases, with a frequency close to that of controls (0.295 vs. 0.298, respectively); indeed, the *TCF7L2* signal was the most significantly different signal between LADA and T2D cases ($P_{difference}$=5.21x10^{-6}). In the GADA+ restricted set, there were no association signals surviving correction for multiple comparisons, and the only signal showing a significant difference between LADA and T2D was the depletion of the *TCF7L2* T allele ($P_{difference}$=5.03x10^{-4}; **Appendix: Supp. Table 3-6**), where the T allele showed modest, albeit non-significant excess when compared to controls (OR=1.088).

Locus	SNP	T2D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T2D *P*-value
			LADA	T2D	Control			
HNF1A	rs12427353	G/C	0.831	0.828	0.787	1.291 [1.256-1.326]	3.42x10^{-4}	0.538
TCF7L2 *	rs7903146	T/C	0.295	0.376	0.298	1.023 [0.994-1.053]	0.702	5.21x10^{-6}

Table 3-3 Association of established T2D loci with LADA. Only T2D variants that significantly associated with LADA after correcting for multiple comparison (p<7.04x10^{-4}) are shown (LADA Association P-value), as well as variants significantly different between LADA and T2D (LADA vs. T2D P-value). The locus reported is the closest, well-known gene of interest to the signal (a full list of genes is provided in Appendix: Supp. Table 3-4). The risk and other alleles reported refer to the alleles in T2D, and the following allele frequencies refer to the frequency of the risk allele reported in T2D, for LADA, T2D, and BMDCS control groups. Odds ratios of the risk allele reported are derived from the BMDCS control data set (n=1,057), the WTCCC T2D (n=1,960), and the LADA cases (n=978). *Although the control risk allele frequency is greater than the case risk allele frequency, the beta calculated from the linear mixed model is adjusted effects after controlling for population stratification, resulting in an OR slightly above 1.

In the restricted set of 309 GADA+ IA2A+ LADA subjects, *HNF1A* continued to yield a significant association (OR=1.47; *P*=2.52x10^{-4}; **Table**

3-4 and Appendix: Supp. Table 3-7). Again, the *TCF7L2* locus was significantly different between LADA and T2D cases ($P_{difference}$=2.56x10^{-7}), with the risk allele frequency even less than that in controls in this restricted case set (allele frequency of 0.251 versus 0.298 in LADA and controls, respectively).

Leveraging 2,820 healthy adult British subjects from the WTCCC as alternative controls, we observed very consistent results overall (**Appendix: Supp. Table 3-8 and Supplemental Results**) despite the array differences for this set.

Locus	SNP	T2D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T2D *P*-value
			LADA	T2D	Control			
HNF1A	rs12427353	G/C	0.857	0.828	0.787	1.474 [1.438-1.511]	2.52x10^{-4}	5.42x10^{-2}
ZBED3	rs6878122	A/G	0.744	0.658	0.706	1.216 [1.184-1.249]	3.86x10^{-2}	1.47x10^{-5}
TCF7L2	rs7903146	T/C	0.251	0.376	0.298	0.852 [0.820-0.883]	8.14x10^{-2}	2.56x10^{-7}

Table 3-4 Association between established T2D loci in LADA cases positive for GAD and IA2 autoantibodies.T2D variants that were significantly associated in LADA cases positive for GAD and IA2 autoantibodies (n=309) (LADA Association P-value) are shown, as well as signals that were significantly different between LADA and T2D cases (LADA vs. T2D P-value). The significance threshold was set to P<7.04x10^{-4} to correct for multiple testing. The locus reported is the closest, well-known gene of interest to the signal (a full list of genes is provided in Appendix: Supp. Table 3-5). T2D risk allele frequencies reported are derived from the BMDCS control data set (n=1,057), the WTCCC T2D (n=1,960), and the LADA cases positive for autoantibodies GADA and IA2A (n=309). The odds ratios for LADA are shown both GEMMA-corrected (for relatedness and batch effects) and uncorrected.

Genetic risk scores

A high T1D GRS implies a high genetic risk for that disease. **Figure 3-1** shows that the T1D GRS better predicted whether a subject is a LADA case or control than the T2D GRS. The areas under the curve (AUC) for

the T1D and the T2D GRS were 0.667 and 0.565, respectively (**Figure 3-1A**). Thus, when considering adult-onset diabetes patients who do not initially require insulin, genetic risk defined for T1D could better identify autoimmune diabetes cases than genetic risk defined for T2D.

This result was more pronounced when considering controls versus 309 GADA+ IA2A+ LADA cases (**Figure 3-1B**) (AUC for T1D GRS=0.760, T2D GRS=0.496). However, these results were less pronounced for the 669 GADA+ only LADA cases versus controls (**Figure 3-1C**) (AUC for T1D GRS=0.623, T2D GRS=0.597). The combined effect of genetic risk using both T1D and T2D SNPs marginally improved classification of LADA cases and controls (AUC=0.673) and classification of GADA+ LADA and controls (AUC=0.635). However, there was no improvement of classification between GADA+ IA2A+ LADA and controls (AUC=0.755) using a combination of T1D-T2D SNPs. To highlight the important role of non-HLA loci in discrimination, we calculated T1D GRS without the HLA region and an HLA only GRS. Additionally, we tested these five models of GRS in discrimination between LADA versus T1D and GADA+ only LADA cases versus GADA+IA2A+ LADA cases (**Appendix: Supp. Figure 3-3**). The HLA alone accounts for a strong difference between all LADA cases and T1D cases (AUC=0.699), especially between T1D and GADA+ only LADA cases (AUC = 0.733). The non-HLA GRS has AUC of 0.655 for distinguishing GADA+IA2A+ LADA cases from BMDCS controls. HLA

only GRS has an AUC of 0.737 for distinguishing GADA+IA2A+ LADA cases from BMDCS controls, but combining these loci, the AUC is 0.76.

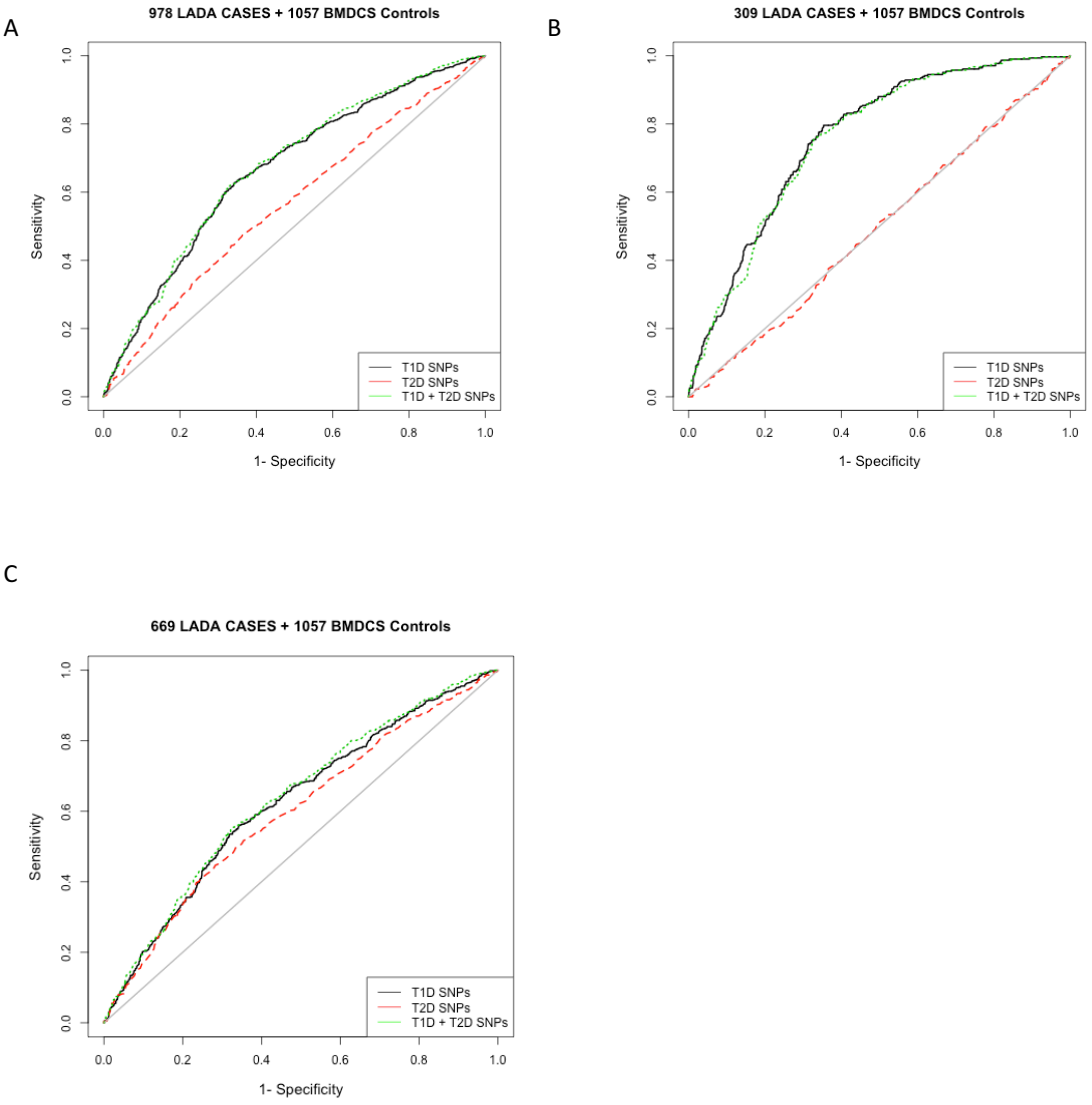


Figure 3-1 T1D and T2D genetic risk scores tested in LADA cases and controls. Weighted genetic risk scores (GRS) for T1D (black) and T2D (red) were calculated by summing over all the risk alleles (T1D/T2D SNPs). The scores were tested in A) 978 LADA cases and 1,057 healthy controls B) 309 autoantibody-positive (GADA, IA2A) LADA cases and 1,057 controls C) 669 GADA only autoantibody positive. The ability of the GRS to discriminate between cases and controls was assessed by receiver operator characteristic (ROC) analysis. The area under the curve (AUC) was 0.667 and 0.565 for T1D and T2D, respectively, in the set with all LADA cases, 0.76 for T1D and 0.496 for T2D in the GADA, IA2Aautoantibody-positive restricted set, and 0.623 for T1D, 0.597 for T2D in the GADA only autoantibody positive restricted set. A combination of T1D and T2D SNPs (green) had an AUC of 0.673 for all samples,0.755 for the GADA+ IA2A+ restricted set and 0·635 for the GADA only restricted set.

Comparison of the T1D SNP-GRS distributions among the six groups (T1D, T2D, LADA, GADA+ IA2A+ LADA, GADA+ only LADA and controls; **Figure 3-2A**) revealed significant differences between all pairs ($P<10^{-5}$), except T2D versus controls. This observation is expected as T2D cases should not harbor a high load of T1D risk alleles. Furthermore, there were only nominally significant differences between LADA and GADA+ only LADA cases. Of particular note, there was a significant difference in the T1D GRS distribution between T1D and GADA+ IA2A+ LADA, highlighting genetic differences between LADA restricted on IA2A+ status and T1D (P=0.0001).

Comparison of the distributions of the T2D SNP-GRS (**Figure 3-2B**) revealed significant differences between LADA and T2D cases ($P=3.50\times10^{-11}$) and between the GADA+ IA2A+ LADA and T2D cases ($P=3.50\times10^{-16}$). These results suggest T2D risk alleles are not enriched in LADA, concordant with the results of our single-SNP analyses. However, the T2D SNP-GRS distribution was also significantly different between LADA and T1D cases ($P=6.10\times10^{-11}$) and controls ($P=8.00\times10^{-6}$). The T2D risk allele load, although not as high as for T2D, is still higher than that seen in T1D or among the healthy population. We observed a nominally significant difference for T2D risk allele load between GADA+ only LADA and T2D cases ($P=5.60\times10^{-3}$) and no statistically significant difference between GADA+ only LADA and overall LADA cases.

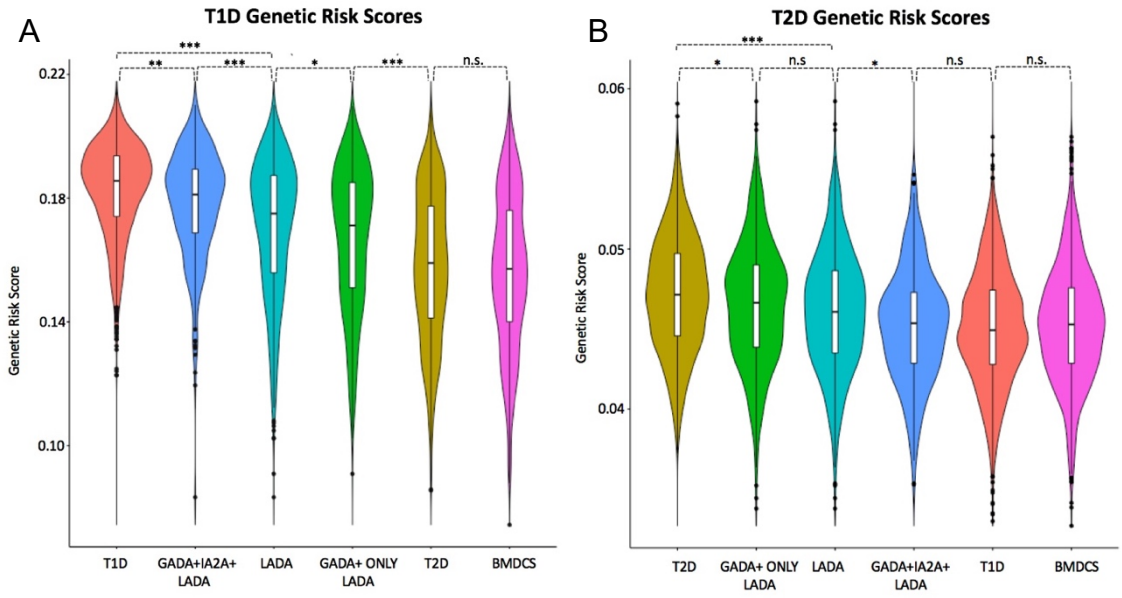


Figure 3-2 GRS distributions between T1D cases, T2D cases, LADA cases, LADA restricted cases, and controls. The GRS distributions were compared across individuals diagnosed with T1D (n=1,990), T2D (n=1,960), LADA (n=978), LADA restricted (n=309), LADA GADA+ only (n=669) and BMDCS controls (n=1,057). (A) Violin plots of the distributions of the GRS calculated using the T1D SNPS for the five groups. A multiple comparison test (Wilcoxon rank sum test) was performed to calculate the significance of pair-wise differences. (B) Violin plots of the distributions of the GRS calculated using the T2D SNPS for the five groups. A multiple comparison test (Wilcoxon rank sum test) was performed to calculate the significance of pair-wise differences. We include some of the significant P-values to highlight key differences. (***P <10-5, **P <0·0001, *P <0·05).

3.4 Discussion

Defining LADA as a distinct form of T1D has two broad benefits. First, it highlights the potential to understand what determines both the degree

and rate of disease progression. Second, it helps define differences between adult-onset autoimmune diabetes, including LADA, and T2D in terms of co-morbidities and putative therapy (Leslie et al., 2016). Leveraging children whose future diabetes risk is unknown represents the most conservative setting in which to conduct this study given they serve as excellent population-based controls in which to contrast the cases; however, the conservative nature of the approach may result in some false negative results.

To shed light on the genetic etiology of LADA, we tested the impact of established T1D and T2D risk loci in the largest set of LADA cases collected to date. Our study differs from a previous association study with GWAS-implicated loci in adult-onset autoimmune diabetes (Howson et al., 2011); first, our LADA cases are distinguished by the fact that they were not treated with insulin upon diagnosis. Furthermore, our study looked at a larger set of T1D and T2D loci, as well as comparing their roles in LADA against T1D and T2D, including taking population substructure into account. As with Howson *et al* (Howson et al., 2011), we observed significant association of the T1D loci *PTPN22, INS, HLA,* and *SH2B3.* However, we did not observe significant association with the *CLEC16A, IL2RA, CTLA4* and *STAT* loci. Despite published data observing the association of T2D locus *TCF7L2* with a subset of T1D patients (Redondo et al., 2017; Redondo et al., 2014), our study did not observe an association of this locus with LADA; one possibility could be that we used

population-based controls, while previous studies may have used a different control strategy where the difference in the risk allele was more evident due to its under-representation in relatively disease-free controls. Our study goes further by leveraging GRS to offer a further line of evidence for the classification of diabetes subtypes, complementing standards for clinical decision making and additional standardized (antibody) testing, each with their strengths and weaknesses.

LADA shows the MHC risk found in adult-onset T1D (Howson et al., 2011) with a reduced genetic susceptibility at this locus compared with childhood-onset T1D. Less clear is whether T2D loci play a role in adult-onset autoimmune diabetes. Our results show that genetic signals implicated in T1D or T2D both play a role in LADA, with four T1D loci and one T2D locus significantly associated with this form of diabetes. LADA is genetically more similar to T1D, especially when cases are constrained on both GADA+ and IA2A+, although LADA shares part of its genetic etiology with T2D. When constrained on GADA+ only, LADA cases became less distinct from T2D, highlighting the importance of IA2A in discriminating LADA within the T1D-T2D spectrum. By implication, a GRS derived from T1D can discriminate, to a degree, non-insulin requiring adult-onset diabetes patients with either autoimmune diabetes or T2D.

Regarding the loci implicated in T1D, our results are consistent with previous studies showing a major role for the MHC, *PTPN22* and *INS* loci in LADA (Cervin et al., 2008; Desai et al., 2006, 2007a). Interestingly, the

risk allele frequency at *INS* (rs689) was even more strongly associated with LADA than with T1D. Therefore, our data strongly points to common insulin-related pathways underpinning autoimmune diabetes irrespective of the age at onset of the disease. Given the evidence that age at diagnosis is genetically determined (Fava et al., 1998), these loci may play a key role in determining the age at disease onset and the rate of disease progression.

While our results suggest LADA is genetically closer to T1D than to T2D, we observed an association at one T2D locus, *HNF1A,* known to be associated with T2D and 'maturity-onset diabetes of the young'; strikingly, the *HNF1A* signal remained significantly associated with LADA even in the cohort enriched for both T1D autoantibodies. Nevertheless, the nature of the role of *HNF1A* in LADA is unclear, although any gene compromising insulin secretory function could predispose to diabetes. This is the first report describing an association between this T2D-associated risk allele and LADA, although this locus has been previously implicated in T1D (Andersen et al., 2014). Additionally, the strongest T2D-associated locus, *TCF7L2*, has been associated with LADA in a Finnish cohort (Andersen et al., 2014; Lukacs et al., 2012; Zampetti et al., 2010), but in our study, the risk allele frequency in LADA was very close to that of controls and lower than controls in GADA+ IA2A+ LADA. Our findings were further supported by leveraging healthy adult British controls from the WTCCC which provided overall consistent results, including for the *HNF1A* signal.

However, given the borderline association of T2D loci identified and the modest power in this single study, these signals must be subjected to replication efforts by independent investigators in order to fully validate these observations.

We found that GRS calculated from T1D and T2D-implicated SNPs, which distinguished LADA cases from controls, with the T1D GRS performing better than the T2D GRS; this difference was particularly striking in GADA+ IA2A+ LADA cases. Comparison of GRS between the six defined groups placed LADA in-between T1D and T2D but closer to T1D. GADA+ IA2A+ LADA was very similar to T1D, primarily because such constraint filters out "T2D like" cases and enriches for "T1D like" cases. The potential for clinical, immunological, or genetic filters to define forms of diabetes is emphasized by the marked overlap in GRS scores, even between T1D and controls.

This study does have limitations. First, GADA+ only LADA cases had a T2D-SNP GRS distribution more similar to T2D than controls. The specific association between the T2D risk score and GADA+ only LADA cases could be in part due to the fact that approximately 1% of all cases are widely considered to be false antibody-positive T2D. Thus, larger studies may resolve whether T2D risk alleles play a role in LADA. Indeed, this study was underpowered to identify specific associations other than for *HNF1A*. Second, two different genotyping arrays were utilized; thus, to correct for potential batch effects due to genotype array differences,

population substructure, and relatedness among samples, we used a linear mixed model, resulting in highly conservative effect estimates. Consequently, it is possible we have missed some true positive associations since we robustly controlled for false positive results.

The current nomenclature to classify diabetes, designating it as 'T1D' or 'T2D', was adopted to foster research and appropriate therapy for different phenotypic presentations. The combination of GRS, age at diagnosis, clinical phenotype, autoantibody assays, and C-peptide estimates as a proxy for insulin secretion, affords a more sophisticated approach with the potential to dissect the heterogeneity of diabetes (Oram et al., 2016). This study highlights the uncertainty of the current classification of diabetes (Schwartz et al., 2016). These results suggest that clinical phenotype alone is insufficient to define the major types of diabetes. To better treat the various diabetes subtypes, we need to integrate the use of clinical phenotype, metabolic status, immune changes, and underlying genetic risk.

3.5 Conclusion

LADA is genetically closer to T1D than T2D, although the genetic load of T1D risk alleles is less than childhood-onset T1D, particularly at the MHC, potentially accounting for the later disease-onset. Our results show that the genetic spectrum of T1D extends into adult-onset diabetes, where it can clinically masquerade as T2D. Furthermore, T2D genetic risk plays a small role in LADA, with a degree of evidence for the *HNF1A* locus,

highlighting the potential for GRS to contribute towards defining diabetes subtypes.

CHAPTER 4. HERITABILITY OF LADA

4.1 Introduction

Estimating heritability is central to studying and assessing the risk for genetic disease pathogenesis. Heritability is the proportion of phenotypic variance attributable to genetic variance (Visscher et al., 2008). Supporting evidence for trait heritability has been demonstrated through family (Falconer, 1965) and twin studies (Boomsma et al., 2002). For example, if the incidence of a given disease is higher among relatives of affected subjects than among individuals from the general population, this suggests heritable factors play a role in the disease. Additional evidence supporting a detectable genetic component is when monozygotic twins show more similarity for a particular disease compared to dizygotic twins. Although twin studies are classical approaches for minimizing environmental effects, obtaining sufficiently large sample sizes of family and twin material can often be relatively challenging. Additionally, observations from twin and family studies alone do not determine heritability estimates.

Alternatively, recent methods have started to leverage high throughput microarray genotype data to detect the fraction of phenotypic differences explained by the additive contributions from single nucleotide polymorphisms (SNPs). Consequently, the estimated heritability by these recent methods is represented by the total variance explained by SNPs ('SNP heritability'). In contrast to family and twin studies, individuals are typically unrelated in these genotyped datasets. It is also important to note that heritability estimates from family studies are represented by narrow-sense heritability, which is phenotypic

variance due to any type of additive genetic variation, not just SNPs. On the other hand,heritability estimates from twin studies are represented by broad–sense heritability, which is the proportion of phenotypic variance due to all genetic variation (i.e. dominance and epistasis) (Falconer, 1965). Therefore, SNP heritability should not be confused with narrow-sense and broad-sense heritability, and comparison of heritability estimates reported by different approaches should be interpreted cautiously. Three tools capable of SNP heritability estimation are: 'genome-wide complex trait analysis' (GCTA) (Yang et al., 2011), 'Linkage-Disequilibrium Adjusted Kinships' (LDAK) (Speed et al., 2012) and 'linkage disequilibrium score regression' (LDSC) (Bulik-Sullivan et al., 2015).

GCTA and LDAK estimate SNP heritability for a given complex trait from individual genotype data, while LDSC requires summary statistics derived from genome-wide association studies (GWAS). GCTA assumes a linear mixed model, in which SNP effects are normally distributed, along with fixed effects such as gender and principal components. A key step in the GCTA model is calculating the genetic relationship between pairwise individuals from all the autosomal SNPs, with this data stored in a genetic relationship matrix (GRM). LDAK is an extension of GCTA, where GRMs (termed kinship matrices in the LDAK model) are adjusted for linkage disequilibrium (LD) by adding weights to SNPs. Both GCTA and LDAK estimatevariance components using the restricted maximum likelihood (REML) method (Patterson & Thompson, 1971). REML is an estimation procedure that maximizes the log-likelihood of the linear mixed model by estimating the unknown parameter, in this case variance components, from the data available. LDSC, on-the-other-hand, assumes a simple polygenic model

and estimates heritability by solving the slope of the regression line produced when regressing test statistics from GWAS against LD score. Each SNP has an LD score which is defined by the sum of its r^2 (Pearson's correlation coefficient) with every SNP within a 1cM window. Under the GCTA and LDSC models, two SNPs in the same region have the same expected contribution to heritability regardless of LD, whereas LDAK models variants in perfect LD as contributing less to heritability than two SNPs exhibiting no LD. Both the GCTA and LDSC are commonly used, while LDAK is a relatively more recent method.

In this study, GCTA, LDAK, and LDSC were used to estimate SNP heritability for latent autoimmune diabetes in adults (LADA) for the first time. LADA is commonly referred to as "type 1.5 diabetes" because of its shared features with type 1 diabetes (T1D) and type 2 diabetes (T2D). T1D and T2D risk are known to be driven by a strong genetic component. In particular, previous studies have shown T1D to be more heritable than T2D (88% vs 26%, respectively), with first-degree relatives of T1D cases at ~15 times greater risk of presenting with the disease than the general population, while first degree relatives of individuals with T2D are only ~3 times more likely to present with the disease than individuals without a family history (Florez et al., 2003; Hyttinen et al., 2003; Poulsen et al., 1999). Despite the number of published heritability studies on these two classic forms of diabetes, there is currently no report of heritability estimates for LADA. Given that LADA shares features of both T1D and T2D, LADA heritability estimates are hypothesized to fall somewhere between T1D and T2D estimates.

4.2 Materials & Methods

Data

The raw dataset consisted of 2,843 LADA cases of European ancestry and 1,885 controls from the Bone Mineral Density Childhood Study (BMDCS). LADA cases were from 'ActionLada+Plus' (British, American, and German), DIREVA (Finnish), BOTNIA (Finnish), ANDIS (Swedish), and SDR (Swedish) (these cohorts are further described in **Appendix: Supp. Table 5-1**). LADA diagnosis criteria was defined by the presence of diabetes-associated autoantibodies and the lack of insulin requirement for at least 6 months after diagnosis. Additionally, diagnosis criteria included an age at diagnosis greater than 35 years old for all cohorts except ActionLada+Plus, where age at diagnosis criteria was between 30 and 70 years old.

Quality Control and Imputation

LADA samples were genotyped on the Illumina Infinium II Omni Express BeadChip and Illumina Infinium Core array (Illumina, San Diego, CA, USA). BMDCS controls were genotyped on the Illumina Infinium II OmniExpress. Post-genotyping quality control was performed using PLINK, including removal of samples with a call rate less than 95%, ambiguous gender, excess or reduced heterozygosity, and related individuals. The threshold for removing individuals with relatedness was pi_hat greater than 0.185, which was based on a pruned dataset. A correlation-squared threshold of 0.5 and a window of 10Mb was used for pruning SNPs. Individuals with divergent ancestry were removed from the analysis. At the end, 2,706 LADA cases and 1,254 controls remained. The

majority of controls were removed due to divergent ancestry and the remaining samples were removed due to relatedness and excess/reduced heterozygosity. Monomorphic SNPs were removed, as well as SNPs with minor allele frequencies (MAF) less than 5% or missingness rate less than 95%. Majority of SNPs were removed when merging all datasets. The Michigan Haplotype Reference Consortium (HRC) imputation service (URL) was utilized to perform imputation for autosomal SNPs only. Sex chromosomes were not included in this analysis.

Post imputation quality control

Post-imputation quality control consisted of retaining only biallelic autosomal SNPs that have a MAF greater than 1% and an information score greater than 0.99 (high quality SNPs). To ensure reliable heritability estimates, principal component analysis was performed and 20 principal components (PCs), suggested by the LDAK protocol, were used as covariates to account for population stratification during downstream analysis. Gender and significantly associated SNPs were also integrated into the model as fixed effects. There were 7,478,177 SNPs remaining after post- imputation quality control.

Heritability estimated by GCTA

After quality control, using GCTA, genetic relationships between pairwise individuals from all the autosomal SNPs were calculated and stored in a genetic relationship matrix (GRM). Finally, a genetic restricted maximum likelihood (GREML) algorithm was used to estimate variance component. SNP heritability was calculated using the GRM with gender and 20 PCs as covariates. The GCTA

tool was downloaded from the GCTA website (http://cnsgenomics.com/software/gcta). The disease prevalence of 0.0028 was specified, so that GCTA could transform the estimate of variance explained on the observed scale to that on the underlying liability scale. The LADA prevalence of 0.0028 was derived from T2D prevalence. Specifically, the expected proportion of T2D cases that are misdiagnosed LADA cases (8%) was multiplied by prevalence for T2D (0.035) (Maahs et al., 2010; Sinnott et al., 2017).

Heritability estimated by LDAK

LDAK was implemented on the same imputed dataset used in the GCTA analysis. After post-imputation quality control, SNP weights were calculated and kinships for each chromosome were computed then merged at the end (these were the most time-consuming steps post-imputation). To test for cryptic relatedness, a test for inflation was performed(Yang et al., 2011). This is done by splitting the whole genome into quarters: the first quarter of the genome (chromosome 1-3), the second quarter of the genome (chromosome 4-7), the third quarter (chromosome 8-11) and the fourth quarter of the genome (chromosome 12-22). The inflation estimate was derived by summing the heritability estimates for each quarter subtracted by the heritability estimate for the whole genome (inflation = $(\sum_{i=1}^{4} \hat{h}_i^2)$-$\hat{h}_{whole}^2$). Finally, SNP heritability was calculated using the kinship matrix with gender, 20 PCs, and top-associated SNPs included as fixed effects. LDAK5 program was downloaded from the LDAK website (http://dougspeed.com/ldak/).

Heritability estimated by LDSC

Summary statistics for unpublished LADA GWAS meta-analysis, as well as published T1D GWAS (Bradfield et al., 2011) and T2D GWAS (Morris et al., 2012) were leveraged to assess LDSC estimates for the three forms of diabetes. The dataset used for GCTA and LDAK is a smaller subset of the dataset used in this LADA GWAS meta-analysis and therefore LDSC cannot be directly compared to GCTA and LDAK estimates. The number of SNPs in the LADA summary statistics was 7,813,592. The LDSC v.1.0.0 python package was used to calculate heritability estimates for LADA, T1D and T2D, both with and without the major histocompatibility complex (MHC) region. Estimates reported for T2D with the MHC region removed were obtained from the publicly available LD-Hub (Zheng et al., 2017)website (http://ldsc.broadinstitute.org/).

Power analysis

GCTA-GREML power calculator was used to assess whether we have a large enough sample size to detect heritability (http://cnsgenomics.com/shiny/gctaPower). The following assumptions were used: 2,706 cases and 1,254 controls, a population prevalence of 0.0028, variance of the SNP-derived genetic relationships of 2×10^{-5} and type 1 error of 0.05. The prevalence estimate for T2D was 0.0316 in 2012 (Sinnott et al., 2017). Prevalence continues to increase so assuming a T2D prevalence of 0.035 we also assume a 0.0028 prevalence for LADA, which is comparable to the T1D prevalence in 2002 (0.00228). Power was calculated for varying heritability estimates: 0.35 (GCTA T2D heritability estimate from published study), 0.73 (GCTA T1D heritability estimate from

published study), 0.54 (average heritability estimate using T1D and T2D estimates).

4.3 Results

Heritability estimates

LADA heritability estimates calculated by all three methods fell between previous estimates for T1D and T2D (**Figure 4-1**). Heritability estimates for LADA according to the LDSC method were 0.392 and 0.256 (with and without the MHC region, respectively) **(Table 4-1).** Indeed, LDSC estimates for LADA are greater than the T2D heritability estimates of 0.167 and 0.087 (with and without the MHC region, respectively) and less than the T1D heritability estimates of 0.785 and 0.282 (with and without the MHC region, respectively). Similarly, GCTA-estimated heritability for LADA (0.533) was lower than what has been reported for T1D (0.73) but higher than what has been reported for T2D (0.35; **Table 4-2**) (Speed et al., 2012). GCTA estimates for LADA are comparable to LDAK's LADA heritability estimate of 0.517, which is also lower than LDAK heritability estimates for T1D (0.74) and higher than LDAK's heritability estimates for T2D (0.44) (Speed et al., 2012). Population structure can inflate SNP-based heritability estimates, and after testing for inflation due to population structure using LDAK, an inflation estimate of 0.022 suggests that inflation due to population structure is responsible for approximately 4.64% of the observed LDAK heritability estimates. According to the power calculator

provided by GCTA, there was enough power to detect at least a heritability of 0.35 (**Table 4-3**).

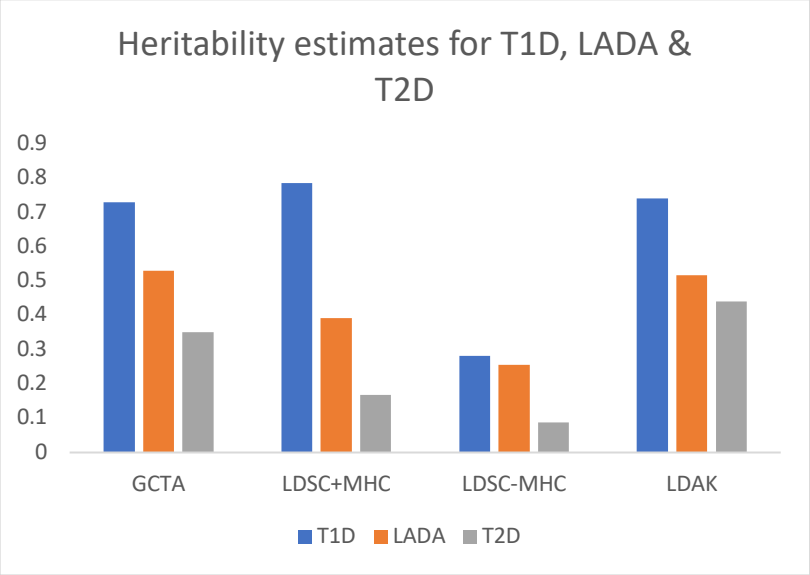


Figure 4-1 Comparison of heritability estimates for T1D (blue), LADA (orange), and T2D (gray) across three tools: GCTA, LDSC, and LDAK. The asterisk (*) represents estimates reported in published data(Speed et al., 2012; Zheng et al., 2017).

	LDSC with MHC	LDSC without MHC
T1D	0.785 (0.546)	0.282 (0.131)
LADA	0.328 (0.096)	0.261 (0.072)
T2D	0.167 (0.017)	0.087 (0.009)*

Table 4-1 Heritability estimates for T1D, LADA, and T2D derived from the LDSC method with and without the MHC. The asterisk (*) represents estimates reported in published data(Zheng et al., 2017). Standard error for each estimate is represented by values within the parentheses.

	GCTA	LDAK
T1D	0.73 (0.06)*	0.74 (0.08)*
LADA	0.533 (0.038)	0.517(0.14)
T2D	0.35 (0.06)*	0.44 (0.08)*

Table 4-2 Heritability estimates for T1D, LADA, and T2D across two tools: GCTA and LDAK. The asterisk (*) represents estimates reported in published data(Speed et al., 2012). Stand deviation for each estimate is represented by values within the parentheses.

Trait heritability (on liability scale, ^{2})	POWER	SE	NCP
0.35	1	0.0389	80.8545
0.54	1	0.0739	192.4668
0.73	1	0.0739	351.7338

Table 4-3 Power calculation for GCTA, using GCTA-GREML to calculate the probability of detecting $^2 > 0$, with 0.05 type 1 error, for LADA cases and controls. Standard error (SE) of the SNP-heritability (^{2}) is shown as well as the non-centrality parameter (NCP) of the chi-squared test statistic. NCP is equal to h^4/SE^2.

4.4 Discussion

Given that LADA appears to be genetically more similar to T1D (**Chapter 3** (Mishra et al., 2017)), yet still share similar clinical features with T2D (adult age of onset and initial insulin independency), heritability estimates are predicted to fall in the range between T1D and T2D estimates. All three methods, GCTA, LDSC, and LDAK, estimated the heritability of LADA to fall between T1D and T2D estimates. In general, unreliable heritability estimates could be due to various reasons including cryptic relatedness, population structure or sample size. According to the power calculator provided by GCTA, there was enough power to detect at least a heritability of 0.35 (based on reports for T2D using GCTA). Additionally, more stringent thresholds during quality control are crucial for reducing cryptic relatedness. After performing stricter QC and including more significant PCs, the inflation due to population structure explained 4.64% of the heritability estimated by LDAK (0.517).

A major limiting factor for GCTA and LDAK is computational efficiency. Imputation and computing SNP weights are the most time-consuming steps of the analysis. In addition to the time-consuming steps introduced by GCTA and LDAK, solving the mixed models via REML is a memory-consuming part of the analysis that we do not encounter when using LDSC (Speed et al., 2017). LDSC is less computationally intensive compared to GCTA and LDAK because the method only relies on GWAS summary level data and assumes a simple polygenic model.

Despite the feasibility and computational efficiency of LDSC, the LDSC method estimates the heritability without the MHC region. However, the MHC region

contributes up to 50% of the genetic risk of T1D. Additionally, the strongest associated signal in LADA and T1D is located in the MHC. The MHC has also been implicated in T2D, which could be a true signal or could be due to the unaccounted presence of autoimmune diabetes (LADA cases) (Scott et al., 2017). Thus, the MHC region should be handled more delicately when applying these heritability calculations to autoimmune diabetes. If this region is completely removed, then careful interpretation is required as SNPs that are highly associated with the disease are no longer accounted for in the analysis. When implementing LDSC with and without the MHC region, we not only see the role the MHC plays in T1D compared to LADA and T2D, but also the MHC region introduces more noise to the estimates for T1D and LADA (standard error: 0.546 and 0.159, respectively). Previous studies have shown that LDSC tends to have large standard deviations, and this is likely due to the difference in parameter estimation. LDSC uses method of moments whereas GCTA and LDAK use maximum likelihood estimators (REML) to estimate parameters (Speed et al., 2017).

However, a proper comparison between LDSC versus GCTA and LDAK cannot be achieved until LDSC is implemented on the same exact dataset used in the GCTA and LDAK analysis. Difference in estimates observed between LDSC and the two alternative methods could be due to the additional samples in the meta-analysis. This study can be furthered developed by performing LDSC on the same dataset utilized by GCTA and LDAK to properly compare estimates between LDSC versus LDAK and GCTA. Additional future steps include repeating LDAK and GCTA analysis using a MAF threshold of 5% during post-imputation quality control to maintain consistency with genotype quality control

and leveraging the Wellcome Trust Case Control Consortium T1D/T2D datasets and compare results to published datasets to ensure proper implementation and analysis of the three methods.

4.5 Conclusion

The SNP heritability estimates for LADA fall between T1D and T2D according to heritability estimates derived by all three methods, thus further supporting the observations in Chapter 3 and pointing to the concept of 'type 1.5 diabetes'. LDSC was the fastest and most straight-forward tool to implement, however exclusion of the MHC results in much lower heritability estimates and more noise, especially when studying autoimmune diseases. Across all three methods, it is not clear how MHC should be handled; thus, heritability estimates should be carefully interpreted.

CHAPTER 5. THE FIRST GENOME-WIDE ASSOCIATION STUDY OF LADA

*The content of this chapter is published in Cousminer DL. *, Ahlqvist E.*, Mishra R. *, Andersen MK.*, Chesi A., Hawa MI., Davis A., Hodge KM., Bradfield JP., Zhou K., Guy VC., Akerlund M., Wod m., Fritsche LG., Vestergaard H., Snyder J., Højlund K., Allan Linneberg, Käräjämäki A., Brandslund I., Kim CE., Witte D., Sørgjerd EP., Brillon DJ.,Pedersen O., Beck-Nielsen H., Grarup N.,5 Pratley RE., Rickels MR., Vella A., Ovalle F., Melander O.,Harris RI., Varvel S., Grill VER., BMDCS, Hakonarson H., Froguel P., Lonsdale JT., Mauricio D., Schloot NC., Khunti K.,Greenbaum CJ, Asvold BO., Yderstræde KB., Pearson ER., Schwartz S., Voight BF., Hansen T., Tuomi T., Boehm BO.*, Groop L.*, Leslie RD.*, Grant SFA*. "First Genome-Wide Association Study of Latent Autoimmune Diabetes in Adults Reveals Novel Insights Linking Immune and Metabolic Diabetes." Diabetes Care (2018) PMID: 3025408. *equal contribution. RM's contribution to the manuscript includes study concept and design, analysis and interpretation of data, drafting, critical revision, and writing sections of the final version of the manuscript. RM had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy for the data analysis.*

5.1 Introduction

The relationship between LADA and both T1D and T2D is not fully elucidated and not appropriately encapsulated in the term 'type 1.5 diabetes' (Ahlqvist et al., 2018; Thomas et al., 2017; Tuomi et al., 2014).

In many populations, LADA is at least as prevalent as childhood-onset T1D (Hawa et al., 2013), but is frequently misdiagnosed as T2D (Laugesen et al., 2015; Tuomi et al., 1999) given its presentation in adults without need for insulin. As such, subjects with LADA could be present in cohort studies for T2D that do not screen out autoantibody-positive cases, potentially resulting in the identification of genetic associations for T2D that are etiologically related to autoimmunity. Furthermore, LADA has a natural history distinct from that of T2D and is likely mismanaged as a result (Laugesen et al., 2015). The challenge to define adult autoimmune diabetes, including LADA, as distinct from the generality of T2D is acute given the increasingly larger datasets assembled to identify additional, common genetic risk factors of increasingly smaller effect sizes. Indeed, reflecting this concern, recent genome-wide association study (GWAS) analyses of T2D have reported associations at T1D-associated regions such as *HLA-DQA1* in populations of European ancestry (Scott et al., 2017) and *HLA-B* and *INS-IGF2* in African ancestry populations (Ng et al., 2014). As such, understanding the genetic etiology of adult autoimmune diabetes will not only aid the characterization of this relatively common form of diabetes, but will also facilitate our understanding of both T1D and T2D.

To date, the relatively limited candidate gene studies carried out for LADA have supported a role for both T1D and T2D risk loci (Andersen et al., 2014; Andersen et al., 2010; Bakhtadze et al., 2008; Desai et al., 2007a;

Hosszúfalusi et al., 2003; Howson et al., 2011; Mishra et al., 2017; Tuomi et al., 2014). Most notable from these previous studies is the implicated role of the key type 2-associated *TCF7L2* locus in the pathogenesis of LADA (Andersen et al., 2014; Bakhtadze et al., 2008; Cervin et al., 2008). More recently, we constructed genetic risk scores combining known T1D and T2D loci and assessed their impact in LADA, and our results implicated a role for both sets of loci (**Chapter 3** (Mishra et al., 2017)). However, no systematic genome-wide appraisal of adult autoimmune diabetes has been performed. Therefore, in this study, we performed the first GWAS of LADA against population controls and further contrasted LADA against T1D and T2D to better understand its genomic signature in comparison to these two better characterized forms of diabetes.

5.2 Methods

Study subjects

Cases diagnosed with LADA were included from cohorts of European ancestry (**Appendix: Supp. Table 5-1**), including 'ActionLada-Plus,' All New Diabetics In Scania (ANDIS), the Botnia Study, Copenhagen LADA (including samples from Danish Centre for strategic Research in Type 2 Diabetes (DD2), Vejle Biobank, Odense University Hospital (OUH), Copenhagen Insulin and Metformin Therapy trial (CIMT), Inter99, and Steno Diabetes Center (SDC)), the Diabetes Registry Vasa (DIREVA), GoDARTS, Nord-Trøndelag Health Study (HUNT), and Scania Diabetes Registry (SDR). Controls were population-based (including samples from the Bone Mineral Density in Childhood Study (BMDCS), Copenhagen

controls (with samples from the 1936 Birth Cohort and ADDITION-PRO), GoDARTS, HUNT, and the Malmö Diet and Cancer study, DIREVA, and SDR). Inclusion and exclusion criteria for LADA, T1D, T2D, and population controls varied by cohort (see **Appendix: Supp. Table 5-1 and Supplemental Note** for details). In general, LADA was defined by an age at diagnosis older than 20, 30 or 35 years, with some cohorts restricting the upper age limit to 70 years; the presence of diabetes-associated autoimmune autoantibodies, in particular GADA-positivity; and the lack of insulin requirement for 6 months or 1 year after diagnosis. In some cases, C-peptide level was also used as a filter.

Genotyping and imputation

Each respective cohort performed genome-wide genotyping on the Illumina Core Exome chip, the Illumina OmniExpressExome BeadChip, or the Affymetrix 6.0 array. Cases and controls from each study center were matched on the same genotyping chip to reduce batch effects. Standard post-genotyping quality control was performed, including sample exclusions for ambiguous gender, call rate < 95%, and any duplicate or related individuals (pi_hat ≥ 0.2), and SNP exclusions for monomorphic SNPs, SNPs with MAF < 0.05, and SNPs with missingness rate > 0.05. The Haplotype Reference Consortium (HRC) imputation service (Michigan imputation server, https://imputationserver.sph.umich.edu/index.html) was utilized to perform imputation for autosomal SNPs.

Genome-wide association and meta-analysis: LADA vs. controls, LADA vs. T1D, and LADA vs. T2D

SNPTEST (Marchini et al., 2007) or EPACTS (http://genome.sph.umich.edu/wiki/EPACTS) was used by each respective cohort to perform case-control GWAS of LADA (n = 2,634) vs. population controls (n = 5,947), LADA (n = 2,454) vs. cases with T1D (n = 968), and LADA (n = 2,779) vs. cases with T2D (n = 10, 396), including sex and principal components as covariates (see **Appendix: Supp. Table 5-1** for cohort-specific covariates). After GWAS, filtering was performed centrally to include only SNPs with a MAF > 0.05, INFO quality score > 0.4, and a Hardy-Weinberg equilibrium $P > 1x10^{-7}$. Meta-analysis was then performed for LADA vs. population controls, LADA vs. T1D, and LADA vs. T2D with GWAMA (Magi & Morris, 2010) with two rounds of genomic control (**Appendix: Supp. Table 5-2; Supp. Figure 5-1 and 5-2**). Signals in the secondary tier ($P = 1 \times 10^{-6} – 5 \times 10^{-8}$) for the LADA vs. population controls analysis were followed up in the GODARTS and HUNT cohorts (LADA, n = 345; controls, n = 1,664) and meta-analyzed with the discovery set (total LADA, n = 2,979; controls, n = 7,611) to assess whether any novel signals would reach genome-wide significance.

Enrichment of directional consistency among T1D/T2D loci in LADA

To estimate whether the concordance in direction of effects for T1D and T2D loci in LADA is significantly different from chance, a binomial test was used assuming a null hypothesis of 50% agreement.

Conditional analysis

Approximate conditional analysis for known T1D-associated loci was carried out for the LADA vs controls summary statistics results for the 10p15.1 locus using

Genome-wide Complex Trait Analysis (GCTA) (Yang et al., 2011). For this locus, LADA vs controls + HUNT summary statistics were conditioned on the following T1D-associated SNPs: rs61839660, rs10795791, rs7090530, rs12251307, rs41295121, and rs11258747(Barrett et al., 2009; Bradfield et al., 2011; Onengut-Gumuscu et al., 2015). For 12q24.3, two of the T1D-associated SNPs (rs3184504 (Barrett et al., 2009) and rs653178 (Onengut-Gumuscu et al., 2015) were in high LD ($r^2 > 0.9$) with our lead SNP, and the MHC, *PTPN22*, and *INS* loci were not conditioned as the top signals were identified as T1D-associated SNPs.

Stratification analysis by GAD autoantibody titer

Cases with LADA are heterogeneous in terms of GAD autoantibody titer (Buzzetti et al., 2007). Therefore, to further understand the genetic landscape of LADA in the context of different GAD levels, we stratified cases into tertiles in ActionLADA and ANDIS, DIREVA, and SDR. We performed three GWAS, on (1) the top tertile with the highest GAD titers (n = 627) vs. population controls (n = 4314); (2) the top two tertiles with the highest GAD titers (n = 1012) vs. population controls (n = 4314); and (3) the bottom tertile with the lowest GAD titers (n = 562) vs. population controls (n = 4314).

LD Score Regression

To test for genetic correlations genome-wide between LADA, T1D (Aylward et al., 2018; Bradfield et al., 2011), and T2D (Mahajan et al., 2018; Morris et al., 2012), we performed LD score (LDSC) regression using the LDSC v.1.0.0 python package (Bulik-Sullivan et al., 2015).

Pathway analysis

DEPICT pathway analysis (Pers et al., 2015) was used to perform gene set enrichment, tissue enrichment, and gene prioritization analyses.

HLA imputation/analysis

The HLA imputation software SNP2HLA (Jia et al., 2013) was used to impute chromosome 6 in ActionLADA-Plus (n = 1,365), Swedish cases with LADA (n = 794), BMDCS (n = 1,056) and T1D cases from the WTCCC (n = 1,990). HLA alleles with 4-digit resolution were imputed. The R package 'BIGDAWG' (https://cran.r-project.org/web/packages/BIGDAWG) (Pappas et al., 2016) was used to test for allele frequency differences for established T1D-associated HLA haplotypes between LADA versus T1D, as well as LADA versus BMDCS. Haplotypes with frequencies less than 1% across LADA, T1D, and BMDCS were removed from the analysis given that rare haplotypes can result in unstable variance estimates and unreliable test statistics.

5.3 Results

Genome-wide association of LADA versus population controls

We first conducted GWAS in patients with LADA (n = 2,634) versus population-based controls (n = 5,947) of European ancestry in a discovery meta-analysis setting (**Appendix: Suppl. Table 5-1;** power calculations can be found in **Supp. Table 5-3**). Four signals achieved genome-wide significance ($P < 5x10^{-8}$), all at established T1D risk loci (*HLA*, *PTPN22*, *INS*, and *SH2B3*; **Table 5-1, Appendix:**

Supp. Figures 5-1 and 5-2). Pathway analysis with DEPICT (Pers et al., 2015) for signals at $P < 10^{-5}$ supported a strong immune role in the pathogenesis of LADA (**Appendix: Supp. Tables 5-4 and 5-5**), with gene set enrichment analysis implicating 'abnormal cytotoxic T cell physiology' (nominal $P = 6.39 \times 10^{-7}$) as well as the 'mTOR subnetwork' ($P = 6.03 \times 10^{-5}$) and 'cell cycle' ($P = 1.67 \times 10^{-5}$), and immune system tissue types, including 'natural killer cells' and 'T lymphocytes' (nominal $P = 0.0079$ and 0.0082, respectively). This is consistent with previous reports of these cell types playing a role in the pathogenesis of T1D and LADA (Radenkovic et al., 2016; Wang et al., 2015).

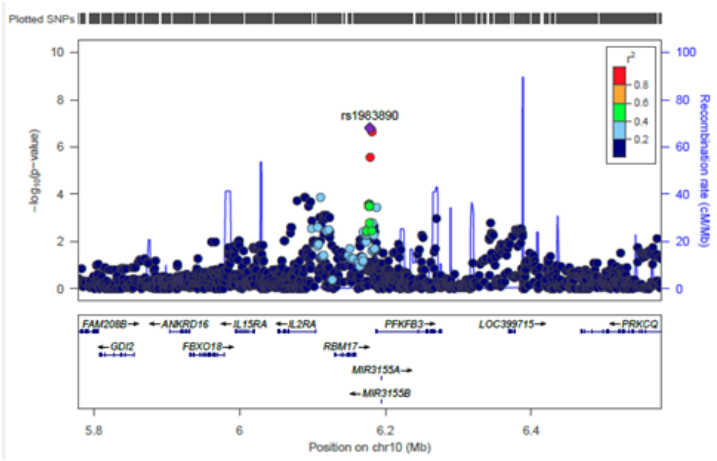
SNP	Chr	Position (b37)	Ref/other allele	Effect allele freq (cases/ctrls)	OR	95% CI	*P*	Gene
LADA (n =2,634) vs. population controls (n = 5,947)								
rs9273368	6	32626475	A/G	0.50/0.28	3.115	2.855-3.398	7.87x10-143	*HLA-DQB1*
rs2476601	1	114377568	A/G	0.159/0.102	1.717	1.539-1.915	7.21x10-22	*PTPN22*
rs689	11	2182224	T/A	0.802/0.726	1.483	1.363-1.613	1.07x10-19	*INS*
rs7310615	12	111865049	C/G	0.553/0.492	1.284	1.193-1.383	4.92x10-11	*SH2B3*
LADA (n = 2,779) vs. T2D cases (n = 10,396)								
rs9273368	6	32626475	A/G	0.43/0.301	2.439	2.222-2.676	3.17x10-78	*HLA-DQB1*
rs689	11	2182224	T/A	0.783/0.715	1.473	1.352-1.605	9.86x10-19	*INS*
rs2476601	1	114377568	A/G	0.173/0.140	1.529	1.38-1.693	4.52x10-16	*PTPN22*
rs3184504	12	111884608	C/T	0.544/0.52	1.24	1.151-1.336	1.77x10-08	*SH2B3*
LADA (n = 2,454) vs. T1D cases (n = 968)								
rs9273368	6	32626475	A/G	0.415/0.65	0.335	0.256-0.385	8.46x10-40	*HLA-DQB1*

Table 5-1 Genome-wide significant signals associated with LADA. We performed three genome-wide association approaches, first for LADA versus population controls (top panel), then for LADA versus T1D (T1D, middle panel) and LADA versus T2D (T2D, lower panel). Odds ratios (ORs) are given for the LADA risk allele except for rs92773368 in LADA vs. T1D, to illustrate that the T1D risk allele was depleted in LADA.

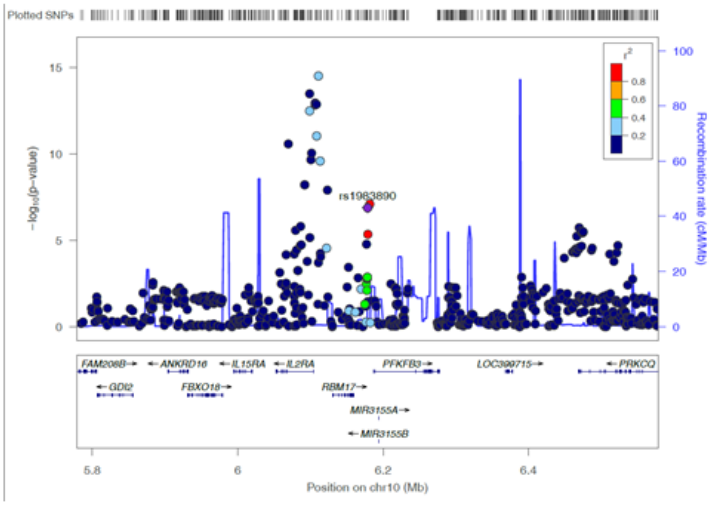
Replication supports a novel locus at *PFKFB3*

Using cases with LADA and population samples from an additional two study centers, we attempted validation of 13 signals with suggestive association ($P < 5x10^{-5}$) (**Appendix: Supp. Table 5-6**). We observed a novel signal at 10p15.1 between the two established T1D loci at *IL2RA* and *PRKCQ*, which achieved genome-wide significance (rs1983890-C, OR (95% CI) = 1.16 (1.14-1.32), $P = 3.02 \times 10^{-8}$) (**Figure 5-1A and 5-1B**). Given that the LADA signal is situated in close proximity to known T1D risk loci and was in moderate to low LD with established T1D-associated alleles (**Appendix: Supp. Table 5-7**), we conditioned on the T1D SNPs and observed that rs1983890 remained strongly associated with LADA (OR (95% CI) = 1.15 (1.13-1.19), $P = 4.35 \times 10^{-8}$) (**Figure 5-1C**). This signal reached suggestive association in a study of T1D ($P = 1.3 \times 10^{-7}$) (Bradfield et al., 2011) and as such may not represent a unique LADA association. DEPICT gene prioritization analysis identified the gene encoding '6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 3' (*PFKFB3*), the nearest gene to the LADA signal, as the most likely functional candidate (**Appendix: Supp. Table 5-8**).

A. LADA



B. T1D



C. LADA conditional

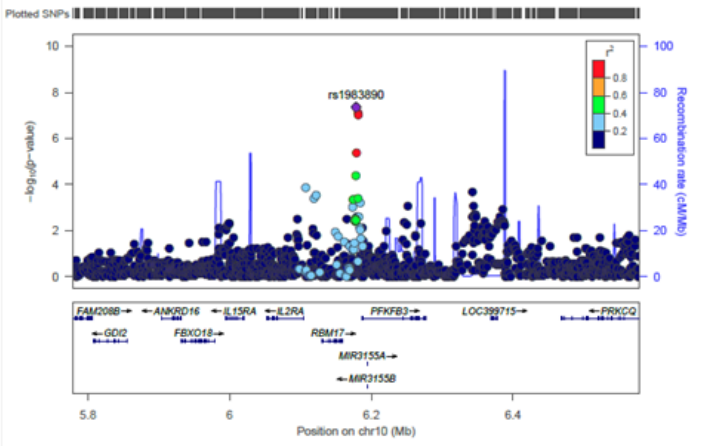


Figure 5-1 LocusZoom plots for the *PFKFB3* locus. (A) In LADA vs. population controls with the addition of replication samples, rs1983890 reached borderline genome-

wide significance. (B) This signal lies in between two T1D-associated loci at 10p15.1(Bradfield et al., 2011). (C) When we conditioned on the two known T1D loci, the signal in LADA remained. LocusZoom plots were constructed to show the association data of SNPs 400kb upstream and downstream of the lead LADA-associated signal at rs1983890.

Candidate loci for T1D and T2D

Some of the loci that were suggestively associated with LADA in this study overlapped previously documented T1D associations, including rs11755527 (*BACH2*) and rs941576 (*DLK1*) (Barrett et al., 2009; Bradfield et al., 2011; Onengut-Gumuscu et al., 2015), and the T2D association at rs11888640 (*THADA*). Taking a candidate gene approach, we extracted 66 established T1D-associated loci from the LADA vs. population controls meta-analysis and found that 17 of these yielded association with LADA after multiple-test correction ($P < 7.6 \times 10^{-4}$, **Appendix: Supp. Table 5-9**). Taking a similar approach with 65 established T2D loci, none surpassed the significance threshold; however, at the nominal significance level ($P < 0.05$), eleven T1D and eleven T2D variants were associated with LADA, all having the same direction of effect as seen for T1D and T2D, respectively, except for the T2D locus *CILP2* (rs10401969-T, OR = 0.820 (0.726-0.927), $P = 0.0016$; **Appendix: Supp. Table 5-10**). On the whole, both T1D and T2D loci had lower *P*-values in LADA than expected by chance (**Appendix: Supp. Figure 5-3**). Approximately 90.6% T1D loci **(Appendix: Supp. Table 5-9)** had directional consistency in LADA **(*P*-value = 4.51×10^{-12})** and 72.3% of T2D loci **(Appendix: Supp. Table 5-10)** had directional consistency in LADA (*P*-value = 2.10×10^{-4}). Combining T1D and T2D loci, 81.4% had directional consistency in LADA (*P*-value = 1.40×10^{-13}). Therefore, we observed a significant enrichment of established T1D and T2D loci having the same directional effect in LADA.

GWAS of LADA versus T2D and T1D

Next, we compared LADA with T2D at the genome-wide level. Similar to the results of LADA vs. population controls, LADA (n = 2,454) vs. T2D (n = 10,396) yielded genome-wide significance for the same four T1D risk loci (**Table 5-1**). We then performed a GWAS of LADA (n = 2,454) vs. T1D (n = 968) to assess whether any differences could be detected. Only the HLA region was significantly different between T1D and LADA, representing a relative depletion of the lead signal in LADA when compared to T1D (rs9273368-A, OR (95% CI) = 0.335 (0.256-0.385), $P = 8.46 \times 10^{-40}$; **Table 5-1**). Leveraging the entire genome-wide summary statistics, genetic correlation analyses showed that LADA was positively correlated with both T1D (with the inclusion of the HLA; r_g (SE) = 0.385 (0.136), $P = 0.0047$) and T2D (without the HLA; r_g (SE) = 0.281 (0.106), $P = 0.008$). Genetic correlations were re-calculated using a more recent T1D GWAS (Aylward et al., 2018) and T2D GWAS (Mahajan et al., 2018), which also showed positive correlation with both T1D (without the HLA; r_g (SE) = 0.475 (0.139), $P = 0.0006$) and T2D (without the HLA; r_g (SE) = 0.371 (0.071), $P = 1.84 \times 10^{-7}$).

Stratified GWAS of LADA by GAD autoantibody tertile

Stratifying LADA cases into tertiles resulted in the detection of the same four loci, although the magnitude of the associations differed between the top tertile vs. population controls, the top 2 tertiles vs. population controls, and the bottom tertile vs. population controls (**Appendix: Supp. 5-11**). As expected, the ORs for the leading loci were strongest in the LADA cases with the highest GAD

autoantibody titers. For example, rs9273368 (*HLA-DQB1*) showed the strongest association with LADA in the analysis including the top tertile of GAD autoantibody titer (OR (95% CI) = 3.30 (2.81-3.88), $P = 1.89 \times 10^{-47}$) and the lowest association in the bottom GAD autoantibody tertile (OR (95% CI) = 2.42 (2.06-2.85), $P = 2.13 \times 10^{-26}$). Furthermore, only the *HLA-DQB1* locus was significantly associated in the LADA cases with the lowest GAD titers, while the *PTPN22*, *INS,* and *SH2B3* loci were only evident among cases with higher GAD titers. Furthermore, rs7903146 at *TCF7L2* had a slightly higher OR in the group with the lowest GAD titer than that with the highest GAD titer (1.09 vs. 1.05, respectively).

HLA haplotype analysis

To further investigate differences in the HLA region between LADA and T1D, we imputed this region using SNP2HLA in 2,159 LADA cases from the ActionLADA + CHOP + Swedish cohorts and 1,990 patients with T1D (WTCCC) and compared the frequencies of the leading T1D-associated HLA haplotypes (**Appendix: Supp. Table 5-12**). After removing haplotypes with less than 1% frequency, fifteen known T1D-associated HLA haplotypes were tested for association in LADA compared to T1D. Eleven T1D haplotypes were significantly different in frequency between LADA and T1D cases after correction for multiple testing ($P < 0.003$), with all but four being protective against T1D. The four T1D susceptibility haplotypes, HLA-DRB1*0301-DQA1*0501-DQB1*0201, HLA-DRB1*0401-DQA1*0301-DQB1*0302, HLA-DRB1*0404-DQA1*0301-DQB1*0302, and

HLA-DRB1*0405-DQA1*0301-DQB1*0302, had significantly lower frequencies in LADA than in T1D.

5.4 Discussion

Taken collectively, GWAS and HLA haplotype analyses based on established associations, along with gene set enrichment analyses, support the hypothesis that the strongest genetic risk loci for LADA are shared with T1D, but that established T2D alleles also play a weaker role, as evidenced by the enrichment of established T2D loci in LADA and the positive genetic correlation between LADA and T2D. The strong T1D-like signature seen here in adult autoimmune diabetes could be explained by the differing genetic architectures between the two main types of diabetes (Timpson et al., 2018), with T1D having multiple low-frequency risk variants with high odds ratios while T2D has many common risk variants with smaller effect sizes. Given these architectural differences, any trait with a T1D-like genetic component will detect T1D signals first, and would only subsequently detect the T2D signals with increased statistical power (**Appendix: Supp. Table 5-3**).

Furthermore, this has important implications for genetic studies of T2DX, in which misdiagnosed autoimmune diabetes cases are not routinely screened out. With increasing sample sizes and the ability to detect additional loci, T2D GWAS that are ‘contaminated’ with adult autoimmune cases will inevitably begin to detect T1D-associated genetic loci, potentially mis-assigning these loci to T2D etiology.

In comparing LADA to the general population, we identified a novel independent genome-wide significant signal at the *PFKFB3* locus that persisted after conditioning on the two nearby T1D-associated signals on chromosome 10p15. Cumulative evidence for the 10p15 locus suggests it is a complex region associated with autoimmune diabetes, given that it already harbors two established risk alleles for T1D as well as our signal for LADA. Previous studies strongly support *PFKFB3* as a plausible biological candidate in diabetes, given its gene product's role as a regulator of glycolysis and insulin signaling (Duran et al., 2009). In mice, a pair of complementary studies showed that disrupted *PFKFB3* in adipose tissue exacerbated insulin resistance and adipose tissue inflammation (Huo et al., 2010), while overexpression of the gene was protective (Huo et al., 2012). Furthermore, *PFKFB3* plays a role in autoimmune diseases; in T cells from rheumatoid arthritis patients, *PFKFB3* is lost leading to decreased T cell glucose consumption and impaired autophagy, which in turn lead to an inability to mount a normal immune response and an increase in T cell apoptosis (Weyand et al., 2013). Further studies are thus warranted to investigate the role of *PFKFB3* in LADA, and to determine whether this signal is truly a distinguishing feature between adult and childhood-onset autoimmune diabetes.

Although the lead genome-wide significant loci are shared with those T1D risk, they clearly have a diminished impact in LADA. To further investigate the differences between LADA and T1D at the HLA region, we performed a comparative haplotype analysis that showed a decreased frequency of T1D-associated risk haplotypes in LADA. This could be partly explained by the established age gradient in HLA frequencies seen in T1D patients; however, HLA risk genotype frequencies have also been shown to differ between LADA patients and T1D patients with age at onset >35 years (Andersen et al., 2010; Luo et al.,

2016). Future in-depth studies of the differences in HLA risk haplotypes between T1D and LADA taking age and ethnicity into account are also warranted.

In terms of T2D-associated loci, our results differ from previous candidate studies. For instance, our previously reported *HNF1A* locus (**Chapter 3** (Mishra et al., 2017)) was not observed in this setting. Furthermore, while previous studies showed an association for the leading T2D risk locus at *TCF7L2* with LADA (Andersen et al., 2014; Cervin et al., 2008), our data shows relatively limited support of this finding (**Appendix: Supp. Table 5-10**) (LADA vs. population controls, rs7903146-T: OR (95% CI) = 1.107 (1.024-1.20), P = 0.011), which may be due to the limited power of our study to detect T2D signals (**Appendix: Supp. Table 5-2**). To understand the evidence supporting the previous association, we examined the allele frequencies of the lead variant in each contributing cohort. This revealed that the difference in risk allele frequency between cases and controls was cohort-specific, with only one case-control set (ActionLADA cases vs BMDCS controls) not supporting this association, principally due to the higher frequency of the risk allele in the control set (**Appendix: Supp. Table 5-13**). One possibility is that inclusion or exclusion of cases with T2D from control cohorts would affect the frequency of the risk allele; however, sensitivity analysis with control sets that either excluded or included cases with diabetes in Swedish and Danish samples showed the persistence of an association (**Appendix: Supp. 5-13**), although not at the genome-wide significance level. Interestingly, a recent study found that the T2D risk allele at the key *TCF7L2* locus was associated with T1D cases who were older than 12 years at onset and were positive for only a single autoimmune antibody (Redondo et al., 2017). That study provides further evidence for a role for T2D

genetic risk in later-onset autoimmune diabetes and resonates with the genome-wide observations we report here in adults.

The precise diagnostic criteria used to distinguish LADA from adult-onset T1D and T2D remain under debate. These differences in opinion have hindered the collection of well-phenotyped, clearly defined LADA cohorts for genetic studies, and are reflected in the cohorts we included in this study, e.g. in terms of heterogeneous age inclusion thresholds and differences in autoantibody testing. In this study, we strove to be inclusive to maximize our sample size and statistical power, but we acknowledge that stringent, deeply phenotyped cohorts are needed to truly address where adult autoimmune diabetes is placed on the diabetes spectrum. Another debate surrounds the idea that LADA cohorts may simply be collections of poorly phenotyped cases with adult-onset T1D and T2D, and refute the idea that LADA is a unique disease entity. However, GAD assays have a specificity of 95–98%, so by implication, some cases with T2D with low-level GAD can be incorrectly classified as LADA cases; these would, however, represent only a very small fraction of cases since the predictive specificity of GAD would have been increased by our cohort enrichment as with any biomarker assay. Conversely, the small percentage of cases with LADA who do not have GAD positivity but have other islet autoantibodies and are misclassified as having T2D, could affect the estimate of genetic correlation between LADA and T2D to a small degree. Future studies should focus on defining the heterogeneity and misdiagnosis rates among patients with LADA.

Despite these limitations, using the definition of LADA presented here, we identified factors which potentially distinguish this form of adult autoimmune diabetes from childhood-onset T1D as well as T2D: (1) a novel signal at the *PFKFB3* locus, and (2) attenuation of T1D-associated HLA risk haplotypes.

Overall, we find the presence of both a T1D-like autoimmune genetic component and a T2D-like metabolic/anthropometric genetic component consistent with the phenotypic features of both main diabetes types, suggesting that LADA as defined here is a hybrid of these two major diseases. Our findings promote the hypothesis that the polygenic component that contributes susceptibility to T2D can act as a modifier to T1D risk, possibly as a 'second hit' in individuals who have moderate underlying autoimmune susceptibility that is insufficient to trigger childhood T1D but greater than that of the general population and sufficient to lead to clinical diabetes in adulthood. Taken together, future studies should examine the role of body mass index, which is lower in T1D and higher among patients with T2D, in adult autoimmune diabetes, as well as further defining the role of factors that potentially distinguish adult autoimmune diabetes from T1D and T2D.

5.5 Conclusion

In this first GWAS of LADA, we show that the leading genome-wide significant signals point towards LADA as being a late-onset form of T1D, albeit with a genetically attenuated potency of key T1D-associated HLA haplotypes, but also with a T2D-like genetic component. Further in-depth studies are necessary to address how LADA and insulin dependence develops and to study the impact of heterogeneity among cases with LADA, as well as a need for functional studies to investigate how the glycolytic regulator *PFKFB3* is situated at the intersection of autoimmune and metabolic diabetes. Furthermore, our LADA dataset should act as a resource to help mitigate the unaccounted presence of autoimmune diabetes in patients masquerading as T2D, with implications both for GWAS and for clinical management.

CHAPTER 6. GENETIC DISCRIMINATION BETWEEN LADA AND CHILDHOOD-ONSET TYPE 1 DIABETES WITHIN THE MHC

6.1 Introduction

'Latent autoimmune diabetes in adults' (LADA) is characterized by initial insulin independency for at least six months after diagnosis and the presence of diabetes associated autoantibodies (Tuomi et al., 1999). Despite such features, autoantibody screening is not routinely carried out in routine clinical practice, resulting in frequent misdiagnoses. For instance, in a cohort of apparent T2D cases, as many as 8-10% can actually represent misdiagnosed autoimmune diabetes cases (Hawa et al., 2013; Mishra et al., 2018). Hence, there is a need to identify biomarkers to aid in accurately diagnosing LADA as well as other diabetes subtypes (Ahlqvist et al., 2018).

The genetic etiology of LADA was, until recently, largely unknown (Cousminer et al., 2018). Previous genetic studies have suggested the condition comprised both T1D and T2D components either because it is an intermediate form of diabetes or because it is a mixture of T2D in a predominantly T1D cohort owing to a high false positive detection rate using autoantibodies when screening. Since, LADA is currently defined as a slowly progressive form of T1D (American Diabetes Association/ADA, 2018), it is crucial to define genetic differences between childhood-onset T1D and LADA if we are to clarify the clinical utility of identifying adult-onset autoimmune diabetes.

Previous genetic studies in LADA have shown a strong association signal in the major histocompatibility complex (MHC), although with diminished effect sizes

compared to observations in childhood onset T1D (Cousminer et al., 2018; Mishra et al., 2017). The MHC region is located on chromosome 6 and harbors over 400 genes, with two main classes, MHC Class I and MHC Class II, which together harbor classic human leukocyte antigen (HLA) genes (*HLA-A, HLA-B, HLA-C and HLA-DRB, HLA-DQA, HLA-DQB, HLA-DPA, and HLA-DPB*). The *HLA* encodes cell surface proteins for antigen presentation and accounts for approximately 50% of the genetic heritability of T1D, with susceptibility principally harbored within the MHC Class II genes *HLA-DQB1* and *HLA-DRB1*. However, in addition to Class II genes, previous studies have also pinpointed MHC Class I genes in susceptibility to T1D (Howson et al., 2009; Noble et al., 2002; Valdes et al., 2005); in particular, variation within the MHC class I genes *HLA-A* and *HLA-B* variation has been shown through conditional analysis to further increase T1D risk (Nejentsev et al., 2007). MHC Class I markers have also been shown to be associated with younger age-at-diagnosis in T1D, and given the adult-onset phenotype of LADA, we hypothesized that this genetic variation will be less enriched in LADA.

In this effort, we first attempted to recapitulate the independent effects of MHC Class I variants using the SNP2HLA imputation tool followed by stepwise forward logistic regression in the same T1D cohort as the previous study (Nejentsev et al., 2007). Subsequently, we set out to identify distinguishing features within the MHC between childhood-onset T1D from adult-onset LADA, by performing the same conditional analysis followed by a replication attempt in a second case/control set.

6.2 Materials & Methods

Study populations

(I) LADA cases: 1,492 LADA cases were derived from multiple cohorts across the United Kingdom, Germany and the United States. Details on the participants can be found in **Appendix: Supp. Table 6-1**. All participants were diagnosed with LADA if they fulfilled the following criteria: aged 30-70 years old, tested positive for diabetes-associated Glutamic Acid Decarboxylase autoantibodies (GADA) and were not on insulin treatment for at least 6 months after diagnosis. **(II) Controls:** The LADA population-based controls comprised of two cohorts (n=2,979). The first cohort consisted of 1,296 non-diabetic children and adolescents of European ancestry, aged 5-20 years, enrolled in the Bone Mineral Density in Childhood Study (BMDCS (Kalkwarf et al., 2007)). The second control cohort consisted of 1,683 adults of European ancestry from a Non-Hodgkin lymphoma GWAS available in dbGaP (www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000818.v2.p1) (Berndt et al., 2013). Details on the control cohorts can be found in **Appendix: Supp. Table 6-1**. **(III) Recapitulating a previous study:** We also leveraged 3,000 healthy adult British Birth Cohort controls, 2,000 individuals with childhood-onset T1D, and 1,999 individuals with T2D from the Wellcome Trust Case Control Consortium (WTCCC) (Wellcome Trust Case Control Consortium, 2007) to recapitulate observations found in a previous study (Nejentsev et al., 2007). Individual data from the WTCCC is available through the Consortium's Data Access Committee (http://www.wtccc.org.uk). More details on cohort information can be found in **Appendix: Supp. Table 5-1**. **(IV) Replication:** A cohort of individuals from Sweden were used for further recapitulation and replication, including case

subjects with T1D (N = 656), LADA cases (n=823) and population-based controls (N=3,218). Details on the participants can be found in **Appendix: Supp. Table 6-1**. See flow chart for overview of datasets and workflow (**Figure 6-1**).

Genotyping

All samples, except the WTCCC data, were genotyped using the Illumina OmniExpress genotyping chip. WTCCC T1D and T2D cases were genotyped using Affymetrix 500K and WTCCC controls were genotyped on the Illumina 1.2M BeadChip. Quality control was performed using PLINK. Individuals with ambiguous sex, genotype missingness >5%, genome-wide heterozygosity (3 standard deviations from the mean), duplicates and related-individuals were excluded (See **Appendix: Supp. Table 6-1** for details). Principal component (PC) analysis was performed using PLINK, and outliers were removed to exclude individuals with non-European ancestry. Single nucleotide polymorphisms (SNPs) with missing rate <5%, minor allele frequency (MAF) <1% and Hardy-Weinberg equilibrium exact test P-value below $1x10^{-5}$ were removed before HLA imputation.

HLA imputation

Starting from the genotyped SNPs, we imputed chromosome 6 using the HLA imputation software SNP2HLA along with the Type 1 Diabetes Genetics Consortium (T1DGC) reference panel (Jia et al., 2013). A marker window size of 1,000 bp and a posterior probability (gprob) threshold of 0.5 were used. The HLA alleles of LADA cases (n = 1,428) and WTCCC T1D cases (n = 1,985) were imputed to both 2-digit resolution and 4-digit resolution for increased coverage

and resolution of HLA alleles. In total, there were 5,698 SNPs, 424 HLA alleles and 1,276 HLA amino acids. In this study, we focused on a subset of SNPs and HLA alleles which had a MAF greater than 1% in all three control cohorts (159 HLA alleles and 5,506 SNPs remained).

Power calculations

Power calculations were performed using the Genetic Association Study (GAS) Power Calculator (http://csg.sph.umich.edu/abecasis/cats/). Assumptions included a multiplicative model, a disease incidence of 0.0036, 1,428 cases and 2,979 controls and a significance level of 8.83×10^{-6} , based on a Bonferroni correction for the 5665 variants tested (**Appendix: Supp. Table 6-2**).

Recapitulation of a previously published conditional analysis for T1D

Logistic regression using SNPTEST (Marchini et al., 2007) was used to test all HLA alleles and SNPs with MAF >1% in all three control cohorts. Sex and the 12 broad geographical regions, provided by the WTCCC, were included as covariates in the analysis. The analyses were performed in the WTCCC T1D vs control datasets using forward stepwise conditional logistic regression until there were no significant signals remaining after correction for multiple testing

Conditional analysis in LADA vs population-based controls

Conditional logisitic regression was performed using SNPTEST in the LADA vs population-based controls, including sex and the first 4 principle components as covariates.

Replication

To further validate MHC Class I independent effects in T1D and lack of MHC Class I independent effects in LADA, we implemented approximate conditional analyses (COJO) in GCTA (Yang et al., 2012) on summary statistics from the Swedish replication cohort. Association analysis was performed using SNPTEST, and sex and the first four PCs were used as covariates. There were 656 cases with T1D vs 3,218 population-based controls, and 823 cases with LADA vs 3,211 population-based controls.

Sensitivity analysis

We performed sensitivity analysis to determine whether the lack of independent T1D-associated signals in MHC Class I genes in LADA cases could be due to a lack of power. We randomly sampled 1,428 T1D cases and 714 T1D cases (subsets equating to the same size as the LADA cohort and half the size of the LADA cohort, respectively) and 2,219 controls to determine whether the T1D-associated signals could be still be detected. Stepwise conditional logisitic regression using SNPTEST was performed as above. Given the hypothesis that LADA is potentially simply a mixture of T1D and T2D cases, we performed a further constrained conditional analysis in 714 randomly sampled T1D cases and and 714 randomly sampled T2D cases (total n = 1,428 cases) and 2,219 WTCCC controls.

Futher validating independent signals

PLINK was used to calculate pair-wise linkage disequilibrium between variants to further validate that the associated variants were truly independent of each other. To confirm independent association of *HLA-B*39*, the specific *HLA-B*39* subtype *HLA-B*3906* was tested in the WTCCC T1D cases (n = 1985) vs controls (n = 2219) dataset using the presence of *DQB1*0402* and *DQB1*0501* as covariates.

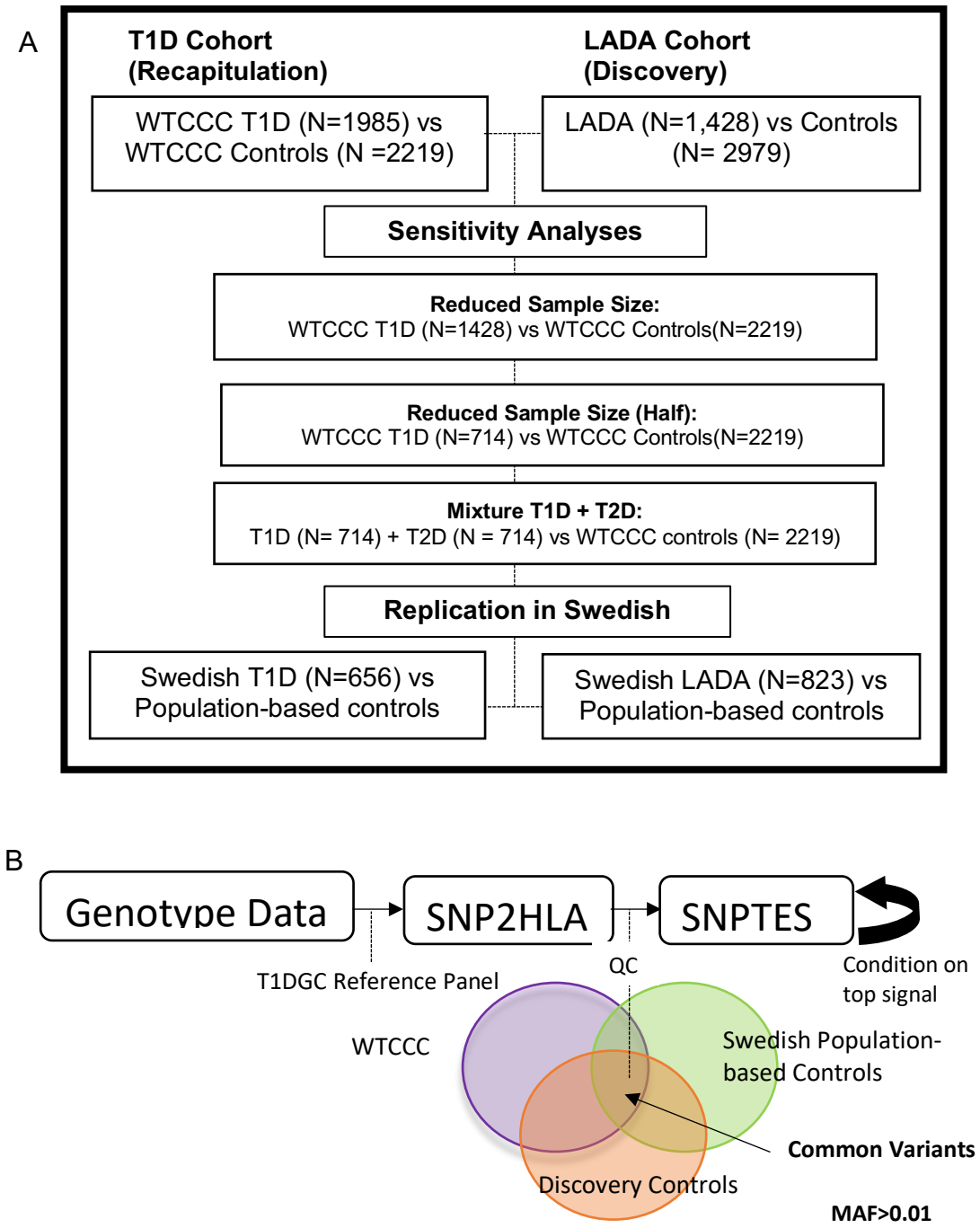


Figure 6-1 Flow Chart of (A) datasets and (B) overall analysis

6.3 Results

Confirming independent effects of MHC Class I signals in WTCCC T1D vs Controls

Although previous studies have reported MHC Class I independent effects in T1D, those studies used directly HLA-typed cases and controls. Given the cost and challenges of direct HLA typing, we elected to utilize the imputation tool SNP2HLA on genotyping data. Furthermore, given this approach differs from the Nejensev *et al.*, it was crucial to first ensure that we could recapitulate the previously reported T1D observations in the same cohort as previously leveraged. Before conditioning, we observed rs3957146 as the strongest association signal in the T1D vs WTCCC controls analysis ($P = 8.94x10^{-165}$; **Figure 6-2A**). rs3957146 is in strong LD with a classical HLA subtype allele, *HLA-DQB1*0302* ($r^2 = 0.99$). After conditioning on the top signal, rs3957146, and subsequent independent MHC Class II signals (*HLA-DQB1*0201* and rs9268633), we observed the reported independent significant association of MHC Class I variants rs1610649 (*HLA-G*, $P = 6.89x10^{-23}$) and *HLA-B*39* ($P = 1.06x10^{-15}$) (**Figure 6-2B**; **Table 6-1**). Conditioning on these variants in addition to the MHC class II variants, also demonstrated significant association with the *HLA-A* locus (rs94/16/19259852, $P=2.04x10^{-8}$). Overall, our results were consistent with the observations seen in Nejensev, *et al.* and subsequent studies (Howson et al., 2009; Nejentsev et al., 2007).

Conditional analysis in LADA vs population-based controls

We then went on to perform stepwise conditional analysis in 1,428 LADA cases and 2,979 controls. Similar to observations in the T1D vs WTCCC controls

dataset, before conditioning on any variants, the strongest association signal in LADA vs population-based controls was also rs3957146 (P = 1.80x10^{-68}; **Figure 6-3A**). Although we had power to detect *HLA-B*39* (**Appendix: Supp. Table 6-2**), when conditioning on the most highly significant MHC Class II alleles (rs3957146, *HLA-DRB1*03,* rs9269081, *DRB1*0404* and *DQB1*0602*), there were no remaining independent signals in the MHC Class I region reaching significance after correction for multiple comparisons (P < 8.83x10^{-6};**Figure 6-3B**). Furthermore, we also noted independent effects in the MHC Class III region (rs2143462, P = 8.24x10^{-8}) and the MHC Class II region (*HLA-DPA1*02*, P = 1.62x10^{-6} and *HLA-DPB1* variant rs3130192, P = 5.32x10^{-6}). Here, *HLA-DPB1* variant is in strong LD with rs2301225 (r^2=0.85) and is independently associated with T1D. MHC Class I variants were not observed to be independently associated with LADA after correcting for multiple comparison (**Table 6-2**).

Sensitivity analysis in reduced sample of T1D vs controls

To ensure that power issues in the relatively smaller LADA sample size, when compared to the T1D sample, did not explain the lack of significant associations in MHC Class I genes, we conducted a sensitivity analysis by systematically decreasing the sample size of the T1D vs WTCCC control cohort to match the size of the discovery LADA cohort (n= 1428 T1D cases and 2219 controls) and performing conditional analysis. Independent significant association signals at *HLA-G* (P = 1.37 x10^{-17}), *HLA-B* (P = 5.58x10^{-14}) and *MUC22* (rs9262545, P = 3.36x10^{-9}; rs9262547, P = 9.69x10^{-14}) were still observed in this reduced T1D sample size (**Table 6-3**), although these signals were missing in the comparatively-sized LADA vs controls dataset. Similarily, independent significant

association signals at *HLA-B* (P = 1.26x10^{-10}), *HLA-G* (P = 0.002), and *MUC22* (rs9262545, P = 1.57x10^{-5}) remained after further reducing the T1D cohort size to equate with half the LADA cohort size (**Table 6-4**).

Sensitivity analysis in an artificially mixed cohort of T1D and T2D cases vs controls

Another explanation for the lack of independent, significant associations across MHC Class I genes in LADA could be due to the possibility that the LADA cohort simply represents an approximalty 50/50 mixture of misdiagnosed T1D and T2D cases. Therefore, we randomly sampled 714 cases with T1D, 714 cases with T2D and 2,219 controls, creating a "mixture" cohort. We performed the same conditional analysis described above and observed that the *HLA-B*, *HLA-G* and two *MUCC2* signals in the MHC Class I regions remained independently significant in this mixed cohort, driven by the T1D case subset (**Table 6-5**).

Replication

We leveraged summary statistics data from Swedish cohorts to attempt replication of our findings. In T1D vs controls, the strongest assocation was rs9275206 (P = 6.35X10^{-89}), which is in strong LD with *HLA-DQB1*0302 (r^2* = 0.99*).* After conditioning on rs9275206 and subsequent top signals (**Table 6-6**), we again observed significant association signals at the *HLA-G* (P = 1.74 x10^{-10}) and *HLA-B* (1.10x10^{-9}) loci . However, as before, when conditional analysis was performed in LADA vs controls, there were no such signals across MHC Class I genes, and very sparse signals in the MHC Class II region (**Table 6-7**).

*Futher validating HLA*B*39*

It has been shown that the *HLA-B*3906* allele is associated with a high risk of diabetes only for specific HLA-DR/DQ haplotypes, *DRB1*0801-DQB1*0402* and *DRB1*0101-DQB1*0501* (Baschal et al., 2011). When specifically conditioning on these HLA-DR/DQ haplotypes in the WTCCC T1D and control cohort, the independent significant association of the more specific *HLA-B*3906* subtype still remained (OR (95% CI) = 4.57 (3.08-6.80); $P = 5.84 \times 10^{-14}$). Furthermore, we observed a significant association signal at the *NOTCH4* (rs397081, $P = 1.11 \times 10^{-10}$) locus, the *MUC22* locus (rs9262545, $P = 7.83 \times 10^{-11}$ and rs9262547, $P = 7.17 \times 10^{-17}$) and the *HLA-A* locus (rs9259852, $P = 5.84 \times 10^{-14}$) (**Figure 6-2C**). Notably, rs9259852 is in strong LD with the classic HLA subtype allele, HLA-A*32 ($r^2=0.96$).

6.4 Discussion

The main objective of this study was to perform conditional analysis of the HLA region in LADA, which has been under explored to date in this disease context. The few genetic studies in LADA (Cousminer et al., 2018; Desai et al., 2007; Luo et al., 2016) only focused on the HLA Class II *DRB1* and *DQB1* haplotypes. Such studies, in populations of both European and Chinese ancestry show that T1D risk haplotypes are less frequent in LADA compared to childhood-onset T1D cases, whereas T1D protective haplotypes are more frequent in LADA, suggesting that LADA is a gentically attentuated form of T1D. By extending the analysis of HLA in LADA beyond MHC Class II region, we were able to observe further genetic differences between LADA and childhood-onset T1D.

First, we leveraged the WTCCC T1D and control dataset, as a positive control , with previous studies identifying MHC Class I independent T1D associations in the MHC Class I region (Eike et al., 2009; Howson et al., 2009; Nejentsev et al., 2007; Noble et al., 2002). Since these studies were reported, imputation tools have allowed the analysis of the HLA region more cheaply and, in general, more practically. Before investigating MHC Class II independent LADA assocations in the MHC Class I region, given the difference in our analytical appraoch, we recapitulated the observations in previous studies (Howson et al., 2009; Nejentsev et al., 2007), by leveraging the same WTCCC T1D and control datasets. We confirmed that MHC Class I variants are significantly associated with T1D, independent of MHC Class II region using this imputation-based approach followed by stepwise conditional logisitic regression. The conditional analysis was repeated in the LADA cohort, which consisted of cases and population-based controls. Crucially, there were no significant independent effects in the MHC Class I region remaining after correction for multiple comparisons; furthermore this observation was replicated in a separate Swedish cohort of T1D cases, LADA cases and population-based controls.

Our observations are further supported through a series of sensitivity analyses by reducing the size of the WTCCC T1D and control cohort, demonstrating that this lack of observation of an MHC Class I independent effect in LADA was not due to reduced power. Additionally, motivated by the hypothesis that results from our previous genetic studies of LADA (Cousminer et al., 2018) can be explained by poor phenotyping of actual T1D and T2D cases, we randomly sampled cases from the WTCCC T1D and T2D cohort to create an artificial LADA cohort under the assumption that LADA would be a “mixture” of actual T1D and T2D cases. In this sensitivity analysis, we still observed the same

independent effects of MHC Class I variants, showing that the T1D signature remained in the “mixture” cohort, and suggesting that our LADA cohorts do not represent a signiifcant proportion of misdiagnosed cases with other forms of diabetes. We recognize that T1D cases were sampled from the cohort of childhood-onset T1D but that could not account for the argument that, in genetic terms, LADA simply represents a mixture of adulthood-onset T1D and T2D.

The MHC Class I variant *HLA-B*39* is an established locus associated with T1D risk (Howson et al., 2009; Nejentsev et al., 2007; Noble et al., 2010). More specifically, studies pointed to a strong association with T1D for the subtype *HLA-B*3906,* which is now used in T1D genetic risk scores to predict T1D diagnosis (Sharp et al., 2019). It has also been shown that the *B*3906* allele significantly enhances the risk of T1D when present on specific *HLA-DR/DQ* haplotypes (e.g. *DRB1*0801-DQB1*0402 and DRB1*0101-DQB1*0501*). The frequency of *HLA-B*3906* is different among different populations, and here did not survive our filter of having a MAF > 1% in the replication control cohort of Swedes. Thus, it was excluded in the analyses across the three cohorts. However, we confirmed that the *HLA-B*3906* allele remained significantly associated with T1D after conditioning on the presence of the *DRB1*0801-DQB1 0402 and DRB1*0101-DQB1*0501* haplotypes. Additionally, *HLA-B*3906* is associated with younger age-at-diagnosis in T1D (Nejentsev et al., 2007; Valdes et al., 2005). A recent study using a NOD mouse model showed that *HLA-B*3906* mediates the development of CD8+ T cells required for T1D onset; moreover, in the context of reduced immunological tolerance to insulin, *HLA-B*3906*-transgenic NOD mice develop T1D at an accelerated rate (Ali et al., 2018). The lack of an independent *HLA-B*39* association observed in the adult-

onset phenotype of LADA further confirms the link between HLA-B*39 with autoimmune progression with earlier onset of clinical disease.

HLA-B associations have been confirmed in a previous study (Eike et al., 2009), as well as associations around *HLA-G,* which is expressed in human pancreas (Cirulli et al., 2006) and may play a role in autoimmune progression (Shiroishi et al., 2003). However, the MHC Class I variant rs1619379, located in *HLA-G* and ~100kb telomeric of *HLA-A*, may be less informative compared to *HLA-A* variants in predicting T1D risk (Howson et al., 2009). This particular MHC Class I variant was independently significant in the downsampled T1D cohort, but is in strong linkage disequilibrium with *HLA-G* variants, rs1610649 and rs2735028, which were significantly associated in the full T1D set, the mixture cohort consisting of T1D and T2D cases, and the T1D Swedish replication cohort. Additionally, the MHC Class I variants located in the *MUC22* locus have not been replicated in separate cohorts, and likely form haplotypes with HLA Class I alleles.

While our study does not include direct HLA typing, leveraging an established HLA imputation method allowed us to investigate HLA associations in the largest LADA cohort to date and to directly compare the results to oberservations made in a childhood-onset T1D cohort. The imputation method SNP2HLA has been commonly used in the field to assess the genetics of autoimmune diseases (Hu et al., 2015; Jia et al., 2013; Karnes et al., 2017; Sharp et al., 2019). One limitation of this study was that we only tested variants with a MAF > 1% in all three control cohorts, which resulted in filtering out many informative alleles such as *HLA-B*3906*. By filtering to include only common alleles we limited potential discrepancies between populations, and were able to

replicate our observations across cohorts with different frequencies of known risk variants.

Future studies are warranted to validate these findings in larger cohorts of LADA, as well as LADA cohorts directly typed for MHC Class II and MHC Class I HLA alleles. Additionally, to futher delineate this putative distinguishing genetic feature between LADA and childhood-onset T1D, it will be crucial to investigate how the HLA profile of LADA compares to children with T1D stratified for different autoantibody positivity status. Studies have shown that children with T1D who are positive for a single autoantibody are more like to show T2D features (Redondo et al., 2017; Redondo et al., 2014), for instance, a significant association with T2D GWAS-implicated variants.

6.5 Conclusion

Overall, our results point to a key differences in the genetic signature in the MHC region, especially Class I markers, between LADA and childhood-onset T1D. This study highlights the clinical utility of genetic screening in adult-onset autoimmune diabetes, the potential of defining those subjects at risk of rapid loss of insulin secretion and the need to consider more tailored approaches to immune therapy according to genetic characteristics.

SNP/HLA-Allele	Locus	Position	Alleles (Major/Minor)	Minor Allele Frequency		Odds Ratio (95% CI)	Single Marker P-value	Conditional			Top Classical HLA allele	LD (r^2)
				Cases	Controls			P-value*	Beta	Standard Error		
rs3957146	*HLA-DQA2*	32789508	T/C	0.385	0.113	4.91 (4.39- 5.50)	8.94E-165	8.94E-165	1.44	0.05	DQB1*0302	0.99
DQB1*0201	*HLA-DQB1*	32739039	A/P	0.338	0.140	3.13 (2.81-3.49)	1.02E-91	7.98E-153	1.56	0.06	DQB1*0201	1.00
rs9268633	*HLA-DRA*	32514451	A/G	0.0174	0.197	0.07 (0.06-0.09)	5.94E-132	5.62E-53	-1.46	0.10	DRB1*1501	0.59
rs1610649	*HLA-G*	29876896	A/G	0.384	0.418	0.87 (0.79-0.94)	5.59E-04	6.89E-23	-0.61	0.06	B*39	0.00
B*39	*HLA-B*	31431272	A/P	0.043	0.016	2.73 (2.07-3.62)	1.82E-12	1.06E-15	1.36	0.17	B*39	1.00
DRB1*0404	*HLA-DRB1*	32660042	A/P	0.082	0.048	1.75 (1.47-2.09)	5.47E-09	6.17E-15	1.04	0.13	DRB1*0404	1.00
rs17427599	*HLA-DQB1*	32775342	C/T	0.151	0.245	0.55 (0.49-0.61)	2.16E-23	7.97E-12	-0.59	0.09	DPB1*0402	0.00
rs2301225	*HLA-DPA1*	33143838	C/T	0.059	0.109	0.51 (0.44-0.60)	1.07E-16	9.56E-12	-0.72	0.11	DPB1*0402	0.96
rs397081	*NOTCH4*	32300595	T/C	0.095	0.045	2.22 (1.86-2.65)	2.20E-18	1.11E-10	0.79	0.12	A*3201	0.00
rs9262545	*MUC22*	31101041	G/A	0.087	0.119	0.70 (0.61 -0.81)	4.89E-06	7.83E-11	-0.67	0.11	A*3201	0.00
rs9262547	*MUC22*	31101206	T/A	0.135	0.119	1.16 (1.02-1.32)	0.034	7.17E-17	1.59	0.19	A*3201	0.00
rs9259852	*HLA-A*	30004400	T/C	0.023	0.041	0.55 (0.43-0.72)	3.04E-06	2.04E-08	-1.00	0.18	A*3201	0.96
rs9269081	*HLA-DRA*	32549078	C/A	0.110	0.265	0.34 (0.30-0.39)	1.19E-62	1.61E-07	-0.57	0.11	DQB1*0602	0.00
rs1978029	*HLA-DQB2*	32839688	T/C	0.349	0.463	0.62 (0.57-0.68)	2.53E-22	1.92E-06	-0.36	0.08	^A*24	0.00

Table 6-1 Independent association signals from the conditional analysis in T1D cases vs WTCCC controls. Conditional P-value calculated from stepwise regression conditional on all SNP/HLA-Alleles in rows above (first column) in 1985 T1D cases versus 2219 controls. *Conditional P-value from stepwise regression conditional on all SNPs/HLA-Alleles (column 1) in rows above. Position is base pair position according to build 36 of the human genome reference. Shaded rows are variants in the MHC Class I region. Linkage disequilibrium (LD) (r^2) is between the SNP/HLA allele and most significant (top) classical HLA allele. "P" in the allele column indicates that the HLA allele is present and "A" indicates absent.^A*24 did not reach significance after conditioning on all variants in the first column.

SNP/HLA-Allele	Locus	Position	Alleles (Major /Minor)	Minor Allele Frequency		Odds Ratio (95% CI)	Single Marker P-value	Conditional			Top Classical HLA allele	LD (r2)
				Cases	Controls			P-value*	Beta	Standard Error		
rs3957146	*HLA-DQA2*	32789508	T/C	0.251	0.101	3.03 (2.63-3.33)	1.80E-68	1.80E-68	1.14	0.06	DQB1*0302	0.99
DRB1*03	*HLA-DRB1*	32660042	A/P	0.209	0.118	1.99 (1.76-2.24)	5.31E-35	3.23E-52	1.07	0.07	DRB1*03	1
rs9269081	*HLA-DRA*	32549078	C/A	0.179	0.315	0.47 (0.42 -0.53)	1.40E-42	3.54E-16	-0.49	0.06	DQB1*0604	0.01
DRB1*0404	*HLA-DRB1*	32660042	A/P	0.037	0.035	1.04 (0.82-1.33)	0.328084	3.08E-11	0.95	0.14	DRB1*0404	1
DQB1*0604	*HLA_DQB1*	32739039	A/P	0.059	0.035	1.75 (1.43-2.17)	5.73E-07	1.55E-10	0.81	0.13	DQB1*0604	1
rs2143462	*C6orf10*	32443182	C/T	0.221	0.167	1.42 (1.26-1.58)	1.39E-08	8.24E-08	0.47	0.09	DPA1*02	0.01
DPA1*02	*HLA-DPA1*	33145064	A/P	0.148	0.186	0.76 (0.67-0.86)	1.23E-04	1.62E-06	-0.34	0.07	DPA1*02	1
rs3130192	*HLA-DPB1*	33169908	C/T	0.066	0.103	0.62 (0.52-0.73)	3.96E-08	5.32E-06	-0.42	0.09	^DPB1*0402	0.85

Table 6-2 Independent association signals from the conditional analysis in LADA cases vs controls. Conditional P-value calculated from stepwise regression conditional on all SNP/HLA-alleles in rows above (first column) in 1428 LADA cases versus 2979 controls. Conditional P-value from stepwise regression conditional on all SNPs/HLA-Alleles (column 1) in rows above. Position is base pair position according to build 36 of the human genome reference. LD (r^2) is between the SNP/HLA allele and the most signficant (top) classical HLA allele. “P” in the allele column indicates the HLA allele is present and “A” indicates absent. ^DPB1*0402 did not reach significance after conditioning on all variants in the first column.

Table 6-3 Independent association signals from the conditional analysis in T1D vs controls in the Swedish replication cohort. Conditional P-value calculated from stepwise regression conditional on all SNP/HLA-Alleles in rows above (first column)

SNP/HLA-Allele	Locus	Position	Alleles (Major /Minor)	Frequency		Odds Ratio (95% CI)	Single Marker P-value	Conditional			Top Classical HLA allele	LD (r2)
				Cases	Controls			P-value*	Beta	Standard Error		
rs9275206	*HLA-DQA1*	32765543	A/G	0.393	0.147	3.70 (3.33-4.34)	6.35E-89	6.35E-89	1.57	0.08	DQB1*0302	0.99
DQB1*0201	*HLA-DQB1*	32739039	A/P	0.312	0.129	3.07 (2.67-3.52)	2.17E-58	1.96E-135	1.98	0.08	DQB1*0201	1.00
DQB1*0602	*HLA-DQB1*	32739039	A/P	0.005	0.138	0.03 (0.02-0.07)	2.41E-43	1.85E-40	-1.26	0.09	DQB1*0602	1.00
rs9269081	*HLA-DRA*	32549078	C/A	0.076	0.262	0.23(0.19-0.29)	1.60E-45	1.99E-30	-0.82	0.07	DRB1*0401	0.06
rs3129871	*HLA-DRA*	32514320	C/A	0.095	0.335	0.21(0.17-0.25)	1.27E-65	3.33E-22	-0.63	0.07	DRB1*0401	0.06
rs9784758	*TAP2*	32896489	T/C	0.212	0.091	2.70(2.27-3.13)	2.02E-34	1.03E-23	0.97	0.10	DRB1*0401	0.11
rs805294	*LY6G6C*	31796196	C/T	0.530	0.359	2.01(1.78-2.27)	6.41E-30	8.02E-20	0.55	0.06	B*39	0.00
rs707919	*LY6G5B*	31749118	T/C	0.333	0.310	1.11(0.98-1.27)	0.04	5.61E-16	0.42	0.05	DRB1*0401	0.13
rs3130192	*HLA-DPB1*	33169908	C/T	0.049	0.107	0.43(0.33-0.56)	6.17E-10	8.83E-15	-0.77	0.10	DPB1*0402	0.88
rs11969522	*PSMB9*	32967803	G/C	0.081	0.033	0.44 (0.39-0.49)	5.98E-16	6.45E-13	-1.10	0.15	B*39	0.00
rs9275425	*HLA-DQB1*	32778852	C/A	0.487	0.293	2.22(2.04-2.56)	1.33E-40	2.54E-12	0.37	0.05	B*39	0.01
rs1053924	*PRRT1*	32228693	G/A	0.144	0.305	0.38(0.32-0.45)	1.16E-32	8.02E-11	-0.43	0.07	B*39	0.00
B*39	*HLA-B*	31431272	A/P	0.034	0.013	2.75(1.90-3.98)	6.80E-09	1.74E-10	1.54	0.24	B*39	1.00
rs2735028	*HLA-G*	29893517	G/A	0.319	0.335	0.93(0.82-1.05)	0.315203	1.10E-09	-0.39	0.06	A*0101	0.42
rs3104407	*HLA-DQA2*	32790430	A/G	0.244	0.457	0.38(0.33-0.44)	5.10E-45	3.01E-10	-0.36	0.06	DRB1*0401	0.12
rs3104406	*HLA-DQA2*	32790421	G/A	0.143	0.339	0.33(0.28-0.38)	2.23E-39	9.71E-11	-0.37	0.06	DQB1*0301	0.11
rs2239803	HLA-DRA	32519811	G/A	0.543	0.436	1.54(1.37-1.72)	3.33E-13	3.67E-09	0.32	0.05	DQB1*0503	0.01
rs3117099	*BTNL2*	32466248	C/T	0.361	0.246	1.73(1.53-1.97)	3.29E-17	7.95E-11	0.35	0.05	DQB1*0604	0.06
rs549182	*NOTCH4*	32313023	G/A	0.040	0.017	2.44(1.75-3.41)	9.09E-09	1.37E-06	0.97	0.20	DQB1*0503	0.01
rs3135392	*HLA-DRA*	32517220	G/T	0.446	0.417	1.13(1.00-1.27)	0.04	3.23E-07	0.15	0.03	DRB1*0401	0.15
rs3132132	*HLA-DMB*	33009912	G/A	0.082	0.116	0.68(0.55-0.84)	6.15E-04	1.81E-06	-0.46	0.10	DRB1*0401	0.01
rs3129727	*HLA-DQA2*	32787668	C/T	0.001	0.021	0.04(0.01-0.26)	1.44E-06	7.31E-06	-1.09	0.24	^DQB1*0503	0.99

in 656 T1D cases versus 3218 controls. Position is base pair position according to build 36 of the human genome reference. Shaded rows denote this signal appeared in the full T1D vs WTCCC dataset. LD (r^2) is between the SNP/HLA allele and top classical HLA allele. “P” in the allele column indicates the HLA allele is present and “A” indicates absent. ^DQB1*0503 did not reach significance after conditioning on all variants in the first column.

SNP/HLA-Allele	Locus	Position	Allele (Major /Minor)	Frequency		Odds Ratio (95% CI)	Single Marker P-value	Conditional			Top Classical HLA allele	LD (r2)
				Cases	Controls			P-value*	Beta	Standard Error		
rs3129882	*HLA-DRA*	32517508	A/G	0.223	0.378	0.47(0.42-0.54)	3.51E-10	3.51E-10	-0.84	0.13	HLA_DRB1_15	0.20
rs2596560	*MICA*	31463297	A/G	0.586	0.747	0.48 (0.43-0.53)	3.67E-08	4.19E-07	-0.67	0.13	HLA_DQB1_0201	0.35

Table 6-4 Independent association signals from the conditional analysis in LADA vs controls in the Swedish replication cohort. Conditional P-value calculated from stepwise regression conditional on all SNP/HLA-Alleles in rows above (first column) in 823 LADA cases versus 3211 controls. “P” in the allele column indicates the HLA allele is present and “A” indicates absent.

SNP/HLA-Allele	Locus	Position	Allele (Major /Minor)	Frequency		Odds Ratio (95% CI)	Single Marker P-value	Conditional			Top Classical HLA allele	LD (r2)
				Cases	Controls			P-value*	Beta	Standard Error		
rs3957146	*HLA-DQA2*	32789508	T/C	0.390	0.113	5.02 (4.45-5.66)	2.98E-157	2.98E-157	1.57	0.06	DQB1*0302	0.99
DQB1*0201	*HLA-DQB1*	32739039	A/P	0.338	0.140	3.14 (2.80-3.528)	7.33E-81	3.27E-136	1.67	0.07	DQB1*0201	1.00
rs9268633	*HLA-DRA*	32514451	G/A	0.020	0.197	12.3 (9.34-16.17)	6.40E-102	2.02E-37	1.35	0.11	DQB1*0602	0.53
rs1619379	*HLA-G*	29893214	G/A	0.398	0.432	0.87 (0.79-0.95)	0.002	1.37E-17	-0.56	0.07	DRB1*0404	0.00
DRB1*0404	*HLA-DRB1*	32660042	A/P	0.083	0.048	1.78 (1.47-2.15)	1.54E-08	9.91E-13	0.98	0.14	DRB1*0404	1.00
B*39	*HLA-B*	31431272	A/P	0.043	0.016	2.79 (2.08-3.75)	5.09E-12	5.58E-14	1.43	0.19	B*39	1.00
rs2301225	*HLA-DPA1*	33143838	C/T	0.057	0.109	0.49 (0.41-0.59)	1.60E-14	3.96E-12	-0.81	0.12	DPB1*0402	0.96
rs9267665	*C2*	31978835	C/T	0.082	0.030	2.92 (2.35-3.64)	1.86E-21	1.73E-09	0.88	0.15	DQB1*0602	0.00
rs9269081	*HLA-DRA*	32549078	C/A	0.110	0.265	0.34 (0.30-0.39)	1.60E-54	1.79E-13	-0.80	0.11	DQB1*0602	0.31
rs9262545	*MUC22*	31101041	G/A	0.087	0.119	0.70 (0.60-0.82)	8.07E-06	3.36E-09	-0.67	0.11	A*24	0.00
rs9262547	*MUC22*	31101206	T/A	0.133	0.119	1.13 (0.98-1.31)	0.16388	9.69E-14	1.64	0.22	B*14	0.00
rs549182	*NOTCH4*	32313023	G/A	0.074	0.020	3.93 (3.05-5.06)	5.59E-28	1.20E-06	1.33	0.27	B*14	0.00
rs2853928	*HLA-C*	31365490	G/T	0.372	0.295	1.42 (1.28-1.57)	3.22E-11	1.26E-06	0.48	0.10	C*07	0.79

Table 6-5 Independent association signals from the conditional analysis in the downsampled cohort of T1D cases vs controls. Conditional P-value calculated from stepwise regression conditional on all SNP/HLA-alleles in rows above (first column) in 1428 T1D cases versus 2219 controls. Shaded rows denote that this MHC Class I signal appeared in the full T1D vs WTCCC dataset but not in LADA vs control set. *rs1619379 is in LD with rs1610649 (r2=0.822), which is consistent in the full T1D vs WTCCC dataset. “P” in the allele column indicates the HLA allele is present and “A” indicates absent. ^C*07 did not reach significance after conditioning on all variants in the first column.

SNP/HLA-Allele	Locus	Position	Allele (Major /Minor)	Allele Frequency		Odds Ratio (95% CI)	Single Marker P-value	Conditional			Top Classical HLA allele	LD (r2)
				Cases	Controls			P-value	Beta	Standard Error		
rs3957146	*HLA-DQA2*	32789508	T/C	0.256	0.112	2.72 (2.40-3.09)	1.16E-52	1.16E-52	0.99	0.07	DQB1*0302	0.99
rs2187668	*HLA-DQA1*	32713862	G/A	0.234	0.140	1.88 (1.66-2.12)	2.71e-27	1.03E-38	0.89	0.07	DRB1*0301	0.99
DRB1*0404	*HLA-DRB1*	32660042	A/P	0.059	0.049	1.22 (0.99-1.51)	0.075	2.11E-10	0.85	0.13	DRB1*0404	1.00
rs1619379	*HLA-G*	29893214	G/A	0.400	0.430	0.88(0.800.97)	0.014	1.60E-07	-0.30	0.06	B*39	0.00
rs9262545	*MUC22*	31101041	G/A	0.083	0.119	0.67 (0.57-0.79)	3.43E-07	6.42E-07	-0.44	0.09	B*39	0.00
rs9262547	*MUC22*	31101206	T/A	0.117	0.118	0.99 (0.85-1.14)	0.548	3.03E-16	1.63	0.20	B*39	0.00
HLA_B_39	*HLA-B*	31431272	A/P	0.030	0.015	1.99 (1.44-2.77)	4.73E-05	2.61E-06	0.88	0.19	^B*39	1.00

Table 6-6 Independent association signals from the conditional analysis in a sample of randomly selected T1D and T2D cases vs controls. Conditional P-value calculated from stepwise regression conditional on all SNP/HLA-alleles in rows above (first column) in 714 T1D cases + 714 T2D cases versus 2219 controls. Shaded rows highlight the MHC Class I signals that appear in the full T1D vs WTCCC conditional analysis. “P” in the allele column indicates that the HLA allele is present and “A” indicates absent. ^B*39 did not reach significance after conditioning on all variants in the first column.

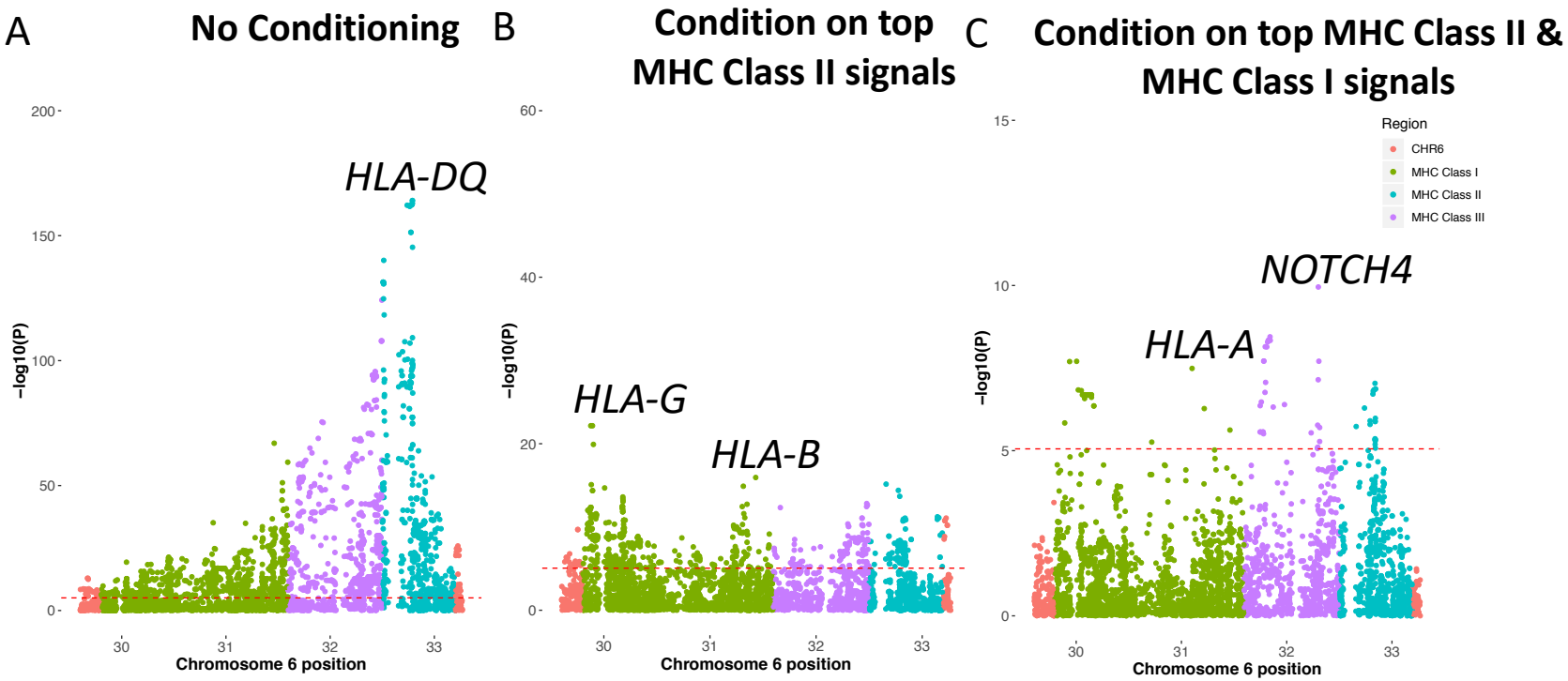


Figure 6-2 Conditional analysis in 1,985 T1D cases and 2,219 WTCCC controls. *A) Logistic regression analysis without conditioning on MHC Class II alleles. B) Logistic regression analysis conditioning on MHC Class II alleles. C) Logistic regression analysis conditioning on MHC Class II and MHC Class I signals.*

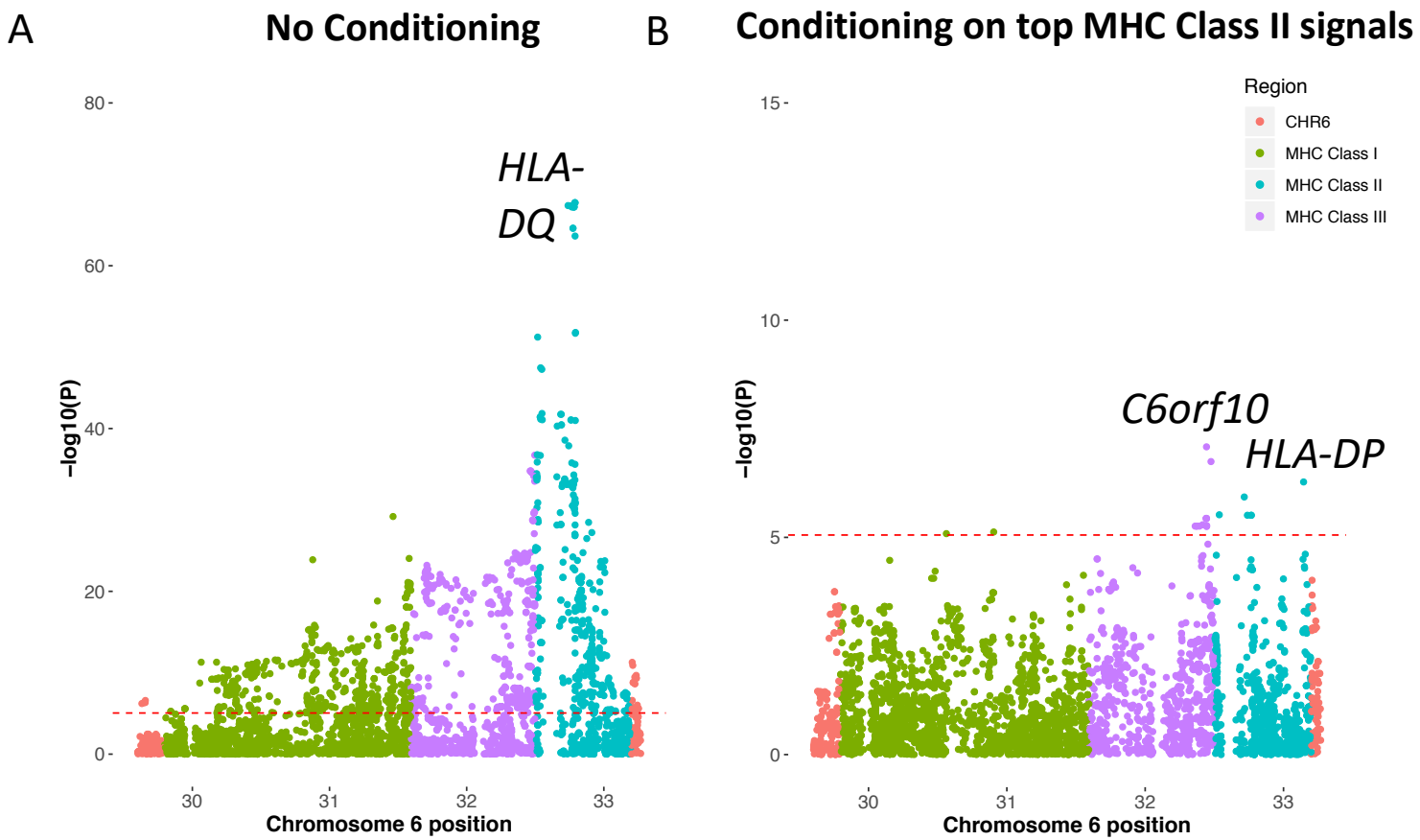


Figure 6-3 Conditional analysis in 1,428 LADA cases and 2,979 WTCCC controls. A) Logistic regression analysis without conditioning on MHC Class II alleles. B) Logistic regression analysis conditioning on MHC Class II alleles.

CHAPTER 7. FUTURE DIRECTIONS

7.1 Brief Summary

While the genetic risk scores and heritability calculations both suggested LADA genetically resides between T1D and T2D, principally T1D GWAS-implicated loci including the MHC region, *PTPN22*, *INS*, and *SH2B3*, were shown to be significantly associated with LADA in both the candidate locus association study and the GWAS meta-analysis. The GWAS also highlighted a novel signal at the *PFKFB3* locus, which had never been reported in previous T1D and T2D GWAS, and could be associated with older age of onset of autoimmune diabetes. Additionally, the MHC association with LADA is attenuated compared to T1D, with established T1D risk haplotypes less frequent in LADA, and conversely protective T1D haplotypes more frequent. When looking further into the HLA region in LADA, contrary to observations in childhood-onset T1D, we did not observe independent effects of MHC Class I variants, particularly the *HLA-B*39* allele, which is consistent with previous studies. Despite the significant role T1D genetics plays in LADA, there is also a positive genetic correlation between LADA and T2D, and significant enrichment of directional consistency for T2D loci in LADA. However, it remains unclear whether these results are at least partially explained by T2D individuals mistakenly classified as “affected with LADA”, and the studies presented in this dissertation have yet to fully answer this outstanding question due to the challenges outlined below.

7.2 Limitations

The limitations of our genetic efforts in LADA include the lack of power to detect most T2D signals that would reach genome-wide significance. If our hypothesis of a T2D component of LADA holds, the T2D loci with relatively weak effect sizes should emerge in a larger cohort given the contrasting genetic architecture between T1D and T2D (Timpson et al., 2018). Additionally, there is a possibility of misdiagnosed individuals in our LADA cohort and in the T1D and T2D reference cohorts, which was a major challenge in all of our analyses. Another limitation is the reference T1D and T2D GWAS that were utilized in our studies. For instance, Bradfield *et al* used T2D cases as part of their controls in a T1D GWAS, which resulted in conservative effects sizes that were used in our initial T1D GRS calculations. Given that T2D genetic studies do not screen for autoantibodies and T1D-associated loci have started to emerge in these studies (Mahajan et al., 2018; Ng et al., 2014; Scott et al., 2017), it is also very possible that our analyses using T2D GWAS data may have been confounded by poor phenotyping. Moving forward, one of our key conclusions is that T2D genetic studies must exclude cases positive for diabetes associated autoantibodies. Additionally, we did not have consistent access across the cohorts to important clinical data such as age of onset, time to insulin and BMI, and systematically collecting this information will be crucial for future studies.

7.3 Is LADA simply a mixture of T1D and T2D?

In the first attempt to answer this question, different proportions of T1D and T2D cases were randomly sampled from the WTCCC dataset to represent an "artificial" model of a possible LADA cohort and the distributions of T1D and T2D GRS were compared to the distribution of T1D and T2D GRS in LADA (n=978) (**Figure 7-1A and 7-1B**). GRS were calculated using 69 T1D loci and 71 T2D GWAS from Chapter 3 (**Appendix: Supp. Table 3-1**). Wilcoxon rank sum test was used to perform a pairwise comparison amongst the six groups (10% T1D cases/90% T2D cases, 25% T1D cases/75% T2D cases, 50% T1D cases/50% T2D cases, 75% T1D cases/25% T2D cases, 100% T1D cases and LADA cases). The T1D GRS distribution for the 50% T1D and 50% T2D mixture cohort appeared to be nearly identical to the T1D GRS distribution for LADA (P = 1.00).

This cocktail analysis was repeated after increasing the LADA sample size to 2,735 cases and using 30 T1D loci and their respective odds ratio from Oram *et al* for T1D GRS (Oram et al., 2016) and 403 T2D loci from the most recent T2D GWAS (Mahajan et al., 2018) for the T2D GRS(**Figure 7-1C and 7-1D**). In this analysis, we still saw an identical distribution of T1D GRS between 50/50 T1D/T2D GRS mixture and LADA, which was consistent with the initial T1D GRS cocktail analysis. However, contrary to the initial T2D GRS cocktail analysis using 71 SNPs, the T2D GRS distribution of LADA was near identical to the T2D GRS distribution of the 75% T1D/25% T2D mixture. This observation was also seen using the same 403 SNP-derived T2D GRS in the 978 LADA cases, which highlights how crucial T2D GWAS references will be for future work. Differences between the two T2D cocktail analyses could be explained by

the potential increase in influence of the misdiagnosis rate resulting in T1D loci (*i.e. HLA, INS*) emerging in the recent T2D GWAS.

To assess whether there is a correlation between the T1D genetics and T2D genetics for an individual with LADA, we plotted the T2D GRS against the T1D GRS in LADA (n=2,735), controls (WTCCC controls and BMDCS, N=4,286), WTCCC T1D (N=1,995), WTCCC T2D (N=1,971), and different mixtures of T1D and T2D cases. We observed a negative correlation between the T2D GRS and T1D GRS for LADA (**Figure 7-2A**); however, this is also seen when randomly sampling different proportions of T1D and T2D cases (**Figure 7-2B**). While the use of T1D and T2D whole-genome polygenic risk scores could be more informative, without a cleaner reference T2D GWAS dataset, we still may not fully answer whether the T2D genetic component of LADA is explained by "T2D contaminates" in the cohort.

A future approach to investigate this outstanding question is to leverage the summary statistics from the T1D, T2D and LADA GWAS. Therefore, future T2D GWAS removing individuals positive for diabetes associated autoantibodies will be imperative. By leveraging summary statistics data, two potential hypotheses can be tested: 1) LADA is simply a form of T1D with low penetrance and 2) LADA is a mixture of T1D and T2D cases. The first hypothesis can be tested by estimating parameters that will obtain a 'low penetrant T1D model' that best fits the association summary statistics based on the LADA GWAS. The second hypothesis can be tested by estimating the proportion of T1D cases + T2D cases that will obtain a model that best fits the data observed from the LADA GWAS. We can further improve this study by leveraging samples from the Type 1 Diabetes Genetic Consortium and the UK Biobank, using strict criteria for T1D

and T2D cases. A potential result from this study would be identifying a model that best fits LADA, however it may not explain the observations we see in the LADA GWAS. This would suggest a genetic signature unique to LADA. The model that best fits LADA should clinically and biologically make sense as well.

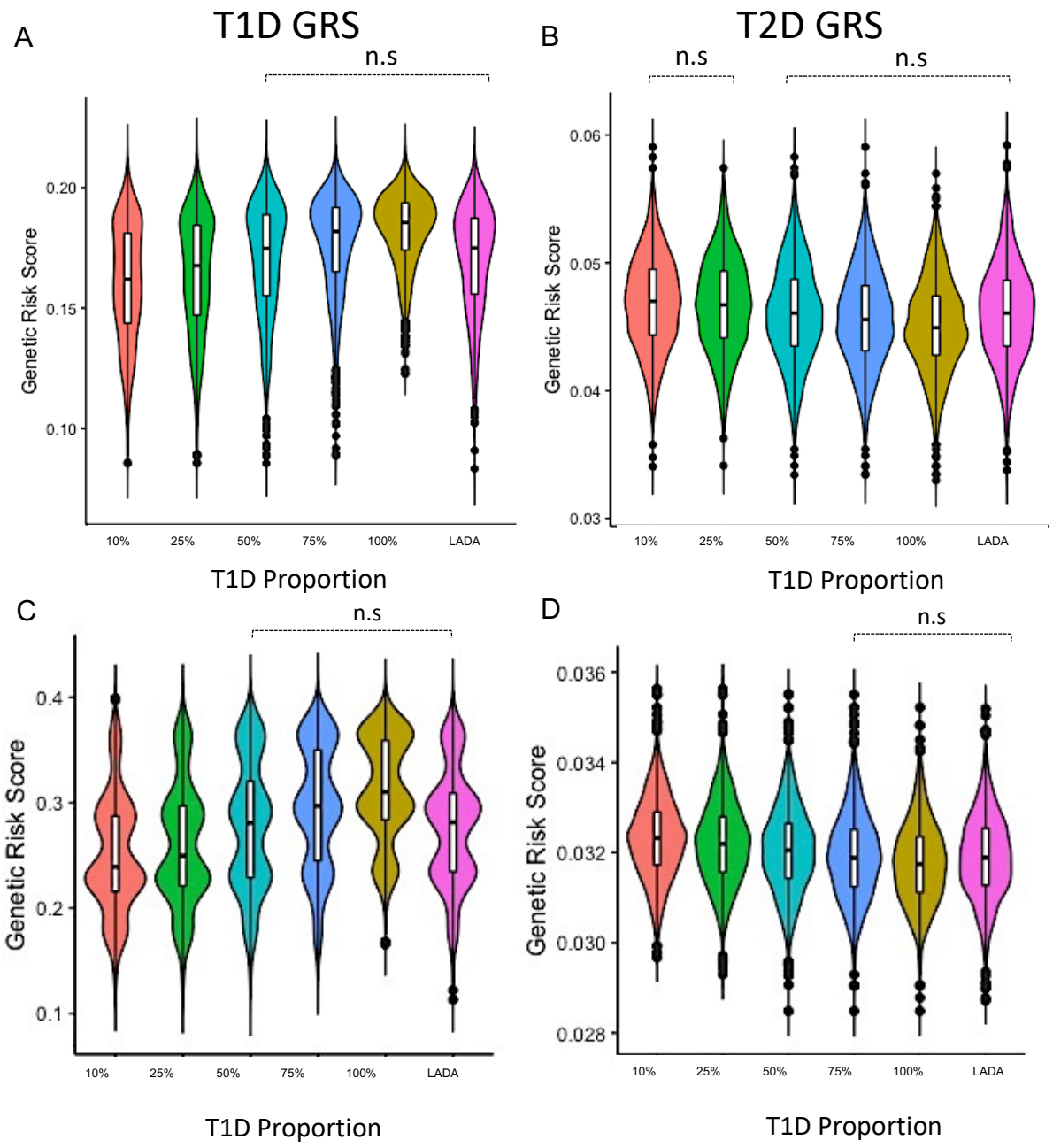


Figure 7-1 GRS distribution of varying proportions of T1D and T2D cases (A) Compared to T1D GRS in LADA (n=978) (B) Compared to T2D GRS in LADA (n=978) (C) Compared to Oram et al T1D GRS in LADA (n=2,735) (D) Compared to updated T2D GRS in LADA (n=2,735)

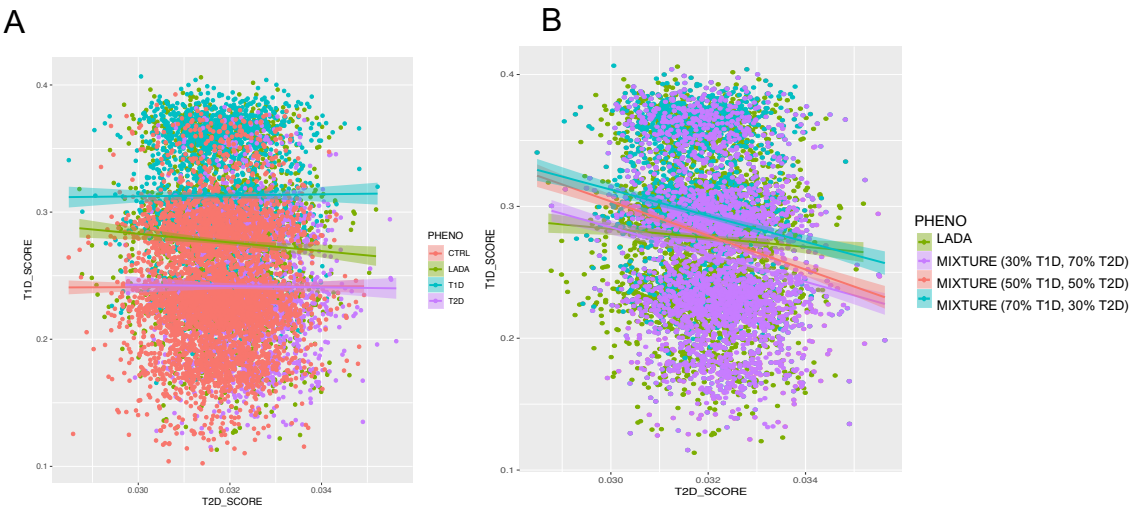


Figure 7-2 T2D GRS plotted against T1D GRS in LADA (n=2,735), controls (N=4,286), T1D (N=1,995), T2D (N=1,971), and different mixtures of T1D and T2D cases.

7.4 Follow-up on genetic markers involved in autoimmune diabetes

The presence of autoantibodies and the T1D genetic component of LADA has a lot of power to discriminate LADA from T2D. However, our work has also pointed to potential genetic discriminators of LADA and childhood-onset T1D, that can help shed light on the progression of autoimmune diabetes. For instance, the novel independent LADA-associated variant, located between two known signals at a T1D locus, points to an interesting functional candidate, '6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 3' (*PFKFB3*), that links immune and metabolic traits. *PFKFB3* is a known regulator of glycolysis and insulin signaling in T2D and inflammation and autophagy in autoimmune diseases. While our LADA GWAS was the first to report this novel signal, the variant subsequently reached genome-wide association in a recent T1D study (Aylward et al., 2018).

One explanation for why this signal has not been reported before could be lack of power for prior T1D studies. However, another explanation, particularly for (Bradfield et al., 2011), could be the inclusion of WTCCC T2D cases in the control group. Thus, the novel variant is not a unique marker for LADA, however it is possible that this variant could be associated with older age of onset and is thus more amenable for detection in a LADA setting; therefore, future genetic studies of autoimmune diabetes in all ages should further investigate this variant and its role in the progression of autoimmune diabetes.

Additionally, genetic variation in the MHC Class I region may also point to potential discriminators between LADA and childhood-onset T1D. Particularly, *HLA-B*3906* has been shown in previous studies to be associated with early age of onset, and our lack of observation of an independent effect of this signal in LADA confirms this relationship. It is also possible that T1D variants, which we had power to detect in our LADA GWAS but did not reach genome-wide significance in the study, could also mediate the progression of autoimmune diabetes. Identifying variants associated with early onset of T1D or conversely delay in progression of T1D, and functionally following up on their involvement in insulin production and beta cell autoimmunity will be an interesting future direction from this work.

7.5 Towards precision medicine in diabetes

While some scientists and clinicians believe LADA is simply a mixture of T1D and T2D, some believe LADA is part of a diabetes continuum with childhood-onset T1D on one end of the spectrum and T2D on the other end of the spectrum. Our genetic studies of LADA, as well as childhood-onset T1D studies, have supported the hypothesis of an interplay between T1D and T2D genetics.

Additionally, individuals can have both T1D and T2D, which has been termed "double diabetes" (Pozzilli et al., 2011). Atypical cases such as double diabetes, and "Flatbush diabetes" demonstrate that diabetes definitions are certainly not clear cut and is extremely heterogenous. The extent of heterogeneity within T2D alone has been demonstrated in previous clustering analyses (Ahlqvist et al., 2018; Udler et al., 2018). As a result of this phenotypic heterogeneity, misclassification may have confounded genetic studies of diabetes which makes it more challenging to understand the genetics and pathogenesis. Large scale genetic data combined with patient data that includes BMI, C-peptide levels, HbA1c, insulin treatment, time to insulin, age of diagnosis, and autoantibody titre will help to further move forward the field of precision medicine in diabetes.

7.6 Conclusion

By conducting the first genome-wide systematic appraisal of LADA, we have demonstrated that the genetic etiology of this relatively common diabetes subtype is largely similar to the genetic etiology of T1D, suggesting that LADA is an age-related extension of the childhood-onset T1D pathogenesis. Although there was clear evidence of a T2D genetic component, further studies are required to investigate the extent to which these observations are due to poor phenotyping. Finally, we did not find strong evidence for novel genetic variants unique to LADA so one of our conclusions is that LADA is very unlikely to be a separate diseases entity, although some MHC signatures may act as a discriminator from T1D. Our genetic efforts in LADA have moved the field forward by sparking the debate surrounding the definition of LADA and highlighting the

need for a better classification system for diabetes. Additionally, this work has highlighted the potential utility of genetics in identifying and monitoring high risk patients with adult-onset autoimmune diabetes, thus leading to better and more refined treatment and therapeutic interventions.

APPENDIX

Chapter 2: A Global Perspective of LADA Supplemental Table

Region	County	2015 National prevalence of diabetes % [uncertainty interval]	Approximate number of Adults with diabetes [aged 20-79]	Approximate number of children with type 1 diabetes [aged 0-14]	Basis of estimate [OGTT, extrapolation, other[1]]
Africa	Nigeria	1.9 [1.4 - 4.6]	1,564,700	14,400	Extrapolation
	Ghana	1.9 [1.5 - 4.7]	266,200	N/A	Extrapolation
	Kenya	2.2 [1.2 -10.2]	478,000	N/A	OGTT
	Ethopia	2.9 [1.7 - 7.6]	1,333,200	578	Extrapolation
Middle East	Iran	8.5 [6.6 -11.5]	4,602,200	3,100	Other
	Saudi Arabia	17.6 [13.5 -19.6]	3,487,300	16,100	OGTT
	United Arab Emirates	14.6 [13.0 - 17.1]	1,086,300	N/A	OGTT
Asia	India	8.7 [7.0 - 10.6]	69,188,600	70,000	OGTT
	Sri Lanka	8.5[6.1 - 12.9]	1,161,700	N/A	OGTT
	China	10.6 [9.6 - 12.9]	109,649,100	30,500	OGTT
	Republic of Korea	8.7 [6.9 - 11.2]	3,369,000	400	Other
	Japan	7.6 [6.5 - 10.1]	7,202,200	2,500	Other
North America	USA	12.8 [12.1 - 13.5]	29,251,600	84,100	OGTT
Europe	UK	6.2 [5.4 - 8.5]	2,858,600	19,800	OGTT
	Finland	9.0 [6.7 - 11.0]	360,000	4,400	OGTT
	Italy	7.9 [7.1 - 9.2]	3,507,700	6,800	OGTT
	Sweden	6.3 [5.2 - 9.5]	446,900	4,400	OGTT
	Norway	7.8 [5.9 - 9.7]	289,600	1,900	Extrapolation
	Spain	10.4 [8.2 - 14.7]	3,576,100	8,800	OGTT
	Turkey	12.5 [11.2 - 14.9]	6,339,000	9,600	OGTT

Supp. Table 2-1. National prevalence of diabetes in 2015

This data is derived from the International Diabetes Federation(Ogurtsova et al., 2017)for highlighted countries in Chapter 1.[1]Other includes 'Fasting blood glucose', 'Self-reported', 'Medical record or clinical diagnosis', and 'HbA1c'; OGTT = oral glucose tolerance test.

Chapter 3: RELATIVE CONTRIBUTION OF T1D AND T2D LOCI TO THE GENETIC ETIOLOGY OF LADA

Supplemental Results

Given our primary control cohort consisted of European-descent children ascertained from the United States, while the LADA cases were adults ascertained from the UK and Germany, we also leveraged 2,820 healthy adult British controls from the WTCCC and, overall, observed consistent results (**Appendix: Supp. Table 3-8**), despite the array differences for this control set, as outlined in the Methods section. The T1D signals at the MHC (OR=1.18, P=1.01x10-4), *PTPN22* (OR=1.33; P=3.86x10-6), *SH2B3* (OR=1.16; P=9.35x10-5), and *INS* (rs689; OR=1.24; P=1.27x10-6) remained strongly associated and were directionally consistent. Similar observations were also seen when comparing the MHC ($P_{difference}$=1.26x10-17), INS (rs689; $P_{difference}$= 3.88x10-4), and SMARCE1 ($P_{difference}$= 6.54x10-4) loci in the LADA cases vs the WTCCC T1D cases. While the *HNF1A* association weakened, it did remain significantly associated (OR=1.11; P=0.036) while the *TCF7L2* signal remained significantly differently associated in LADA cases vs T1D cases ($P_{difference}$=5.21x10-6).

When constrained on GADA+ only LADA cases, the signal at the MHC was no longer significantly associated while *TCF7L2* remined non-significant for cases vs WTCCC controls; however, they both remained

significantly different between LADA cases vs T1D cases ($P_{difference}$=1.99x10-24 and 5.03x10-4, respectively). Furthermore, when constrained on LADA cases positive for GADA+ and IA2A+ antibodies, the T1D association signals remained all significant, although not all were significantly different from T1D, i.e. MHC (OR=1.59, P=3.34x10-9, $P_{difference}$=4.01x10-3), *PTPN22* (OR=1.86; P=1.88x10-7, $P_{difference}$= N.S), *SH2B3* (OR=1.32; P=1.95x10-4, $P_{difference}$= N.S), and *INS* (rs689; OR=1.48; P=2.07x10-6 , $P_{difference}$=1.68x10-6) showing again that this set of LADA cases is overall more similar to T1D cases. The *HNF1A* signal remained significantly associated in these restricted LADA cases (OR=1.27; P=0.011), while the *ZBED3* and *TCF7L2* loci continued to be significantly different in LADA cases vs T1D cases ($P_{difference}$=1.47x10-5 and 2.56x10-7, respectively).

Tables

T1D SNPS	Gene	Odds Ratio	Weight	Effect Allele
rs2187668, rs7454108	DR3/DR4-DQ8	48.18	3.87	
	DR3/DR3	21.12	3.05	
	DR4-DQ8/DR4-DQ8	21.98	3.09	
	DR4-DQ8/X	7.03	1.95	
	DR3/X	4.53	1.51	
rs1264813	HLA_A_24	1.54	0.43	T
rs2395029	HLA_B_5701	2.51	0.92	T
rs3129889	HLA_DRB1_15	14.88	2.7	A
rs10272724	*IKZF1, FIGNL1, DDC, GRB10*	1.15	0.14	T
rs10492166	*CD69*	1.15	0.14	G
rs10509540	*RNLS*	1.14	0.13	T
rs10517086	*SLC34A2, SEL1L3, SMIM20, RBPJ, CCKAR, TBC1D19*	1.09	0.09	A
rs10795791	*IL2RA, RBM17*	1.16	0.15	G
rs11170466	*ITGB7*	1.19	0.17	T
rs11203203	*UBASH3A*	1.14	0.13	A
rs113010081	*CCR5*	1.19	0.17	T
rs11571316	*CTLA4, ICOS*	1.22	0.2	G
rs11755527	*BACH2*	1.13	0.12	G
rs11954020	*IL7R*	1.11	0.1	G
rs12148472	*CHRNB4, ADAMTS7, MORF4L1, CTSH, RASGRF1*	1.20	0.18	T
rs12453507	*FBXL20, MED1, CDK12, NEUROD2, PPP1R1B, STARD3, TCAP, PNMT, PGAP3, ERBB2, MIEN1, GRB7, IKZF3, ZPBP2, GSDMB, ORMDL3, LRRC3C, GSDMA, PSMD3, CSF3, MED24, THRA*	1.11	0.1	G
rs12708716	*CLEC16A,CIITA, DEXI, RMI2, SOCS1, TNP2, PRM3, PRM2, PRM1, CTD-3088G3.8*	1.23	0.21	A
rs12908309	*RASGRP1*	1.19	0.17	G
rs1456988	*C14orf64*	1.12	0.11	G
rs1538171	*CENPW*	1.12	0.11	G
rs1615504	*DOK6, CD226*	1.13	0.12	T
rs1738074	*RSPH3, TAGAP*	1.09	0.09	C
rs17696736	*SH2B3,NAA25, CUX2, FAM109A, ATXN2, BRAP, ACAD10, RP11-162P23.2 (ENSG00000257767), ALDH2, MAPKAPK5, TMEM116, ERP29, TRAFD1, HECTD4, RPL6, PTPN11, RPH3A*	1.34	0.29	G
rs193778	*CLEC16A,CIITA, DEXI, RMI2, SOCS1, TNP2, PRM3, PRM2, PRM1, CTD-3088G3.8*	1.14	0.13	G
rs1990760	*GCG, FAP, IFIH1, GCA, KCNH7*	1.15	0.14	T
rs2281808	*SIRPD, RP4-576H24.4 (ENSG00000260861), SIRPB1, SIRPG*	1.11	0.1	C
rs2304256	*ICAM1, ICAM4, ICAM5, ZGLP1, FDX1L, FDX1L, CTD-2369P2.12, RAVER1, ICAM3, TYK2, CDC37, PDE4A, KEAP1, S1PR5*	1.16	0.15	C
rs3024493	*MAPKAPK2, IL10, IL19, IL20*	1.22	0.2	C
rs3825932	*CHRNB4, ADAMTS7, MORF4L1, CTSH, RASGRF1*	1.16	0.15	T
rs402072	*DACT3, PRKD2, STRN4, FKRP, SLC1A5*	1.15	0.14	T
rs4763879	*CD69, KLRB1, CLEC2D, CLECL1*	1.09	0.09	A
rs478222	*DNMT3A,CENPO, ADCY3, DNAJC27, EFR3B, POMC*	1.15	0.14	A
rs4849135	*ACOXL,LINC00116, LIMS3, LIMS3L, RGPD6, BUB1, ACOXL*	1.12	0.11	G
rs4948088	*COBL*	1.30	0.26	C
rs5753037	*NF2, CABP7, ZMAT5, UQCR10, ASCC2, MTMR3, HORMAD2, LIF, OSM*	1.11	0.1	T
rs597325	*BACH2*	1.19	0.17	G
rs601338	*FUT2,SULT2B1, FAM83E, SPACA4, RPL18, SPHK2, DBP, CA11, NTN5, MAMSTR, RASIP1, IZUMO1, FUT1, FGF21*	1.34	0.29	A
rs602662	*FUT2,SULT2B1, FAM83E, SPACA4, RPL18, SPHK2, DBP, CA11, NTN5, MAMSTR, RASIP1, IZUMO1, FUT1, FGF21*	1.12	0.11	A
rs61839660	*IL2RA, RBM17, PFKFB3*	1.62	0.48	C
rs6476839	*GLIS3*	1.12	0.11	T
rs6679677	*PTPN22, PHTF1*	1.82	0.6	A
rs6691977	*KIF14, DDX59, CAMSAP2, GPR25, C1orf106*	1.13	0.12	C
rs689	*INS,IGF2, INS-IGF2,TH*	2.39	0.87	T
rs6920220	*TNFAIP3*	1.12	0.11	A
rs705705	*IKZF4,PMEL, CDK2, RAB5B, SUOX, RPS26, ERBB3, RP11-603J24.9, PA2G4, RPL41, ZC3H10, ESYT1, MYL6B, MYL6, SMARCC2, RNF41, NABP2, SLC39A5, ANKRD52, COQ10A, CS, RP11-977G19.10, CNPY2, PAN2, IL23A, STAT2, APOF*	1.25	0.22	C
rs7090530	*IL2RA, RBM17, PFKFB3*	1.22	0.2	A
rs7221109	*CCR7, SMARCE1, KRT222, KRT222, KRT24*	1.05	0.05	C
rs72727394	*RASGRP1, C15orf53*	1.15	0.14	T

rs72928038	*BACH2*	1.20	0.18	A
rs7804356	*SKAP2, C7orf71, HOXA1, HOXA2, HOXA3, HOXA4, HOXA5, HOXA6, HOXA7, HOXA9*	1.14	0.13	T
rs7928968	*INS,IGF2, INS-IGF2,TH*	1.25	0.22	T
rs911263	*RAD51B, ZFP36L1*	1.12	0.11	T
rs924043	*WDR27, C6orf120, PHF10, TCTE3, ERMARD, DLL1, FAM120B, PSMB1, TBP, PDCD2*	1.19	0.17	C
rs9272346	*MHC*	5.58	1.72	A
rs9585056	*UBAC2, GPR18, GPR183, TM9SF2*	1.12	0.11	C
rs9924471	*SBK1, NPIPB6, EIF3CL, NPIPB7, CLN3, CLN3, APOBR, IL27, NUPR1, CCDC101, SULT1A2, SULT1A1, NPIPB8, EIF3C, NPIPB9, ATXN2L, TUFM, SH2B1, ATP2A1, RABEP2, CD19, NFATC2IP, SPNS1, LAT*	1.25	0.22	A
rs10877012	*CYP27B1*	1.22	0.2	G
rs11258747	*PRKCQ*	1.45	0.37	G
rs1465788	*RAD51B, ZFP36L1*	1.16	0.15	C
rs229541	*IL2RB, C1QTNF6, SSTR3, RAC2*	1.12	0.11	A
rs2542151	*PTPN2*	1.30	0.26	G
rs2611215	*TMEM192, KLHL2, MSMO1, CPE, TLL1*	1.19	0.17	A
rs6827756	*KIAA1109, ADAD1, IL2, IL21*	1.13	0.12	T
rs694739	*BAD,MACROD1, FLRT1, STIP1, FERMT3, TRPT1, NUDT22, DNAJC4, VEGFB, FKBP2, PPP1R14B, PLCB3, GPR137, KCNK4, TEX40, ESRRA, TRMT112, PRDX5, CCDC88B, RPS6KA4*	1.05	0.05	A
rs722988	*ITGB1, NRP1*	1.11	0.1	C
rs941576	*DLK1*	1.14	0.13	A
rs9653442	*AFF3*	1.11	0.1	C
Type 2 Diabetes SNP	**Gene**	**Odds Ratio**	**Weight**	**Effect Allele**
rs7903146	*TCF7L2*	1.40	0.34	T
rs17791513	*TLE4*	1.21	0.19	A
rs7756992	*CDKAL1*	1.20	0.18	G
rs10811661	*CDKN2A/B*	1.19	0.17	T
rs3802177	*SLC30A8*	1.16	0.15	G
rs2261181	*HMGA2*	1.16	0.15	T
rs1111875	*HHEX/IDE*	1.15	0.14	C
rs10203174	*THADA*	1.15	0.14	C
rs9936385	*FTO*	1.13	0.12	C
rs6878122	*ZBED3*	1.13	0.12	G
rs4430796	*HNF1B*	1.13	0.12	G
rs4402960	*IGF2BP*	1.13	0.12	T
rs17168486	*DGKB*	1.13	0.12	T
rs1552224	*ARAP1 (CENTD2)*	1.13	0.12	A
rs10401969	*CILP2*	1.13	0.12	C
rs849135	*JAZF1*	1.12	0.11	G
rs12427353	*HNF1A*	1.12	0.11	G
rs7593730	*RBMS1*	1.11	0.1	C
rs10830963	*MTNR1B*	1.11	0.1	G
rs7612463	*UBE2E2*	1.11	0.1	C
rs516946	*ANK1*	1.11	0.1	C
rs17106184	*FAF1*	1.11	0.1	G
rs1359790	*SPRY2*	1.11	0.1	G
rs13233731	*KLF14*	1.11	0.1	G
rs11063069	*CCND2*	1.11	0.1	G
rs10923931	*NOTCH2*	1.11	0.1	T
rs7955901	*TSPAN8*	1.09	0.09	C
rs4458523	*WFS1*	1.09	0.09	G
rs2943640	*IRS1*	1.09	0.09	C
rs243088	*BCL11A*	1.09	0.09	T
rs163184	*KCNQ1*	1.09	0.09	G
rs12899811	*PRC1*	1.09	0.09	G
rs12571751	*ZMIZ1*	1.09	0.09	A
rs11717195	*ADCY5*	1.09	0.09	T
rs11634397	*ZFAND6*	1.09	0.09	G
rs10842994	*KLHDC5*	1.09	0.09	C
rs7845219	*TP53INP1*	1.08	0.08	T
rs7178572	*HMG20A*	1.08	0.08	G
rs702634	*ARL15*	1.08	0.08	A
rs6813195	*TMEM154*	1.08	0.08	C
rs6808574	*LPP*	1.08	0.08	C
rs5215	*KCNJ11*	1.08	0.08	C
rs2075423	*PROX1*	1.08	0.08	G
rs12970134	*MC4R*	1.08	0.08	A

rs9505118	*SSR1/RREB1*	1.07	0.07	A
rs6795735	*ADAMTS9*	1.07	0.07	C
rs4812829	*HNF4A*	1.07	0.07	A
rs4275659	*MPHOSPH9*	1.07	0.07	C
rs2796441	*TLE1*	1.07	0.07	G
rs2334499	*DUSP8*	1.07	0.07	T
rs8108269	*GIPR*	1.06	0.06	G
rs7163757	*C2CD4A*	1.06	0.06	C
rs11257655	*CDC123*	1.06	0.06	T
rs7041847	*GLIS3*	1.05	0.05	A
rs459193	*ANKRD55*	1.05	0.05	G
rs10278336	*GCK*	1.05	0.05	A
rs780094	*GCKR*	1.04	0.04	C
rs3923113	*GRB14*	1.04	0.04	A
rs2028299	*AP3S2*	1.04	0.04	C
rs831571	*PSMD6*	1.03	0.03	C
rs16861329	*ST64GAL1*	1.03	0.03	C
rs7403531	*RASGRP1*	1.02	0.02	T
rs3786897	*PEPD*	1.02	0.02	A
rs1802295	*VPS26A*	1.02	0.02	T
rs9470794	*ZFAND3*	1.01	0.01	T
rs6723108	*TMEM163*	1.01	0.01	T
rs6467136	*GCC1*	1.01	0.01	A

Supp. Table 3-1. Established SNPs for GRS Calculation

RSID	Locus	Chr	T1D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T1D *P*-value
				LADA	T1D	Control			
rs9272346	MHC	6	A/G	0.686	0.818	0.579	1.455 [1.427-1.483]	9.64E-11	1.26E-17
rs6679677	*PTPN22*	1	A/C	0.143	0.17	0.093	1.469 [1.427-1.510]	6.38E-06	2.61E-02
rs17696736	*SH2B3*	12	G/A	0.515	0.503	0.44	1.277 [1.250-1.304]	1.10E-05	5.42E-01
rs689	*INS,IGF2*	11	T/A	0.796	0.741	0.73	1.265 [1.234-1.296]	2.39E-04	3.88E-04
rs597325	*BACH2*	6	G/A	0.646	0.647	0.583	1.192 [1.164-1.220]	2.11E-03	5.06E-01
rs7804356	*SKAP2*	7	T/C	0.791	0.768	0.751	1.217 [1.185-1.249]	2.77E-03	5.87E-02
rs11954020	*IL7R*	5	G/C	0.438	0.415	0.387	1.174 [1.146-1.202]	5.44E-03	2.03E-01
rs4505848	*KIAA1109*	4	G/A	0.36	0.363	0.321	1.160 [1.131-1.189]	1.23E-02	7.92E-01
rs9585056	*UBAC2*	13	C/T	0.246	0.266	0.22	1.174 [1.142-1.206]	1.54E-02	3.44E-02
rs11755527	*BACH2*	6	G/C	0.481	0.5	0.43	1.139 [1.112-1.166]	1.75E-02	2.36E-01
rs229541	*IL2RB*	22	A/G	0.436	0.468	0.412	1.136 [1.108-1.164]	2.81E-02	6.70E-02
rs1538171	*CENPW*	6	G/C	0.514	0.475	0.472	1.127 [1.099-1.155]	3.64E-02	1.38E-01
rs402072	*DACT3*	19	T/C	0.861	0.862	0.841	1.173 [1.134-1.212]	4.41E-02	6.13E-01
rs12908309	*RASGRP1*	15	G/A	0.779	0.78	0.76	1.142 [1.107-1.176]	5.96E-02	3.07E-01
rs9653442	*AFF3*	2	C/T	0.485	0.502	0.454	1.109 [1.082-1.137]	6.53E-02	2.71E-01
rs7928968	*INS,IGF2*	11	T/A	0.259	0.267	0.232	1.144 [1.108-1.180]	6.72E-02	4.27E-01
rs72928038	*BACH2*	6	A/G	0.193	0.21	0.168	1.162 [1.122-1.202]	6.74E-02	9.66E-03
rs601338	*FUT2*	19	A/G	0.451	0.456	0.507	0.905 [0.878-0.932]	7.33E-02	8.10E-01
rs6476839	*GLIS3*	9	T/A	0.448	0.435	0.413	1.096 [1.067-1.124]	1.14E-01	5.32E-01
rs911263	*RAD51B*	14	T/C	0.706	0.74	0.691	1.091 [1.062-1.121]	1.49E-01	2.38E-01
rs602662	*FUT2*	19	G/A	0.523	0.518	0.513	1.084 [1.057-1.111]	1.49E-01	5.72E-01
rs2542151	*PTPN2*	18	G/T	0.148	0.202	0.166	0.897 [0.860-0.934]	1.53E-01	2.58E-03
rs4763879	*CD69*	12	A/G	0.397	0.392	0.376	1.085 [1.056-1.113]	1.58E-01	8.13E-01
rs694739	*BAD,*	11	A/G	0.64	0.639	0.608	1.085 [1.056-1.113]	1.64E-01	7.21E-01
rs941576	*DLK1*	14	A/G	0.574	0.602	0.56	1.084 [1.055-1.112]	1.64E-01	3.81E-01
rs10509540	*RNLS*	10	T/C	0.741	0.75	0.716	1.086 [1.056-1.116]	1.76E-01	7.69E-01
rs4849135	*ACOXL*	2	G/T	0.73	0.727	0.717	1.076 [1.045-1.107]	2.42E-01	9.59E-01
rs10877012	*CYP27B1*	12	T/G	0.336	0.313	0.311	1.071 [1.042-1.101]	2.48E-01	2.91E-01
rs3024493	*MAPKAPK2*	1	C/A	0.855	0.856	0.839	1.088 [1.050-1.125]	2.73E-01	7.29E-01
rs4948088	*COBL*	7	C/A	0.965	0.962	0.958	1.158 [1.087-1.229]	3.11E-01	8.67E-01
rs11571316	*CTLA4*	2	G/A	0.602	0.623	0.588	1.060 [1.032-1.088]	3.14E-01	2.43E-01
rs7111341	*INS,IGF2*	11	C/T	0.757	0.75	0.73	1.063 [1.032-1.094]	3.35E-01	6.90E-01
rs72727394	*RASGRP1*	15	T/C	0.195	0.223	0.185	1.071 [1.036-1.107]	3.39E-01	5.47E-02
rs6691977	*KIF14*	1	C/T	0.213	0.219	0.207	1.065 [1.032-1.098]	3.56E-01	5.00E-01
rs1456988	*C14orf64*	14	G/T	0.276	0.301	0.26	1.060 [1.029-1.091]	3.57E-01	3.37E-01
rs1465788	*RAD51B*	14	C/T	0.735	0.733	0.728	1.059 [1.027-1.090]	3.70E-01	8.25E-01
rs1615504	*DOK6*	18	T/C	0.492	0.504	0.473	1.051 [1.023-1.080]	3.90E-01	3.12E-01
rs7221109	*CCR7*	17	C/T	0.621	0.687	0.632	0.954 [0.925-0.983]	4.23E-01	6.54E-04
rs12148472	*CHRNB4*	15	T/C	0.879	0.899	0.884	0.935 [0.893-0.978]	4.40E-01	1.64E-02
rs1738074	*RSPH3*	6	C/T	0.607	0.568	0.585	1.045 [1.017-1.073]	4.43E-01	2.67E-02
rs6920220	*TNFAIP3*	6	A/G	0.22	0.248	0.205	1.053 [1.019-1.087]	4.53E-01	1.21E-01
rs1990760	*GCG*	2	T/C	0.633	0.651	0.618	1.043 [1.015-1.071]	4.59E-01	6.66E-01
rs2611215	*TMEM192*	4	A/G	0.175	0.173	0.165	1.054 [1.018-1.091]	4.81E-01	7.07E-01
rs61839660	*IL2RA*	10	C/T	0.91	0.923	0.899	1.064 [1.017-1.111]	5.19E-01	2.09E-03
rs7090530	*IL2RA*	10	A/C	0.605	0.645	0.611	0.964 [0.936-0.993]	5.34E-01	1.12E-02
rs2281808	*SIRPD*	20	C/T	0.677	0.666	0.659	1.037 [1.007-1.067]	5.47E-01	5.34E-01
rs722988	*ITGB1*	10	C/T	0.391	0.401	0.387	1.035 [1.006-1.063]	5.55E-01	6.23E-01
rs113010081	*CCR5*	3	T/C	0.887	0.901	0.882	1.052 [1.008-1.097]	5.77E-01	2.05E-01
rs5753037	*NF2*	22	T/C	0.381	0.415	0.375	1.032 [1.004-1.060]	5.83E-01	4.47E-01
rs2304256	*ICAM1*	19	C/A	0.729	0.73	0.722	1.033 [1.003-1.064]	5.98E-01	8.94E-01
rs11203203	*UBASH3A*	21	A/G	0.389	0.401	0.371	1.031 [1.002-1.059]	6.01E-01	3.78E-01
rs705705	*IKZF*	12	C/G	0.341	0.386	0.331	1.031 [1.002-1.061]	6.12E-01	1.82E-01
rs478222	*DNMT3A*	2	A/T	0.564	0.628	0.581	0.971 [0.943-1.000]	6.17E-01	5.05E-02
rs11171710	*IKZF4*	12	A/G	0.458	0.397	0.447	1.025 [0.998-1.053]	6.58E-01	1.90E-02
rs7202877	*CTRB2*	16	G/T	0.111	0.115	0.106	1.038 [0.995-1.082]	6.72E-01	2.50E-01
rs11258747	*PRKCQ*	10	G/T	0.77	0.761	0.77	1.028 [0.996-1.061]	6.73E-01	6.77E-01
rs12708716	*CLEC16A*	16	A/G	0.637	0.702	0.633	1.024 [0.996-1.052]	6.84E-01	5.55E-03
rs10795791	*IL2RA*	10	G/A	0.434	0.456	0.418	1.022 [0.994-1.049]	7.06E-01	4.83E-02
rs3825932	*CHRNB4*	15	T/C	0.311	0.29	0.309	1.023 [0.993-1.053]	7.06E-01	2.23E-01
rs10517086	*SLC34A2*	4	A/G	0.303	0.326	0.305	1.021 [0.991-1.050]	7.33E-01	4.80E-01
rs10492166	*CD69*	12	G/A	0.508	0.549	0.506	1.019 [0.991-1.046]	7.40E-01	2.10E-02
rs9924471	*SBK1*	16	A/G	0.18	0.18	0.176	1.019 [0.983-1.055]	7.97E-01	7.30E-01
rs193778	*CLEC16A*	16	G/A	0.239	0.284	0.244	1.015 [0.983-1.047]	8.17E-01	5.64E-02
rs11170466	*ITGB7*	12	T/C	0.06	0.056	0.058	0.978 [0.921-1.035]	8.46E-01	4.86E-01
rs924043	*WDR27*	6	C/T	0.852	0.867	0.856	0.989 [0.949-1.029]	8.94E-01	3.16E-01
rs12453507	*FBXL20*	17	G/C	0.524	0.541	0.522	1.005 [0.977-1.034]	9.28E-01	8.88E-01
rs10272724	*IKZF1*	7	T/C	0.73	0.749	0.733	1.000 [0.969-1.031]	9.99E-01	6.14E-02

Supp. Table 3-2. LADA (n= 978) vs T1D and BMDCS controls testing T1D candidae loci

RSID	Locus	Chr	T1D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T1D *P*-value
				LADA	T1D	Control			
rs9272346	MHC	6	A/G	0.651	0.818	0.579	1.296 [1.266-1.326]	6.84E-05	1.99E-24
rs17696736	*SH2B3*	12	G/A	0.503	0.503	0.44	1.226 [1.197-1.256]	1.23E-03	0.99
rs6679677	*PTPN22*	1	A/C	0.131	0.170	0.093	1.383 [1.336-1.429]	1.27E-03	0.01
rs9585056	*UBAC2*	13	C/T	0.256	0.266	0.22	1.244 [1.209-1.279]	3.48E-03	0.25
rs597325	*BACH2*	6	G/A	0.646	0.647	0.583	1.207 [1.177-1.237]	3.67E-03	0.56
rs11954020	*IL7R*	5	G/C	0.444	0.415	0.387	1.191 [1.160-1.221]	0.01	0.17
rs689	*INS,IGF2*	11	T/A	0.782	0.741	0.73	1.208 [1.175-1.241]	0.01	0.01
rs229541	*IL2RB*	22	A/G	0.446	0.468	0.412	1.169 [1.138-1.199]	0.02	0.14
rs4505848	*KIAA1109*	4	G/A	0.363	0.363	0.321	1.170 [1.139-1.201]	0.02	0.79
rs1538171	*CENPW*	6	G/C	0.518	0.475	0.472	1.156 [1.126-1.186]	0.02	0.15
rs7804356	*SKAP2*	7	T/C	0.781	0.768	0.751	1.166 [1.132-1.201]	0.04	0.31
rs402072	*DACT3*	19	T/C	0.861	0.862	0.841	1.185 [1.143-1.226]	0.06	0.59
rs601338	*FUT2*	19	A/G	0.446	0.456	0.493	0.888 [0.858-0.918]	0.06	0.95
rs11755527	*BACH2*	6	G/C	0.474	0.5	0.43	1.119 [1.090-1.148]	0.07	0.14
rs10877012	*CYP27B1*	12	T/G	0.344	0.313	0.311	1.120 [1.089-1.152]	0.09	0.17
rs12908309	*RASGRP1*	15	G/A	0.779	0.78	0.76	1.137 [1.100-1.175]	0.11	0.43
rs602662	*FUT2*	19	G/A	0.528	0.518	0.487	1.104 [1.074-1.133]	0.12	0.74
rs2542151	*PTPN2*	18	G/T	0.149	0.202	0.166	0.887 [0.847-0.927]	0.16	1.20E-03
rs3024493	*MAPKAPK2*	1	C/A	0.857	0.856	0.839	1.118 [1.077-1.158]	0.20	0.84
rs7928968	*INS,IGF2*	11	T/A	0.251	0.267	0.232	1.112 [1.073-1.152]	0.21	0.22
rs941576	*DLK1*	14	A/G	0.574	0.602	0.56	1.085 [1.055-1.116]	0.21	0.39
rs2611215	*TMEM192*	4	A/G	0.187	0.173	0.165	1.110 [1.070-1.149]	0.22	0.42
rs12148472	*CHRNB4*	15	T/C	0.873	0.899	0.884	0.887 [0.842-0.933]	0.22	2.36E-03
rs4763879	*CD69*	12	A/G	0.394	0.392	0.376	1.078 [1.048-1.108]	0.25	0.81
rs4948088	*COBL*	7	C/A	0.967	0.962	0.958	1.206 [1.130-1.283]	0.25	0.83
rs7090530	*IL2RA*	10	A/C	0.591	0.645	0.611	0.927 [0.897-0.958]	0.26	2.88E-03
rs72928038	*BACH2*	6	A/G	0.184	0.21	0.168	1.110 [1.066-1.153]	0.27	3.33E-03
rs6476839	*GLIS3*	9	T/A	0.44	0.435	0.413	1.070 [1.040-1.101]	0.30	0.94
rs10509540	*RNLS*	10	T/C	0.738	0.75	0.716	1.073 [1.041-1.105]	0.31	0.63
rs5753037	*NF2*	22	T/C	0.386	0.415	0.375	1.063 [1.033-1.093]	0.34	0.82
rs1456988	*C14orf64*	14	G/T	0.274	0.301	0.26	1.067 [1.033-1.100]	0.37	0.24
rs9653442	*AFF3*	2	C/T	0.472	0.502	0.454	1.057 [1.028-1.087]	0.38	0.11
rs6920220	*TNFAIP3*	6	A/G	0.22	0.248	0.205	1.061 [1.024-1.097]	0.45	0.12
rs11171710	*IKZF4*	12	A/G	0.466	0.397	0.447	1.049 [1.019-1.079]	0.45	4.31E-03
rs4849135	*ACOXL*	2	G/T	0.724	0.727	0.717	1.054 [1.021-1.087]	0.46	1.00
rs694739	*BAD*	11	A/G	0.626	0.639	0.608	1.049 [1.018-1.079]	0.47	0.32
rs1465788	*RAD51B*	14	C/T	0.734	0.733	0.728	1.052 [1.019-1.085]	0.48	0.96
rs10492166	*CD69*	12	G/A	0.515	0.549	0.506	1.044 [1.015-1.074]	0.49	0.13
rs722988	*ITGB1*	10	C/T	0.396	0.401	0.387	1.041 [1.010-1.072]	0.54	0.75
rs11258747	*PRKCQ*	10	G/T	0.774	0.761	0.77	1.042 [1.007-1.077]	0.58	0.47
rs924043	*WDR27*	6	C/T	0.845	0.867	0.856	0.951 [0.908-0.994]	0.59	0.13
rs72727394	*RASGRP1*	15	T/C	0.194	0.223	0.185	1.045 [1.007-1.083]	0.59	0.05
rs2281808	*SIRPD*	20	C/T	0.656	0.666	0.659	0.965 [0.934-0.997]	0.60	0.66
rs7202877	*CTRB2*	16	G/T	0.114	0.115	0.106	1.053 [1.006-1.100]	0.61	0.31
rs7111341	*INS,IGF2*	11	C/T	0.731	0.75	0.73	0.970 [0.937-1.003]	0.67	0.23
rs113010081	*CCR5*	3	T/C	0.888	0.901	0.882	1.044 [0.996-1.092]	0.68	0.15
rs1990760	*GCG*	2	T/C	0.629	0.651	0.618	1.027 [0.997-1.057]	0.68	0.67
rs6691977	*KIF14*	1	C/T	0.208	0.219	0.207	1.030 [0.994-1.066]	0.71	0.31
rs10795791	*IL2RA*	10	G/A	0.418	0.456	0.418	0.976 [0.946-1.006]	0.71	0.01
rs1738074	*RSPH3*	6	C/T	0.601	0.568	0.585	1.024 [0.993-1.054]	0.72	0.08
rs1615504	*DOK6*	18	T/C	0.488	0.504	0.473	1.023 [0.993-1.054]	0.73	0.22
rs61839660	*IL2RA*	10	C/T	0.905	0.923	0.899	1.037 [0.987-1.087]	0.74	3.47E-03
rs193778	*CLEC16A*	16	G/A	0.236	0.284	0.244	1.023 [0.989-1.058]	0.76	0.11
rs10272724	*IKZF1*	7	T/C	0.726	0.749	0.733	0.979 [0.946-1.013]	0.77	0.06
rs911263	*RAD51B*	14	T/C	0.69	0.74	0.691	1.016 [0.984-1.048]	0.82	0.05
rs3825932	*CHRNB4*	15	T/C	0.309	0.29	0.309	1.016 [0.984-1.049]	0.82	0.24
rs11203203	*UBASH3A*	21	A/G	0.385	0.401	0.371	1.013 [0.983-1.044]	0.84	0.38
rs7221109	*CCR7*	17	C/T	0.624	0.687	0.632	0.989 [0.958-1.021]	0.87	4.39E-03
rs2304256	*ICAM1*	19	C/A	0.724	0.73	0.722	1.009 [0.976-1.041]	0.90	0.74
rs12453507	*FBXL20*	17	G/C	0.517	0.541	0.522	0.992 [0.962-1.023]	0.91	0.73
rs9924471	*SBK1*	16	A/G	0.174	0.18	0.824	0.992 [0.953-1.031]	0.92	0.48
rs705705	*IKZF4*	12	C/G	0.331	0.386	0.331	0.994 [0.962-1.026]	0.93	0.09
rs10517086	*SLC34A2*	4	A/G	0.3	0.326	0.305	1.006 [0.974-1.038]	0.93	0.20
rs12708716	*CLEC16A*	16	A/G	0.632	0.702	0.633	1.004 [0.973-1.034]	0.96	0.01
rs11571316	*CTLA4*	2	G/A	0.587	0.623	0.588	1.002 [0.972-1.032]	0.97	0.10
rs478222	*DNMT3A*	2	A/T	0.574	0.628	0.581	1.002 [0.971-1.033]	0.98	0.11
rs11170466	*ITGB7*	12	T/C	0.061	0.056	0.058	0.998 [0.936-1.060]	0.99	0.38

Supp. Table 3-3. LADA cases positive for GADA only (n= 669) vs T1D and BMDCS controls testing T1D candidate loci

RSID	Locus	Chr	T1D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T1D *P*-value
				LADA	T1D	Control			
rs9272346	MHC	6	A/G	0.763	0.818	0.579	1.983 [1.954-2.012]	1.20E-15	4.01E-03
rs6679677	*PTPN22*	1	A/C	0.17	0.17	0.093	1.864 [1.819-1.909]	2.19E-06	6.03E-01
rs17696736	*SH2B3*	12	G/A	0.542	0.503	0.44	1.481 [1.452-1.511]	5.93E-06	1.80E-01
rs689	*INS,IGF2*	11	T/A	0.824	0.741	0.73	1.440 [1.407-1.474]	1.90E-04	1.68E-06
rs7111341	*INS,IGF2*	11	C/T	0.812	0.75	0.73	1.360 [1.327-1.394]	1.82E-03	2.39E-04
rs7804356	*SKAP2*	7	T/C	0.814	0.768	0.751	1.354 [1.320-1.388]	2.27E-03	1.22E-02
rs2281808	*SIRPD*	20	C/T	0.723	0.666	0.659	1.276 [1.244-1.308]	9.17E-03	9.53E-03
rs694739	*BAD*	11	A/G	0.672	0.639	0.608	1.253 [1.223-1.283]	1.06E-02	3.07E-01
rs911263	*RAD51B*	14	T/C	0.739	0.74	0.691	1.265 [1.233-1.297]	1.09E-02	6.27E-01
rs72928038	*BACH2*	6	A/G	0.213	0.21	0.168	1.367 [1.324-1.410]	1.31E-02	4.29E-01
rs9653442	*AFF3*	2	C/T	0.515	0.502	0.454	1.235 [1.205-1.264]	1.37E-02	5.59E-01
rs11571316	*CTLA4*	2	G/A	0.634	0.623	0.588	1.238 [1.208-1.268]	1.48E-02	5.03E-01
rs11755527	*BACH2*	6	G/C	0.495	0.5	0.43	1.229 [1.200-1.258]	1.54E-02	6.10E-01
rs7928968	*INS,IGF2*	11	T/A	0.276	0.267	0.232	1.292 [1.253-1.331]	2.32E-02	5.67E-01
rs597325	*BACH2*	6	G/A	0.646	0.647	0.583	1.216 [1.186-1.246]	2.68E-02	6.67E-01
rs10795791	*IL2RA*	10	G/A	0.469	0.456	0.418	1.200 [1.171-1.229]	3.39E-02	6.80E-01
rs6476839	*GLIS3*	9	T/A	0.465	0.435	0.413	1.188 [1.158-1.219]	5.29E-02	2.04E-01
rs4505848	*KIAA1109*	4	G/A	0.354	0.363	0.321	1.161 [1.130-1.192]	1.01E-01	9.74E-01
rs705705	*IKZF4*	12	C/G	0.364	0.386	0.331	1.158 [1.126-1.190]	1.14E-01	7.76E-01
rs11954020	*IL7R*	5	G/C	0.424	0.415	0.387	1.149 [1.118-1.179]	1.18E-01	5.20E-01
rs4763879	*CD69*	12	A/G	0.405	0.392	0.376	1.139 [1.108-1.169]	1.48E-01	7.25E-01
rs61839660	*IL2RA*	10	C/T	0.921	0.923	0.899	1.223 [1.173-1.273]	1.66E-01	2.66E-01
rs1738074	*RSPH3*	6	C/T	0.62	0.568	0.585	1.119 [1.089-1.148]	1.93E-01	1.26E-02
rs402072	*DACT3*	19	T/C	0.861	0.862	0.841	1.168 [1.127-1.209]	1.93E-01	7.25E-01
rs10509540	*RNLS*	10	T/C	0.748	0.75	0.716	1.123 [1.090-1.155]	2.19E-01	9.90E-01
rs478222	*DNMT3A*	2	A/T	0.542	0.628	0.581	0.896 [0.865-0.927]	2.24E-01	1.32E-02
rs7221109	*CCR7*	17	C/T	0.615	0.687	0.632	0.898 [0.867-0.928]	2.27E-01	1.96E-03
rs4849135	*ACOXL*	2	G/T	0.744	0.727	0.717	1.122 [1.090-1.155]	2.28E-01	6.52E-01
rs2542151	*PTPN2*	18	G/T	0.147	0.202	0.166	0.877 [0.838-0.916]	2.50E-01	4.07E-02
rs1615504	*DOK6*	18	T/C	0.499	0.504	0.473	1.106 [1.076-1.136]	2.50E-01	6.01E-01
rs6691977	*KIF14*	1	C/T	0.223	0.219	0.207	1.124 [1.089-1.160]	2.58E-01	8.94E-01
rs9924471	*SBK1*	16	A/G	0.193	0.18	0.176	1.133 [1.095-1.172]	2.65E-01	7.40E-01
rs1990760	*GCG*	2	T/C	0.642	0.651	0.618	1.100 [1.070-1.130]	2.69E-01	7.60E-01
rs12908309	*RASGRP1*	15	G/A	0.778	0.78	0.76	1.124 [1.087-1.160]	2.69E-01	3.31E-01
rs1538171	*CENPW*	6	G/C	0.505	0.475	0.472	1.099 [1.069-1.128]	2.76E-01	4.97E-01
rs2304256	*ICAM1*	19	C/A	0.739	0.73	0.722	1.107 [1.074-1.140]	2.92E-01	3.85E-01
rs12708716	*CLEC16A*	16	A/G	0.647	0.702	0.633	1.093 [1.063-1.124]	3.15E-01	9.47E-02
rs7090530	*IL2RA*	10	A/C	0.636	0.645	0.611	1.089 [1.059-1.119]	3.32E-01	9.76E-01
rs1456988	*C14orf64*	14	G/T	0.278	0.301	0.26	1.097 [1.065-1.130]	3.32E-01	7.46E-01
rs2611215	*TMEM192*	4	A/G	0.149	0.173	0.165	0.898 [0.858-0.938]	3.57E-01	2.38E-01
rs72727394	*RASGRP1*	15	T/C	0.198	0.223	0.185	1.103 [1.065-1.140]	3.75E-01	1.82E-01
rs11203203	*UBASH3A*	21	A/G	0.396	0.401	0.371	1.079 [1.049-1.109]	3.82E-01	9.08E-01
rs941576	*DLK1*	14	A/G	0.574	0.602	0.56	1.079 [1.049-1.109]	3.84E-01	6.79E-01
rs601338	*FUT2*	19	A/G	0.461	0.456	0.493	0.935 [0.906-0.964]	4.30E-01	7.86E-01
rs3024493	*MAPKAPK2*	1	C/A	0.85	0.856	0.839	1.087 [1.047-1.126]	4.73E-01	7.16E-01
rs924043	*WDR27*	6	C/T	0.866	0.867	0.856	1.084 [1.041-1.127]	5.19E-01	7.56E-01
rs229541	*IL2RB*	22	A/G	0.414	0.468	0.412	1.057 [1.026-1.089]	5.39E-01	2.07E-02
rs12148472	*CHRNB4*	15	T/C	0.893	0.899	0.884	1.085 [1.038-1.131]	5.48E-01	8.32E-01
rs10517086	*SLC34A2*	4	A/G	0.309	0.326	0.305	1.054 [1.023-1.085]	5.61E-01	9.95E-01
rs1465788	*RAD51B*	14	C/T	0.736	0.733	0.728	1.056 [1.023-1.090]	5.76E-01	9.36E-01
rs11171710	*IKZF4*	12	A/G	0.44	0.397	0.447	0.953 [0.923-0.983]	5.78E-01	3.13E-01
rs6920220	*TNFAIP3*	6	A/G	0.218	0.248	0.205	1.059 [1.023-1.095]	5.85E-01	2.27E-01
rs602662	*FUT2*	19	G/A	0.513	0.518	0.487	1.045 [1.016-1.074]	6.04E-01	6.69E-01
rs12453507	*ERBB2*	17	G/C	0.538	0.541	0.522	1.041 [1.011-1.072]	6.48E-01	7.67E-01
rs9585056	*UBAC2*	13	C/T	0.227	0.266	0.22	1.045 [1.010-1.080]	6.68E-01	2.18E-02
rs3825932	*CHRNB4*	15	T/C	0.314	0.29	0.309	1.041 [1.009-1.074]	6.70E-01	3.32E-01
rs10272724	*IKZF1*	7	T/C	0.738	0.749	0.733	1.025 [0.992-1.058]	8.00E-01	3.56E-01
rs193778	*CLEC16A*	16	G/A	0.244	0.284	0.244	1.024 [0.990-1.059]	8.10E-01	1.62E-01
rs10492166	*CD69*	12	G/A	0.49	0.549	0.506	0.981 [0.951-1.010]	8.20E-01	1.82E-02
rs11170466	*ITGB7*	12	T/C	0.06	0.056	0.058	0.965 [0.904-1.026]	8.42E-01	9.89E-01
rs722988	*ITGB1*	10	C/T	0.379	0.401	0.387	1.016 [0.985-1.046]	8.59E-01	5.64E-01
rs11258747	*PRKCQ*	10	G/T	0.764	0.761	0.77	0.989 [0.954-1.024]	9.11E-01	8.21E-01
rs10877012	*CYP27B1*	12	T/G	0.32	0.313	0.311	1.007 [0.976-1.038]	9.39E-01	5.86E-01
rs5753037	*NF2*	22	T/C	0.371	0.415	0.375	0.994 [0.964-1.024]	9.45E-01	2.26E-01
rs4948088	*COBL*	7	C/A	0.96	0.962	0.958	1.007 [0.933-1.081]	9.74E-01	3.03E-01
rs113010081	*CCR5*	3	T/C	0.886	0.901	0.882	0.998 [0.951-1.046]	9.91E-01	1.85E-01
rs7202877	*CTRB2*	16	G/T	0.105	0.115	0.106	1.001 [0.954-1.048]	9.94E-01	3.51E-01

Supp. Table 3-4. LADA cases positive for GADA and IA2A (n= 309) vs T1D and BMDCS controls testing T1D candidate loci

RSID	Locus	Chr	T2D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T2D *P*-value
				LADA	T2D	Control			
rs12427353	*HNF1A*	12	G/C	0.831	0.828	0.787	1.291 [1.256-1.326]	3.42E-04	5.38E-01
rs849135	*JAZF1*	7	G/A	0.526	0.529	0.485	1.165 [1.137-1.193]	7.99E-03	8.10E-01
rs6723108	*TMEM163*	2	T/G	0.617	0.521	0.553	1.149 [1.121-1.177]	1.58E-02	2.21E-03
rs11634397	*ZFAND6*	15	G/A	0.683	0.679	0.653	1.152 [1.123-1.181]	1.67E-02	9.25E-01
rs8108269	*GIPR*	19	G/T	0.338	0.31	0.292	1.153 [1.123-1.183]	1.96E-02	1.70E-01
rs7178572	*HMG20A*	15	G/A	0.702	0.73	0.73	0.870 [0.839-0.900]	2.58E-02	8.38E-02
rs12970134	*MC4R*	18	A/G	0.28	0.291	0.247	1.151 [1.120-1.181]	2.59E-02	5.59E-01
rs16861329	*ST64GAL1*	3	C/T	0.882	0.872	0.861	1.186 [1.145-1.226]	3.92E-02	3.61E-01
rs1111875	*HHEX/IDE*	10	C/T	0.625	0.634	0.59	1.115 [1.087-1.142]	5.36E-02	3.24E-01
rs17106184	*FAF1*	1	G/A	0.91	0.918	0.895	1.191 [1.146-1.237]	6.00E-02	6.55E-01
rs2943640	*IRS1*	2	C/A	0.662	0.675	0.634	1.118 [1.089-1.148]	6.58E-02	8.53E-01
rs163184	*KCNQ1*	11	G/T	0.497	0.49	0.463	1.104 [1.077-1.132]	8.09E-02	6.43E-01
rs10923931	*NOTCH2*	1	T/G	0.117	0.117	0.106	1.161 [1.119-1.204]	8.48E-02	9.67E-01
rs6813195	*TMEM154*	4	C/T	0.74	0.737	0.718	1.110 [1.079-1.141]	1.03E-01	6.24E-01
rs7041847	*GLIS3*	9	A/G	0.535	0.516	0.505	1.097 [1.068-1.125]	1.12E-01	2.24E-01
rs5215	*KCNJ11*	11	C/T	0.382	0.382	0.362	1.097 [1.068-1.125]	1.15E-01	4.90E-01
rs17168486	*DGKB*	7	T/C	0.187	0.172	0.163	1.120 [1.085-1.156]	1.17E-01	3.41E-01
rs4430796	*HNF1B*	17	G/A	0.512	0.479	0.484	1.092 [1.064-1.121]	1.26E-01	2.51E-02
rs10811661	*CDKN2A/B*	9	T/C	0.851	0.858	0.831	1.120 [1.082-1.157]	1.39E-01	6.04E-01
rs10203174	*THADA*	2	C/T	0.887	0.897	0.871	1.136 [1.094-1.178]	1.40E-01	9.93E-02
rs11063069	*CCND2*	12	G/A	0.21	0.223	0.221	0.911 [0.878-0.944]	1.65E-01	2.69E-01
rs10886471	*GRK5*	10	T/C	0.469	0.477	0.489	0.927 [0.899-0.954]	1.78E-01	3.66E-01
rs11717195	*ADCY5*	3	T/C	0.783	0.777	0.77	1.096 [1.062-1.130]	1.84E-01	3.07E-01
rs391300	*SRR*	17	C/T	0.632	0.638	0.61	1.079 [1.051-1.107]	1.88E-01	7.14E-01
rs780094	*GCKR*	2	C/T	0.571	0.618	0.6	0.928 [0.900-0.956]	1.88E-01	3.15E-02
rs7403531	*RASGRP1*	15	T/C	0.221	0.229	0.204	1.091 [1.058-1.124]	1.97E-01	3.29E-01
rs1802295	*VPS26A*	10	T/C	0.302	0.321	0.322	0.925 [0.895-0.956]	2.12E-01	1.21E-01
rs17584499	*PTPRD*	9	T/C	0.212	0.17	0.193	1.084 [1.051-1.118]	2.35E-01	1.54E-03
rs516946	*ANK1*	8	C/T	0.767	0.782	0.754	1.080 [1.048-1.112]	2.42E-01	7.69E-02
rs3786897	*PEPD*	19	A/G	0.608	0.578	0.589	1.070 [1.041-1.098]	2.48E-01	6.75E-02
rs2261181	*HMGA2*	12	T/C	0.115	0.109	0.101	1.111 [1.066-1.155]	2.54E-01	9.26E-01
rs4275659	*MPHOSPH9*	12	C/T	0.715	0.735	0.694	1.072 [1.042-1.102]	2.55E-01	8.06E-01
rs3923113	*GRB14*	2	A/C	0.646	0.656	0.632	1.062 [1.034-1.091]	3.03E-01	1.76E-01
rs7955901	*TSPAN8*	12	C/T	0.461	0.487	0.438	1.058 [1.031-1.086]	3.11E-01	1.29E-01
rs3802177	*SLC30A8*	8	G/A	0.684	0.706	0.699	0.941 [0.911-0.971]	3.23E-01	5.96E-02
rs4458523	*WFS1*	4	G/T	0.611	0.614	0.59	1.056 [1.028-1.084]	3.36E-01	9.89E-01
rs7163757	*C2CD4A*	15	C/T	0.573	0.59	0.588	0.950 [0.923-0.978]	3.65E-01	3.76E-02
rs6878122	*ZBED3*	5	A/G	0.718	0.658	0.706	1.051 [1.021-1.081]	4.20E-01	8.81E-04
rs17791513	*TLE4*	9	A/G	0.946	0.952	0.939	1.096 [1.037-1.154]	4.45E-01	6.81E-01
rs459193	*ANKRD55*	5	G/A	0.753	0.771	0.741	1.050 [1.018-1.082]	4.55E-01	1.43E-01
rs13233731	*KLF14*	7	G/A	0.544	0.538	0.527	1.043 [1.015-1.072]	4.67E-01	9.79E-01
rs7593730	*RBMS1*	2	C/T	0.793	0.811	0.796	1.052 [1.018-1.085]	4.68E-01	4.54E-01
rs2075423	*PROX1*	1	G/T	0.643	0.678	0.653	0.959 [0.930-0.988]	4.75E-01	2.41E-01
rs10842994	*KLHDC5*	12	C/T	0.803	0.814	0.794	1.050 [1.016-1.083]	4.84E-01	4.23E-01
rs1359790	*SPRY2*	13	G/A	0.746	0.741	0.736	1.043 [1.011-1.075]	5.16E-01	5.07E-01
rs243088	*BCL11A*	2	T/A	0.494	0.492	0.485	1.036 [1.008-1.064]	5.32E-01	4.08E-01
rs10401969	*CILP2*	19	C/T	0.077	0.088	0.067	1.069 [1.016-1.122]	5.37E-01	7.38E-02
rs11257655	*CDC123*	10	T/C	0.194	0.236	0.213	0.962 [0.927-0.997]	5.85E-01	9.13E-02
rs7756992	*CDKAL1*	6	G/A	0.306	0.316	0.289	1.033 [1.003-1.062]	5.96E-01	6.09E-01
rs2796441	*TLE1*	9	G/A	0.607	0.582	0.601	1.031 [1.003-1.058]	5.97E-01	4.85E-02
rs7202877	*BCAR1*	16	T/G	0.889	0.914	0.894	0.963 [0.920-1.007]	6.72E-01	4.14E-02

rs4812829	*HNF4A*	20	G/A	0.836	0.837	0.831	1.033 [0.995-1.070]	6.74E-01	3.23E-01
rs7612463	*UBE2E2*	3	C/A	0.887	0.894	0.89	0.963 [0.918-1.007]	6.75E-01	1.19E-01
rs10830963	*MTNR1B*	11	G/C	0.292	0.27	0.29	0.974 [0.944-1.005]	6.76E-01	2.28E-01
rs10278336	*GCK*	7	A/G	0.58	0.577	0.579	1.024 [0.996-1.053]	6.82E-01	3.21E-01
rs12899811	*PRC1*	15	G/A	0.324	0.315	0.306	1.024 [0.994-1.053]	6.98E-01	9.54E-01
rs7903146	*TCF7L2*	10	T/C	0.295	0.376	0.298	1.023 [0.994-1.053]	7.02E-01	5.21E-06
rs9936385	*FTO*	16	C/T	0.4	0.454	0.394	1.022 [0.994-1.050]	7.09E-01	1.80E-03
rs6795735	*ADAMTS9*	3	C/T	0.571	0.618	0.575	0.982 [0.955-1.009]	7.45E-01	1.26E-02
rs9505118	*SSR1/RREB1*	6	A/G	0.595	0.612	0.592	1.018 [0.989-1.046]	7.60E-01	1.91E-01
rs831571	*PSMD6*	3	C/T	0.829	0.81	0.827	1.023 [0.986-1.059]	7.64E-01	1.74E-01
rs7845219	*TP53INP1*	8	T/C	0.515	0.525	0.515	0.983 [0.955-1.011]	7.65E-01	3.86E-01
rs4402960	*IGF2BP*	3	T/G	0.305	0.351	0.314	0.983 [0.952-1.013]	7.82E-01	1.01E-02
rs12571751	*ZMIZ1*	10	A/G	0.544	0.583	0.549	0.988 [0.960-1.015]	8.27E-01	1.02E-02
rs2334499	*DUSP8*	11	T/C	0.409	0.418	0.41	1.011 [0.983-1.039]	8.52E-01	1.93E-01
rs6467136	*GCC1*	7	A/G	0.479	0.472	0.477	1.008 [0.981-1.036]	8.85E-01	7.75E-01
rs9470794	*ZFAND3*	6	T/C	0.908	0.911	0.908	1.014 [0.966-1.061]	8.89E-01	3.93E-01
rs2028299	*AP3S2*	15	C/A	0.284	0.297	0.291	0.995 [0.965-1.026]	9.42E-01	7.89E-01
rs1552224	*ARAP1 (CENTD2)*	11	A/C	0.851	0.849	0.85	1.005 [0.967-1.043]	9.45E-01	8.96E-01
rs702634	*ARL15*	5	A/G	0.306	0.289	0.299	0.997 [0.967-1.027]	9.56E-01	5.92E-01
rs6808574	*LPP*	3	C/T	0.621	0.604	0.613	1.003 [0.975-1.031]	9.60E-01	7.07E-01

Supp. Table 3-5. LADA (n= 978) vs T2D and BMDCS controls testing T2D candidate loci

RSID	Locus	Chr	T2D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T2D *P*-value
				LADA	T2D	Control			
rs1111875	*HHEX/IDE*	10	C/T	0.643	0.634	0.59	1.205 [1.175-1.234]	3.63E-03	0.75
rs12970134	*MC4R*	18	A/G	0.291	0.291	0.247	1.216 [1.183-1.249]	0.01	0.97
rs12427353	*HNF1A*	12	G/C	0.818	0.828	0.787	1.224 [1.187-1.260]	0.01	0.72
rs11634397	*ZFAND6*	15	G/A	0.69	0.679	0.653	1.184 [1.153-1.215]	0.01	0.80
rs16861329	*ST64GAL1*	3	C/T	0.886	0.872	0.861	1.251 [1.208-1.294]	0.02	0.29
rs849135	*JAZF1*	7	G/A	0.527	0.529	0.485	1.168 [1.138-1.198]	0.02	0.97
rs7178572	*HMG20A*	15	G/A	0.699	0.73	0.73	0.851 [0.818-0.884]	0.02	0.08
rs4458523	*WFS1*	4	G/T	0.628	0.614	0.59	1.143 [1.113-1.173]	0.04	0.31
rs6723108	*TMEM163*	2	T/G	0.613	0.521	0.553	1.144 [1.114-1.174]	0.04	4.27E-03
rs3786897	*PEPD*	19	A/G	0.625	0.578	0.589	1.143 [1.112-1.173]	0.04	0.02
rs11717195	*ADCY5*	3	T/C	0.796	0.777	0.77	1.162 [1.126-1.199]	0.06	0.18
rs8108269	*GIPR*	19	G/T	0.336	0.31	0.292	1.142 [1.110-1.175]	0.06	0.26
rs10203174	*THADA*	2	C/T	0.89	0.897	0.871	1.200 [1.154-1.246]	0.06	0.18
rs6813195	*TMEM154*	4	C/T	0.748	0.737	0.718	1.141 [1.107-1.175]	0.07	0.39
rs10811661	*CDKN2A/B*	9	T/C	0.86	0.858	0.831	1.170 [1.130-1.211]	0.07	0.64
rs163184	*KCNQ1*	11	G/T	0.498	0.49	0.463	1.124 [1.094-1.154]	0.07	0.39
rs17106184	*FAF1*	1	G/A	0.913	0.918	0.895	1.207 [1.158-1.256]	0.07	0.57
rs5215	*KCNJ11*	11	C/T	0.386	0.382	0.362	1.112 [1.081-1.142]	0.11	0.41
rs391300	*SRR*	17	C/T	0.638	0.638	0.61	1.109 [1.078-1.139]	0.11	0.98
rs4430796	*HNF1B*	17	G/A	0.517	0.479	0.484	1.108 [1.078-1.138]	0.12	0.03
rs17584499	*PTPRD*	9	T/C	0.217	0.17	0.193	1.130 [1.094-1.167]	0.12	9.54E-04
rs10923931	*NOTCH2*	1	T/G	0.12	0.117	0.106	1.164 [1.118-1.209]	0.12	0.91
rs2943640	*IRS1*	2	C/A	0.656	0.675	0.634	1.108 [1.076-1.140]	0.13	0.47
rs516946	*ANK1*	8	C/T	0.775	0.782	0.754	1.116 [1.082-1.151]	0.14	0.22
rs1359790	*SPRY2*	13	G/A	0.76	0.741	0.736	1.104 [1.070-1.139]	0.18	0.75
rs10401969	*CILP2*	19	C/T	0.085	0.088	0.067	1.172 [1.115-1.228]	0.19	0.35
rs7903146	*TCF7L2*	10	T/C	0.315	0.376	0.298	1.088 [1.056-1.119]	0.21	5.03E-04
rs3923113	*GRB14*	2	A/C	0.65	0.656	0.632	1.081 [1.050-1.112]	0.24	0.36
rs2261181	*HMGA2*	12	T/C	0.114	0.109	0.101	1.129 [1.080-1.177]	0.25	0.95
rs780094	*GCKR*	2	C/T	0.573	0.618	0.6	0.929 [0.899-0.959]	0.25	0.04
rs17791513	*TLE4*	9	A/G	0.949	0.952	0.939	1.168 [1.105-1.231]	0.25	0.81
rs17168486	*DGKB*	7	T/C	0.181	0.172	0.163	1.100 [1.061-1.139]	0.25	0.53
rs13233731	*KLF14*	7	G/A	0.551	0.538	0.527	1.077 [1.047-1.108]	0.26	0.47
rs10886471	*GRK5*	10	T/C	0.47	0.477	0.489	0.932 [0.903-0.962]	0.27	0.36
rs10830963	*MTNR1B*	11	G/C	0.282	0.27	0.29	0.932 [0.899-0.965]	0.32	0.64
rs7955901	*TSPAN8*	12	C/T	0.464	0.487	0.438	1.061 [1.032-1.090]	0.34	0.16
rs4275659	*MPHOSPH9*	12	C/T	0.715	0.735	0.694	1.065 [1.033-1.097]	0.36	0.56
rs11063069	*CCND2*	12	G/A	0.214	0.223	0.221	0.932 [0.897-0.968]	0.36	0.43
rs7041847	*GLIS3*	9	A/G	0.527	0.516	0.505	1.060 [1.029-1.090]	0.37	0.40
rs7756992	*CDKAL1*	6	G/A	0.312	0.316	0.289	1.059 [1.027-1.091]	0.40	0.79
rs9470794	*ZFAND3*	6	T/C	0.915	0.911	0.908	1.095 [1.044-1.147]	0.41	0.65
rs1802295	*VPS26A*	10	T/C	0.304	0.321	0.322	0.944 [0.912-0.977]	0.42	0.25
rs7403531	*RASGRP1*	15	T/C	0.218	0.229	0.204	1.064 [1.029-1.100]	0.42	0.22
rs3802177	*SLC30A8*	8	A/G	0.31	0.294	0.301	1.055 [1.022-1.087]	0.45	0.11
rs4812829	*HNF4A*	20	G/A	0.841	0.837	0.831	1.068 [1.027-1.109]	0.45	0.37
rs243088	*BCL11A*	2	T/A	0.496	0.492	0.485	1.040 [1.010-1.070]	0.54	0.41
rs12899811	*PRC1*	15	G/A	0.329	0.315	0.306	1.042 [1.010-1.073]	0.55	0.87
rs12571751	*ZMIZ1*	10	A/G	0.564	0.583	0.549	1.040 [1.009-1.070]	0.55	0.14
rs2334499	*DUSP8*	11	T/C	0.409	0.418	0.41	1.040 [1.009-1.070]	0.55	0.30
rs7593730	*RBMS1*	2	C/T	0.793	0.811	0.796	1.048 [1.011-1.084]	0.55	0.49
rs9936385	*FTO*	16	C/T	0.402	0.454	0.394	1.039 [1.008-1.069]	0.56	1.18E-03
rs6795735	*ADAMTS9*	3	C/T	0.567	0.618	0.575	0.964 [0.934-0.993]	0.56	0.01
rs10278336	*GCK*	7	A/G	0.586	0.577	0.579	1.038 [1.008-1.069]	0.57	0.27
rs831571	*PSMD6*	3	C/T	0.835	0.81	0.827	1.047 [1.008-1.086]	0.59	0.07
rs7202877	*BCAR1*	16	T/G	0.886	0.914	0.894	0.950 [0.902-0.997]	0.61	0.04
rs7612463	*UBE2E2*	3	C/A	0.886	0.894	0.89	0.955 [0.907-1.003]	0.65	0.17
rs4402960	*IGF2BP*	3	T/G	0.315	0.351	0.314	1.029 [0.997-1.062]	0.68	0.05
rs702634	*ARL15*	5	A/G	0.711	0.711	0.701	1.027 [0.994-1.060]	0.71	0.88
rs2028299	*AP3S2*	15	C/A	0.291	0.297	0.291	1.026 [0.993-1.059]	0.72	0.69
rs2796441	*TLE1*	9	G/A	0.609	0.582	0.601	1.023 [0.992-1.053]	0.73	0.14
rs11257655	*CDC123*	10	T/C	0.196	0.236	0.213	0.975 [0.937-1.012]	0.75	0.12
rs6467136	*GCC1*	7	A/G	0.472	0.472	0.477	0.980 [0.950-1.010]	0.75	0.60
rs1552224	*ARAP1 (CENTD2)*	11	A/C	0.848	0.849	0.85	0.975 [0.934-1.016]	0.77	0.73
rs9505118	*SSR1/RREB1*	6	A/G	0.592	0.612	0.592	1.016 [0.986-1.047]	0.81	0.17
rs7163757	*C2CD4A*	15	C/T	0.583	0.59	0.588	0.988 [0.958-1.018]	0.85	0.15
rs7845219	*TP53INP1*	8	T/C	0.524	0.525	0.515	1.009 [0.980-1.039]	0.88	0.45
rs6808574	*LPP*	3	C/T	0.62	0.604	0.613	1.009 [0.979-1.040]	0.89	0.65
rs459193	*ANKRD55*	5	G/A	0.748	0.771	0.741	1.008 [0.974-1.043]	0.91	0.13
rs6878122	*ZBED3*	5	A/G	0.706	0.658	0.706	0.995 [0.963-1.028]	0.95	0.01
rs10842994	*KLHDC5*	12	C/T	0.797	0.814	0.794	1.004 [0.968-1.041]	0.96	0.27
rs2075423	*PROX1*	1	G/T	0.66	0.678	0.653	0.999 [0.968-1.030]	0.99	0.56

Supp. Table 3-6. LADA cases positive for GADA only (n= 669) vs T2D and BMDCS controls testing T2D candidate loci

RSID	Locus	Chr	T2D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T2D *P*-value
				LADA	T2D	Control			
rs12427353	*HNF1A*	12	G/C	0.857	0.828	0.787	1.474 [1.438-1.511]	2.52E-04	5.42E-02
rs8108269	*GIPR*	19	G/T	0.341	0.31	0.292	1.221 [1.189-1.253]	3.25E-02	1.77E-01
rs6878122	*ZBED3*	5	A/G	0.744	0.658	0.706	1.216 [1.184-1.249]	3.86E-02	1.47E-05
rs2943640	*IRS1*	2	C/A	0.674	0.675	0.634	1.209 [1.177-1.241]	4.04E-02	9.59E-01
rs6723108	*TMEM163*	2	T/G	0.625	0.521	0.553	1.193 [1.164-1.223]	4.20E-02	6.22E-03
rs7041847	*GLIS3*	9	A/G	0.554	0.516	0.505	1.188 [1.157-1.218]	5.27E-02	1.68E-01
rs7903146	*TCF7L2*	10	T/C	0.251	0.376	0.298	0.852 [0.820-0.883]	8.14E-02	2.56E-07
rs12571751	*ZMIZ1*	10	A/G	0.503	0.583	0.549	0.866 [0.836-0.895]	9.35E-02	1.02E-03
rs2075423	*PROX1*	1	G/T	0.608	0.678	0.653	0.867 [0.837-0.898]	1.11E-01	1.14E-02
rs7403531	*RASGRP1*	15	T/C	0.228	0.229	0.204	1.174 [1.138-1.209]	1.22E-01	7.04E-01
rs849135	*JAZF1*	7	G/A	0.523	0.529	0.485	1.141 [1.112-1.171]	1.30E-01	9.02E-01
rs17168486	*DGKB*	7	T/C	0.199	0.172	0.163	1.170 [1.131-1.208]	1.63E-01	1.92E-01
rs4402960	*IGF2BP*	3	T/G	0.283	0.351	0.314	0.876 [0.843-0.908]	1.65E-01	4.11E-03
rs7163757	*C2CD4A*	15	C/T	0.55	0.59	0.588	0.893 [0.864-0.922]	1.87E-01	4.06E-02
rs11063069	*CCND2*	12	G/A	0.201	0.223	0.221	0.877 [0.841-0.912]	1.98E-01	3.36E-01
rs459193	*ANKRD55*	5	G/A	0.764	0.771	0.741	1.133 [1.099-1.167]	2.09E-01	6.51E-01
rs10842994	*KLHDC5*	12	C/T	0.816	0.814	0.794	1.141 [1.104-1.177]	2.15E-01	8.39E-01
rs702634	*ARL15*	5	A/G	0.659	0.711	0.701	0.895 [0.863-0.927]	2.31E-01	1.05E-01
rs163184	*KCNQ1*	11	G/T	0.497	0.49	0.463	1.104 [1.074-1.133]	2.47E-01	5.68E-01
rs1802295	*VPS26A*	10	T/C	0.296	0.321	0.322	0.899 [0.867-0.932]	2.59E-01	2.46E-01
rs9470794	*ZFAND3*	6	T/C	0.892	0.911	0.908	0.851 [0.801-0.900]	2.64E-01	1.80E-01
rs17106184	*FAF1*	1	G/A	0.905	0.918	0.895	1.165 [1.117-1.213]	2.73E-01	5.08E-01
rs3802177	*SLC30A8*	8	G/A	0.672	0.706	0.699	0.910 [0.878-0.942]	3.13E-01	8.75E-02
rs7178572	*HMG20A*	15	G/A	0.707	0.73	0.73	0.910 [0.877-0.943]	3.26E-01	2.48E-01
rs11257655	*CDC123*	10	T/C	0.189	0.236	0.213	0.900 [0.863-0.937]	3.29E-01	1.22E-01
rs2028299	*AP3S2*	15	C/A	0.269	0.297	0.291	0.911 [0.878-0.944]	3.34E-01	2.02E-01
rs4275659	*MPHOSPH9*	12	C/T	0.714	0.735	0.694	1.092 [1.060-1.124]	3.49E-01	7.39E-01
rs10923931	*NOTCH2*	1	T/G	0.11	0.117	0.106	1.134 [1.088-1.181]	3.52E-01	9.38E-01
rs780094	*GCKR*	2	C/T	0.566	0.618	0.6	0.923 [0.894-0.953]	3.55E-01	1.77E-01
rs4458523	*WFS1*	4	G/T	0.574	0.614	0.59	0.930 [0.900-0.959]	3.97E-01	1.20E-01
rs1359790	*SPRY2*	13	G/A	0.717	0.741	0.736	0.925 [0.892-0.958]	4.21E-01	8.02E-02
rs7845219	*TP53INP1*	8	T/C	0.495	0.525	0.515	0.933 [0.903-0.963]	4.24E-01	3.70E-01
rs10886471	*GRK5*	10	T/C	0.468	0.477	0.489	0.935 [0.906-0.965]	4.32E-01	5.75E-01
rs3786897	*PEPD*	19	A/G	0.57	0.578	0.589	0.935 [0.905-0.965]	4.47E-01	5.84E-01
rs16861329	*ST64GAL1*	3	C/T	0.872	0.872	0.861	1.097 [1.055-1.140]	4.50E-01	6.59E-01
rs10830963	*MTNR1B*	11	G/C	0.314	0.27	0.29	1.070 [1.037-1.103]	4.77E-01	4.90E-02
rs5215	*KCNJ11*	11	C/T	0.374	0.382	0.362	1.065 [1.034-1.095]	4.85E-01	9.76E-01
rs11634397	*ZFAND6*	15	G/A	0.668	0.679	0.653	1.062 [1.032-1.093]	4.97E-01	4.43E-01
rs10401969	*CILP2*	19	C/T	0.06	0.088	0.067	0.888 [0.828-0.948]	4.99E-01	2.40E-02
rs4430796	*HNF1B*	17	G/A	0.5	0.479	0.484	1.060 [1.029-1.090]	5.09E-01	3.44E-01
rs6467136	*GCC1*	7	A/G	0.492	0.472	0.477	1.054 [1.024-1.083]	5.41E-01	8.66E-01
rs2261181	*HMGA2*	12	T/C	0.115	0.109	0.101	1.084 [1.036-1.133]	5.69E-01	5.69E-01
rs10203174	*THADA*	2	C/T	0.879	0.897	0.871	1.065 [1.021-1.109]	6.21E-01	1.05E-01
rs11717195	*ADCY5*	3	T/C	0.757	0.777	0.77	0.951 [0.915-0.987]	6.29E-01	6.77E-01
rs1111875	*HHEX/IDE*	10	C/T	0.587	0.634	0.59	0.960 [0.931-0.990]	6.35E-01	9.02E-03
rs243088	*BCL11A*	2	T/A	0.489	0.492	0.485	1.041 [1.011-1.070]	6.41E-01	4.54E-01
rs6813195	*TMEM154*	4	C/T	0.723	0.737	0.718	1.044 [1.011-1.077]	6.52E-01	8.01E-01
rs516946	*ANK1*	8	C/T	0.751	0.782	0.754	1.037 [1.004-1.071]	7.08E-01	9.86E-02
rs10278336	*GCK*	7	A/G	0.567	0.577	0.579	0.968 [0.937-0.998]	7.09E-01	6.17E-01
rs3923113	*GRB14*	2	A/C	0.639	0.656	0.632	1.033 [1.003-1.064]	7.14E-01	2.16E-01
rs391300	*SRR*	17	C/T	0.618	0.638	0.61	1.032 [1.002-1.061]	7.20E-01	5.31E-01
rs7593730	*RBMS1*	2	C/T	0.794	0.811	0.796	1.039 [1.002-1.075]	7.23E-01	4.01E-01
rs7612463	*UBE2E2*	3	C/A	0.887	0.894	0.89	0.953 [0.904-1.001]	7.31E-01	3.48E-01
rs7955901	*TSPAN8*	12	C/T	0.456	0.487	0.438	1.028 [0.999-1.057]	7.47E-01	8.39E-02
rs12970134	*MC4R*	18	A/G	0.257	0.291	0.247	1.032 [0.998-1.066]	7.50E-01	1.88E-01
rs6795735	*ADAMTS9*	3	C/T	0.579	0.618	0.575	1.027 [0.997-1.056]	7.57E-01	2.46E-01
rs4812829	*HNF4A*	20	G/A	0.826	0.837	0.831	0.970 [0.931-1.010]	7.97E-01	5.08E-01
rs831571	*PSMD6*	3	C/T	0.817	0.81	0.827	0.974 [0.935-1.012]	8.13E-01	9.28E-01
rs1552224	*ARAP1 (CENTD2)*	11	A/C	0.859	0.849	0.85	1.025 [0.983-1.066]	8.40E-01	3.23E-01
rs17584499	*PTPRD*	9	T/C	0.202	0.17	0.193	1.020 [0.984-1.056]	8.49E-01	1.43E-01
rs2796441	*TLE1*	9	G/A	0.602	0.582	0.601	1.015 [0.986-1.045]	8.61E-01	2.25E-01
rs9505118	*SSR1/RREB1*	6	A/G	0.602	0.612	0.592	1.015 [0.985-1.045]	8.65E-01	6.60E-01
rs6808574	*LPP*	3	C/T	0.623	0.604	0.613	1.009 [0.980-1.039]	9.14E-01	8.13E-01
rs10811661	*CDKN2A/B*	9	T/C	0.832	0.858	0.831	1.011 [0.972-1.049]	9.26E-01	3.70E-01
rs9936385	*FTO*	16	C/T	0.396	0.454	0.394	1.008 [0.978-1.037]	9.30E-01	2.42E-02
rs12899811	*PRC1*	15	G/A	0.314	0.315	0.306	0.995 [0.964-1.026]	9.56E-01	9.26E-01
rs13233731	*KLF14*	7	G/A	0.529	0.538	0.527	1.004 [0.974-1.034]	9.66E-01	2.41E-01
rs2334499	*DUSP8*	11	T/C	0.409	0.418	0.41	0.997 [0.967-1.027]	9.75E-01	2.40E-01
rs7756992	*CDKAL1*	6	G/A	0.294	0.316	0.289	1.001 [0.969-1.033]	9.89E-01	7.67E-01
rs7202877	*BCAR1*	16	T/G	0.895	0.914	0.894	0.999 [0.952-1.046]	9.94E-01	2.00E-01
rs17791513	*TLE4*	9	A/G	0.939	0.952	0.939	1.001 [0.941-1.062]	9.94E-01	3.10E-01

Supp. Table 3-7. LADA cases positive for GADA and IA2A (n= 309) vs T2D and BMDCS controls testing T2D candidate loci

A) MHC, *PTPN22, SH2B3, INS, SMARCE1* association replicated in 978 LADA cases and 2,820 WTCCC healthy British controls

Locus	SNP	T1D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T1D *P*-value
			LADA	T1D	Control			
MHC	rs9272346	A/G	0.763	0.818	0.620	1.591 [1.577-1.604]	3.34×10^{-9}	4.01×10^{-3}
PTPN22	rs6679677	A/C	0.17	0.17	0.097	1.855 [1.834-1.875]	1.88×10^{-7}	0.603
SH2B3	rs17696736	G/A	0.542	0.503	0.44	1.316 [1.303-1.329]	1.95×10^{-4}	0.180
INS	rs689	T/A	0.824	0.741	0.712	1.479 [1.465-1.494]	2.07×10^{-6}	1.68×10^{-6}
INS	rs7111341	C/T	0.812	0.75	0.734	1.272 [1.258-1.287]	4.05×10^{-3}	2.39×10^{-4}
SMARCE1	rs7221109	C/T	0.621	0.687	0.648	1.010[0.995-1.025]	0.810	6.54×10^{-4}

B) MHC, *PTPN22, SH2B3, and INS* association replicated in 309 LADA cases positive for autoantibodies GADA and IA2 and 2,820 WTCCC healthy British controls

Locus	SNP	T1D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T1D *P*-value
			LADA	T1D	Control			
MHC	rs9272346	A/G	0.686	0.818	0.620	1.177[1.161-1.192]	1.01×10^{-4}	1.26×10^{-17}
PTPN22	rs6679677	A/C	0.143	0.17	0.097	1.329[1.306-1.352]	3.86×10^{-6}	2.61×10^{-2}
SH2B3	rs17696736	G/A	0.515	0.503	0.44	1.162[1.148-1.177]	9.35×10^{-5}	0.542
INS	rs689	T/A	0.796	0.741	0.712	1.235[1.219-1.252]	1.27×10^{-4}	3.88×10^{-4}

C) *HNF1A* and *TCF7L2* association replicated in 978 LADA cases and 2,820 WTCCC healthy British controls

Locus	SNP	T2D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T2D *P*-value
			LADA	T2D	Control			
HNF1A	rs12427353	G/C	0.831	0.828	0.807	1.111 [1.092-1.129]	3.57×10^{-2}	0.538
TCF7L2	rs7903146	T/C	0.295	0.376	0.300	0.988 [0.972-1.004]	0.702	5.21×10^{-6}

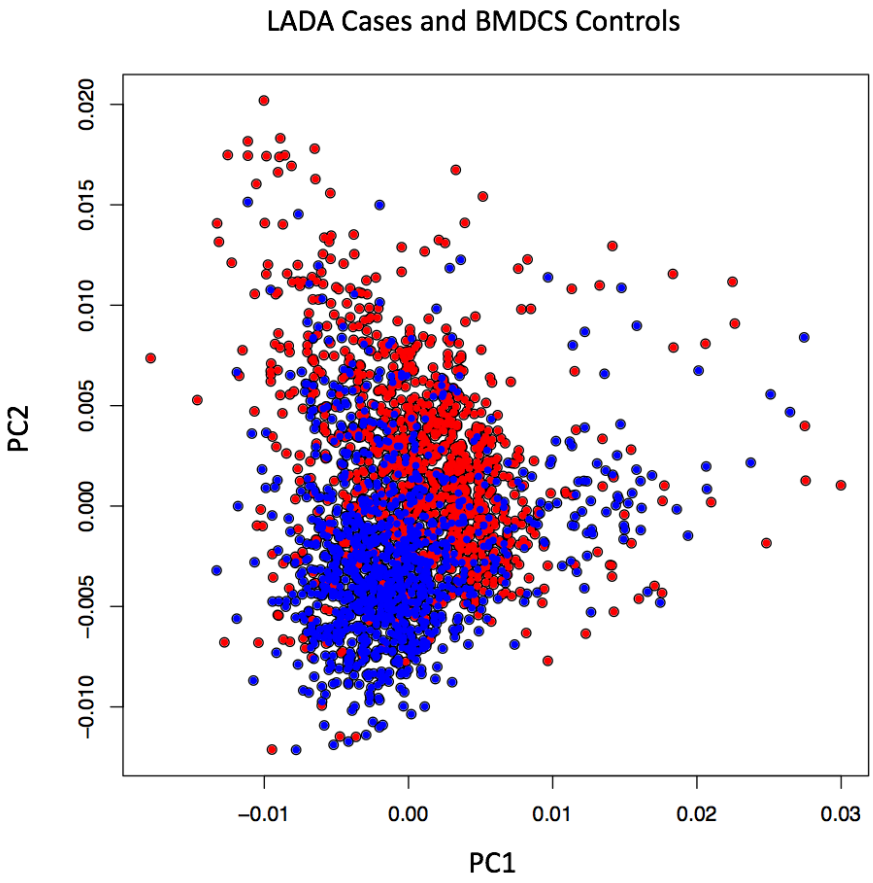
D) *HNF1A, ZBED3, and TCF7L2* association replicated in 309 LADA cases positive for autoantibodies GADA and IA2 and 2,820 WTCCC healthy British controls

Locus	SNP	T2D Alleles Risk/Other	Risk Allele Frequency			LADA Odds Ratio	LADA *P*-value	LADA vs T2D *P*-value
			LADA	T2D	Control			
HNF1A	rs12427353	G/C	0.857	0.828	0.807	1.274 [1.258-1.290]	1.05×10^{-2}	5.42×10^{-2}
ZBED3	rs6878122	A/G	0.744	0.658	0.658	1.132 [1.118-1.146]	0.123	1.47×10^{-5}
TCF7L2	rs7903146	T/C	0.251	0.376	0.300	0.851 [0.837-0.865]	4.62×10^{-2}	2.56×10^{-7}

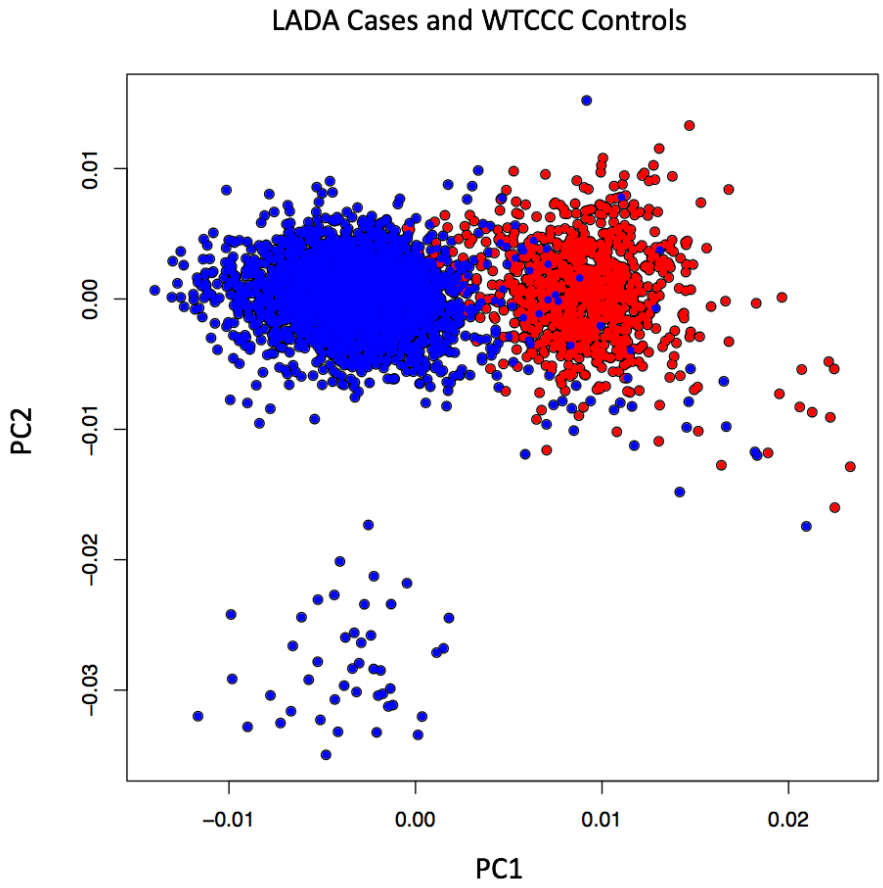
Supp. Table 3-8. Follow-up analysis of significant T1D and T2D loci in LADA using WTCCC controls

Supplemental Figures

A

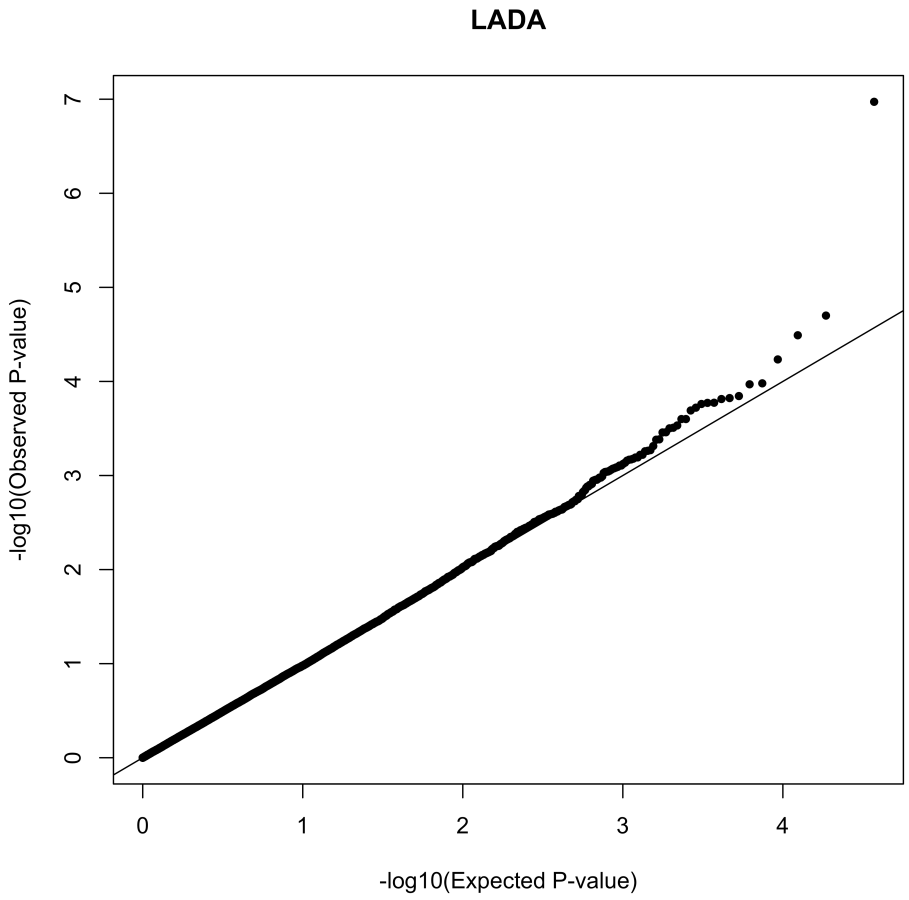


B



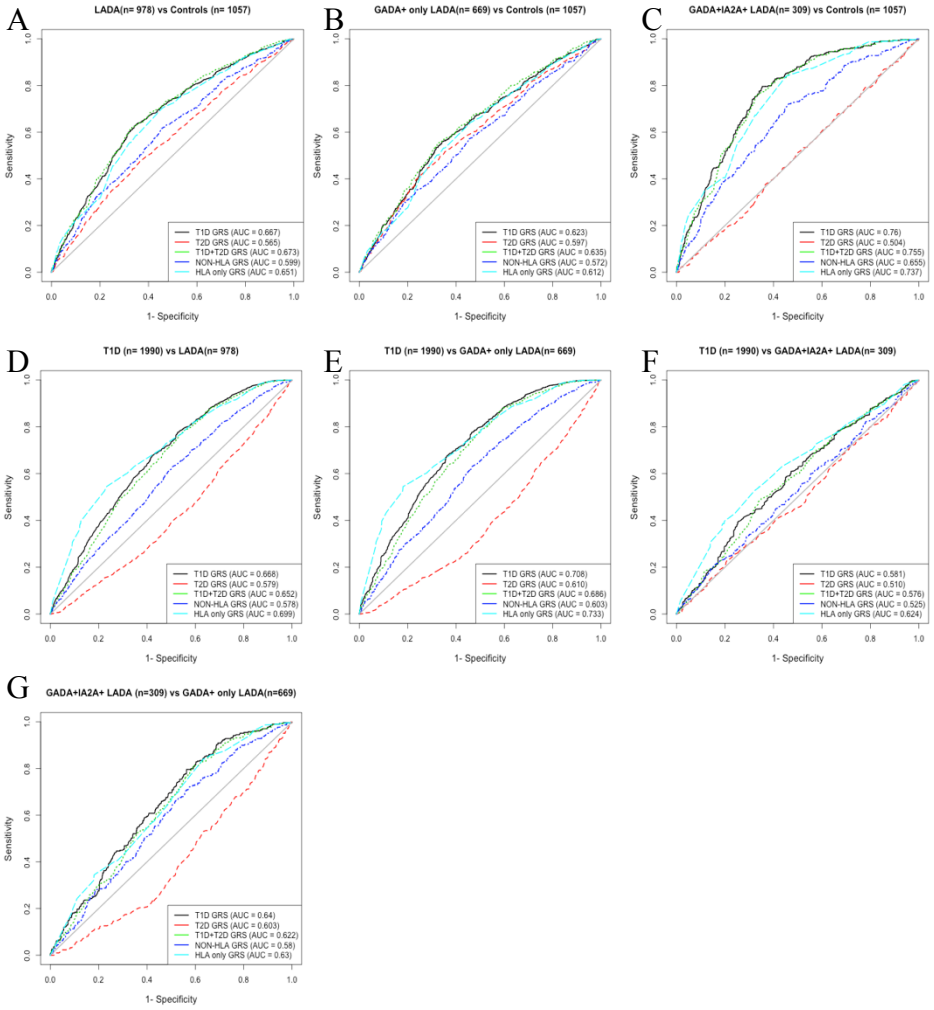
Supp. Figure 3-1. Principle Component Analyses.

A) 978 LADA cases (red) vs 1,057 BMDCS controls (blue), and B) 978 LADA cases (red) vs 2,820 WTCCC controls (blue).



Supp. Figure 3-2. Quantile-quantile plot

Quantile-quantile plot of pruned markers used for PCs (no HLA signals) for LADA cases and BMDCS controls. The genomic inflation factor is 0.966.



Supp. Figure 3-3. ROC plots

Extended set of ROC plots including Non-HLA and HLA only GRS also tested in LADA vs T1D and GADA+ only LADA vs GADA+IA2A LADA. Weighted genetic risk scores (GRS) for T1D GRS (black), T1D without HLA GRS (blue) , HLA only GRS (cyan), T2D GRS (red) and both T1D and T2D GRS (green) were calculated and tested in A) 978 LADA cases and 1,057 healthy controls B) 669 GADA only autoantibody- positive and 1,057 healthy controls C) 309 autoantibody-positive (GADA, IA2A) LADA and 1,057 controls D) 978 LADA and 1,990 WTCCC T1D cases E) 669 GADA only autoantibody- positive LADA and 1,990 WTCCC T1D F) 309 autoantibody- positive (GADA, IA2A) LADA and 1,990 WTCCC T1D. The T2D GRS predict GADA+ only LADA from GADA+IA2A+ LADA and LADA from T1D, resulting in T2D GRS performance dropping below the y=x line.

Chapter 5: THE FIRST GENOME-WIDE ASSOCIATION STUDY IN LADA – SUPPLEMENTARY MATERIAL

Supplemental Tables

Cohort	Population	N	Age (mean, range)	% Male/Female	GADA measurement method, notes	Genotyping chip	Genotyping QC	Imputation panel and service	Imputation QC
ActionLada (LADA)	British, German	1051	47.7, 46.99-48.47	M57%/F37% * 6% missing info.	In House RIA	Illumina Infinium II Omni Express	Individuals with ambiguous sex, genotype missingness > 5%, and relatedness (pi_hat > 0.2) were excluded. PCA was performed to exclude individuals of non-European ancestry.	HRC r1.1.2016 (Michigan)	SNP missing call rate <95%, SNP missing rate <95%, MAF <0.05, non-European individuals
ActionLada 'Plus' (LADA)	American and British	441	55.33 years (25 years-88 years)	55.6%/44.1%	RIA at Northwest Lipid, Metabolism, and Diabetes Research Laboratories, Seattle Washington	Illumina Infinium II Omni Express	Individuals with ambiguous sex, genotype missingness > 5%, and relatedness (pi_hat > 0.2) were excluded. PCA was performed to exclude individuals of non-European ancestry.	HRC r1.1.2016 (Michigan)	SNP missing call rate <95%, SNP missing rate <95%, MAF <0.05, non-European individuals
Bone Mineral Density in Childhood Study (population controls)	American of diverse ethnic backgrounds	1056	5-20 yrs at baseline	49%/51% at baseline	Not done	Illumina Infinium II Omni Express	Individuals with ambiguous sex, genotype missingness > 5%, and relatedness (pi_hat > 0.2) were excluded. PCA was performed to exclude individuals of non-European ancestry.	HRC r1.1.2016 (Michigan)	SNP missing call rate <95%, SNP missing rate <95%, MAF <0.05, non-European individuals
Botnia study (LADA)	Finnish (Botnia region, Western Finland)	157	54.82 ± 11-72 yrs	42% Male	in-house RIP, RSR EIA	Illumina Human Core Exome	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); relatedness (pi-hat >= 0.2); average pi-hat outliers; population outliers.Exclude SNPs: monomorphic; (MAF >= 0.05 and SNP missing rate > 0.05); (MAF < 0.05 and SNP missing rate > 0.01); (MAF >= 0.05 and HWE <= 0.00000057); (MAF < 0.05 and HWE <= 0.0001)	HRC r1.1.2016 (Michigan)	Include: MAF > 0.01; INFO > 0.4; (MAF < 0.05 and HWE > 0.0001 or MAF >= 0.05 and HWE > 0.00000057)
Genetics of Diabetes Audit and Research (GoDARTS) (LADA/T2D/Population	Scottish	206	mean 58.54 STD 12.42	49.0% Male	RSR ELISA (RS-GDE96-F)	Affymatrix 6; Illumina Infinium II Omni Express	Individuals with ambiguous sex or geno-pheno sex discrepancy, genotype missingness > 2%, and relatedness (pi_hat > 0.125) were excluded. PCA was performed to exclude individuals of non-European ancestry.	HRC r1.1.2016 (Sanger)	None
The Nord-Trøndelag Health Study (LADA)	Norwegian	139	Mean 67.5 and range 57.4	53.2% Male	AntiGAD was measured by immuno-precipitation using translation labeled 3H-GAD65 as labeled reagent (Novo Nordisk, Denmark).	IlluminaHumanCore arrays (UM_HUNT_Biobank_11788091, HumanCoreExome-12-v1-0, and HumanCoreE xome-12-v1-1)	Samples that failed to reach a 99% call rate, had contamination > 2.5% as estimated with BAF Regress, large chromosomal copy number variants, lower call rate of a technical duplicate pair and twins, gonosomal constellations other than XX and XY, or whose inferred sex contradicted the reported gender, were excluded. Samples that passed quality control were analysed in a second round of genotype calling following the Genome Studio quality control protocol.	chr1-22: HRC.r1-1 + 2,200 whole-genome sequenced HUNT samples; chrX: HRC.r1-1 (Michigan)	Imputation was performed on the samples of recent European ancestry using Minimac3 (v2.0.1, http://genome.sph.umich.edu/wiki/Minimac3) and the Haplotype Reference Consortium reference panel (release version 1.1). A maximal set of relatively unrelated individuals (kinship coefficient < 0.0884) was chosen using KING and FastIndep.
The Nord-Trøndelag Health Study (T2D)		695	Mean 67.2 and range 55.0	53.2% Male		IlluminaHumanCore arrays (UM_HUNT_Biobank_11788091, HumanCoreExome-12-v1-0, and HumanCoreE xome-12-v1-1)		chr1-22: HRC.r1-1 + 2,200 whole-genome sequenced HUNT samples; chrX: HRC.r1-1 (Michigan)	
The Nord-Trøndelag Health Study (Non-diabetic population controls)		695	Mean 67.5 and range 57.8	53.2% Male		IlluminaHumanCore arrays (UM_HUNT_Biobank_11788091, HumanCoreExome-12-v1-0, and HumanCoreE xome-12-v1-1)		chr1-22: HRC.r1-1 + 2,200 whole-genome sequenced HUNT samples; chrX: HRC.r1-1 (Michigan)	
All New Diabetics In Scania (ANDIS)(LADA)	Swedish	440	LADA: 59(35-94)	55/45	Enzyme-Linked Immunosorbent Assay (ELISA) or Radioimmunoassay at Clinical Chemistry in Malmö	Illumina Infinium Omni Express Exome	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); duplicates or related individuals (pi-hat >= 0.2); average pi-hat outliers;population outliers.Exclude SNPs: monomorphic; (MAF >= 0.05 and SNP missing rate > 0.05); (MAF < 0.05 and SNP missing rate > 0.01); (MAF >= 0.05 and HWE <= 0.00000057); (MAF < 0.05 and HWE <= 0.0001)	HRC r1.1.2016 (Michigan)	Include: MAF > 0.01; INFO > 0.4; (MAF < 0.05 and HWE > 0.0001 or MAF >= 0.05 and HWE > 0.00000057)
Diabetes Registry Vasa (LADA / T1D / T2D)	Finnish	3290 (LADA 139, T1D 365 , T2D 2786)	LADA 57.84 ±10.22; T1D 15.80 ±9.03; T2D 59.63 ± 10.29	LADA 46.3% Male; T1D 53.0% Male; T2D 56.5% Male	Enzyme Immunoassay (EIA)	Illumina Human CoreExome	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); duplicates or related individuals (pi-hat >= 0.2); average pi-hat outliers; population outliers.Exclude SNPs: monomorphic; (MAF >= 0.05 and SNP missing rate > 0.05); (MAF < 0.05 and SNP missing rate > 0.01); (MAF >= 0.05 and HWE <= 0.00000057);(MAF < 0.05 and HWE <= 0.0001)	HRC r1.1.2016(Michigan)	Include: MAF > 0.01; INFO > 0.4; (MAF < 0.05 and HWE > 0.0001 or MAF >= 0.05 and HWE > 0.00000057)

Malmö controls (Non-diabetic population controls)	Swedish	3126	72.5(5.6) range 61-85	41/59	Not done	Illumina Infinium Omni Express Exome	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); duplicates or related individuals (pi-hat >= 0.2); population outliers. Exclude SNPs: monomorphic; SNP missing rate > 0.05; HWE <= 10e-6	HRC r1.1.2016(Michigan)	Include: MAF > 0.01; INFO > 0.4; (MAF < 0.05 and HWE > 0.0001 or MAF >= 0.05 and HWE > 0.00000057)
Scania Diabetes Registry (LADA cases/ T1D cases/ T2D cases)	Swedish	3567	LADA only: 59(36-90)	57/43	Radioimmunoassay	Illumina Infinium Omni Express Exome + ???	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); duplicates or related individuals (pi-hat >= 0.2); average pi-hat outliers; population outliers. Exclude SNPs: monomorphic; (MAF >= 0.05 and SNP missing rate > 0.05); (MAF < 0.05 and SNP missing rate > 0.01); (MAF >= 0.05 and HWE <= 0.00000057); (MAF < 0.05 and HWE <= 0.0001)	HRC r1.1.2016(Michigan)	Include: MAF > 0.01; INFO > 0.4; (MAF < 0.05 and HWE > 0.0001 or MAF >= 0.05 and HWE > 0.00000057)
Copenhagen controls (population controls)	Danish	1974	64.42 (range, 34.44)	49.9% male	Not done	Illumina Human Core Exome	Prior to imputation we removed variants that had a missingness of more than 5% (across batches), a minor allele frequency of less than 5%, or a hardy-weinberg equilibrium pvalue < 10e-5.	Haplotype Reference Consortium (r1.1) (Sanger)	None
The 1936 birth cohort (population controls)	Danish	624 non-diabetic individuals (502 NGTs,	See above	See above	See above	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above
ADDITION-PRO (controls)	Danish	1350 non-diabetic individuals (812 NGTs, 538 IFG/IGTs)	See above	See above	See above	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above
Danish Centre for strategic Research in Type 2 Diabetes (DD2) (LADA)	Danish	539	58.32 (range, 67.31)	56.2% male		Illumina Human Core Exome	Prior to imputation we removed variants that had a missingness of more than 5% (across batches), a minor allele frequency of less than 5%, or a hardy-weinberg equilibrium pvalue < 10e-5.	Haplotype Reference Consortium (r1.1) (Sanger)	None
Danish Centre for strategic Research in Type 2 Diabetes (DD2) (LADA)	Danish	158	See above	See above	AESKULISA	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above
Vejle Biobank (LADA)	Danish	124	See above	See above	AESKULISA	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above
OUH (LADA)	Danish	66	See above	See above	RSR RIA	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above
CIMT (LADA)	Danish	31	See above	See above	RSR ELISA	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above
Inter99 (LADA)	Danish	19	See above	See above	RSR ELISA	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above
SDC (LADA)	Danish	141	See above	See above	RSR ELISA (n=52) or AESKULISA (n=89)	See above	See above	Haplotype Reference Consortium (r1.1) (Sanger)	See above

Supp. Table 5-1. Cohort Information

	Directly Genotypes Markers		Imputed Markers				
Cohort	**Lambda**	**Count**	**Lambda**	**Count**	**N LADA**	**N Ctrls**	**total markers**
LADACTRL							
Swedish	1.0297	575268	1.0131	7131633	805	3126	7706901
ActionLADA	1.1662	5397740	1.0000	0	1353	1014	5397740
Danish	1.0000	0	1.0274	5271474	476	1807	5271474
				totals	2634	5947	7868002
LADAT1D							
Cohort	**Lambda**	**Count**	**Lambda**	**Count**	**N LADA**	**N T1D**	**total markers**
Swedish + ActionLADA	1.0301	575963	1.0286	7137816	2158	603	7713779
Finnish	1.0359	254991	1.0314	7607304	296	365	7862295
				totals	2454	968	8328843
LADAT2D							
Cohort	**Lambda**	**Count**	**Lambda**	**Count**	**N LADA**	**N T2D**	**total markers**
Swedish + ActionLADA	1.0621	575855	1.0399	7142626	2158	2599	7718481
Finnish	1.0251	255224	1.0216	7614587	296	2786	7869811
GODARTS-Illumina	1.0000	0	0.9854	5386729	97	2098	5386729
GODARTS-Affymetrix	1.0000	0	0.9774	5383813	89	2218	5383813
HUNT	1	0	1.0242	5546184	139	695	5546184
				totals	2779	10396	8465706

Supp. Table 5-2. Cohort information for GWAMA

LADA	Assuming disease prevalence of 0.0036				
MAF	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**
0.1	0	0.017	0.254	0.758	0.976
0.2	0.001	0.13	0.764	0.992	1
0.3	0.002	0.27	0.917	0.999	1
0.4	0.003	0.346	0.948	1	1
0.5	0.003	0.342	0.939	0.999	1
T1D	Assuming disease prevalence of 0.0033				
	1.1	**1.2**	**1.3**	**1.4**	**1.5**
0.1	0	0.017	0.253	0.758	0.976
0.2	0.001	0.129	0.763	0.992	1
0.3	0.002	0.27	0.917	0.999	1
0.4	0.003	0.346	0.948	1	1
0.5	0.003	0.342	0.939	0.999	1
T2D	Assuming disease prevalence of 0.045				
	1.1	**1.2**	**1.3**	**1.4**	**1.5**
0.1	0	0.017	0.253	0.758	0.976
0.2	0.001	0.174	0.84	0.997	1
0.3	0.003	0.345	0.954	1	1
0.4	0.005	0.43	0.974	1	1
0.5	0.005	0.426	0.968	1	1

Supp. Table 5-3. Power Calculations.

Power was calculated using CaTS calculator. (http://csg.sph.umich.edu/abecasis/cats/)
Number of cases: 2700, number of controls: 5500, significance level: 0.00000005

Original gene set ID	Original gene set description	Nominal P value	False discovery rate < 5%
MP:0005078	abnormal cytotoxic T cell physiology	6.392E-07	No
ENSG00000131153	GINS2 subnetwork	6.832E-07	No
ENSG00000149554	CHEK1 subnetwork	2.411E-06	No
KEGG_CELL_CYCLE	KEGG_CELL_CYCLE	1.167E-05	No
ENSG00000123374	CDK2 subnetwork	1.513E-05	No
ENSG00000141510	TP53 subnetwork	3.328E-05	No
ENSG00000170312	CDK1 subnetwork	4.780E-05	No
ENSG00000198793	MTOR subnetwork	6.033E-05	No
MP:0004045	abnormal cell cycle checkpoint function	8.531E-05	No
ENSG00000134259	NGF subnetwork	9.766E-05	No
REACTOME_P75NTR_SIGNALS_VIA_NF:KB	REACTOME_P75NTR_SIGNALS_VIA_NF:KB	9.993E-05	No
GO:0007260	tyrosine phosphorylation of STAT protein	1.006E-04	No
MP:0003333	liver fibrosis	1.122E-04	No
REACTOME_NF:KB_IS_ACTIVATED_AND_SIGNALS_SURVIVAL	REACTOME_NF:KB_IS_ACTIVATED_AND_SIGNALS_SURVIVAL	1.184E-04	No
ENSG00000127314	RAP1B subnetwork	1.267E-04	No
GO:0051247	positive regulation of protein metabolic process	1.375E-04	No
REACTOME_PI3K_CASCADE	REACTOME_PI3K_CASCADE	1.381E-04	No
MP:0008501	increased IgG2b level	1.798E-04	No
MP:0001654	hepatic necrosis	1.906E-04	No
ENSG00000180228	PRKRA subnetwork	2.149E-04	No
GO:0042509	regulation of tyrosine phosphorylation of STAT protein	2.225E-04	No
ENSG00000168040	FADD subnetwork	2.517E-04	No
ENSG00000120008	WDR11 subnetwork	2.669E-04	No
GO:0032270	positive regulation of cellular protein metabolic process	2.820E-04	No

MP:0002020	increased tumor incidence	2.837E-04	No
ENSG00000171861	RNMTL1 subnetwork	3.582E-04	No
ENSG00000138376	BARD1 subnetwork	3.638E-04	No
MP:0002834	decreased heart weight	3.735E-04	No
MP:0002495	increased IgA level	3.805E-04	No
ENSG00000127191	TRAF2 subnetwork	3.925E-04	No
ENSG00000153201	RANBP2 subnetwork	4.834E-04	No
GO:0042516	regulation of tyrosine phosphorylation of Stat3 protein	5.024E-04	No
ENSG00000185658	BRWD1 subnetwork	5.105E-04	No
ENSG00000163539	CLASP2 subnetwork	5.416E-04	No
ENSG00000134057	CCNB1 subnetwork	5.663E-04	No
ENSG00000066117	SMARCD1 subnetwork	5.726E-04	No
ENSG00000132646	PCNA subnetwork	5.766E-04	No
ENSG00000142856	ITGB3BP subnetwork	5.893E-04	No
MP:0001861	lung inflammation	5.987E-04	No
ENSG00000104823	ECH1 subnetwork	6.190E-04	No
REACTOME_ASSEMBLY_OF_HIV_VIRION	REACTOME_ASSEMBLY_OF_HIV_VIRION	6.410E-04	No
ENSG00000123219	CENPK subnetwork	6.448E-04	No
ENSG00000115163	CENPA subnetwork	6.679E-04	No
ENSG00000206215	ENSG00000206215 subnetwork	6.811E-04	No
ENSG00000206287	RING1 subnetwork	6.811E-04	No
ENSG00000204227	RING1 subnetwork	6.811E-04	No
ENSG00000100368	CSF2RB subnetwork	6.882E-04	No
GO:0001819	positive regulation of cytokine production	6.893E-04	No
MP:0005334	abnormal fat pad morphology	7.367E-04	No
ENSG00000198961	PJA2 subnetwork	8.300E-04	No

ENSG00000130338	TULP4 subnetwork	8.563E-04	No
MP:0005166	decreased susceptibility to injury	8.792E-04	No
ENSG00000153044	CENPH subnetwork	9.019E-04	No
MP:0011099	complete lethality throughout fetal growth and development	9.193E-04	No
ENSG00000137285	TUBB2B subnetwork	9.232E-04	No
ENSG00000157106	SMG1 subnetwork	9.847E-04	No
ENSG00000166793	YPEL4 subnetwork	9.959E-04	No

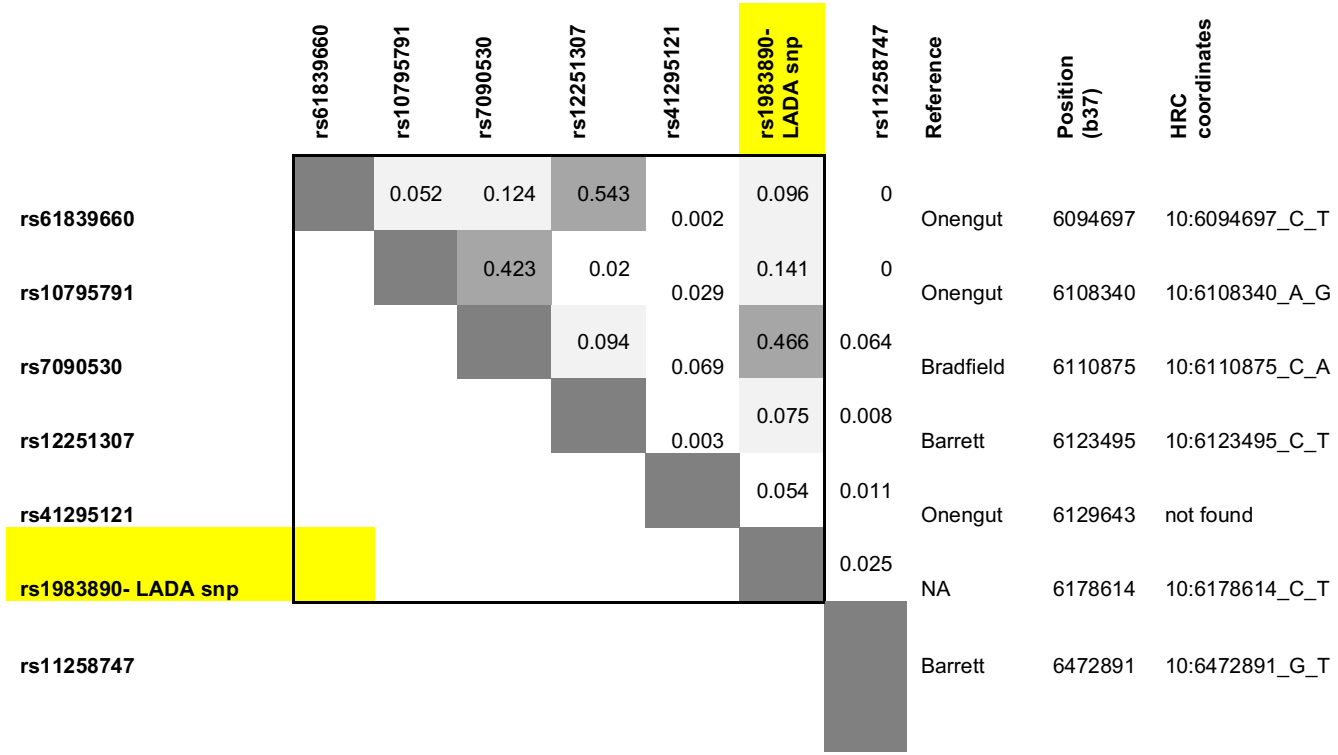
Supp. Table 5-4. DEPICT gene set enrichment analysis p<0.001

MeSH term	MeSH first level term	MeSH second level term	Nominal P value	False discovery rate < 5%
A15.382.490.555.567.537	Killer Cells Natural	Hemic and Immune Systems	0.008	No
A11.118.637.555.567.569	T Lymphocytes	Cells	0.008	No
A15.145	Blood	Hemic and Immune Systems	0.010	No
A15.145.229	Blood Cells	Hemic and Immune Systems	0.010	No
A11.118.637	Leukocytes	Cells	0.011	No
A11.872.378	Hematopoietic Stem Cells	Cells	0.015	No
A11.118.637.555.567.569.200.700	T Lymphocytes Regulatory	Cells	0.017	No
A15.145.229.637.555.567.569.200	CD4 Positive T Lymphocytes	Hemic and Immune Systems	0.018	No
A15.145.229.334	Erythrocytes	Hemic and Immune Systems	0.020	No
A15.382.520	Lymphatic System	Hemic and Immune Systems	0.021	No
A10.549	Lymphoid Tissue	Tissues	0.021	No
A02.835.583.443.800.800	Synovial Fluid	Musculoskeletal System	0.021	No
A11.118.637.555.567.562.440	Precursor Cells B Lymphoid	Cells	0.023	No
A11.872.378.294	Lymphoid Progenitor Cells	Cells	0.023	No
A15.378.316	Bone Marrow Cells	Hemic and Immune Systems	0.025	No
A15.378	Hematopoietic System	Hemic and Immune Systems	0.025	No
A11.443	Erythroid Cells	Cells	0.025	No
A10.549.400	Lymph Nodes	Tissues	0.026	No
A11.627	Myeloid Cells	Cells	0.028	No
A15.145.229.637.555	Leukocytes Mononuclear	Hemic and Immune Systems	0.032	No
A11.118.637.415	Granulocytes	Cells	0.032	No
A15.382.680	Phagocytes	Hemic and Immune Systems	0.035	No
A15.382.490.555.567.622	Lymphocytes Null	Hemic and Immune Systems	0.035	No
A15.382.490.315.583	Neutrophils	Hemic and Immune Systems	0.036	No
A11.627.635	Myeloid Progenitor Cells	Cells	0.042	No
A15.382	Immune System	Hemic and Immune Systems	0.045	No
A15.382.520.604.700	Spleen	Hemic and Immune Systems	0.049	No

Supp. Table 5-5. DEPICT tissue enrichment

	Initial GWAS (LADACTRL)										With HUNT & GoDARTS Replication						
gene	Chromosome	Position	SNP	Reference Allele	Other allele	eaf	OR	p-value	N	effects	OR	OR_95L	OR_95U	p-value	n_studies	n_samples	effects
PFKFB3	10	61786144	rs1983890	C	T	0.642	1.225	2.69E-07	8152	---	1.23	1.14	1.32	3.02E-08	5	10590	+++++
BACH2	6	90958231	rs11755527	G	C	0.444	1.201	6.96E-07	8152	+++	1.20	1.13	1.29	1.02E-07	5	10589	+++++
ACTN1	14	69473004	rs446091	G	A	0.532	1.209	4.97E-07	8151	+++	1.17	1.09	1.25	1.19E-05	5	10589	+++++
TNFRSF11B	8	119886923	rs2055101	T	C	0.517	1.200	1.19E-06	8151	+++	1.15	1.07	1.23	7.47E-05	5	10589	+++-+-
ELL	19	18637610	rs4808814	T	C	0.315	0.829	3.33E-06	8151	---	0.85	0.79	0.91	1.09E-05	5	10590	-----
THADA	2	43806382	rs11888640	T	C	0.518	0.844	6.32E-06	8151	---	0.85	0.79	0.91	1.94E-06	5	10589	----+
LOC100289230	5	98450813	rs10900893	A	G	0.518	0.844	4.99E-06	8151	---	0.86	0.80	0.92	1.13E-05	5	10589	----+
DLK1	14	101306045	rs941576	G	A	0.445	0.846	9.51E-06	8151	---	0.87	0.81	0.93	1.01E-04	5	10589	----+
SPIRE1	18	12687969	rs113916314	G	A	0.082	1.388	1.14E-06	8151	+++	1.29	1.14	1.46	5.45E-05	5	10589	+++--
LOC285889	7	156016226	rs1922086	C	T	0.566	0.840	8.19E-06	8152	---	0.85	0.79	0.92	1.56E-05	5	10590	-----
LOC441178	6	168111238	rs619330	A	G	0.523	1.196	8.98E-06	8151	+++	1.18	1.10	1.27	1.25E-05	5	10589	+++-+-
C6orf132	6	42075522	rs9349218	A	G	0.182	1.235	7.93E-06	8152	+++	1.18	1.08	1.28	2.73E-04	5	10590	+++--
TMEM106B	7	12077764	rs7800617	G	A	0.588	1.204	5.03E-06	8151	+++	1.15	1.07	1.24	1.74E-04	5	10589	+++--

Supp. Table 5-6. Initial GWAS (LADACTRL) and with HUNT & GoDARTS Replication



	rs61839660	rs10795791	rs7090530	rs12251307	rs41295121	rs1983890- LADA snp	rs11258747	Reference	Position (b37)	HRC coordinates
rs61839660		0.052	0.124	0.543	0.002	0.096	0	Onengut	6094697	10:6094697_C_T
rs10795791			0.423	0.02	0.029	0.141	0	Onengut	6108340	10:6108340_A_G
rs7090530				0.094	0.069	0.466	0.064	Bradfield	6110875	10:6110875_C_A
rs12251307					0.003	0.075	0.008	Barrett	6123495	10:6123495_C_T
rs41295121						0.054	0.011	Onengut	6129643	not found
rs1983890- LADA snp							0.025	NA	6178614	10:6178614_C_T
rs11258747								Barrett	6472891	10:6472891_G_T

Supp. Table 5-7.Linkage Disequilibrium in Chr10p15.

Genome-wide significant signals associated with T1D from various publications (and top LADA signal) - pairwise r^2

Locus	# of genes in locus	Chromosome and position	Ensembl Gene ID	Gene symbol	P value	Gene closest to lead SNP	Gene bio-type	Top cis eQTL SNP (Westra et al. Nature Genetics 2014)	False discovery rate < 5%
rs4808814	1	chr19:18553475-18632937	ENSG00000105656	ELL	0.0032	true	Protein coding	rs10164319	No
rs10858008; rs2476601	7	chr1:113933371-114456708	ENSG00000118655	DCLRE1B	0.0094	false	Protein coding+ processed transcript	rs878129	No
rs10858008; rs2476601	7	chr1:113933371-114456708	ENSG00000134242	PTPN22	0.0286	true	protein_coding+ retained_intron+ nonsense_mediated_decay+ processed_transcript	rs7555634; rs1217397	No
rs11888640	1	chr2:43393800-43823185	ENSG00000115970	THADA	0.0531	true	protein_coding+ nonsense_mediated_decay+ processed_transcript+ retained_intron	-	No
rs313568	2	chr4:25121627-25280714	ENSG00000109618	SEPSECS	0.0557	false	protein_coding+ nonsense_mediated_decay+ processed_transcript+ retained_intron	rs4697567	No
rs2926016	1	chr11:81590893-82429124	ENSG00000245832	-	0.1533	true	lincRNA	-	No
rs10900893	1	chr5:98190908-98262240	ENSG00000153922	CHD1	0.1650	true	retained_intron+protein_coding + nonsense_mediated_decay+ processed_transcript	rs16779	No
rs1983890	1	chr10:6186881-6277495	ENSG00000170525	PFKFB3	0.2428	true	protein_coding+ nonsense_mediated_decay+ processed_transcript	-	No
rs10858008; rs2476601	7	chr1:113933371-114456708	ENSG00000081019	RSBN1	0.2672	false	protein_coding+ nonsense_mediated_decay+ processed_transcript	rs2185827	No
rs11755527	1	chr6:90636248-91006627	ENSG00000112182	BACH2	0.2677	true	protein_coding+ processed_transcript	rs10944479; rs10944479	No
rs10858008; rs2476601	7	chr1:113933371-114456708	ENSG00000134262	AP4B1	0.2781	false	protein_coding+ processed_transcript	rs1217397	No
rs10858008; rs2476601	7	chr1:113933371-114456708	ENSG00000081026	MAGI3	0.4521	true	protein_coding+ processed_transcript	-	No
rs2055101	2	chr8:119935796-120118821	ENSG00000184374	COLEC10	0.4643	false	processed_transcript+ protein_coding	rs4295687	No
rs313568	2	chr4:25121627-25280714	ENSG00000038210	PI4K2B	0.6484	true	processed_transcript+ protein_coding	rs12498160	No
rs10858008; rs2476601	7	chr1:113933371-114456708	ENSG00000116793	PHTF1	0.6504	false	processed_transcript+ protein_coding	rs4838993	No
rs10858008; rs2476601	7	chr1:113933371-114456708	ENSG00000188761	BCL2L15	0.6931	false	retained_intron+protein_coding + processed_transcript	-	No
rs1922086	1	chr7:156230483-156238282	ENSG00000233878	-	0.8493	true	lincRNA	-	No
rs941576	3	chr14:101245747-101327368	ENSG00000211574	MIR770	0.8758	false	miRNA	-	No
rs941576	3	chr14:101245747-101327368	ENSG00000252701	-	0.8847	false	snoRNA	-	No
rs941576	3	chr14:101245747-101327368	ENSG00000214548	MEG3	0.8934	true	lincRNA	-	No
rs2055101	2	chr8:119935796-120118821	ENSG00000164761	TNFRSF11B	0.9327	true	protein_coding+retained_intron + nonsense_mediated_decay	rs1389541	No
rs4409356	1	chr7:54610018-54638773	ENSG00000170419	VSTM2A	0.9582	true	protein_coding+ retained_intron+ processed_transcript	-	No

Supp. Table 5- 8. DEPICT gene prioritization

SNP	chr:pos	reference allele	Other allele	eaf	OR	OR_se	z	P	_-log10_p-value	q_statistic	q_p-value	i2	n_studies	n_samples	effects	AverageCall	Rsq	Genotyped
rs7454108	6:32681483	C	T	0.169	2.744	0.123	21.551	**1.84E-101**	100.736	6.108	0.106	0.509	4	9415	++++	0.999	0.993	Imputed
rs9272346	6:32604372	A	G	0.576	1.951	0.074	16.914	**3.31E-63**	62.481	6.941	0.074	0.568	4	9415	++++	0.918	0.758	Imputed
rs2187668	6:32605884	T	C	0.152	2.479	0.141	15.097	**1.02E-50**	49.989	5.950	0.114	0.496	4	9414	++++	0.918	0.577	Imputed
rs3129889	6:32413545	A	G	0.880	2.440	0.146	14.008	**6.57E-44**	43.183	14.168	0.001	0.859	3	8581	+++?	1.000	1.000	Genotyped
rs6679677	1:114303808	A	C	0.121	1.667	0.085	9.490	**4.82E-21**	20.317	3.365	0.339	0.108	4	9414	++++	1.000	0.999	Genotyped
rs689	11:2182224	T	A	0.707	1.459	0.058	9.111	**1.60E-19**	18.795	1.974	0.578	0.000	4	9414	++++	0.980	0.921	Imputed
rs17696736	12:11248681 8	G	A	0.453	1.216	0.042	5.457	**6.22E-08**	7.206	1.091	0.779	0.000	4	9415	++++	0.758	0.357	Imputed
rs11755527	6:90958231	G	C	0.443	1.201	0.041	5.143	**3.38E-07**	6.471	0.074	0.995	0.000	4	9414	++++	0.980	0.934	Imputed
rs597325	6:91002494	G	A	0.580	1.195	0.042	4.864	**1.40E-06**	5.853	4.291	0.232	0.301	4	9415	++++	0.975	0.902	Imputed
rs72928038	6:90976768	A	G	0.182	1.231	0.056	4.386	**1.36E-05**	4.867	5.003	0.172	0.400	4	9414	++++	0.941	0.700	Imputed
rs229541	22:37591318	A	G	0.422	1.165	0.041	4.215	**2.90E-05**	4.537	1.945	0.584	0.000	4	9414	++?-	1.000	0.999	Genotyped
rs941576	14:10130604 5	A	G	0.448	1.162	0.030	-4.143	**3.95E-05**	4.403	1.727	0.631	0.000	4	9414	----	0.924	0.791	Imputed
rs11571316	2:20473108 9	G	A	0.379	1.161	0.031	-4.052	**5.83E-05**	4.234	1.589	0.662	0.000	4	9415	----	0.989	0.960	Imputed
rs7090530	10:61110875	A	C	0.593	1.157	0.042	3.868	**1.24E-04**	3.906	2.489	0.477	0.000	4	9414	++++	0.999	0.996	Imputed
rs72727394	15:38847022	T	C	0.198	1.178	0.050	3.679	**2.63E-04**	3.580	1.222	0.748	0.000	4	9415	++++	0.991	0.959	Imputed
rs61839660	10:6094697	C	T	0.083	1.273	0.049	-3.628	**3.19E-04**	3.496	1.408	0.704	0.000	4	9414	----	0.998	0.972	Imputed
rs4763879	12:9910164	A	G	0.385	1.141	0.041	3.582	**3.80E-04**	3.421	1.063	0.786	0.000	4	9415	++++	0.989	0.966	Imputed
rs9653442	2:100825367	C	T	0.525	1.124	0.031	-3.286	0.000 1115	2.953	1.145	0.766	0.000	4	9414	----	1.000	1.000	Genotyped
rs10492166	12:9885999	G	A	0.478	1.104	0.032	-2.740	0.006 5577	2.183	2.869	0.412	0.000	4	9414	----	0.993	0.978	Imputed
rs12708716	16:11179873	A	G	0.344	1.109	0.033	-2.717	0.007 0225	2.153	3.698	0.296	0.189	4	9414	---+	1.000	1.000	Genotyped
rs10795791	10:6108340	G	A	0.425	1.102	0.039	2.681	0.007 7794	2.108	0.160	0.984	0.000	4	9415	++++	0.999	0.996	Imputed
rs1738074	6:159465977	C	T	0.582	1.100	0.039	2.608	0.009 656	2.015	2.289	0.515	0.000	4	9415	++++	1.000	0.999	Genotyped
rs1615504	18:67526644	T	C	0.513	1.096	0.032	-2.557	0.011 174	1.952	1.283	0.733	0.000	4	9414	----	0.993	0.976	Imputed
rs12453507	17:38053207	G	C	0.507	0.924	0.032	-2.204	0.028 717	1.542	1.215	0.749	0.000	4	9414	----	0.998	0.994	Imputed
rs9585056	13:10008176 6	C	T	0.710	1.091	0.037	-2.063	0.040 561	1.392	3.083	0.379	0.027	4	9415	----	0.987	0.949	Imputed
rs2542151	18:12779947	G	T	0.771	1.099	0.040	-2.051	0.040 174 2	1.379	4.995	0.172	0.399	4	9414	-+--	1.000	0.999	Genotyped
rs1990760	2:163124051	T	C	0.608	1.078	0.039	2.031	0.043 78	1.359	4.845	0.183	0.381	4	9415	++?-	0.809	0.456	Imputed
rs7804356	7:26891665	T	C	0.224	1.091	0.038	-2.029	0.043 977	1.357	3.102	0.376	0.033	4	9414	---+	1.000	1.000	Genotyped
rs10509540	10:90023033	T	C	0.291	1.072	0.035	-1.771	0.078 735	1.104	1.906	0.592	0.000	4	9415	---+	1.000	0.999	Genotyped
rs3024493	1:206943968	C	A	0.156	1.088	0.043	-1.722	0.087 412	1.058	2.247	0.523	0.000	4	9414	---+	1.000	0.999	Genotyped
rs11170466	12:53585859	T	C	0.066	1.127	0.075	1.674	0.096 555	1.015	0.870	0.833	0.000	4	9414	++++	0.991	0.883	Imputed
rs911263	14:68753593	T	C	0.671	1.063	0.040	1.561	0.121 385	0.916	5.098	0.165	0.411	4	9414	++?-	0.989	0.958	Imputed

rs6476839	9:42908823	T	A	0.414	1.054	0.037	1.450	0.149926	0.824	1.493	0.684	0.000	4	9415	++++	0.999	0.997	Imputed
rs12908309	15:38928677	G	A	0.239	1.063	0.038	-1.450	0.150153	0.823	3.234	0.357	0.072	4	9414	--++	0.972	0.885	Imputed
rs193778	16:11351211	G	A	0.250	1.059	0.042	1.398	0.165352	0.782	2.559	0.465	0.000	4	9415	+++-	0.999	0.997	Imputed
rs11203203	21:43836186	A	G	0.377	1.051	0.038	1.352	0.179448	0.746	4.701	0.195	0.362	4	9414	++-+	0.995	0.983	Imputed
rs5753037	22:30581722	T	C	0.380	1.051	0.037	1.346	0.181418	0.741	3.821	0.281	0.215	4	9414	+++-	1.000	1.000	Genotyped
rs1456988	14:98488007	G	T	0.681	1.053	0.036	-1.306	0.19481	0.710	3.983	0.263	0.247	4	9415	+----	1.000	1.000	Genotyped
rs11258747	10:6472891	T	G	0.232	1.054	0.043	1.221	0.222551	0.647	1.127	0.771	0.000	4	9415	++-+	0.991	0.961	Imputed
rs602662	19:49206985	G	A	0.488	0.960	0.036	1.136	0.259598	0.586	5.381	0.146	0.443	4	9415	+-+-	0.997	0.991	Imputed
rs2281808	20:16105551	C	T	0.634	1.043	0.038	1.099	0.275131	0.560	3.133	0.372	0.043	4	9414	++--	1.000	0.998	Genotyped
rs10877012	12:58162085	G	T	0.338	0.963	0.038	0.975	0.332948	0.478	0.618	0.892	0.000	4	9415	++++	0.995	0.980	Imputed
rs2611215	4:166574267	A	G	0.770	1.045	0.043	-0.936	0.353012	0.452	0.502	0.918	0.000	4	9414	---+	0.985	0.922	Imputed
rs11954020	5:35883251	G	C	0.406	1.034	0.037	0.902	0.370518	0.431	10.783	0.013	0.722	4	9414	++--	0.998	0.993	Imputed
rs402072	19:47219122	T	C	0.153	1.044	0.046	-0.864	0.390923	0.408	3.173	0.366	0.055	4	9414	+-+-	0.995	0.967	Imputed
rs1538171	6:126752884	G	C	0.466	1.031	0.036	0.858	0.394335	0.404	3.559	0.313	0.157	4	9414	+++-	1.000	0.999	Imputed
rs10272724	7:50477213	T	C	0.268	0.968	0.040	0.816	0.417739	0.379	2.767	0.429	0.000	4	9414	+-+-	0.999	0.994	Imputed
rs11301008 1	3:46457412	T	C	0.120	1.044	0.051	-0.773	0.442786	0.354	0.858	0.836	0.000	4	9414	--+-	0.992	0.935	Imputed
rs7928968	11:2050299	T	A	0.245	1.033	0.043	0.760	0.450346	0.346	4.883	0.181	0.386	4	9414	-++ +	0.932	0.725	Imputed
rs3825932	15:79235446	C	T	0.646	0.972	0.036	-0.752	0.455487	0.342	0.631	0.889	0.000	4	9414	----	0.895	0.671	Imputed
rs2304256	19:10475652	C	A	0.280	1.029	0.037	-0.708	0.481983	0.317	2.145	0.543	0.000	4	9414	+--+	1.000	0.999	Genotyped
rs10517086	4:26085511	A	G	0.293	1.027	0.039	0.685	0.496398	0.304	0.720	0.868	0.000	4	9414	+-+-	0.999	0.997	Genotyped
rs601338	19:49206674	G	A	0.460	0.978	0.035	0.622	0.537088	0.270	5.594	0.133	0.464	4	9414	+-+-	0.999	0.995	Imputed
rs9924471	16:28591530	A	G	0.153	1.030	0.049	0.605	0.548339	0.261	1.206	0.752	0.000	4	9415	++-+	0.938	0.669	Imputed
rs6691977	1:2008 14959	C	T	0.206	0.974	0.041	-0.587	0.559972	0.252	3.537	0.316	0.152	4	9414	-++ +	0.995	0.980	Imputed
rs924043	6:170379025	C	T	0.803	0.974	0.049	-0.500	0.619885	0.208	1.750	0.626	0.000	4	9414	-++-	0.984	0.917	Imputed
rs478222	2:25301755	A	T	0.420	1.018	0.035	-0.489	0.626 74	0.202	3.245	0.355	0.075	4	9414	-+-+	0.994	0.978	Imputed
rs722988	10:33426147	C	T	0.376	1.014	0.037	0.379	0.706743	0.151	2.264	0.519	0.000	4	9415	+--+	0.995	0.984	Imputed
rs4849135	2:111615079	G	T	0.673	1.014	0.039	0.350	0.728215	0.138	1.533	0.675	0.000	4	9414	-++ +	0.999	0.997	Imputed
rs6827756	4:123184411	T	C	0.600	0.987	0.036	0.348	0.730151	0.137	4.384	0.223	0.316	4	9414	+-++	0.998	0.993	Imputed
rs6920220	6:138006504	A	G	0.219	1.014	0.042	0.308	0.759487	0.119	3.607	0.307	0.168	4	9415	-++-	1.000	1.000	Genotyped
rs694739	11:64097233	A	G	0.378	1.011	0.035	-0.307	0.76021	0.119	4.411	0.220	0.320	4	9414	+---	1.000	1.000	Genotyped
rs12148472	15:79231478	T	C	0.124	1.014	0.051	-0.254	0.80096	0.096	2.997	0.392	0.000	4	9415	-++-	0.926	0.464	Imputed
rs1465788	14:69263599	C	T	0.689	1.010	0.039	0.250	0.803759	0.095	1.438	0.697	0.000	4	9414	-++ +	1.000	1.000	Genotyped
rs4948088	7:51027194	C	A	0.843	1.021	0.088	0.221	0.826422	0.083	1.021	0.600	0.000	3	7047	+?-+	1.000	0.998	Genotyped
rs7221109	17:38770286	C	T	0.615	1.002	0.036	0.040	0.96852	0.014	1.662	0.645	0.000	4	9414	+-++	1.000	0.998	Genotyped

Supp. Table 5-9A. T1D loci in LADACTRL

SNP	LOCUS	Risk Allele	Disease risk allele frequency in LADA cohort	Published effect size	Power to detect signal	
					α = 0.00000005	α = 0.05
rs7454108	HLA_DQB1	--	--	--	--	--
rs9272346	*MHC*	A	0.58	5.58	**100%**	**100%**
rs2187668	HLA_DRB1	--	--	--	--	--
rs3129889	HLA_DRB1_15	A	0.88	14.88	**100%**	**100%**
rs6679677	*PTPN22*	A	0.12	1.82	**100%**	**100%**
rs689	*INS,IGF2*	T	0.71	2.39	**100%**	**100%**
rs17696736	*SH2B3*	G	0.45	1.34	**100%**	**100%**
rs11755527	*BACH2*	G	0.44	1.13	3.80%	**95.70%**
rs597325	*BACH2*	G	0.58	1.19	38.60%	**100%**
rs72928038	*BACH2*	A	0.18	1.2	12.70%	**99.10%**
rs229541	*IL2RB*	A	0.42	1.12	2%	**92%**
rs941576	*DLK1*	A	0.45	1.14	6.60%	**97.50%**
rs11571316	*CTLA4*	G	0.38	1.22	66.40%	**100%**
rs7090530	*IL2RA*	A	0.59	1.22	66.30%	**100%**
rs72727394	*RASGRP1*	T	0.20	1.15	2.10%	**92.80%**
rs61839660	*IL2RA*	C	0.08	1.62	**100%**	**100%**
rs4763879	*CD69*	A	0.38	1.09	-	71.70%
rs9653442	*AFF3*	C	0.52	1.11	1.10%	**88.20%**
rs10492166	*CD69*	G	0.48	1.15	10.90%	**98.80%**
rs12708716	*CLEC16A*	A	0.34	1.23	70.90%	**100%**
rs10795791	*IL2RA*	G	0.42	1.16	15.70%	**99.30%**
rs1738074	*RSPH3*	C	0.58	1.09	-	**72.70%**
rs1615504	*DOK6*	T	0.51	1.13	3.90%	**95.80%**
rs12453507	*FBXL20*	G	0.51	1.11	1.10%	**88.30%**
rs9585056	*UBAC2*	C	0.71	1.12	-	**86.90%**
rs2542151	*PTPN2*	G	0.77	1.3	**85.30%**	**100%**
rs1990760	*GCG, FAP, IFIH1, GCA, KCNH7*	T	0.61	1.15	9%	**98.40%**
rs7804356	*SKAP2*	T	0.22	1.14	1.70%	**97.50%**
rs10509540	*RNLS*	T	0.29	1.14	3.40%	**95.20%**
rs3024493	*MAPKAPK2*	C	0.16	1.22	16.00%	**99.40%**
rs11170466	*ITGB7*	T	0.07	1.19	-	**78.40%**
rs911263	*RAD51B*	T	0.67	1.12	1.20%	**89.20%**
rs6476839	*GLIS3*	T	0.41	1.12	1.90%	**92.20%**
rs12908309	*RASGRP1*	G	0.24	1.19	18.40%	**99.50%**
rs193778	*CLEC16A*	G	0.25	1.14	2.30%	**93.40%**
rs11203203	*UBASH3A*	A	0.38	1.14	5.50%	**97.10%**
rs5753037	*NF2*	T	0.38	1.11	-	**86.70%**

rs1456988	*C14orf64*	G	0.68	1.12	1.10%	**88.70 %**
rs11258747	*PRKCQ*	G	0.77	1.45	**100%**	**100%**
rs602662	*FUT2*	A	0.51	1.12	2.10%	**92.80 %**
rs2281808	*SIRPD*	C	0.63	1.11	-	**85.80 %**
rs10877012	*CYP27B1*	G	0.34	1.22	61.40%	**100%**
rs2611215	*TMEM192*	A	0.77	1.19	14%	**99.20 %**
rs11954020	*IL7R*	G	0.41	1.11	-	**87.50 %**
rs402072	*DACT3*	T	0.15	1.15	-	**86.50 %**
rs1538171	*CENPW*	G	0.47	1.12	2%	**92.80 %**
rs10272724	*IKZF1*	T	0.27	1.15	4.70%	**96.60 %**
rs11301008 1	*CCR5*	T	0.12	1.19	2.50%	**93.70 %**
rs7928968	*INS,IGF2*	T	0.25	1.25	67.60%	**100%**
rs3825932	*CHRNB4*	T	0.35	1.16	12.80%	**99.10 %**
rs2304256	*ICAM1*	C	0.28	1.16	8.30%	**98.30 %**
rs10517086	*SLC34A2*	A	0.29	1.09	-	66.10 %
rs601338	*FUT2*	A	0.54	1.34	**100%**	**100%**
rs9924471	*SBK1*	A	0.15	1.25	31.60%	**99.90 %**
rs6691977	*KIF14*	C	0.21	1.13	-	**86%**
rs924043	*WDR27*	C	0.80	1.19	9.00%	**98.50 %**
rs478222	*DNMT3A*	A	0.42	1.15	10.20%	**98.7**
rs722988	*ITGB1*	C	0.38	1.11	-	**86.70 %**
rs4849135	*ACOXL*	G	0.67	1.12	1.20%	**89.20 %**
rs6827756	*KIAA1109*	T	0.60	1.13	3.20%	**95%**
rs6920220	*TNFAIP3*	A	0.22	1.12	-	**81.70 %**
rs694739	*BAD*	A	0.38	1.05	-	30%
rs12148472	*CHRNB4*	T	0.12	1.2	3.70%	**95.50 %**
rs1465788	*RAD51B*	C	0.69	1.16	9.00%	**98.40 %**
rs4948088	*COBL*	C	0.84	1.3	55.50%	**100%**
rs7221109	*CCR7*	C	0.61	1.05	-	30%

Supp. Table 5-9B. Power to detect T1D loci in LADACTR

SNP	chr:pos	reference_allele	other_allele	eaf	OR	OR_se	OR_95L	OR_95U	z	P	-log10_p-value	q_statistic	q_p-value	I2	n_studies	n_samples	effects	LOCUS	AverageCall	Rsq	Genotyped
rs10842994	12:27965150	C	T	0.199	1.169	0.037	1.070	1.278	-3.444	0.001	3.198	1.896	0.594	0.000	4	9415	---	*KLHDC5*	0.998	0.987	Imputed
rs10401969	19:19407718	T	C	0.088	0.820	0.072	0.726	0.927	3.176	0.002	2.788	0.806	0.848	0.000	4	9415	++++	*CILP2*	0.999	0.988	Imputed
rs10203174	2:43690030	C	T	0.110	1.193	0.046	1.064	1.339	-3.016	0.003	2.558	3.069	0.381	0.023	4	9414	---	*THADA*	0.999	0.993	Imputed
rs11634397	15:80432222	G	A	0.642	1.121	0.041	1.040	1.208	2.976	0.003	2.502	5.130	0.162	0.415	4	9415	+++-	*ZFAND6*	0.977	0.923	Imputed
rs8108269	19:46158513	G	T	0.308	1.112	0.042	1.030	1.200	2.716	0.007	2.153	2.231	0.526	0.000	4	9414	+++-	*GIPR*	1.000	0.999	Genotyped
rs7903146	10:114758349	T	C	0.280	1.108	0.042	1.025	1.197	2.575	0.011	1.974	4.975	0.174	0.397	4	9415	+-++	*TCF7L2*	0.972	0.896	Imputed
rs10923931	1:120517959	T	G	0.102	1.160	0.064	1.034	1.302	2.522	0.012	1.909	4.832	0.185	0.379	4	9415	++-+	*NOTCH2*	0.990	0.930	Imputed
rs849135	7:28196413	G	A	0.476	1.088	0.032	1.014	1.167	-2.349	0.020	1.704	5.835	0.120	0.486	4	9414	+---	*JAZF1*	0.999	0.998	Imputed
rs9936385	16:53819169	C	T	0.411	1.088	0.038	1.014	1.168	2.335	0.020	1.689	4.717	0.194	0.364	4	9415	++++	*FTO*	1.000	0.999	Genotyped
rs163184	11:2847069	G	T	0.488	1.087	0.038	1.013	1.166	2.326	0.021	1.678	2.877	0.411	0.000	4	9415	++++	*KCNQ1*	0.976	0.925	Imputed
rs7403531	15:38822905	T	C	0.713	1.111	0.045	1.002	1.231	-2.006	0.047	1.332	2.755	0.252	0.274	3	7047	-?--	*RASGRP1*	0.969	0.861	Imputed
rs4275659	12:123447928	C	T	0.675	1.081	0.042	0.999	1.169	1.943	0.054	1.270	7.983	0.046	0.624	4	9415	-+++	*MPHOSPH9*	0.970	0.886	Imputed
rs516946	8:41519248	C	T	0.724	1.086	0.045	0.998	1.183	1.909	0.058	1.235	3.287	0.350	0.087	4	9414	++++	*ANK1*	0.970	0.886	Imputed
rs2796441	9:84308948	G	A	0.406	1.072	0.033	0.998	1.152	-1.907	0.058	1.233	4.573	0.206	0.344	4	9414	-+--	*TLE1*	1.000	0.999	Genotyped
rs17168486	7:14898282	T	C	0.178	1.092	0.049	0.996	1.196	1.877	0.063	1.204	3.763	0.288	0.203	4	9415	+++-	*DGKB*	0.995	0.978	Imputed
rs2943640	2:227093585	C	A	0.623	1.069	0.039	0.993	1.150	1.779	0.077	1.111	0.775	0.855	0.000	4	9414	++++	*IRS1*	0.998	0.991	Imputed
rs6878122	5:76427311	G	A	0.681	1.072	0.036	0.991	1.159	-1.744	0.083	1.079	7.855	0.049	0.618	4	9415	-++-	*ZBED3*	0.991	0.966	Imputed
rs4458523	4:6289986	G	T	0.570	1.066	0.038	0.992	1.145	1.743	0.084	1.078	0.094	0.993	0.000	4	9414	++++	*WFS1*	0.996	0.987	Imputed
rs6723108	2:135479980	T	G	0.486	1.065	0.038	0.991	1.144	1.709	0.090	1. 047	2.095	0.553	0.000	4	9415	++++	*TMEM163*	0.939	0.803	Imputed

rs4402960	3:185511687	T	G	0.301	1.068	0.040	0.988	1.153	1.663	0.099	1.005	1.513	0.679	0.000	4	9415	+-++	*IGF2BP*	0.999	0.997	Imputed
rs7756992	6:20679709	G	A	0.283	1.068	0.041	0.988	1.154	1.661	0.099	1.003	1.454	0.693	0.000	4	9415	+++-	*CDKAL1*	0.997	0.989	Imputed
rs1552224	11:72433098	A	C	0.171	0.926	0.049	0.843	1.018	1.595	0.113	0.946	5.509	0.138	0.455	4	9415	++-+	*ARAP1 (CENTD2)*	1.000	1.000	Genotyped
rs2334499	11:1696849	T	C	0.424	0.947	0.033	0.882	1.016	-1.511	0.134	0.874	1.056	0.788	0.000	4	9414	---	*DUSP8*	0.970	0.906	Imputed
rs10811661	9:22134094	T	C	0.160	1.079	0.044	0.978	1.190	-1.509	0.134	0.873	4.491	0.213	0.332	4	9414	+---	*CDKN2A/B*	0.998	0.988	Imputed
rs243088	2:60568745	T	A	0.475	1.054	0.037	0.981	1.131	1.440	0.153	0.816	3.395	0.335	0.116	4	9414	+-+-	*BCL11A*	0.995	0.984	Imputed
rs6813195	4:153520475	T	C	0.284	0.945	0.036	0.874	1.021	-1.430	0.156	0.808	4.347	0.226	0.310	4	9414	---+	*TMEM154*	0.999	0.996	Genotyped
rs12970134	18:57884750	A	G	0.277	1.055	0.041	0.975	1.142	1.327	0.188	0.727	3.175	0.365	0.055	4	9415	++--	*MC4R*	1.000	0.999	Genotyped
rs12899811	15:91544076	A	G	0.311	0.952	0.039	0.883	1.028	1.257	0.212	0.673	1.999	0.573	0.000	4	9414	+-++	*PRC1*	0.920	0.740	Imputed
rs17106184	1:50909985	G	A	0.104	1.077	0.052	0.959	1.209	-1.254	0.213	0.671	1.154	0.764	0.000	4	9414	---	*FAF1*	0.997	0.976	Imputed
rs2075423	1:214154719	G	T	0.345	1.047	0.035	0.973	1.127	-1.236	0.220	0.658	2.434	0.487	0.000	4	9414	-+--	*PROX1*	0.987	0.956	Imputed
rs17791513	9:81905590	A	G	0.071	1.091	0.060	0.950	1.251	-1.235	0.220	0.657	0.917	0.821	0.000	4	9414	--+-	*TLE4*	0.992	0.903	Imputed
rs7845219	8:95937502	T	C	0.451	1.044	0.033	0.972	1.120	-1.186	0.239	0.621	1.336	0.720	0.000	4	9414	---	*TP53INP1*	0.998	0.995	Imputed
rs6808574	3:187740523	T	C	0.594	0.959	0.037	0.892	1.030	1.148	0.254	0.594	4.030	0.258	0.256	4	9414	++-+	*LPP*	0.988	0.963	Imputed
rs7041847	9:4287466	A	G	0.496	1.039	0.033	0.968	1.115	-1.060	0.293	0.533	4.199	0.241	0.286	4	9414	+---	*GLIS3*	0.993	0.977	Imputed
rs7163757	15:62391608	C	T	0.444	0.965	0.036	0.899	1.036	0.979	0.331	0.480	4.202	0.240	0.286	4	9414	+++-	*C2CD4A*	1.000	1.000	Genotyped
rs9470794	6:38106844	C	T	0.080	0.938	0.058	0.826	1.067	-0.971	0.335	0.475	0.901	0.825	0.000	4	9414	--+-	*ZFAND3*	0.995	0.953	Imputed
rs6795735	3:64705365	T	C	0.404	0.966	0.034	0.899	1.037	-0.960	0.341	0.468	3.986	0.263	0.247	4	9414	+---	*ADAMTS9*	0.994	0.981	Imputed
rs459193	5:55806751	G	A	0.703	1.040	0.041	0.959	1.127	0.946	0.348	0.459	4.414	0.220	0.320	4	9414	++-+	*ANKRD55*	1.000	0.999	Genotyped
rs831571	3:64048297	C	T	0.761	1.045	0.047	0.953	1.146	0.940	0.351	0.455	0.331	0.954	0.000	4	9414	++++	*PSMD6*	0.990	0.954	Imputed
rs2261181	12:66212318	T	C	0.086	1.058	0.064	0.933	1.199	0.883	0.381	0.420	0.891	0.828	0.000	4	9415	+++-	*HMGA2*	0.997	0.979	Imputed
rs7178572	15:77747190	G	A	0.674	0.966	0.037	0.894	1.044	-0.878	0.383	0.416	5.408	0.144	0.445	4	9414	+-+-	*HMG20A*	0.997	0.988	Imputed
rs11717195	3:123082398	T	C	0.238	1.034	0.039	0.952	1.122	-0.788	0.434	0.362	1.328	0.723	0.000	4	9414	---+	*ADCY5*	0.996	0.981	Imputed

rs3786897	19:33893008	A	G	0.428	1.029	0.034	0.958	1.105	-0.783	0.437	0.360	0.472	0.925	0.000	4	9414	--+-	*PEPD*	0.964	0.883	Imputed
rs1359790	13:80717156	G	A	0.265	0.970	0.040	0.895	1.050	0.755	0.454	0.343	0.545	0.909	0.000	4	9414	+-++	*SPRY2*	1.000	0.999	Genotyped
rs10830963	11:92708710	C	G	0.284	1.031	0.037	0.953	1.115	-0.755	0.454	0.343	0.220	0.974	0.000	4	9414	---+	*MTNR1B*	0.967	0.874	Imputed
rs1111875	10:94462882	C	T	0.408	1.027	0.034	0.957	1.103	-0.744	0.460	0.337	5.037	0.169	0.404	4	9415	--++	*HHEX/IDE*	0.999	0.995	Imputed
rs4430796	17:36098040	A	G	0.525	1.026	0.036	0.956	1.102	0.723	0.473	0.325	3.262	0.353	0.080	4	9415	--++	*HNF1B*	0.981	0.941	Imputed
rs7612463	3:23336450	A	C	0.125	1.038	0.054	0.933	1.155	0.684	0.497	0.303	0.730	0.866	0.000	4	9415	-+++	*UBE2E2*	1.000	0.999	Genotyped
rs1802295	10:70931474	T	C	0.340	0.975	0.036	0.905	1.050	-0.667	0.508	0.294	0.997	0.802	0.000	4	9414	--++	*VPS26A*	1.000	0.999	Genotyped
rs12427353	12:121426901	G	C	0.198	1.023	0.042	0.938	1.117	-0.519	0.607	0.217	13.097	0.004	0.771	4	9414	+-++	*HNF1A*	0.998	0.987	Imputed
rs3802177	8:118185025	G	A	0.312	0.981	0.038	0.910	1.058	0.496	0.622	0.206	5.188	0.159	0.422	4	9415	++-+	*SLC30A8*	0.928	0.753	Imputed
rs702634	5:53271420	A	G	0.657	0.982	0.037	0.911	1.059	-0.466	0.644	0.191	0.849	0.838	0.000	4	9414	+-+-	*ARL15*	0.999	0.998	Genotyped
rs11063069	12:4374373	G	A	0.206	1.019	0.043	0.934	1.111	0.422	0.675	0.170	1.818	0.611	0.000	4	9414	+-++	*CCND2*	0.908	0.592	Imputed
rs4812829	20:42989267	A	G	0.170	0.981	0.044	0.893	1.076	-0.413	0.682	0.166	1.074	0.783	0.000	4	9414	---+	*HNF4A*	0.994	0.962	Imputed
rs7955901	12:71433293	T	C	0.545	1.015	0.035	0.946	1.088	0.401	0.690	0.161	6.318	0.097	0.525	4	9414	+-++	*TSPAN8*	0.998	0.993	Imputed
rs13233731	7:130437689	G	A	0.495	0.986	0.035	0.919	1.058	0.382	0.704	0.152	2.679	0.444	0.000	4	9414	+-+-	*KLF14*	1.000	1.000	Imputed
rs6467136	7:127164958	G	A	0.529	1.007	0.035	0.939	1.080	0.192	0.849	0.071	3.862	0.277	0.223	4	9414	+---	*GCC1*	0.998	0.995	Imputed
rs10278336	7:44245363	G	A	0.435	1.006	0.035	0.938	1.080	0.176	0.861	0.065	1.572	0.666	0.000	4	9414	-+++	*GCK*	0.994	0.980	Imputed
rs780094	2:27741237	C	T	0.607	1.004	0.036	0.934	1.081	0.118	0.906	0.043	0.486	0.922	0.000	4	9415	+++-	*GCKR*	1.000	0.999	Genotyped
rs5215	11:17408630	T	C	0.597	0.996	0.035	0.926	1.070	-0.116	0.908	0.042	1.988	0.575	0.000	4	9414	+-+-	*KCNJ11*	0.998	0.994	Imputed
rs2028299	15:90374257	C	A	0.675	1.004	0.038	0.929	1.086	-0.108	0.915	0.039	0.135	0.987	0.000	4	9414	-+-+	*AP3S2*	0.995	0.982	Imputed
rs16861329	3:186666461	C	T	0.122	1.004	0.051	0.903	1.117	-0.077	0.939	0.027	4.751	0.191	0.369	4	9414	+-+-	*ST64GAL1*	0.980	0.872	Imputed
rs9505118	6:7290437	G	A	0.412	1.002	0.035	0.933	1.076	0.057	0.955	0.020	2.601	0.457	0.000	4	9414	+++-	*SSR1/RREB1*	1.000	1.000	Genotyped
rs7593730	2:161171454	C	T	0.745	0.999	0.043	0.916	1.090	-0.020	0.984	0.007	0.853	0.837	0.000	4	9415	-++-	*RBMS1*	1.000	1.000	Genotyped
rs12571751	10:80942631	A	G	0.463	1.000	0.035	0.932	1.073	0.004	0.997	0.001	2.753	0.431	0.000	4	9414	++-+	*ZMIZ1*	1.000	0.999	Genotyped

Supp. Table 5-10A. T2D loci in LADACTR

SNP	LOCUS	AverageCall	Rsq	Genotyped	Risk Allele	Disease risk allele frequency in LADA cohort	Published effect size	Power to detect signal	
								α = 0.00000005	α = 0.05
rs10842994	*KLHDC5*	0.99757	0.98748	Imputed	C	0.20	1.09	0.10%	55.60%
rs10401969	*CILP2*	0.99885	0.98822	Imputed	C	0.91	1.13	-	54.70%
rs10203174	*THADA*	0.99902	0.99292	Imputed	C	0.11	1.15	-	76.70%
rs11634397	*ZFAND6*	0.97661	0.92318	Imputed	G	0.64	1.09	-	70.10%
rs8108269	*GIPR*	0.99968	0.99879	Genotyped	G	0.31	1.06	-	37.20%
rs7903146	*TCF7L2*	0.97237	0.89642	Imputed	T	0.28	1.4	**100.00%**	**100%**
rs10923931	*NOTCH2*	0.99031	0.93035	Imputed	T	0.10	1.11	-	48.30%
rs849135	*JAZF1*	0.99948	0.99825	Imputed	G	0.48	1.12	2.10%	**92.80%**
rs9936385	*FTO*	0.9998	0.99944	Genotyped	C	0.41	1.13	3.50%	**95.40%**
rs163184	*KCNQ1*	0.97639	0.92469	Imputed	G	0.49	1.09	-	74%
rs7403531	*RASGRP1*	0.96912	0.86089	Imputed	T	0.71	1.02	-	8.40%
rs4275659	*MPHOSPH9*	0.97022	0.88621	Imputed	C	0.67	1.07	-	48.20%
rs516946	*ANK1*	0.97022	0.88621	Imputed	C	0.67	1.07	-	48.20%
rs2796441	*TLE1*	0.99962	0.99875	Genotyped	G	0.41	1.07	-	52.20%
rs17168486	*DGKB*	0.99547	0.97845	Imputed	T	0.18	1.13	-	**81.90%**
rs2943640	*IRS1*	0.99758	0.99089	Imputed	C	0.62	1.09	-	71.20%
rs6878122	*ZBED3*	0.99149	0.96632	Imputed	G	0.68	1.13	3.00%	92.80%

rs4458523	*WFS1*	0.99578	0.98712	Imputed	G	0.57	1.09	-	73%
rs6723108	*TMEM163*	0.93865	0.80301	Imputed	T	0.49	1.01	-	6%
rs4402960	*IGF2BP*	0.99909	0.99666	Imputed	T	0.30	1.13	2.90%	**92.70%**
rs7756992	*CDKAL1*	0.9969	0.98911	Imputed	G	0.28	1.2	33.00%	**99.90%**
rs1552224	*ARAP1 (CENTD2)*	0.99994	0.99965	Genotyped	A	0.17	1.13	-	**80.20%**
rs2334499	*DUSP8*	0.96986	0.90614	Imputed	T	0.42	1.07	-	52.40%
rs10811661	*CDKN2A/B*	0.99828	0.98842	Imputed	T	0.16	1.19	6.40%	**97.50%**
rs243088	*BCL11A*	0.99521	0.98421	Imputed	T	0.47	1.09	-	73.90%
rs6813195	*TMEM154*	0.99889	0.99601	Genotyped	C	0.72	1.08	-	54.60%
rs12970134	*MC4R*	0.99981	0.99923	Genotyped	A	0.28	1.08	-	55.50%
rs12899811	*PRC1*	0.91964	0.74026	Imputed	G	0.69	1.09	-	66.80%
rs17106184	*FAF1*	0.99712	0.97558	Imputed	G	0.10	1.11	-	48.30%
rs2075423	*PROX1*	0.98715	0.95577	Imputed	G	0.35	1.08	-	60.80%
rs17791513	*TLE4*	0.99184	0.90305	Imputed	A	0.07	1.21	1.10%	**85.50%**
rs7845219	*TP53INP1*	0.99821	0.99509	Imputed	T	0.45	1.08	-	63.90%
rs6808574	*LPP*	0.98842	0.96312	Imputed	C	0.41	1.08	-	63%
rs7041847	*GLIS3*	0.99314	0.97714	Imputed	A	0.50	1.05	-	31.40%
rs7163757	*C2CD4A*	0.99994	0.99978	Genotyped	C	0.44	1.06	-	41.70%
rs9470794	*ZFAND3*	0.99457	0.95266	Imputed	T	0.92	1.01	-	5.30%
rs6795735	*ADAMTS9*	0.99419	0.981	Imputed	C	0.40	1.07	-	51.90%

rs459193	*ANKRD55*	0.9996	0.9985 2	Genotyped	G	0.70	1.05	-	27%
rs831571	*PSMD6*	0.99034	0.9543 1	Imputed	C	0.76	1.03	-	11.80%
rs2261181	*HMGA2*	0.99715	0.9792 5	Imputed	T	0.09	1.16	-	74.40%
rs7178572	*HMG20A*	0.99704	0.9877 9	Imputed	G	0.67	1.08	-	58.60%
rs1171719 5	*ADCY5*	0.99606	0.9811 7	Imputed	T	0.24	1.09	-	61%
rs3786897	*PEPD*	0.96442	0.8830 8	Imputed	A	0.43	1.02	-	9.10%
rs1359790	*SPRY2*	0.99976	0.9990 1	Genotyped	G	0.27	1.11	-	**80.50%**
rs1083096 3	*MTNR1B*	0.96662	0.8736 1	Imputed	G	0.72	1.11	-	**80.20%**
rs1111875	*HHEX/IDE*	0.99865	0.9948 3	Imputed	C	0.41	1.15	10.00%	**98.60%**
rs4430796	*HNF1B*	0.98054	0.9407 5	Imputed	G	0.47	1.13	3.90%	**95.80%**
rs7612463	*UBE2E2*	0.99979	0.9988 1	Genotyped	C	0.88	1.11	-	52.50%
rs1802295	*VPS26A*	0.99982	0.9991 8	Genotyped	T	0.34	1.02	-	8.80%
rs1242735 3	*HNF1A*	0.9977	0.9870 3	Imputed	G	0.20	1.12	-	**95.30%**
rs3802177	*SLC30A8*	0.92842	0.7528 9	Imputed	G	0.31	1.16	14.10%	**98.70%**
rs702634	*ARL15*	0.99949	0.9981 3	Genotyped	A	0.66	1.08	-	59.20%
rs1106306 9	*CCND2*	0.90818	0.5922 4	Imputed	G	0.21	1.11	-	73.70%
rs4812829	*HNF4A*	0.99378	0.9615 8	Imputed	A	0.17	1.07	-	34%
rs7955901	*TSPAN8*	0.99777	0.9933 5	Imputed	C	0.46	1.09	-	73.90%
rs1323373 1	*KLF14*	0.99989	0.9996 3	Imputed	G	0.50	1.11	1.10%	**88.30%**
rs6467136	*GCC1*	0.99828	0.9948 5	Imputed	A	0.47	1.01	-	6%

rs10278336	*GCK*	0.99398	0.98044	Imputed	A	0.44	1.05	-	31%
rs780094	*GCKR*	0.99984	0.99937	Genotyped	C	0.61	1.04	-	21.10%
rs5215	*KCNJ11*	0.9982	0.9941	Imputed	C	0.40	1.08	-	62.70%
rs2028299	*AP3S2*	0.99535	0.98243	Imputed	C	0.67	1.04	-	19.90%
rs16861329	*ST64GAL1*	0.97976	0.87244	Imputed	C	0.12	1.03	-	9%
rs9505118	*SSR1/RREB1*	0.99994	0.99978	Genotyped	A	0.59	1.07	-	51.80%
rs7593730	*RBMS1*	0.99999	0.99996	Genotyped	C	0.75	1.11	-	77.20%
rs12571751	*ZMIZ1*	0.99952	0.99862	Genotyped	A	0.46	1.09	-	73.80%

Supp. Table 5-10B. Power to detect T2D loci in LADACTRL

Stratification by top 1 GAD tertile, top 2 GAD tertiles, or bottom 1 GAD tertile, each vs. population controls, in ActionLADA and Swedish LADA cases. Only genome-wide significant loci are shown.																
Top 1 GAD titer tertile																
Rs number	**Ref. allele**	**Other allele**	**eaf**	**OR**	**OR se**	**OR 95L**	**OR 95U**	**z**	**P**	**-log10 P**	**Q statistic**	**Q P**	**i2**	**N studies**	**N samples**	**effects**
6:32626475	A	G	0.313	3.30	0.25	2.81	3.88	14.53	1.89E-47	46.72	81.60	0.000	0.988	2.00	4940	++
12:112553032	A	T	0.450	1.56	0.10	1.37	1.77	6.73	2.01E-11	10.70	0.04	0.836	0.000	2.00	4940	++
11:2181060	G	T	0.732	1.55	0.11	1.34	1.79	5.96	2.95E-09	8.53	1.11	0.293	0.096	2.00	4940	++
1:114415368	C	G	0.780	0.64	0.05	0.55	0.75	-5.84	5.95E-09	8.23	0.62	0.430	0.000	2.00	4940	--
Top 2 GAD titer tertiles																
6:32626475	A	G	0.326008	3.110	0.192	2.733	3.538	17.23	7.86E-66	65.105	105.897	0.000	0.991	2.000	5425.000	++
11:2182224	T	A	0.738589	1.563	0.087	1.393	1.754	7.60	4.18E-14	13.379	2.612	0.106	0.617	2.000	5425.000	++
1:114377568	G	A	0.883128	0.589	0.044	0.503	0.689	-6.60	5.38E-11	10.269	4.866	0.027	0.794	2.000	5425.000	--
12:112553032	A	T	0.452865	1.362	0.068	1.228	1.510	5.86	5.57E-09	8.254	0.062	0.803	0.000	2.000	5425.000	++
Bottom 1 GAD titer tertile																
6:32626475	A	G	0.294	2.421	0.186	2.06	2.85	10.64	2.13E-26	25.672	7.158	0.007	0.860	2.000	4875.000	++
TCF7L2(10:114758349) in GAD tertile analyses																
Top 1 GAD titer tertile	T	C	0.278649	1.050	0.070	0.914	1.207	0.687	0.493	0.307	0.087	0.768	0.000	2.000	4941.000	++
Top 2 GAD titer tertiles	T	C	0.280189	1.063	0.058	0.950	1.189	1.067	0.288	0.540	0.899	0.343	0.000	2.000	5426.000	+-
Bottom 1 GAD titer tertile	T	C	0.273435	1.093	0.076	0.944	1.265	1.186	0.236	0.628	3.734	0.053	0.732	2.000	4876.000	+-

Supp. Table 5-11. GAD tertiles

DRB1-DQA1-DQB1	BMDCS	LADA	T1D	LADA vs BMDCS OR	LADA vs BMDCS *p*	LADA vs T1D OR	LADA vs T1D *p*
01:01-01:01-05:01	0.097	0.069	0.054	0.7[0.58-0.84]	0.00012548	0.76[0.63-0.91]	0.0029333
03:01-05:01-02:01	0.117	0.223	0.334	2.17[1.86-2.53]	< 2.22e-16	1.75[1.58-1.93]	< 2.22e-16
04:01-03:01-03:02	0.053	0.193	0.228	4.32[3.51-5.35]	< 2.22e-16	1.23[1.11-1.37]	0.00011523
04:04-03:01-03:02	0.038	0.043	0.082	1.13[0.86-1.5]	0.36897	2.01[1.66-2.44]	7.05E-14
04:05-03:01-03:02	0.004	0.009	0.022	2.27[1.04-5.66]	0.030787	2.56[1.71-3.88]	9.08E-07
07:01-02:01-02:02	0.093	0.074	0.041	0.78[0.65-0.94]	0.0086422	0.54[0.44-0.66]	3.07E-10
07:01-02:01-03:03	0.041	0.014	0.002	0.33[0.23-0.47]	1.19E-11	0.16[0.07-0.33]	5.53E-09
11:01-05:01-03:01	0.071	0.027	0.007	0.36[0.28-0.47]	< 2.22e-16	0.24[0.15-0.37]	1.47E-12
13:01-01:03-06:03	0.058	0.049	0.012	0.84[0.67-1.07]	0.14127	0.23[0.16-0.31]	< 2.22e-16
14:01-01:01-05:03	0.018	0.006	0.001	0.31[0.18-0.53]	1.52E-06	0.17[0.04-0.5]	0.00022447
15:01-01:02-06:02	0.139	0.046	0.003	0.3[0.25-0.36]	< 2.22e-16	0.05[0.02-0.1]	< 2.22e-16

Supp. Table 5-12. Established T1D HLA haplotype frequencies in LADA, T1D and controls

rs7903146-T (*TCF7L2*)

	Ncases	Ncontrols	MAF LADA cases	MAF controls	OR	P
LADACTRL_SWE_SWE	805	3126	0.301	0.271	1.182	0.009
LADACTRL_ACLADA_BMDCS	1353	1014	0.296	0.297	0.986	0.825595
LADACTRL_DAN_DAN	476	1807	0.298	0.276	1.114	0.247267
LADACTRL_HUNT	139	695	0.295	0.242	1.300	0.066926
LADA vs CTRLS meta-analysis (+HUNT)	2773	6642	0.298	0.271	1.107	0.01606

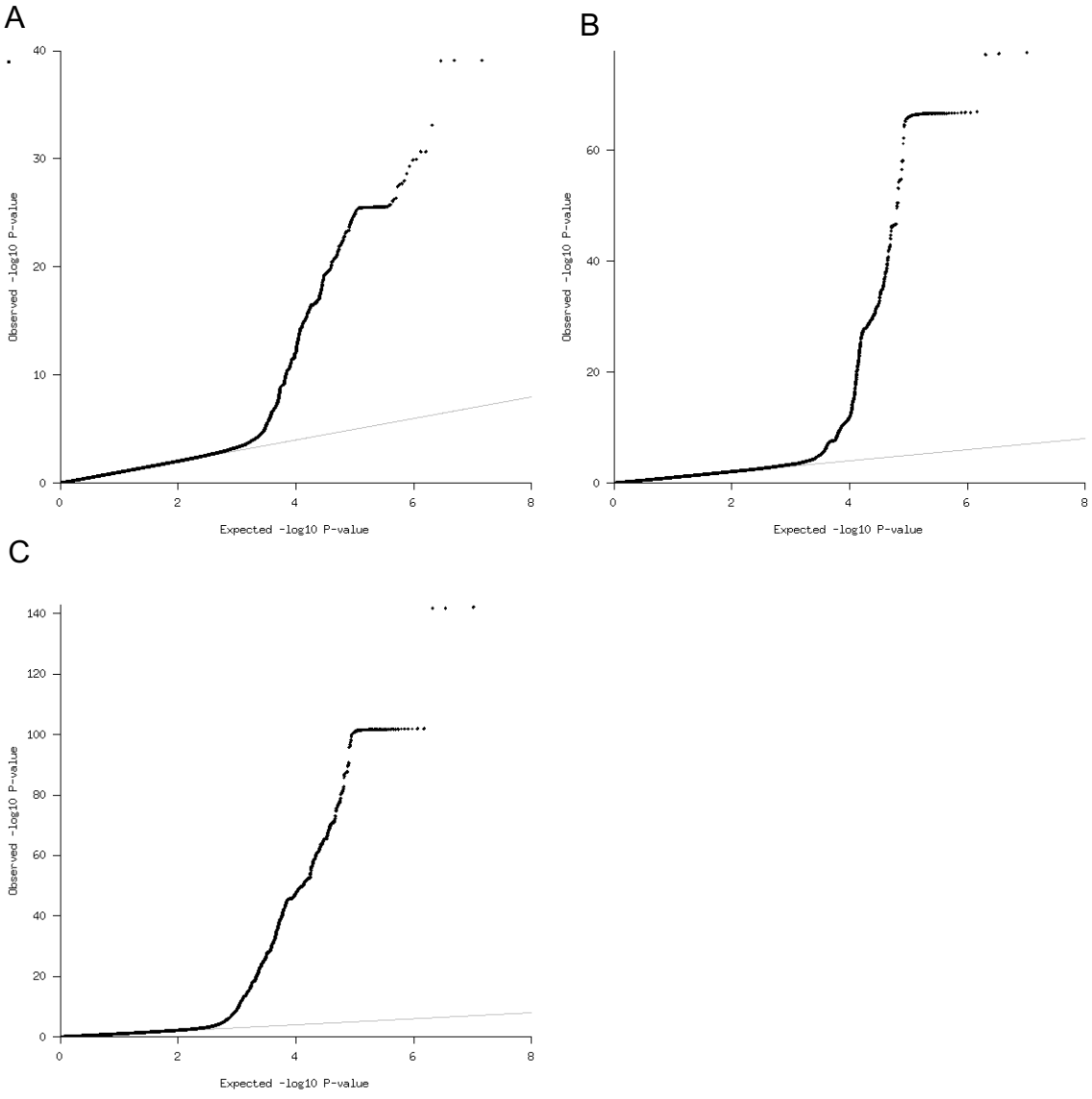
	Ncases	Ncontrols	MAF LADA cases	MAF T1D cases	OR	P
LADAT1D_FIN_FIN	296	365	0.221	0.193	1.148	0.345333
LADAT1D_SWE-CHOP_SWE	2158	603	0.298	0.242	1.300	0.0018
LADA vs T1D meta-analysis	2454	968	0.260	0.218	1.260	0.001794

	Ncases	Ncontrols	MAF LADA cases	MAF T2D cases	OR	P
LADAT2D_FIN_FIN	296	2785	0.221	0.223	0.931	0.501235
LADAT2D_SWE-CHOP_SWE	2158	2599	0.298	0.337	0.793	5.84E-06
LADAT2D_GoDARTS1	97	2098	0.298	0.345	0.806	0.166333
LADAT2D_GoDARTS2	89	2218	0.264	0.339	0.700	0.0287021
LADAT2D_HUNT	139	695	0.295	0.340	0.781	0.10056
LADA vs T2D meta-analysis	2779	10395	0.275	0.317	0.807	4.02E-07

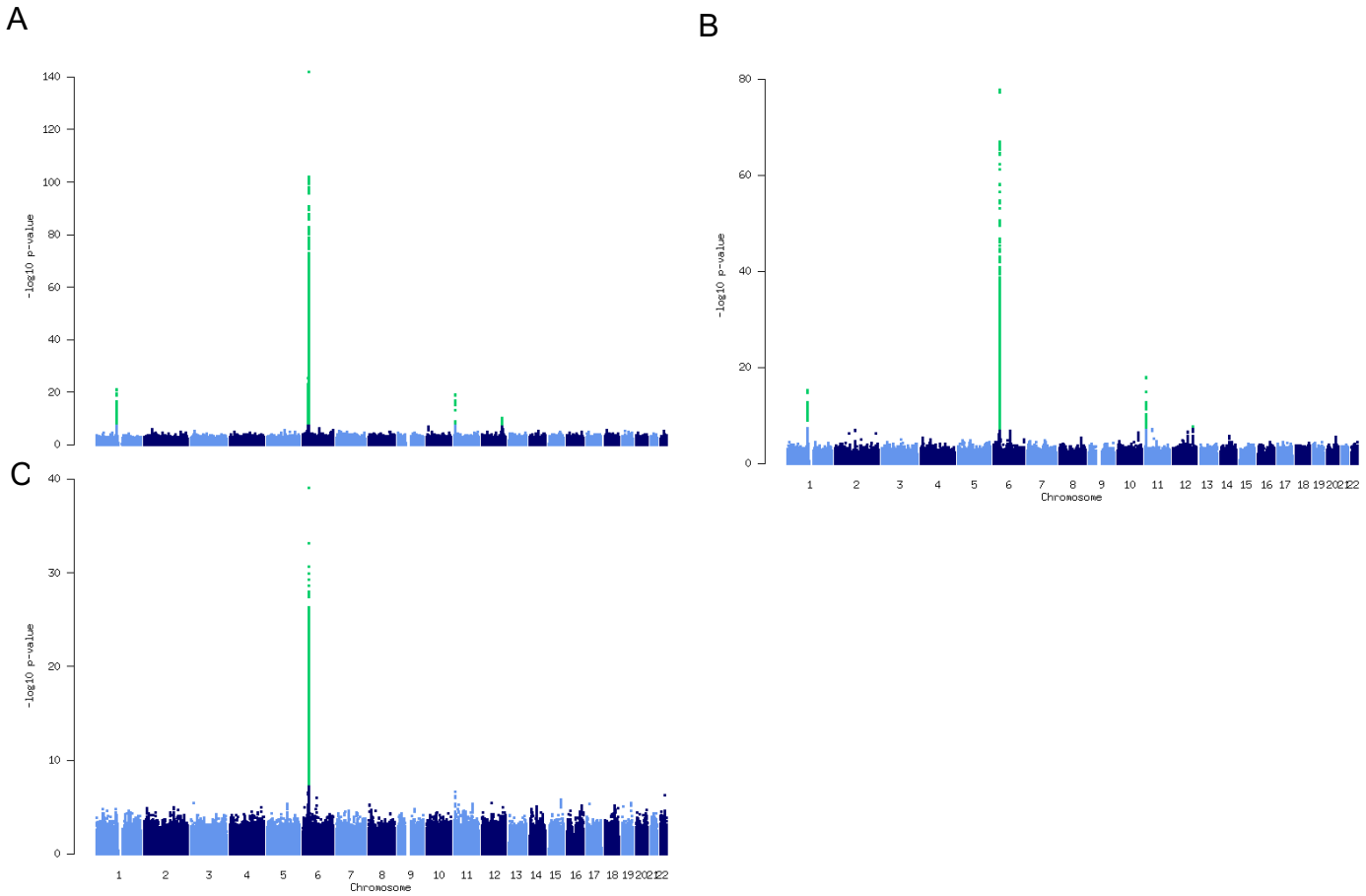
Additional sensitivity analyses	Ncases	Ncontrols	MAF LADA cases	MAFcontrols	OR	P
Danish LADA vs controls (with normal fasting glucose)	476	8341	0.3	0.26	1.18	0.017
Swedish LADA vs controls (incl. diabetics)	884	3126	0.31	0.27	1.2	0.0038
Swedish LADA vs non-diabetic controls	884	2618	0.31	0.26	1.28	0.0001

Supp. Table 5-13. *TCF7L2* associations

Supplemental Figures



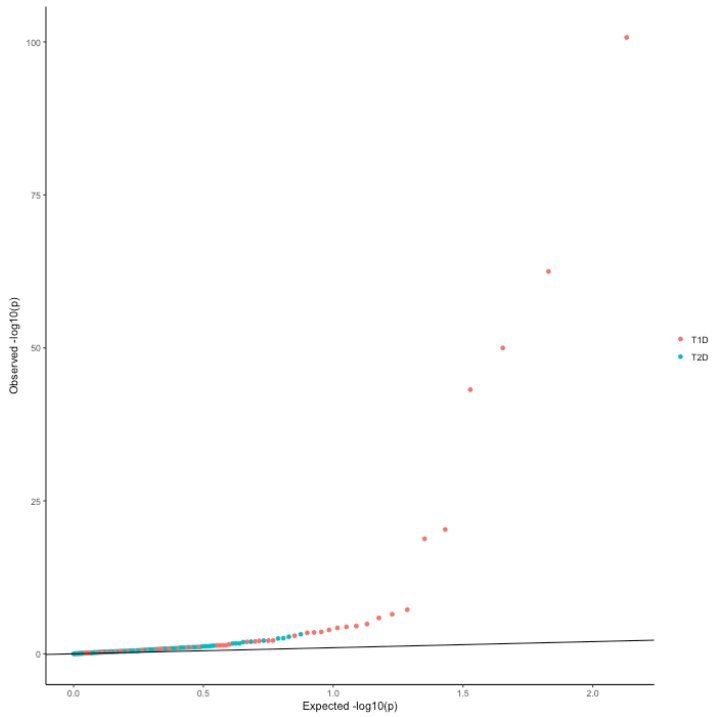
Supp. Figure 5-1. QQ plots for (A) LADA vs. population controls, (B) LADA vs. T1D, and (C) LADA vs. T2D



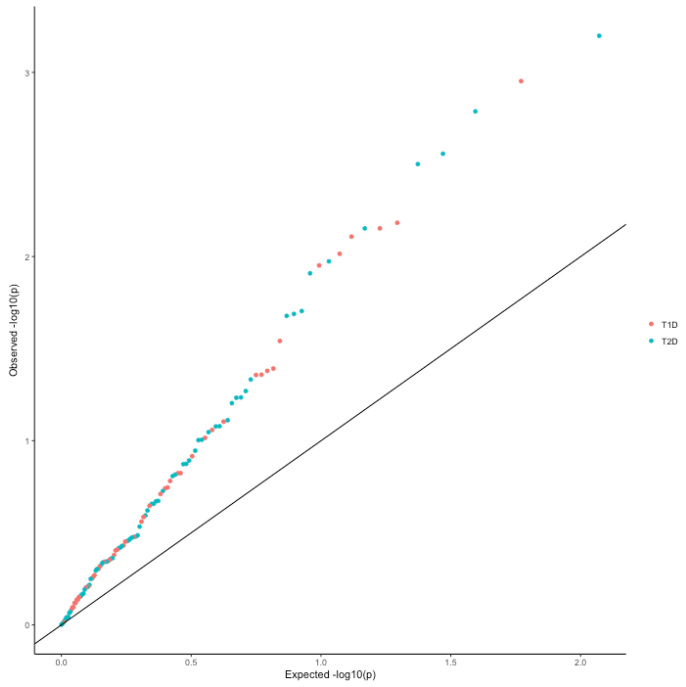
Supp. Figure 5-2. Manhattan plots for (A) LADA vs. population controls, (B) LADA vs. T1D, and (C) LADA vs. T2D

A

.



B



.

Supp. Figure 5-3. QQ plots showing established T1D (red) and T2D (blue) loci in LADA vs. population controls. A) All loci; B) zoomed in at 1 > P > 10^{-4} (excluding top T1D signals).

Supplemental Note: Cohort Information

This study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113.

Cohort name: Action Lada

Cohort type: LADA cases

Inclusion/exclusion criteria: Patients were designated with diabetes according to standard criteria, and LADA was defined as follows: patients 1) aged 30–70 years, 2) with diabetes associated autoantibodies, and 3) who did not require insulin treatment for at least 6 months post diagnosis. Type 1 autoimmune diabetic patients were defined as case subjects with diabetes and with diabetes-associated autoantibodies where insulin was started at diagnosis or within 1 month of diagnosis. Inclusion criteria for all patients were that patients have diabetes (with at least two recorded fasting blood glucose measurements >7 mmol/L), that time from diagnosis was 5 years for all patients, and that patients were aged 30–70 years at the time of recruitment. Exclusion criteria were insufficient dataset, current pregnancy, renal disease with raised creatinine or proteinuria, or acute illness at the time of testing.

Recruitment Location: Queen Mary University of London and the Action Lada consortium in Europe

Number of study subjects: 1098

Acknowledgements: We would like to acknowledge the Action Lada consortium.

Funding: This study was partially funded by the 5th Framework Programme of the European Union.

Cohort reference: REC Reference P/02/240

Conflicts of interest: No potential conflicts of interest relevant to this work.

Cohort name: Action Lada 'Plus'

Cohort type: LADA cases

Inclusion/exclusion criteria: Patients were designated with diabetes according to standard criteria, and LADA was defined as follows: patients 1) aged 30–70 years, 2) with diabetes associated autoantibodies, and 3) who did not require insulin treatment for at least 6 months post diagnosis. Type 1 autoimmune diabetic patients were defined as case subjects with diabetes and with diabetes-associated autoantibodies where insulin was started at diagnosis or within 1 month of diagnosis. Inclusion criteria for all patients were that patients have diabetes (with at least two recorded fasting blood glucose measurements >7 mmol/L), that time from diagnosis was 5 years for all patients, and that

patients were aged 30–70 years at the time of recruitment. Exclusion criteria were insufficient dataset, current pregnancy, renal disease with raised creatinine or proteinuria, or acute illness at the time of testing.

Recruitment Location: The Mayo, Rochester, MN, USA; Cornell Medical College, New York City, NY, USA; University of Alabama, Birmingham, AL, USA; University of Pennsylvania, Philadelphia, PA, USA; MODEL Clinical Research, Baltimore, MD, USA; Adventist Health System, Sunbelt Inc. d/b/a Florida Hospital, Orlando, FL, USA; Atlanta Diabetes Associates, Atlanta, GA, USA; Geisinger Health System, Danville, PA, USA; University of Leicester, Leicester, UK; T1D Exchange, Benaroya Research Institute, Seattle, WA; Health Diagnostic Lab Inc., Richmond, VA; National Disease Research Interchange, Philadelphia, PA.

Number of study subjects: 441

Acknowledgements:

Funding:

Cohort reference:

Conflicts of interest:

Cohort name: All New Diabetics In Scania (ANDIS)

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases

Inclusion/exclusion criteria:

GAD

ELISA: Negative:< 5 kE/L, Positive:=> 10 kE/L

RIA: Negative:0-34 U/ml, Positive:> 50 U/ml

LADA

Age at onset ≥ 35 years

GAD (ELISA) > 10 kE/L

GAD (RIA) >50 U/ml

Non-Scandinavian individuals excluded

Recruitment Location: Scania, Sweden

Number of study subjects: 440

Acknowledgements: We thank all the patients and the health care providers across Scania and Ostrobothnia for their support and their willingness to participate. We would also like to thank Johan Hultman, Jasmina Kravic, Maria Fälemark, Christina Rosborn, Gabriella Gremsperger, Maria Sterner, Malin Neptin, Lisa Sundman, Paula Kokko, and Ulrika Blom-Nilsson for excellent technical and administrative support. Finally we would

like to thank Rita Jedlert and Region Skåne (Scania County) as well as the ANDIS steering committee for their support.

Funding: Supported by grants from the Swedish Research Council (including project grants Dnr.521-2010-3490 and infrastructure grants Dnr. 2010-5983 and Dnr 2012-5538 to LG), Linnéus grant 349-2006-237, a strategic research grant (Exodiab Dnr 2009-1039), an ERC Advanced Research grant (GA 269045) and Academy of Finland (grants no. 263401 and 267882) to LG.

Cohort reference: http://andis.ludc.med.lu.se/

Conflicts of interest: None

Cohort name: Bone Mineral Density in Childhood Study (BMDCS)

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): population controls

Cohort description: The Bone Mineral Density in Childhood Study is a multicenter, longitudinal study of bone accrual in healthy children.

Inclusion/exclusion criteria: Only individuals of European ancestry were included.

Recruitment Location: Children's Hospital of Los Angeles (Los Angeles, CA), Cincinnati Children's Hospital Medical Center (Cincinnati, OH), Creighton University (Omaha, NE), Children's Hospital of Philadelphia (Philadelphia, PA), and Columbia University (New York, NY)

Number of study subjects: 1056

Acknowledgements: We appreciate the dedication of the study participants and their families, and the support of Dr. Karen Winer, Scientific Director of the Bone Mineral Density in Childhood Study.

Funding: The BMDCS was funded by the National Institutes of Health [grant number R01 HD58886 to B.Z. and S.G.]; the Eunice Kennedy Shriver National Institute of Child Health and Human Development [grant numbers N01-HD-1-3228, -3329, -3330, -3331, -3332, -3333]; the Clinical and Translational Science Awards Program [grant number 8 UL1 TR000077]; American Diabetes Association Grant 1-17-PDF-077 [to D.C.]; and the Institute for Translational Medicine and Theraputics (ITMAT) Transdisciplinary Program in Translational Medicine and Therapeutics (to D.C., B.V., and S.G.). The project described was supported by the National Center for Research Resources [grant number UL1RR024134], and is now at the National Center for Advancing Translational Sciences [grant number UL1TR000003]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Conflicts of interest: None

Study cohort: Copenhagen Controls

The Copenhagen Control sample is collected from two cohorts (The 1936 birth cohort and ADDITION-PRO), and comprises 1974 non-diabetic adults. The control subjects had a mean age of 64.42 (range, 34.44), and 49.9% were male.

Cohort name: The 1936 birth cohort

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): Population controls

Cohort description/inclusion/exclusion criteria: The cohort consists of all subjects born in 1936, who, on 2 April 1976, were resident in one of four municipalities nearby Glostrup Hospital, Denmark (n=695). The cohort was collected to assess the age-specific prevalence of diabetes mellitus and impaired glucose tolerance in 60-year-old individuals in 1996/97.

Recruitment Location: The samples were collected from four municipalities nearby Glostrup Hospital, Denmark (1936 control cohort). The samples were collected in 2009–2011 from four Danish research centres (Steno Diabetes Center, Aarhus University Hospital, Holstebro Hospital, and Hospital of South West Jutland, Esbjerg).

Number of study subjects: 624 non-diabetic individuals (502 NGTs, 122 IFG/IGTs)

Acknowledgements: The authors are grateful to the staff at the Centre of Preventive Medicine, and to MD, general practitioner, Professor Hanne Hollnagel Dr Med. Sci., who initiated the study of the 1936 cohort.

Funding: The collection of the cohort was financially supported by The Danish Heart Foundation and The Danish Medical Research Council.

Cohort reference: Drivsholm T. Increasing prevalence of diabetes mellitus and impaired glucose tolerance among 60-year-old Danes. *Diabet Med* 2001.

Conflicts of interest: NA

Cohort name: ADDITION-PRO

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): Controls

Cohort description/inclusion/exclusion criteria: ADDITION-PRO is a longitudinal cohort study of 2082 adults (>45 years) collected to have IGT, IFG, or NGT either with high or low risk of developing type 2 diabetes (based on information about age, sex, gestational diabetes, family history of diabetes, hypertension, BMI, and level of physical activity).

The samples were collected in 2009–2011from four Danish research centres (Steno Diabetes Center, Aarhus University Hospital, Holstebro Hospital, and Hospital of South West Jutland, Esbjerg).

Number of study subjects: 1350 non-diabetic individuals (812 NGTs, 538 IFG/IGTs)
Acknowledgements: The ADDITION-PRO study is managed by the ADDITION-DK steering committee (Torsten Lauritzen, Knut Borch-Johnsen, Annelli Sandbæk, Marit E. Jørgensen, and Daniel Witte).
Funding: The ADDITION-PRO study was funded by an unrestricted grant from the European Foundation for the Study of Diabetes/Pfizer for Research into Cardiovascular Disease Risk Reduction in Patients with Diabetes (74550801), the Danish Council for Strategic Research, internal research and equipment funds from Steno Diabetes Center and supported by research grants from the Novo Nordisk Foundation.
Cohort reference: Johansen et al. Protocol for ADDITION-PRO: a longitudinal cohort study of the cardiovascular experience of individuals at high risk for diabetes recruited from Danish primary care. *BMC Public Health* 2012.
Conflicts of interest: NA

Study cohort: Copenhagen LADA
The Copenhagen LADA sample (n=539) is collected from six cohorts (DD2, Vejle Biobank, OUH, CIMT, Inter99, and SDC). The LADA patients had a mean age of 58.32 (range, 67.31), and 56.2% were male.
The following inclusion criteria for LADA have been applied in all sub-cohorts: GADA positive, ≥20 years at the time of diagnosis, and treated without insulin for the first year after diagnosis or having fasting serum C-peptide ≥300 pmol/L at the time of investigation.
Cohort name: Danish Centre for strategic Research in Type 2 Diabetes (DD2)
Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases
Cohort description/inclusion/exclusion criteria: DD2 is nationwide cohort, enrolling patients with newly diagnosed type 2 diabetes from general practitioners and hospital specialist outpatient clinics since 2010. GADA was measured in 5966 patients, with an AESKULISA assay.
Number of study subjects: 158 LADA cases
Acknowledgements: The DD2-project partners are listed on the website ww.DD2.nu.
Funding: The DD2 study is supported by the Danish Agency for Science (grant no. 09-067009 and 09-075724), the Danish Health and Medicines Authority, the Danish Diabetes Association, and an unrestricted donation from Novo Nordisk A/S.
Cohort reference: Thomsen et al. The Danish Centre for Strategic Research in Type 2 Diabetes (DD2): Organization of diabetes care in Denmark and supplementary data sources for data collection among DD2 study participants. *Clin Epidemiol* 2012.
Conflicts of interest: NA

Cohort name: Vejle Biobank

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases

Cohort description/inclusion/exclusion criteria: The Vejle Diabetes Biobank was established as a regional Bio bank and comprises individuals with diabetes and a gender- and age-matched control population. All participants were aged between 25 and 75 years (both ages included) and residing in the former County of Vejle area on December 31, 2006. Altogether, 3320 patients with type 2 diabetes or type 1 diabetes were recruited from the central database at Vejle Hospital Laboratory Center. GADA were measured in all 3320 patients, with an AESKULISA assay.

Number of study subjects: 124 LADA cases

Acknowledgements: The laboratory technologists Britta Kristensen, Lene Juul Hansen, Annette Kaaris, Jan Johannsen, Merete Willumsen, Birgitte Henriksen, Camilla Davidsen, and Sara Egsgaard are acknowledged for their continued engagement and dedicated work.

Funding: The Vejle Biobank project was funded by the Danish Council for Independent Research/Medical Sciences, the Research Council of Vejle Hospital, the Department of Internal Medicine, Vejle Hospital, Vejle County, the Danish Research Fund, the Lions Club International Denmark, and anonymous donations.

Cohort reference: Petersen et al. Vejle Diabetes Biobank – a resource for studies of the etiologies of diabetes and its comorbidities. *Clin Epidemiol*. 2016.

Conflicts of interest: NA

Cohort name: OUH

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases

Cohort description/inclusion/exclusion criteria: The OUH LADA cohort is collected from a database of patients with diabetes newly referred to Odense University Hospital (OUH), Denmark, between 1997 and 2011. GAD autoantibodies were measured in 5,671 patients with diabetes, applying an RSR RIA assay, 279 were GADA positive, above 30 years of age, and had fasting C-peptide above 300 pmol/l. Of these DNA was available for 66.

Number of study subjects: 66 LADA cases

Acknowledgements: Department of Endocrinology, Odense University Hospital, Denmark, is acknowledged for their collection of the OUH cohort.

Funding: NA

Cohort reference: NA

Conflicts of interest: NA

Cohort name: CIMT

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases

Cohort description/inclusion/exclusion criteria: The CIMT trial is a multicenter randomized placebo controlled superiority trail conducted from 2008 to 2012 at eight hospitals in the capital region of Denmark. Inclusion criteria included diagnosis of type 2 diabetes, >30 years at diagnosis, BMI >25 kg/m^2, HbA_{1c}>7.5%, treatment with oral anti-diabetic drugs for ≥1 year, and/or insulin treatment for ≥3 months. Exclusion criteria included: major cardiovascular disease within the past 3 months, carotid artery stenosis >70%, heart failure, recent cancer, renal or liver disease, alcohol or drug abuse, unstable retinopathy, pregnancy, breastfeeding, fertile women not using anticonception, or allergy towards trial medication. Altogether, 412 type 2 diabetes patients were included in the trial and were screened for the presence of GADA with an RSR ELISA kit.

Number of study subjects: 31 LADA cases

Acknowledgements: The CIMT trial group is acknowledged for their effort in collecting and characterizing the cohort.

Funding: The CIMT study was funded by an unrestricted grant from Novo Nordisk A/S.

Cohort reference: Lundby et al. Study rationale and design of the CIMT trial: the Copenhagen Insulin and Metformin Therapy trial. *Diabetes Obes Metab* 2009.

Conflicts of interest: NA

Cohort name: Inter99

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases

Cohort description/inclusion/exclusion criteria: Inter99 is a population based intervention cohort, comprising individuals from the Copenhagen area. Altogether 6784 individuals participated in the baseline examination. GADA were measured in 2531 individuals, with an RSR ELISA kit.

Number of study subjects: 19 LADA cases

Acknowledgements: The staff from Research Centre for Prevention and Health, The capital region, Glostrup, Denmark is acknowledged their effort in making the Inter99 study possible.

Funding: The Inter99 study is funded by The Danish Medical Research Council, The Danish Centre for Evaluation and Health Technology Assessment, Novo Nordisk, Copenhagen County, The Danish Heart Foundation, The Danish Pharmaceutical Association, Augustinus foundation, Ib Henriksen foundation and Becket foundation.

Cohort reference: Jørgensen et al. A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. *Eur J Cardiovasc Prev Rehabil*. 2003.

Conflicts of interest: NA

Cohort name: SDC
Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases
Cohort description/inclusion/exclusion criteria: The SDC cohort comprises patients >18 years with type 2 diabetes (n=1676) recruited from the outpatient clinic at Steno Diabetes Center, Gentofte, Denmark. Individuals in pregnancy, having another cause of diabetes or being of another ethnicity than Danish were excluded. GADA were measured in 1595 individuals. Of the 141 LADA patients, GADA were measured with RSR ELISA in 52 patients, and with AESKULISA in 89 patients.
Number of study subjects: 141 LADA cases
Acknowledgements: NA
Funding: NA
Cohort reference: NA
Conflicts of interest: NA

Copenhagen general acknowledgements: Novo Nordisk Foundation Center for Basic Metabolic Research is an independent research center at the University of Copenhagen and is partly funded by an unrestricted donation from the Novo Nordisk Foundation. This work was supported by a research grant from the Danish Diabetes Academy supported by the Novo Nordisk Foundation, and grants from The Danish Council for Independent Research - Medical Sciences.

Cohort name: Diabetes Registry Vasa (DIREVA)
Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases/ T1D cases/ T2D cases
Inclusion/exclusion criteria: LADA
Age at onset ≥ 35 years
C-peptide (KLU) > 0.2 nmol/L
GAD65a (EIA) ≥ 10 U/ml

T1D
Age at onset < 35 years

C-peptide (KLU) < 0.2 nmol/L

T2D
Age at onset ≥ 35 years
C-peptide (KLU) ≥ 0.2 nmol/L
GAD65a (EIA) <10 U/ml

Recruitment Location: Vasa, Finland
Number of study subjects: 3290
Acknowledgements: Same as for ANDIS
Funding: DIREVA was supported by the Vasa Hospital district.
+funding overlapping with ANDIS
Cohort reference: NA
Conflicts of interest: None

Cohort name: GoDARTS
Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases):

- population controls (replication study only) 969
- LADA cases 206
- T2D cases 4413

Inclusion/exclusion criteria:

- Age diagnosis <35
- No insulin within 1 year diagnosis
- GADA positive

Recruitment Location: Univeersity of Dundee
Number of study subjects: (see above)
Acknowledgements: The Wellcome Trust United Kingdom Type 2 Diabetes Case Control Collection (GoDARTS) cohort collection was funded by The Wellcome Trust and informatics support is provided by the Chief Scientist Office, Scotland. E.R.P. holds a Wellcome Trust New Investigator Award (102820/Z/13/Z).
Funding: NA
Cohort reference: Diabetes Care. 2014;37(3):718-24. (PMID: 24186880)
Conflicts of interest: NA

Cohort name: HUNT
Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA and T2D cases and non-diabetic population controls

Inclusion/exclusion criteria: LADA: Self-reported yes to having diabetes, positive for GAD antibodies, initial age at diagnosis >30 years old and no insulin treatment within one year of diagnosis.
T2D: Self-reported yes to having diabetes, GAD antibodies negative, initial age at diagnosis >30 years old and no insulin treatment within one year of diagnosis. Age and gender matched to the LADA cases.
Non-diabetic controls: Self-reported no to ever having diabetes and had non-fasting serum glucose <7.0 mmol/l. Age and gender matched to the LADA cases.
Recruitment Location: HUNT Research Centre, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, the Norwegian University of Science and Technology, Norway
Number of study subjects: 139 LADA, 695 T2D and 695 non-diabetic controls
Acknowledgements: The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology), Nord-Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health.
Funding: The K.G. Jebsen Center for Genetic Epidemiology is financed by Stiftelsen Kristian Gerhard Jebsen; Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology; the Liaison Committee between the Central Norway Regional Health Authority and NTNU; and the Liaison Committee between St. Olavs Hospital and the Faculty of Medicine and Health Sciences at NTNU.

The genotyping was financed by the National Institute of Health (NIH), University of Michigan, The Norwegian Research Council, the Liaison Committee between the Central Norway Regional Health Authority and NTNU, and the Liaison Committee between St. Olavs Hospital and the Faculty of Medicine and Health Sciences at NTNU.
Cohort references:
Holmen J, Midthjell K, Krüger Ø, Langhammer A, Holmen TL, Bratberg GH, Vatten L, Lund-Larsen PG. *The Nord-Trøndelag Health Study 1995–97 (HUNT2): objectives. contents. methods and participation.* Norsk Epidemiologi 2003. **13**(1): p. 19-32.

Krokstad S, et al. *Cohort Profile: the HUNT Study. Norway.* Int J Epidemiol. 2013. **42**(4): p. 968-77.

Nielsen JB et al (2018) Genome-wide Study of Atrial Fibrillation Identifies Seven Risk Loci and Highlights Biological Pathways and Regulatory Elements Involved in Cardiac Development. Am J Hum Genet. 102(1):103-115 [29290336].

Conflicts of interest: There are no disclosures to report.

Cohort name: Malmö Diet and Cancer study

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): Non-diabetic controls

Inclusion/exclusion criteria:

Diabetes

Number of study subjects: 3126

Acknowledgements: NA

Funding: NA

Cohort reference: Berglund G, Nilsson P, Eriksson KF, Nilsson JA, Hedblad B, Kristenson H, et al. Long-term outcome of the Malmo preventive project: mortality and cardiovascular morbidity. J Intern Med. 2000;247(1):19-29. Epub 2000/02/15.

Conflicts of interest: None

Cohort name: Scania Diabetes Registry (SDR)

Cohort type (population controls/ LADA cases/ T1D cases/ T2D cases): LADA cases/ T1D cases/ T2D cases

Inclusion/exclusion criteria:

SDR

GAD (Wallenberg lab (AU ref < 5.0)

GAD (Wallenberg lab (IU/ml ref <32)

C-peptide (Klin kem (RIA) ref 0.25-0.75)

C-peptide (Klin kem ref 0.3-1.3)

C-peptide (Lund (ref 0.25-0.75)

LADA

Age at onset ≥ 35

GAD ≥ 10 AU

GAD ≥ 50 IU/ml

T1D

Age at onset < 35

GAD ≥ 20 AU

GAD ≥ 100 IU/ml
C-peptide ≥ 0.25 (Klin kem (RIA))
C-peptide ≥ 0.3 (Klin kem)
C-peptide ≥ 0.25 (Lund)

T2D
BMI > 25
GAD < 5 AU
GAD ≤ 34 IU/ml
C-peptide (Klin kem (RIA)) ≥ 0.75
C-peptide (Klin kem) ≥ 1.3
C-peptide (Lund) ≥ 0.75
For patients that did not fulfill the criteria for any of the above, the diagnosis given by their physician was used

Non-Scandinavian individuals excluded
Number of study subjects: 3567
Acknowledgements: NA
Funding: NA
Cohort reference: Lindholm E, Agardh E, Tuomi T, Groop L, Agardh CD. Classifying diabetes according to the new WHO clinical stages. Eur J Epidemiol. 2001;17(11):983-9. Epub 2002/10/17.
Conflicts of interest: None

Chapter 6. Genetic Discrimination between LADA and childhood-onset type 1 diabetes within the MHC

Supplemental Tables

Cohort	Details	Genotyping chip	Quality Control (QC)
ActionLada (LADA Cases , N=1051)	Populations: British, German Aged 30-70 Patients were designated with diabetes according to standard criteria, and LADA was defined as follows: patients 1) aged 30–70 years, 2) with diabetes associated autoantibodies, and 3) who did not require insulin treatment for at least 6 months post diagnosis. Only samples of European ancestry were included. Cohort overlaps with Cousminer DL et al. Diabetes Care (2018)	Illumina Infinium II Omni Express	Genotyping QC: Individuals with ambiguous sex, genotype missingness > 5%, and relatedness (pi_hat > 0.2) were excluded. PCA was performed to exclude individuals of non-European ancestry. Imputation QC: MAF <0.01
ActionLada 'Plus' (LADA Cases N=441)	Populations: American and British LADA was defined as follows: patients 1) aged 30–70 years, 2) with diabetes associated autoantibodies, and 3) who did not require insulin treatment for at least 6 months post diagnosis. A description of participants and study design can be found in Hawa MI et al. Diabetes Care (2013). More details such as exclusion criteria, etc can be found in Cousminer DL et al. Diabetes Care (2018)	Illumina Infinium II Omni Express	Genotyping QC: Individuals with ambiguous sex, genotype missingness > 5%, and relatedness (pi_hat > 0.2) were excluded. PCA was performed to exclude individuals of non-European ancestry. Imputation QC: MAF <0.01

Bone Mineral Density in Childhood Study (BMDCS) (Population controls,N=1296)	Populations: American of diverse ethnic backgrounds Samples of European descent only were included. More details such as recruitment, exclusion criteria, etc can be found in Cousminer DL et al. Diabetes Care (2018)	Illumina Infinium II Omni Express	Genotyping QC: Individuals with ambiguous sex, genotype missingness > 5%, and relatedness (pi_hat > 0.2) were excluded. PCA was performed to exclude individuals of non-European ancestry. Imputation QC: MAF <0.01
Non-hodgkin lyphomia Controls (dbGaP) (Population controls, N=1683)	Populations: USA, Multiple European countries, Australia More details such as recruitment, inclusion and exlusion, etc can be found in Berndt et al	Illumina Omni Express	Genotyping QC: Individuals with genotype missingness > 5% were excluded. PCA was performed to exclude individuals of non-European ancestry. Imputation QC: MAF <0.01
Wellcome Trust Case Control Consortium (WTCCC) **1958 British Birth Control Cohort** (N=3000)	Populations: England, Scotland and Wales More details such as recruitment, inclusion and exclusion criteria, etc can be found in WTCCC Nature (2007)	Illumina 1.2 M	Genotyping QC: Individuals with ambiguous sex, genotype missingness > 5%, and duplicates and relatedness (pi_hat > 0.2) were excluded. PCA was performed to exclude individuals of non-European ancestry. Single nucleotide polymorphisms (SNPs) with a call rate <95%,

			Hardy-Weinberg equilibrium P<10-5 and with A/T and G/C alleles were removed. Imputation QC: MAF <0.01
WTCCC T1D (N=2000)		Affymetrix 500 M	Genotyping QC: Individuals with ambiguous sex, genotype missingness > 5%, and rduplicates were excluded. PCA was performed to exclude individuals of non-European ancestry. Single nucleotide polymorphisms (SNPs) with a call rate <95%, Hardy-Weinberg equilibrium P<10-5 and with A/T and G/C alleles were removed. Imputation QC: MAF <0.01
WTCCC T1D (N=1990)		Affymetrix 500 M	Genotyping QC: Individuals with ambiguous sex, duplicates, and genotype missingness > 5%. PCA was performed to exclude individuals of non-European ancestry. Single

			nucleotide polymorphisms (SNPs) with a call rate <95%, Hardy-Weinberg equilibrium P<10-5 and with A/T and G/C alleles were removed. Imputation QC: MAF <0.01
Swedish Replication LADA(N=823) All New Diabetics In Scania (ANDIS), and Scania Diabetes Registry (SDR)	Population : Scania, Sweden **ANDIS**: GAD (ELISA: Negative:< 5 U/ml, Positive:=> 10 U/ml, RIA: Negative:0-34 U/ml, Positive:> 50 U/ml); LADA: Age at onset ≥ 35 years, GAD (ELISA) > 10 kE/L, GAD (RIA) >50 U/ml, Cohort reference: http://andis.ludc.med.lu.se/; **SDR**: GAD (Wallenberg lab (AU ref < 5.0), GAD (Wallenberg lab (IU/ml ref <32); LADA: Age at onset ≥ 35, GAD ≥ 10 AU, GAD ≥ 50 IU/ml, For patients that did not fulfill the criteria for any of the above, the diagnosis given by their physician was used, Cohort reference: Lindholm E, Agardh E, Tuomi T, Groop L, Agardh CD. Classifying diabetes according to the new WHO clinical stages. Eur J Epidemiol. 2001;17(11):983-9. Epub 2002/10/17	Illumina Infinium Omni Express Exome	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); duplicates or related individuals (pi-hat >= 0.2); average pi-hat outliers; population outliers. Exclude SNPs: monomorphic; (MAF >= 0.05 and SNP missing rate > 0.05); (MAF < 0.05 and SNP missing rate > 0.01); (MAF >= 0.05 and HWE <= 0.00000057); (MAF < 0.05 and HWE <= 0.0001) Imputation QC: MAF <0.01

Swedish Replication T1D (N=656) Scania Diabetes Registry (SDR)	Population : Scania, Sweden GAD (Wallenberg lab (AU ref < 5.0), GAD (Wallenberg lab (IU/ml ref <32), C-peptide (Klin kem (RIA) ref 0.25-0.75), C-peptide (Klin kem ref 0.3-1.3), C-peptide (Lund (ref 0.25-0.75); T1D: Age at onset < 35, GAD ≥ 20 AU, GAD ≥ 100 IU/ml, C-peptide ≥ 0.25 (Klin kem (RIA)), C-peptide ≥ 0.3 (Klin kem), C-peptide ≥ 0.25 (Lund), For patients that did not fulfill the criteria for any of the above, the diagnosis given by their physician was used, Cohort reference: Lindholm E, Agardh E, Tuomi T, Groop L, Agardh CD. Classifying diabetes according to the new WHO clinical stages. Eur J Epidemiol. 2001;17(11):983-9. Epub 2002/10/17	Illumina Infinium Omni Express Exome	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); duplicates or related individuals (pi-hat >= 0.2); average pi-hat outliers; population outliers. Exclude SNPs: monomorphic; (MAF >= 0.05 and SNP missing rate > 0.05); (MAF < 0.05 and SNP missing rate > 0.01); (MAF >= 0.05 and HWE <= 0.00000057); (MAF < 0.05 and HWE <= 0.0001)

			Imputation QC: MAF <0.01
Swedish Replication Population-based controls (N=3218) Malmö Diet and Cancer Study–Cardiovascular Cohort (MDC-CC)	Population: Malmö, Sweden More details such as recruitment, inclusion and exclusion criteria, etc can be found in Rosvall M, Persson M, Ostling G, et al. Risk factors for the progression of carotid intima-media thickness over a 16-year follow-up period: the Malmo Diet and Cancer Study. Atherosclerosis. Apr 2015;239(2):615-621.	Illumina Infinium Omni Express Exome	Exclude individuals: Call rate < 95%; ambiguous gender; genome-wide heterozygosity (3 SD from mean); duplicates or related individuals (pi-hat >= 0.2); population outliers. Exclude SNPs: monomorphic; SNP missing rate > 0.05; HWE <= 10e-6 Imputation QC: MAF <0.01

Supp. Table 6-1. Cohort information

Odds Ratio						
MAF	**3**	**2**	**1.8**	**1.6**	**1.4**	**1.35**
0.01	**89.70%**	14.50%	5.30%	1.30%	0.20%	0.10%
0.03	**100%**	**90.90%**	64.90%	25.80%	3.80%	1.90%
0.05	**100%**	**99.80%**	**95.60%**	65.10%	14.70%	7.60%
0.1	**100%**	**100%**	**100%**	**98.50%**	58.10%	37.60%
0.2	**100%**	**100%**	**100%**	**100%**	**95.40%**	**84.30%**
0.3	**100%**	**100%**	**100%**	**100%**	**99.4%**	**95.9%**

Supp. Table 6-2. Power Calculations for detecting MHC Class I variants in LADA. Assumptions include a multiplicative model for non-MHC Class I signals, a significance level of 8.93×10^{-6},a disease prevalence of 0.0036, and 1,428 LADA cases and 2,979 controls.

BIBLIOGRAPHY

Adeleye, O. O., Ogbera, A. O., Fasanmade, O., Ogunleye, O. O., Dada, A. O., Ale, A. O., & Abatan, F. M. (2012). Latent autoimmune diabetes mellitus in adults (LADA) and its characteristics in a subset of Nigerians initially managed for type 2 diabetes. *International Archives of Medicine*. https://doi.org/10.1186/1755-7682-5-23

Agyei-Frempong, M. T., Titty, F. V. K., Owiredu, W. K. B. A., & Eghan, B. A. (2008). The prevalence of autoimmune diabetes among diabetes mellitus patients in Kumasi, Ghana. *Pakistan Journal of Biological Sciences*. https://doi.org/10.3923/pjbs.2008.2320.2325

Ahlqvist, E., Storm, P., Karajamaki, A., Martinell, M., Dorkhan, M., Carlsson, A., … Groop, L. (2018). Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol*. https://doi.org/10.1016/S2213-8587(18)30051-2

Alberti, K. G., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med*, *15*(7), 539–553. https://doi.org/10.1002/(SICI)1096-9136(199807)15:7<539::AID-DIA668>3.0.CO;2-S

Ali, R., DiLorenzo, T. P., Racine, J. J., Chapman, H. D., Serreze, D. V., & Schloss, J. (2018). HLA-B*39:06 Efficiently Mediates Type 1 Diabetes in a Mouse Model Incorporating Reduced Thymic Insulin Expression. *The Journal of Immunology*. https://doi.org/10.4049/jimmunol.1701652

American Diabetes, A. (2009). Standards of medical care in diabetes--2009. *Diabetes Care*, *32 Suppl 1*, S13-61. https://doi.org/10.2337/dc09-S013

American Diabetes Association/ADA. (2018). American Diabetes Association. 2. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes - 2018. *Diabetes Care*. https://doi.org/10.2337/dc18-S002

American Diabetes Association. (2014). *National Diabetes Statistics Report , 2014 Estimates of Diabetes and Its Burden in the United States*. *National Diabetes Statistics Report*. https://doi.org/Centers for Disease Control and Prevention. National Diabetes Statistics Report, 2017. Atlanta, GA: Centers for Disease Control and Prevention, U.S. Dept of Health and Human Services; 2017.

Andersen, C. D., Bennet, L., Nystrom, L., Lindblad, U., Lindholm, E., Groop, L., & Rolandsson, O. (2013). Worse glycaemic control in LADA patients than in those with type 2 diabetes, despite a longer time on insulin therapy. *Diabetologia*, *56*(2), 252–258.

https://doi.org/10.1007/s00125-012-2759-y
Andersen, M. K., Lundgren, V., Turunen, J. A., Forsblom, C., Isomaa, B., Groop, P. H., … Tuomi, T. (2010). Latent autoimmune diabetes in adults differs genetically from classical type 1 diabetes diagnosed after the age of 35 years. *Diabetes Care*. https://doi.org/10.2337/dc09-2188
Andersen, M. K., Sterner, M., Forsen, T., Karajamaki, A., Rolandsson, O., Forsblom, C., … Tuomi, T. (2014). Type 2 diabetes susceptibility gene variants predispose to adult-onset autoimmune diabetes. *Diabetologia*, *57*(9), 1859–1868. https://doi.org/10.1007/s00125-014-3287-8
Aylward, A., Chiou, J., Okino, M.-L., Kadakia, N., & Gaulton, K. J. (2018). Shared genetic risk contributes to type 1 and type 2 diabetes etiology. *Human Molecular Genetics*. https://doi.org/10.1093/hmg/ddy314
Bach, J. F. (2002). The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med*, *347*(12), 911–920. https://doi.org/10.1056/NEJMra020100
Bakhtadze, E., Cervin, C., Lindholm, E., Borg, H., Nilsson, P., Arnqvist, H. J., … Groop, L. C. (2008). Common variants in the TCF7L2 gene help to differentiate autoimmune from non-autoimmune diabetes in young (15-34 years) but not in middle-aged (40-59 years) diabetic patients. *Diabetologia*. https://doi.org/10.1007/s00125-008-1161-2
Barinas-Mitchell, E., Pietropaolo, S., Zhang, Y. J., Henderson, T., Trucco, M., Kuller, L. H., & Pietropaolo, M. (2004). Islet cell autoimmunity in a triethnic adult population of the Third National Health and Nutrition Examination Survey. *Diabetes*, *53*(5), 1293–1302. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15111499
Barrett, J. C., Clayton, D. G., Concannon, P., Akolkar, B., Cooper, J. D., Erlich, H. A., … Rich, S. S. (2009). Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nature Genetics*. https://doi.org/10.1038/ng.381
Baschal, E. E., Baker, P. R., Eyring, K. R., Siebert, J. C., Jasinski, J. M., & Eisenbarth, G. S. (2011). The HLA-B*3906 allele imparts a high risk of diabetes only on specific HLA-DR/DQ haplotypes. *Diabetologia*. https://doi.org/10.1007/s00125-011-2161-1
Berndt, S. I., Skibola, C. F., Joseph, V., Camp, N. J., Nieters, A., Wang, Z., … Slager, and S. L. (2013). Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nature Genetics*. https://doi.org/10.1038/ng.2652
Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg932
Bosi, E. P., Garancini, M. P., Poggiali, F., Bonifacio, E., & Gallus, G. (1999). Low prevalence of islet autoimmunity in adult diabetes and low predictive value of islet autoantibodies in the general adult population of northern Italy. *Diabetologia*, *42*(7), 840–844. Retrieved

from https://www.ncbi.nlm.nih.gov/pubmed/10440126
Bradfield, J. P., Qu, H. Q., Wang, K., Zhang, H., Sleiman, P. M., Kim, C. E., … Hakonarson, H. (2011). A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet*, *7*(9), e1002293. https://doi.org/10.1371/journal.pgen.1002293
Bulik-Sullivan, B., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., … O'Donovan, M. C. (2015). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*. https://doi.org/10.1038/ng.3211
Buzzetti, R., Di Pietro, S., Giaccari, A., Petrone, A., Locatelli, M., Suraci, C., … Non Insulin Requiring Autoimmune Diabetes Study, G. (2007). High titer of autoantibodies to GAD identifies a specific phenotype of adult-onset autoimmune diabetes. *Diabetes Care*, *30*(4), 932–938. https://doi.org/10.2337/dc06-1696
Buzzetti, R., Zampetti, S., & Maddaloni, E. (2017). Adult-onset autoimmune diabetes: current knowledge and implications for management. *Nat Rev Endocrinol*, *13*(11), 674–686. https://doi.org/10.1038/nrendo.2017.99
Cervin, C., Lyssenko, V., Bakhtadze, E., Lindholm, E., Nilsson, P., Tuomi, T., … Groop, L. (2008). Genetic similarities between latent autoimmune diabetes in adults, type 1 diabetes, and type 2 diabetes. *Diabetes*, *57*(5), 1433–1437. https://doi.org/10.2337/db07-0299
Cheng, C. Y., Reich, D., Haiman, C. A., Tandon, A., Patterson, N., Elizabeth, S., … Kao, W. H. L. (2012). African ancestry and its correlation to type 2 diabetes in african americans: A genetic admixture analysis in three U.S. population cohorts. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0032840
Cirulli, V., Zalatan, J., McMaster, M., Prinsen, R., Salomon, D. R., Ricordi, C., … Crisa, L. (2006). The class I HLA repertoire of pancreatic islets comprises the nonclassical class Ib antigen HLA-G. *Diabetes*. https://doi.org/10.2337/db05-0731
Cooke, J. N., Ng, M. C., Palmer, N. D., An, S. S., Hester, J. M., Freedman, B. I., … Bowden, D. W. (2012). Genetic risk assessment of type 2 diabetes-associated polymorphisms in African Americans. *Diabetes Care*, *35*(2), 287–292. https://doi.org/10.2337/dc11-0957
Cousminer, D. L., Ahlqvist, E., Mishra, R., Andersen, M. K., Chesi, A., Hawa, M. I., … Grant, S. F. A. (2018). First Genome-Wide Association Study of Latent Autoimmune Diabetes in Adults Reveals Novel Insights Linking Immune and Metabolic Diabetes. *Diabetes Care*, *41*(November), dc181032. https://doi.org/10.2337/dc18-1032
Cowie, C. C., Rust, K. F., Byrd-Holt, D. D., Eberhardt, M. S., Flegal, K. M., Engelgau, M. M., … Gregg, E. W. (2006). Prevalence of diabetes and impaired fasting glucose in adults in the U.S. population: National Health And Nutrition Examination Survey 1999-2002. *Diabetes Care*, *29*(6), 1263–1268. https://doi.org/10.2337/dc06-0062

Dabelea, D., Mayer-Davis, E. J., Saydah, S., Imperatore, G., Linder, B., Divers, J., … Hamman, R. F. (2014). Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA - Journal of the American Medical Association*. https://doi.org/10.1001/jama.2014.3201

De Luca, M., Angrisani, L., Himpens, J., Busetto, L., Scopinaro, N., Weiner, R., … Shikora, S. (2016). Indications for Surgery for Obesity and Weight-Related Diseases: Position Statements from the International Federation for the Surgery of Obesity and Metabolic Disorders (IFSO). *Obes Surg*, *26*(8), 1659–1696. https://doi.org/10.1007/s11695-016-2271-4

Dennis, J. M., Shields, B. M., Hill, A. V, Knight, B. A., McDonald, T. J., Rodgers, L. R., … consortium, M. (2018). Precision Medicine in Type 2 Diabetes: Clinical Markers of Insulin Resistance Are Associated With Altered Short- and Long-Term Glycemic Response to DPP-4 Inhibitor Therapy. *Diabetes Care*. https://doi.org/10.2337/dc17-1827

Desai, M., Zeggini, E., Horton, V. A., Owen, K. R., Hattersley, A. T., Levy, J. C., … Clark, A. (2006). The variable number of tandem repeats upstream of the insulin gene is a susceptibility locus for latent autoimmune diabetes in adults. *Diabetes*, *55*(6), 1890–1894. https://doi.org/10.2337/db06-0089

DIAGRAM, C., AGENTD, C., SATD, C., MATD, C., TDGEbN-gsim-ES, C., Mahajan, A., … Morris, A. P. (2014). Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet*, *46*(3), 234–244. https://doi.org/10.1038/ng.2897

Donnelly, L. A., Zhou, K., Doney, A. S. F., Jennison, C., Franks, P. W., & Pearson, E. R. (2018). Rates of glycaemic deterioration in a real-world population with type 2 diabetes. *Diabetologia*, *61*(3), 607–615. https://doi.org/10.1007/s00125-017-4519-5

Duran, J., Obach, M., Navarro-Sabate, A., Manzano, A., Gómez, M., Rosa, J. L., … Bartrons, R. (2009). Pfkfb3 is transcriptionally upregulated in diabetic mouse liver through proliferative signals. *FEBS Journal*. https://doi.org/10.1111/j.1742-4658.2009.07161.x

Eike, M. C., Becker, T., Humphreys, K., Olsson, M., & Lie, B. A. (2009). Conditional analyses on the T1DGC MHC dataset: Novel associations with type 1 diabetes around HLA-G and confirmation of HLA-B. *Genes and Immunity*. https://doi.org/10.1038/gene.2008.74

Eisenbarth, G. S., Orban, T., Fain, P. R., Jeffrey, J., & Redondo, M. J. (2008). Concordance for Islet Autoimmunity among Monozygotic Twins. *New England Journal of Medicine*. https://doi.org/10.1056/nejmc0805398

Falconer, D. S. (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics*. https://doi.org/10.1111/j.1469-1809.1965.tb00500.x

Fava, D., Gardner, S., Pyke, D., & Leslie, R. D. (1998). Evidence that the age at diagnosis of IDDM is genetically determined. *Diabetes Care*, *21*(6), 925–929. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9614609

Florez, J. C., Hirschhorn, J., & Altshuler, D. (2003). The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits. *Annu Rev Genomics Hum Genet*, *4*, 257–291. https://doi.org/10.1146/annurev.genom.4.070802.110436

Frohnert, B. I., Ide, L., Dong, F., Baron, A. E., Steck, A. K., Norris, J. M., & Rewers, M. J. (2017). Late-onset islet autoimmunity in childhood: the Diabetes Autoimmunity Study in the Young (DAISY). *Diabetologia*, *60*(6), 998–1006. https://doi.org/10.1007/s00125-017-4256-9

Gambelunghe, G., Ghaderi, M., Cosentino, A., Falorni, A., Brunetti, P., Falorni, A., & Sanjeevi, C. B. (2000). Association of MHC Class I chain-related A (MIC-A) gene polymorphism with Type I diabetes. *Diabetologia*. https://doi.org/10.1007/s001250051336

Garg, S. K., Henry, R. R., Banks, P., Buse, J. B., Davies, M. J., Fulcher, G. R., … Strumph, P. (2017). Effects of Sotagliflozin Added to Insulin in Patients with Type 1 Diabetes. *N Engl J Med*, *377*(24), 2337–2348. https://doi.org/10.1056/NEJMoa1708337

Genovese, S., Bazzigaluppi, E., Goncalves, D., Ciucci, A., Cavallo, M. G., Purrello, F., … Pozzilli, P. (2006). Clinical phenotype and beta-cell autoimmunity in Italian patients with adult-onset diabetes. *Eur J Endocrinol*, *154*(3), 441–447. https://doi.org/10.1530/eje.1.02115

Hawa, M. I., Kolb, H., Schloot, N., Beyan, H., Paschou, S. A., Buzzetti, R., … Action, L. consortium. (2013). Adult-onset autoimmune diabetes in Europe is prevalent with a broad clinical phenotype: Action LADA 7. *Diabetes Care*, *36*(4), 908–913. https://doi.org/10.2337/dc12-0931

Hjort, R., Alfredsson, L., Andersson, T., Carlsson, P. O., Grill, V., Groop, L., … Carlsson, S. (2017). Family history of type 1 and type 2 diabetes and risk of latent autoimmune diabetes in adults (LADA). *Diabetes and Metabolism*. https://doi.org/10.1016/j.diabet.2017.05.010

Horton, V., Stratton, I., Bottazzo, G. F., Shattock, M., Mackay, I., Zimmet, P., … Turner, R. (1999). Genetic heterogeneity of autoimmune diabetes: age of presentation in adults is influenced by HLA DRB1 and DQB1 genotypes (UKPDS 43). UK Prospective Diabetes Study (UKPDS) Group. *Diabetologia*, *42*(5), 608–616. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10333055

Hossein, G. M., Maryam, S., Saeed, E., Hamid, N., & Mahdieh, M. (2015). The prevalence of latent autoimmune diabetes in adults and its correlates in patients with type 2 diabetes in Kerman, Iran [2011]. *Diabetes Metab Syndr*, *9*(2), 104–107. https://doi.org/10.1016/j.dsx.2014.04.014

Hosszúfalusi, N., Vatay, A., Rajczy, K., Prohászka, Z., Pozsonyi, E.,

Horváth, L., … Panczél, P. (2003). Similar genetic features and different islet cell autoantibody pattern of latent autoimmune diabetes in adults (LADA) compared with adult-onset type 1 diabetes with rapid progression. *Diabetes Care*. https://doi.org/10.2337/diacare.26.2.452

Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., & Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet*, *44*(8), 955–959. https://doi.org/10.1038/ng.2354

Howson, J. M. M., Walker, N. M., Clayton, D., & Todd, J. A. (2009). Confirmation of HLA class II independent type 1 diabetes associations in the major histocompatibility complex including HLA-B and HLA-A. *Diabetes, Obesity and Metabolism*. https://doi.org/10.1111/j.1463-1326.2008.01001.x

Howson, J. M., Rosinger, S., Smyth, D. J., Boehm, B. O., Group, A.-E. S., & Todd, J. A. (2011). Genetic analysis of adult-onset autoimmune diabetes. *Diabetes*, *60*(10), 2645–2653. https://doi.org/10.2337/db11-0364

Hu, X., Deutsch, A. J., Lenz, T. L., Onengut-Gumuscu, S., Han, B., Chen, W. M., … Raychaudhuri, S. (2015). Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nature Genetics*. https://doi.org/10.1038/ng.3353

Huang, G., Xiang, Y., Pan, L., Li, X., Luo, S., & Zhou, Z. (2013). Zinc transporter 8 autoantibody (ZnT8A) could help differentiate latent autoimmune diabetes in adults (LADA) from phenotypic type 2 diabetes mellitus. *Diabetes Metab Res Rev*, *29*(5), 363–368. https://doi.org/10.1002/dmrr.2396

Huo, Y., Guo, X., Li, H., Wang, H., Zhang, W., Wang, Y., … Wu, C. (2010). Disruption of inducible 6-phosphofructo-2-kinase ameliorates diet-induced adiposity but exacerbates systemic insulin resistance and adipose tissue inflammatory response. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M109.058446

Huo, Y., Guo, X., Li, H., Xu, H., Halim, V., Zhang, W., … Wu, C. (2012). Targeted overexpression of inducible 6-phosphofructo-2-kinase in adipose tissue increases fat deposition but protects against diet-induced insulin resistance and inflammatory responses. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M112.370379

Hwangbo, Y., Kim, J. T., Kim, E. K., Khang, A. R., Oh, T. J., Jang, H. C., … Cho, Y. M. (2012). Prevalence and clinical characteristics of recently diagnosed type 2 diabetes patients with positive anti-glutamic Acid decarboxylase antibody. *Diabetes Metab J*, *36*(2), 136–143. https://doi.org/10.4093/dmj.2012.36.2.136

Hyttinen, V., Kaprio, J., Kinnunen, L., Koskenvuo, M., & Tuomilehto, J. (2003). Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study. *Diabetes*, *52*(4), 1052–1055. Retrieved from

https://www.ncbi.nlm.nih.gov/pubmed/12663480
Ipadeola, A., Adeleye, J. O., & Akinlade, K. S. (2015). Latent autoimmune diabetes amongst adults with type 2 diabetes in a Nigerian tertiary hospital. *Primary Care Diabetes*. https://doi.org/10.1016/j.pcd.2014.09.003
Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W. M., Concannon, P. J., Rich, S. S., … de Bakker, P. I. (2013). Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One*, *8*(6), e64683. https://doi.org/10.1371/journal.pone.0064683
Kalkwarf, H. J., Zemel, B. S., Gilsanz, V., Lappe, J. M., Horlick, M., Oberfield, S., … Shepherd, J. A. (2007). The bone mineral density in childhood study: bone mineral content and density according to age, sex, and race. *J Clin Endocrinol Metab*, *92*(6), 2087–2099. https://doi.org/10.1210/jc.2006-2553
Kanungo, A., & Sanjeevi, C. B. (2003). IA-2 autoantibodies are predominant in latent autoimmune diabetes in adults patients from Eastern India. In *Annals of the New York Academy of Sciences*. https://doi.org/10.1196/annals.1288.065
Karnes, J. H., Steiner, H. E., Phillips, E. J., Roden, D. M., Bastarache, L., Mosley, J. D., … Gaudieri, S. (2017). Comparison of HLA allelic imputation programs. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0172444
Katulanda, P., Shine, B., Katulanda, G. W., Silva, A., Asfir, E. L., Sheriff, R., … Matthews, D. R. (2008). Diabetes mellitus among young adults in Sri Lanka-role of GAD antibodies in classification and treatment: The Sri Lanka Young Diabetes study. *Diabetologia*. https://doi.org/10.1007/s00125-008-1052-6
Kobayashi, T., Maruyama, T., Shimada, A., Kasuga, A., Kanatsuka, A., Takei, I., … Yokoyama, J. (2002). Insulin intervention to preserve beta cells in slowly progressive insulin-dependent (type 1) diabetes mellitus. *Ann N Y Acad Sci*, *958*, 117–130. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12021091
Kundu, S., Aulchenko, Y. S., van Duijn, C. M., & Janssens, A. C. (2011). PredictABEL: an R package for the assessment of risk prediction models. *Eur J Epidemiol*, *26*(4), 261–264. https://doi.org/10.1007/s10654-011-9567-4
Laugesen, E., Østergaard, J. A., & Leslie, R. D. G. (2015). Latent autoimmune diabetes of the adult: Current knowledge and uncertainty. *Diabetic Medicine*. https://doi.org/10.1111/dme.12700
Leslie, R. D., Palmer, J., Schloot, N. C., & Lernmark, A. (2016). Diabetes at the crossroads: relevance of disease classification to pathophysiology and treatment. *Diabetologia*, *59*(1), 13–20. https://doi.org/10.1007/s00125-015-3789-z
Liu, L., Li, X., Xiang, Y., Huang, G., Lin, J., Yang, L., … Group, L. C. S. (2015). Latent autoimmune diabetes in adults with low-titer GAD

antibodies: similar disease progression with type 2 diabetes: a nationwide, multicenter prospective study (LADA China Study 3). *Diabetes Care*, *38*(1), 16–21. https://doi.org/10.2337/dc14-1770
Lukacs, K., Hosszufalusi, N., Dinya, E., Bakacs, M., Madacsy, L., & Panczel, P. (2012). The type 2 diabetes-associated variant in TCF7L2 is associated with latent autoimmune diabetes in adult Europeans and the gene effect is modified by obesity: a meta-analysis and an individual study. *Diabetologia*, *55*(3), 689–693. https://doi.org/10.1007/s00125-011-2378-z
Luo, S., Lin, J., Xie, Z., Xiang, Y., Zheng, P., Huang, G., … Zhou, Z. (2016). HLA genetic discrepancy between latent autoimmune diabetes in adults and type 1 diabetes: LADA China study No. 6. *Journal of Clinical Endocrinology and Metabolism*. https://doi.org/10.1210/jc.2015-3771
Maahs, D. M., West, N. A., Lawrence, J. M., & Mayer-Davis, E. J. (2010). Epidemiology of type 1 diabetes. *Endocrinol Metab Clin North Am*, *39*(3), 481–497. https://doi.org/10.1016/j.ecl.2010.05.011
Maddaloni, E., Lessan, N., Al Tikriti, A., Buzzetti, R., Pozzilli, P., & Barakat, M. T. (2015). Latent Autoimmune Diabetes in Adults in the United Arab Emirates: Clinical Features and Factors Related to Insulin-Requirement. *PLoS One*, *10*(8), e0131837. https://doi.org/10.1371/journal.pone.0131837
Magi, R., & Morris, A. P. (2010). GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics*, *11*, 288. https://doi.org/10.1186/1471-2105-11-288
Mahajan, A., Taliun, D., Thurner, M., Robertson, N. R., Torres, J. M., Rayner, N. W., … McCarthy, M. I. (2018). Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nature Genetics*. https://doi.org/10.1038/s41588-018-0241-6
Maioli, M., Pes, G. M., Delitala, G., Puddu, L., Falorni, A., Tolu, F., … Cucca, F. (2010). Number of autoantibodies and HLA genotype, more than high titers of glutamic acid decarboxylase autoantibodies, predict insulin dependence in latent autoimmune diabetes of adults. *Eur J Endocrinol*, *163*(4), 541–549. https://doi.org/10.1530/EJE-10-0427
Manan, H., Angham, A. M., & Sitelbanat, A. (2010). Genetic and diabetic auto-antibody markers in Saudi children with type 1 diabetes. *Human Immunology*. https://doi.org/10.1016/j.humimm.2010.09.008
Marchini, J., Howie, B., Myers, S., McVean, G., & Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*, *39*(7), 906–913. https://doi.org/10.1038/ng2088
Martinell, M., Pingel, R., Hallqvist, J., Dorkhan, M., Groop, L., Rosengren, A., … Stalhammar, J. (2017). Education, immigration and income as risk factors for hemoglobin A1c >70 mmol/mol when diagnosed with

type 2 diabetes or latent autoimmune diabetes in adult: a population-based cohort study. *BMJ Open Diabetes Res Care*, *5*(1), e000346. https://doi.org/10.1136/bmjdrc-2016-000346

Mayer, D, E. J., Lawrence, J., Dabelea, D., Divers, J., Isom, S., Dolan, L., … Wagenknecht, L. (2017). Incidence Trends of Type 1 and Type 2 Diabetes among Youths, 2002–2012. *New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa1610187

Mbanya, J. C. N., Motala, A. A., Sobngwi, E., Assah, F. K., & Enoru, S. T. (2010). Diabetes in sub-Saharan Africa. *Lancet (London, England)*, *375*(9733), 2254–2266. https://doi.org/10.1016/S0140-6736(10)60550-8

Mishra, R., Chesi, A., Cousminer, D. L., Hawa, M. I., Bradfield, J. P., Hodge, K. M., … Grant, S. F. A. (2017). Relative contribution of type 1 and type 2 diabetes loci to the genetic etiology of adult-onset, non-insulin-requiring autoimmune diabetes. *BMC Med*, *15*(1), 88. https://doi.org/10.1186/s12916-017-0846-0

Mishra, R., Hodge, K. M., Cousminer, D. L., Leslie, R. D., & Grant, S. F. A. (2018). A Global Perspective of Latent Autoimmune Diabetes in Adults. *Trends in Endocrinology and Metabolism*, *29*(9), 638–650. https://doi.org/10.1016/j.tem.2018.07.001

Morris, A. P., Voight, B. F., Teslovich, T. M., Ferreira, T., Segre, A. V, Steinthorsdottir, V., … Meta-analysis, C. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*, *44*(9), 981–990. https://doi.org/10.1038/ng.2383

Motala, A. A., Omar, M. A. K., & Pirie, F. J. (2003). Epidemiology of type 1 and type 2 diabetes in africa. *Journal of Cardiovascular Risk*. https://doi.org/10.1097/00043798-200304000-00002

Muazu, S. B., Okpe, I., & Anumah, F. (2016). The prevalence and characteristics of latent autoimmune diabetes in adults subset among type two diabetes mellitus patients in Northern Nigeria. *Ann Afr Med*, *15*(4), 163–170. https://doi.org/10.4103/1596-3519.194277

Nejentsev, S., Howson, J. M., Walker, N. M., Szeszko, J., Field, S. F., Stevens, H. E., … Wellcome Trust Case Control, C. (2007). Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature*, *450*(7171), 887–892. https://doi.org/10.1038/nature06406

Ng, M. C., Shriner, D., Chen, B. H., Li, J., Chen, W. M., Guo, X., … Consortium, M. E. of type 2 Di. in A. A. (2014). Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genet*, *10*(8), e1004517. https://doi.org/10.1371/journal.pgen.1004517

Noble, J. A., Valdes, A. M., Bugawan, T. L., Apple, R. J., Thomson, G., & Erlich, H. A. (2002). The HLA class I A locus affects susceptibility to type 1 diabetes. *Human Immunology*. https://doi.org/10.1016/S0198-

8859(02)00421-4
Noble, J. A., Valdes, A. M., Varney, M. D., Carlson, J. A., Moonsamy, P., Fear, A. L., … Erlich, H. A. (2010). HLA class I and genetic susceptibility to type 1 diabetes: results from the Type 1 Diabetes Genetics Consortium. *Diabetes*. https://doi.org/10.2337/db10-0699
O'Connell, J., Gurdasani, D., Delaneau, O., Pirastu, N., Ulivi, S., Cocca, M., … Marchini, J. (2014). A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genet*, *10*(4), e1004234. https://doi.org/10.1371/journal.pgen.1004234
Ogurtsova, K., da Rocha Fernandes, J. D., Huang, Y., Linnenkamp, U., Guariguata, L., Cho, N. H., … Makaroff, L. E. (2017). IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract*, *128*, 40–50. https://doi.org/10.1016/j.diabres.2017.03.024
Onengut-Gumuscu, S., Chen, W. M., Burren, O., Cooper, N. J., Quinlan, A. R., Mychaleckyj, J. C., … Rich, S. S. (2015). Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nature Genetics*. https://doi.org/10.1038/ng.3245
Ong, Y. H., Koh, W. C. A., Ng, M. L., Tam, Z. Y., Lim, S. C., Boehm, B. O., & Adult-Onset Autoimmune Diabetes Mellitus, C. (2017). Glutamic acid decarboxylase and islet antigen 2 antibody profiles in people with adult-onset diabetes mellitus: a comparison between mixed ethnic populations in Singapore and Germany. *Diabet Med*, *34*(8), 1145–1153. https://doi.org/10.1111/dme.13358
Oram, R. A., Patel, K., Hill, A., Shields, B., McDonald, T. J., Jones, A., … Weedon, M. N. (2016). A Type 1 Diabetes Genetic Risk Score Can Aid Discrimination Between Type 1 and Type 2 Diabetes in Young Adults. *Diabetes Care*, *39*(3), 337–344. https://doi.org/10.2337/dc15-1111
Otieno, C. F., Huho, A. N., Omonge, E. O., Amayo, A. A., & Njagi, E. (2008). Type 2 diabetes mellitus: clinical and aetiologic types, therapy and quality of glycaemic control of ambulatory patients. *East Afr Med J*, *85*(1), 24–29. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18543523
Pappas, D. J., Marin, W., Hollenbach, J. A., & Mack, S. J. (2016). Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline. *Hum Immunol*, *77*(3), 283–287. https://doi.org/10.1016/j.humimm.2015.12.006
Park, Y., Hong, S., Park, L., Woo, J., Baik, S., Nam, M., … Group, K. collaboratory. (2011). LADA prevalence estimation and insulin dependency during follow-up. *Diabetes Metab Res Rev*, *27*(8), 975–979. https://doi.org/10.1002/dmrr.1278
Parker, J., Talih, M., Malec, D., Beresovsky, V., Carroll, M., Gonzalez, J., … Ward, B. (2017). National Center for Health Statistics Data

Presentation Standards for Proportions. *Vital Health Statistics 2.*, (175), 1–22.
Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*. https://doi.org/10.1093/biomet/58.3.545
Pers, T. H., Karjalainen, J. M., Chan, Y., Westra, H. J., Wood, A. R., Yang, J., … Franke, L. (2015). Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*. https://doi.org/10.1038/ncomms6890
Poulsen, P., Kyvik, K. O., Vaag, A., & Beck-Nielsen, H. (1999). Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia*, *42*(2), 139–145. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10064092
Pozzilli, P., Guglielmi, C., Caprio, S., & Buzzetti, R. (2011). Obesity, autoimmunity, and double diabetes in youth. *Diabetes Care*. https://doi.org/10.2337/dc11-s213
Pugliese, A., Boulware, D., Yu, L., Babu, S., Steck, A. K., Becker, D., … Type 1 Diabetes TrialNet Study, G. (2016). HLA-DRB1*15:01-DQA1*01:02-DQB1*06:02 Haplotype Protects Autoantibody-Positive Relatives From Type 1 Diabetes Throughout the Stages of Disease Progression. *Diabetes*, *65*(4), 1109–1119. https://doi.org/10.2337/db15-1105
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., … Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, *81*(3), 559–575. https://doi.org/10.1086/519795
Qi, X., Sun, J., Wang, J., Wang, P. P., Xu, Z., Murphy, M., … Xu, W. (2011). Prevalence and correlates of latent autoimmune diabetes in adults in Tianjin, China: a population-based cross-sectional study. *Diabetes Care*, *34*(1), 66–70. https://doi.org/10.2337/dc10-0488
Radenkovic, M., Silver, C., Arvastsson, J., Lynch, K., Lernmark, Harris, R. A., … Cilio, C. M. (2016). Altered regulatory T cell phenotype in latent autoimmune diabetes of the adults (LADA). *Clinical and Experimental Immunology*. https://doi.org/10.1111/cei.12834
Radtke, M. A., Midthjell, K., Nilsen, T. I., & Grill, V. (2009). Heterogeneity of patients with latent autoimmune diabetes in adults: linkage to autoimmunity is apparent only in those with perceived need for insulin treatment: results from the Nord-Trondelag Health (HUNT) study. *Diabetes Care*, *32*(2), 245–250. https://doi.org/10.2337/dc08-1468
Redondo, M. J., Geyer, S., Steck, A. K., Sosenko, J., Anderson, M., Antinozzi, P., … and the Type 1 Diabetes TrialNet Study, G. (2017). TCF7L2 Genetic Variants Contribute to Phenotypic Heterogeneity of Type 1 Diabetes. *Diabetes Care*. https://doi.org/10.2337/dc17-0961
Redondo, M. J., Grant, S. F., Davis, A., Greenbaum, C., & Biobank, T. D.

E. (2017). Dissecting heterogeneity in paediatric Type 1 diabetes: association of TCF7L2 rs7903146 TT and low-risk human leukocyte antigen (HLA) genotypes. *Diabet Med*, *34*(2), 286–290. https://doi.org/10.1111/dme.13123

Redondo, M. J., Muniz, J., Rodriguez, L. M., Iyer, D., Vaziri-Sani, F., Haymond, M. W., … Balasubramanyam, A. (2014). Association of TCF7L2 variation with single islet autoantibody expression in children with type 1 diabetes. *BMJ Open Diabetes Res Care*, *2*(1), e000008. https://doi.org/10.1136/bmjdrc-2013-000008

Redondo, M. J., Steck, A. K., Sosenko, J., Anderson, M., Antinozzi, P., Michels, A., … Geyer, S. (2018). Transcription factor 7-like 2 (TCF7L2) gene polymorphism and progression from single to multiple autoantibody positivity in individuals at risk for type 1 diabetes. *Diabetes Care*. https://doi.org/10.2337/dc18-0861

Sachan, A., Zaidi, G., Sahu, R. P., Agrawal, S., Colman, P. G., & Bhatia, E. (2015). Low prevalence of latent autoimmune diabetes in adults in northern India. *Diabet Med*, *32*(6), 810–813. https://doi.org/10.1111/dme.12644

Schwartz, S. S., Epstein, S., Corkey, B. E., Grant, S. F., Gavin 3rd, J. R., & Aguilar, R. B. (2016). The Time Is Right for a New Classification System for Diabetes: Rationale and Implications of the beta-Cell-Centric Classification Schema. *Diabetes Care*, *39*(2), 179–186. https://doi.org/10.2337/dc15-1585

Scott, R. A., Scott, L. J., Magi, R., Marullo, L., Gaulton, K. J., Kaakinen, M., … Meta-analysis, C. (2017). An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes*. https://doi.org/10.2337/db16-1253

Shaltout, A. A., Wake, D., Thanaraj, T. A., Omar, D. M., Al-AbdulRazzaq, D., Channanath, A., … Steering Group for the Study of Childhood Diabetes in, K. (2017). Incidence of type 1 diabetes has doubled in Kuwaiti children 0-14 years over the last 20 years. *Pediatr Diabetes*, *18*(8), 761–766. https://doi.org/10.1111/pedi.12480

Sharp, S. A., Rich, S. S., Wood, A. R., Jones, S. E., Beaumont, R. N., Harrison, J. W., … Oram, R. A. (2019). Development and standardization of an improved type 1 diabetes genetic risk score for use in newborn screening and incident diagnosis. *Diabetes Care*, *42*(February), 200–207. https://doi.org/10.2337/dc18-1785

Shiroishi, M., Tsumoto, K., Amano, K., Shirakihara, Y., Colonna, M., Braud, V. M., … Maenaka, K. (2003). Human inhibitory receptors Ig-like transcript 2 (ILT2) and ILT4 compete with CD8 for MHC class I binding and bind preferentially to HLA-G. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1431057100

Sinnott, S. J., McHugh, S., Whelton, H., Layte, R., Barron, S., & Kearney, P. M. (2017). Estimating the prevalence and incidence of type 2 diabetes using population level pharmacy claims data: a cross-

sectional study. *BMJ Open Diabetes Res Care*, *5*(1), e000288. https://doi.org/10.1136/bmjdrc-2016-000288
Siraj, E. S., Gupta, M. K., Yifter, H., Ahmed, A., Kebede, T., Reja, A., & Abdulkadir, J. (2016). Islet cell-associated autoantibodies in Ethiopians with diabetes mellitus. *Journal of Diabetes and Its Complications*. https://doi.org/10.1016/j.jdiacomp.2016.05.005
Soriguer-Escofet, F., Esteva, I., Rojo-Martinez, G., Ruiz De Adana, S., Catalá, M., Merelo, M. J., … Freire, J. M. (2002). Prevalence of latent autoimmune diabetes of adults (LADA) in Southern Spain. *Diabetes Research and Clinical Practice*. https://doi.org/10.1016/S0168-8227(02)00002-5
Speed, D., Cai, N., Johnson, M. R., Nejentsev, S., & Balding, D. J. (2017). Reevaluation of SNP heritability in complex human traits. *Nature Genetics*. https://doi.org/10.1038/ng.3865
Speed, D., Hemani, G., Johnson, M. R., & Balding, D. J. (2012). Improved heritability estimation from genome-wide SNPs. *American Journal of Human Genetics*. https://doi.org/10.1016/j.ajhg.2012.10.010
Steck, A. K., Xu, P., Geyer, S., Redondo, M. J., Antinozzi, P., Wentworth, J. M., … Type 1 Diabetes TrialNet Study, G. (2017). Can Non-HLA Single Nucleotide Polymorphisms Help Stratify Risk in TrialNet Relatives at Risk for Type 1 Diabetes? *J Clin Endocrinol Metab*, *102*(8), 2873–2880. https://doi.org/10.1210/jc.2016-4003
Takeda, H., Kawasaki, E., Shimizu, I., Konoue, E., Fujiyama, M., Murao, S., … Ehime, S. (2002). Clinical, autoimmune, and genetic characteristics of adult-onset diabetic patients with GAD autoantibodies in Japan (Ehime Study). *Diabetes Care*, *25*(6), 995–1001. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12032105
Tasker, R. C., & Acerini, C. L. (2014). Cerebral edema in children with diabetic ketoacidosis: vasogenic rather than cellular? *Pediatr Diabetes*, *15*(4), 261–270. https://doi.org/10.1111/pedi.12153
Thomas, N. J., Jones, S. E., Weedon, M. N., Shields, B. M., Oram, R. A., & Hattersley, A. T. (2017). Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *Lancet Diabetes Endocrinol*. https://doi.org/10.1016/S2213-8587(17)30362-5
Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg.2017.101
Todd, J. A. (2010). Etiology of Type 1 Diabetes. *Immunity*. https://doi.org/10.1016/j.immuni.2010.04.001
Torn, C., Mueller, P. W., Schlosser, M., Bonifacio, E., Bingley, P. J., & Participating, L. (2008). Diabetes Antibody Standardization Program: evaluation of assays for autoantibodies to glutamic acid decarboxylase and islet antigen-2. *Diabetologia*, *51*(5), 846–852.

https://doi.org/10.1007/s00125-008-0967-2
Tull, E., & Roseman, J. (1995). Diabetes in African Americans. In *Diabetes in America* (2nd ed., pp. 613–630). Bethesda,MD: National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.
Tuomi, T., Carlsson, A., Li, H., Isomaa, B., Miettinen, A., Nilsson, A., … Groop, L. C. (1999). Clinical and genetic characteristics of type 2 diabetes with and without GAD antibodies. *Diabetes*, *48*(1), 150–157. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9892237
Tuomi, T., Groop, L. C., Zimmet, P. Z., Rowley, M. J., Knowles, W., & Mackay, I. R. (1993). Antibodies to glutamic acid decarboxylase reveal latent autoimmune diabetes mellitus in adults with a non-insulin-dependent onset of disease. *Diabetes*, *42*(2), 359–362. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8425674
Tuomi, T., Santoro, N., Caprio, S., Cai, M., Weng, J., & Groop, L. (2014). The many faces of diabetes: A disease with increasing heterogeneity. *The Lancet*. https://doi.org/10.1016/S0140-6736(13)62219-9
Turner, R., Stratton, I., Horton, V., Manley, S., Zimmet, P., Mackay, I. R., … Holman, R. (1997). UKPDS 25: autoantibodies to islet-cell cytoplasm and glutamic acid decarboxylase for prediction of insulin requirement in type 2 diabetes. UK Prospective Diabetes Study Group. *Lancet*, *350*(9087), 1288–1293. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9357409
Udler, M. S., Kim, J., von Grotthuss, M., Bonàs-Guarch, S., Cole, J. B., Chiou, J., … Florez, J. C. (2018). Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.1002654
Unnikrishnan, A. G., Singh, S. K., & Sanjeevi, C. B. (2004). Prevalence of GAD65 antibodies in lean subjects with type 2 diabetes. *Ann N Y Acad Sci*, *1037*, 118–121. https://doi.org/10.1196/annals.1337.018
Valdes, A. M., Erlich, H. A., & Noble, J. A. (2005). Human leukocyte antigen class I B and C loci contribute to Type 1 Diabetes (T1D) susceptibility and age at T1D onset. *Human Immunology*. https://doi.org/10.1016/j.humimm.2004.12.001
Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era - Concepts and misconceptions. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2322
Wang, Y., Yuan, W., Guo, H., & Jiang, Y. (2015). High frequency of activated NKp46^{+} natural killer cells in patients with new diagnosed of latent autoimmune diabetes in adults. *Autoimmunity*. https://doi.org/10.3109/08916934.2014.990629
Wellcome Trust Case Control, C. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, *447*(7145), 661–678.

https://doi.org/10.1038/nature05911
Weyand, C. M., Fujii, H., Mohan, S. V., Goronzy, J. J., & Yang, Z. (2013). Phosphofructokinase deficiency impairs ATP generation, autophagy, and redox balance in rheumatoid arthritis T cells. *The Journal of Experimental Medicine*. https://doi.org/10.1084/jem.20130252
Winkler, C., Krumsiek, J., Buettner, F., Angermuller, C., Giannopoulou, E. Z., Theis, F. J., … Bonifacio, E. (2014). Feature ranking of type 1 diabetes susceptibility genes improves prediction of type 1 diabetes. *Diabetologia*, *57*(12), 2521–2529. https://doi.org/10.1007/s00125-014-3362-1
Wod, M., Thomsen, R. W., Pedersen, L., Yderstraede, K. B., Beck-Nielsen, H., & Hojlund, K. (2018). Lower mortality and cardiovascular event rates in patients with latent autoimmune diabetes in adults (LADA) as compared with type 2 diabetes and insulin deficient diabetes: A Cohort Study of 4,368 Patients. *Diabetes Res Clin Pract*. https://doi.org/10.1016/j.diabres.2018.02.042
Xiang, Y., Huang, G., Shan, Z., Pan, L., Luo, S., Yang, L., … Zhou, Z. (2015). Glutamic acid decarboxylase autoantibodies are dominant but insufficient to identify most Chinese with adult-onset non-insulin requiring autoimmune diabetes: LADA China study 5. *Acta Diabetol*, *52*(6), 1121–1127. https://doi.org/10.1007/s00592-015-0799-8
Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Madden, P. A. F., Heath, A. C., … Visscher, P. M. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics*. https://doi.org/10.1038/ng.2213
Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*. https://doi.org/10.1016/j.ajhg.2010.11.011
Yang, J., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M., … Visscher, P. M. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet*, *43*(6), 519–525. https://doi.org/10.1038/ng.823
Zaharieva, E. T., Velikova, T. V, Tsakova, A. D., & Kamenov, Z. A. (2017). Prevalence of Positive Diabetes-Associated Autoantibodies among Type 2 Diabetes and Related Metabolic and Inflammatory Differences in a Sample of the Bulgarian Population. *J Diabetes Res*, *2017*, 9016148. https://doi.org/10.1155/2017/9016148
Zampetti, S., Spoletini, M., Petrone, A., Capizzi, M., Arpi, M. L., Tiberti, C., … Nirad Study, G. (2010). Association of TCF7L2 gene variants with low GAD autoantibody titre in LADA subjects (NIRAD Study 5). *Diabet Med*, *27*(6), 701–704. https://doi.org/10.1111/j.1464-5491.2010.02997.x
Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., … Neale, B. M. (2017). LD Hub: A centralized

database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btw613

Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*, *44*(7), 821–824. https://doi.org/10.1038/ng.2310

Zhou, Z., Xiang, Y., Ji, L., Jia, W., Ning, G., Huang, G., … Group, L. C. S. (2013). Frequency, immunogenetics, and clinical characteristics of latent autoimmune diabetes in China (LADA China study): a nationwide, multicenter, clinic-based cross-sectional study. *Diabetes*, *62*(2), 543–550. https://doi.org/10.2337/db12-0207

Zinman, B., Kahn, S. E., Haffner, S. M., O'Neill, M. C., Heise, M. A., Freed, M. I., & Group, A. S. (2004). Phenotypic characteristics of GAD antibody-positive recently diagnosed patients with type 2 diabetes in North America and Europe. *Diabetes*, *53*(12), 3193–3200. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15561950