Genetics of Low Polygenic Risk Score Type 1 Diabetes Patients: rare variants in 22 novel loci

Jingchun Qu[1], Hui-Qi Qu[1], Jonathan Bradfield[2], Joseph Glessner[1], Xiao Chang[1], Lifeng Tian[1], Michael March[1], Jeffrey D Roizen[3], Patrick Sleiman[1,3,4], Hakon Hakonarson[1,3,4,5†].

Affiliations:

[1]The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, 19104, USA.

[2]Quantinuum Research LLC, San Diego, California, 92101, USA

[3]Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, 19104, USA.

[4]Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, 19104, USA.

[5]Division of Pulmonary Medicine Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, 19104, USA.

† Corresponding authors:

* Corresponding author:
Dr. Hakon Hakonarson
Center for Applied Genomics
3615 Civic Center Blvd
Abramson Building
Philadelphia, PA 19104, United States of America
Telephone: 267-426-0088
Fax: 267-426-0363
Email: hakonarson@email.chop.edu

Abstract

With polygenic risk score (PRS) for autoimmune type 1 diabetes (T1D), this study identified T1D cases with low T1D PRS and searched for susceptibility loci in these cases. Our hypothesis is that genetic effects (likely mediated by relatively rare genetic variants) of non-mainstream (or non-autoimmune) T1D might have been diluted in the previous studies on T1D cases in general. Two cohorts for the PRS modeling and testing respectively were included. The first cohort consisted of 3,356 T1D cases and 6,203 controls, and the independent second cohort consisted of 3,355 T1D cases and 6,203 controls. Cases with low T1D PRS were identified using PRSice-2 and compared to controls with low T1D PRS by genome-wide association (GWA) test. Twenty-six genetic loci with SNPs/SNVs associated with low PRS T1D at genome-wide significance ($P \leq 5.0xE-08$) were identified, including 4 established T1D loci, as well as 22 novel loci represented by rare SNVs. For the 22 novel loci, 12 regions have been reported of association with obesity related traits by previous GWA studies. Five loci encoding long intergenic non-protein coding RNAs (lncRNA), two loci involved in N-linked glycosylation, two loci encoding GTPase activators, and two ciliopathy genes, are also highlighted in this study.

Key words: Genome-wide Association Study; N-linked glycosylation; Non-autoimmune; Long intergenic non-protein coding RNA; Polygenic Risk Score; Type 1 Diabetes

Introduction

Type 1 diabetes (T1D) is caused by T-cell mediated autoimmune destruction of pancreatic β-cells(1). There is no cure for T1D to date. The molecular mechanisms underlying T1D are complex and not completely understood. Human genetic studies have uncovered multiple T1D genes that contribute to our understanding of the pathogenesis ofT1D(2-7). With the rapid advances in human genomics technology in recent years, over 70 T1D loci have been identified(8) (https://www.ebi.ac.uk/gwas/). While these discoveries of T1D-associated genes have greatly increased our knowledge of T1D, our current genetic knowledge on T1D is far from complete, and a large number of T1D genes remain uncovered(9). A key bottleneck for the GWAS approach is limitation of sample size even with the presense of collaborative international consortia(10). The phenotype of type 1 diabetes has been regarded as heterogeneous. While the majority of T1D patients have autoimmune disease, 5–10% of Caucasian diabetic subjects with recent-onset T1D do not have islet cell antibodies, often referred to as T1bD(11). Due to different pathogenesis, T1bD cases may be associated with different genetic loci from autoimmune T1D, or T1aD. However, the smaller proportion of T1bD cases suggests that T1bD-related genetic effects have been diluted in the previous studies with T1D cases studied in general. Besides T1bD, the non-autoimmune and monogenic form of pediatric diabetes, maturity-onset diabetes of the young (MODY) cases, may be misdiagnosed as T1D(12), which further contributes to the heterogeneity of the T1D phenotype.

With numerous genetic loci for many human complex diseases identified to date, polygenic risk scores (PRS) aggregate the effects of many genetic variants across the human genome into a single score, an approach that has been shown of improve disease prediction and differential diagnosis(13). The T1D loci identified by the GWAS studies to date are mainly associated with the genetic susceptibility of the major component of the heterogeneous T1D phenotype, i.e. T1aD, while the genetic susceptibility of the minor non-autoimmune components (e.g. T1bD and misdiagnosed MODY) are undere-represented in those results likely as a result of being diluted In this study, we propose that a high T1D PRS score predicts or suggests a T1aD case, whereas a low T1D PRS score in a T1D case suggests the opposite and represents our major interest in this study. Our aim in this study is to identify low PRS T1D cases and to run a seprate GWAS in an attempt to uncover genetic loci associated with T1bD patients.

Methods

Subjects: 6,711 European T1D cases and 12,406 European controls were included in this study. The T1D cases were from the Children's Hospital of Philadelphia (CHOP)(14), The Diabetes Control and Complications Trial – Epidemiology of Diabetes Interventions and Complications (DCCT-EDIC) cohort (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000086.v2.p1), the Type 1 Diabetes Genetics Consortium (T1DGC, http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000180.v1.p1), and later recruited subjects at CHOP, respectively. The genotyping was done by the Illumina Human Hap550 Genotyping BeadChip or a newer version of Illumina Genotyping BeadChip. Other demographic, phenotypic and genotypic details about these individuals were described in our previous publication(15). Imputation of 39,131,579 single nucleotide polymorphisms (SNP) on auto-chromosomes was done using the Sanger Imputation Service (https://www.sanger.ac.uk/tool/sanger-imputation-service/) based on the Haplotype Reference

Consortium (HRC) r1.1 reference panel (HRC.r1-1.GRCh37.wgs.mac5.sites.tab), with the quality filters of $R^2 \geq 0.4$. Altogether, 32,251,301 autosomal single nucleotide variants (SNV) with quality $R^2 \geq 0.4$ were included in this study. Population stratification was assessed by principal component analysis (PCA), and genetic association tests were corrected by the first 10 principal components (PC). The association test was done using PLINK1.9 software(16).

Polygenic risk scores (PRS): To avoid the issue of overfitting for PRS scoring, the subjects were randomly splitted into two independent cohorts without duplication, i.e. the PRS training cohort including 3,356 T1D cases and 6,203 controls, and the PRS testing cohort including 3,355 T1D cases and 6,203 controls. PRSs of the test cohort were calculated using the Polygenic Risk Score software (PRSice-2)(17), based on the statistics of the training group. The performance of a series of cutoff of T1D association P-values (including 10^{-10}, 10^{-9} , 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 0.001, 0.01, 0.05, 0.1, 0.2, and 1) for selection of SNP markers was assessed by the Area Under the ROC Curve (AUC). The P-value cutoff with the largest AUC was adopted.

GWAS of T1D patients with low PRS: According to the PRS values, the T1D patients were separated into two groups, i.e. a low PRS group and a high PRS group. The PRS cutoff was determined by the maximum Matthews correlation coefficient (MCC). Using the same PRS cutoff, health controls with low T1D PRS were identified. The GWAS of T1D patients with low PRS was performed by comparing to health controls with low T1D PRS. The Manhattan plots were done using the web-based FUMA platform(18). Genetic association signals within each locus were plotted by LocusZoom(19).

Data and Resource Availability: The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Results

AUC of different cutoffs of T1D association P-values for SNP selection and PRS

The AUCs of different cutoffs of T1D association P-values for selection of SNP sets are shown in Table 1a. The best AUC (0.8607) is seen at the cutoff of P-value≤1E-05, which suggests that stricter cutoff may cause the missing of informative SNPs, while looser may introduce noise by including SNPs with spurious T1D association. Based on the SNP markers with T1D association P-value≤1E-05, a PRS score was acquired for each individual in the independent test cohort. By the maximum MCC (Supplementary Table 1), a PRS cutoff of 1.11E-03 has the maximum MCC (0.6294). A PRS≤1.11E-03 was defined as low risk, and a PRS>1.11E-03 was defined as high risk. With this threshold, the sensitivity (True positive rate, TPR) for T1D prediction is 75.9%, and the specificity (True negative rate, TFR) for T1D prediction is 86.4%. By PRS≤1.11E-03, 810 (24.1%, including 408 males, 400 females, and 2 cases with undetermined sex) out of 3,355 T1D cases had low PRS; and 5,358 (86.4%, including 2,893 males, 2,453 females, and 12 cases with undetermined sex) out of 6,203 controls had low PRS.

GWAS of T1D patients with low PRS

The GWAS of T1D patients with low T1D PRS compared to controls with low T1D PRS identified a large number of SNPs associated with T1D with genome-wide significance (P≤5.0xE-08), from 7 genetic loci (Supplementary Table 2, Figure 1). Among these 7 genetic loci,

3 loci have been established of T1D association by previous studies, including *HLA*, *INS*, and *PTPN22* (Table 2a). By looking at the established leading T1D signal of each locus, the frequencies of the predisposing alleles of *HLA* and *PTPN22* were lower in the low T1D PRS cohort, while the protective allele of *INS* were higher in the low T1D PRS cohort. The effect sizes of *HLA* (P=2.72E-06) and *PTPN22* (P=0.047) were significantly smaller in the low PRS cases. Besides these 3 established T1D loci, 4 novel loci associated with low PRS T1D were identified (Table 3a). LocusZoom plots for genetic association signals within each locus are shown in Supplementary Figure 1-4. The association signals of these loci are only seen in low PRS T1D cases, but not in the T1D cases overall, and were missed previously due to diluted genetic effects.

Replication of the PRS model and additional novel loci

Consequently, we switched the two cohorts, i.e. using the second cohort for the statistics of PRS modelling, then we tested the PRS models in the first cohort. The AUCs of different cutoffs of T1D association P-values for selection of SNP sets are shown in Table 1b. The best AUC (0.8654) is seen at the cutoff of P-value≤1E-05, which repeated the PRS model in the above step. Based on the SNP markers with T1D association P-value≤1E-05, a PRS score was acquired for each individual in the independent test cohort. By the maximum MCC (Supplementary Table 3), a PRS cutoff of 1.24E-03 has the maximum MCC (0.6294). A PRS≤7.18E-04 was defined as low risk, and a PRS>7.18E-04 was defined as high risk. With this threshold, the sensitivity (True positive rate, TPR) for T1D prediction is 66.0%, and the specificity (True negative rate, TFR) for T1D prediction is 93.6%. By PRS≤7.18E-04, 918 (27.4%, including 437 males, 479 females, and 2 cases with undetermined sex) out of 3,356 T1D cases had low PRS; and 5,585 (90.0%, including 3,008 males, 2,565 females, and 12 cases with undetermined sex) out of 6,203 controls had low PRS.

As expected from the above results, in the switched cohort, the GWAS of T1D patients with low T1D PRS compared to controls with low T1D PRS identified a large number of SNPs associated with T1D with genome-wide significance (P≤5.0xE-08) as well (Supplementary Table 4, Figure 2). Among these loci, 4 loci have been established of T1D association by previous studies, including *HLA*, *INS*, *PTPN22*, and the *IKZF4/RPS26/ERBB3* locus (Table 2b). Consistent to the first GWAS results listed above, by looking at the established leading T1D signal of each locus, the frequencies of the predisposing alleles of *HLA*, *PTPN22* and *IKZF4* were lower in the low T1D PRS cohort, while the protective allele of *INS* were higher in the low T1D PRS cohort. The effect size of the leading *HLA* SNP was significantly smaller in the low PRS cases (P=1.05E-11). Besides these established T1D loci, 18 novel loci associated with low PRS T1D were identified in this cohort (Table 3b). LocusZoom plots for genetic association signals within each locus are shown in Supplementary Figure 5-22.

Discussion

Altogether, rare variants (MAF<5%) from 22 novel loci were identified in the low PRS T1D cases with genome-wide significance (P<5.00E-08), in addition to the 4 established T1D loci with smaller genetic effects in these cases. The association signals of these loci are only seen in low PRS T1D cases, but not in the T1D cases overall, and were missed previously due to rare allele frequencies and diluted genetic effects in the general T1D cohort. Among the 22 loci, two

genetic regions have been reported of association with diabetes, i.e. the region containing the *DLL1/ FAM120B* locus associated with type 1 diabetes in Caucasian by our previous study(20), and the region containing the *TICRR* locus associated Type 2 diabetes in African population(21). In addition, a number of genetic associations with body mass index (BMI), obesity, and autoimmunity, have been reported in the flanking regions of 300kb on each side of the new loci according to the GWAS Catalog (https://www.ebi.ac.uk/gwas/, Supplementary materials). Further details on these 22 loci are described below.

***LINC01865/LINC01874* tagged by rs186500234**

The long intergenic non-protein coding RNA 1865 gene (*LINC01865*) has low expression observed in testis, brain, and duodenum. The long intergenic non-protein coding RNA 1874 gene (*LINC01874*) has restricted expression toward kidney(22). This genetic region has been reported of association with body mass index (BMI) by previous study(23).

***LOC730100* tagged by rs28957087**

LOC730100 encodes a long non-coding RNA (ncRNA), a competing endogenous RNA for human microRNA 760 (miR-760)(24). The latter inhibits the expression of the Forkhead Box A1 gene (*FOXA1*). As a hepatocyte nuclear factor, *FOXA1*, also known as *HNF3A* or *TCF3A*, regulates tissue-specific gene expression in liver and many other tissues(25). FoxA1 is essential for normal pancreatic and ß-cell function and a negative regulator of the hepatocyte nuclear factor-1 (HNF1) homeobox A gene (*HNF1A*) and the hepatocyte nuclear factor 4, alpha gene (*HNF4A*)(26) (27). *HNF1A* and *HNF4A* are established genes causing maturity-onset diabetes of the young (MODY). The *FOXA1* mutation Ser448Asn has been suggested of association with impaired glucose homeostasis(27).

***B3GNT2/TMEM17* tagged by rs75634056**

The UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2 gene (*B3GNT2*) encodes an enzyme involved in the biosynthesis of poly-N-acetyllactosamine chains. The gene plays an important role in immunological biofunctions, and its deficiency causes hyperactivation of lymphocytes in mice(28). The transmembrane protein 17 gene (*TMEM17*) encodes a critical component of a protein complex at the base of cilia. Previous GWAS studies have reported association with Crohn's disease, ankylosing spondylitis, and hypothyroidism in this genetic region.

***FAM136A/TGFA* tagged by rs77418738**

The family with sequence similarity 136 member A gene (*FAM136A*) encodes a mitochondrially localized protein. The transforming growth factor alpha gene (TGFA) mediates cell-cell adhesion and activates cell proliferation, differentiation and development. This region has been reported of association with obesity-related traits(29).

***GCC2/EDAR* tagged by rs922452**

The GRIP and coiled-coil domain containing 2 (GCC2) encodes a long coiled-coil protein, also known as GCC185, which is localized to the trans-Golgi network with critical function in

maintaining Golgi structure and tethering transport vesicle(30). The ectodysplasin A receptor gene (EDAR) encodes a member of the tumor necrosis factor receptor family with a key role in ectodermal differentiation. Association with low birth weight at this region has been reported(31).

***SEL1L3* tagged by rs6842426**

The locus SEL1L family member 3 gene (*SEL1L3*) is a paralog of the SEL1L adaptor subunit of ERAD E3 ubiquitin ligase gene (*SEL1L*). The latter is highly expressed in pancreas and thyroid, and is crucial for misfolded proteins in the endoplasmic reticulum being discharged into the cytosol and degraded by the proteasome(32). This gene region has been reported association with obesity-related traits(29) and non-alcoholic fatty liver disease(33).

***TBC1D1/LINC01258* tagged by rs4833044**

This genetic locus contains two genes, the TBC1 domain family member 1 gene (*TBC1D1*) and the long intergenic non-protein coding RNA 1258 gene (*LINC01258*). A common variant it this locus has been reported to be associated with childhood obesity(29; 34), triacylglycerol 54:5 levels(35), lymphocyte percentage of leukocytes(36) by previous studies. Acting as a GTPase activator, the *TBC1D1* protein plays a role in regulating cell growth and differentiation. Rare mutations in *TBC1D1* have been reported to be associated with congenital anomalies of the kidney and urinary tract.

***LINC02432/IL15* tagged by rs9790756**

The long intergenic non-protein coding RNA 2432 gene (*LINC02432*) has higher expression in kidney and pancreas. Interleukin 15 (IL-15) encoded by the gene *IL15* is essential for regulating activation and proliferation of T and natural killer cells, and supporting lymphoid homeostasis. IL-15 and interleukine 2 (IL-2) share many biological activities and receptor components with IL-2. IL-2 is a powerful growth factor for both T and B lymphocytes. Both IL2 and the α chain of the IL2 receptor complex gene (*IL2RA*) has been established of genetic association with T1D by previous studies(37-39).

***DEK/RNF144B* tagged by rs16880565**

The DEK proto-oncogene gene (*DEK*) encodes a site-specific DNA binding protein and a component of the pre-mRNA splicing complex, and is involved in transcriptional regulation and pre-mRNA splicing. *DEK* encoded protein is also an autoantigen in patients with pauciarticular onset juvenile rheumatoid arthritis. The ring finger protein 144B gene (*RNF144B*) encoded protein inhibits LPS-induced inflammatory responses by binding with TANK binding kinase 1 (TBK1) and causing interferon regulatory factor 3 (IRF3) dephosphorylation and interferon β (IFN-β) reduction. This region has been reported of association with BMI by previous studies(23; 40).

***RGS17* tagged by rs80292134**

The regulator of G protein signaling 17 gene (*RGS17*) encodes a member of the regulator of G-protein signaling family. This genetic region has been established association with BMI by previous studies(23; 40).

***DLL1/FAM120B* tagged by rs3800237**

The delta like canonical Notch ligand 1 (DLL1) encodes a Notch ligand with a role in cell-fate decision processes in lymphopoiesis. This Notch ligand can completely inhibit the differentiation of human hematopoietic progenitors into the B cell lineage while promoting the generation of T cell/natural killer (NK) precursors(41). The family with sequence similarity 120B gene (FAM120B) encodes a constitutive coactivator of peroxisome proliferator-activated receptor γ (PPARγ, a major therapeutic target for insulin sensitivity) and promotes adipogenesis(42). The region containing the *DLL1/ FAM120B* genes has been reported of association with T1D in Caucasian by our previous study(20).

***NME8/GPR141* tagged by rs12532321**

The NME/NM23 family member 8 gene (*NME8*) encodes an axoneme protein, and its mutation may cause primary ciliary dyskinesia. The G protein-coupled receptor 141 gene (*GPR141*) at the upstream of *NME8* is highly expressed in bone marrow. This genetic region has been reported of association with obesity-related traits in Hispanic children(29).

***CALN1* tagged by rs118182411**

The calneuron 1 gene (*CALN1*), encoding a protein with high similarity to the calcium-binding proteins of calmodulin, is highly expressed in brain and adrenal. This genetic region has established association with BMI by previous studies(23; 40).

***ZNF804B* tagged by rs77205087**

The zinc finger protein 804B gene (*ZNF804B*) has been reported of association with N-linked glycosylation of human immunoglobulin G (IgG), which modulates its binding to Fc receptors(43). N-glycosylation of cytokines and proteases is also a regulatory mechanism in inflammation and autoimmunity(44). Changes in N-glycosylation have been associated with different autoimmune diseases, including rheumatoid arthritis(45), type 1 diabetes(46), Crohn's disease(47).

***NFIB* tagged by rs10961435**

The nuclear factor I B gene (*NFIB*) encodes a transcription factor in the FOXA1 transcription factor network. NFIB has been shown to play critical roles in lung and brain development. A previous study has shown that NFIB can bind with FoxA1 and modulate the transcriptional activity of FoxA1(48), while the later has been suggested to play a role in pancreatic and ß-cell function and non-autoimmune diabetes as discussed above.

***TBC1D2/GABBR2* tagged by rs11559334**

This genetic locus contains two protein-coding genes, the TBC1 domain family member 2 gene (*TBC1D2*) and the gamma-aminobutyric acid type B receptor subunit 2 gene (*GABBR2*, encoding a member of the G-protein coupled receptor 3 family). As discussed above, this study identified an association signal in the *TBC1D1* region, and the *TBC1D1* locus has been reported of association with childhood obesity(29; 34).

***LINC00841/C10orf142* tagged by rs746298**

The two genes at this locus, *LINC00841/C10orf142,* encode two long intergenic non-protein coding RNAs (lincRNA). While the function of these two genes remain unknown, this locus has been reported of association with obesity-related traits(29).

***SYT10/ALG10* tagged by rs10506114**

The synaptotagmin 10 gene (*SYT10*) encodes a membrane protein of secretory vesicles expressed in pancreas, lung and kidney(49). The ALG10 alpha-1,2-glucosyltransferase gene (*ALG10*) encodes a membrane-associated protein that adds the third glucose residue to the lipid-linked oligosaccharide precursor for N-glycosylation in endoplasmic reticulum (ER)(50). As discussed above in the *ZNF804B* locus, N-glycosylation of IgG, cytokines and proteases is also a regulatory mechanism in inflammation and autoimmunity(43; 44) associated with different autoimmune diseases. This region has established association with waist-hip ratio by previous study(40).

***CHST11* tagged by rs75438334**

The carbohydrate sulfotransferase 11 gene (*CHST11*) encodes a member of the sulfotransferase 2 family catalyzing chondroitin sulfate synthesis. This genetic region has been reported of association with waist circumference adjusted for body mass index by previous study(51).

***CHFR/LOC101928530/ZNF605* tagged by rs12230138**

The checkpoint with forkhead and ring finger domains gene (*CHFR*) encodes an E3 ubiquitin-protein ligase and is involved in the DNA damage response and checkpoint regulation. The structure and function of the gene *LOC101928530* is still uncharacterized. The function of the zinc finger protein 605 gene (*ZNF605*) may be related to Herpes Simplex Virus 1 infection (https://pathcards.genecards.org/card/herpes_simplex_virus_1_infection). This region has been reported of association with BMI by previous study(52).

***TICRR/ KIF7* tagged by rs2197053**

The TOPBP1 interacting checkpoint and replication regulator gene (*TICRR*) encodes Treslin, which is involved in triggering the initiation of DNA replication. The kinesin family member 7 gene (*KIF7*) in this region encodes a cilia-associated protein of the kinesin family, with its mutations causing ciliopathies. The region containing the *TICRR* gene has been reported of association with T2D in African population(21), BMI(23; 52), and obesity-related traits(29) by previous studies.

***LINC01695/LINC00161* tagged by rs7278151**

Function of the long intergenic non-protein coding RNA 1695 gene (*LINC01695*) is still uncharacterized. The long intergenic non-protein coding RNA 161 gene (*LINC00161*) encodes a functional RNA that regulates Mitogen-activated protein kinase 1 (MAPK1) expression. The MAPK1/STAT3 pathway has been proposed as a novel diabetes target for its critical role in glucose homeostasis(53).

In summary, in the genetic regions containing the 22 novel loci disclosed by this study, more than half of these regions have been reported of association with obesity-related traits, BMI, or waist circumference. The correlation with obesity related traits or impaired glucose homeostasis is in keeping with non-autoimmune roles in the diabetes patients with low T1D PRS. Interestingly, genes related N-linked glycosylation, e.g. *ZNF804B* and *ALG10*, are highlighted in this study, which may suggest the role of N-glycosylation bridging impaired glucose homeostasis and autoimmune diabetes. N-glycosylation is commonly altered in diabetes(54). This particular locus supports an interesting hypothesis of T1D pathogenesis, i.e. the accelerator hypothesis, which implies that increasing obesity-associated insulin resistance accelerates the disease process of type 1 diabetes(55; 56). Insulin resistance-related mechanisms might thus be able to serve as potential novel therapeutic targets for these patients with low T1D PRS.

In addition, 5 loci encoding long intergenic non-protein coding RNAs (lncRNA) identified in this study emphasize the importance of lncRNAs in these diabetes patients. This study identified 2 loci containing *TBC1D1* and *TBC1D2* respectively, encoding two GTPase activators. *TBC1D1* has been suggested as a novel obesity gene by previous study(34). Two loci containing the *TMEM17* and *KIF7* genes corrected with ciliopathies suggest a role of primary cilia in diabetes(57). However, we admit that this study has limitations related to the bottleneck of sample size and data resources. The novel loci reported in this study still need replication in independent samples. In addition, the functional mechanisms of these genetic loci in diabetes warrant experimental investigation.

Acknowledgement: The authors apologize that many important references listed in the supplementary materials cannot be cited in the main text because of page limitation. The study was supported by Institutional Development Funds from the Children's Hospital of Philadelphia to the Center for Applied Genomics and The Children's Hospital of Philadelphia Endowed Chair in Genomic Research to HH. Dr. Hakon Hakonarson is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests: none to declare.

Reference:

1. Atkinson MA, Eisenbarth GS, Michels AW: Type 1 diabetes. The Lancet 2014;383:69-82
2. Todd JA, Bell JI, McDevitt HO: HLA-DQ[beta] gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 1987;329:599-604
3. Baisch JM, Weeks T, Giles R, Hoover M, Stastny P, Capra JD: Analysis of HLA-DQ genotypes and susceptibility in insulin-dependent diabetes mellitus. N Engl J Med 1990;322:1836-1841
4. Todd JA: Genetic Analysis of Type 1 Diabetes Using Whole Genome Approaches. PNAS 1995;92:8560-8565
5. Noble JA, Valdes AM, Cook M, Klitz W, Thomson G, Erlich HA: The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. Am J Hum Genet 1996;59:1134-1148
6. she J-X: Susceptibility to type I diabetes: HLA-DQ and DR revisited. Immunology Today 1996;17:323
7. Bell GI, Horita S, Karam JH: A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. Diabetes 1984;33:176-183
8. Onengut-Gumuscu S, Chen W-M, Burren O, Cooper NJ, Quinlan AR, Mychaleckyj JC, Farber E, Bonnie JK, Szpak M, Schofield E: Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal

variants with lymphoid gene enhancers. Nature genetics 2015;47:381-386
9. Polychronakos C, Li Q: Understanding type 1 diabetes through genetics: advances and prospects. Nature Reviews Genetics 2011;12:781-792
10. Rich SS, Concannon P, Erlich H, Julier C, Morahan G, Nerup J, Pociot F, Todd JA: The type 1 diabetes genetics consortium. Annals of the New York Academy of Sciences 2006;1079:1-8
11. Leslie RD, Atkinson MA, Notkins AL: Autoantigens IA-2 and GAD in Type I (insulin-dependent) diabetes. Diabetologia 1999;42:3-14
12. Ehtisham S, Hattersley A, Dunger D, Barrett T: First UK survey of paediatric type 2 diabetes and MODY. Archives of disease in childhood 2004;89:526-529
13. Lambert SA, Abraham G, Inouye M: Towards clinical utility of polygenic risk scores. Human Molecular Genetics 2019;28:R133-R142
14. Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Casalunovo T, Taback SP, Frackelton EC, Lawson ML, Robinson LJ, Skraban R, Lu Y, Chiavacci RM, Stanley CA, Kirsch SE, Rappaport EF, Orange JS, Monos DS, Devoto M, Qu HQ, Polychronakos C: A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448:591-594
15. Bradfield JP, Qu H-Q, Wang K, Zhang H, Sleiman PM, Kim CE, Mentch FD, Qiu H, Glessner JT, Thomas KA: A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. PLoS genetics 2011;7:e1002293
16. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ: Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:s13742-13015-10047-13748
17. Choi SW, O'Reilly PF: PRSice-2: Polygenic Risk Score software for biobank-scale data. GigaScience 2019;8
18. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D: Functional mapping and annotation of genetic associations with FUMA. Nature communications 2017;8:1-11
19. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ: LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics (Oxford, England) 2010;26:2336-2337
20. Bradfield JP, Qu HQ, Wang K, Zhang H, Sleiman PM, Kim CE, Mentch FD, Qiu H, Glessner JT, Thomas KA, Frackelton EC, Chiavacci RM, Imielinski M, Monos DS, Pandey R, Bakay M, Grant SF, Polychronakos C, Hakonarson H: A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. PLoS Genet 2011;7:e1002293
21. Chen J, Sun M, Adeyemo A, Pirie F, Carstensen T, Pomilla C, Doumatey AP, Chen G, Young EH, Sandhu M, Morris AP, Barroso I, McCarthy MI, Mahajan A, Wheeler E, Rotimi CN, Motala AA: Genome-wide association study of type 2 diabetes in Africa. Diabetologia 2019;62:1204-1211
22. Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, Asplund A, Sjöstedt E, Lundberg E, Szigyarto CA, Skogs M, Takanen JO, Berling H, Tegel H, Mulder J, Nilsson P, Schwenk JM, Lindskog C, Danielsson F, Mardinoglu A, Sivertsson A, von Feilitzen K, Forsberg M, Zwahlen M, Olsson I, Navani S, Huss M, Nielsen J, Ponten F, Uhlén M: Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Molecular & cellular proteomics : MCP 2014;13:397-406
23. Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, O'Connor LJ, Himes BE, Gazal S, Hasegawa K, Camargo CA, Jr., Qi L, Moffatt MF, Hu FB, Lu Q, Cookson WOC, Liang L: Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. The Journal of allergy and clinical immunology 2020;145:537-549
24. Li Q, Lu J, Xia J, Wen M, Wang C: Long non-coding RNA LOC730100 enhances proliferation and invasion of glioma cells through competitively sponging miR-760 from FOXA1 mRNA. Biochemical and biophysical research communications 2019;512:558-563
25. Lee CS, Friedman JR, Fulmer JT, Kaestner KH: The initiation of liver development is dependent on Foxa transcription factors. Nature 2005;435:944-947
26. Duncan SA, Navas MA, Dufort D, Rossant J, Stoffel M: Regulation of a transcription factor network required for differentiation and metabolism. Science (New York, NY) 1998;281:692-695
27. Navas MA, Vaisse C, Boger S, Heimesaat M, Kollee LA, Stoffel M: The human HNF-3 genes: cloning, partial sequence and mutation screening in patients with impaired glucose homeostasis. Human heredity 2000;50:370-381
28. Togayachi A, Kozono Y, Kuno A, Ohkura T, Sato T, Hirabayashi J, Ikehara Y, Narimatsu H: Beta3GnT2 (B3GNT2), a major polylactosamine synthase: analysis of B3GNT2-deficient mice. Methods in enzymology 2010;479:185-204
29. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, Butte NF: Novel genetic loci identified

for the pathophysiology of childhood obesity in the Hispanic population. PLoS One 2012;7:e51954
30. Brown FC, Schindelhaim CH, Pfeffer SR: GCC185 plays independent roles in Golgi structure maintenance and AP-1-mediated vesicle tethering. The Journal of cell biology 2011;194:779-787
31. Plotnikov D, Williams C, Guggenheim JA: Association between birth weight and refractive error in adulthood: a Mendelian randomisation study. The British journal of ophthalmology 2020;104:214-219
32. Mueller B, Klemm EJ, Spooner E, Claessen JH, Ploegh HL: SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. Proc Natl Acad Sci U S A 2008;105:12325-12330
33. Chalasani N, Guo X, Loomba R, Goodarzi MO, Haritunians T, Kwon S, Cui J, Taylor KD, Wilson L, Cummings OW, Chen YD, Rotter JI: Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. Gastroenterology 2010;139:1567-1576, 1576.e1561-1566
34. Stone S, Abkevich V, Russell DL, Riley R, Timms K, Tran T, Trem D, Frank D, Jammulapati S, Neff CD, Iliev D, Gress R, He G, Frech GC, Adams TD, Skolnick MH, Lanchbury JS, Gutin A, Hunt SC, Shattuck D: TBC1D1 is a candidate for a severe obesity gene and evidence for a gene/gene interaction in obesity predisposition. Hum Mol Genet 2006;15:2709-2720
35. Rhee EP, Ho JE, Chen MH, Shen D, Cheng S, Larson MG, Ghorbani A, Shi X, Helenius IT, O'Donnell CJ, Souza AL, Deik A, Pierce KA, Bullock K, Walford GA, Vasan RS, Florez JC, Clish C, Yeh JR, Wang TJ, Gerszten RE: A genome-wide association study of the human metabolome in a community-based cohort. Cell metabolism 2013;18:130-143
36. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR, Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A, Kamat M, Kim B, Mandoli A, Marchini J, Martens JHA, Meacham S, Megy K, O'Connell J, Petersen R, Sharifi N, Sheard SM, Staley JR, Tuna S, van der Ent M, Walter K, Wang SY, Wheeler E, Wilder SP, Iotchkova V, Moore C, Sambrook J, Stunnenberg HG, Di Angelantonio E, Kaptoge S, Kuijpers TW, Carrillo-de-Santa-Pau E, Juan D, Rico D, Valencia A, Chen L, Ge B, Vasquez L, Kwan T, Garrido-Martín D, Watt S, Yang Y, Guigo R, Beck S, Paul DS, Pastinen T, Bujold D, Bourque G, Frontini M, Danesh J, Roberts DJ, Ouwehand WH, Butterworth AS, Soranzo N: The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell 2016;167:1415-1429.e1419
37. Vella A, Cooper JD, Lowe CE, Walker N, Nutland S, Widmer B, Jones R, Ring SM, McArdle W, Pembrey ME, Strachan DP, Dunger DB, Twells RC, Clayton DG, Todd JA: Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. Am J Hum Genet 2005;76:773-779
38. Qu H-Q, Montpetit A, Ge B, Hudson TJ, Polychronakos C: Toward Further Mapping of the Association Between the IL2RA Locus and Type 1 Diabetes. Diabetes 2007;56:1174-1176
39. Plagnol V, Howson JM, Smyth DJ, Walker N, Hafler JP, Wallace C, Stevens H, Jackson L, Simmonds MJ, Bingley PJ, Gough SC, Todd JA: Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. PLoS Genet 2011;7:e1002216
40. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, Yengo L, Ferreira T, Marouli E, Ji Y, Yang J, Jones S, Beaumont R, Croteau-Chonka DC, Winkler TW, Hattersley AT, Loos RJF, Hirschhorn JN, Visscher PM, Frayling TM, Yaghootkar H, Lindgren CM: Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet 2019;28:166-174
41. Jaleco AC, Neves H, Hooijberg E, Gameiro P, Clode N, Haury M, Henrique D, Parreira L: Differential effects of Notch ligands Delta-1 and Jagged-1 in human lymphoid differentiation. J Exp Med 2001;194:991-1002
42. Li D, Kang Q, Wang DM: Constitutive coactivator of peroxisome proliferator-activated receptor (PPARgamma), a novel coactivator of PPARgamma that promotes adipogenesis. Molecular endocrinology (Baltimore, Md) 2007;21:2320-2333
43. Lauc G, Huffman JE, Pučić M, Zgaga L, Adamczyk B, Mužinić A, Novokmet M, Polašek O, Gornik O, Krištić J, Keser T, Vitart V, Scheijen B, Uh HW, Molokhia M, Patrick AL, McKeigue P, Kolčić I, Lukić IK, Swann O, van Leeuwen FN, Ruhaak LR, Houwing-Duistermaat JJ, Slagboom PE, Beekman M, de Craen AJ, Deelder AM, Zeng Q, Wang W, Hastie ND, Gyllensten U, Wilson JF, Wuhrer M, Wright AF, Rudd PM, Hayward C, Aulchenko Y, Campbell H, Rudan I: Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. PLoS Genet 2013;9:e1003225
44. Van den Steen P, Rudd PM, Dwek RA, Van Damme J, Opdenakker G: Cytokine and protease glycosylation as a regulatory mechanism in inflammation and autoimmunity. In *Glycoimmunology 2*, Springer, 1998, p. 133-143
45. Nakagawa H, Hato M, Takegawa Y, Deguchi K, Ito H, Takahata M, Iwasaki N, Minami A, Nishimura S: Detection of altered N-glycan profiles in whole serum from rheumatoid arthritis patients. Journal of chromatography B, Analytical technologies in the biomedical and life sciences 2007;853:133-137

46. Bermingham ML, Colombo M, McGurnaghan SJ, Blackbourn LAK, Vučković F, Pučić Baković M, Trbojević-Akmačić I, Lauc G, Agakov F, Agakova AS, Hayward C, Klarić L, Palmer CNA, Petrie JR, Chalmers J, Collier A, Green F, Lindsay RS, Macrury S, McKnight JA, Patrick AW, Thekkepat S, Gornik O, McKeigue PM, Colhoun HM: N-Glycan Profile and Kidney Disease in Type 1 Diabetes. Diabetes Care 2018;41:79-87

47. Trbojević Akmačić I, Ventham NT, Theodoratou E, Vučković F, Kennedy NA, Krištić J, Nimmo ER, Kalla R, Drummond H, Štambuk J, Dunlop MG, Novokmet M, Aulchenko Y, Gornik O, Campbell H, Pučić Baković M, Satsangi J, Lauc G: Inflammatory bowel disease associates with proinflammatory potential of the immunoglobulin G glycome. Inflammatory bowel diseases 2015;21:1237-1247

48. Boachie AM, Degraff D, Yu X, Sun Q, Friedman D, Gronostajski R, Matusik R: Abstract 1231: Nuclear Factor I family members interact with FoxA1 to regulate androgen responsive promoters. Cancer Research 2010;70:1231-1231

49. Zhao E, Li Y, Fu X, Zeng L, Zeng H, Jin W, Chen J, Yin G, Qian J, Ying K, Xie Y, Zhao RC, Mao Y: Cloning and characterization of human synaptotagmin 10 gene. DNA sequence : the journal of DNA sequencing and mapping 2003;14:393-398

50. Burda P, Aebi M: The ALG10 locus of Saccharomyces cerevisiae encodes the alpha-1,2 glucosyltransferase of the endoplasmic reticulum: the terminal glucose of the lipid-linked oligosaccharide is required for efficient N-linked glycosylation. Glycobiology 1998;8:455-462

51. Tachmazidou I, Süveges D, Min JL, Ritchie GRS, Steinberg J, Walter K, Iotchkova V, Schwartzentruber J, Huang J, Memari Y, McCarthy S, Crawford AA, Bombieri C, Cocca M, Farmaki AE, Gaunt TR, Jousilahti P, Kooijman MN, Lehne B, Malerba G, Männistö S, Matchan A, Medina-Gomez C, Metrustry SJ, Nag A, Ntalla I, Paternoster L, Rayner NW, Sala C, Scott WR, Shihab HA, Southam L, St Pourcain B, Traglia M, Trajanoska K, Zaza G, Zhang W, Artigas MS, Bansal N, Benn M, Chen Z, Danecek P, Lin WY, Locke A, Luan J, Manning AK, Mulas A, Sidore C, Tybjaerg-Hansen A, Varbo A, Zoledziewska M, Finan C, Hatzikotoulas K, Hendricks AE, Kemp JP, Moayyeri A, Panoutsopoulou K, Szpak M, Wilson SG, Boehnke M, Cucca F, Di Angelantonio E, Langenberg C, Lindgren C, McCarthy MI, Morris AP, Nordestgaard BG, Scott RA, Tobin MD, Wareham NJ, Burton P, Chambers JC, Smith GD, Dedoussis G, Felix JF, Franco OH, Gambaro G, Gasparini P, Hammond CJ, Hofman A, Jaddoe VWV, Kleber M, Kooner JS, Perola M, Relton C, Ring SM, Rivadeneira F, Salomaa V, Spector TD, Stegle O, Toniolo D, Uitterlinden AG, Barroso I, Greenwood CMT, Perry JRB, Walker BR, Butterworth AS, Xue Y, Durbin R, Small KS, Soranzo N, Timpson NJ, Zeggini E: Whole-Genome Sequencing Coupled to Imputation Discovers Genetic Signals for Anthropometric Traits. Am J Hum Genet 2017;100:865-884

52. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL: Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am J Hum Genet 2019;104:65-75

53. Kinoshita T, Doi K, Sugiyama H, Kinoshita S, Wada M, Naruto S, Tomonaga A: Knowledge-Based Identification of the ERK2/STAT3 Signal Pathway as a Therapeutic Target for Type 2 Diabetes and Drug Discovery. Chemical Biology & Drug Design 2011;78:471-476

54. Rudman N, Gornik O, Lauc G: Altered N-glycosylation profiles as potential biomarkers and drug targets in diabetes. FEBS Letters 2019;593:1598-1615

55. Wilkin TJ: The accelerator hypothesis: weight gain as the missing link between Type I and Type II diabetes. Diabetologia 2001;44:914-922

56. Kibirige M, Metcalf B, Renuka R, Wilkin T: Testing the accelerator hypothesis: the relationship between body mass and age at diagnosis of type 1 diabetes. Diabetes care 2003;26:2865-2870

57. Volta F, Gerdes JM: The role of primary cilia in obesity and diabetes. Annals of the New York Academy of Sciences 2017;1391:71-84

Table 1 The AUCs of different cutoffs of T1D association P-values

a. First cohort	
P value*	**AUC****
≤1.00E-10	0.8462
≤1.00E-09	0.8487
≤1.00E-08	0.8518
≤1.00E-07	0.8565
≤1.00E-06	0.8604
≤1.00E-05	0.8607
≤1.00E-04	0.8590
≤0.001	0.8561
≤0.01	0.8546
≤0.05	0.8502
≤0.1	0.8508
≤0.2	0.8530
≤0.5	0.8563
≤1	0.8579
b. Switched cohort	
P value*	**AUC****
≤1.00E-10	0.8576
≤1.00E-09	0.8589
≤1.00E-08	0.8588
≤1.00E-07	0.8609
≤1.00E-06	0.8633
≤1.00E-05	0.8654
≤1.00E-04	0.8618
≤0.001	0.8555
≤0.01	0.8470
≤0.05	0.8441
≤0.1	0.8446
≤0.2	0.8467
≤0.5	0.8521
≤1	0.8533

* The P values are based on the statistics of the PRS training cohort;

** The AUCs are the PRS performances in the independent testing cohort.

Table 2 Leading SNPs at three loci have been established of T1D association

a. First cohort						*Low PRS cases vs Low PRS controls*							*All cases vs all controls in the test cohort*					OR heterogeneity P
CHR	SNP	BP	Gene.ref	GA1	Quality Sc	MAF	*n*	OR	L95	U95	P	MAF	*n*	OR	L95	U95	P	
1	rs2476601	114,377,568	PTPN22	A	0.991604	0.08694	6168	1.859	1.586	2.179	1.96E-14	0.1117	9558	2.237	2.044	2.449	3.73E-68	0.047
6	rs9273368	32,626,475	HLA-DQB1	A	0.774174	0.2585	6168	3.581	3.189	4.022	5.83E-103	0.3849	9558	4.972	4.622	5.349	*<1E-350*	2.72E-06
11	rs689	2,182,224	INS	A	0.936309	0.2572	6168	0.3855	0.3304	0.4499	1.09E-33	0.2328	9558	0.4463	0.4119	0.4837	3.78E-86	0.099
b. Switched cohort						*Low PRS cases vs Low PRS controls*							*All cases vs all controls in the test cohort*					
CHR	SNP	BP	Gene.ref	GA1	Quality Sc	MAF	*n*	OR	L95	U95	P	MAF	*n*	OR	L95	U95	P	OR heterogeneity P
1	rs2476601	114,377,568	PTPN22	A	0.991604	0.09534	6503	2.239	1.943	2.581	8.22E-29	0.1154	9559	2.265	2.068	2.48	1.14E-69	0.893
6	rs9273368	32,626,475	HLA-DQB1	A	0.774174	0.259	6503	3.05	2.741	3.395	7.76E-93	0.3773	9559	4.768	4.438	5.122	*<1E-350*	1.04619E-11
11	rs689	2,182,224	INS	A	0.936309	0.2586	6503	0.4353	0.3788	0.5002	8.84E-32	0.2358	9559	0.4858	0.4492	0.5254	5.99E-73	0.178
12	rs1702877	56,427,808	IKZF4	T	0.989502	0.3279	6503	1.353	1.22	1.499	8.87E-09	0.3435	9559	1.371	1.288	1.46	5.78E-23	0.83

Table 3 Novel loci associated with low PRS T1D

a. First cohort					Low PRS cases vs Low PRS controls						All cases vs all controls in the test cohort						
CHR	SNP	BP	Gene.ref	CA1	MAF	n	OR	L95	U95	P	MAF	n	OR	L95	U95	P	Quality Sc
2	rs2222778	51,708,739	LOC73010(	A	0.007734	6168	3.395	2.225	5.182	1.45E-08	0.006035	9558	1.472	1.024	2.116	0.0366	0.979255
2	rs28957422	51,710,022	LOC73010(	T	0.007734	6168	3.395	2.225	5.182	1.45E-08	0.006035	9558	1.472	1.024	2.116	0.0366	0.975325
2	rs75496629	51,722,547	LOC73010(	G	0.008385	6168	3.404	2.275	5.095	2.60E-09	0.006628	9558	1.606	1.14	2.261	0.00674	0.974373
2	rs114137458	51,723,592	LOC73010(	A	0.008385	6168	3.404	2.275	5.095	2.60E-09	0.006628	9558	1.606	1.14	2.261	0.00674	0.977535
2	rs146290742	51,725,134	LOC73010(	A	0.007734	6168	3.41	2.246	5.178	8.58E-09	0.006035	9558	1.545	1.078	2.214	0.01792	0.962422
2	rs28957070	51,726,189	LOC73010(	T	0.007571	6168	3.512	2.306	5.35	4.86E-09	0.005927	9558	1.593	1.109	2.289	0.01184	0.97098
2	rs148736250	51,727,908	LOC73010(	T	0.007408	6168	3.638	2.38	5.56	2.42E-09	0.00582	9558	1.641	1.139	2.364	0.007909	0.965001
2	rs115578007	51,728,692	LOC73010(	A	0.009606	6168	3.064	2.092	4.488	8.93E-09	0.007759	9558	1.489	1.084	2.045	0.014	0.905771
2	rs57623361	51,731,591	LOC73010(	T	0.008222	6168	3.358	2.235	5.046	5.56E-09	0.00652	9558	1.601	1.134	2.261	0.007513	0.974803
2	rs28957081	51,732,492	LOC73010(	G	0.008385	6168	3.404	2.275	5.095	2.60E-09	0.006628	9558	1.606	1.14	2.261	0.00674	0.973265
2	rs28958299	51,740,567	LOC73010(	A	0.007408	6168	3.556	2.331	5.425	3.95E-09	0.00582	9558	1.621	1.127	2.334	0.009278	0.978082
2	rs28957085	51,747,733	LOC73010(	T	0.007408	6168	3.556	2.331	5.426	3.94E-09	0.00582	9558	1.622	1.127	2.335	0.009201	0.982582
2	rs1528792	51,753,028	LOC73010(	G	0.008141	6168	3.392	2.263	5.085	3.32E-09	0.006358	9558	1.583	1.119	2.239	0.009496	0.990323
2	rs28957087	51,755,192	LOC73010(	C	0.008385	6168	3.567	2.388	5.327	5.20E-10	0.00652	9558	1.685	1.194	2.377	0.003	0.950904
2	rs1406418	51,755,731	LOC73010(	C	0.008385	6168	3.42	2.287	5.114	2.10E-09	0.00652	9558	1.634	1.158	2.306	0.00523	0.948292
2	rs28958318	51,757,154	LOC73010(	G	0.008141	6168	3.199	2.126	4.814	2.48E-08	0.006358	9558	1.554	1.097	2.202	0.01305	0.939698
2	rs28957091	51,758,251	LOC73010(	A	0.007489	6168	3.332	2.184	5.082	2.31E-08	0.005873	9558	1.575	1.097	2.261	0.0138	0.932196
4	rs4833044	38,417,429	TBC1D1;LI	A	0.003582	6168	5.354	2.962	9.68	2.80E-08	0.002748	9558	2.013	1.187	3.412	0.009411	0.926446
9	rs12685465	14,228,376	NFIB	T	0.004803	6168	4.45	2.671	7.413	9.83E-09	0.003664	9558	1.614	1.016	2.564	0.04245	0.919389
9	rs10961435	14,229,049	NFIB	C	0.004722	6168	4.568	2.733	7.637	6.87E-09	0.003664	9558	1.706	1.075	2.709	0.02348	0.924563
10	rs746298	44,756,039	LINC00841	T	0.005617	6168	4.311	2.622	7.085	8.31E-09	0.004365	9558	2.275	1.467	3.53	0.000243	0.972703
b. Switched cohort					Low PRS cases vs Low PRS controls						All cases vs all controls in the test cohort						
CHR	dbSNP	BP	Gene.ref	CA1	MAF	n	OR	L95	U95	P	MAF	n	OR	L95	U95	P	Quality Sc
2	rs186500234	351,860	LINC0186	A	0.01083	6503	2.843	1.977	4.088	1.74E-08	0.01108	9559	1.678	1.277	2.203	0.000199	0.76609
2	rs77155228	62,636,361	B3GNT2;TI	G	0.00317	6503	5.962	3.213	11.06	1.51E-08	0.002528	9559	2.005	1.129	3.561	0.01764	0.717752
2	rs75233229	62,660,029	B3GNT2;TI	A	0.004098	6503	4.814	2.801	8.271	1.27E-08	0.003442	9559	1.985	1.222	3.226	0.005617	0.845001
2	rs75634056	62,660,518	B3GNT2;TI	C	0.004021	6503	4.979	2.881	8.603	8.83E-09	0.003388	9559	2.05	1.256	3.347	0.004086	0.847746
2	rs76505469	62,673,171	B3GNT2;TI	T	0.004176	6503	4.742	2.76	8.15	1.75E-08	0.003334	9559	1.794	1.104	2.917	0.01837	0.887349
2	rs17040236	70,563,856	FAM136A;	A	0.002474	6503	8.063	3.893	16.7	1.91E-08	0.002044	9559	2.107	1.103	4.026	0.02405	0.909783
2	rs57971004	70,563,888	FAM136A;	A	0.002474	6503	8.063	3.893	16.7	1.91E-08	0.002044	9559	2.107	1.103	4.026	0.02405	0.903266
2	rs1382458	70,573,576	FAM136A;	T	0.004176	6503	5.112	2.865	9.121	3.33E-08	0.003765	9559	2.2	1.333	3.633	0.002056	0.923724
2	rs116533147	70,575,312	FAM136A;	G	0.004176	6503	5.12	2.869	9.136	3.25E-08	0.003603	9559	1.971	1.18	3.291	0.009479	0.902899
2	rs77418738	70,578,049	FAM136A;	T	0.002706	6503	7.634	3.792	15.37	1.25E-08	0.002151	9559	1.892	1.006	3.558	0.04796	0.87356
2	rs75502807	70,578,550	FAM136A;	C	0.004176	6503	5.12	2.869	9.136	3.25E-08	0.003711	9559	2.134	1.287	3.539	0.00329	0.884187
2	rs116081627	70,579,366	FAM136A;	G	0.002552	6503	7.668	3.727	15.78	3.14E-08	0.002044	9559	1.886	0.9861	3.608	0.05516	0.872094
2	rs11123695	109,082,052	GCC2	T	0.01175	6503	2.56	1.832	3.579	3.73E-08	0.009358	9559	1.093	0.8102	1.475	0.56	0.807184
2	rs3827760	109,513,601	EDAR	G	0.01307	6503	2.537	1.836	3.505	1.69E-08	0.01043	9559	1.104	0.8308	1.468	0.4943	0.847028
2	rs922452	109,543,883	EDAR	T	0.01461	6503	2.483	1.822	3.383	8.38E-09	0.01183	9559	1.117	0.8548	1.46	0.4171	0.876009
4	rs6842426	25,812,477	SEL1L3	A	0.003325	6503	5.618	3.055	10.33	2.81E-08	0.002474	9559	1.991	1.121	3.537	0.01878	0.693839
4	rs72615957	142,499,563	LINC0243	A	0.00317	6503	5.651	3.032	10.53	4.98E-08	0.002474	9559	2.513	1.402	4.504	0.001972	0.896354
4	rs9790756	142,501,470	LINC0243	T	0.004253	6503	4.809	2.79	8.289	1.58E-08	0.003442	9559	2.121	1.285	3.501	0.003267	0.916406
6	rs10046450	18,338,709	DEK;RNF1	A	0.03704	6503	1.84	1.48	2.287	4.00E-08	0.03576	9559	1.191	1.018	1.392	0.02874	0.985992
6	rs72830389	18,344,198	DEK;RNF1	T	0.03526	6503	1.864	1.495	2.325	3.28E-08	0.03415	9559	1.206	1.029	1.414	0.02057	0.984113
6	rs16880565	18,348,630	DEK;RNF1	G	0.03619	6503	1.867	1.501	2.321	1.98E-08	0.03496	9559	1.216	1.039	1.423	0.01503	0.977553
6	rs80292134	153,424,759	RGS17	A	0.002552	6503	7.115	3.606	14.04	1.52E-08	0.001936	9559	2.265	1.182	4.343	0.01378	0.866733
6	rs77992292	153,426,323	RGS17	G	0.002552	6503	7.115	3.606	14.04	1.52E-08	0.001936	9559	2.265	1.182	4.343	0.01378	0.862918
6	rs3734776	170,592,945	DLL1	T	0.007191	6503	3.289	2.149	5.035	4.23E-08	0.006292	9559	1.743	1.219	2.492	0.00232	0.891623
6	rs3800237	170,596,266	DLL1	A	0.009434	6503	3.084	2.126	4.475	3.01E-09	0.008175	9559	1.664	1.218	2.274	0.001395	0.882698
6	rs76430845	170,694,803	FAM120B	T	0.007037	6503	3.451	2.229	5.342	2.78E-08	0.006131	9559	1.793	1.243	2.587	0.001785	0.985089
7	rs77713312	37,915,974	NME8	T	0.003866	6503	5.088	2.878	8.994	2.18E-08	0.002958	9559	1.77	1.038	3.017	0.03593	0.910206
7	rs3778716	37,916,799	NME8	G	0.003866	6503	5.088	2.878	8.994	2.18E-08	0.002958	9559	1.77	1.038	3.017	0.03593	0.908477
7	rs12532321	37,922,589	NME8	A	0.003944	6503	5.317	3.029	9.334	5.92E-09	0.003012	9559	1.837	1.083	3.114	0.02407	0.904163
7	rs78142343	37,928,948	NME8	G	0.003944	6503	5.317	3.029	9.334	5.92E-09	0.003012	9559	1.837	1.083	3.114	0.02407	0.911494
7	rs74721191	37,930,404	NME8	G	0.003944	6503	5.317	3.029	9.334	5.92E-09	0.003012	9559	1.837	1.083	3.114	0.02407	0.897322
7	rs2100250	37,931,481	NME8	G	0.003944	6503	5.317	3.029	9.334	5.92E-09	0.003012	9559	1.837	1.083	3.114	0.02407	0.901506
7	rs35928775	71,446,334	CALN1	A	0.005181	6503	3.993	2.435	6.547	4.09E-08	0.004034	9559	1.44	0.9136	2.271	0.1162	0.896408
7	rs118182411	71,475,692	CALN1	C	0.004949	6503	4.252	2.554	7.08	2.64E-08	0.003818	9559	1.528	0.9524	2.452	0.0788	0.878019
7	rs76060515	88,938,393	ZNF804B	T	0.007191	6503	3.382	2.232	5.126	9.17E-09	0.005755	9559	1.295	0.8917	1.882	0.1743	0.930748
7	rs77205087	88,948,091	ZNF804B	T	0.007269	6503	3.501	2.315	5.296	2.95E-09	0.005755	9559	1.35	0.9298	1.959	0.1148	0.917186
9	rs11559334	101,117,596	GABBR2	T	0.002784	6503	6.58	3.395	12.75	2.40E-08	0.002205	9559	2.445	1.335	4.478	0.003793	0.856568
12	rs10506114	33,829,465	SYT10;ALC	T	0.009666	6503	2.77	1.934	3.967	2.74E-08	0.007691	9559	1.127	0.8161	1.555	0.4686	0.986083
12	rs4142676	33,833,469	SYT10;ALC	A	0.009743	6503	2.742	1.916	3.924	3.51E-08	0.007744	9559	1.115	0.8082	1.538	0.5073	0.98933
12	rs11052843	33,845,949	SYT10;ALC	C	0.009743	6503	2.742	1.916	3.924	3.51E-08	0.007744	9559	1.115	0.8082	1.538	0.5073	0.993611
12	rs12228218	33,848,259	SYT10;ALC	A	0.009743	6503	2.742	1.916	3.924	3.51E-08	0.007744	9559	1.115	0.8082	1.538	0.5073	0.993278
12	rs11052847	33,848,927	SYT10;ALC	T	0.009743	6503	2.742	1.916	3.924	3.51E-08	0.007744	9559	1.115	0.8082	1.538	0.5073	0.993816
12	rs2087269	33,850,143	SYT10;ALC	C	0.009743	6503	2.742	1.916	3.924	3.51E-08	0.007744	9559	1.115	0.8082	1.538	0.5073	0.993816
12	rs11052850	33,850,615	SYT10;ALC	C	0.009743	6503	2.742	1.916	3.924	3.51E-08	0.007744	9559	1.115	0.8082	1.538	0.5073	0.993795
12	rs1352395	33,850,931	SYT10;ALC	A	0.009743	6503	2.742	1.916	3.924	3.51E-08	0.007744	9559	1.115	0.8082	1.538	0.5073	0.989424
12	rs11052881	33,889,887	SYT10;ALC	T	0.009511	6503	2.778	1.928	4.003	4.14E-08	0.007691	9559	1.128	0.8152	1.56	0.4681	0.981571
12	rs75438334	105,130,915	CHST11	T	0.003402	6503	5.828	3.194	10.63	9.15E-09	0.00285	9559	2.405	1.414	4.092	0.001211	0.888074
12	rs78308059	133,456,478	CHFR	G	0.006573	6503	3.605	2.297	5.657	2.46E-08	0.005378	9559	1.494	0.9952	2.243	0.05277	0.902349
12	rs12230138	133,474,618	LOC10192	C	0.006186	6503	3.933	2.492	6.207	4.09E-09	0.004786	9559	1.387	0.9105	2.113	0.1277	0.905817
12	rs11147161	133,491,783	LOC10192	A	0.005877	6503	3.894	2.433	6.232	1.47E-08	0.004571	9559	1.469	0.9542	2.261	0.08064	0.917771
15	rs2197053	90,143,033	TICRR	G	0.004253	6503	4.702	2.75	8.039	1.55E-08	0.003872	9559	2.105	1.326	3.339	0.001584	0.791548
21	rs16997642	29,578,961	LINC0169	C	0.006341	6503	3.56	2.276	5.569	2.65E-08	0.006023	9559	2.03	1.401	2.942	0.000184	0.921
21	rs2831662	29,594,709	LINC0169	C	0.006418	6503	3.487	2.234	5.445	3.91E-08	0.00597	9559	1.827	1.264	2.642	0.001343	0.926048
21	rs73897628	29,597,973	LINC0169	T	0.005877	6503	3.763	2.365	5.987	2.25E-08	0.005324	9559	1.92	1.299	2.837	0.00107	0.904896
21	rs73897685	29,698,384	LINC0169	G	0.007269	6503	3.329	2.173	5.101	3.29E-08	0.006346	9559	1.699	1.182	2.441	0.00419	0.970546
21	rs7278151	29,724,671	LINC0169	T	0.01114	6503	2.904	2.016	4.182	1.02E-08	0.009358	9559	1.37	1.006	1.865	0.0459	0.90377
21	rs145901638	29,801,999	LINC0169	A	0.005181	6503	3.999	2.446	6.538	3.29E-08	0.004302	9559	1.87	1.217	2.873	0.004312	0.791071

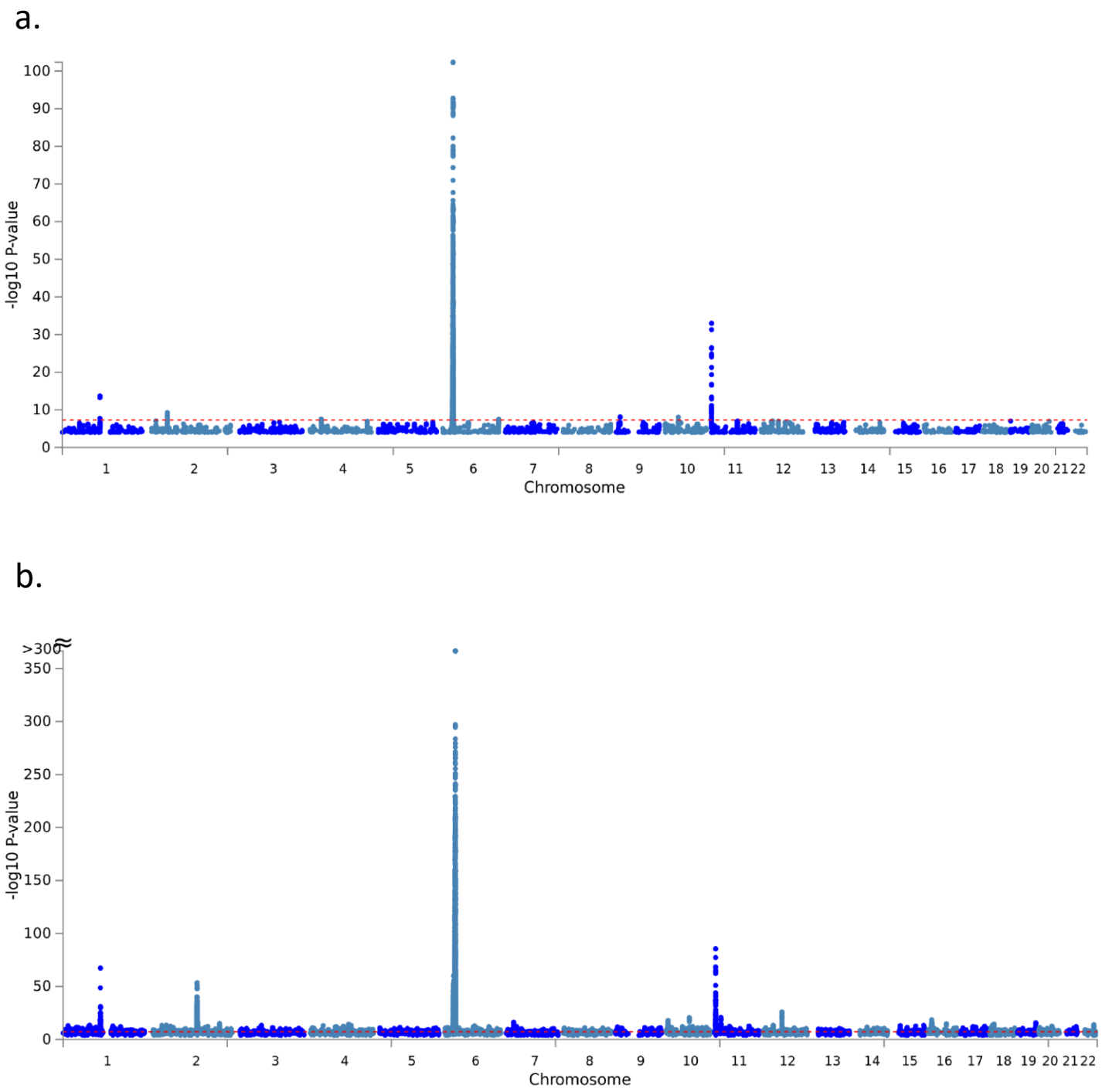
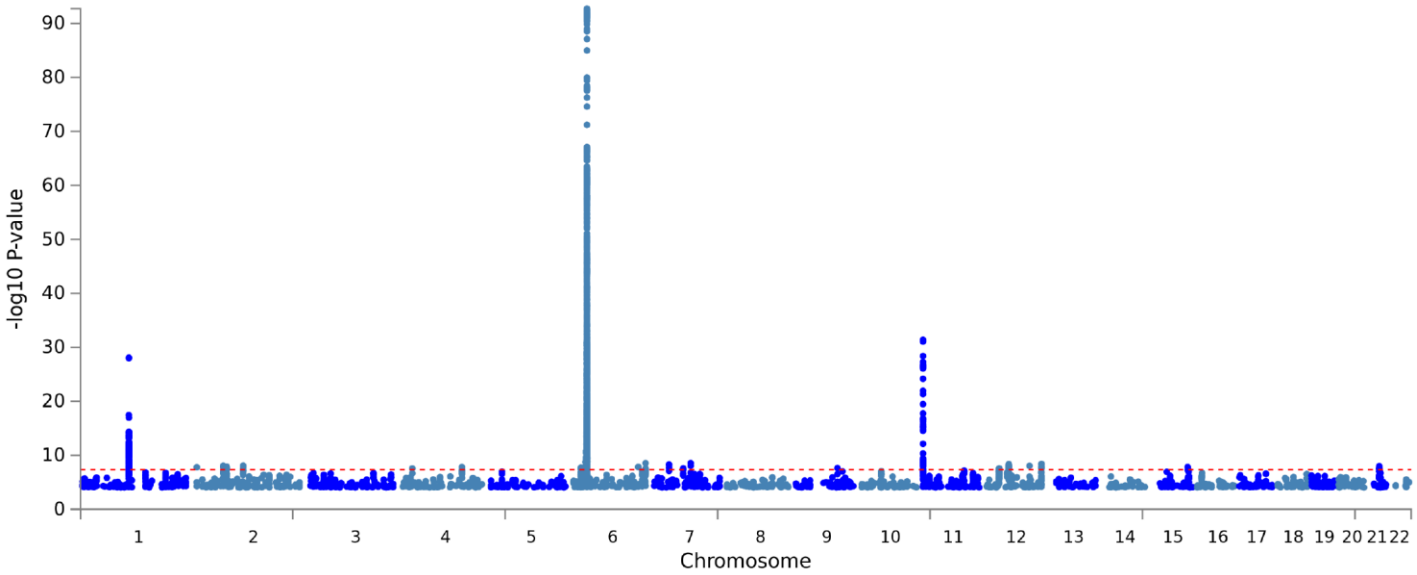


Figure 1. The Manhattan plots of the first cohort. (a) The plot of the GWAS of T1D patients with low T1D PRS compared to controls with low T1D PRS (810 cases vs. 5358 controls); (b) The plot of the GWAS of all T1D patients compared to all controls (3355 cases vs. 6203 controls).

a.



b.

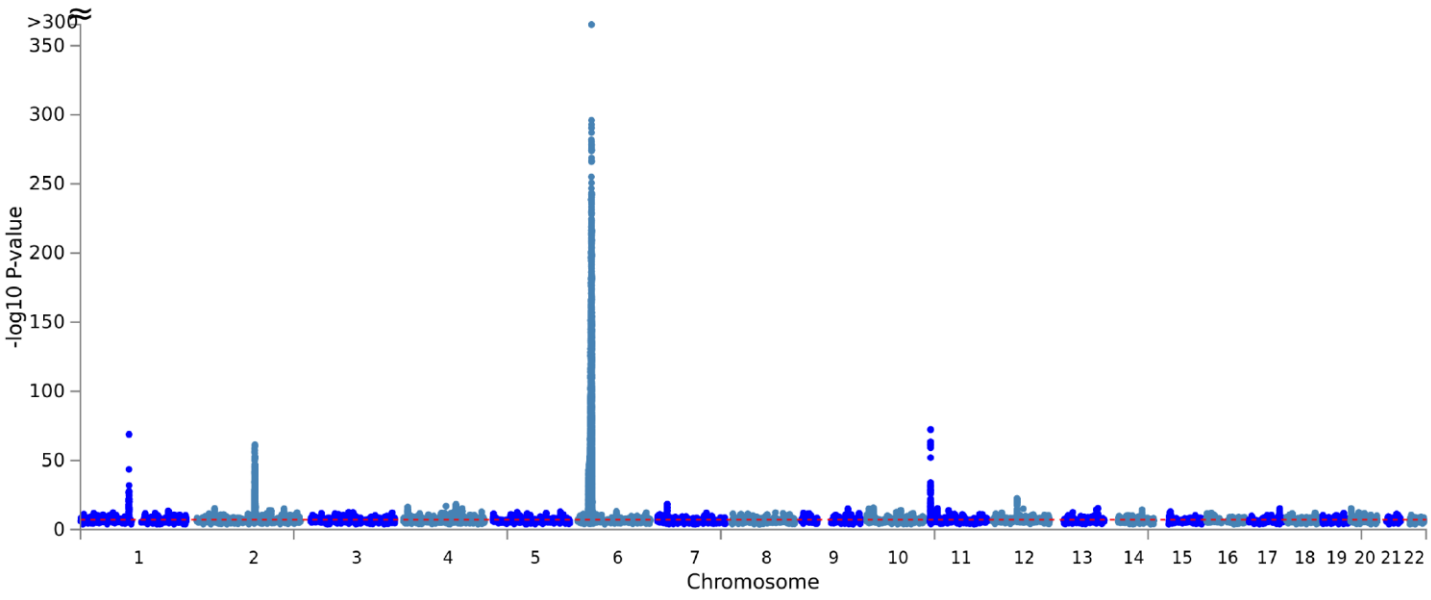


Figure 2. The Manhattan plots of the second cohort. (a) The plot of the GWAS of T1D patients with low T1D PRS compared to controls with low T1D PRS (918 cases vs. 5585 controls); (b) The plot of the GWAS of all T1D patients compared to all controls (3356 cases vs. 6203 controls).