



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.



THE UNIVERSITY *of* EDINBURGH

**Identifying core genes, proteins and
possible drug targets for type 1 diabetes
with novel statistical genetics methods
and large-scale omics data**

Xuan Zhou

Thesis submitted for the degree of

Doctor of Philosophy

The University of Edinburgh

2025

Declaration

I, Xuan Zhou, hereby declare that the thesis has been composed by myself. All the research work in this thesis was conducted on my own, except where input from others was included as part of a co-authored scientific paper, which has been published. In this case, a detailed contribution has been clearly indicated in **Chapter 3**. This thesis has not been submitted for any other degree or professional qualification.

Signature:

Date: March 2, 2026

Publications

The following publication is included in the **Chapter 3** in this thesis.

- **Xuan Zhou** [#], Andrii Iakovliev [#], Stuart McGurnaghan, Buddhiprabha Erabadda, Caroline Hayward, Paul M. McKeigue, Athina Spiliopoulou, Helen M. Colhoun; Genome-Wide Aggregated Trans Effects Analysis for Circulating Proteins Indicates a Key Role of Immune Checkpoints in Type 1 Diabetes. *Diabetes* 2025;74(10):1873–1884. ([#]Joint first author)

Other publications during the PhD programme:

- **Xuan Zhou**, Nan Yang, Wei Xu, Xue Li, Athina Spiliopoulou, Evropi Theodoratou; Associations of genetic factors with vascular diabetes complications: an umbrella review. *Journal of Global Health* 2025;15:04081.
- **Xuan Zhou**, Qian Xiao, Fangyuan Jiang, Jing Sun, Lijuan Wang, Lili Yu, Yajing Zhou, Jianhui Zhao, Han Zhang, Shuai Yuan, Maria Timofeeva, Athina Spiliopoulou, Ines Mesa-Eguiagaray, Susan M. Farrington, Philip J. Law, Richard S. Houlston, Kefeng Ding, Malcolm G. Dunlop, Evropi Theodoratou, Xue Li; Dissecting the pathogenic effects of smoking and its hallmarks in blood DNA methylation on colorectal cancer risk. *British Journal of Cancer* 2023;129:1306–1313.

Conference presentations during the PhD programme:

- Associations of genetic factors with diabetes complications: an umbrella review. European Human Genetics Conference, 2024. Poster presentation.
- Discovery of core genes for type 1 diabetes via genome-wide aggregation of *trans*-effects using proteomics studies. American Society of Human Genetics Annual Meeting, 2024. Poster presentation.
- Genome-wide aggregation of *trans*-effects from proteomic studies indicates a key role of immune checkpoints in type 1 diabetes. Alliance for Complex Trait Genetics 20th Meeting, 2025. Oral presentation.

- Plasma protein biomarkers of type 1 diabetes (T1D) and genetic evidence for their causality. European Human Genetics Conference, 2025. Poster presentation.
- Plasma protein profiles of type 1 diabetes in the UK Biobank Pharma Proteomics Project (UKB-PPP). American Diabetes Association 85th Scientific Sessions, 2025. Poster presentation.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, **Dr Athina Spiliopoulou**, **Prof Evropi Theodoratou**, and **Prof Paul McK-eigue** for their unwavering support, insightful guidance, and constant encouragement throughout the course of my PhD. A special thanks goes to **Prof Helen Colhoun**, who provided invaluable advice and support on the GATE project and enormous help with the logical structure of this thesis. It has been a privilege to work under their mentorship; their expertise and scholarly attitude have been a constant source of inspiration.

I would also like to extend my heartfelt thanks to my colleagues. In particular, I would like to thank **Dr Andrii Iakovliev** and **Dr Buddhiprabha Erabadda** for hands-on help with the analysis pipeline; **Dr Luke Blackburn** and **Dr Stuart McGurnaghan** for managing the data and server; **Andrew Wilson** for his administrative help. My gratitude also goes to all the other members in the Diabetes Medical Informatics and Epidemiology group for fostering such a friendly and collaborative working environment. I also appreciate the contributions of **Dr Wei Xu** and **Nan Yang** to the umbrella review project, although it is not included in this thesis.

I am profoundly grateful to the **University of Edinburgh** and **Dr Athina Spiliopoulou** for funding my PhD through the Chancellor's Fellow studentship.

Finally, my deepest thanks go to my family and friends in Edinburgh, Florida, and back home in China, for their support and encouragement along the way. A special tribute to my parents, my sister, and my partner, for their boundless love, care, understanding, and support.

Abstract

Background

Type 1 diabetes (T1D) is an autoimmune disease in which the body's immune system destroys insulin-producing β cells in the pancreas. 9.5 million people are living with clinically diagnosed T1D globally in 2025, and this figure is projected to increase to 14.7 million in 2040. People with T1D require lifelong insulin therapy and glucose monitoring and have an increased risk for both acute and chronic complications. There is currently no cure for T1D. Only one agent—teplizumab, an anti-CD3 antibody—has been licensed for delaying progression from islet autoimmunity to clinical T1D but has a median delay of 32.5 months. Stem cell therapy for T1D has shown some promise, but it is still at early-stage clinical trials with tiny sample sizes; its long-term effects in a wider population remain unknown.

Human genetic evidence in support of a causal role for a drug target in a given indication increases the probability of success in clinical trials by 2.6-fold. T1D has a strong genetic predisposition, with a sibling recurrence risk ratio estimated between 15 and 20. Around half of the total genetic component is attributable to variants in the human leukocyte antigen (HLA) region. Genome-wide association studies (GWAS) have identified around a hundred genomic regions associated with T1D outside the HLA region, which collectively explain another 25% of the genetic component. However, the translation of GWAS findings to drug targets has been limited due to the complexity of mapping GWAS hits to the causal genes and functional mechanisms mediating the genetic effects.

The aim of this thesis was to leverage a novel framework of genetic analysis together with available large-scale omics data to identify core genes proximal to T1D pathogenesis and thus to provide human genetic evidence for prioritising drug targets for T1D.

Methods

To detect core genes and proteins for T1D, I first performed genome-wide aggregated trans-effects (GATE) analysis in two T1D datasets: a Scottish dataset and UK Biobank. For each individual in these two T1D datasets, I computed genotypically predicted protein levels by taking a weighted sum of their genotypes over *trans*-protein quantitative trait loci (*trans*-pQTL). Summary statistics for *trans*-pQTL were extracted from three large proteomics studies: the UK Biobank Pharma Proteomics Project (UKB-PPP), deCODE, and INTERVAL. Genotypically predicted levels for each protein (GATE scores) were tested for association with T1D. Individuals in the UKB-PPP dataset were not included in the UK Biobank subset used to test associations with T1D. Putative core proteins were selected based on the strength of the association (p -value $< 10^{-6}$) and the number of effective *trans*-pQTLs (> 5) contributing to the GATE score.

While these stringent thresholds were applied to limit detection of false positive findings, they may also result in missing true effects that are smaller or noisier. To get a more comprehensive list of core proteins for T1D, I subsequently performed a second study adopting an alternative strategy. Using the plasma protein measurements from the UKB-PPP dataset in 190 T1D cases and 36870 controls without diabetes, I first tested the measured levels of 2919 proteins for association with T1D and then validated significantly associated proteins using their GATE scores in the Scottish dataset and UK Biobank. Proteins associated with T1D both at the measured level ($p < 10^{-6}$) and the genetically predicted level (p -value < 0.001 in either dataset) with a consistent direction of effect and a larger magnitude of effect size at the measured level were declared as potential core proteins.

In both studies, a series of additional, prespecified lines of evidence was evaluated to support or refute a causal role for each identified protein in T1D, including Mendelian randomisation (MR) analysis with *trans*-pQTLs as instruments, annotation of monogenic causes of autoimmune diabetes, *cis*-score association with T1D, annotation of known T1D GWAS hits, protein association with T1D, and review of reported evidence in experimental animal models and drug effects in humans.

Having identified core proteins for T1D from my two studies, I then performed a follow-up study to evaluate if the transcript and protein levels of these proteins were different in individuals who progressed to islet autoimmunity (IA) or

overt T1D compared with those who did not. I leveraged transcriptomic and proteomic data in the Environmental Determinants of Diabetes in the Young (TEDDY) study, which included 418 pairs of persistent IA and 114 pairs of clinical T1D in a nested case-control design with incidence density sampling. Average transcript (measured by microarray and RNA sequencing) and protein levels (measured by the Olink Target 96 Inflammation panel) before the age of one year were calculated from available measurements for each individual. Comparison of the measured activity of core genes/proteins between those who progressed to IA or T1D within the next six years of follow-up and those who did not was performed using conditional logistic regression.

Results

In total, 2864 scores with more than 5 *trans*-pQTLs corresponding to 2340 unique proteins were tested for association with T1D in the Scottish dataset and 3212 scores for 2580 unique proteins in UK Biobank. The GATE analysis identified 25 putative core proteins in the Scottish dataset ($p < 10^{-6}$), 12 of them replicated in the UK Biobank ($p < 0.001$). Two additional putative core proteins were identified in the UK Biobank ($p < 10^{-6}$), and both replicated in the Scottish dataset ($p < 0.001$). The strongest association was with *PDCD1* (log odds ratio = 0.21, $p = 6 \times 10^{-24}$). Of the 27 putative core proteins, 11 were supported by a dose-response relationship in the MR analysis; rare variants in *PDCD1* had been discovered to cause autoimmune diabetes; *CXCL9* was supported with a *cis*-score association of T1D; 4 are located within 200 kb from previously reported T1D GWAS hits for T1D; 11 were supported by protein associations; 9 were supported by experimental perturbation in mouse models; and 1 was supported by drug effects in humans. Functionally, four proteins (*PDCD1*, *LAG3*, *CD5*, and *TIGIT*) are immune checkpoint receptors, four (*CXCL9*, *CXCL10*, *CCL19*, *CXCL11*) are chemokines, two (*NCR1* and *KLRB1*) are receptors expressed on natural killer cells, and eight (*PRSS27*, *AGR2*, *FGF19*, *GCG*, *CCK*, *CFC1*, *TCN1*, *REG1B*) are relevant to pancreas and gut functions.

In the UKB-PPP dataset, 257 proteins were found to be associated with T1D ($p < 10^{-6}$) after adjusting for age, sex, ethnicity, kidney filtration, liver function, and medication use. Among these proteins, 255 had available GATE scores and were tested for association with T1D in the UK Biobank and Scottish dataset. Fourteen proteins were further supported by their GATE score association with

T1D in the UK Biobank and 13 in the Scottish dataset. Together, 24 unique proteins were identified as potential core proteins. Among these, *CFC1* and *KLRB1* were detected as putative core proteins in the original GATE analysis, while the remaining 22 proteins represented novel findings. Among the novel hits, *CD28*, *ADA2*, and *ADGRE2* were supported with a dose-response effect by MR analysis; *CD274* had been reported as a monogenic cause of autoimmune diabetes; *CD28* is located proximate to known GWAS hits; *NOTCH1*, *CD28*, *TNFRSF4*, *ITGB2*, *ALCAM*, and *CD274* were supported by experimental evidence in animal models; and *CD274* and *CD28* were supported by drug effects in humans. Functionally, *CD28* and *TNFRSF4* are co-stimulatory immune checkpoint receptors, *CD274* encodes a ligand to *PDCD1*, *PRSS2*, *PNLIPRP1*, and *MUC2* are gut and pancreatic proteins, and *NOTCH1* and its ligand *DLL1* are involved in both embryonic development and immune regulation.

In total, the previous two analyses identified 49 unique candidate proteins. Transcripts for all 49 proteins were available in the TEDDY study, while serum protein levels for only eight of them (*CXCL11*, *CXCL9*, *CCL19*, *CD5*, *CXCL10*, *FGF19*, *LTA*, *CD274*) were measured. With the microarray data, average transcript levels before the age of one for *PDCD1* and *CD48* were nominally associated with incident IA developed within the next six years of follow-up (N = 104 pairs), and those for *PDCH17* and *AGR2* were nominally associated with incident T1D (N = 54 pairs). Using the RNA-seq data, average transcript levels for six genes were associated with incident IA (N = 274 pairs), with only the association from *SIGLEC6* remaining statistically significant after multiple testing correction. Average serum protein levels before the age of one for *LTA* and *CCL19* levels were associated with incident IA (N = 307 pairs), and the *LTA* association remained significant after correction for multiple testing.

Discussion

This thesis presents a comprehensive investigation of core genes and proteins proximal to T1D pathogenesis and identification of plausible drug targets using novel statistical genetic approaches and multi-omics data. The findings highlight the primary roles of immune checkpoint signalling, reinforce the contribution of innate immunity, and support the potential involvement of the exocrine pancreas and gut in the pathogenesis of T1D.

By systematically assessing several lines of evidence in support of causality of each of the 49 putative core proteins in T1D and reviewing what is known about their safety and technical feasibility as possible drug targets, the 49 proteins were classified into four categories: (1) strong causal evidence with existing drugs in clinical pipelines (*PDCD1*, *CD274*, *LAG3*, *CD28*, and *TNFRSF4*); (2) strong causal evidence with agents at preclinical stages (*LGALS9*, *IDO1*, and *TIGIT*); (3) strong causal evidence but requiring further evaluation of technical feasibility (*CD5*, *CD48*, *NCR1*, *KLRB1*, *CXCL9*, *CXCL10*, *CXCL11*, *CCL19*, *FASLG*, *NOTCH1*, and *DLL1*); and (4) requiring further validation of causal relationships (including all the pancreatic and gut proteins and other immune proteins).

Studying the activity level of these proteins in prospective cohorts for progression to T1D will be important to strengthen the evidence of causality and to decipher the direction of effect on disease risk. The study performed in this thesis using the TEDDY cohort was constrained by the small number of proteins with available proteomic measurements and by low statistical power to detect signals due to the limited sample size and the technical noise in transcript and protein measurements. Additional prospective cohorts with proteomics data, such as the Global Platform for the Prevention of Autoimmune Diabetes (GPPAD) study, the All Babies in Southeast Sweden (ABIS) study, and the Innovative approaches to understanding and arresting type 1 diabetes (INNODIA) study, are useful resources in which to validate putative core genes and proteins for T1D in the future.

Overall, this work advances understanding of the genetic architecture of T1D, establishes a framework to identify core genes and proteins proximal to T1D pathogenesis, and provides human genetic evidence for prioritising drug targets for T1D.

Lay Summary

Type 1 diabetes (T1D) is a disease where the immune system mistakenly attacks and destroys the cells in the pancreas that produce insulin. More than 9 million people live with T1D worldwide, and this number is rising. People with T1D need to take insulin shots and monitor their blood sugar levels every day for their whole life. They also face a higher risk of various health issues. There is no cure for T1D yet, even though several drugs have been developed with the aim of delaying the need for insulin injections. So far, only one drug, teplizumab, has been approved for this purpose. A few successful cases of reversing T1D with a special cell transplant have been reported, but this treatment requires ongoing medication to prevent rejection of the transplant, and the long-term effects are unknown.

Genetic research shows that having a genetic link to a drug target increases the chances of success in trials. T1D has a strong genetic component. Siblings of T1D patients have a 6 to 16 times higher risk of getting T1D compared to the general population. About half of this genetic risk is linked to genes in a specific region in the human DNA, called the HLA region. Studies have also found about a hundred other DNA regions related to T1D, but turning these findings into drugs is challenging because it's difficult to pinpoint the exact genes and understand how they work.

The “omnigenic” hypothesis suggests that small changes in many genes in many DNA regions combine to affect a few core genes and their proteins that directly impact the disease. This thesis aimed to use a new method based on this idea, along with large-scale biological data, to find key proteins involved in T1D and help identify potential drug targets.

Specifically, I used two search strategies to find key proteins for T1D, and then I evaluated if the proteins I found were different between children who went on to develop T1D compared to children that did not develop T1D. My two search strategies were:

1. A genome-wide aggregated *trans*-effects (GATE) analysis, where I used DNA to predict the protein levels for 5130 proteins circulating in the blood and then tested if these predicted protein levels were different between 6,828 people with T1D and 416K people without any diabetes.
2. An exploratory analysis of directly measured protein levels of 2919 proteins circulating in blood, where I compared if the measured protein levels were different between 190 people with T1D and 36870 people without any diabetes. I then examined if the proteins I found were also supported by genetic evidence.

The GATE analysis found 27 proteins linked to T1D, including immune system proteins and those related to the pancreas and gut. Some of these proteins were confirmed by looking at earlier studies and experiments, though many weren't previously known to be linked to T1D. In the exploratory analysis, 257 proteins were linked to T1D, and 24 proteins were further supported genetically. In the follow-up study, molecular activity for certain immune and gut proteins was linked to early signs of T1D.

These findings provide new insights into how T1D develops, especially regarding immune system proteins. These findings also suggest potential drug targets, PD-1/PD-L1, LAG-3, CD28, and OX40, worth investigating in clinical trials of T1D.

Table of contents

Declaration	ii
Publications	iii
Acknowledgements	v
Abstract	vi
Lay Summary	xi
List of Figures	xvii
List of Tables	xxi
Glossary	xxiv
Abbreviations	xxiv
1 Introduction	1
1.1 Type 1 diabetes	2
1.2 Epidemiology of T1D	3
1.3 Genetic predisposition of T1D	5
1.3.1 HLA region	6
1.3.2 Insulin (<i>INS</i>) gene	7
1.3.3 <i>CTLA4</i>	8
1.3.4 <i>PTPN22</i>	8
1.3.5 <i>IL2RA</i> (<i>CD25</i>)	8
1.3.6 <i>IFIH1</i>	9
1.3.7 Large-scale GWAS of T1D	9
1.3.8 Rare variants associated with monogenic autoimmune di- abetes	11
1.4 Disease-modifying therapy of T1D	12

1.4.1	Immunotherapies targeting specific immune cells	13
1.4.2	Cytokine-directed interventions	17
1.4.3	Antigen-specific therapy	18
1.4.4	β cell survival	20
1.4.5	Others	20
1.4.6	Stem cell therapy	21
1.5	Current challenges and motivation	22
1.6	Aims and objectives	26
2	Methods	27
2.1	T1D target datasets	28
2.1.1	Scottish dataset	28
2.1.2	UK Biobank	29
2.2	Computation of genome-wide aggregated <i>trans</i> -effects (GATE) scores for proteins in target datasets	30
2.2.1	Constructing locus-specific vectors of weights of SNPs as- sociated with proteins	30
2.2.2	Computation of scores in target datasets	31
2.3	Testing association of GATE scores with T1D	32
2.3.1	Statistical model for test of association of GATE scores with T1D	32
2.3.2	Filtering criteria for reporting GATE associations with T1D	32
2.4	Evaluation of supportive evidence for causality of GATE score as- sociations with T1D	33
2.5	Ethical considerations	35
3	Identification of putative core proteins for T1D using the GATE anal- ysis	36
3.1	Introduction	37
3.2	Accepted manuscript	39
3.2.1	Abstract	39
3.2.2	Article Highlights	39
3.2.3	Introduction	40
3.2.4	Research Design and Methods	41
3.2.5	Results	45
3.2.6	Discussion	51
3.2.7	Conclusions	56
3.2.8	Supplementary Materials	58

3.3	Characteristics of the GATE analysis in the Scottish and UK Biobank datasets	77
3.4	Pleiotropy of master regulators	79
3.5	Technical challenges of protein assays and their impact on GATE analysis	83
3.6	Conclusion	87
4	Plasma proteins associated with T1D and their potential as drug targets	88
4.1	Introduction	89
4.2	Methods	90
4.2.1	Study population	90
4.2.2	Statistical analysis	90
4.3	Results	93
4.3.1	Characteristics of the UKB-PPP dataset	93
4.3.2	Proteins associated with T1D	95
4.3.3	Genetic validation of protein associations using GATE scores	96
4.3.4	Supportive evidence for causality and druggability	97
4.3.5	Association of potential core proteins with HbA1c and T1D duration	99
4.4	Discussion	99
4.4.1	Immune checkpoint receptor and ligand	100
4.4.2	Proteins involved in immune cell adhesion and trafficking	101
4.4.3	Neuroimmune proteins	102
4.4.4	Other immune proteins	102
4.4.5	Gut and pancreatic proteins	103
4.4.6	Notch signalling pathway	103
4.5	Supplementary materials	104
5	Prospective validation of proteins associated with T1D in high-risk individuals.	116
5.1	Introduction	117
5.2	Methods	117
5.2.1	Study population	117
5.2.2	Data cleaning and statistical analyses	120
5.3	Results	121
5.3.1	Characteristics of the TEDDY nested case-control study	121

5.3.2	Association of transcript and protein levels with incident IA and T1D	124
5.4	Discussion	127
5.4.1	Methodological limitations	127
5.4.2	Prospective datasets for the development of T1D	129
5.5	Supplementary materials	133
6	Discussion	150
6.1	Statement of overall aim	151
6.2	Statement of principal findings	152
6.3	Implications of key pathways in T1D pathogenesis and their therapeutic potential	153
6.3.1	Strong causal evidence with existing drugs in clinical pipelines	153
6.3.2	Strong causal evidence with agents at preclinical stages .	162
6.3.3	Strong causal evidence but requiring further evaluation of technical feasibility	164
6.3.4	Requiring further validation of causal relationships	168
6.4	Strengths and limitations	170
6.5	Future directions	173
6.6	Conclusion	175
7	Bibliography	176

List of Figures

1.1	Three distinct stages of T1D. Adapted from Insel et al., Diabetes Care 2015; 38: 1964–1974. © 2015 American Diabetes Association. Used under educational, non-profit use with proper citation. .	3
1.2	Geometric mean plasma C-peptide (pmol/l) by age at onset and duration in individuals with T1D. Log C-peptide levels were smoothed using locally estimated scatterplot smoothing (LOESS) regression (polynomial degree 2, span 0.25). In each panel, C-peptide level (pmol/l) is represented both by colour intensity and by its coordinate along the vertical axis. Reproduced from McKeigue et al., BMC Medicine 2019; 17(1):165. CC BY 4.0.	4
3.1	Graphic illustration of the study design. Left (input data): summary-level data from pQTL GWAS studies and individual-level genotypic and case/control data from target datasets. Right (analysis steps): pGATE scores are computed using pQTL summary statistics and individual-level genotypes. Computed scores are tested for association with the disease (type 1 diabetes). For identified putative core proteins, additional supporting evidence for causality of the encoding gene with the disease is evaluated.	41
3.2	Heatmap of correlations among pGATE scores for putative core proteins for type 1 diabetes ordered by hierarchical clustering. Correlations were computed using non-diabetic controls in the Scottish dataset. Two clusters of correlated scores are apparent: a cluster of moderately correlated pGATE scores for proteins encoded by genes known to be highly expressed in gut (top left corner) and a cluster of lower correlations among pGATE scores for immune-related proteins (bottom right corner).	48

- 3.3 Network plot of *trans*-pQTLs contributing to pGATE scores for putative core proteins for type 1 diabetes. *Trans*-pQTLs are grouped into clumps by genomic location. Putative core proteins are shown in spheres, grouped by functional role. The size of the sphere is proportional to the effect of the pGATE score on type 1 diabetes risk. *Trans*-pQTL clumps are represented by stars if a known type 1 diabetes GWAS hit is located within 200 kb from the clump and circles if not. *Trans*-pQTL clumps located within 200 kb from a known type 1 diabetes GWAS hit and contributing to pGATE scores of more than 5 core proteins (i.e., ‘master regulators’) are labelled by the gene name the GWAS hit was assigned to. All other clumps are labelled by the clump number as shown in Table S3.2. The size of the star/circle is proportional to the number of target proteins. 49
- S3.1 Mendelian randomisation analysis of the effect of the putative core proteins on type 1 diabetes in the Scottish dataset. All *trans*-pQTLs contributing to the pGATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and type 1 diabetes are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate. 59
- S3.2 Mendelian randomisation analysis of the effect of the putative core proteins on type 1 diabetes in the Scottish dataset, except for *CD48* and *LAG3*, which are represented by the UK Biobank data due to a higher number of *trans*-pQTLs for these two putative core proteins in this dataset. All *trans*-pQTLs contributing to the pGATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and type 1 diabetes are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate. 60

S3.3	Manhattan plot of GATE scores with at least five effective <i>trans</i> -loci in the Scottish dataset.	78
S3.4	Manhattan plot of GATE scores with at least five effective <i>trans</i> -loci in the UK Biobank dataset.	79
S3.5	<i>Trans</i> -pQTLs that are also T1D GWAS hits contributing to the GATE scores of the putative core proteins. <i>Trans</i> -pQTLs overlapping T1D GWAS hits are listed on the x-axis and the putative core proteins are shown on the y-axis. Each dot represents a <i>trans</i> -pQTL; colour indicates the direction of effect of the <i>trans</i> -pQTL on T1D risk (red, increase risk; green, decrease risk) and point size reflects the effect size of the <i>trans</i> -pQTL on circulating protein levels.	80
S3.6	Heatmap of raw correlations among GATE scores for 19 putative core proteins with <i>trans</i> -pQTLs in or nearby the <i>SH2B3</i> gene. Correlations were computed using non-diabetic controls in the Scottish dataset.	83
S3.7	Heatmap of partial correlations among GATE scores for 19 putative core proteins with <i>trans</i> -pQTLs in or nearby the <i>SH2B3</i> gene, conditioning on the genotype of rs3184504*T. Correlations were computed using non-diabetic controls in the Scottish dataset. . .	84
4.1	Flowchart of the study design. The upper panel outlines the association analyses between protein levels and T1D case-control status, HbA1c, and T1D duration in the UKB-PPP dataset. Models with basic adjustment included age at assessment, sex, and ethnicity as covariates, and fully adjusted models additionally accounted for kidney function, liver function, and medication use. For the analysis involving T1D duration, protein levels were specified as the dependent variable, duration as the primary independent variable, and age at onset, sex, ethnicity, kidney function, liver function, and medication use as covariates. The lower panel shows the genetic validation of protein associations using GATE scores in UK Biobank and the Scottish dataset, followed by evaluation of causal relationships and druggability.	94
4.2	Heatmap of correlations among the top 30 T1D-associated proteins, identified in the fully adjusted model and ordered by hierarchical clustering. Proteins are labelled using gene symbol, chromosome, and transcriptional start position (in megabases), separated by underscore (“_”). Correlations were calculated using protein data from non-diabetic controls in the UKB-PPP dataset. .	96

S4.1 Suggestive dose-response effects of protein levels on T1D supported by the Mendelian randomisation analysis in the UK Biobank. All <i>trans</i> -pQTLs contributing to the GATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and T1D are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate.	105
S4.2 Suggestive dose-response effects of protein levels on T1D supported by the Mendelian randomisation analysis in the Scottish dataset. All <i>trans</i> -pQTLs contributing to the GATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and T1D are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate.	106
5.1 The development of persistent islet autoimmunity (IA) and type 1 diabetes (T1D) in the TEDDY nested case-control study. The age at onset ranged from 2.6 to 72.4 months for IA cases (mean = 23.8 months) and from 8.2 to 75.1 months for T1D cases (mean = 32.1 months).	122
S5.1 The availability of transcripts measured by microarray across time points (3 - 72 months) in 2013 blood samples from 401 subjects.	133
S5.2 The availability of transcripts measured by RNA-seq across time points (3 - 78 months) in 5034 blood samples from 899 subjects.	134
S5.3 The availability of serum protein data measured by Olink Target 96 Inflammation panel across time points (3 - 75 months) in 6652 blood samples from 934 subjects.	135

List of Tables

3.1	Putative core proteins for type 1 diabetes based on aggregated <i>trans</i> -effects on protein levels (pGATE).	47
3.2	Summary of supportive evidence for the putative core proteins for type 1 diabetes.	50
3.3	Associations of type 1 diabetes with <i>cis</i> -scores and measured protein levels for putative core proteins.	52
S3.1	Gene name, protein name synonyms and biological functions of the putative core proteins for type 1 diabetes.	61
S3.2	Clumps (genomic locations) harbouring <i>trans</i> -pQTLs contributing to pGATE scores of the putative core proteins for type 1 diabetes in the Scottish dataset.	65
S3.3	Mendelian randomisation analysis of the effect of the putative core proteins on type 1 diabetes.	76
S3.4	Detailed supportive evidence of putative core proteins through perturbation in experimental animal models or drug effect in humans.	77
S3.5	Comparison GATE score associations for 19 core proteins with <i>trans</i> -pQTLs in or nearby the <i>SH2B3</i> gene between the original GATE analysis and the sensitivity analysis in the Scottish dataset.	82
S3.6	Comparison of GATE scores computed with pQTLs from UKB-PPP and deCODE for associations with T1D in the Scottish and UK Biobank datasets.	85
4.1	Characteristics of T1D cases and controls without diabetes in UKB-PPP dataset at study recruitment.	95
4.2	Potential core proteins supported by both protein associations and GATE score associations.	98
4.3	Summary of supportive evidence for the potential core proteins for T1D.	100
S4.1	Code and detailed name for lipid-lowering, antihypertensive, and antithrombotic medications in UK Biobank.	106

S4.2 Mendelian randomisation analysis of the effect of the potential core proteins on T1D.	110
S4.3 <i>Cis</i> -score association for the potential core proteins with T1D. . .	111
S4.4 Associations with T1D, HbA1c and duration for the potential core proteins in the UKB-PPP dataset.	112
S4.5 Gene name, protein name synonyms and biological functions of the potential core proteins for T1D.	113
5.1 HLA haplotype eligibility for inclusion in the TEDDY study.	118
5.2 Characteristics of persistent islet autoimmunity (IA) and type 1 diabetes (T1D) case-control pairs in the TEDDY nested case-control study.	122
5.3 Nominal associations of average transcripts and protein levels before the age of one with the risk of incident IA and T1D within the next six years of follow-up.	125
5.4 Characteristics of prospective datasets for the development of T1D with proteomics data.	131
5.5 Characteristics of T1D screening and intervention programmes with available samples for future omics investigations.	132
S5.1 Detailed probe information for the transcripts of genes corresponding to the candidate core proteins and housekeeping genes. . . .	136
S5.2 Correlations between multiple probes for the same gene in the TEDDY microarray gene expression data.	138
S5.3 Correlations between RNA-seq and microarray measurements for the same gene in the TEDDY study.	139
S5.4 Correlations of microarray measurements between different timepoints in the TEDDY study.	141
S5.5 Correlations of RNA-seq measurements between different timepoints in the TEDDY study.	143
S5.6 Correlations of protein measurements between different timepoints in the TEDDY study.	144
S5.7 Associations of average transcripts levels measured by microarray before the age of one with the risk of incident IA.	145
S5.8 Associations of average transcripts levels measured by microarray before the age of one with the risk of incident T1D.	146
S5.9 Associations of average transcripts levels measured by RNA-seq before the age of one with the risk of incident IA.	147

S5.10	Associations of average transcripts levels measured by RNA-seq before the age of one with the risk of incident T1D.	148
S5.11	Associations of average proteins levels before the age of one with the risk of incident IA.	149
S5.12	Associations of average protein levels before the age of one with the risk of incident T1D.	149
6.1	Potential issues and solutions for differentiating causality from reverse causation and confounding.	172

Glossary

Abbreviations

APC: antigen-presenting cell

cOR: conditional odds ratio

GADA: glutamic acid decarboxylase autoantibody

GATE: genome-wide aggregated *trans*-effects

GWAS: genome-wide association study

HLA: human leukocyte antigen

IA: islet autoimmunity

IA-2A: tyrosine phosphatase-like protein IA-2 autoantibody

IAA: insulin autoantibody

Ig: immunoglobulin

IVW: inverse-variance weighted

LD: linkage disequilibrium

mAb: monoclonal antibody

MHC: major histocompatibility complex

MR: Mendelian randomisation

NK: natural killer

NOD mice: non-obese diabetic mice

OR: odds ratio

pQTL: protein quantitative trait loci

PTMs: post-translational modifications

SD: standard deviation

SNP: single nucleotide polymorphism

T1D: type 1 diabetes

Treg: regulatory T cell

TWAS: transcriptome-wide association study

ZnT8A: zinc transporter 8 autoantibody

1

Introduction

1.1 Type 1 diabetes

Type 1 diabetes (T1D) is a chronic autoimmune disease in which the body's immune system attacks insulin-producing β cells in the pancreatic islet, resulting in insulin deficiency. As shown in Figure 1.1, the progression to T1D has three distinct stages [Insel et al., 2015]. Stage 1 is characterised by the development of two or more islet autoantibodies, including autoantibodies to insulin (IAA), glutamic acid decarboxylase (GADA), tyrosine phosphatase-like protein IA-2 (IA-2A), and/or zinc transporter 8 (ZnT8A). Individuals at stage 1 still have normal glycaemia. In stage 2, individuals with islet autoantibodies progress to glucose intolerance or dysglycaemia due to the loss of functional β cell mass while still being presymptomatic. Dysglycaemia in this stage is defined by impaired fasting plasma glucose (i.e., fasting plasma glucose 100-125 mg/dL, 5.6-6.9 mmol/L), impaired glucose tolerance (i.e., 2-h plasma glucose 140-199 mg/dL, 7.8-11.0 mmol/L during oral glucose tolerance test), or glycated haemoglobin (HbA1c, a blood test that measures the average blood sugar levels over the past 2-3 months) 5.7-6.4%, 39-47 mmol/mol, or $\geq 10\%$ increase in HbA1c. Stage 3 is the clinical diagnosis of diabetes, marked by hyperglycaemia with the presence of some or all typical symptoms, including excessive thirst (polydipsia), frequent urination (polyuria) or bedwetting, lack of energy or fatigue, constant hunger, sudden weight loss, blurred vision, and more severe conditions such as diabetic ketoacidosis and/or hyperglycaemic hyperosmolar state [Magliano and Boyko, 2025]. The up-to-date American Diabetes Association diagnostic criteria for T1D are as follows [American Diabetes Association Professional Practice Committee, 2024]:

- HbA1c $\geq 6.5\%$, ≥ 48 mmol/mol **OR**
- Fasting plasma glucose ≥ 126 mg/dL, ≥ 7.0 mmol/L **OR**
- 2-h plasma glucose ≥ 200 mg/dL, ≥ 11.1 mmol/L during oral glucose tolerance test **OR**
- In an individual with classic symptoms of hyperglycaemia or hyperglycaemia crisis, a random plasma glucose ≥ 200 mg/dL, ≥ 11.1 mmol/L

After the diagnosis of T1D, pancreatic β cell function typically continues to decline over the first several years before stabilising at a persistently low level [Shields et al., 2018]. C-peptide, a byproduct of endogenous insulin synthesis, serves as a biomarker of residual β cell function. Although long-term longitudinal data on C-peptide levels are sparse, a characteristic trajectory – an initial rapid decline

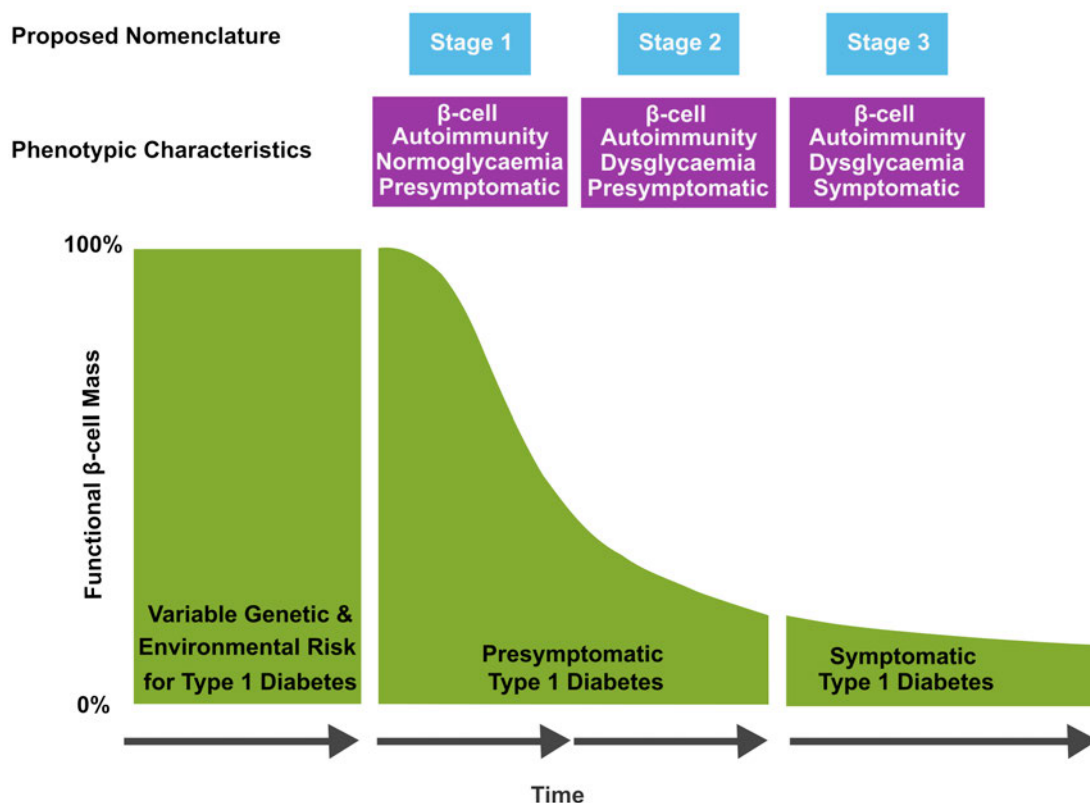


Figure 1.1: Three distinct stages of T1D. Adapted from Insel et al., Diabetes Care 2015; 38: 1964–1974. © 2015 American Diabetes Association. Used under educational, non-profit use with proper citation.

followed by a prolonged stable phase (Figure 1.2) – has been inferred based on cross-sectional measurements in individuals with varying ages at onset and disease durations [Shields et al., 2018, McKeigue et al., 2019].

1.2 Epidemiology of T1D

According to the Type 1 Diabetes Index, a joint initiative of Breakthrough T1D, Life for a Child, the International Diabetes Federation, and the International Society for Paediatric and Adolescent Diabetes, there are 9.5 million individuals living with clinically diagnosed T1D worldwide in 2025, and this number is projected to reach 14.7 million by 2040 [Ogle et al., 2025]. The prevalence of T1D varies substantially across geographic regions. Europe and Central Asia have the highest number of people with T1D (3.64 million), followed by North America (2.19 million), South Asia (1.97 million), Middle East and North Africa (1.79 million), Latin America and Caribbean (1.42 million), East Asia and Pacific (1.35 million),

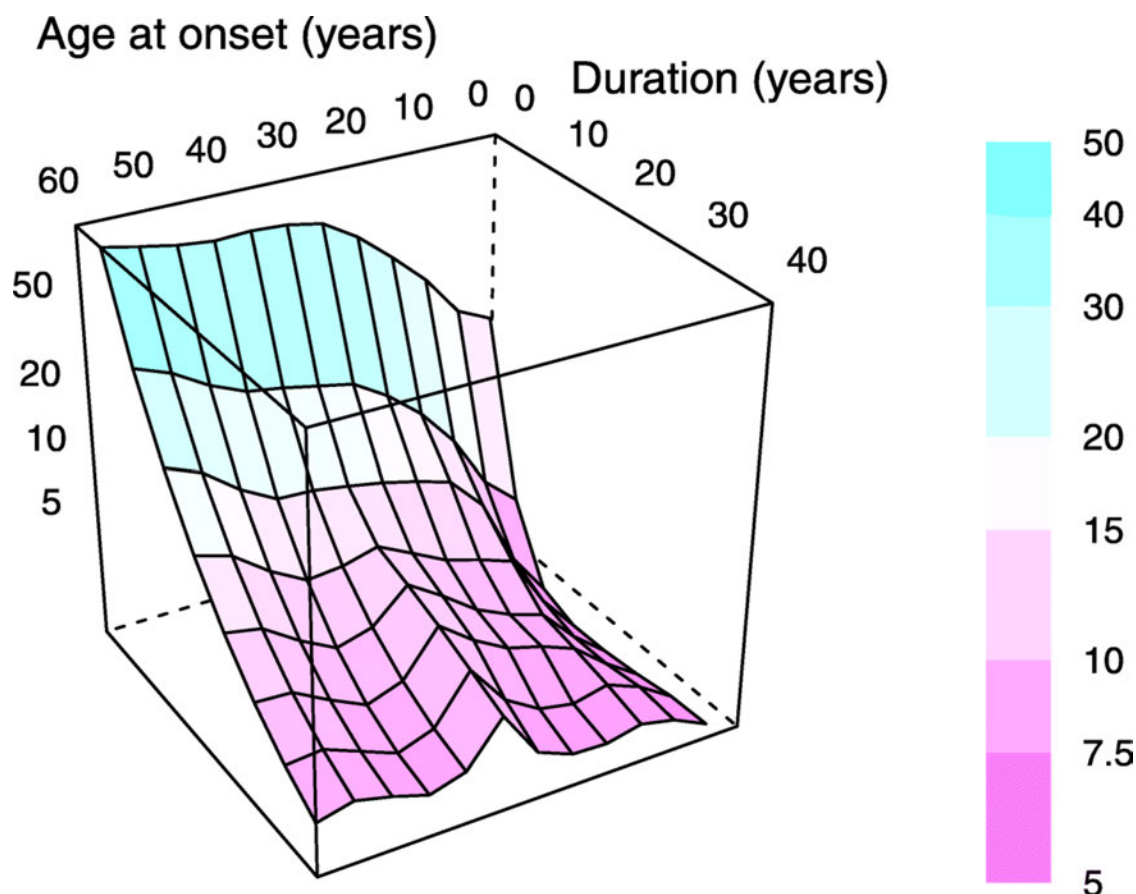


Figure 1.2: Geometric mean plasma C-peptide (pmol/l) by age at onset and duration in individuals with T1D. Log C-peptide levels were smoothed using locally estimated scatterplot smoothing (LOESS) regression (polynomial degree 2, span 0.25). In each panel, C-peptide level (pmol/l) is represented both by colour intensity and by its coordinate along the vertical axis. Reproduced from McKeigue et al., BMC Medicine 2019; 17(1):165. CC BY 4.0.

and Sub-Saharan Africa (1.22 million). Of this total population of 9.5 million, 2.12 million (22.3%) were younger than 20 years old, 8.86 million (93.3%) were between 20 and 59 years old, and 2.6 million (27.4%) were 60 or older. In 2025, 513,125 people were newly diagnosed with T1D around the world, and 173,516 lives were lost to T1D, of which 29,821 were never diagnosed.

People with T1D require insulin injections and blood glucose monitoring to maintain glucose concentrations within normal ranges. Over the past decades, new insulin analogues with various onsets and durations of action, non-insulin medications, continuous glucose monitoring (CGM) and intermittently viewed CGM devices have been widely used and have advanced diabetes care. However, among people newly diagnosed with T1D in 2025, there are still 34% of them only having access to minimal care (i.e., receiving human premixed insulin and having

limited access to self-management of blood glucose and complications screening). For a 10-year-old child diagnosed with T1D in 2024, the estimated average remaining life expectancy ranges from less than 10 years in Sub-Saharan African countries to over 60 years in some high-income countries. A high proportion of individuals with T1D still struggle to maintain their blood glucose concentrations and face an elevated risk for a series of acute and chronic complications. Acute complications include hypoglycaemia, diabetic ketoacidosis, and hyperglycaemic hyperosmolar state, which require immediate recognition, diagnosis, and intervention. Chronic complications include microvascular complications (i.e., retinopathy, nephropathy, and neuropathy) and macrovascular complications (i.e., cardiovascular disease and peripheral artery disease). Moreover, people with T1D are more likely to develop other autoimmune diseases, particularly autoimmune thyroid disease, coeliac disease, and pernicious anaemia [Nederstigt et al., 2019].

1.3 Genetic predisposition of T1D

T1D has a strong genetic component, as evident by its higher prevalence in relatives of affected individuals compared to the general population. While the prevalence of T1D is around 0.5% in the general population [Carstensen et al., 2020, Xu et al., 2018, Gajewska et al., 2020], the risk increases significantly among family members. Monozygotic twins share 100% of their genetic variants, and dizygotic twins or non-twin siblings share approximately 50% of their genetic variants. When there is a proband twin diagnosed with T1D, the observed concordance rate of a co-twin developing T1D ranges from 23% to 61% for monozygotic twins and from 0% to 27% for dizygotic twins [Giwa et al., 2020]. Compared with dizygotic twins and non-twin siblings, monozygotic twin siblings of probands with T1D are also more likely to develop islet autoantibodies. The sibling recurrence risk ratio (λ_s) – defined as the ratio of risk for siblings of individuals with a disease to population prevalence – was estimated between 15 and 20 [Risch, 1987, Rich, 1990], further underscoring the strong familial aggregation of T1D. The heritability of T1D (proportion of phenotypic variance explained by the genetic component) ranges from 0.5 to 0.88 based on family and twin studies. Moreover, the heritability is higher for childhood-onset T1D (0.81) than adult-onset T1D (0.56) [Wei et al., 2024].

Over the past decades, the major histocompatibility complex (MHC), also known as the human leukocyte antigen (HLA) region in humans, has been established

as the strongest contributor to the genetic predisposition of T1D, accounting for approximately 50% of its heritability [Redondo et al., 2023]. Beyond the HLA region, candidate gene and linkage studies have identified additional susceptibility loci, including *INS*, *CTLA4*, and *PTPN22*. Prior to the large-scale genome-wide association studies (GWASs), the tag single nucleotide polymorphism (SNP) approach and genome-wide nonsynonymous SNP scan identified *IL2RA* (*CD25*) and *IFIH1* as being associated with T1D [Vella et al., 2005, Smyth et al., 2006]. So far, GWASs have uncovered around a hundred loci containing common variants with relatively small effects [Onengut-Gumuscu et al., 2025]. Meanwhile, the development of next-generation DNA sequencing techniques has advanced the discovery of rare variants with large effects [Forgetta et al., 2020].

1.3.1 HLA region

The HLA region, located on the short arm of human chromosome 6p21.3, is a highly polymorphic area with strong linkage disequilibrium (LD) between particular loci. It is divided into three sub-regions encoding class I (HLA-A, -B, and -C), class II (HLA-DP, -DQ, and -DR), and class III molecules. HLA class I molecules consist of a single polypeptide chain, while class II molecules are heterodimers formed by the products of two genes (A and B, encoding the α and β polypeptide chains, respectively) [Noble, 2024]. Both HLA class I and class II molecules play crucial roles in antigen presentation, whereas HLA class III molecules are primarily involved in immune responses via the complement system and inflammatory pathways.

Polymorphisms of HLA class II genes encoding DQ and DR are the major genetic determinants of T1D. The DR3 haplotype (DRB1*03:01~DQA1*05:01~DQB1*02:01) consistently increases T1D risk across diverse populations, though the strength of the association varies [Noble, 2024]. In contrast, DR4 haplotypes (DRB1*04:XX) exhibit more complex associations, with a clear risk hierarchy: highest risk – DRB1*04:05, moderate risk – DRB1*04:01, DRB1*04:02, and DRB1*04:04, and protective effect – DRB1*04:03. The DR3/DR4 heterozygous genotype, with one chromosome carrying the DR3 haplotype and the other carrying a high-risk DR4 haplotype, confers a much higher risk of T1D than the homozygous DR3/DR3 or DR4/DR4 genotypes, indicating a non-additive interaction [Noble and Valdes, 2011]. In a study of 607 Caucasian families by the Type 1 Diabetes Genetics Consortium (T1DGC), the following haplotypes presented the strongest risk or protection [Erlich et al., 2008].

Risk haplotypes

- DRB1*04:05~DQA1*03:01~DQB1*03:02, odds ratio (OR) = 11.37
- DRB1*04:01~DQA1*03:01~DQB1*03:02, OR = 8.39
- DRB1*03:01~DQA1*05:01~DQB1*02:01, OR = 3.64
- DRB1*04:02~DQA1*03:01~DQB1*03:02, OR = 3.63
- DRB1*04:04~DQA1*03:01~DQB1*03:02, OR = 1.59

Protective haplotypes

- DRB1*07:01~DQA1*02:01~DQB1*03:03, OR = 0.02
- DRB1*14:01~DQA1*01:01~DQB1*05:03, OR = 0.02
- DRB1*15:01~DQA1*01:02~DQB1*06:02, OR = 0.03
- DRB1*11:04~DQA1*05:01~DQB1*03:01, OR = 0.07
- DRB1*13:03~DQA1*05:01~DQB1*03:01, OR = 0.08

HLA class I genes exhibit strong LD with class II DR- and DQ-encoding genes [Noble, 2024]. When accounting for LD, A*24, specifically A*24:02, is the most consistently reported risk allele for T1D at the HLA-A locus, and B*39:06 is the most commonly reported risk allele at the HLA-B locus. HLA-C has been less extensively investigated, and no allele has shown consistent associations with T1D.

1.3.2 Insulin (*INS*) gene

The strongest non-HLA association with T1D was discovered at the insulin (*INS*) gene on chromosome 11p15.5. This association is driven by a variable number tandem repeat (VNTR) located 596 base pairs (bp) upstream of the transcriptional start site of the *INS* gene [Bell et al., 1984, Bennett et al., 1995, Durinovic-Belló et al., 2010]. The VNTR is formed by a 14-bp consensus sequence with three allelic classes defined by repeat length: class I alleles (short, 26-63 repeats), class II alleles (intermediate, 63-140 repeats), and class III alleles (long, 141-209 repeats). The class I alleles are associated with a higher susceptibility to T1D, while class III alleles are protective against T1D. Although the precise mechanism is still unknown, it has been discovered that VNTR can regulate *INS* mRNA expression in a tissue-specific pattern. Class III alleles are associated with lower *INS* mRNA levels in the pancreas but higher levels in the thymus, while class I alleles are associated with higher *INS* transcription in the pancreas.

but lower in the thymus [Vafiadis et al., 1997, Pugliese et al., 1997]. In the thymus, the low levels of *INS* transcription may promote the positive selection of T cells and cause migration of CD4⁺ proinsulin-specific T cells to the periphery, whereas the high levels of *INS* transcription will foster negative selection of T cells and thus induce immune tolerance [Xie et al., 2014].

1.3.3 *CTLA4*

The *CTLA4* gene, located on chromosome 2q33, encodes a critical immune checkpoint protein that downregulates T cell activation. An initial study identified an A/G transition at position 49 in exon 1, resulting in a threonine/alanine substitution in the leader peptide, as associated with T1D [Nisticò et al., 1996]. However, subsequent studies discovered an LD block within the *CTLA4* region and rejected the +49G>A variant being causal, as the peak association with T1D was from the ~43 kilobase (kb) region at ~2.8 kb downstream of the 3' region flanking *CTLA4* [Ueda et al., 2003, Qu et al., 2009]. Further comprehensive sequencing and fine mapping would be required to localise the causal variant(s) in this region.

1.3.4 *PTPN22*

Located on human chromosome 1p13, the *PTPN22* gene encodes lymphoid tyrosine phosphatase (LYP), a negative regulator of T cell receptor (TCR) signalling. A nonsynonymous SNP, 1858C>T (also known as R620W, rs2476601), causing an arginine/tryptophan substitution in codon 620, was first discovered to be associated with T1D as well as Graves's disease [Smyth et al., 2004]. This association was subsequently replicated in a few other autoimmune diseases, such as rheumatoid arthritis, juvenile idiopathic arthritis, and systemic lupus erythematosus [Bottini et al., 2006]. By gene editing, it was confirmed that this variant promotes autoimmunity by enhancing TCR signalling in mildly self-reactive T cells [Anderson et al., 2023].

1.3.5 *IL2RA (CD25)*

The *IL2RA (CD25)* gene, located on chromosome 10p15, encodes the α subunit of the IL-2 receptor complex. This receptor plays a pivotal role in immune regulation by modulating the differentiation and homeostasis of both pro- and anti-inflammatory T cells [Ross and Cantrell, 2018]. A rare mutation in *CD25* has

been reported previously as a monogenic cause of severe autoimmune disease [Sharfe et al., 1997]. With the tag SNP approach in the *IL2RA* region among 7457 T1D cases and controls and 725 multiplex families, strong evidence supported the genetic association with T1D [Vella et al., 2005].

1.3.6 *IFIH1*

Located on chromosome 2q24.3, *IFIH1* encodes melanoma differentiation-associated protein 5 (MDA5), an intracellular sensor of viral RNA that triggers the innate immune response. In a genome-wide scan of over 6500 nsSNPs in an interim analysis of 2029 cases and 1755 controls, rs1990760 (A946T) emerged as the third most strongly associated nsSNP with T1D [Smyth et al., 2006]. To address the concern of limited statistical power, this association was further validated in 4253 cases and 5842 controls as well as 2134 parent-child trio genotypes. A robust association of the rs1990760 G allele with T1D was confirmed.

1.3.7 Large-scale GWAS of T1D

The first large-scale GWAS of T1D was conducted by the Wellcome Trust Case-Control Consortium (WTCCC), analysing ~2000 T1D cases and ~3000 controls from the British population [Burton et al., 2007]. This study identified seven independent hits meeting a p -value threshold of 5×10^{-7} , with five previously reported associations (HLA, *CTLA4*, *PTPN22*, *IL2RA*, and *IFIH1*) replicated at p -value ≤ 0.001 . Subsequent analysis of 4000 T1D cases, 5000 controls, and 2997 family trios confirmed four of these seven regions: 12q24, 12q13, 16p13, and 18p11 [Todd et al., 2007].

In 2009, a meta-analysis combining data from the T1DGC, WTCCC, and the Genetics of Kidneys in Diabetes/National Institute of Mental Health (GoKinD/NIMH) studies (7514 T1D cases and 9045 controls) identified 41 distinct genomic locations associated with T1D (p -value $< 10^{-6}$) [Barrett et al., 2009]. Of these, 27 regions had not been reported previously and were taken forward for a replication in an independent set of 4267 cases, 4463 controls, and 2319 affected sib-pair families. Eighteen of these 27 regions reached a p -value < 0.01 in the replication analysis, and another four regions met the nominal p -value < 0.05 threshold. Together, this revealed several new candidate genes that had not been identified earlier, such as *IL10*, *IL19*, *IL20*, *GLIS3*, *CD69* and *IL27*.

Two years later, an updated meta-analysis incorporated two additional studies – the Diabetes Control and Complications Trial-Epidemiology of Diabetes Interventions and Complications (DCCT-EDIC) cohort and the Children’s Hospital of Philadelphia (CHOP) study – expanding the discovery set to 9934 cases and 16956 controls [Bradfield et al., 2011]. 53 SNPs that were deemed significant ($p < 0.00001$) in the fixed-effects meta-analysis were taken forward into replication in 1120 affected trios. Combining together, three novel loci reached genome-wide significance (5×10^{-8}): rs539514 in an intronic region of the *LMO7* gene on 13q22, rs478222 in an intronic region of the *EFR3B* gene on 2p23, and rs924043 in an intergenic region on 6q27. However, the latter two loci reside in regions with complex LD patterns, complicating the identification of causal genes.

To refine known T1D risk loci, map regulatory sequences in relevant tissues and cell types, and uncover additional risk loci shared with other autoimmune diseases, T1DGC designed the ImmunoChip – a targeted genotyping array comprising approximately 186000 SNPs clustered in 186 genomic regions previously implicated in one or more of 12 autoimmune conditions [Onengut-Gumuscu et al., 2025]. Using this array, a fine-mapping study was conducted involving 6670 T1D cases, 12262 controls, 2601 affected sib-pairs, and 69 trio families [Onengut-Gumuscu et al., 2015]. This analysis identified 44 regions significantly associated with T1D at the p -value threshold of 3.23×10^{-7} (ImmunoChip Bonferroni-corrected $p < 0.05$), including four novel loci: 1q32.1 (index SNP rs6691977), 2q13 (rs4849135), 4q32.3 (rs2611215), and 5p13.2 (rs11954020). Notably, the emerging T1D-associated loci exhibited a genetic architecture similar to other autoimmune diseases characterised by autoantibodies, such as juvenile idiopathic arthritis, rather than autoinflammatory disorders like ulcerative colitis and Crohn’s disease. Additionally, T1D-associated SNPs were enriched in enhancer sequences in immunologically relevant tissues, especially the thymus, T cells, B cells and CD34⁺ stem cells.

This ImmunoChip was subsequently applied to two prospective cohorts – the Environmental Determinants of Diabetes in the Young (TEDDY) study and the TrialNet Pathway to Prevention (TrialNet PTP) study – to investigate associations between non-HLA SNPs and the progression to islet autoimmunity (IA) and T1D. In the TEDDY study, which included 5806 subjects, three genomic regions were associated with the development of any persistent confirmed islet autoantibody: two previously known loci near *SH2B3* and *PTPN22* and a novel region near *PPIL2* [Sharma et al., 2018]. For the development of multiple autoantibodies,

associations were identified at two known loci (*PTPN22* and *INS*) and one novel locus (*PXK/PDHB*). Additional associations were observed depending on the first appearing autoantibody: the *INS* and *TTC34/PRDM16* regions were linked to IAA, while *RBFOX1* was related to GADA. Regarding progression to T1D, four regions showed suggestive associations: one known region (*INS*) and three novel regions (*RNASET2*, *PLEKHA1*, and *PPIL2*). In the TrialNet PTP study of 1016 participants with a single positive autoantibody, two variants – rs6476839 (*GLIS3*) and rs3184504 (*SH2B3*) – were associated with progression to diabetes; another variant, rs3087243 (*CTLA4*), was associated with progression to multiple autoantibodies [Steck et al., 2017].

More recently, multi-ancestry and multi-omics approaches for GWAS have been applied to T1D. A meta-GWAS of 18942 T1D cases and 501638 controls of European ancestry from nine cohorts identified 81 loci reaching genome-wide significance, including 48 of 59 previously reported loci and 33 novel loci [Chiou et al., 2021]. These T1D-associated variants were enriched in *cis*-regulatory elements active in T cells, as well as acinar and ductal cells of the exocrine pancreas. In another meta-analysis incorporating 16159 T1D cases, 25386 controls, and 6143 trio families from European, African admixed, East Asian, and other admixed populations with ImmunoChip genotyping data, 78 genomic regions were identified as genome-wide significant, 36 of which were novel [Robertson et al., 2021]. These risk variants were enriched in open chromatin in multiple primary immune cell types, particularly stimulated CD4⁺ effector T cells. Colocalisation analysis revealed five loci where T1D risk variants overlapped with chromatin accessibility quantitative trait loci (caQTLs). Moreover, for the *BACH2* locus, both enhancer caQTL and expression QTL (eQTL) colocalise with T1D association. In a recent multi-ancestry GWAS involving 13412 individuals of European, African and admixed ancestry, seven loci reached genome-wide significance for T1D risk, and four loci were associated with age at onset of T1D at suggestive levels of genome-wide significance (5×10^{-7}) [Michalek et al., 2024].

1.3.8 Rare variants associated with monogenic autoimmune diabetes

Rare variants are too infrequent to be captured by common SNP chips in GWAS studies, and in linkage studies, their detection is hindered by the combination of their rarity and typically modest effect sizes [Pang et al., 2021, 2022]. However,

the rapid development of next-generation DNA sequencing techniques – such as whole-genome sequencing (WGS), whole-exome sequencing (WES), and targeted sequencing – has greatly enhanced the ability to detect rare variants. Using deep sequencing of exons and splice sites in 10 candidate genes, four rare variants within the *IFIH1* locus were discovered as being associated with reduced T1D risk by altering the expression and structure of MDA5 [Nejentsev et al., 2009]. Similarly, targeted deep sequencing of 301 protein-coding genes located in 49 previously reported T1D risk loci uncovered four rare deleterious variants in *PTPN22* (two novel frameshift mutations and two missense variants) significantly associated with T1D [Ge et al., 2015]. One of these variants, rs56048322, affects the splicing of *PTPN22* and results in the production of two *PTPN22* transcripts and a novel LYP isoform, which competes with wild-type LYP and leads to CD4⁺ T cell hyporesponsiveness to antigen stimulation. Based on deep imputation of genotyped data in GWAS, three novel variants outside the HLA region with minor allele frequency < 0.05 were identified [Forgetta et al., 2020]. Among them, rs60587303, which resides in the first intron of *STK39*, was prioritised via in silico functional analysis, and an in vitro follow-up functional experiment deemed that inhibition of *STK39* may augment the sensitivity of activated T cells to IL-2.

In addition, several rare mutations known to cause monogenic autoimmune diseases have also been implicated in autoimmune diabetes. For instance, rare mutations in *AIRE*, *IL2RA*, *LRBA*, and *STAT5B* are inherited in an autosomal recessive manner, mutations in *STAT1* and *STAT3* follow an autosomal dominant pattern, and mutations in *FOXP3* exhibit X-linked recessive inheritance [Bonnetfond et al., 2023]. According to the DiabetesGenes website (<https://www.diabetesgenes.org/tests-for-diabetes-subtypes/monogenic-autoimmunity/>), genetic testing for 15 known monogenic autoimmune genes is recommended for individuals diagnosed with diabetes and at least one additional autoimmune condition before the age of five. Lately, two pairs of siblings with loss-of-function mutations in *PDCD1* (encoding PD-1) and *CD274* (encoding the ligand of PD-1, PD-L1) were reported to exhibit severe autoimmunity, including extremely early-onset T1D [Ogishi et al., 2021, Johnson et al., 2024].

1.4 Disease-modifying therapy of T1D

In clinical practice, insulin remains the main treatment for T1D. Although insulin therapy can effectively control blood glucose, it does not delay, halt or reverse

the progressive β cell destruction in T1D. This limitation has driven the pursuit of disease-modifying therapy that can alter the course of the disease or prevent its onset. Given the autoimmune nature of T1D, immunotherapy has been a focus of this exploration. In the 1980s, cyclosporine – a calcineurin inhibitor primarily targeted at T cells – was the first immunotherapy tested on recent-onset T1D patients. It managed to increase the rate and length of remission (i.e., good metabolic control in the absence of insulin treatment) [Feutren et al., 1986]. However, disease progression resumed after treatment ceased, and no lasting improvement in β cell function was observed compared to placebo [Martin et al., 1991]. Prolonged treatment with cyclosporine was also limited due to the risk of nephrotoxicity and malignancy. Similarly, immunosuppression with azathioprine and prednisone improved metabolic control in some newly diagnosed T1D patients but caused strong side effects [Silverstein et al., 1988]. These limitations have shifted the research focus toward therapies that promote immune tolerance over broad immunosuppression. The current goal of immunotherapy is to develop short-term therapies capable of re-educating the immune system to induce long-term remission with minimal adverse events [Warshauer et al., 2020].

1.4.1 Immunotherapies targeting specific immune cells

1.4.1.1 T cells

Anti-CD3 immunotherapy: CD3, a protein complex and T cell co-receptor, has been a key drug target because of its pivotal role in activating both the CD8⁺ cytotoxic T cells and CD4⁺ T helper cells.

Ortho Kung T3 (OKT3) is the first Fc receptor (FcR)-binding monoclonal antibody (mAb) targeting the ϵ domain of CD3 that was approved to treat acute kidney allograft rejection, but its use in T1D was hindered due to the severe side effect caused by cytokine storms [Gaglia and Kissler, 2019].

Teplizumab, an FcR non-binding mAb targeting the ϵ domain of CD3, is the first and only immunotherapy that has been approved to delay the onset of T1D by the United States Food and Drug Administration (FDA). The initial phase 2 clinical trials were conducted in stage 3 T1D to evaluate the safety, tolerability, and efficacy of a single 14-day course of teplizumab administration [Herold et al., 2002, 2005]. The success in improving C-peptide responses and clinical parameters in people with T1D for at least two years without continued immunosuppressive

medications prompted another phase 2 clinical trial (AbATE), in which a second course of teplizumab was administered after one year to prolong the duration of response [Herold et al., 2013b]. Two subsequent phase 3 clinical trials (Protege and Encore) tested three dosing regimens of teplizumab given in two courses six months apart. The primary outcome – the proportion of patients with insulin use of less than 0.5 U/kg per day and HbA1c of less than 6.5% at one year – was not met in the Protege study, leading to the termination of further development [Sherry et al., 2011]. Afterwards, a phase 2 trial (Delay) assessed the efficacy of a single 14-day course of teplizumab treatment in patients diagnosed with T1D for at least 4 but not more than 12 months who still had clinically significant levels of stimulated C-peptide [Herold et al., 2013a]. Taking these clinical trials in stage 3 T1D together, the efficacy in preservation of β cell function was consistent across multiple trials; lymphopenia, rash, and headache were the most common adverse events [Herold et al., 2023]. Additionally, among relatives of people with T1D who were at stage 2, teplizumab significantly delayed the progression to stage 3 by a median of 32.5 months [Herold et al., 2019, Sims et al., 2021].

Otelixizumab is another Fc-receptor non-binding mAb targeting the ϵ domain of CD3. A total dose of 48-64 mg otelixizumab showed efficacy in preserving residual β cell function for 18 months in recent-onset T1D patients. However, 75% of patients developed a “flu-like” syndrome and symptoms of Epstein-Barr viral (EBV) mononucleosis [Keymeulen et al., 2005]. To mitigate the side effects, low-dose otelixizumab (3.1 mg) was evaluated in two subsequent trials [Aronson et al., 2014, Ambery et al., 2014]. Although low-dose otelixizumab was well tolerated in adolescents and adults with recent-onset T1D, it failed to preserve β cell function. These findings prompted further investigation to identify the optimal dose balancing efficacy, safety, and tolerability [Keymeulen et al., 2021]. A dose-response manner was observed in compromising immunocompetence; 18 mg was the maximum tolerated dose, and 9 mg emerged as the optimal therapeutic dose with the efficacy in preserving β cell function over 18 months but without EBV reactivation. This effect appeared to be related to a higher baseline β cell residual function, which indicated that otelixizumab might be more effective in early-stage T1D when overt hyperglycaemia is not yet apparent. Currently, there are no ongoing clinical trials of otelixizumab.

Anti-CD2 therapy: CD2 is a surface protein expressed mainly on effector memory T cells and central memory T cells. CD58 (lymphocyte function-associated antigen 3, LFA3), mainly expressed on antigen-presenting cells (APCs), is a ligand of CD2. Alefacept, a dimeric fusion protein of LFA3 and immunoglobulin (Ig),

was the first drug approved by the FDA for the treatment of moderate-to-severe plaque psoriasis. Given that T1D is likely mediated by effector memory T cells, alefacept was trialled in newly diagnosed T1D patients [Rigby et al., 2013, 2016]. Alefacept showed efficacy in preserving C-peptide secretion, lowering insulin use, and reducing hypoglycaemic events in these phase 2 clinical trials. However, this drug was voluntarily discontinued by its manufacturer, and its trial was terminated. Siplizumab, a humanised anti-CD2 IgG1 mAb, has been examined for treating graft-versus-host disease, psoriasis, and T cell lymphoma and is currently being investigated in a dose-finding phase 1b trial in newly diagnosed T1D patients aged 8-45 (ClinicalTrials.gov ID: NCT05574335).

Anti-thymocyte globulin (ATG) therapy: ATGs are purified rabbit IgG antibodies directed against human T cells without a specific epitope. Historically, ATG has been used after bone marrow transplants, solid organ transplants, and aplastic anaemia. In the initial clinical trial in T1D, high-dose (9 mg/kg for day 1 and 3 mg/kg for days 2-4) ATG was administered to recent-onset T1D patients [Saudek et al., 2004]. Despite the preservation of C-peptide levels and an improvement in HbA1c, significant adverse effects – transient fever and moderate symptoms of serum sickness – were also observed. Subsequent trials shifted the focus to low-dose ATG and ATG plus granulocyte colony-stimulating factor (GCSF) to balance the strong immunosuppressive effect of ATG alone. In individuals with established T1D (duration between four months and two years), ATG/GCSF combination treatment tended to preserve β cell function [Haller et al., 2015], while in new-onset T1D patients (duration less than 100 days), ATG slowed the decline of C-peptide and reduced HbA1c, but GCSF did not enhance C-peptide preservation [Haller et al., 2018]. Current trials are investigating the effect of low-dose ATG with other agents such as verapamil or adalimumab on preserving β cell functions in newly diagnosed T1D patients (ClinicalTrials.gov ID: NCT07061574, NCT06455319) as well as its own effect on delaying or preventing progression to stage 3 T1D (ClinicalTrials.gov ID: NCT04291703).

Cytotoxic T-lymphocyte protein 4 (CTLA4) analogues: CTLA4 is an immune checkpoint, defined as an essential regulator on immune cells that can either activate or inhibit an immune response. In the context of cancer, cancer cells can exploit immune checkpoints to evade the immune system. It has been proposed that T cell activation requires two signals: one from the TCR interacting with MHC-peptides on the APC surface and another from co-stimulatory molecules on APCs, such as the ligands CD80 and CD86 on APCs binding to CD28 on T

cells [Marelli-Berg et al., 2007]. CTLA4 is a competitive inhibitor of CD28 with higher affinity to the shared ligands on APCs. Abatacept is a CTLA4-Ig fusion protein that prevents T cell activation by blocking CD80 and CD86 on APCs. In recent-onset T1D patients, continued administration of abatacept over two years slowed reduction in β cell function and improved HbA1c [Orban et al., 2011]. This beneficial effect is retained for at least one year after cessation of abatacept or three years after T1D diagnosis [Orban et al., 2014]. Nevertheless, abatacept treatment for one year did not significantly delay progression to stage 2 or 3 among individuals at stage 1 [Russell et al., 2023]. At present, abatacept is being trialled in combination with insulin (ClinicalTrials.gov ID: NCT05742243) or rituximab (ClinicalTrials.gov ID: NCT03929601) in recent-onset T1D.

Regulatory T cell (Treg) therapy: Tregs are a subset of T cells that maintain immune tolerance by potentially suppressing the activation and effector function of other immune cells [Gitelman and Bluestone, 2016]. Tregs develop both in the thymus and periphery via interleukin-2 (IL-2) and transforming growth factor- β (TGF- β) signalling. Tregs act as a protective factor for T1D, with evidence from both NOD mice and humans. In a phase 1 trial, Tregs were isolated from patients with recent-onset T1D, ex vivo-expanded, and infused back [Bluestone et al., 2015]. A subset of infused Tregs were long-lived in the circulation at one year after transfer, and no infusion reactions or severe cell therapy-related adverse events were observed. To enhance the survival and expansion of infused Tregs, another phase 1 trial combining Tregs and low-dose IL-2 was conducted [Dong et al., 2021]. Due to the dual functionality of IL-2 in both promoting the differentiation and homeostasis of Tregs and stimulating the proliferation, differentiation, and growth of effector T cells, this combination strategy expanded both infused and endogenous Tregs but also resulted in a substantial increase in cytotoxic cells. Now, there are a few clinical trials investigating the safety and efficacy of autologous Tregs or umbilical cord blood Tregs on established T1D (ClinicalTrials.gov ID: NCT06708780, NCT02932826, NCT03011021), as well as the combination therapy with rituximab in preventing T1D (ClinicalTrials.gov ID: NCT06688331). As a stimulating factor for Treg development, low-dose IL-2 alone has also been tested in a few phase 1/2 clinical trials and shown to be well tolerated and promote Tregs in newly diagnosed T1D patients [Hartemann et al., 2013, Rosenzwajg et al., 2020]. A phase 2 trial that uses ciclosporin to block effector T cells and then low-dose IL-2 to promote Tregs is underway (ClinicalTrials.gov ID: NCT05153070).

1.4.1.2 B cells

Although T1D is considered a T-cell-mediated autoimmune disease, B cells also play an important role as APCs in T1D.

Anti-CD20 therapy: CD20 is a protein expressed on the surface of normal and malignant B cells and is involved in their proliferation and differentiation. Rituximab is an anti-CD20 mAb that depletes B cells. The first clinical trial in newly diagnosed T1D patients revealed that a four-dose rituximab treatment within three months of diagnosis can practically preserve β cell function over a period of one year [Pescovitz et al., 2009]. However, in the follow-up study of 30 months after randomisation, the rate of fall in C-peptide was similar between patients treated with rituximab and those with placebo, indicating that rituximab was unable to alter the fundamental pathophysiology of disease (i.e., ongoing β cell destruction) [Pescovitz et al., 2014]. Additionally, rituximab was also combined with Tregs in paediatric T1D patients and showed efficacy in delaying the progression of T1D with regard to C-peptide levels and remission, and this effect was superior to monotherapy of rituximab or Tregs [Zieliński et al., 2022]. Currently, there are three ongoing clinical trials: one on rituximab monotherapy in adolescents with recent-onset T1D (ClinicalTrials.gov ID: NCT07041268), one on the combined therapy with Tregs in children at stage 1 (ClinicalTrials.gov ID: NCT06688331), and one on the sequential therapy of rituximab and abatacept in new-onset T1D patients (ClinicalTrials.gov ID: NCT03929601).

1.4.2 Cytokine-directed interventions

As an alternative to directly targeting immune cells, interventions targeting generalised inflammatory responses via cytokines have also been investigated.

Interleukin-1 (IL-1) inhibition IL-1 is a vital immune mediator in the pathogenesis of T1D by triggering β cell apoptosis and inducing local inflammation [Lu et al., 2020]. An anti-IL-1 mAb, canakinumab, and an antagonist of the IL-1 receptor, anakinra, have been trialled in recent-onset T1D patients [Moran et al., 2013]. Regretfully, canakinumab and anakinra failed to preserve C-peptide secretion.

Tumour necrosis factor (TNF) inhibition: Produced by dendritic cells and macrophages in pancreatic islet infiltrates, TNF accelerates β cell apoptosis [Lu et al., 2020]. Blockade of TNF with a soluble TNF receptor fusion protein (etanercept) was first attempted in children newly diagnosed with T1D and showed

promising results in lowering HbA1c and increasing endogenous insulin production [Mastrandrea et al., 2009]. Golimumab, an mAb targeting TNF, exhibited efficacy in preserving β cell function in a phase 2 trial of children and young adults with newly diagnosed T1D [Quattrin et al., 2020]. It is also under investigation in phase 1b to prevent T1D among children, adolescents and young adults at stage 2 T1D (ClinicalTrials.gov ID: NCT03298542).

Interleukin-12 (IL-12) and interleukin-23 (IL-23) inhibition: IL-12 and IL-23 are two members of the IL-12 family, with probable multifaceted roles in T1D [Lu et al., 2020]. Ustekinumab, an mAb used for psoriasis, which inhibits the IL-12 and IL-23 signalling pathways, showed safety in T1D at the doses for psoriasis treatment [Marwaha et al., 2022]. A subsequent phase 2 trial in adolescents with recent-onset T1D suggested that the efficacy of ustekinumab in preserving β cell function could be attributable to the reduction of a subset of activated T helper 17.1 cells [Tatovic et al., 2024]. Another trial with a similar protocol in adults with recent-onset T1D is underway (ClinicalTrials.gov ID: NCT03941132).

Interleukin-21 (IL-21) inhibition: IL-21 is a pleiotropic cytokine that affects the differentiation and function of a wide range of immune cells and has been recognised as promoting disease progression in T1D [Lu et al., 2020]. Anti-IL-21 mAb was assessed in combination with liraglutide (glucagon-like peptide 1 receptor agonist) among adults with recent-onset T1D [von Herrath et al., 2021]. This combination appeared to have similar efficacy as other disease-modifying interventions but a safer profile.

Interleukin-6 (IL-6) inhibition: IL-6 is a mediator in the development and progression of autoimmune diseases [Lu et al., 2020]. Tocilizumab, an mAb targeting the IL-6 receptor and indicated to treat rheumatoid arthritis, failed to modulate CD4⁺ T cell phenotypes or preserve β cell function in individuals with newly diagnosed T1D [Greenbaum et al., 2021]. There has also been a report of a rheumatoid arthritis patient developing T1D after taking tocilizumab [Kawasaki et al., 2022].

1.4.3 Antigen-specific therapy

As described earlier, T1D is characterised by the presence of autoantibodies against autoantigens in the islet. The idea of antigen-specific therapy is to avoid the off-target effect and induce immune tolerance by balancing the T cell population between effector and regulatory T cells.

Peptide immunotherapy: Insulin and GAD65 are considered the primary antigen targets. A dose of 7.5 mg oral insulin per day failed to delay or prevent the development of T1D in relatives of T1D patients who are positive for autoantibodies [The Diabetes Prevention Trial-Type 1 Study Group, 2005, Writing Committee for the Type 1 Diabetes TrialNet Oral Insulin Study Group, 2017]. In two pilot studies among children who are autoantibody-negative but are at high risk defined by family history and HLA haplotypes, dose-increasing oral insulin intervention succeeded in inducing an immune response without hypoglycaemia for those aged two to seven years but not for those aged six months to two years [Bonifacio et al., 2015, Assfalg et al., 2021]. Exploratory analysis indicated that the *INS* AA genotype was associated with immune responses to insulin for children aged six months to two years [Assfalg et al., 2021], which prompted a follow-up trial on infants with a genetically predicted risk of > 10% to develop autoimmunity by the age of six years [Ziegler et al., 2019], but its results are not available yet. Moreover, intranasal and subcutaneous insulin were also tested in seropositive individuals with high genetic risk but did not manage to prevent or delay the onset of T1D [Näntö-Salonen et al., 2008, Vandemeulebroucke et al., 2009]. In people with new-onset T1D, neither the altered insulin B (9-23) peptide ligand (NBI-6024) nor the proinsulin peptide had shown success in preserving β cell function [Walter et al., 2009, Alhadj Ali et al., 2017].

Initial trials of aluminium-formulated GAD65 (GAD-alum) in recent-onset T1D patients yielded discordant results in the efficacy of preserving β cell function [Ludvigsson et al., 2008, Wherrett et al., 2011, Ludvigsson et al., 2012]. A Bayesian meta-analysis based on these three trials revealed that it is highly likely that GAD-alum vaccination has a positive biological effect, while it only confers relatively small clinical benefits [Beam et al., 2017]. Since intralymphatic injections are more effective than conventional administration, GAD-alum was then tested intralymphatically in a pilot study of recent-onset T1D patients, which showed promising results [Ludvigsson et al., 2017, Casas et al., 2020]. Meanwhile, a stratification analysis based on HLA haplotypes confirmed that individuals with the HLA DR3~DQ2 haplotype but not HLA DR4~DQ8 are responders for GAD-alum therapy [Hannelius et al., 2020]. Exciting results in preserving endogenous insulin secretion in recent-onset T1D patients with the HLA DR3~DQ2 haplotype were reported in a phase 2 study of intralymphatic GAD-alum injection and vitamin D supplementation [Ludvigsson et al., 2021]. Now, a phase 3 trial in recent-onset T1D (ClinicalTrials.gov ID: NCT05018585) [Ludvigsson et al., 2022]

and a phase 2 trial in individuals at stage 1 or stage 2 (ClinicalTrials.gov ID: NCT05683990) are underway for GAD-alum.

Additionally, multiple peptides derived from β cell autoantigens demonstrated promising effects in inducing Treg immune tolerance in a phase 1 study of T1D patients [Liu et al., 2022]. IMCY-0098, a modified peptide derived from human proinsulin, presented a promising safety profile and preliminary clinical maintenance of β cell function in a phase 1 study in recent-onset T1D [Van Rangelbergh et al., 2023]. Further phases of clinical trials are warranted to confirm the safety and efficacy of these agents. Additionally, advances in medical nanotechnology have illuminated immunotherapy approaches that target antigens associated with T1D and have shown positive results [Pang et al., 2025]. In humans, an emulsion-based adjuvant nanoparticulate combined with the insulin B chain is being evaluated in a phase 1 trial of T1D (ClinicalTrials.gov ID: NCT03624062). Engineered DNA plasmids that encode proinsulin also proved the ability to suppress proinsulin-specific CD8⁺ T cells and preserve C-peptide secretion in a phase 1 trial of T1D and are waiting for validation in larger studies [Roep et al., 2013].

1.4.4 β cell survival

Verapamil is a calcium channel blocker being used as an antihypertensive agent. In T1D murine models, β cells overexpress thioredoxin-interacting protein (TXNIP) as the disease progresses and in turn induce β cell apoptosis [Xu et al., 2012]. Surprisingly, oral verapamil administration can reduce TXNIP expression, promote β cell survival and improve endogenous insulin production. A consistent effect was seen in two phase 2 clinical trials of new-onset T1D in adults as well as children and adolescents [Ovalle et al., 2018, Forlenza et al., 2023]. Currently, verapamil is being studied both alone (ClinicalTrials.gov ID: NCT04545151) and in combination with ATG (ClinicalTrials.gov ID: NCT06455319).

1.4.5 Others

Tyrosine kinase inhibition: Tyrosine kinases are a broad category of enzymes that catalyse the phosphorylation of tyrosine in proteins, and Janus kinases (JAK) are a family of non-receptor tyrosine kinases. Tyrosine kinases have multiple functions in cellular signalling, and the inhibitor – imatinib – has been used to treat cancer. JAKs are key components of the JAK/STAT signalling pathway,

which is involved in various inflammatory and autoimmune diseases. A selective JAK inhibitor – baricitinib – is used for the treatment of rheumatoid arthritis and alopecia areata. In two phase 2 trials among recent-onset T1D patients, both imatinib and baricitinib demonstrated efficacy in preserving β cell function, although the effect of imatinib was not sustained out to 24 months [Gitelman et al., 2021, Waibel et al., 2023]. Currently, another two JAK inhibitors (abrocitinib and ritlecitinib) are under investigation in a phase 2 trial in recent-onset T1D patients (ClinicalTrials.gov ID: NCT05743244).

1.4.6 Stem cell therapy

Beyond immunotherapy aimed at protecting residual β cells from immune attack, an alternative therapeutic strategy in T1D is the replacement or regeneration of β cells. While pancreatic islet transplantation can effectively restore endogenous insulin production, its broader application is limited because of insufficient donor supply. This shortage has driven extensive research into generating insulin-producing cells from renewable sources, particularly pluripotent stem cells, with the goal of providing an unlimited source for β cell replacement.

One of the most notable advancements is VX-800 (zimislecel), a therapy consisting of fully differentiated islet cells derived from pluripotent stem cells. In a recent phase 1–2 clinical trial involving 14 T1D patients with impaired awareness of hypoglycaemia and severe hypoglycaemia events, VX-800 was transplanted into the portal system. Participants received either a half dose (4×10^8 cells) for safety evaluation or a full dose (8×10^8 cells) for safety and efficacy assessment [Reichman et al., 2025]. Remarkably, all participants regained endogenous insulin production. Among those receiving the full dose, all achieved good glycaemic control, and 10 out of 12 maintained insulin independence one year post-transplantation. VX-800 is now progressing to a phase 3 clinical trial. However, this approach requires concomitant immunosuppression, limiting its widespread applicability. To address this challenge, two modified approaches are under investigation: VX-264, which encapsulates the same VX-800 cells within an immunoisolation device (ClinicalTrials.gov ID: NCT05791201), and VCTX210A, which combines genetically engineered immune-evasive pancreatic endoderm cells with a removable protective device (ClinicalTrials.gov ID: NCT05210530).

Autologous strategies have also been explored. In a phase 1 trial, adipose-derived mesenchymal stromal cells from a patient with T1D were reprogrammed

into pluripotent stem cells, differentiated into functional islets, and transplanted back into the same individual [Wang et al., 2024]. Insulin dependence was achieved at 75 days after transplantation and maintained for one year. However, because this patient was already under regular immunosuppression due to prior liver transplantation, it remains unclear how such an approach would perform in the broader T1D population. Furthermore, this evidence is limited to a single case, and long-term safety and efficacy remain to be unveiled.

Another innovative strategy involves the use of genetically modified allogeneic donor islets. UP421, consisting of gene-edited islets transplanted intramuscularly, recently demonstrated promising results: a T1D patient exhibited stable endogenous insulin production without immunosuppression over 12 weeks after transplantation [Carlsson et al., 2025]. Similarly, this too remains a single case report with unknown long-term outcomes in a wider population.

In addition, mesenchymal stem cells (MSCs) have been studied for their dual immunomodulatory and regenerative functions. An umbrella review concluded that MSC-based treatments can improve C-peptide levels and lower HbA1c, although their impact on exogenous insulin requirements is inconsistent across studies [Koishybayeva et al., 2025].

In summary, current disease-modifying therapies for T1D have yielded some success but are still at very early stages; more clinical trials are warranted to confirm their effects. The majority of disease-modifying therapies focus on preserving β cell function in stage 3 T1D patients, and more therapeutic targets are needed to prevent the progression to T1D in high-risk individuals at stage 1 or 2.

1.5 Current challenges and motivation

Across all drug development programmes that reached clinical phases between 2011 and 2020, only 7.9% of molecular entities ultimately received FDA approval [Thomas et al., 2021]. A separate analysis of monotherapy programmes since 2000 found that drugs with genetic support for causality in the target indication pair are 2.6-fold more likely to succeed compared with those lacking such evidence [Minikel et al., 2024]. This enrichment appears particularly important for disease-modifying therapies, where targets are more often disease-specific, in contrast to symptom-managing agents that frequently act pleiotropically across multiple indications.

In the context of T1D, despite significant advances in the discovery of GWAS hits and rare variants, translation of genetic findings into effective therapeutic targets has been limited. For instance, the HLA region is not druggable due to its high polymorphism, difficulty in binding with small molecules, and essential roles in immune responses. Among all the above disease-modifying therapies, only treatment with CTLA4-Ig, low-dose IL-2 and IL-6 inhibitors is supported by genetic evidence. This limited translation might stem from two fundamental challenges: (1) the complexity of attributing causal genes for GWAS loci and (2) the difficulty of characterising the functional mechanisms mediating the effects of the genetic variants.

GWAS loci typically contain multiple variants in strong LD with each other, all of which are statistically associated with the trait. A range of post-GWAS strategies have been developed to interpret these loci and prioritise the most probable causal variants and effector genes. Fine-mapping, which uses conditional association analysis or Bayesian modelling to differentiate causal variants from neighbouring non-causal variants, is usually the first step [Uffelmann et al., 2021]. However, results from different fine-mapping tools often show limited concordance, and most loci still yield a small set of candidate variants rather than a single definitive causal SNP [Pérez-Granado et al., 2022, Uffelmann et al., 2021]. Because the majority of fine-mapped variants lie in non-coding regions, researchers frequently attribute signals to the nearest gene(s), which may not be the true effector gene(s) [Uffelmann et al., 2021]. Alternative approaches attempt to link GWAS variants to target genes through three-dimensional chromatin contacts, proximity to regulatory elements, or colocalisation with molecular quantitative trait loci (QTLs) [Costanzo et al., 2025]. These approaches generally assume that the genes of interest, i.e., the genes driving the pathogenic processes leading to disease, are the genes located within or near GWAS-associated loci, and thus focus their attention on *cis*-effects of GWAS-identified genetic variants on molecular traits. Despite the availability of multiple colocalisation methods, their results are often inconsistent due to differing assumptions, with a recent benchmarking study showing that the most widely used colocalisation methods were outperformed by the simple heuristic of assigning the signal to the closest gene when trying to colocalise protein-QTLs with expression-QTLs [Tambets et al., 2024].

Mendelian randomisation (MR) is an approach that uses genetic variants as instrumental variables to infer the causal effects of exposures on outcomes [Burgess et al., 2024]. Most MR methods make three common assumptions that must

be met to get a valid causal estimate. Genetic variants must (1) be associated with the exposure (relevance); (2) not be associated with the outcome via confounding pathways (exchangeability); and (3) influence the outcome only through the exposure (exclusion restriction). While relevance can be verified empirically, the latter two assumptions are difficult to prove. Horizontal pleiotropy — where instruments affect the outcome through pathways not mediated by exposure — is the major challenge for causal inference in MR. In addition, it has been reported that the most commonly used inverse-variance weighted (IVW) model tends to underestimate standard errors and inflate statistical significance [Tambets et al., 2024].

Together, these challenges motivate the development of better strategies and improved methods of genetic analysis for identifying causal genes for T1D. For complex traits, it has been observed that (1) genome-wide significant loci typically account for only a small fraction of the total trait heritability; (2) a large proportion of the heritability is attributable to a polygenic background of common variants with tiny individual effect sizes that are spread across the genome; (3) genes with relevant biological functions contribute only slightly more to the heritability than random genes; (4) the heritability is primarily driven by common variants, while rare variants with large effects usually do not explain much of the overall phenotypic variance; (5) the majority of heritability lies in non-coding variants, especially those in gene regulatory regions; (6) SNPs in tissue-specific regulatory elements explain only modestly more heritability than those in broadly active regulatory elements [Boyle et al., 2017, Liu et al., 2019]. These observations gave rise to the “omnigenic” model, which proposes that most of the genetic effects on a typical complex trait are mediated through weak *trans*-effects of common variants that coalesce on the expression of a relatively sparse set of “core genes” in relevant tissues. This is supported by empirical evidence that ~70% of the heritability for gene expression is attributable to *trans*-effects [Ouwens et al., 2020]. Under the omnigenic model, core genes are proximal to the phenotype with direct effects, whereas all the other genes are considered as “peripheral genes” and only affect the phenotype indirectly through the regulation of core genes. If a peripheral gene can regulate multiple downstream core genes, it is termed a “peripheral master regulator”.

In T1D, whilst around half of the heritability is attributable to the HLA region, a substantial contribution also arises from outside the HLA region across the genome. Leukocytes in blood are a relevant tissue for T1D, and considerable

resources now exist for evaluating *trans*-effects on gene expression and circulating protein levels [Iakovliev et al., 2023]. Motivated by the omnigenic model, our research group developed the genome-wide aggregated *trans*-effects (GATE) method, which aggregates *trans*-effects of common SNPs on gene expression and on protein levels to pinpoint core effector genes and proteins for complex diseases. For each individual in a case-control dataset, a genotypic score for predicted gene expression or protein level based on *trans*-effects of common SNPs is computed for each gene or protein and tested for association with disease. An initial application of GATE analysis to T1D identified nine putative core genes based on *trans*-effects on gene expression and 14 putative core proteins based on *trans*-effects on circulating protein levels. The results provided a strong foundation for further refinement of the method and validation and expansion of the findings.

Computationally, GATE analysis is similar to the transcriptome-wide association study (TWAS): a genotypic score reflecting the predicted level of gene expression based on genotypes is constructed for each gene and tested for association with the disease. However, TWAS are usually restricted to *cis*-acting variants [Gamazon et al., 2015, Gusev et al., 2016], similar to colocalisation analysis and to MR studies where the exposure is a molecular QTL [Zheng et al., 2020, Schmidt et al., 2020, Burgess et al., 2023]. Even in TWAS studies that are not restricted to *cis*-QTLs, usually only a few unlinked *trans*-QTLs are included for each transcript [Luningham et al., 2020, Head et al., 2024]. Because disease-relevant genes are depleted of *cis*-eQTLs that have effects large enough to be detectable [Mostafavi et al., 2023], restriction to *cis*-eQTLs focuses investigation on the genes that are least relevant to disease risk. On the contrary, GATE analysis focuses on *trans*-effects of common variants on expression of each gene. *Trans*-effects are generally weak and pleiotropic, and thus less rapidly removed by purifying selection.

Another approach motivated by the omnigenic model is the “expression quantitative trait score” (eQTS) method, which constructs polygenic risk scores for a given complex trait using summary-level GWAS data and tests the association of the score with measured gene expression [Võsa et al., 2021]. However, a key limitation of the eQTS method is that the inclusion of SNPs affecting the phenotype through other pathways into the polygenic risk score will dilute the association between the polygenic risk score and gene expression, thereby reducing statistical power to detect core genes. In contrast, by directly aggregating

weak *trans*-effects across the genome, the GATE method achieves higher statistical power to detect core genes and proteins relevant to the disease. A second practical limitation for the eQTS method is the limited availability of individual-level datasets with genotypes and measured gene expression of large enough sample sizes.

1.6 Aims and objectives

The aim of my PhD was to harness the novel GATE method of genetic analysis and newer GWAS summary data from large-scale proteomics studies that became available at the onset of my PhD to systematically identify core genes and proteins that directly influence the pathogenesis of T1D, thus providing human genetic evidence to support development and facilitate prioritisation of therapeutic targets for T1D.

To achieve this aim, my PhD had the following specific objectives:

1. To identify putative core proteins for T1D using the GATE analysis, by computing GATE scores for plasma proteins with summary statistics for SNP-protein associations from three large proteomics studies; testing the associations of these GATE scores with T1D in two case-control studies; and strengthening causal inference with a novel MR approach robust to horizontal pleiotropy and a predefined causality assessment framework.
2. To identify potential core proteins for T1D using a cross-sectional analysis of measured protein levels from the UK Biobank proteomics study and a confirmatory GATE analysis in two case-control studies as in Objective 1, and evaluate proteins showing a consistent direction of effect with respect to their measured protein level and their genetically predicted level using the same MR approach and causality assessment framework as in Objective 1.
3. To evaluate if transcripts and protein levels for proteins identified in Objectives 1 and 2 were dysregulated during progression to IA or T1D by comparing individuals who progressed to those who did not in a nested case-control dataset from the TEDDY cohort.

2

Methods

This chapter describes the details of GATE methods and datasets used in the analyses of Chapter 3 and part of Chapter 4. Data sources and methods used for other analyses will be presented separately in corresponding chapters.

2.1 T1D target datasets

Two T1D target datasets were available with individual-level genome-wide SNP data and T1D phenotype data.

2.1.1 Scottish dataset

The Scottish dataset includes T1D cases from the Scottish Diabetes Research Network Type 1 Bioresource (SDRNT1BIO) cohort and controls without diabetes from the background population obtained from the Generation Scotland: the Scottish Family Health Study (GS:SFHS).

The SDRNT1BIO cohort is a consented cohort of 6127 people clinically diagnosed with T1D and recruited between 2010 and 2013, which comprises one-third of the adult population with T1D in Scotland [Akbar et al., 2017]. The inclusion criteria were defined as age ≥ 16 years, not pregnant at recruitment, clinical diagnosis of T1D or monogenic diabetes or latent autoimmune diabetes in adults, insulin initiation within one year of diagnosis, and current use of insulin for those who were diagnosed as T1D. Baseline data was linked to the Scottish Care Information Diabetes Collaboration electronic health records (EHRs) via the Community Health Index (CHI) number to capture both retrospective and prospective clinical data. GADA, IA-2A, ZnT8A, and plasma C-peptide were measured in this cohort [McKeigue et al., 2019]. Participants with plasma C-peptide > 600 pmol/L and negative for all three autoantibodies were classified as “possible type 2” and excluded. Definite T1D cases were defined as $\geq$ two insulin prescriptions within one year of diagnosis and no reported diagnosis of monogenic diabetes on the questionnaire.

The GS:SFHS is a family-based, intensively phenotyped cohort representing the general population across Scotland. Participants and their relatives were recruited mainly through primary care [Smith et al., 2006]. The eligibility criteria were age between 35 and 65 years at recruitment, at least one first-degree relative aged ≥ 18 years and at least one full sibling group in the participating family group. Similarly, baseline data was linked to routine healthcare datasets

through the CHI number. Approximately 24,000 volunteers aged 18-98 years were enrolled between 2006 and 2011 [Smith et al., 2013]. Controls without diabetes were defined as individuals with no record of diabetes based on self-report, linkage to hospital records, or any prescription of insulin or anti-diabetic oral medications.

The SDRNT1BIO and GS:SFHS cohorts were genotyped with the Illumina Human Core Exome 24 and the Illumina OmniExpressExome 8V 1-2A bead array, respectively [Iakovliev et al., 2023]. A standard pre-imputation quality control (QC) pipeline was applied at the variant level and sample level. SNPs were filtered out based on call rate, minor allele frequency, and Hardy-Weinberg equilibrium. Samples were excluded based on missing rates and heterozygosity. Genotypes from case and control cohorts were merged and phased for chromosomes 1 to 22 and X. Imputation was conducted against the UK10K and 1000 Genomes reference panels with the Sanger Institute's online imputation service.

2.1.2 UK Biobank

UK Biobank is a population-based biomedical cohort that includes genetic and phenotypic information from about 500,000 UK participants aged 40–69 at enrolment between 2006 and 2010 [Bycroft et al., 2018]. Baseline data was linked to EHRs, death registries, and cancer registries. I defined T1D case/control status based on two sources: (a) “first occurrence” data, which was in turn based on the self-reported, primary care, and secondary care data and coded using the International Classification of Diseases (ICD-10) code E10; (b) self-reported data relevant to diabetes collected by questionnaires and data relevant to insulin use collected by verbal interview at baseline. Specifically, T1D cases were defined as individuals who were under prescription of insulin at baseline and either had a diagnosis of T1D before the age of 50 in the “first occurrence” data or self-reported a diagnosis of diabetes before the age of 50 with initiation of insulin within 1 year of diagnosis. Non-diabetic controls were defined as individuals who had no records of diabetes nor were under a prescription for insulin based on the two aforementioned sources.

The UK Biobank genetic data includes genotypes from 488,377 participants, which were obtained using two very similar genotyping arrays: the Applied Biosystems UK BiLEVE Axiom Array by Affymetrix for a subset of 49,950 participants and the Applied Biosystems UK Biobank Axiom Array for the remaining 438,427

participants [Bycroft et al., 2018]. Pre-imputation QC included tests for consistency in genotype calling across batches and plates, as well as assessments of Hardy-Weinberg equilibrium at the SNP level, missing rates and heterozygosity at the sample level. Genotypes were phased for autosomes and imputed against the Haplotype Reference Consortium and the merged UK10K and 1000 Genomes phase 3 reference panels.

For both the Scottish dataset and UK Biobank, imputed genotypes were lifted to genome build hg38 for the following analyses. Additional pruning was performed to eliminate any pairs of individuals exhibiting a kinship coefficient exceeding 0.05, and principal components (PCs) were calculated among the unrelated individuals.

2.2 Computation of genome-wide aggregated *trans*-effects (GATE) scores for proteins in target datasets

2.2.1 Constructing locus-specific vectors of weights of SNPs associated with proteins

Publicly available regression coefficients for the association of genome-wide SNP data with protein levels (pQTL GWAS summary statistics) from three proteomics studies were used to construct locus-specific vectors of weights for SNPs associated with proteins: (1) the INTERVAL study, which measured 2,994 circulating proteins on the SomaLogic panel based on plasma samples from 3,301 participants [Sun et al., 2018]; (2) the deCODE study, which measured 4,719 plasma proteins on the SomaLogic v4 panel in 35,559 individuals [Ferkingstad et al., 2021]; (3) the UK Biobank Pharma Proteomics Project (UKB-PPP) study, which measured 2,923 plasma proteins on the Olink Explore 3072 panel in 54,219 UK Biobank participants [Sun et al., 2023]. A *trans*-pQTL originating at the transcription site of *CFH*, encoding complement factor H, was detected for 1,975 proteins in the deCODE study. This is likely caused by cross-reactivity between complement factor H and the respective aptamers [Pietzner et al., 2020]. As this cross-reactivity may induce noise in the measurement of the protein level, I excluded these 1,975 proteins from the following analyses.

For each protein, the univariate coefficients of regression between SNPs and protein level, restricted to SNPs either typed or imputed in each target dataset,

were extracted from the summary statistics of each pQTL GWAS together with their p -values. SNPs were grouped into loci, where a locus was defined as containing at least one SNP with a p -value $< 10^{-6}$ and separated by at least 0.5 megabase (Mb) from other such loci. For each locus, the regression coefficients for all SNPs in that locus with a p -value $< 10^{-5}$ were used to generate a “locus-specific” weights vector. This was done by premultiplying the vector of univariate coefficients for the selected SNPs in that locus by the inverse correlation matrix between SNP genotypes, computed in the 1000 Genomes reference panel. This operation approximates the multivariate coefficients that would be obtained by fitting all the selected SNPs in that locus jointly in a multiple regression analysis of protein levels if individual-level data from the pQTL GWAS dataset were available. The operation was implemented using a pseudo-inverse solution to handle ill-conditioned matrices in loci with highly correlated SNPs. This process was implemented in the GENOSCORES platform [Spiliopoulou et al., 2019], generating three lists of locus-specific weight vectors (INTERVAL, deCODE, and UK Biobank) for each of the two target datasets (the Scottish dataset and UK Biobank).

2.2.2 Computation of scores in target datasets

A GATE score for each protein was computed for each individual in the Scottish and UK Biobank datasets by adding up their genotypes at *trans*-pQTLs across the genome, weighted by the vector of locus-specific weights derived in the above step for each locus. Each locus was classified as *cis* if the distance from the locus to the transcription site of the respective gene encoding that protein was less than or equal to 50 kb, *cis-x* if the distance was between 50 kb and 5 Mb, and *trans* if the distance was more than 5 Mb. Only genotypes at *trans*-loci were included in the GATE score. Genotypes at the *cis*-locus for each protein, weighted by the corresponding locus-specific weights vector, were added up to compute a locus-specific *cis*-score for that protein. The start and end positions of the transcription site of each gene were obtained from Ensembl [Martin et al., 2023]. Mapping of proteins to the gene encoding them was obtained from UniProt [The UniProt Consortium, 2023]. As the HLA region is a hotspot for *trans*-QTLs for genes involved in immunity and inflammation [Fehrmann et al., 2011], and associations of these *trans*-QTLs with T1D are heavily confounded by the direct effects of HLA antigens, the HLA region (from 25 to 34 Mb on chromosome 6) was excluded from the computation of genome-wide *trans*-scores.

2.3 Testing association of GATE scores with T1D

2.3.1 Statistical model for test of association of GATE scores with T1D

To test the association between GATE scores and T1D, I used a logistic regression model adjusted for age in the year 2016, sex, and the first three PCs in the Scottish dataset and age at assessment, sex, and the first 20 PCs in the UK Biobank dataset. As the pQTLs are derived from summary statistics of European ancestry, the GATE analysis in both datasets was restricted to participants with European ancestry. After applying these exclusions, the sample size for the Scottish dataset was 4964 T1D cases and 7474 non-diabetes controls; the sample size for the UK Biobank dataset was 1638 cases and 365K non-diabetes controls when testing GATE scores computed from the UKB-PPP pQTL GWAS study (UKB-PPP participant data were used to learn the SNP effects on protein levels and therefore excluded from the GATE-T1D association analysis) and 1864 cases and 408K non-diabetes controls when testing GATE scores computed from the other pQTL GWAS studies. Similarly, I also tested the association of *cis*-scores for proteins with T1D.

The proteomics studies and the target datasets used in the GATE analysis recruited participants independently. As deCODE was conducted among Icelanders, overlap with UK-based datasets is highly unlikely. When testing for GATE scores computed from UKB-PPP in UK Biobank, those who participated in UKB-PPP were explicitly removed. Potential overlap may remain between the Scottish dataset and UKB-PPP. Ideally, such individuals would be identified (e.g., by constructing a genetic relationship matrix across studies) and excluded from the GATE analysis in the Scottish dataset. However, this is not feasible because the individual-level data are held in different trusted environments and cannot be transferred without additional governance approvals. Overall, inflation of type 1 error in the GATE analysis is expected to be low.

2.3.2 Filtering criteria for reporting GATE associations with T1D

To select GATE scores associated with T1D, I filtered on two criteria: (1) *p*-value less than 1×10^{-6} ; (2) at least 5 effective *trans*-loci contributing to the GATE score. The rationale for this is that if a GATE score is dominated by one or

more *trans*-pQTLs with a direct effect on T1D not mediated through that protein, the association could be confounded. The greater the diversity of locus-specific *trans*-scores contributing to a GATE score, the less likely this confounding is. This diversity was quantified as the effective number of loci, which was calculated from the variances of locus-specific *trans*-scores [Iakovliev et al., 2023]. For example, there are K locus-specific *trans*-scores contributing to a GATE score, with their variances being $\sigma_1, \dots, \sigma_K$; the effective number of loci is calculated as $2^{-\sum p_i \log_2 p_i}$, where $p_i = \sigma_i^2 / \sum \sigma_i^2$. Therefore, the effective number of loci reflects both the number of contributing loci and the variance contributed by each locus. Its value ranges from 1, if one of the pQTLs contributing to the GATE score has much larger variance than the others, to K if the variances of the locus-specific *trans*-scores are equal.

2.4 Evaluation of supportive evidence for causality of GATE score associations with T1D

To evaluate causality over confounding for GATE score associations with T1D, a series of criteria was pre-specified. The first two of these, described below, are based on *trans*-pQTLs, and the others are independent assessments that are not based on *trans*-pQTLs.

1. The *trans*-pQTLs contributing to the GATE score do not contribute much to GATE scores for other proteins that are also associated with T1D. To assess this visually, I computed correlations between the GATE scores associated with T1D among non-diabetic controls, reordered the rows and columns of the correlation matrix to make it block-diagonal and examined a heatmap of this matrix to identify blocks corresponding to GATE scores with pleiotropic effects. I also grouped *trans*-pQTLs that contributed to the GATE scores associated with T1D into clumps if their genomic locations overlapped or were within 200 kb of each other and examined which clumps contained *trans*-pQTLs shared by many proteins, i.e., “peripheral master regulators”.
2. There is a dose-response relationship between the effect of the *trans*-pQTLs on the protein and the effect of the *trans*-pQTLs on T1D. A two-sample Mendelian randomisation (MR) analysis is used to test such a dose-response relationship. Specifically, a scalar genetic instrument from each *trans*-pQTL

was constructed; a regularised horseshoe prior to the direct (pleiotropic) effects of the instruments on the outcome was set, and the likelihood of the causal effect parameter was computed by marginalising over the posterior distribution of these pleiotropic effects [McKeigue et al., 2024]. This procedure controls for confounding, even though the confounders (direct effects of instruments on outcomes) are not measured.

3. Rare variants in the gene encoding GATE-associated proteins or in other genes in the same pathway cause monogenic autoimmune diabetes.
4. Association of T1D of *cis*-score. Recent studies have shown that disease-relevant genes are enriched with redundant enhancer domains and depleted of *cis*-QTLs of large effect, as would be expected if perturbation of these genes has large effects on fitness [Mostafavi et al., 2023]. Therefore, it is not expected to detect many associations with *cis*-scores for disease-relevant genes. However, if an association between the *cis*-score and T1D is detected, this is independent, supportive evidence of the role of the gene in T1D since, by design, the *cis*-signal does not contribute to the GATE score.
5. Association of T1D with common variants within 200 kb of the transcription site of the gene encoding GATE-associated proteins in published GWAS. I extracted previously reported GWAS hits from two meta-GWAS of T1D; one included 16,159 unrelated T1D cases, 25,386 unrelated controls and 6,143 affected trio families (i.e., an affected child and both of his/her parents) [Robertson et al., 2021], and the other included 18,942 T1D cases and 501,638 controls [Chiou et al., 2021]. This resulted in a list of 98 genomic regions with T1D GWAS hits.
6. Association with T1D with measured levels of the proteins in the UKB-PPP study (N=232), which was tested using a logistic regression model adjusted for age at assessment and sex. In line with the GATE analysis, only individuals with European ancestry were included in this analysis. Whilst proteins involved in the causal pathway of T1D will not necessarily show association in a prevalent case-control dataset, finding such an association supports the relevance of the protein to disease pathogenesis. Because GATE scores typically explain only a small proportion of the variance of the measured protein, the association of disease with the measured protein should be far stronger than the association with the GATE score if the association with the protein is causal.

7. Published experimental evidence that perturbations in the gene (e.g., knock-out, transduction or overexpression) or perturbations in the gene product with an inhibitor or agonist can alter the development of T1D in animal models (e.g., non-obese diabetic mice, NOD mice). The PubMed database was searched with the search term as follows: “(gene symbol [other term] OR gene symbol *abstract* OR protein name [other term] OR protein name *abstract*) AND (type 1 diabetes) AND (NOD mice)”.
8. Drugs targeting the gene product, its ligands, or receptors cause or ameliorate T1D. This was investigated by querying the DrugBank [Knox et al., 2024], ClinicalTrials.gov, and the ChEMBL databases [Zdrazil et al., 2024].

2.5 Ethical considerations

This thesis did not involve any direct contact with participants or generation of new biological samples. All analyses were conducted using existing datasets (i.e., SDRNT1BIO, GS:SFHS, UK Biobank and TEDDY), for which informed consent had been obtained, and appropriate ethical approvals were granted by relevant research ethics committees. Access to these datasets was authorised by respective data governance bodies. All participant data were de-identified, and no attempts were made to re-identify individuals. Analyses were performed within secure computing environments, and only summary-level results were reported. In addition, care was taken in the interpretation and presentation of findings in the following chapters of the thesis to avoid genetic determinism or overstatement of clinical implications.

3

Identification of putative core proteins
for T1D using the GATE analysis

3.1 Introduction

This chapter aimed to use the novel genetic analysis method, genome-wide aggregated *trans*-effects (GATE) analysis, to identify core proteins proximal to T1D pathogenesis and provide human genetic evidence for potential drug targets. Summary statistics for SNP-protein associations in three large cohorts, UKB-PPP, deCODE, and INTERVAL, were used to compute GATE scores for plasma proteins, and these scores were then tested for association with T1D in a Scottish dataset comprising 4964 T1D cases from the SDRNT1BIO cohort and 7474 controls without diabetes from the GS:SFHS cohort and a held-out dataset with 1638 T1D cases and 365K controls without diabetes in the UK Biobank. Putative core proteins were selected based on the strength of the association (p -value $< 1 \times 10^{-6}$) and the number of effective *trans*-loci (at least 5) contributing to the GATE score. A series of criteria were prespecified to assess the causal roles of proteins whose GATE scores were associated with T1D.

This chapter has been published in the journal *Diabetes* as “Xuan Zhou, Andrii Iakovliev, Stuart McGurnaghan, Buddhiprabha Erabadda, Caroline Hayward, Paul M. McKeigue, Athina Spiliopoulou, Helen M. Colhoun; Genome-Wide Aggregated Trans Effects Analysis for Circulating Proteins Indicates a Key Role of Immune Checkpoints in Type 1 Diabetes. *Diabetes* 2025; db250067. <https://doi.org/10.2337/db25-0067>”.

As the joint first author of this published work, I contributed to the methodology of the study, data curation and analysis, validation, visualisation, interpretation of results, drafting of the manuscript, and addressing reviewers’ comments. The initial software code for the analysis pipeline was developed by my supervisors (Prof Paul McKeigue and Dr Athina Spiliopoulou) and two postdoctoral researchers (Dr Andrii Iakovliev and Dr Buddhiprabha Erabadda) in the group. During my PhD, we further developed the methodological framework and refined the pipeline to ensure greater robustness and reproducibility, to which I contributed through implementation and testing.

For this specific paper, I computed scores based on the new GWAS summary statistics for SNP-protein associations and conducted the analysis in the Scottish dataset. I adapted the pipeline to run in the UK Biobank dataset, which served as an independent replication cohort. I also wrote code to systematically assess the causality for all putative core proteins identified in both datasets using the novel Mendelian randomisation (MR) approach developed by our team. In

addition, I applied conventional MR methods, including the fixed- and random-effects inverse variance weighted (IVW) and weighted median approaches, and compared their results with those obtained using the novel method. Although these results were not included in the manuscript, they are discussed at the end of this chapter. I also conducted an extensive literature review to annotate the biological functions for each putative core protein and to collate supporting evidence from experimental animal models and human studies. Having done all the analyses, I summarised the results, wrote the first draft of the manuscript, created figures and tables, and formatted it in accordance with the journal's submission guidelines. My supervisors and our collaborator (Prof Helen Colhoun) provided feedback and revised the manuscript.

3.2 Accepted manuscript

3.2.1 Abstract

The “omnigenic” hypothesis postulates that polygenic effects of common variants on typical complex traits coalesce via *trans*-effects on the expression of a relatively sparse set of “core” effector genes and their encoded proteins in relevant tissues. The objective of this study was to identify core proteins for type 1 diabetes. We used summary statistics for SNP associations with plasma levels of 5130 proteins in three large cohorts to compute genome-wide aggregated *trans*-effects (GATE) scores for protein levels in two type 1 diabetes case-control studies (6828 cases, 416,000 controls). GATE scores for 27 proteins were associated with type 1 diabetes. Fourteen were replicated between datasets, 11 had support in Mendelian randomisation analysis and 9 had experimental support in mouse models of autoimmune diabetes. The strongest associations were for immune checkpoints (*PDCD1*, *CD5*, *TIGIT*, *LAG3*), chemokines and innate immune system proteins (*NCR1*, *KLRB1*). Whilst *PDCD1* is a known cause of monogenic autoimmune diabetes, neither it nor most of the core proteins identified here were previously reported as GWAS hits for type 1 diabetes. These results identify possible drug targets with genetic support for causality and suggest that PD-1 agonists under development for other indications should be trialled for type 1 diabetes prevention.

3.2.2 Article Highlights

- Demonstrating genetic evidence for a role of a protein in disease gives important support for its potential as a drug target.
- We aimed to identify proteins that have genetic evidence to support a causal role in the pathogenesis of type 1 diabetes.
- 27 core proteins had genetic evidence of causality for type 1 diabetes. Top hits included immune checkpoints (*PDCD1*, *CD5*, *TIGIT*, and *LAG3*) and innate immune system proteins (*NCR1* and *KLRB1*).
- These results identify possible drug targets and suggest that PD-1 agonists should be trialled for type 1 diabetes prevention.

3.2.3 Introduction

Studies of the immunology of type 1 diabetes have yielded many potential drug targets [Herold et al., 2024] but only one licensed disease-modifying drug, teplizumab, that delays progression to stage 3 type 1 diabetes by a median of 32.5 months. In drug development programmes for all diseases, only 8% of molecular entities entering clinical stage gain final approval. Genetic evidence for a drug target, such as from genome-wide association studies (GWAS) or monogenic discovery in humans, increases the probability of success by a factor of 2.6 [Minikel et al., 2024].

Most of the heritability of gene expression is attributable to *trans* rather than to *cis* effects [Liu et al., 2019]. The “omnigenic” hypothesis postulates that the weak polygenic effects of common variants on a typical complex trait coalesce via *trans*-effects on the expression of a relatively sparse set of “core” effector genes and their encoded proteins in relevant tissues [Boyle et al., 2017]. Drug targets are more likely to be found among these core genes that directly influence disease risk than among “peripheral master regulator” genes that have *trans*-effects on expression of many genes.

To identify core genes and proteins, we developed Genome-wide Aggregated *Trans*-Effects (GATE) analysis [Iakovliev et al., 2023]. This uses summary statistics from large genome-wide studies of SNP-transcript associations (eQTL GWAS) or SNP-protein associations (pQTL GWAS) to compute genome-wide aggregated *trans*-scores for the predicted level of each transcript (eGATE) or protein (pGATE) that are tested for association with disease in target case-control datasets. Using eQTL data and a limited set of pQTL data, we previously reported the identification of a sparse set of putative core genes for type 1 diabetes [Iakovliev et al., 2023]. Almost all these genes were in the immune system. Most had not been identified previously as GWAS hits for type 1 diabetes. For some of these genes, causality was supported by perturbation in experimental models affecting diabetes, and some are known genes for monogenic autoimmune diabetes.

The objective of this study was to identify and validate additional core proteins for type 1 diabetes by exploiting additional summary-level data for SNP-protein associations available in the UK Biobank Pharma Proteomics Project (UKB-PPP) [Sun et al., 2023].

3.2.4 Research Design and Methods

Methods for genome-wide aggregated *trans*-effects analysis have been described previously [Iakovliev et al., 2023, Spiliopoulou et al., 2025]. An overview of the study design is shown in Figure 3.1. Ethical approval was obtained from the Tayside Research Ethics Committee (reference: 10/S1402/43) for Scottish Diabetes Research Network Type 1 Bioresource (SDRNT1BIO), the Tayside Committee on Medical Research Ethics for Generation Scotland Family Health Study (GS:SFHS), the North West Centre for Research Ethics Committee (11/NW/0382) for UK Biobank. Informed consent was obtained for all study participants. The work here was approved by the UK Biobank (application 23652).

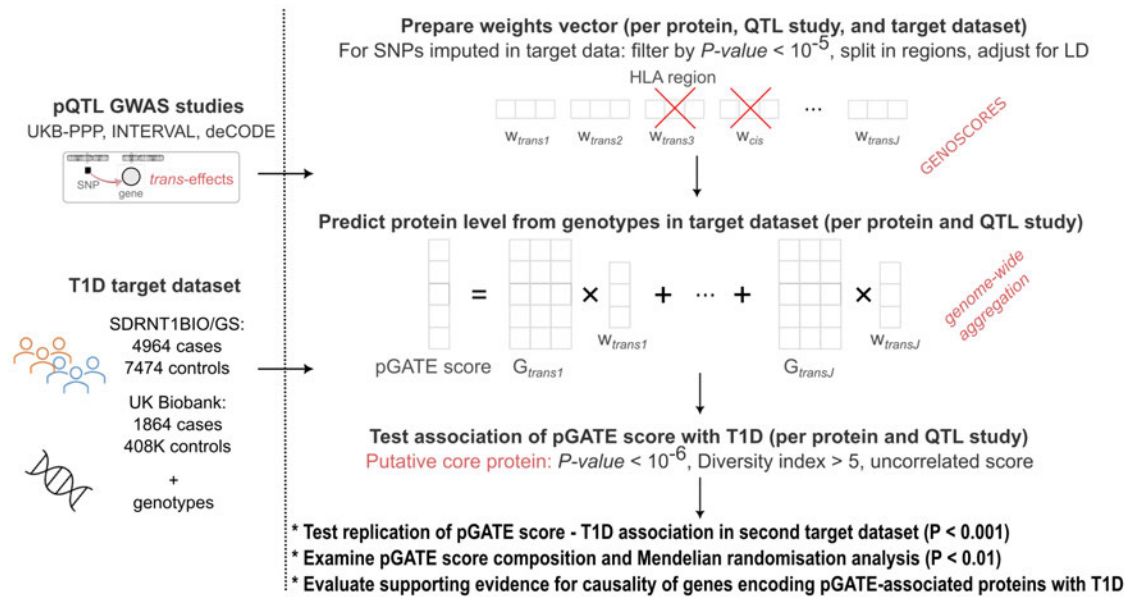


Figure 3.1: Graphic illustration of the study design. Left (input data): summary-level data from pQTL GWAS studies and individual-level genotypic and case/control data from target datasets. Right (analysis steps): pGATE scores are computed using pQTL summary statistics and individual-level genotypes. Computed scores are tested for association with the disease (type 1 diabetes). For identified putative core proteins, additional supporting evidence for causality of the encoding gene with the disease is evaluated.

3.2.4.1 Study populations

Individual-level genome-wide SNP data and type 1 diabetes phenotype data for testing the association of pGATE scores with type 1 diabetes were available from a Scottish case-control dataset and the UK Biobank dataset.

The Scottish cases from the SDRNT1BIO cohort comprise 6127 people clinically diagnosed as having type 1 diabetes and aged ≥ 16 years at recruitment between 2010 and 2013 [Akbar et al., 2017]. Cases were those who started insulin within a year of diagnosis and were not diagnosed with monogenic diabetes. Those with plasma C-peptide greater than 600 pmol/L and negative for three autoantibodies (glutamic acid decarboxylase, tyrosine phosphatase-related islet antigen 2, and zinc transporter 8) were excluded as “possible type 2” [McKeigue et al., 2019]. Scottish controls were from GS:SFHS, a family-based cohort of 24,000 people in Scotland aged ≥ 18 years recruited between 2006 and 2011 [Smith et al., 2006]. Controls were restricted to individuals with no record of diabetes based on self-report, linkage to hospital records, and never prescribed insulin or anti-diabetic oral medications. Genotyping, imputation and quality control procedures for the Scottish dataset have been described previously [Iakovliev et al., 2023].

The UK Biobank cohort comprises approximately 500,000 volunteers from the UK aged 40-69 at recruitment (between 2006 and 2010) [Bycroft et al., 2018]. Cases of type 1 diabetes were defined by a diagnosis of type 1 diabetes before age 50 (based on self-report, hospital admissions or primary care records) and initiation of insulin within 1 year of diagnosis. Genotyping, imputation and quality control procedures for the UK Biobank dataset have been described previously [Bycroft et al., 2018]. Related individuals were removed until no pairs remained with a kinship coefficient exceeding 0.05. Genotypic principal components were calculated from the remaining individuals.

3.2.4.2 Computation of protein genome-wide aggregated *trans*-effects (pGATE) scores in target datasets

We used publicly available regression coefficients for the association of genome-wide SNP data with protein levels (pQTL GWAS summary statistics) from three proteomics studies to construct vectors of weights for SNPs associated with plasma proteins: 2,994 proteins on the SomaLogic panel measured in 3,301 participants in the INTERVAL study [Sun et al., 2018], 4,719 proteins on the SomaLogic v4 panel in 35,559 participants in the deCODE study [Feringstad et al., 2021] and 2,923 proteins on the Olink Explore panel in 54,219 UK Biobank participants – the UK Biobank Pharma Proteomics Project (UKB-PPP) [Sun et al., 2023]. Procedures for quality control of the proteomics data and the estimation of SNP-protein coefficients have been described previously [Sun et al., 2018,

2023, Ferkingstad et al., 2021]. Additionally, 1975 aptamers that appeared to cross-react with *CFH* [Pietzner et al., 2020] were excluded from the deCODE study.

For each region containing at least one SNP associated with the protein at $p < 10^{-6}$, SNPs associated with the protein at $p < 10^{-5}$ were grouped as a quantitative trait locus that was classified as *cis* if the distance from the locus to the transcription site of the respective protein-coding gene was less than 50 kb, *cis-x* if the distance was between 50 kb and 5 Mb, or *trans* if the distance was more than 5 Mb. Univariate regression coefficients were premultiplied by the inverse covariance matrix between genotypes to calculate weights adjusted for linkage disequilibrium. The pGATE score for each protein was computed for each individual in the target dataset as the dot product of the genotypes and the adjusted weights at *trans*-QTLs. Each individual's pGATE score is thus the predicted protein level given their genotype at all *trans*-QTLs. As associations of *trans*-QTLs in the HLA region with type 1 diabetes are heavily confounded by the direct effects of HLA antigens on diabetes, the HLA region (chromosome 6: 25 - 34 Mb) was excluded from the computation of pGATE scores.

3.2.4.3 Testing association of pGATE scores with type 1 diabetes

Tests for association with pGATE scores were restricted to individuals of European ancestry, as the summary statistics for pGATE scores were computed from individuals of European ancestry. For each pGATE score, a logistic regression model with type 1 diabetes as dependent variable was fitted, adjusted for age, sex, and principal components (first three in the Scottish dataset, first 20 in UK Biobank). Associations of type 1 diabetes with *cis*-scores were tested separately. When testing for association of type 1 diabetes with pGATE scores in UK Biobank, individuals who were in the UKB-PPP were excluded. To select putative core proteins, we filtered on two criteria: association of pGATE score with type 1 diabetes at $p < 10^{-6}$, and effective number of *trans*-loci contributing to the pGATE score ≥ 5 [Iakovliev et al., 2023].

3.2.4.4 Assessment of criteria supporting causality over confounding for association of pGATE scores with type 1 diabetes

Association of a pGATE score for a given protein with type 1 diabetes may be confounded by direct effects of genetic variants on the disease that are not

mediated through the protein under study. We assessed criteria that support causality over confounding as follows:

1. The effective number of *trans*-pQTLs contributing to the pGATE score is large: this makes it less plausible that the association is attributable to direct effects of a few “peripheral master regulators”; thus, we filtered associations as described above.
2. The *trans*-pQTLs contributing to the pGATE score do not contribute much to type 1 diabetes-associated pGATE scores for other proteins: To assess this visually, we computed correlations between the diabetes-associated pGATE scores, reordered the rows and columns of the correlation matrix to make it block-diagonal, and examined a heatmap of this matrix to identify blocks corresponding to pGATE scores with pleiotropic effects. We also grouped *trans*-pQTLs that contributed to type 1 diabetes-associated pGATE scores into clumps if their genomic locations overlapped or were within 200 kb of each other and examined which clumps contained *trans*-pQTLs shared by many proteins, i.e., “peripheral master regulators”.
3. There is a dose-response relationship between the effect of the *trans*-pQTLs on the protein and the effect of the *trans*-pQTLs on type 1 diabetes: The method of 2-sample Mendelian randomisation (MR) analysis is effectively a test for such a dose-response relationship. To assess this for each putative core protein, we plotted *trans*-pQTL-disease associations against *trans*-pQTL-protein associations. We tested formally for causality by marginalising over the distribution of (unobserved) direct effects to compute the likelihood of the causal effect parameter given the data. This is similar to a Bayesian method described elsewhere [Grant and Burgess, 2024] except that the model includes a parameter that regularises the size of the nonzero direct effects [Piironen and Vehtari, 2017]. The Stan probabilistic programming language was used to generate the joint posterior distribution. The posterior distribution of the causal effect parameter was divided by the prior to obtain the marginal likelihood, from which a maximum likelihood estimate and *p*-value were obtained.

3.2.4.5 Assessment of other evidence (not based on *trans*-QTLs) supporting causality for associated pGATE scores

We sought evidence of the following:

1. Monogenic autoimmune diabetes caused by rare variants in the gene or in other genes in the same pathway.
2. Association of type 1 diabetes with common variants within 200 kb of the transcription site of the gene, in our dataset or in published meta-GWAS [Robertson et al., 2021, Chiou et al., 2021]. As *cis*-SNPs (within 5 Mb of the transcription site) were excluded from GATE scores, any association of disease with SNPs near the transcription site is independent validation of an effect of the gene.
3. Association of type 1 diabetes (n=232) with measured levels of the protein in the UKB-PPP. Because pGATE scores typically explain only a small proportion of the variance of the measured protein, the association of disease with the measured protein should be far stronger than the association with the pGATE score if the association with the protein is causal.
4. Published evidence that perturbing the gene by knockout, transduction or overexpression, or perturbing the gene product by an inhibitor or an agonist, altered the severity of disease in an experimental model.
5. Drugs targeting the gene product, its ligand or its receptor cause or ameliorate the disease.

3.2.5 Results

3.2.5.1 Association of pGATE scores with type 1 diabetes in the Scottish and UK Biobank datasets

Association of type 1 diabetes with pGATE scores was tested in 4,964 type 1 diabetes cases and 7,474 controls in the Scottish dataset and 1,864 cases and 408K controls in UK Biobank. 2793 pGATE scores for 2774 proteins were computed from UK Biobank Olink pQTL data, 2549 scores for 2475 proteins from deCODE SomaLogic pQTL data, and 3036 scores for 2781 proteins from INTERVAL SomaLogic pQTL data. Some proteins were common to more than one pQTL dataset such that there was more than one score for a given protein in a given dataset. Altogether, 5130 unique proteins were evaluated.

In the Scottish dataset, 25 proteins met our reporting threshold of the pGATE score being associated with type 1 diabetes at p -value $< 10^{-6}$ and having at least 5 effective loci contributing to the score (Table 3.1). The pGATE score for *PDCD1* displayed the strongest association with type 1 diabetes risk in the Scottish dataset (log OR=0.21, p -value = 6×10^{-24}). Of these 25 proteins, 12 had a pGATE score for that protein replicated at p -value < 0.001 in UK Biobank. The direction of effect was consistent in the target datasets. pGATE scores for 2 additional proteins (*TIGIT* and *LAG3*) met reporting criteria in UK Biobank, both of which had a pGATE score for that protein that replicated in the Scottish dataset at p -value < 0.001 . For brevity throughout, we have used HUGO symbols for the genes encoding the pGATE-detected proteins rather than full-length UniProt names. Gene names, protein name synonyms and biological functions are given in Table S3.1.

3.2.5.2 Composition of pGATE scores associated with type 1 diabetes and instrumental variable analysis

Figure 3.2 shows that the correlations between pGATE scores in controls fall into two blocks. One block is formed by pGATE scores for pancreatic and gut associated proteins (*REG1B*, *FGF19*, *TCN1*, *PRSS27*), and the other by pGATE scores for immune-related proteins (e.g., *SDK2* and *CD48*, *CXCL9* and *CXCL10*). Correlations between pGATE scores indicate sharing of *trans*-pQTLs. We grouped overlapping *trans*-pQTLs into clumps, tabulated in Table S3.2 and visualised in Figure 3.3. As highlighted in bold in Table S3.2, many clumps harbour known GWAS hits for type 1 diabetes, with several contributing to *trans*-pQTLs for many proteins, i.e., “peripheral master regulators”, including *PTPN22*, *STAT4*, *CTLA4* and *SH2B3*.

For 11 of the 27 putative core proteins, marginalising over pleiotropic effects supported a dose-response effect of the protein on type 1 diabetes at $p < 0.01$ in either the Scottish or UK Biobank datasets (Table S3.3). Apart from *PRSS27*, all of these are immune-related proteins. Scatter plots of the coefficients for the effects of the *trans*-pQTLs for each protein on type 1 diabetes against their coefficients for the effects on the circulating protein level are shown in Figure S3.1 and Figure S3.2. As shown in Table S3.3, of the remaining 16 proteins without causal support from MR, only five – *SELL*, *LTA*, *TCN*, *CXCL11*, *FCRL6* – met the criterion of effective number of *trans*-pQTLs > 20 for adequate statistical power to detect a dose-response relationship. The conventional random-effects inverse variance weighted (IVW) test for MR was consistent with causality for all 11 proteins other than *PRSS27* (p -value = 0.14) and *NCR1* (p -value = 0.086).

Table 3.1: Putative core proteins for type 1 diabetes based on aggregated *trans*-effects on protein levels (pGATE).

Gene symbol	Transcription site		Scottish (4964 cases) ¹				UK Biobank (1638/1864 cases) ²			
	Chr	Start position (Mb)	pQTL study	Effective number of pQTLs	Log OR	p-value	pQTL study	Effective number of pQTLs	Log OR	p-value
Putative core proteins identified in the Scottish dataset ³										
<i>PDCD1</i>	2	241.85	UKB	30.9	0.21	6×10^{-24}	UKB	20.8	0.21	6×10^{-17}
<i>CXCL9</i>	4	76.00	UKB	11.2	0.19	1×10^{-21}	UKB	7.0	0.09	1×10^{-4}
<i>CXCL10</i>	4	76.02	UKB	14.4	0.16	2×10^{-15}	UKB	6.8	0.12	2×10^{-6}
<i>CRTAM</i>	11	122.84	DeC	10.7	0.16	4×10^{-15}	DeC	9.9	0.12	3×10^{-7}
<i>CCL19</i>	9	34.69	DeC	10.2	0.16	4×10^{-15}	DeC	9.0	0.19	8×10^{-16}
<i>LGALS9</i>	17	27.63	UKB	13.1	0.16	2×10^{-14}	UKB	8.0	0.09	2×10^{-4}
<i>CD5</i>	11	61.10	UKB	16.1	0.14	1×10^{-12}	UKB	10.7	0.12	2×10^{-6}
<i>PRSS27</i>	16	2.71	UKB	6.0	0.13	2×10^{-10}	UKB	6.7	0.03	0.2
<i>FASLG</i>	1	172.66	UKB	15.8	0.12	8×10^{-10}	UKB	17.8	0.08	0.001
<i>IDO1</i>	8	39.90	UKB	11.6	0.12	1×10^{-9}	UKB	6.6	0.10	9×10^{-5}
<i>AGR2</i>	7	16.79	UKB	8.0	-0.12	2×10^{-9}	UKB	7.0	-0.02	0.5
<i>FGF19</i>	11	69.70	UKB	6.2	-0.12	7×10^{-9}	UKB	5.7	-0.05	0.03
<i>SDK2</i>	17	73.33	UKB	5.8	0.12	7×10^{-9}	UKB	5.7	0.09	2×10^{-4}
<i>SELL</i>	1	169.69	UKB	20.1	0.11	4×10^{-8}	UKB	11.4	0.07	0.008
<i>GCG</i>	2	162.14	DeC	8.0	0.11	1×10^{-7}	DeC	7.9	-0.02	0.5
<i>KLRB1</i>	12	9.59	UKB	20.3	0.11	1×10^{-7}	UKB	18.3	0.06	0.01
<i>NCR1</i>	19	54.91	UKB	33.7	0.11	2×10^{-7}	UKB	25.7	0.06	0.02
<i>CD48</i>	1	160.68	DeC	12.2	0.11	2×10^{-7}	UKB	13.2	0.12	9×10^{-7}
<i>CKK</i>	3	42.26	DeC	6.0	0.10	3×10^{-7}	DeC	6.5	-0.04	0.07
<i>CFC1</i>	2	130.59	UKB	12.8	-0.10	5×10^{-7}	UKB	15.3	-0.12	4×10^{-6}
<i>LTA</i>	6	31.57	UKB	11.5	0.10	5×10^{-7}	UKB	9.6	0.07	0.006
<i>TCN1</i>	11	59.85	UKB	6.7	0.10	5×10^{-7}	UKB	7.7	0.02	0.4
<i>CXCL11</i>	4	76.03	UKB	17.3	0.10	6×10^{-7}	DeC	16.3	-0.03	0.2
<i>REG1B</i>	2	79.09	UKB	7.8	-0.10	7×10^{-7}	UKB	6.5	-0.02	0.5
<i>FCRL6</i>	1	159.80	UKB	14.3	0.10	7×10^{-7}	UKB	11.9	0.08	9×10^{-4}
Additional putative core proteins identified in the UK Biobank dataset ³										
<i>TIGIT</i>	3	114.28	DeC	4.9	0.17	7×10^{-17}	DeC	5.9	0.23	2×10^{-24}
<i>LAG3</i>	12	6.77	DeC	15.4	0.07	4×10^{-4}	UKB	18.6	0.13	4×10^{-8}

Chr, chromosome; Mb, megabase; Log OR: log(odds ratio) of type 1 diabetes risk per standard deviation increase in the pGATE score; UKB, UK Biobank Pharma Proteomics Project; DeC, deCODE genetics.

¹For the analysis in the Scottish dataset, 4964 cases (2779 males, 2185 females) and 7474 controls (3173 males, 4301 females) were included.

²For the analysis in UK Biobank, 1638 cases (909 males, 729 females) and 365K controls (162K males, 203K females) were included when testing pGATE scores computed based on the UKB-PPP pQTL GWAS summary statistics, and 1864 cases (1037 males, 827 females) and 408K controls (181K males, 227K females) were included when testing pGATE scores computed from other pQTL GWAS studies.

³Twenty five proteins were identified as putative core proteins for type 1 diabetes in the Scottish dataset (top panel). Twelve of these replicated at p -value < 0.001 in the UK Biobank dataset. An additional two proteins were identified as putative core proteins in UK Biobank and both replicated at p -value < 0.001 in the Scottish dataset (bottom panel). Our reporting criteria for putative core proteins were strength of association (p -value $< 10^{-6}$) with at least 5 effective *trans*-pQTLs contributing to the pGATE score in either dataset. For proteins with multiple pGATE scores (e.g. based on different pQTL GWAS), the score with the strongest association is shown in this table.

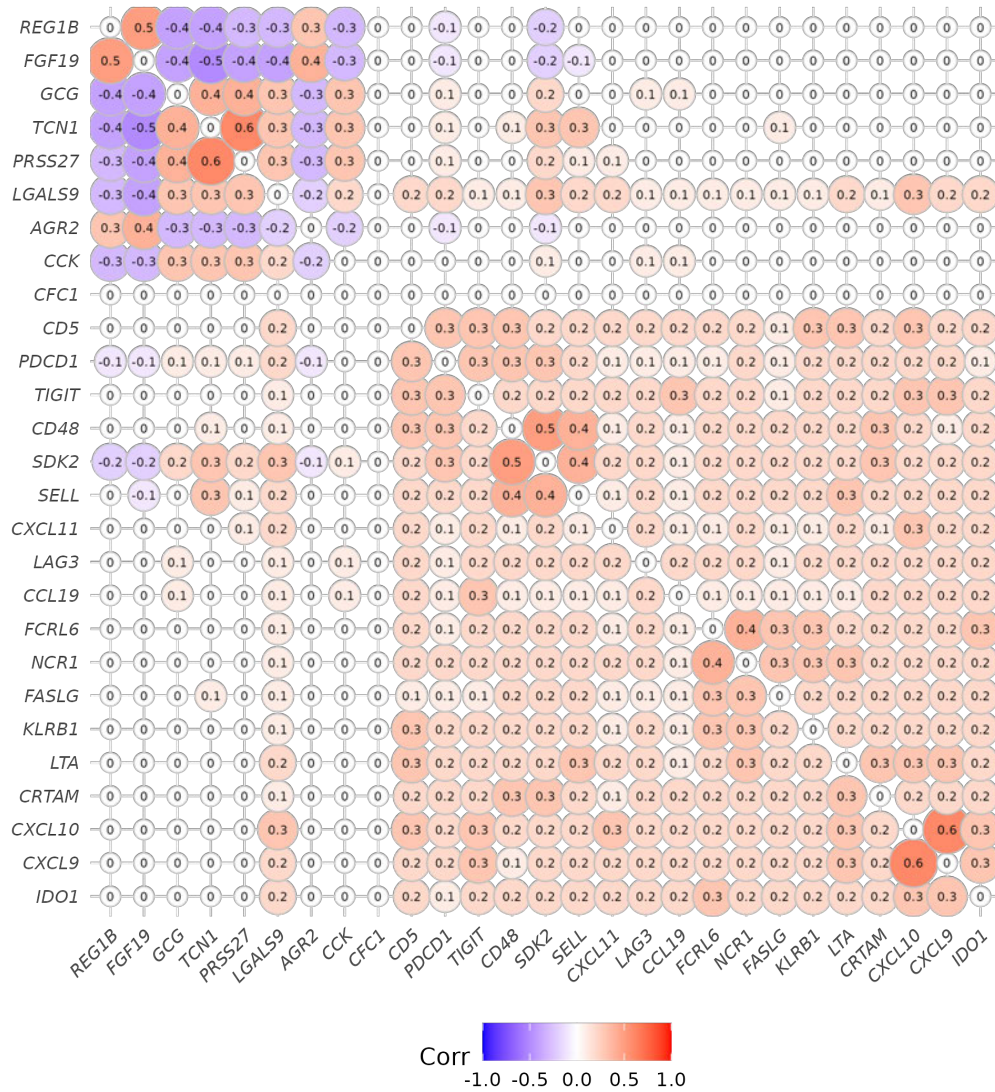


Figure 3.2: Heatmap of correlations among pGATE scores for putative core proteins for type 1 diabetes ordered by hierarchical clustering. Correlations were computed using non-diabetic controls in the Scottish dataset. Two clusters of correlated scores are apparent: a cluster of moderately correlated pGATE scores for proteins encoded by genes known to be highly expressed in gut (top left corner) and a cluster of lower correlations among pGATE scores for immune-related proteins (bottom right corner).

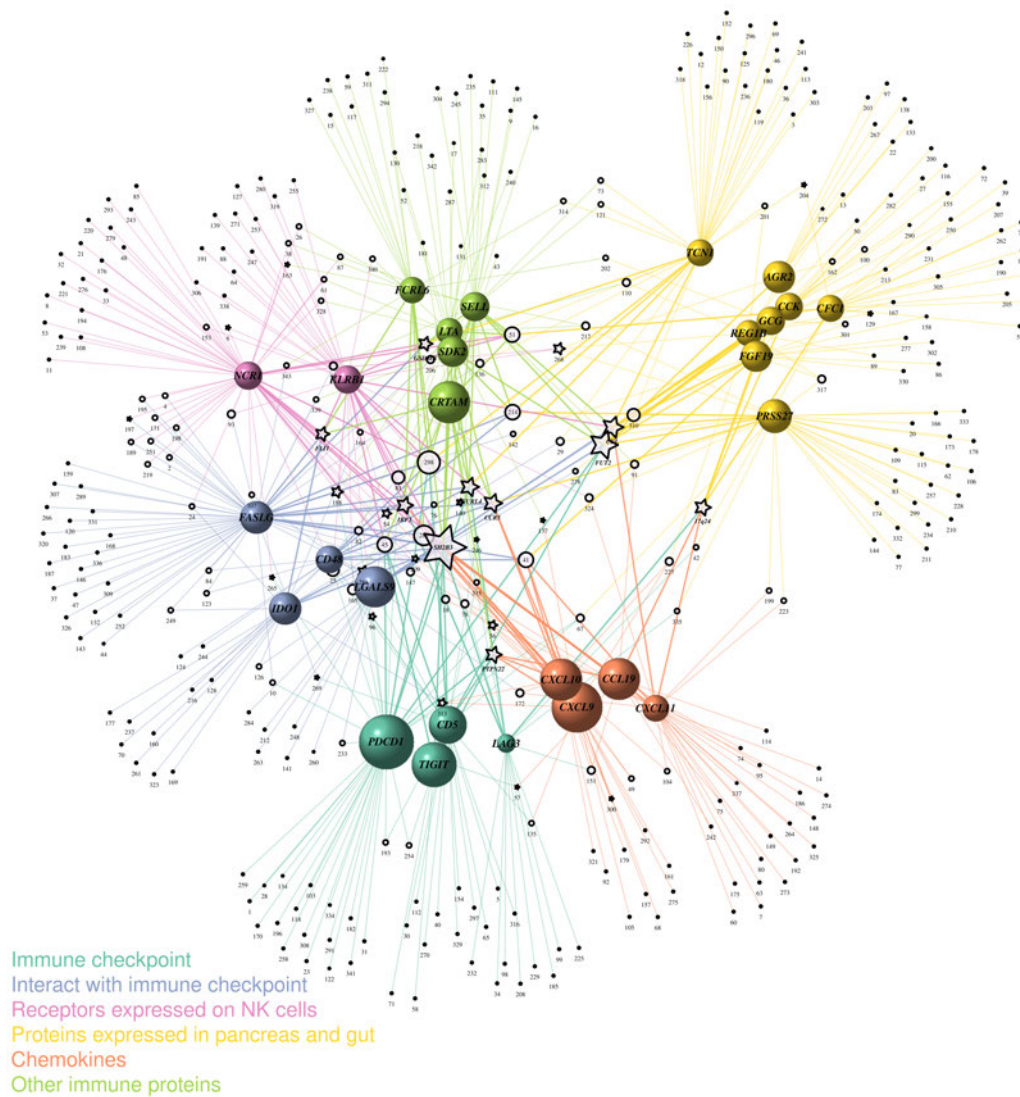


Figure 3.3: Network plot of *trans*-pQTLs contributing to pGATE scores for putative core proteins for type 1 diabetes. *Trans*-pQTLs are grouped into clumps by genomic location. Putative core proteins are shown in spheres, grouped by functional role. The size of the sphere is proportional to the effect of the pGATE score on type 1 diabetes risk. *Trans*-pQTL clumps are represented by stars if a known type 1 diabetes GWAS hit is located within 200 kb from the clump and circles if not. *Trans*-pQTL clumps located within 200 kb from a known type 1 diabetes GWAS hit and contributing to pGATE scores of more than 5 core proteins (i.e., ‘master regulators’) are labelled by the gene name the GWAS hit was assigned to. All other clumps are labelled by the clump number as shown in Table S3.2. The size of the star/circle is proportional to the number of target proteins.

Table 3.2: Summary of supportive evidence for the putative core proteins for type 1 diabetes.

Gene symbol	Mendelian randomisation ¹	Monogenic cause ²	Cis-score association ³	GWAS hit ⁴	Protein association ⁵	Experimental validation ⁶	Drug effect ⁷
<i>PDCD1</i>	+	+	.	.	+	+	+
<i>CXCL9</i>	.	.	+	.	.	.	.
<i>CXCL10</i>	+	.	.	.	.	+	.
<i>CRTAM</i>	.	.	.	.	+	.	.
<i>CCL19</i>	.	.	.	.	.	.	.
<i>LGALS9</i>	+	.	.	.	+	+	.
<i>CD5</i>	+	.	.	+	+	.	.
<i>PRSS27</i>	+	.	.	.	+	.	.
<i>FASLG</i>	+	.	.	+	.	+	.
<i>IDO1</i>	+	.	.	.	.	+	.
<i>AGR2</i>	.	.	.	.	.	.	.
<i>FGF19</i>	.	.	.	.	.	.	.
<i>SDK2</i>	.	.	.	.	+	.	.
<i>SELL</i>	.	.	.	.	.	.	.
<i>GCG</i>	.	.	.	+	.	.	.
<i>KLRB1</i>	+	.	.	+	+	.	.
<i>NCR1</i>	+	.	.	.	+	+	.
<i>CD48</i>	+	.	.	.	+	.	.
<i>CCK</i>	.	.	.	.	.	.	.
<i>CFC1</i>	.	.	.	.	+	.	.
<i>LTA</i>	.	.	.	.	.	.	.
<i>TCN1</i>	.	.	.	.	.	.	.
<i>CXCL11</i>	.	.	.	.	.	.	.
<i>REG1B</i>	.	.	.	.	.	+	.
<i>FCRL6</i>	.	.	.	.	.	.	.
<i>TIGIT</i>	.	.	.	.	.	+	.
<i>LAG3</i>	+	.	.	.	+	+	.

¹ '+' indicates that the criterion is met; '.' indicates that it is not. Criteria are defined as follows:

¹ Mendelian randomisation based on *trans*-pQTLs and marginalizing over the distribution of pleiotropic effects supports a causal effect for the protein on type 1 diabetes at p -value < 0.01 .

² There is a rare variant in the gene encoding the protein that causes monogenic autoimmune diabetes.

³ A *cis*-score for the protein is associated with type 1 diabetes at p -value $< 0.05/22$.

⁴ There is a known GWAS hit within 200 kb of the gene encoding the protein.

⁵ The measured protein level is associated with type 1 diabetes at p -value $< 10^{-4}$ and with direction of effect consistent to the pGATE score effect on type 1 diabetes.

⁶ There is evidence of causality through perturbation of gene encoding the protein or its product in an animal model of type 1 diabetes.

⁷ An existing drug (either investigational or approved) that targets the protein or its ligand/receptor cause or ameliorate type 1 diabetes.

3.2.5.3 Supportive evidence for the role of putative core proteins for type 1 diabetes

Table 3.2 summarises for each putative core protein the results of the five pre-specified criteria other than *trans*-pQTL effects that support causality of the protein for type 1 diabetes.

- For *PDCD1*, which encodes PD-1 (programmed cell death protein 1), and *PDL1* which encodes its ligand, rare loss-of-function variants have been reported to cause monogenic autoimmune diabetes [Ogishi et al., 2021, Johnson et al., 2024].
- For *CXCL9*, the *cis*-score was significantly associated with type 1 diabetes in the Scottish dataset (Table 3.3). For *CD5*, *FASLG*, *GCG*, and *KLRB1*,

associations of type 1 diabetes with SNPs within 200 kb of the transcription site have been reported [Robertson et al., 2021, Chiou et al., 2021].

- For 13 of the putative core proteins, measured levels in the UKB-PPP study were associated with type 1 diabetes at $p < 10^{-4}$ (Table 3.3). The directions of these associations with prevalent type 1 diabetes were consistent with the GATE analysis except for *AGR2* and *REG1B*.
- For 9 of the core proteins, we identified supportive evidence for a causal relationship based on perturbation in a mouse model of type 1 diabetes. Details are summarised in Table S3.4. For *PDCD1*, *TIGIT*, *LAG3*, *LGALS9*, and *IDO1*, the directions of effect of the genetically predicted levels on diabetes in the GATE analysis are opposite to what might be expected from animal models. For *CXCL10*, *FASLG*, *REG1B*, and *NCR1*, their directions of effect in the GATE analysis are the same as in animal models.
- About 1% of cancer patients treated with PD-1 or PD-L1 inhibitors develop type 1 diabetes [Cho and Jung, 2023].

3.2.6 Discussion

3.2.6.1 Statement of principal findings

This study identifies strong associations between type 1 diabetes and genetically predicted levels of proteins based on *trans*-pQTLs detected in large GWAS of high-dimensional proteomics data. For the top 27 proteins that passed our reporting threshold, a dose-response MR analysis provided further support for a causal effect for 11 proteins. Thus, there is strong genetic evidence that these proteins have a causal relationship with disease.

pGATE scores are based on circulating protein levels, and these may have a positive or inverse association with cellular or tissue-specific levels or activity. For example, it is well described that alternatively spliced secreted isoforms of receptors can act as a decoy for the receptor's ligand [Tzaban et al., 2025]. Further studies in relevant tissues and cell types are needed to unravel mechanism. Supportive evidence for causality from gene knockout, overexpression, inhibition or agonism in animal models and from monogenic disease in humans helps to clarify the likely direction of effect. For 7 of the 11 proteins with MR support, experimental studies of perturbation in a mouse model have been reported to alter the development of autoimmune diabetes, but such studies have not been

Table 3.3: Associations of type 1 diabetes with *cis*-scores and measured protein levels for putative core proteins.

Gene symbol	<i>Cis</i> -score (Scottish) ¹			<i>Cis</i> -score (UK Biobank) ¹			Measured protein level (UKB-PPP) ²			
	pQTL study	Log OR	<i>p</i> -value	pQTL study	Log OR	<i>p</i> -value	Cases	Controls	Log OR	<i>p</i> -value
<i>PDCD1</i>	UKB	-0.01	0.5	UKB	0.00	0.9	227	43458	0.52	4×10^{-17}
<i>CXCL9</i>	INT	-0.07	0.001	DeC	-0.04	0.1	226	43373	0.26	1×10^{-4}
<i>CXCL10</i>	UKB	0.05	0.02	DeC	-0.03	0.2	226	43373	0.08	0.3
<i>CRTAM</i>	UKB	-0.02	0.4	DeC	-0.05	0.04	222	42849	0.45	3×10^{-11}
<i>CCL19</i>	DeC	0.03	0.2	UKB	-0.03	0.2	229	43556	0.16	0.01
<i>LGALS9</i>	UKB	0.03	0.2	UKB	0.02	0.4	225	43161	0.30	2×10^{-5}
<i>CD5</i>	UKB	-0.02	0.4	UKB	0.01	0.7	227	43466	0.41	5×10^{-12}
<i>PRSS27</i>	UKB	-0.02	0.4	UKB	-0.02	0.3	223	43035	0.47	1×10^{-12}
<i>FASLG</i>	UKB	0.00	1	UKB	-0.02	0.5	226	43638	0.22	0.001
<i>IDO1</i>	UKB	-0.02	0.3	DeC	0.01	0.6	177	36403	0.20	0.006
<i>AGR2</i>	UKB	0.00	0.9	UKB	0.03	0.2	214	41864	0.69	4×10^{-31}
<i>FGF19</i>	.	.	.	.	.	.	226	43373	-0.16	0.02
<i>SDK2</i>	UKB	0.00	0.8	UKB	-0.04	0.1	177	36589	0.27	7×10^{-5}
<i>SELL</i>	DeC	0.02	0.3	INT	-0.03	0.1	178	37456	-0.01	0.9
<i>GCG</i>	.	.	.	.	.	.	222	42213	0.06	0.4
<i>KLRB1</i>	UKB	-0.02	0.3	UKB	0.03	0.2	226	43638	0.80	2×10^{-36}
<i>NCR1</i>	UKB	0.03	0.1	UKB	-0.03	0.3	225	43061	0.53	9×10^{-17}
<i>CD48</i>	DeC	-0.02	0.3	DeC	-0.03	0.2	226	43303	0.48	1×10^{-14}
<i>CCK</i>	.	.	.	.	.	.	.	.	.	.
<i>CFC1</i>	.	.	.	.	.	.	222	42213	-0.94	4×10^{-56}
<i>LTA</i>	.	.	.	.	.	.	225	43117	0.11	0.1
<i>TCN1</i>	DeC	-0.02	0.3	INT	0.04	0.07	174	36691	0.22	0.001
<i>CXCL11</i>	DeC	0.03	0.2	DeC	0.04	0.06	229	43556	0.18	0.007
<i>REG1B</i>	DeC	-0.03	0.2	DeC	0.04	0.1	227	43687	0.29	8×10^{-6}
<i>FCRL6</i>	INT	0.00	0.8	UKB	-0.03	0.3	225	43117	0.14	0.04
<i>TIGIT</i>	UKB	0.02	0.4	UKB	-0.04	0.1	175	36722	0.12	0.1
<i>LAG3</i>	UKB	0.02	0.3	DeC	0.01	0.5	221	42980	0.47	3×10^{-14}

Log OR: log(odds ratio) of type 1 diabetes risk per standard deviation increase in the *cis*-score or the measured protein level. UKB, UK Biobank Pharma Proteomics Project; INT, INTERVAL; DeC, deCODE genetics. '.' indicates relevant data is not available.

¹*Cis*-scores for putative core proteins were computed when at least one *cis*-SNP was associated with the protein level at *p*-value $< 10^{-6}$ in the pQTL GWAS. Associations of the *cis*-scores with type 1 diabetes were tested in the Scottish and UK Biobank datasets. None of the putative core proteins would be detected by looking at *cis*-effects on protein level.

²Association of type 1 diabetes with the measured protein level for putative core proteins was tested in the UKB-PPP study. The measured levels for thirteen proteins were significantly associated with type 1 diabetes risk at *p*-value $< 10^{-4}$. The direction of effect of the measured protein levels on type 1 diabetes were consistent with the effects of the pGATE scores on type 1 diabetes, except for *AGR2* and *REG1B* where a lower pGATE score was associated with higher type 1 diabetes risk, but higher circulating levels were associated with type 1 diabetes cross-sectionally.

reported for the other 4 [Wang et al., 2005, Lasch et al., 2015, Chou et al., 2009, Nakayama et al., 2002, Mondanelli et al., 2017, Gur et al., 2010, Unno et al., 2002, Trivedi et al., 2024, Zhang et al., 2017]. These results provide strong evidence that GATE analysis is detecting proteins that are causal. This supports further investigation of other putative core proteins that have not yet been studied in type 1 diabetes.

Conventional GWASs have identified *cis* GWAS hits in the genes for only four of these proteins; only one protein had a *cis*-score association with type 1 diabetes in our analysis. The lack of *cis*-variants in core proteins is consistent with recent studies that have shown that disease-relevant genes are enriched with redundant enhancer domains and depleted of *cis*-QTLs of large effect, as would

be expected if perturbation of these genes has large effects on fitness [Mostafavi et al., 2023].

Immune checkpoints

Table S3.1 lists reported biological functions of the putative core proteins for type 1 diabetes. Four of the top proteins we identified – *PDCD1*, *LAG3*, *CD5*, *TIGIT* – are immune checkpoints, defined broadly as receptors on immune cells that are exploited by cancer cells to escape the immune response. For the top hit *PDCD1*, the causal role of PD-1 signalling is unequivocally established by the effects of rare variants and the adverse effects of PD-1 inhibitors; loss of PD-1 signalling causes type 1 diabetes [Cho and Jung, 2023]. For soluble PD-1, the positive association with type 1 diabetes of the pGATE score and the measured level of the circulating protein is in the opposite direction to that expected from the above experimental evidence. This positive association with measured PD-1 levels has also been shown for incident type 1 diabetes in children with islet-cell antibodies [Bruzzaniti et al., 2024]. A possible explanation is that the soluble form of PD-1 can act as a decoy for the ligand for the cellular receptor [Gu et al., 2018, Khan et al., 2020] but further mechanistic studies are needed. We recently reported *PDCD1*, *CD5* and *TIGIT* as core proteins for rheumatoid arthritis [Spiliopoulou et al., 2025].

Deletion of *LAG3* in CD8+ T cells accelerated development of diabetes in non-obese diabetic (NOD) mice [Grebinski et al., 2022]. *Lag3*^{-/-} NOD mice developed substantially accelerated diabetes with 100% incidence [Bettini et al., 2011]. *TIGIT* inhibition resulted in rapid diabetes progression but only in the presence of partial inhibition of PD-1 [Trivedi et al., 2024]; *TIGIT* is involved in instability of regulatory T cells (Tregs) in NOD mice [Thirawatananond et al., 2023].

Other immune proteins

The top proteins also included the chemokines encoded by *CXCL9*, *CXCL10*, *CXCL11* and *CCL19*, with *CXCL10* having the strongest evidence for causality from MR and experimental perturbation. Other proteins with strong genetic evidence for causality include those encoded by *LGALS9*, *FASLG*, *IDO1* and *CD48*, which have been reported to interact with immune checkpoints.

LGALS9 encodes galectin-9, a ligand for the immune checkpoint receptor TIM-3 [Lv et al., 2023]. The role of *FASLG* (tumour necrosis factor ligand superfamily member 6/Fas ligand) in type 1 diabetes has been controversial [Strasser

et al., 2009] with conflicting NOD mouse studies [Trivedi et al., 2019]. IDO-1 (indoleamine 2,3-dioxygenase 1) is important in the balance between immune tolerance or activation by plasmacytoid dendritic cells [Pallotta et al., 2022] and modulates immune checkpoint-mediated immunosuppression [Yan et al., 2024]. There is supportive evidence for its role in the NOD mouse [Pallotta et al., 2022]. CD48 can activate PD-1 in human T cells by cross-linking and phosphorylation [White et al., 2023].

Two of the top associated pGATE scores with MR support were *NCR1* and *KLRB1*, which encode natural cytotoxicity triggering receptor 1 and killer cell lectin-like receptor subfamily B member 1 expressed on natural killer (NK) cells. This points to a role of NK cells in type 1 diabetes, supported for *NCR1* by experimental perturbation [Gur et al., 2010].

Other immune proteins among the top 27, but less strongly supported by other evidence for causality, include those encoded by *LTA* (lymphotoxin- α), *CRTAM* (cytotoxic and regulatory T cell molecule), *FCRL6* (Fc receptor-like protein 6), *SELL* (L-selectin) and *SDK2* (protein sidekick 2).

Pancreatic and gut proteins

Among the top 27 genetically predicted protein levels that passed our reporting threshold, several were known to be gut associated and with no reported immune role (*PRSS27*, *GCG*, *CCK*, *TCN1*, positively associated; *AGR2*, *FGF19*, *CFC1*, *REG1B*, inversely associated). Of these, only *REG1B* (lithostathine-1- β /regenerating islet-derived protein 1- β) has supportive experimental evidence [Unno et al., 2002], and only the *GCG* gene (encoding glucagon, glucagon-like peptide 1 and other peptides) contains a previously reported type 1 diabetes GWAS hit. None are known monogenic causes of type 1 diabetes.

The associations with type 1 diabetes of *PRSS27* (serine protease 27/pancreasin) [Bhagwandin et al., 2003] and *REG1B* are noteworthy given the interest in the role of the exocrine pancreas in type 1 diabetes. However, as the *FUT2* locus, a known GWAS hit for type 1 diabetes [Smyth et al., 2011], is a strong master regulator of all these gut associated proteins except *CFC1*, causality cannot be distinguished from confounding with current data, but further studies are warranted. The MR analysis provided further support for causality only for *PRSS27*.

3.2.6.2 Strengths and limitations

Strengths are that this study uses a novel approach which harnesses *trans*-effects that are missed by conventional methods that focus on *cis*-effects: the key results are supported by a formal test of causality that allows for pleiotropic effects and by replication in two independent case-control datasets.

A limitation is that further studies are needed to clarify mechanism and direction of effect. Also, the SNP associations with proteins were estimated in datasets that did not exclude people with type 1 diabetes. However, this is unlikely to make any material difference to SNP-protein regression coefficients, as the prevalence of type 1 diabetes is only about 0.5%. For the UK Biobank dataset where individual-level proteomics data are now available, we confirmed that the exclusion of those with type 1 diabetes did not appreciably alter the regression coefficients.

GATE analysis is still constrained by the limited size of the datasets for learning *trans*-pQTLs or *trans*-eQTLs, which are typically weak. Even with the relatively large sample size of 45,397 participants in UKB-PPP, pQTL GWAS summary statistics capture only about 10%-20% of the estimated SNP heritability attributable to *trans*-effects [Sun et al., 2023]. If all *trans*-effects were captured, the effect sizes for pGATE scores would be several times larger than those given in Table 3.1, assuming the detected associations are causal.

For some proteins in the UKB-PPP dataset measured by the Olink platform, a high fraction of measured values is below the limit of detection (LOD). However, of the 27 top proteins based on pGATE scores, the median proportion of values below the LOD was 0.03%; only one (*AGR2*) had a high proportion of values below LOD (28%). The agreement between measurements of the same proteins on Somalogic and Olink platforms (both in terms of pQTLs and of the measured protein levels) is poor for some proteins [Sun et al., 2018, 2023, Pietzner et al., 2020, 2021, Eldjarn et al., 2023, Wang et al., 2025, Rooney et al., 2024]. However, proteins that are poorly measured with low reliability or accuracy, regardless of platform, will have few detected *trans*-pQTLs, and would thus be excluded by our filtering of GATE scores with effective number of *trans*-pQTLs < 5. Of the top 27 core proteins reported, for 9 of them there was an adequate score evaluable for both platforms (effective number of *trans*-pQTLs > 5). For all of these 9, the direction of association was the same for both UKB-PPP and deCODE platforms, though the magnitude of the odds ratio and significance level varied. Finally, we note that the use of GWAS proteomics datasets such

as UKB-PPP that are not representative of the total population in age or other demographic factors could mean that some SNP-protein associations remain undetected.

3.2.7 Conclusions

This study provides genetic evidence for causality of proteins regulating T cell responses – specifically immune checkpoints – in type 1 diabetes and also highlights the importance of innate immune pathways acting on NK cells. These results should help to focus drug development on targets with a high chance of success at clinical stage. We suggest that PD-1 agonists, which are being developed for other indications and have been found to be well-tolerated in phase 2 trials, should be trialled for type 1 diabetes prevention.

3.2.7.1 Acknowledgements

The authors would like to express their gratitude to all study participants, and medical and administrative staff who helped with participant recruitment, and sample and data collection, and were thus an integral part to the delivery of this study.

3.2.7.2 Funding

This study was supported by Diabetes UK (15/0005301). The establishment of the SDRN Type 1 Bioresource was supported by the Chief Scientist Office of the Scottish Government Health Directorates (ETM/47), by Diabetes UK (10/0004010), and by in-kind contributions from the Scottish Diabetes Research Network. Genotyping was supported by the Juvenile Diabetes Research Fund (17-2013-7). Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006] and is currently supported by the Wellcome Trust [216767/Z/19/Z]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Edinburgh Clinical Research Facility, University of Edinburgh, Scotland, and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award “STratifying Resilience and Depression Longitudinally” (STRADL) Reference 104036/Z/14/Z). The development of the GENOSCORES platform was supported by a Springboard Award (SBF006/1109) from the Academy of Medical Sciences, supported in turn by the Wellcome Trust,

the UK Government Department of Business, Energy and Industrial Strategy, the British Heart Foundation, and Diabetes UK.

AI was supported by the Medical Research Council Cross Disciplinary Fellowship (XDF) Programme (MC_FE_00035). CH was supported by the Medical Research Council Human Genetics Unit programme grant (U. MC_UU_00007/10).

3.2.7.3 Data availability

The code used in this study is available in our GitHub repository (<https://github.com/molepi-precmed/trans-qtls>), and summary-level data for the association of type 1 diabetes with all tested pGATE scores has been deposited to Zenodo (<https://doi.org/10.5281/zenodo.16738221>). Applicants who wish to access SDRNT1BIO data can submit an initial request by emailing SDRNT1BIO@ed.ac.uk. The access procedure for the Generation Scotland data is described at <https://genscot.ed.ac.uk/for-researchers/access>. UK Biobank data may be accessed by completing an application at <https://www.ukbiobank.ac.uk/>. The platform used to compute locus-specific genotypic scores and the database of published GWAS summary statistics are accessible on the <https://genoscores.cphs.mvm.ed.ac.uk/> platform. Access requests to the GENOSCORES platform and the corresponding R package should be directed to the corresponding author.

3.2.7.4 Author contributions

Xuan Zhou: Methodology; Software; Validation; Formal analysis; Data curation; Writing - original draft; Writing - review & editing; Visualization.

Andrii Iakovliev: Conceptualization; Methodology; Software; Validation; Formal analysis; Data curation; Writing - original draft; Writing - review & editing; Supervision; Visualization; Funding acquisition.

Stuart J McGurnaghan: Software; Validation; Data curation; Writing - review & editing.

Buddhiprabha Erabadda: Software; Validation; Formal analysis; Data curation; Writing - review & editing; Visualization.

Caroline Hayward: Validation; Resources; Data curation; Writing - review & editing; Funding acquisition.

Paul M McKeigue: Conceptualization; Methodology; Software; Validation; Formal analysis; Data curation; Writing - original draft; Writing - review & editing; Supervision; Visualization; Funding acquisition; Project administration.

Athina Spiliopoulou: Conceptualization; Methodology; Software; Validation; Formal analysis; Data curation; Writing - original draft; Writing - review & editing; Supervision; Visualization; Funding acquisition.

Helen M Colhoun: Conceptualization; Methodology; Validation; Resources; Data curation; Writing - original draft; Writing - review & editing; Supervision; Visualization; Funding acquisition.

3.2.7.5 Guarantor Data Access and Responsibility Statement

P.M.M. and H.M.C. are the guarantors of this work, had full access to all the data, and take full responsibility for the integrity of the data and the accuracy of the data analysis.

3.2.7.6 Duality of Interest

No potential conflicts of interest relevant to this article were reported.

3.2.7.7 Prior Presentation information

Results of this study have been previously presented at the annual conference of the European Society of Human Genetics (ESHG 2024) held in Berlin, Germany, 1-4 June 2024; the Immunology of Diabetes Society Congress (IDS 2024) held in Bruges, Belgium, 4-8 November 2024; and the American Society of Human Genetics (ASHG 2024) held in Denver, USA, 5-9 November 2024. Conference abstracts are usually published 6 months after the conference and a DOI will be added accordingly.

3.2.8 Supplementary Materials

3. Identification of putative core proteins for T1D using the GATE analysis 59

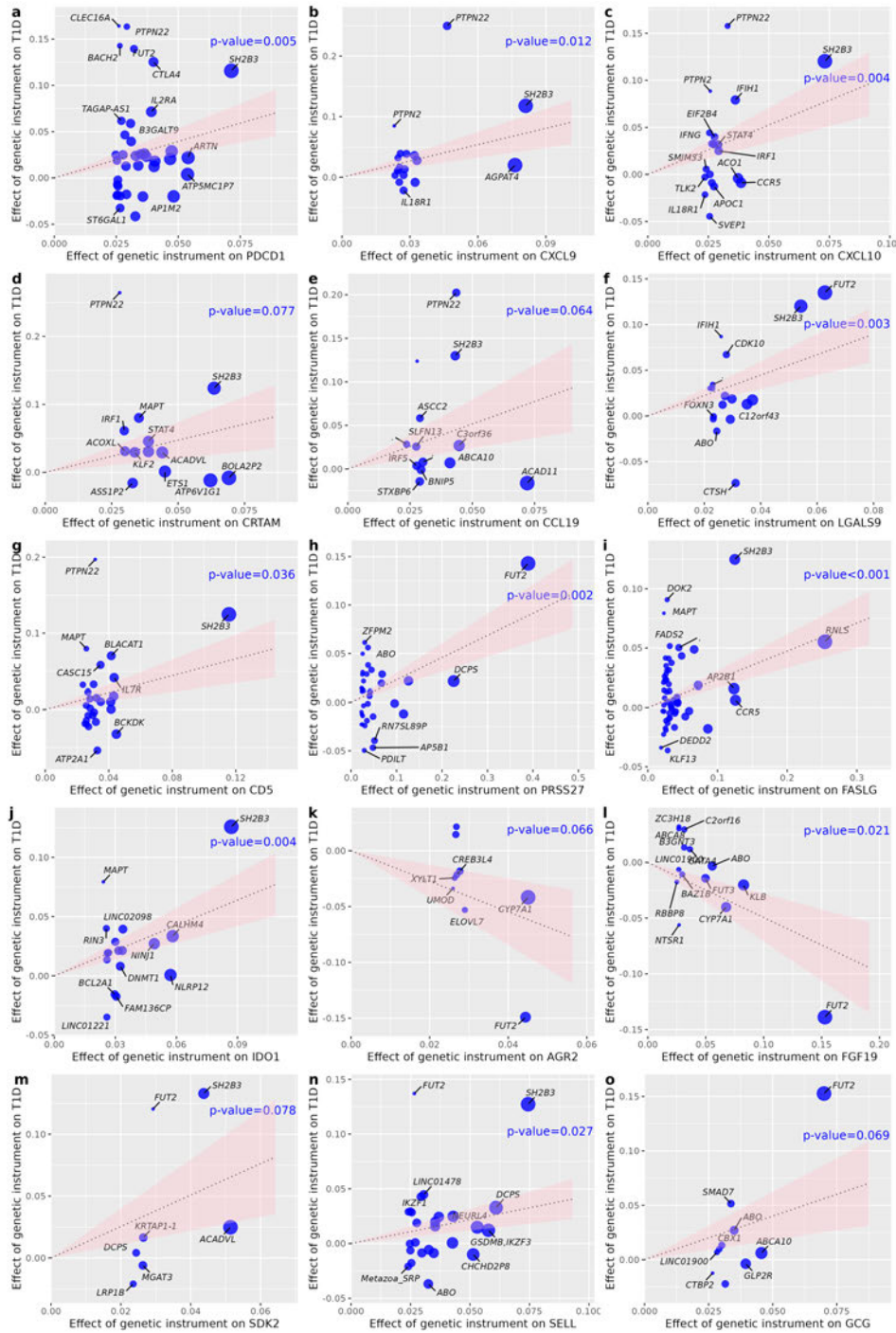


Figure S3.1: Mendelian randomisation analysis of the effect of the putative core proteins on type 1 diabetes in the Scottish dataset. All *trans*-pQTLs contributing to the pGATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and type 1 diabetes are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate.

3. Identification of putative core proteins for T1D using the GATE analysis 60

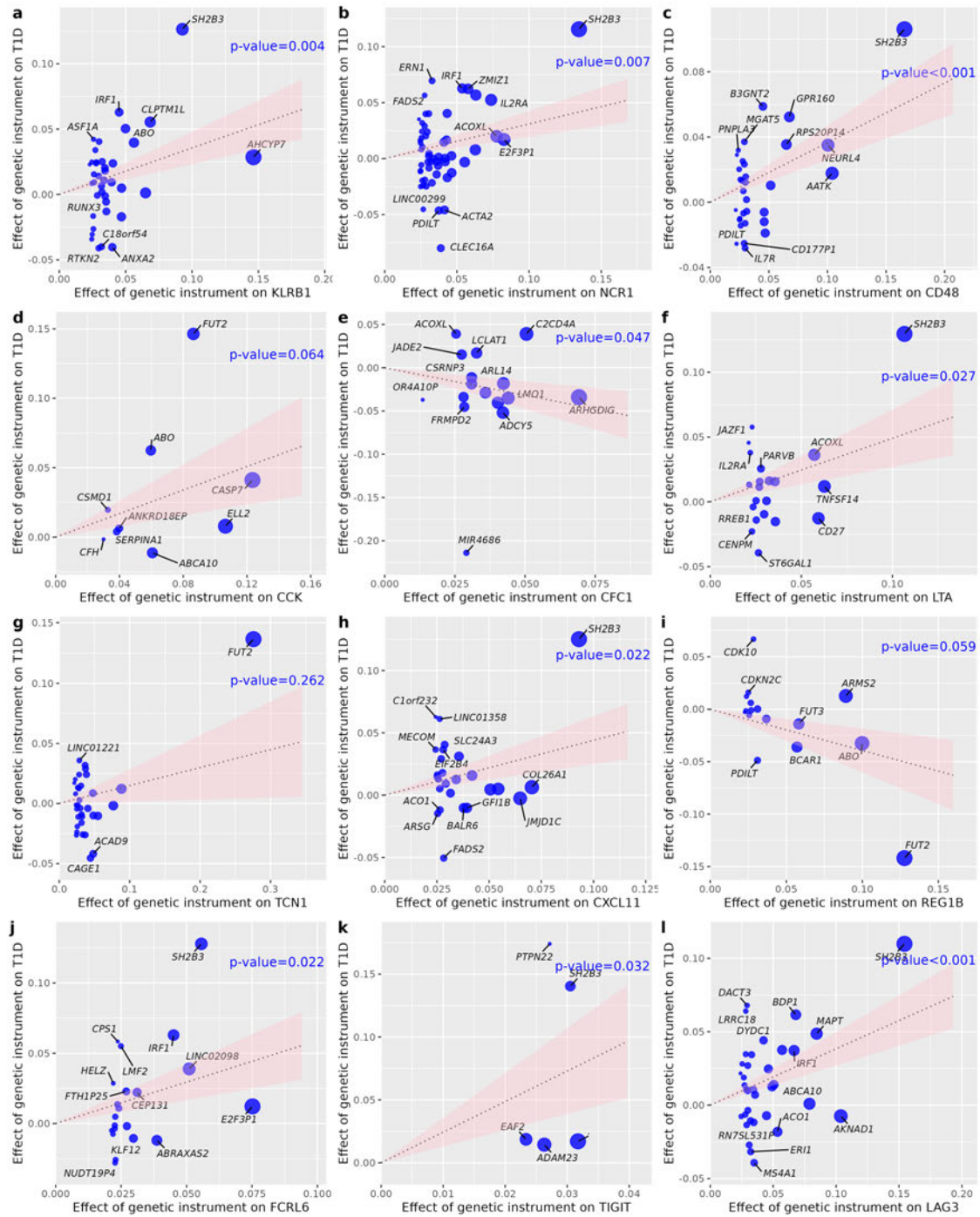


Figure S3.2: Mendelian randomisation analysis of the effect of the putative core proteins on type 1 diabetes in the Scottish dataset, except for *CD48* and *LAG3*, which are represented by the UK Biobank data due to a higher number of *trans*-pQTLs for these two putative core proteins in this dataset. All *trans*-pQTLs contributing to the pGATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and type 1 diabetes are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate.

3. Identification of putative core proteins for T1D using the GATE analysis 61

Table S3.1: Gene name, protein name synonyms and biological functions of the putative core proteins for type 1 diabetes.

UniProt ID	Gene names	Protein names	Biological functions
Q15116	<i>PDCD1</i> <i>PD1</i>	Programmed cell death protein 1 (Protein PD-1) (hPD-1) (CD antigen CD279)	Inhibitory receptor on T cells that acts as an immune checkpoint and helps cancer cells to escape immune surveillance [Gu et al., 2018, Ren et al., 2018].
P18627	<i>LAG3</i> <i>FDC</i>	Lymphocyte activation gene 3 protein (LAG-3) (CD antigen CD223) [Cleaved into: Secreted lymphocyte activation gene 3 protein (sLAG-3)]	Immune checkpoint receptor protein that expresses on activated immune cells, negatively regulates the function of T cells, and promotes tumour cells immune escape in the tumour microenvironment [Huo et al., 2022].
Q495A1	<i>TIGIT</i> <i>VSIG9</i> <i>VSTM3</i>	T-cell immunoreceptor with Ig and ITIM domains (V-set and immunoglobulin domain-containing protein 9) (V-set and transmembrane domain-containing protein 3)	Exerts its inhibitory effects on innate and adaptive immunity by including triggering T/NK cell intrinsic inhibition, inducing immunosuppressive DCs, inhibiting CD226 signaling, enhancing immunosuppression of Tregs and promoting Fap2-induced T/NK cells inhibition [Tang et al., 2023].
P06127	<i>CD5</i> <i>LEU1</i>	T-cell surface glycoprotein CD5 (Lymphocyte antigen T1/Leu-1) (CD antigen CD5)	Transmembrane glycoprotein expressed on T cells and some B cells. Using anti-CD5 monoclonal antibodies can promote anti-tumour T cell immunity [Alotaibi et al., 2024].
Q07325	<i>CXCL9</i> <i>CMK</i> <i>MIG</i> <i>SCYB9</i>	C-X-C motif chemokine 9 (Gamma-interferon-induced monokine) (Monokine induced by interferon-gamma) (HuMIG) (MIG) (Small-inducible cytokine B9)	Belongs to the CXC chemokines subfamily, binds to CXCR3, can recruit and migrate different categories of immune cells, regulate tumor pathological behaviors like proliferation, invasion and metastasis, activate angiogenesis, etc [Zhou et al., 2023]. Discriminates between those with and without type 1 diabetes [Pan et al., 2022].
P02778	<i>CXCL10</i> <i>INP10</i> <i>SCYB10</i>	C-X-C motif chemokine 10 (10 kDa interferon gamma-induced protein) (Gamma-IP10) (IP-10) (Small-inducible cytokine B10) [Cleaved into: CXCL10(1-73)]	Belongs to the CXC chemokines subfamily, binds to CXCR3, can recruit and migrate different categories of immune cells, regulate tumor pathological behaviors like proliferation, invasion and metastasis, activate angiogenesis, etc [Zhou et al., 2023]. It has been suggested to be one of the major chemokines expressed in the islet environment [Pan et al., 2022].

3. Identification of putative core proteins for T1D using the GATE analysis 62

O14625	CXCL11 ITAC SCYB11 SCYB9B	C-X-C motif chemokine 11 (Beta-R1) (H174) (Interferon gamma-inducible protein 9) (IP-9) (Interferon-inducible T-cell alpha chemoattractant) (I-TAC) (Small-inducible cytokine B11)	Belongs to the CXC chemokines subfamily, binds to CXCR3, can recruit and migrate different categories of immune cells, regulate tumor pathological behaviors like proliferation, invasion and metastasis, activate angiogenesis, etc [Zhou et al., 2023]. It is implicated in the pathogenesis of autoimmune diseases [Pan et al., 2022].
Q99731	CCL19 ELC MIP3B SCYA19	C-C motif chemokine 19 (Beta-chemokine exodus-3) (CK beta-11) (Epstein-Barr virus-induced molecule 1 ligand chemokine) (EBI1 ligand chemokine) (ELC) (Macrophage inflammatory protein 3 beta) (MIP-3-beta) (Small-inducible cytokine A19)	May play a role not only in inflammatory and immunological responses but also in normal lymphocyte recirculation and homing. Binds to chemokine receptor CCR7.
O00182	LGALS9	Galectin-9 (Gal-9) (Ecalectin) (Tumor antigen HOM-HD-21)	One of four ligands for the immune checkpoint receptor TIM-3 [Lv et al., 2023].
P48023	FASLG APT1LG1 CD95L FASL TNFSF6	Tumor necrosis factor ligand superfamily member 6 (Apoptosis antigen ligand) (APTL) (CD95 ligand) (CD95-L) (Fas antigen ligand) (Fas ligand) (FasL) (CD antigen CD178) [Cleaved into: Tumor necrosis factor ligand superfamily member 6, membrane form; Tumor necrosis factor ligand superfamily member 6, soluble form (Receptor-binding FasL ectodomain) (Soluble Fas ligand) (sFasL); ADAM10-processed FasL form (APL); FasL intracellular domain (FasL ICD) (SPPL2A-processed FasL form) (SPA)]	Expresses on cytotoxic T cells and NK cells and contributes to their ability to kill target cells.
P14902	IDO1 IDO INDO	Indoleamine 2,3-dioxygenase 1 (IDO-1) (EC 1.13.11.52) (Indoleamine-pyrrole 2,3-dioxygenase)	A tryptophan metabolising enzyme whose induction by interferon- γ is important in the balance between immune tolerance or activation by plasmacytoid dendritic cells [Pallotta et al., 2022, Mondanelli et al., 2017] also modulates immune checkpoint-mediated immunosuppression [Zhai et al., 2018, Yan et al., 2024].
P09326	CD48 BCM1 BLAST1	CD48 antigen (B-lymphocyte activation marker BLAST-1) (BCM1 surface antigen) (Leukocyte antigen MEM-102) (SLAM family member 2) (SLAMF2) (Signaling lymphocytic activation molecule 2) (TCT.1) (CD antigen CD48)	A glycosylphosphatidylinositol-linked glycoprotein molecule expresses on T cells, can activate PD-1 in human T cells by cross-linking and phosphorylation [White et al., 2023].

3. Identification of putative core proteins for T1D using the GATE analysis 63

O76036	NCR1 LY94	Natural cytotoxicity triggering receptor 1 (Lymphocyte antigen 94 homolog) (NK cell-activating receptor) (Natural killer cell p46-related protein) (NK-p46) (NKp46) (hNKp46) (CD antigen CD335)	A cytotoxicity-activating receptor on NK cells and can recognize an unknown ligand expressed on β -cells [Gardner and Fraker, 2021].
Q12918	KLRB1 CLEC5B NKR P1A	Killer cell lectin-like receptor subfamily B member 1 (C-type lectin domain family 5 member B) (HNKR-P1a) (NKR-P1A) (Natural killer cell surface protein P1A) (CD antigen CD161)	Recognizes lectin-like transcript-1 (LLT1) and mediates inhibitory signalling for NK cells cytotoxicity [Bláha et al., 2022].
P01374	LTA TNFB TNFSF1	Lymphotoxin-alpha (LT-alpha) (TNF-beta) (Tumor necrosis factor ligand superfamily member 1)	Ligand for the immune checkpoint receptor TNFRSF14 [von Hoff et al., 2019, Aubert et al., 2021].
O95727	CRTAM	Cytotoxic and regulatory T-cell molecule (Class-I MHC-restricted T-cell-associated molecule) (CD antigen CD355)	Expresses upon T cell activation on CD8+ T and NK cells and modulates cytotoxicity [Vázquez-Martínez et al., 2020, Fuchs et al., 2019].
Q6DN72	FCRL6 FCRH6	Fc receptor-like protein 6 (FcR-like protein 6) (FcRL6) (Fc receptor homolog 6) (FcRH6) (IFGP6)	Expresses on mature NK and T cells with cytotoxic potential [Davis, 2020].
P14151	SELL LNHR LYAM1	L-selectin (CD62 antigen-like family member L) (Leukocyte adhesion molecule 1) (LAM-1) (Leukocyte surface antigen Leu-8) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (Lymph node homing receptor) (TQ1) (gp90-MEL) (CD antigen CD62L)	A homing receptor for lymphocyte migration [Ivetic et al., 2019].
Q58EX2	SDK2 KIAA1514	Protein sidekick-2	Belongs to the immunoglobulin superfamily, distinctively participates in the formation and/or maintenance of neural circuits in the retina [Yamagata, 2020].
Q9BQR3	PRSS27 MPN	Serine protease 27 (EC 3.4.21.-) (Marapsin) (Pancreasin)	A secreted tryptic serine protease that is expressed mainly in the pancreas [Bhagwandin et al., 2003] and oesophagus [Kataoka et al., 2021].
O95994	AGR2 AG2	Anterior gradient protein 2 homolog (AG-2) (hAG-2) (HPC8) (Secreted cement gland protein XAG-2 homolog)	A protein disulfide isomerase that is crucial for proper processing of gel-forming mucins in the gut and is widely expressed. It is over-expressed in many cancers including pancreatic ductal adenocarcinoma [Salu et al., 2025].
P48304	REG1B PSPS2 REGL	Lithostathine-1-beta (Pancreatic stone protein 2) (PSP-2) (Regenerating islet-derived protein 1-beta) (REG-1-beta) (Regenerating protein I beta)	One of four REG proteins expressed in the exocrine pancreas and elsewhere in the gut [Chen et al., 2019].

3. Identification of putative core proteins for T1D using the GATE analysis 64

P01275	GCG	Glucagon [Cleaved into: Glicentin; Glicentin-related polypeptide (GRPP); Oxyntomodulin (OXM) (OXY); Glucagon; Glucagon-like peptide 1 (GLP-1) (Incretin hormone); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)]	A well known pancreatic α-cell protein and GWAS hits for T1D.
P06307	CCK	Cholecystokinin (CCK) [Cleaved into: Cholecystokinin-58 (CCK58); Cholecystokinin-58 desnonopeptide ((1-49)-CCK58); Cholecystokinin-39 (CCK39); Cholecystokinin-33 (CCK33); Cholecystokinin-25 (CCK25); Cholecystokinin-18 (CCK18); Cholecystokinin-12 (CCK12); Cholecystokinin-8 (CCK8); Cholecystokinin-7 (CCK7); Cholecystokinin-5 (CCK5)]	Produced by the enteroendocrine cells of the duodenum and stimulates both bile release from the gallbladder and exocrine pancreatic enzymes.
P20061	TCN1 TC1	Transcobalamin-1 (TC-1) (Haptocorrin) (HC) (Protein R) (Transcobalamin I) (TC I) (TCI)	Produced by salivary glands and transports Vitamin-B12 (cobalamin) to the intestine.
O95750	FGF19	Fibroblast growth factor 19 (FGF-19)	A gut hormone secreted by ileal enterocytes in response to feeding, has insulin-like effect and participate in bile homeostasis.
P0CG37	CFC1	Cryptic protein (Cryptic family protein 1)	Participates in patterning the left-right axis in embryonic development.

The Hugo Gene names for the genes encoding the putative core proteins along with the corresponding UniProt protein names are shown in bold.

3. Identification of putative core proteins for T1D using the GATE analysis 65

Table S3.2: Clumps (genomic locations) harbouring *trans*-pQTLs contributing to pGATE scores of the putative core proteins for type 1 diabetes in the Scottish dataset.

Clump	Chr	Start position (Mb)	End position (Mb)	Target proteins	Number of target proteins	Nearby genes	Monogenic genes
1	1	1.21	1.37	<i>PDCD1</i>	1	<i>ACAP3, B3GALT6, C1QTNF12, CPTP, DVL1, ...</i>	
2	1	2.22	2.32	<i>FASLG, NCR1</i>	2	<i>SKI</i>	
3	1	8.32	8.65	<i>TCN1</i>	1	<i>RERE, RERE-AS1, SLC45A1</i>	
4	1	20.72	21.21	<i>FASLG, NCR1</i>	2	<i>EIF4G3, HP1BP3, HSPE1P27, MIR1256, RNU7-200P, ...</i>	
5	1	24.15	24.16	<i>CD5</i>	1	<i>IFNLR1</i>	
6	1	24.91	24.98	<i>KLRB1, NCR1</i>	2	<i>RUNX3</i> , <i>MIR6731</i>	
7	1	26.00	26.33	<i>CXCL11</i>	1	<i>C1orf232, CATSPER4, CD52, CEP85, CNKSR1, ...</i>	
8	1	30.75	30.78	<i>NCR1</i>	1	<i>LAPTM5</i>	
9	1	36.39	36.49	<i>SELL</i>	1	<i>CSF3R, LSM10, MRPS15, OSCP1, RNU4-27P, ...</i>	
10	1	43.48	44.12	<i>FASLG, PDCD1</i>	2	<i>ARTN, ATP6V0B, B4GALT2, CCDC24, DPH2, ...</i>	
11	1	45.33	45.54	<i>NCR1</i>	1	<i>CCDC163, HMGB1P48, MMACHC, MUTYH, PPIAP36, ...</i>	
12	1	47.20	47.23	<i>TCN1</i>	1	<i>TAL1</i>	
13	1	50.56	51.01	<i>REG1B</i>	1	<i>CDKN2C, FAF1, MRPS6P2, PHB1P12, RNU6-1026P</i>	
14	1	58.95	59.00	<i>CXCL11</i>	1	<i>LINC01358, PHB1P3</i>	
15	1	71.61	71.66	<i>FCRL6</i>	1	<i>NEGR1</i>	
16	1	91.95	92.05	<i>SELL</i>	1	<i>BRDT, EPHX4</i>	
17	1	92.48	92.82	<i>LTA</i>	1	<i>CCNJP2, EVI5, GFI1, H3P3, HMGB3P9, ...</i>	
18	1	109.24	109.50	<i>CXCL11, LAG3, NCR1</i>	3	<i>AMIGO1, ATXN7L2, CELSR2, CYB561D1, MYBPHL, ...</i>	
19	1	113.53	114.13	<i>CCL19, CD5, CRTAM, CXCL10, CXCL9, PDCD1, TIGIT</i>	7	<i>PTPN22</i> , <i>AP4B1, AP4B1-AS1, BCL2L15, ...</i>	
20	1	117.60	117.62	<i>PRSS27</i>	1	<i>TENT5C, TENT5C-DT</i>	
21	1	150.76	150.98	<i>NCR1</i>	1	<i>ARNT, CERS2, CTSK, CTSS, CTXND2, ...</i>	
22	1	153.79	154.11	<i>AGR2</i>	1	<i>CREB3L4, CRTC2, DENND4B, GATAD2B, JTB, ...</i>	
23	1	156.77	156.84	<i>PDCD1</i>	1	<i>HDGF, NTRK1, PRCC, SH2D2A</i>	
24	1	161.19	161.65	<i>IDO1, NCR1</i>	2	<i>ADAMTS4, APOA2, CFAP126, FCER1G, FCGR2A, ...</i>	

3. Identification of putative core proteins for T1D using the GATE analysis 66

25	1	179.31	179.79	CD48, CD5, KLRB1, NCR1, PDCD1	5	AXDND1, FAM163A, HNRNPA1P54, MEF2AP1, NPHS2, ...	
26	1	184.70	185.46	FCRL6, NCR1	2	EDEM3, FTH1P25, GS1-279B7.1, IVNS1ABP, LINC01633, ...	
27	1	196.68	196.81	CCK	1	CFH, CFHR3	
28	1	198.30	198.46	PDCD1	1	NEK7	
29	1	198.99	199.07	IDO1, PRSS27, TCN1	3	LINC01221, LINC01222	
30	1	205.24	205.84	CD5	1	BLACAT1, CDK18, ELK4, KLHDC8A, LEMD1, ...	
31	1	206.50	206.51	PDCD1	1	C1orf147, IKBKE, RASSF5	
32	1	207.74	207.88	NCR1	1	CD46, CDCA4P4, MIR29B2, MIR29B2CHG, MIR29C	
33	1	220.82	220.92	NCR1	1	HLX, HLX-AS1, LINC01352, MTARC1, RNU6ATAC35P	
34	1	226.72	226.90	LAG3	1	ITPKB, PSEN2, RPS27P5	
35	1	229.43	229.56	SELL	1	ABCB10, ACTA1, NUP133, NUP133-DT, R5SP78, ...	
36	1	234.55	234.56	TCN1	1		
37	1	247.86	247.88	FASLG	1	TRIM58	
38	2	8.30	8.32	FCRL6, NCR1	2	LINC00299	
39	2	17.44	17.81	CFC1	1	GEN1, RAD51AP2, SMC6, VSNL1	
40	2	25.27	25.28	CD5	1	DNMT3A	
41	2	27.33	28.02	CD48, CXCL10, CXCL11, CXCL9, FASLG, FGF19	6	BABAM2, C2orf16, CCDC121, EIF2B4, FAM133EP, ...	
42	2	30.20	30.67	CFC1, LAG3	2	H3P5, LBH, LCLAT1	
43	2	38.69	38.82	CRTAM	1	ASS1P2, DHX57, GALM, GEMIN6, NPLP1, ...	
44	2	60.39	60.49	FASLG	1	BCL11A , MIR4432, MIR4432HG	
45	2	62.26	62.57	CD5, KLRB1, NCR1, PDCD1, SELL, TIGIT	6	RN7SL18P, RN7SL51P, RSL24D1P2, TMEM17	
46	2	64.86	64.86	TCN1	1	LINC01800	
47	2	68.33	68.44	FASLG	1	PLEK	
48	2	95.73	96.11	NCR1	1	ADRA2B, ANKRD36C, FAHD2CP, GPAT2, GPAT2P1, ...	
49	2	102.30	102.48	CXCL10, CXCL9	2	IL18R1, IL18RAP, IL1RL1, MIR4772, SDR42E1P5, ...	
50	2	105.87	105.87	REG1B	1	NCK2	
51	2	110.80	111.26	CFC1, CRTAM, KLRB1, LTA, NCR1, SELL	6	ACOXL, ACOXL-AS1, BCL2L11, MIR4435-2HG, RPL5P9	
52	2	141.32	141.32	SDK2	1	LRP1B	
53	2	144.58	144.64	NCR1	1		
54	2	162.10	162.47	CXCL10, KLRB1, LGALS9, NCR1	4	IFIH1 , FAP, GCA, GCG, KCNH7	
55	2	165.68	165.78	CFC1	1	CSRNP3, GALNT3	
56	2	190.39	191.14	CRTAM, CXCL10, CXCL9, PDCD1	4	STAT4 , STAT4-AS1, GLS, MFS6, B1, ...	STAT1
57	2	203.83	203.93	CCL19, PDCD1	2	CTLA4	CTLA4
58	2	206.52	206.52	TIGIT	1	ADAM23	
59	2	210.68	210.68	FCRL6	1	CPS1	
60	2	224.01	224.02	CXCL11	1	SERPINE2	

3. Identification of putative core proteins for T1D using the GATE analysis 67

61	2	231.33	231.90	LTA, NCR1	2	ARMC9, B3GNT7, COPS7B, HIGD2AP1, LINC00471, ...
62	2	240.57	240.60	PRSS27	1	CAPN10, CAPN10-DT, RNPEPL1
63	3	16.84	18.33	CXCL11	1	BALR6, MIR3714, PDCL3P3, PLCL2, PLCL2-AS1, ...
64	3	23.89	23.98	KLRB1	1	NKIRAS1, NR1D2, RPL15
65	3	25.29	25.29	CD5	1	RARB
66	3	45.81	50.41	CD48, CRTAM, CXCL10, CXCL9, FASLG, KLRB1, LTA, PRSS27	8	CCR5, ACTL11P, ALS2CL, AMIGO3, AMT, ...
67	3	56.81	56.92	CD5, PDCD1, PRSS27	3	ARHGEF3
68	3	58.20	58.20	CXCL9	1	DSE1L3
69	3	69.75	69.86	TCN1	1	MITF, RN7SL418P
70	3	101.19	101.60	IDO1	1	FAM136CP, FAM172BP, IMPG2, PCNP, SENP7, ...
71	3	121.79	121.98	TIGIT	1	EAF2, IQCB1, SLC15A2
72	3	123.22	123.42	CFC1	1	ADCY5, PDIA5, SEC22A
73	3	128.49	128.94	SELL, TCN1	2	ACAD9, ACAD9-DT, CFAP92, FTH1P4, GATA2, ...
74	3	132.40	133.20	CCL19	1	ACAD11, ACKR4, DJC13, HSPA8P19, NIP7P2, ...
75	3	133.82	134.04	CCL19	1	C3orf36, RAB6B, SLCO2A1, SRPRB
76	3	136.08	136.87	CD5, FCRL6	2	HMGN1P10, MSL2, NCK1, NCK1-DT, PCCB, ...
77	3	141.31	141.62	PRSS27	1	KRT18P35, PXYLP1, RASA2, RASA2-IT1, YWHAQP6, ...
78	3	159.91	160.01	CXCL9, KLRB1, LAG3	3	IL12A, IL12A-AS1
79	3	160.23	160.68	CFC1	1	ARL14, B3GAT3P1, IFT80, KP4, KRT8P12, ...
80	3	169.07	169.15	CXCL11	1	MECOM
81	3	169.76	170.24	CD48, KLRB1, LTA, PDCD1, SELL	5	ACTRT3, GPR160, KMT5AP3, LRRC34, MYNN, ...
82	3	186.94	187.04	FASLG, LTA, PDCD1	3	ST6GAL1
83	3	187.91	187.96	PRSS27	1	
84	4	3.43	3.48	CD48, FASLG	2	DOK7, HGFAC, RGS12
85	4	13.02	13.10	NCR1	1	U6
86	4	39.34	39.79	FGF19	1	KLB, LIAS, MIR5591, RFC1, RN7SL558P, ...
87	4	54.48	54.68	NCR1, SELL	2	KIT, LINC02260
88	4	55.36	55.62	KLRB1	1	CLOCK, LINC02928, NMU, PDCL2, RN7SKP30, ...
89	4	76.44	76.49	REG1B	1	SHROOM3
90	4	86.35	87.23	TCN1	1	AFF1, C4orf36, GAPDHP60, KLHL8, MAPK10, ...
91	4	104.80	105.31	CCL19, PRSS27, SELL	3	RN7SL89P, RNU6-351P, RPL6P14, TET2, TET2-AS1
92	4	121.25	121.26	CXCL10	1	
93	4	141.46	141.89	FASLG, KLRB1, NCR1	3	IL15, LINC02276
94	5	1.28	1.37	FASLG, KLRB1, SELL	3	CLPTM1L, LINC01511, MIR4457, TERT

3. Identification of putative core proteins for T1D using the GATE analysis 68

95	5	30.74	30.75	<i>CCL19</i>	1	
96	5	35.73	36.03	<i>CD5, KLRB1, PDCD1</i>	3	IL7R , <i>CAPSL</i> , <i>SPEF2</i> , <i>UGT3A1</i>
97	5	60.75	60.82	<i>AGR2</i>	1	<i>ELOVL7</i>
98	5	69.19	69.55	<i>LAG3</i>	1	<i>AK6, CCDC125, CDK7</i> , <i>CENPH, CFL1P5</i> , ...
99	5	71.37	71.75	<i>LAG3</i>	1	<i>BDP1, CARTPT</i> , <i>HMGN1P12, LINC02197</i> , <i>MCCC2</i> , ...
100	5	95.82	96.49	<i>CCK, GCG</i>	2	<i>ELL2, FABP5P5, GGCTP1</i> , <i>GLRX</i> , ...
101	5	130.72	132.80	<i>CRTAM, CXCL10, FASLG</i> , <i>FCRL6, KLRB1, LAG3</i> , <i>NCR1</i>	7	IRF1 , <i>IRF1-AS1, ACSL6</i> , <i>ACSL6-AS1</i> , ...
102	5	134.49	134.53	<i>CFC1</i>	1	<i>JADE2, LINC01843</i> , <i>RN7SL541P, RNU6-456P</i>
103	5	138.25	138.29	<i>PDCD1</i>	1	<i>CDC25C, GFRA3</i> , <i>RPS27AP18</i>
104	5	139.24	139.63	<i>CXCL10, CXCL11</i>	2	<i>DJC18, ECSCR, H3P25</i> , <i>MATR3, MZB1</i> , ...
105	5	159.34	159.40	<i>CXCL9</i>	1	
106	5	172.87	172.88	<i>PRSS27</i>	1	<i>ERGIC1</i>
107	5	177.11	177.39	<i>LGALS9, NCR1</i>	2	<i>LMAN2, MXD3, NSD1</i> , <i>PRELID1, PRMT1P1</i> , ...
108	5	179.98	180.07	<i>NCR1</i>	1	<i>MIR340, RNF130</i>
109	6	1.72	2.46	<i>PRSS27</i>	1	<i>GMDS, GMDS-DT</i> , <i>HMGN2P28</i>
110	6	6.99	7.33	<i>LTA, REG1B, TCN1</i>	3	<i>CAGE1, RREB1, SSR1</i>
111	6	10.52	10.54	<i>SELL</i>	1	<i>GCNT2</i>
112	6	21.86	21.90	<i>CD5</i>	1	<i>CASC15</i>
113	6	22.29	22.38	<i>TCN1</i>	1	<i>CASC15, PRL</i>
114	6	36.32	36.33	<i>CCL19</i>	1	<i>BNIP5</i>
115	6	36.75	36.77	<i>PRSS27</i>	1	<i>CPNE5</i>
116	6	39.06	39.34	<i>CCK</i>	1	<i>ANKRD18EP, GLP1R</i> , <i>KCNK16, KCNK17</i> , <i>KCNK5</i> , ...
117	6	44.73	45.04	<i>FCRL6</i>	1	<i>NUDT19P4, SUPT3H</i>
118	6	51.28	51.30	<i>PDCD1</i>	1	
119	6	52.52	52.60	<i>TCN1</i>	1	<i>EFHC1, TRAM2</i> , <i>TRAM2-AS1</i>
120	6	66.60	66.62	<i>FASLG</i>	1	
121	6	87.07	87.71	<i>SELL, TCN1</i>	2	<i>AKIRIN2, C6orf163</i> , <i>CFAP206, CGA, GJB7</i> , ...
122	6	90.27	90.30	<i>PDCD1</i>	1	BACH2
123	6	95.44	95.72	<i>CD48, FASLG</i>	2	<i>CYCSP17, MANEA</i> , <i>MANEA-DT</i>
124	6	106.18	106.33	<i>CD48</i>	1	<i>ATG5, RN7SL47P</i> , <i>RNU6-344P, U4</i>
125	6	109.33	109.78	<i>TCN1</i>	1	<i>AK9, CCDC162P, CD164</i> , <i>FIG4, MICAL1</i> , ...
126	6	115.86	116.76	<i>IDO1, LGALS9</i>	2	<i>CALHM4, CALHM5</i> , <i>CALHM6, CALHM6-AS1</i> , <i>CBX3P9</i> , ...
127	6	118.81	118.89	<i>KLRB1</i>	1	<i>ASF1A, MCM9</i>
128	6	119.33	119.55	<i>CD48</i>	1	<i>MAN1A1</i>
129	6	126.34	127.20	<i>CFC1, REG1B</i>	2	CENPW , <i>MIR588</i> , <i>PRELID1P1, RNU6-200P</i> , <i>RPS4XP9</i> , ...
130	6	129.54	129.55	<i>LTA</i>	1	

3. Identification of putative core proteins for T1D using the GATE analysis 69

131	6	137.64	137.69	<i>CRTAM</i>	1	<i>LINC03004</i>	<i>TNFAIP3</i>
132	6	139.52	139.52	<i>FASLG</i>	1		
133	6	148.07	148.10	<i>AGR2</i>	1		
134	6	159.05	159.09	<i>PDCD1</i>	1	<i>TAGAP-AS1</i>	
135	6	159.70	161.17	<i>CD5, CXCL9</i>	2	<i>ACAT2, AGPAT4, AIRN, CHP1P2, HNRNPH1P1, ...</i>	
136	7	6.46	6.51	<i>FASLG, SELL, TCN1</i>	3	<i>DAGLB, GRID2IP, KDELR2</i>	
137	7	28.10	28.17	<i>CXCL9, LTA</i>	2	<i>JAZF1</i> , <i>RNU6-979P</i>	
138	7	31.44	31.44	<i>AGR2</i>	1		
139	7	44.96	45.09	<i>KLRB1</i>	1	<i>CCM2, MYO1G, CAD, SNHG15, SNORA9</i>	
140	7	50.22	50.32	<i>KLRB1, LAG3, SELL</i>	3	<i>IKZF1</i>	
141	7	64.46	64.62	<i>LGALS9</i>	1	<i>BNIP3P42, HNRNCP7, ZNF680, ZNF680P1</i>	
142	7	73.40	73.64	<i>FASLG, FGF19</i>	2	<i>BAZ1B, BCL7B, FZD9, MLXIPL, RNU6-1198P, ...</i>	
143	7	75.62	75.64	<i>FASLG</i>	1	<i>HIP1</i>	
144	7	80.54	80.66	<i>PRSS27</i>	1	<i>CD36</i>	
145	7	92.78	92.78	<i>SELL</i>	1	<i>CDK6</i>	
146	7	99.28	99.77	<i>FASLG</i>	1	<i>ARPC1A, ARPC1B, ATP5MF, ATP5MF-PTCD1, BUD31, ...</i>	
147	7	100.22	100.92	<i>CXCL11, FASLG, NCR1</i>	3	<i>ACHE, ACTL6B, AGFG2, C7orf61, CASTOR3, ...</i>	
148	7	101.40	102.30	<i>CXCL11</i>	1	<i>COL26A1, CUX1, LINC01007, MIR4285, MYL10, ...</i>	
149	7	106.67	106.74	<i>CXCL11</i>	1		
150	7	121.32	121.40	<i>TCN1</i>	1	<i>FAM3C, WNT16</i>	
151	7	128.93	129.14	<i>CCL19, CXCL10, LAG3</i>	3	<i>CYCSP20, IRF5, ODCP, RN7SL306P, TNPO3, ...</i>	
152	7	148.94	149.21	<i>TCN1</i>	1	<i>COX6B1P1, GHET1, PDIA4, RN7SL521P, RNU6-650P, ...</i>	
153	7	154.72	154.74	<i>KLRB1, NCR1</i>	2	<i>DPP6</i>	
154	7	156.46	156.47	<i>CD5</i>	1	<i>RNF32-DT</i>	
155	8	4.77	4.84	<i>CCK</i>	1	<i>CSMD1, PAICSP4</i>	
156	8	6.94	6.96	<i>TCN1</i>	1	<i>DEFA8P</i>	
157	8	8.67	8.78	<i>CXCL9</i>	1	<i>CLDN23</i>	
158	8	11.63	11.68	<i>FGF19</i>	1	<i>GATA4</i>	
159	8	21.76	22.00	<i>FASLG</i>	1	<i>DOK2, GFRA2, OR6R2P, XPO7</i>	
160	8	23.54	23.56	<i>IDO1</i>	1	<i>SLC25A37</i>	
161	8	42.12	42.15	<i>CXCL9</i>	1	<i>AP3M2</i>	
162	8	58.19	58.66	<i>AGR2, FGF19</i>	2	<i>CYP7A1, FAM110B, NSMAF, PPIAP85, SDCBP, ...</i>	
163	8	59.11	59.28	<i>FCRL6, NCR1</i>	2	<i>TOX</i> , <i>TOX-DT</i>	
164	8	60.68	60.84	<i>IDO1, SELL</i>	2	<i>CHD7</i>	
165	8	78.43	78.94	<i>CD48, CD5, KLRB1, LGALS9</i>	4	<i>C4orf46P3, IL7, LINC02605, PKIA, PKIA-AS1, ...</i>	
166	8	105.52	105.58	<i>PRSS27</i>	1	<i>ZFPM2, ZFPM2-AS1</i>	
167	8	115.95	116.00	<i>REG1B</i>	1	<i>LINC00536</i>	
168	8	129.58	129.61	<i>FASLG</i>	1	<i>CCDC26</i>	
169	8	143.90	144.01	<i>IDO1</i>	1	<i>GRI, MIR661, PARP10, PLEC</i>	
170	9	5.48	5.50	<i>PDCD1</i>	1		
171	9	16.84	16.98	<i>FASLG, NCR1</i>	2	<i>BNC2</i>	

3. Identification of putative core proteins for T1D using the GATE analysis 70

172	9	32.20	32.58	<i>CXCL10, CXCL11, LGALS9</i>	3	<i>ACO1, DDX58, DFFBP1, NDUFB6, R5SP281, ...</i>	
173	9	33.09	33.18	<i>PRSS27</i>	1	<i>B4GALT1, B4GALT1-AS1, RNU4ATAC15P</i>	
174	9	35.72	36.52	<i>PRSS27</i>	1	<i>CCIN, CLTA, CREB3, FAM221B, GBA2, ...</i>	
175	9	89.00	89.00	<i>CXCL11</i>	1	<i>S1PR3</i>	
176	9	91.14	91.39	<i>NCR1</i>	1	<i>AUH, LINC00484, LINC02937</i>	
177	9	93.12	93.31	<i>IDO1</i>	1	<i>NINJ1, WNK2</i>	
178	9	97.67	98.13	<i>PRSS27</i>	1	<i>ANP32B, CORO2A, FOXE1, HEMGN, KRT18P13, ...</i>	
179	9	110.57	110.60	<i>CXCL10</i>	1	<i>SVEP1</i>	
180	9	113.21	113.37	<i>TCN1</i>	1	<i>BSPRY, CDC26, FKBP15, PRPF4, RNF183, ...</i>	
181	9	114.57	114.94	<i>CRTAM</i>	1	<i>ATP6V1G1, DELEC1, TEX48, TEX53, TMEM268, ...</i>	
182	9	120.75	120.96	<i>PDCD1</i>	1	<i>B3GALT9, C5, CUTALP, FBXW2, PHF19, ...</i>	
183	9	122.84	123.28	<i>FASLG</i>	1	<i>GPR21, MIR600, MIR600HG, RABGAP1, RC3H2, ...</i>	
184	9	132.99	133.49	<i>CCK, CXCL11, FGF19, GCG, KLRB1, LGALS9, PRSS27, REG1B, SELL</i>	9	<i>ABO, ADAMTS13, CACFD1, GF11B, LCN1P1, ...</i>	
185	9	136.41	136.49	<i>LAG3</i>	1	<i>C9orf163, ENTR1, INPP5E, PMPCA, SEC16A</i>	
186	10	3.10	3.10	<i>CXCL11</i>	1	<i>PFKP</i>	
187	10	5.20	5.28	<i>FASLG</i>	1	<i>AKR1C4, AKR1C7P, ARL4AP3, LINC02561</i>	
188	10	5.69	6.14	<i>FASLG, KLRB1, LTA, NCR1, PDCD1</i>	5	<i>IL2RA</i> , <i>ANKRD16, FBH1, GDI2, IL15RA, ...</i>	<i>IL2RA</i>
189	10	8.41	8.45	<i>FASLG, NCR1</i>	2		
190	10	48.15	48.16	<i>CFC1</i>	1	<i>FRMPD2</i>	
191	10	62.10	62.20	<i>KLRB1</i>	1	<i>RTKN2</i>	
192	10	63.06	63.73	<i>CXCL11</i>	1	<i>JMJD1C, JMJD1C-AS1, MIR1296, MRPL35P2, NRBF2, ...</i>	
193	10	69.38	69.54	<i>CD5, PDCD1</i>	2	<i>ATP5MC1P7, HK1, TACR2, TMEM256P1, TSPAN15</i>	
194	10	71.75	71.80	<i>NCR1</i>	1	<i>CDH23, MIR7152, VSIR</i>	
195	10	79.15	79.30	<i>FASLG, NCR1</i>	2	<i>ZMIZ1</i>	
196	10	80.28	80.52	<i>PDCD1</i>	1	<i>DYDC1, DYDC2, MAT1A, PRXL2A, TSPAN14, ...</i>	
197	10	88.25	89.25	<i>FASLG, NCR1</i>	2	<i>RNLS</i> , <i>ACTA2, ACTA2-AS1, ANKRD22, CH25H, ...</i>	
198	10	91.68	92.80	<i>FASLG, NCR1</i>	2	<i>BTAFF1, CPEB3, EIF2S2P3, EIF4A1P8, FAF2P1, ...</i>	
199	10	102.06	103.24	<i>CXCL11, PRSS27</i>	2	<i>ACTR1A, ARL3, AS3MT, BORCS7, BORCS7-ASMT, ...</i>	
200	10	113.31	113.88	<i>CCK</i>	1	<i>CASP7, DCLRE1A, HABP2, MIR4483, NHLRC2, ...</i>	

3. Identification of putative core proteins for T1D using the GATE analysis 71

201	10	122.31	122.74	REG1B, TCN1	2	ARMS2, BTBD16, C10orf120, DMBT1, DMBT1L1, ...
202	10	124.61	125.05	FCRL6, GCG	2	ABRAXAS2, CTBP2, EEF1AKMT2, FAM53B, FAM53B-AS1, ...
203	11	1.15	1.15	AGR2	1	
204	11	2.17	2.31	CFC1, TCN1	2	C11orf21, MIR4686, TSPAN32
205	11	8.22	8.24	CFC1	1	LMO1
206	11	46.93	47.69	FASLG, NCR1, TCN1	3	ACP2, AGBL2, ARFGAP2, C1QTNF4, CELF1, ...
207	11	54.57	55.74	CFC1	1	OR4A10P, OR4A11P, OR4A12P, OR4A13P, OR4A15, ...
208	11	60.35	60.41	LAG3	1	MS4A14, MS4A6E, MS4A7
209	11	61.77	61.87	CXCL11, FASLG, NCR1	3	FADS2, FADS1, FADS3, FEN1, MIR1908, ...
210	11	65.48	66.01	PRSS27	1	AP5B1, BANF1, C11orf68, CCDC85B, CFL1, ...
211	11	95.51	95.54	PRSS27	1	
212	11	95.76	95.86	LGALS9	1	CEP57, FAM76B, MTMR2, R5SP345
213	11	100.77	100.83	REG1B	1	ARHGAP42
214	11	126.11	126.55	CD48, FASLG, PRSS27, SDK2, SELL, TCN1	6	DCPS, FAM118B, FOXRED1, GSEC, ...
215	11	128.18	128.54	CRTAM, FASLG, FCRL6, IDO1, KLRB1, NCR1	6	ETS1, ETS1-AS1, LINC02098, LINC02725, MIR6090
216	11	134.74	134.74	CD48	1	LINC02714
217	12	0.47	0.55	FCRL6, LTA, PRSS27	3	B4GALNT3
218	12	6.33	6.58	LTA	1	CD27, CD27-AS1, CHD4, GAPDH, GAPDH-DT, ...
219	12	10.34	10.49	FASLG, NCR1	2	KLRC1, KLRC2, KLRC3, KLRC4, KLRC4-KLRK1, ...
220	12	45.69	45.99	NCR1	1	ARID2, KNOP1P2, LINC00938, RN7SL246P, SCAF11
221	12	51.28	51.38	NCR1	1	BIN2, CELA1, GALNT6
222	12	53.11	53.11	FCRL6	1	SOAT2
223	12	54.20	54.36	CXCL11, PRSS27	2	CBX5, COPZ1, HNRNPA1, MIR148B, MIR3198-2, ...
224	12	56.47	56.53	CFC1	1	GLS2, HSPD1P4, MIP, RBMS2, SPRYD4
225	12	56.99	57.10	LAG3	1	GPR182, MYO1A, B2, NEMP1, STAT6, ...
226	12	62.89	62.92	TCN1	1	PPM1H
227	12	68.12	68.35	CXCL10, CXCL9, REG1B	3	IFNG, IFNG-AS1, IL22, IL26, LINC02384, ...
228	12	89.20	89.24	PRSS27	1	LINC02458
229	12	101.47	101.49	LAG3	1	SPIC
230	12	109.72	112.85	CCL19, CD48, CD5, CRTAM, CXCL10, CXCL11, CXCL9, FASLG, FCRL6, IDO1, KLRB1, LAG3, LGALS9, LTA, NCR1, PDCD1, SDK2, SELL, TIGIT	19	SH2B3, ACAD10, ADAM1A, ADAM1B, ALDH2, ...
231	12	115.43	115.50	GCG	1	

3. Identification of putative core proteins for T1D using the GATE analysis 72

232	12	118.58	118.60	LAG3	1	
233	12	120.70	121.03	LGALS9, PDCD1	2	ACADS, ARF1P2, C12orf43, CLIC1P1, HNF1A, ...
234	12	123.26	123.40	PRSS27	1	CDK2AP1, KMT5A, MIR8072, MTRFR, R5SP375, ...
235	13	28.02	28.12	SELL	1	CHCHD2P8, FLT3, KATNBL1P1
236	13	40.56	40.94	TCN1	1	CYCSP34, ELF1, FOXO1, Metazoa SRP, MIR320D1, ...
237	13	72.25	72.32	IDO1	1	
238	13	74.11	74.13	FCRL6	1	KLF12
239	13	75.25	75.26	NCR1	1	LINC01078
240	13	108.19	108.35	LTA	1	ABHD13, LIG4, R5SP39, TNFSF13B
241	14	23.11	23.13	TCN1	1	CEBPE, SLC7A8
242	14	25.03	25.04	CCL19	1	STXBP6
243	14	64.43	64.67	NCR1	1	AKAP5, HSPA2, HSPA2-AS1, MTHFD1, NUP50P1, ...
244	14	65.34	65.69	CD48	1	EIF1AXP2, FUT8, FUT8-AS1, MIR625, RPL21P8
245	14	67.49	67.52	SELL	1	TMEM229B
246	14	68.61	68.85	CD5, LTA	2	RAD51B
247	14	77.08	77.14	KLRB1	1	CIPC, TMEM63C, ZDHHC22
248	14	89.44	89.44	LGALS9	1	FOXN3
249	14	92.54	92.56	FASLG, IDO1	2	RIN3
250	14	94.32	94.39	CCK	1	SERPI1, SERPI2, SERPI6
251	14	99.26	99.39	FASLG, NCR1	2	BCL11B
252	15	31.34	31.37	FASLG	1	KLF13
253	15	49.95	51.33	KLRB1	1	AHCYP7, AP4E1, ATP8B4, CYP19A1, DCAF13P3, ...
254	15	58.46	58.50	CD5, PDCD1	2	ALDH1A2, LIPC, LIPC-AS1
255	15	60.39	60.52	KLRB1	1	ANXA2, ICE2, RORA, RORA-AS1
256	15	61.82	62.39	CFC1	1	C2CD4A, C2CD4B, GOLGA2P11, HMGN1P26, Metazoa SRP, ...
257	15	64.05	64.81	PRSS27	1	CIAO2A, CSNK1G1, DAPK2, GAPDHP61, MIR1272, ...
258	15	69.69	69.75	PDCD1	1	DRAIC
259	15	77.15	77.26	PDCD1	1	PEAK1
260	15	78.93	78.96	LGALS9	1	CTSH, RASGRF1
261	15	79.96	79.99	IDO1	1	BCL2A1
262	16	0.16	0.52	CFC1	1	ARHGDIG, AXIN1, C4orf46P1, DECR2, FAM234A, ...
263	16	8.48	8.48	LGALS9	1	
264	16	8.93	8.96	CXCL11	1	USP7
265	16	11.02	11.35	NCR1, PDCD1	2	CLEC16A, HNRNPCP4, MIR548H2, PRM1, PRM2, ...
266	16	11.57	11.63	FASLG	1	LITAF
267	16	17.27	17.28	AGR2	1	XYLT1

3. Identification of putative core proteins for T1D using the GATE analysis 73

268	16	20.34	20.40	AGR2, KLRB1, NCR1, PRSS27, REG1B	5	UMOD, PDILT
269	16	28.50	28.93	CD5, FASLG	2	IL27, ATP2A1, ATP2A1-AS1, ATXN2L, CDC37P1, ...
270	16	30.44	31.36	CD5	1	BCKDK, BCL7C, CFAP119, CTF1, CTF2P, ...
271	16	57.04	57.05	KLRB1	1	NLRC5
272	16	75.20	75.79	REG1B	1	BCAR1, ADAT1, ATP5PBP7, CFDP1, CHST5, ...
273	16	81.82	81.86	CXCL11	1	PLCG2
274	16	85.38	85.38	CXCL11	1	GSE1
275	16	85.65	85.66	CXCL9	1	GSE1, RN7SL381P
276	16	85.96	85.98	NCR1	1	
277	16	88.58	88.59	FGF19	1	ZC3H18
278	16	89.63	89.80	LGALS9, REG1B	2	CDK10, CHMP1A, DPEP1, FANCA, LINC02166, ...
279	17	2.24	2.30	NCR1	1	SMG6
280	17	5.38	5.51	KLRB1	1	C1QBP, DERL2, DHX33, DHX33-DT, MIS12, ...
281	17	6.83	7.63	CD48, CRTAM, FASLG, LAG3, PDCD1, SDK2, SELL, TCN1	8	NEURL4, ACADVL, ACAP1, ALOX12, ALOX12-AS1, ...
282	17	9.86	9.90	GCG	1	GLP2R, RCVRN
283	17	16.55	16.90	LTA	1	CCDC144A, COTL1P1, FAM106C, KRT16P2, KRT16P6, ...
284	17	21.54	21.58	LGALS9	1	LINC02693
285	17	35.29	35.66	CCL19, CXCL10, CXCL9, FASLG, FCRL6, IDO1, KLRB1, NCR1	8	AP2B1, E2F3P1, PEX12, SLFN11, SLFN12, ...
286	17	39.42	40.39	FASLG, KLRB1, LTA, NCR1, SELL, TCN1	6	GSDMB, CASC3, CDC6, CDK12, CSF3, ...
287	17	40.99	41.07	SDK2	1	KRTAP1-1, KRTAP1-3, KRTAP1-4, KRTAP1-5, KRTAP2-1, ...
288	17	45.12	46.79	CD48, CD5, CRTAM, FASLG, IDO1	5	MAPT, MAPT-AS1, MAPT-IT1, ACBD4, ARHGAP27, ...
289	17	47.31	47.72	FASLG	1	EFCAB13, EFCAB13-DT, ITGB3, KPNB1, KPNB1-DT, ...
290	17	48.07	48.25	GCG	1	CBX1, MIR1203, RNU6-1152P, RNU6-1201P, SKAP1, ...
291	17	49.24	49.40	PDCD1	1	FLJ40194, MIR6129, ZNF652, ZNF652-AS1
292	17	62.48	62.60	CXCL10	1	TLK2
293	17	64.04	64.12	NCR1	1	ERN1
294	17	67.09	67.26	FCRL6	1	HELZ, HELZ-AS1, RPL36AP48
295	17	68.28	69.55	CCK, CCL19, CFC1, CXCL11, FGF19, GCG, LAG3	7	ABCA10, ABCA5, ABCA6, ABCA8, ABCA9, ...
296	17	76.60	76.66	TCN1	1	RNU6-227P, SNHG16, ST6GALC1
297	17	78.12	78.15	CD5	1	C17orf99, TMC6, TMC8

3. Identification of putative core proteins for T1D using the GATE analysis 74

298	17	81.17	81.30	CD48, CD5, FASLG, FCRL6, KLRB1, NCR1, PDCD1, SELL, TCN1	9	CEP131, NDUFAF8, PVALEF, SLC38A10, TEPSIN	
299	18	9.60	9.62	PRSS27	1	PPP4R1, PPP4R1-AS1, RNU6-903P	
300	18	12.77	12.88	CXCL10, CXCL9	2	PTPN2, LINC01882	
301	18	22.03	22.10	FGF19, GCG	2	LINC01900	
302	18	23.01	23.03	FGF19	1	RBBP8	
303	18	25.90	26.22	TCN1	1	DHFRP1, NPM1P2, PSMA8, RPS24P18, SS18	
304	18	44.51	44.64	SELL	1	LINC01478, MLECP1	
305	18	48.87	48.93	GCG	1	SMAD7	
306	18	54.25	54.42	KLRB1	1	C18orf54, CUPIN1P, POLI, SNRPGP2, STARD6	
307	19	0.81	0.81	FASLG	1	PTBP1	
308	19	3.95	4.12	PDCD1	1	DAPK3, EEF2, MAP2K2, MIR637, PIAS4, ...	
309	19	4.35	4.52	FASLG	1	CHAF1A, HDGFL2, MIR4746, MPND, PLIN4, ...	
310	19	5.76	5.85	FGF19, LGALS9, PRSS27, REG1B, TCN1	5	CATSPERD, DUS3L, FUT3, FUT6, NRTN, ...	
311	19	6.21	6.28	FCRL6	1	MLLT1	
312	19	6.61	6.68	LTA	1	TNFSF14	
313	19	10.21	11.08	CXCL10, CXCL9, IDO1, PDCD1	4	TYK2, AP1M2, ATG4D, C19orf38, CARM1, ...	
314	19	13.77	13.85	FCRL6, TCN1	2	C19orf53, MIR23A, MIR23AHG, MIR24-2, MIR27A, ...	
315	19	16.30	16.46	CD5, CRTAM	2	EPS15L1, KLF2, KLF2-DT	
316	19	17.38	17.43	CD5	1	BISPR, BST2, CCDC194, MVB12A, PLVAP	
317	19	17.79	17.81	FGF19, PRSS27, REG1B	3	B3GNT3	
318	19	33.26	33.27	TCN1	1		
319	19	38.77	38.80	KLRB1	1	CAPN12, LGALS7, LGALS7B, RNU6-140P	
320	19	42.03	42.25	FASLG	1	DEDD2, ERF, GRIK5, GSK3A, MIR4323, ...	
321	19	44.89	44.92	CXCL10	1	APOC1, APOE, NECTIN2, TOMM40	
322	19	48.10	48.97	AGR2, CCK, CD5, FGF19, GCG, LGALS9, PDCD1, PRSS27, REG1B, SDK2, SELL, TCN1	12	FUT2, BAX, BCAT2, CA11, CARD8, ...	
323	19	53.79	53.83	IDO1	1	NLRP12	
324	19	54.86	55.06	CXCL11, KLRB1, PRSS27	3	FCAR, GP6, GP6-AS1, KIR3DL2, NLRP2, ...	
325	20	19.26	19.35	CXCL11	1	SLC24A3, SLC24A3-AS1	
326	20	33.93	35.24	FASLG	1	ACSS2, AHCY, ASIP, CDC42P1, DYNLRB1, ...	ITCH
327	20	36.88	36.94	FCRL6	1	SAMHD1, TLDC2	
328	20	44.00	44.05	LTA, NCR1	2	RN7SL443P, TOX2	
329	20	58.65	58.70	CD5	1	NPEPL1, STX16, STX16-NPEPL1	
330	20	62.71	62.71	FGF19	1	NTSR1	
331	21	13.89	13.99	FASLG	1	ANKRD20A11P, RHOT1P2, RNU6-954P, SNX18P13	

3. Identification of putative core proteins for T1D using the GATE analysis 75

332	21	39.35	39.70	<i>PRSS27</i>	1	<i>B3GALT5, B3GALT5-AS1, GET1, GET1-SH3BGR, HMGN1, ...</i>	
333	21	41.34	41.36	<i>PRSS27</i>	1	<i>FAM3B</i>	
334	21	44.29	44.29	<i>PDCD1</i>	1	AIRE	AIRE
335	22	25.11	25.24	<i>LAG3, PRSS27</i>	2	<i>KIAA1671, KIAA1671-AS1</i>	
336	22	28.16	28.74	<i>FASLG</i>	1	<i>CHEK2, MIR5739, RN7SL162P, SNORD42, TTC28</i>	
337	22	29.73	30.20	<i>CCL19</i>	1	<i>ASCC2, CNN2P1, HORMAD2, HORMAD2-AS1, MIR6818, ...</i>	
338	22	38.11	38.25	<i>KLRB1</i>	1	<i>BAIAP2L2, MAFF, PLA2G6, RN7SL704P, TMEM184B</i>	
339	22	39.26	39.46	<i>FASLG, SDK2</i>	2	<i>MGAT3, RPL3, SNORD43, SNORD83A, SNORD83B, ...</i>	
340	22	41.02	41.96	<i>KLRB1, LTA</i>	2	<i>ACO2, ACTBP15, C22orf46, CCDC134, CENPM, ...</i>	
341	22	42.19	42.29	<i>PDCD1</i>	1	<i>OGFRP1, TCF20</i>	
342	22	43.93	44.00	<i>LTA</i>	1	<i>PARVB, PNPLA3, RPL35AP36, SAMM50</i>	
343	22	50.45	50.53	<i>FASLG, FCRL6</i>	2	<i>LMF2, NCAPH2, ODF3B, SBF1, SCO2, ...</i>	

For each clump that harbours *trans*-pQTLs for a target protein, a locus-specific *trans*-score for that clump and that protein is computed. The pGATE score for each protein is then computed by adding up these locus-specific *trans*-scores across all clumps. Genes within 200 kb from the clump are displayed as nearby genes, among which previously reported type 1 diabetes GWAS hits are shown in bold. Only the first 5 genes are listed for each clump, and three dots denote the presence of more than 5 genes in that location. A single dot denotes that the clump cannot be annotated by genes in Ensembl. Monogenic causes for autoimmune diabetes within 200 kb from the clump are displayed as monogenic genes.

Table S3.3: Mendelian randomisation analysis of the effect of the putative core proteins on type 1 diabetes.

Gene symbol	Scottish				UK Biobank			
	Raw number of pQTLs	Estimate	Standard error	p-value	Raw number of pQTLs	Estimate	Standard error	p-value
<i>PDCD1</i>	37	0.788	0.280	0.005	39	0.501	0.193	0.009
<i>CXCL9</i>	18	0.897	0.356	0.01	20	0.370	0.238	0.1
<i>CXCL10</i>	19	1.051	0.368	0.004	19	0.504	0.270	0.06
<i>CRTAM</i>	13	0.947	0.536	0.08	12	0.561	0.445	0.2
<i>CCL19</i>	13	1.033	0.557	0.06	15	0.934	0.434	0.03
<i>LGALS9</i>	17	1.116	0.374	0.003	16	0.437	0.248	0.08
<i>CD5</i>	27	0.551	0.263	0.04	29	0.304	0.115	0.008
<i>PRSS27</i>	35	0.229	0.073	0.002	38	-0.001	0.062	1
<i>FASLG</i>	56	0.237	0.071	8×10^{-4}	59	0.142	0.050	0.005
<i>IDO1</i>	17	0.710	0.244	0.004	18	0.386	0.140	0.006
<i>AGR2</i>	9	-1.369	0.744	0.07	9	-0.112	0.380	0.8
<i>FGF19</i>	14	-0.494	0.213	0.02	14	-0.133	0.126	0.3
<i>SDK2</i>	7	1.271	0.720	0.08	8	0.697	0.481	0.1
<i>SELL</i>	26	0.436	0.197	0.03	26	0.197	0.136	0.1
<i>GCG</i>	10	0.802	0.441	0.07	10	-0.145	0.281	0.6
<i>KLRB1</i>	39	0.356	0.124	0.004	42	0.165	0.119	0.2
<i>NCR1</i>	55	0.309	0.114	0.007	56	0.135	0.089	0.1
<i>CD48</i>	16	0.533	0.238	0.03	32	0.366	0.105	5×10^{-4}
<i>CCK</i>	9	0.424	0.229	0.06	9	-0.154	0.176	0.4
<i>CFC1</i>	16	-0.635	0.320	0.05	19	-0.534	0.241	0.03
<i>LTA</i>	21	0.491	0.221	0.03	23	0.232	0.167	0.2
<i>TCN1</i>	34	0.149	0.133	0.3	37	0.055	0.068	0.4
<i>CXCL11</i>	26	0.440	0.192	0.02	24	-0.155	0.138	0.3
<i>REG1B</i>	16	-0.394	0.208	0.06	16	-0.043	0.151	0.8
<i>FCRL6</i>	21	0.590	0.257	0.02	23	0.384	0.217	0.08
<i>TIGIT</i>	5	2.431	1.132	0.03	6	2.465	1.090	0.02
<i>LAG3</i>	19	0.391	0.263	0.1	36	0.384	0.114	8×10^{-4}

All *trans*-pQTLs contributing to the pGATE score of a putative core protein are used as genetic instruments for that protein. The maximum likelihood estimate, standard error, and *p*-value for the causal effect of each protein on type 1 diabetes is reported. Estimation was performed by marginalising over the posterior distribution of pleiotropic effects of the *trans*-pQTLs.

Table S3.4: Detailed supportive evidence of putative core proteins through perturbation in experimental animal models or drug effect in humans.

Gene	Experiment	Perturbation	Drug effect
<i>PDCD1</i>	[Wang et al., 2005]	<i>Pdcd1</i> knockout led to accelerated type 1 diabetes	[Cho and Jung, 2023]
<i>CXCL10</i>	[Lasch et al., 2015]	Anti-CXCL10 antibody induced remission of type 1 diabetes	.
<i>LGALS9</i>	Chou et al. [2009]	Overexpression of <i>Lgals9</i> (gal-9) reduced incidence of diabetes in NOD mice	.
<i>FASLG</i>	[Nakayama et al., 2002]	Anti-FASLG antibody prevented diabetes in NOD mice	.
<i>IDO1</i>	[Mondanelli et al., 2017]	Bortezomib increased IDO1 and prevented diabetes onset in NOD mice but the effect was absent in <i>Ido1</i> ^{-/-} NOD mice	.
<i>NCR1</i>	[Gur et al., 2010]	<i>Ncr1</i> knockout mice had less development of STZ induced type 1 diabetes	.
<i>REG1B</i>	[Unno et al., 2002]	In NOD mice carrying the <i>Ins-Reg</i> transgene development of diabetes was reduced	.
<i>TIGIT</i>	[Trivedi et al., 2024]	Partial antibody mediated inhibition TIGIT in combination with PD1-PDL1 inhibition resulted in rapid diabetes in NOD mice	.
<i>LAG3</i>	[Bettini et al., 2011]	<i>Lag3</i> knockout NOD mice developed substantially accelerated diabetes	.

NOD, non-obese diabetic mouse model, is a spontaneous model for type 1 diabetes. Mouse gene names are spelled using only uppercase the first letter and lowercase for the rest. *Ins-Reg* is the rat insulin 2 promoter/mouse *Reg* hybrid gene, which is used to target the overexpression of *Reg* in pancreatic β cells in transgenic mice.

The following sections provide additional descriptions of the GATE analyses in two datasets, and considerations of technical aspects that influence the properties and structure of the data I analysed as part of this chapter, their possible impact on GATE associations and how these impacts could be mitigated.

3.3 Characteristics of the GATE analysis in the Scottish and UK Biobank datasets

In the Scottish dataset, 4964 individuals with T1D and 7474 non-diabetic controls were included in the GATE analysis. Among T1D cases, the mean age at diagnosis was 23 years (SD = 14). The mean age at study recruitment was 49 years (SD = 15) for T1D cases and 58 years (SD = 13) for non-diabetic controls. A total of 2864 GATE scores had at least five effective *trans*-loci, including 71 GATE scores whose corresponding protein-coding genes are located on the X or Y chromosome. The genomic locations of the genes encoding proteins with a GATE score and the *p*-values of association of these GATE scores with T1D are shown in Figure S3.3.

In the UK Biobank dataset, 1638 T1D cases and 365K non-diabetic controls were included in the GATE analysis when testing GATE scores computed based on the pQTLs from the UKB-PPP dataset. The mean age at study recruitment

was 54 years (SD = 8) for T1D cases and 56 years (SD = 8) for non-diabetic controls, while the mean age at diagnosis was 29 years (SD = 13) among T1D cases. When testing GATE scores computed based on pQTLs from the other two proteomics studies, 1864 T1D cases and 408K non-diabetic controls were included. Among them, the mean age at study recruitment was 54 years (SD = 8) for T1D cases and 56 years (SD = 8) for controls, and the mean age at diagnosis was 29 years (SD = 13) for T1D cases. Overall, 3212 GATE scores had at least five effective *trans*-loci, including 82 GATE scores corresponding to protein-coding genes on the X or Y chromosome. The genomic locations of the genes encoding proteins with a GATE score and the *p*-values of association of these GATE scores with T1D are shown in Figure S3.4.

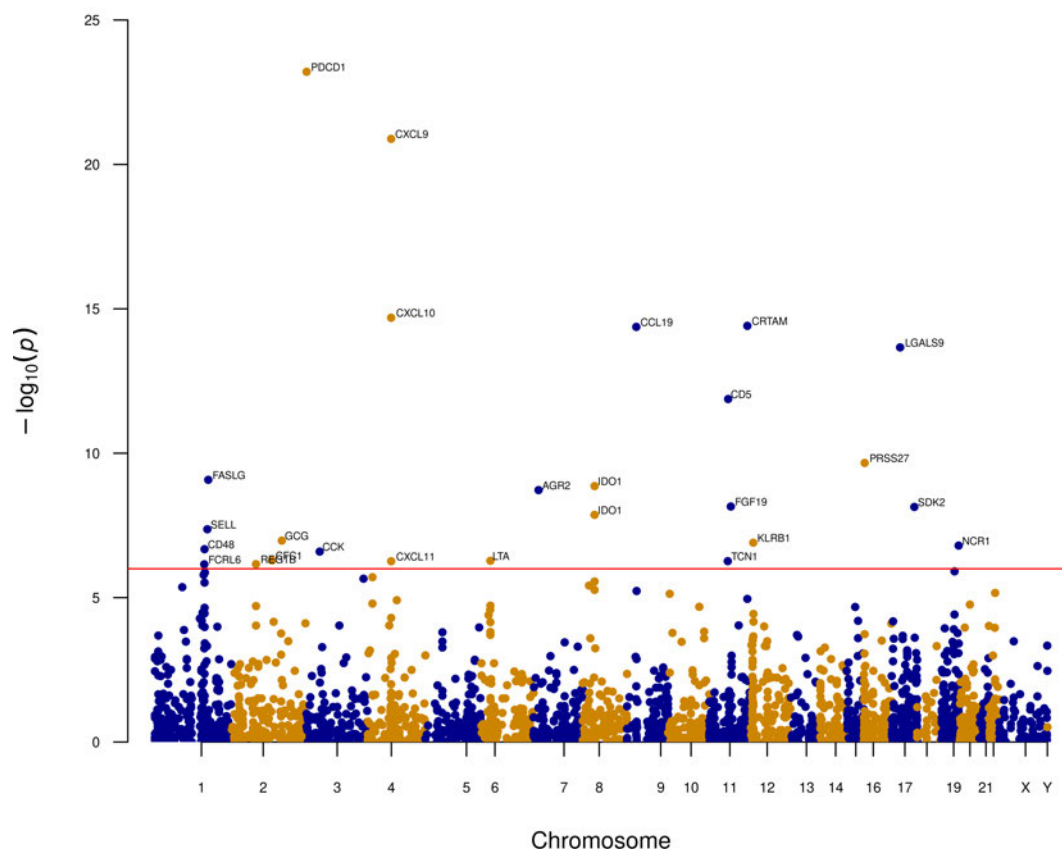


Figure S3.3: Manhattan plot of GATE scores with at least five effective *trans*-loci in the Scottish dataset.

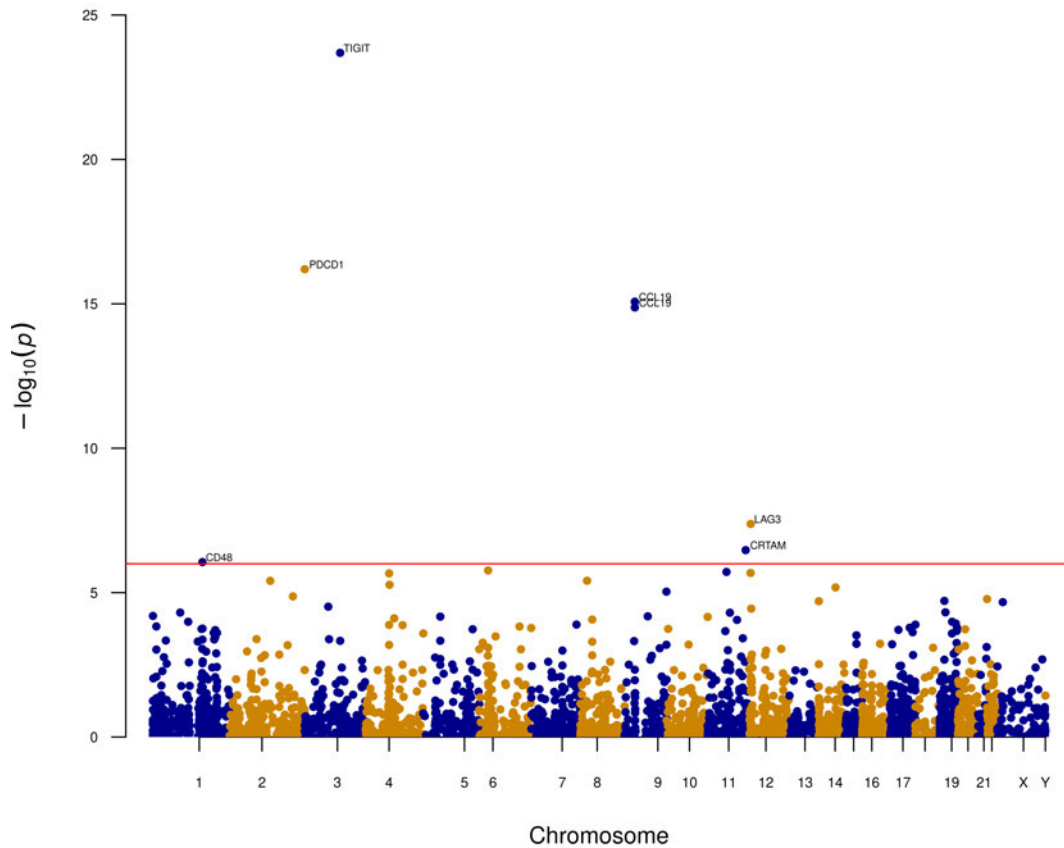


Figure S3.4: Manhattan plot of GATE scores with at least five effective *trans*-loci in the UK Biobank dataset.

3.4 Pleiotropy of master regulators

In this chapter, 27 putative core proteins were identified through the GATE analysis. As shown in Figure 3.2, GATE scores for these 27 core proteins formed two correlation blocks, driven by shared *trans*-pQTLs. In order to identify the genomic locations harbouring these shared signals, I clumped *trans*-pQTLs together if they were within 200 kb of each other and I annotated these clumps if they were within 200 kb from a known GWAS hit for T1D (Figure S3.5). Two master regulatory regions were identified, based on the corresponding clump having *trans*-effects on more than 10 putative core proteins. Both of these regions overlap known T1D GWAS hits. The first region is a *trans*-clump spanning from 109.72-112.85 Mb on chromosome 12, around the *SH2B3* gene, which contributes to the GATE scores of 19 immune-related proteins. The second spans from 48.10-

48.97 Mb on chromosome 19, around the *FUT2* gene, and contributes to the GATE scores of seven pancreatic proteins and five immune proteins.

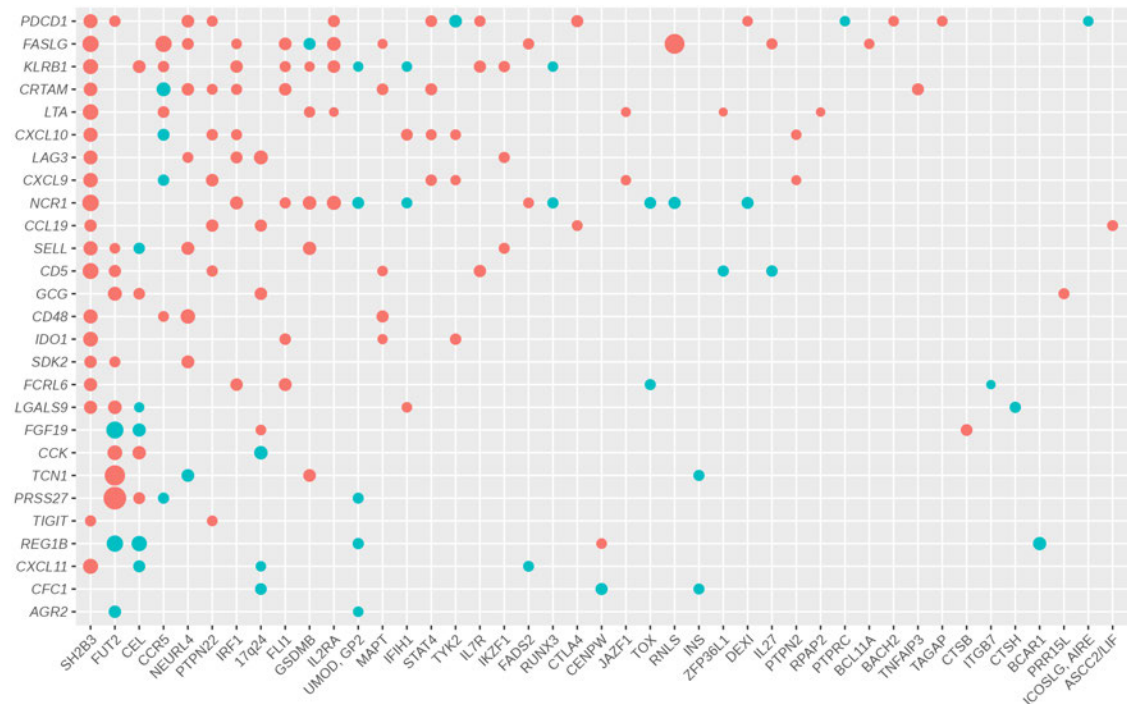


Figure S3.5: *Trans*-pQTLs that are also T1D GWAS hits contributing to the GATE scores of the putative core proteins. *Trans*-pQTLs overlapping T1D GWAS hits are listed on the x-axis and the putative core proteins are shown on the y-axis. Each dot represents a *trans*-pQTL; colour indicates the direction of effect of the *trans*-pQTL on T1D risk (red, increase risk; green, decrease risk) and point size reflects the effect size of the *trans*-pQTL on circulating protein levels.

As reported in both the deCODE and UKB-PPP pQTL studies, the *SH2B3* and *FUT2* loci are among the most pleiotropic *trans*-pQTLs [Ferkingstad et al., 2021, Sun et al., 2023]. However, the biological mechanisms underlying their pleiotropy are likely to be distinct. *SH2B3* encodes SH2B Adaptor Protein (also known as lymphocyte adaptor protein, LNK), which acts as a negative regulator of cytokine and growth factor signalling [Morris et al., 2021]. Through interactions with phosphorylated sites on Janus Kinase family proteins, including JAK2 and JAK3, SH2B3 plays a critical role in regulating the Janus Kinase-Signal Transducer and Activator of Transcription (JAK-STAT) signalling pathway, and consequently, hematopoiesis. Mutations in *SH2B3* gene have been associated with a wide spectrum of diseases, including cardiovascular diseases, cancers, and autoimmune conditions. In contrast, *FUT2* encodes α (1,2) fucosyltransferase that catalyses the addition of fucose to terminal glycans on glycoproteins and glycolipids, thereby determining secretor status in bodily fluids and regulating

host-microbiome interactions [Tong et al., 2014]. Therefore, pleiotropy at the *SH2B3* locus likely reflects its central role in intracellular signalling pathways, while pleiotropy at the *FUT2* locus is more plausibly driven by widespread fucosylation, a form of post-translational modification (PTM).

In order to characterise the effect of these master regulatory regions on the GATE associations, I first examined the results from the MR analyses. As described in section 3.2.5.2 of the manuscript, I used a novel MR approach that explicitly models pleiotropic (direct) effects of the instruments (derived from *trans*-pQTLs) on the outcome, and then marginalises over the pleiotropic effects to assess causality of the exposure (protein level) on the outcome (T1D risk). As shown in Figure S3.1 and Figure S3.2, the instruments at *SH2B3* and *FUT2* are “outliers” with respect to the maximum likelihood estimate for the causal effect (straight line), suggesting that a non-zero direct effect is inferred for these instruments, down-weighting their contribution through the exposure. Among the 19 immune proteins with *trans*-pQTLs in or near *SH2B3*, ten (*PDCD1*, *CXCL10*, *LGALS9*, *CD5*, *FASLG*, *IDO1*, *KLRB1*, *NCR1*, *CD48*, *LAG3*) were supported by MR analysis with the novel approach. This means that the remaining *trans*-pQTLs contributing to the GATE scores for these 19 proteins support a dose-response relationship, even when the contribution of *SH2B3* is downweighted. For the 12 core proteins with *trans*-pQTLs in or near *FUT2*, only three immune proteins (*PDCD1*, *LGALS9*, *CD5*) and one pancreatic protein (*PRSS27*) showed evidence of causality. This suggests that the signal from the remaining *trans*-pQTLs for the pancreatic proteins is not strong enough to support a causal effect when the contribution of *FUT2* is downweighted by modelling a direct effect on T1D risk (which could for instance be mediated by fucosylation in general rather than a specific protein).

I also used conventional MR methods, including the fixed- and random-effects inverse variance weighted (IVW) and weighted median approaches. While the novel MR approach showed evidence of causal effects for 11 of the 27 putative core proteins, the conventional IVW fixed-effects method suggested causal effects for all 27 proteins, suggesting inflation of type I error. The IVW random-effects and weighted median approaches supported 11 and 13 proteins ($p < 0.01$), respectively, of which six and nine overlapped with those identified by the novel MR method. Among the 11 proteins supported by the novel MR approach, nine and ten met the $p < 0.05$ threshold under the IVW random-effects and weighted median methods, respectively. Overall, these results indicated that the novel MR approach controlled the inflated type I error more effectively.

In addition to the MR analysis described above, I conducted a sensitivity analysis to examine the impact of directly accounting for the effect of *SH2B3* in the GATE analysis. Given that the direct effect of *SH2B3* on T1D is primarily driven by a missense variant (rs3184504*T) [Watson et al., 2025], I conducted a sensitivity analysis in the Scottish dataset to further account for its pleiotropy by adjusting for the genotype of this variant in the logistic regression model in the GATE analysis. As shown in Table S3.5, although effect sizes for these 19 proteins were attenuated relative to the original analysis, the directions of effect remained consistent and the average reduction of effect sizes is around one third. To further confirm that pleiotropic effects were adequately controlled, I compared the raw correlations between the GATE scores for these 19 core proteins with the partial correlations conditioning on the genotype of rs3184504. As illustrated in Figure S3.6 and Figure S3.7, the mean correlation was substantially reduced from 0.21 to 0.07, with over one third of correlations reduced to zero, suggesting that accounting for the missense variant removes the shared (pleiotropic) signal at the *SH2B3* locus.

Table S3.5: Comparison GATE score associations for 19 core proteins with *trans*-pQTLs in or nearby the *SH2B3* gene between the original GATE analysis and the sensitivity analysis in the Scottish dataset.

Gene symbol	Transcription site		pQTL study	Effective number of pQTLs	Original analysis ¹		Sensitivity analysis ²	
	Chr	Start position (Mb)			Log OR	p-value	Log OR	p-value
<i>PDCD1</i>	2	241.85	UKB	30.9	0.21	6×10^{-24}	0.18	1×10^{-17}
<i>CXCL9</i>	4	76.00	UKB	11.2	0.19	1×10^{-21}	0.16	4×10^{-13}
<i>CXCL10</i>	4	76.02	UKB	14.4	0.16	2×10^{-15}	0.12	7×10^{-8}
<i>CRTAM</i>	11	122.84	DeC	10.7	0.16	4×10^{-15}	0.12	9×10^{-9}
<i>CCL19</i>	9	34.69	DeC	10.2	0.16	4×10^{-15}	0.13	1×10^{-9}
<i>LGALS9</i>	17	27.63	UKB	13.1	0.16	2×10^{-14}	0.12	1×10^{-8}
<i>CD5</i>	11	61.10	UKB	16.1	0.14	1×10^{-12}	0.10	1×10^{-5}
<i>FASLG</i>	1	172.66	UKB	15.8	0.12	8×10^{-10}	0.10	4×10^{-6}
<i>IDO1</i>	8	39.90	UKB	11.6	0.12	1×10^{-9}	0.07	0.002
<i>SDK2</i>	17	73.33	UKB	5.8	0.12	7×10^{-9}	0.07	0.004
<i>SELL</i>	1	169.69	UKB	20.1	0.11	4×10^{-8}	0.07	0.001
<i>KLRB1</i>	12	9.59	UKB	20.3	0.11	1×10^{-7}	0.07	8×10^{-4}
<i>NCR1</i>	19	54.91	UKB	33.7	0.11	2×10^{-7}	0.07	0.001
<i>CD48</i>	1	160.68	DeC	12.2	0.11	2×10^{-7}	0.06	0.01
<i>LTA</i>	6	31.57	UKB	11.5	0.10	5×10^{-7}	0.04	0.07
<i>CXCL11</i>	4	76.03	UKB	17.3	0.10	6×10^{-7}	0.06	0.01
<i>FCRL6</i>	1	159.80	UKB	14.3	0.10	7×10^{-7}	0.06	0.003
<i>TIGIT</i>	3	114.28	DeC	4.9	0.17	7×10^{-17}	0.13	1×10^{-8}
<i>LAG3</i>	12	6.77	DeC	15.4	0.07	4×10^{-4}	0.02	0.3

Chr, chromosome; Mb, megabase; Log OR: log(odds ratio) of type 1 diabetes risk per standard deviation increase in the pGATE score; UKB, UK Biobank Pharma Proteomics Project; DeC, deCODE genetics.

In both the original and sensitivity analyses, 4964 T1D cases (2779 males, 2185 females) and 7474 non-diabetic controls (3173 males, 4301 females) were included.

¹Model in the original analysis: T1D ~ GATE score + age + sex + PCs

²Model in the sensitivity analysis: T1D ~ GATE score + rs3184504*T + age + sex + PCs

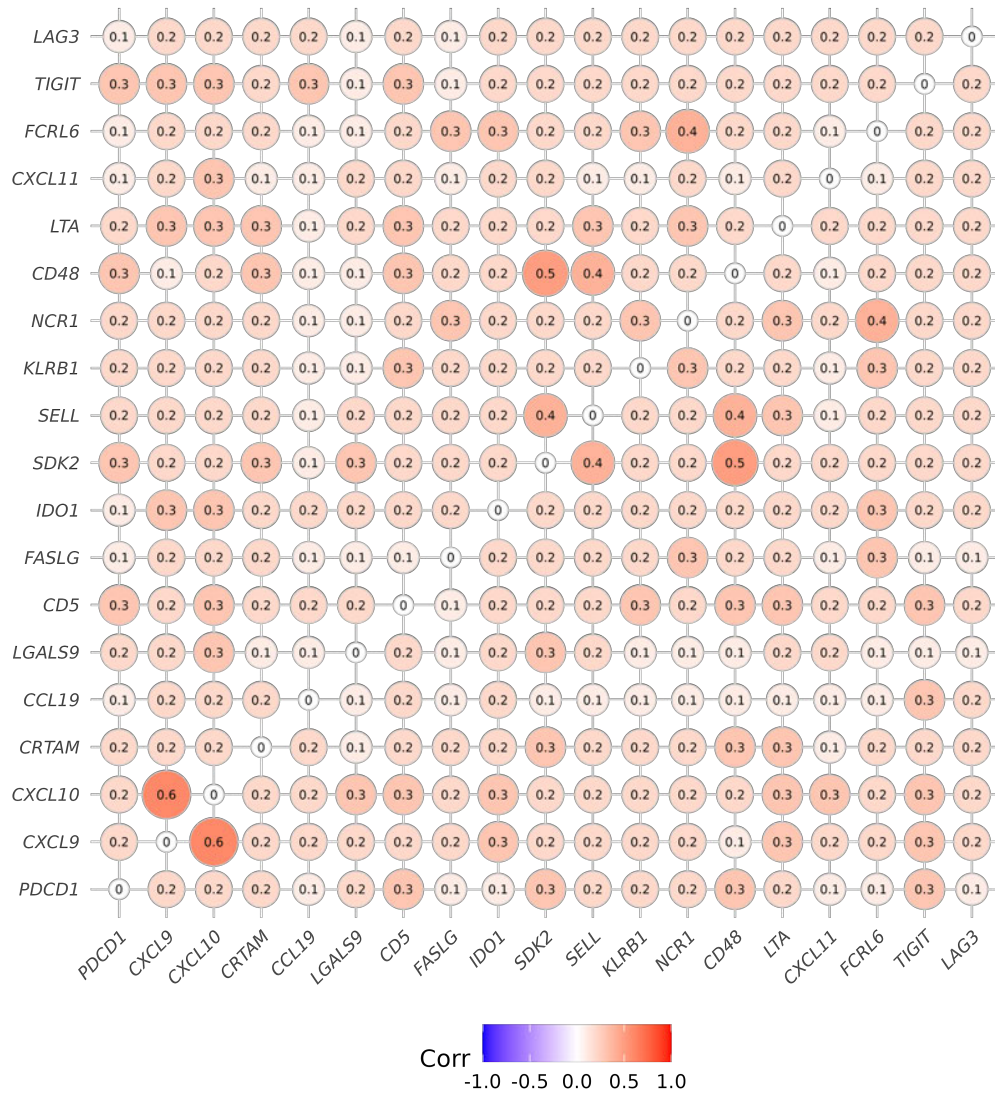


Figure S3.6: Heatmap of raw correlations among GATE scores for 19 putative core proteins with *trans*-pQTLs in or nearby the *SH2B3* gene. Correlations were computed using non-diabetic controls in the Scottish dataset.

3.5 Technical challenges of protein assays and their impact on GATE analysis

The GATE analysis leveraged SNP-protein summary statistics from three large-scale pQTL studies: UKB-PPP (2923 proteins measured by the Olink Explore 3072 panel in 54219 participants), deCODE (4719 proteins measured by SomaLogic v4 panel in 35559 participants), and INTERVAL (2994 proteins measured by SomaLogic panel in 3301 participants). Olink utilises paired antibodies labelled with single oligonucleotides to bind different epitopes on target proteins,

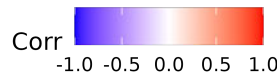


Figure S3.7: Heatmap of partial correlations among GATE scores for 19 putative core proteins with *trans*-pQTLs in or nearby the *SH2B3* gene, conditioning on the genotype of rs3184504*T. Correlations were computed using non-diabetic controls in the Scottish dataset.

whereas SomaLogic employs engineered single-stranded oligonucleotides, aptamers [Pietzner et al., 2021]. Despite these differences, both platforms are affinity-based and rely on conserved binding regions of target proteins for accurate quantification. Protein-altering variants affecting these regions may introduce binding artefacts.

Multiple studies with overlapping measurements in the same population have systematically evaluated cross-platform correlations, genetic associations (i.e., pQTLs), and disease relevance [Pietzner et al., 2021, Katz et al., 2022, Eldjarn

et al., 2023, Rooney et al., 2024, Wang et al., 2025]. The overall cross-platform concordance is moderate and highly variable. Proteins with well-characterised epitope targets, high abundance, and strong genetic effects tend to show higher cross-platform agreement, whereas those with low abundance, or susceptible to off-target binding of the antibody or to epitope effects that affect the accessibility of the epitope, such as post-translational modifications tend to display discordance. Moreover, these two techniques usually have different sensitivities towards protein-altering variants.

In the GATE analysis, for the overlapping proteins that were measured by different platforms, only the strongest associations with at least five effective *trans*-loci were retained (Table 3.1). Of the 27 putative core proteins, 21 were detected based on *trans*-pQTLs from UKB-PPP, and 6 were detected based on *trans*-pQTLs from deCODE. Among the 27 putative core proteins, only 4 had a GATE score with at least 5 effective *trans*-loci computed with pQTLs from INTERVAL, and all of these GATE score associations were surpassed by those from UKB-PPP or deCODE. Nine of the 27 core proteins (*CXCL9*, *CRTAM*, *SELL*, *GCG*, *CD48*, *TCN1*, *CXCL11*, *TIGIT*, *LAG3*) had GATE scores with at least five effective *trans*-loci computed with pQTLs from both UKB-PPP and deCODE. All of them showed consistent direction of effect in the associations with T1D, although the effect sizes and *p*-values varied (Table S3.6).

Table S3.6: Comparison of GATE scores computed with pQTLs from UKB-PPP and deCODE for associations with T1D in the Scottish and UK Biobank datasets.

Gene symbol	UKB, Scottish dataset ¹			DeC, Scottish dataset ²			UKB, UK Biobank ³			DeC, UK Biobank ⁴		
	Effective number of pQTLs	Log OR	<i>p</i> -value	Effective number of pQTLs	Log OR	<i>p</i> -value	Effective number of pQTLs	Log OR	<i>p</i> -value	Effective number of pQTLs	Log OR	<i>p</i> -value
<i>PDCD1</i>	30.9	0.21	6×10^{-24}	.	.	.	20.8	0.21	6×10^{-17}	.	.	.
<i>CXCL9</i>	11.2	0.19	1×10^{-21}	5.8	0.08	5×10^{-5}	7.0	0.09	1×10^{-4}	5.5	0.08	6×10^{-4}
<i>CXCL10</i>	14.4	0.16	2×10^{-15}	4.8	0.11	1×10^{-7}	6.8	0.12	2×10^{-6}	4.8	0.12	4×10^{-7}
<i>CRTAM</i>	31.3	0.09	1×10^{-5}	10.7	0.16	4×10^{-15}	20.9	0.04	0.07	9.9	0.12	3×10^{-7}
<i>CCL19</i>	2.1	-0.36	5×10^{-55}	10.2	0.16	4×10^{-15}	8.2	0.20	1×10^{-15}	9.0	0.19	8×10^{-16}
<i>LGALS9</i>	13.1	0.16	2×10^{-14}	.	.	.	8.0	0.09	2×10^{-4}	.	.	.
<i>CD5</i>	16.1	0.14	1×10^{-12}	4.7	0.03	0.2	10.7	0.12	2×10^{-6}	3.9	0.02	0.5
<i>PRSS27</i>	6.0	0.13	2×10^{-10}	2.8	0.13	6×10^{-10}	6.7	0.03	0.2	2.7	0.00	0.8
<i>FASLG</i>	15.8	0.12	8×10^{-10}	1.2	0.00	1	17.8	0.08	0.001	1.2	-0.04	0.1
<i>IDO1</i>	11.6	0.12	1×10^{-9}	1.4	0.04	0.07	6.6	0.10	9×10^{-5}	1.3	0.02	0.5
<i>AGR2</i>	8.0	-0.12	2×10^{-9}	.	.	.	7.0	-0.02	0.5	.	.	.
<i>FGF19</i>	6.2	-0.12	7×10^{-9}	4.9	-0.11	6×10^{-8}	5.7	-0.05	0.03	4.6	-0.02	0.4
<i>SDK2</i>	5.8	0.12	7×10^{-9}	.	.	.	5.7	0.09	2×10^{-4}	.	.	.
<i>SELL</i>	20.1	0.11	4×10^{-8}	23.3	0.03	0.1	11.4	0.07	0.008	20.5	0.03	0.2
<i>GCG</i>	11.3	0.05	0.007	8.0	0.11	1×10^{-7}	9.9	0.01	0.6	7.9	-0.02	0.5
<i>KLRB1</i>	20.3	0.11	1×10^{-7}	.	.	.	18.3	0.06	0.01	.	.	.
<i>NCR1</i>	33.7	0.11	2×10^{-7}	.	.	.	25.7	0.06	0.02	.	.	.
<i>CD48</i>	19.1	0.09	3×10^{-6}	12.2	0.11	2×10^{-7}	13.2	0.12	9×10^{-7}	10.0	0.08	4×10^{-4}
<i>CKK</i>	.	.	.	6.0	0.10	3×10^{-7}	.	.	.	6.5	-0.04	0.07
<i>CFC1</i>	12.8	-0.10	5×10^{-7}	4.6	0.00	0.9	15.3	-0.12	4×10^{-6}	8.3	0.04	0.1
<i>LTA</i>	11.5	0.10	5×10^{-7}	.	.	.	9.6	0.07	0.006	.	.	.
<i>TCN1</i>	6.7	0.10	5×10^{-7}	7.2	0.05	0.02	7.7	0.02	0.4	5.4	-0.01	0.8
<i>CXCL11</i>	17.3	0.10	6×10^{-7}	14.1	0.03	0.1	11.2	0.02	0.5	16.3	-0.03	0.2
<i>REG1B</i>	7.8	-0.10	7×10^{-7}	4.6	-0.10	4×10^{-7}	6.5	-0.02	0.5	4.3	0.00	0.8
<i>FCRL6</i>	14.3	0.10	7×10^{-7}	.	.	.	11.9	0.08	9×10^{-4}	.	.	.
<i>TIGIT</i>	4.7	0.09	3×10^{-6}	4.9	0.17	7×10^{-17}	5.5	0.09	5×10^{-4}	5.9	0.23	2×10^{-24}
<i>LAG3</i>	23.1	0.05	0.03	15.4	0.07	4×10^{-4}	18.6	0.13	4×10^{-8}	18.1	0.06	0.01

Log OR: log(odds ratio) of type 1 diabetes risk per standard deviation increase in the pGATE score; UKB, UK Biobank Pharma Proteomics Project; DeC, deCODE genetics.

¹GATE scores were computed with *trans*-pQTLs from UKB-PPP and then tested for associations in the Scottish dataset (4964 T1D cases and 7474 non-diabetic controls).

²GATE scores were computed with *trans*-pQTLs from deCODE and then tested for associations in the Scottish dataset (4964 T1D cases and 7474 non-diabetic controls).

³GATE scores were computed with *trans*-pQTLs from UKB-PPP and then tested for associations in the remaining participants of UK Biobank (1638 T1D cases and 365K non-diabetic controls).

⁴GATE scores were computed with *trans*-pQTLs from deCODE and then tested for associations in the all participants of UK Biobank (1864 T1D cases and 408K non-diabetic controls).

As discussed earlier, PTMs represent an important technical challenge for protein assays and the master regulator *FUT2* is an example of their potential impact on the GATE analysis. PTMs are chemical modifications to proteins after synthesis that affect activity, localisation, stability, and interactions of proteins [Zhong et al., 2023]. Over 650 types of PTMs have been identified, including phosphorylation, ubiquitination, glycosylation, methylation, short-chain and long-chain fatty acid modifications, SUMOylation, carbamylation, and redox modifications. In theory, *trans*-pQTLs might reflect genetic predispositions to certain PTMs of target proteins rather than true abundance changes. This in turn could introduce confounding to the associations from the GATE analysis, if for instance, specific PTMs have a strong effect on the disease, rather than the protein upon which these PTMs are found. An ideal way to study if the GATE associations are influenced by PTMs is to leverage both SNP-transcript and SNP-protein summary statistics to compute GATE scores for transcripts and corresponding proteins as validations for each other. However, due to the limited coverage of SNPs in eQTLGen Phase 1, GATE scores for the transcripts of the 27 putative core proteins did not have adequate effective *trans*-loci. Another complication is that the SNP-transcript associations are measured with respect to transcript levels in blood, while the circulating proteins could be produced in other tissues. *Trans*-effects on transcript levels can be context-specific and thus can differ between expression in blood and expression in tissue. In this case, the blood-based pQTL-GATE score would contain the context-specific *trans*-effects that led to the production of the protein in the tissue, while the blood-based eQTL-GATE score (even with sufficient *trans*-loci) would not. Instead, a few other strategies were employed to mitigate the impact of PTMs on the GATE analysis. Firstly, GATE scores were computed via aggregation of multiple *trans*-loci and the effective number of *trans*-loci for each GATE score was calculated as a weighted index of the variances of the locus-specific scores at each *trans*-locus. For a GATE score with K *trans*-loci, this index takes the value of 1 if one of the *trans*-loci has much larger variance than the others (i.e., dominates the GATE score), and the value of K if the variances across the *trans*-loci are equal (i.e., equal contributions to the GATE score). For the declaration of putative core proteins, an index of at least five effective *trans*-loci was required. Thus, the GATE score associations were not driven by any single *trans*-pQTL. In addition, a series of validation criteria, including the MR analysis based on *trans*-pQTLs, as well as six further criteria independent of *trans*-pQTLs, were applied to prioritise the most biologically plausible effector proteins and potential drug targets for T1D.

3.6 Conclusion

In this chapter, I leveraged the GATE methods and identified 27 putative core proteins for T1D. Incorporating other lines of evidence, immune proteins, especially immune checkpoints (*PDCD1*, *CD5*, *TIGIT* and *LAG3*) and innate immune proteins (*NCR1* and *KLRB1*), stood out as the top hits with the most convincing evidence. Given that PD-1 and LAG-3 agonists are in clinical trials for other autoimmune conditions, they should be trialled for preventing T1D. For the eight pancreatic and gut proteins, despite meeting the selection criteria for putative core proteins, supportive evidence was generally limited.

Of these 27 putative core proteins, only four (*CD5*, *FASLG*, *GCG*, *KLRB1*) are proximate to previously reported GWAS hits for T1D, and one (*CXCL9*) was supported by a *cis*-score association with T1D. By aggregating *trans*-effects across the genome, the GATE approach is capable of detecting disease-relevant genes that might be missed or obscured in conventional GWAS analysis.

4

Plasma proteins associated with T1D
and their potential as drug targets

4.1 Introduction

In Chapter 3, I identified 27 putative core proteins for T1D by first carrying out a GATE analysis. In this approach, an arbitrary p -value less than 1×10^{-6} was applied to minimise false positive association. However, all such p -value thresholds are arbitrary and will, of course, exclude some true associations. In this chapter, I adopted a different strategy to identify further likely causal proteins that did not meet the $p < 1 \times 10^{-6}$ criterion by starting with testing for association between disease and circulating protein level first and then evaluating these associations for causality against several criteria, including a confirmatory GATE score – disease association. For this, I used the direct measurements of plasma protein levels in the UKB-PPP dataset, in which a subset of 257 people had type 1 diabetes (232 are of European ancestry).

Specifically, the objectives of this chapter were to

- test the association between plasma protein levels and T1D in the UKB-PPP dataset;
- validate protein associations with genetic proxies, i.e., GATE scores;
- assess for causality and potential as drug targets with pre-specified criteria, i.e., a dose-response effect supported by MR analysis, reported monogenic causes of autoimmune diabetes, *cis*-score association, known T1D GWAS hits, experimental evidence in animal models, and drug effects in humans.

In this chapter, I wrote the initial analysis plan which I then revised together with my supervisors and our collaborator (Prof Helen Colhoun). I wrote code and conducted the cross-sectional analyses. The genetic validation of proteins followed the same analytical pipeline used in Chapter 3, which was originally developed by my supervisors and postdoctoral researchers within the group. I refined and adapted this pipeline, and wrote additional code to perform the Mendelian randomisation analysis, cross-reference results with monogenic causes for autoimmune diabetes and GWAS hits for T1D, annotate biological functions, and collate supporting evidence from experimental animal models and human studies. I wrote the original draft of the chapter, which was subsequently reviewed by my supervisors and Prof Colhoun, and I made all revisions based on their feedback.

4.2 Methods

4.2.1 Study population

The UKB-PPP dataset is a subset of 54,219 participants from the UK Biobank cohort with plasma protein profiling [Sun et al., 2023]. I used the same definition of T1D cases and controls in the UKB-PPP dataset as presented in Chapter 2 and identified 257 T1D cases and 47170 controls without diabetes. Plasma protein quantification was performed using the Olink Explore 3072 platform, targeting 2923 unique proteins across eight panels – inflammation (INF and INF II), oncology (ONC and ONC II), cardiometabolic (CAR and CAR II) and neurological (NEU and NEU II) [Sun et al., 2023]. Protein levels were quantified using Proximity Extension Assay (PEA) technology, with which matched antibody pairs labelled with unique complementary oligonucleotides (proximity probes) bind to their respective target proteins, followed by DNA amplification and next-generation sequencing. Data were quality controlled using internal and external controls, and Normalised Protein eXpression (NPX) values were generated through standard normalisation and batch-adjustment procedures. No significant batch or plate effects were detected.

4.2.2 Statistical analysis

4.2.2.1 Protein data cleaning and imputation

To address missingness in the protein data, I firstly excluded proteins with more than 20% missingness across individuals and individuals with more than 10% missingness across proteins, which resulted in a dataset of 2919 proteins and 41,415 individuals. Next, missing values were imputed using a principal components analysis-based approach [Brini and van den Heuvel, 2024]. Specifically, I Gaussianised and standardised the non-missing NPX values for each protein and imputed the missing values using the “MIPCA” function in the missMDA R package [Josse and Husson, 2016].

Part of the missingness might be attributable to measurements below the limit of detection (LOD). An alternative strategy for handling this would be to use a statistical model such as Tobit regression which explicitly models data truncation of the dependent variable [Tobin, 1958]. In this case, the protein level is the dependent variable (outcome) and T1D case/control status is the independent

variable (predictor). While this approach explicitly accounts for LOD-related missingness, it was not adopted here due to the high dimensionality of the proteomic data and the multiple underlying causes of missingness.

4.2.2.2 Consideration of reverse causation and confounding

To identify causal protein associations with T1D, ideally one would have proteins measured in samples from persons then followed up for the development of disease, i.e., an incident cohort. Instead, the UKB-PPP is a prevalent cohort. Indeed, most T1D cases in the UKB-PPP dataset have long-standing disease; the median age at clinical diagnosis of T1D is typically under 30 years [Ogle et al., 2025], whilst participants in the UKB-PPP dataset were middle-aged (40 - 69 years) at recruitment. To minimise associations due to reverse causation, i.e., protein levels that are secondarily elevated by diabetes or its complications or its treatment, I adjusted for several covariates in the analysis.

Firstly, the kidney is of great importance in maintaining protein homeostasis by retaining essential proteins in the bloodstream and filtering waste and excess fluid into urine. In people with T1D, high blood glucose may damage renal blood vessels, leading to impaired kidney filtration and subsequent alterations in plasma proteins. Secondly, the liver is the primary organ for nutrient metabolism and protein synthesis, and altered liver function has been reported to be more common in T1D patients [Mertens et al., 2021]. Thirdly, circulating protein levels can be directly or indirectly modulated by medications. Apart from insulin, T1D patients are more frequently prescribed lipid-lowering, antihypertensive, and antithrombotic drugs to prevent cardiovascular diseases than those in the general population [Manrique-Acevedo et al., 2024].

Therefore, I calculated the estimated glucose filtration rate (eGFR) from serum creatinine using the Modification of Diet in Renal Disease (MDRD) equation as an index of kidney function. Liver function markers, including alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), γ -glutamyl transferase (GGT), bilirubin, and albumin, were derived. The prescription of lipid-lowering medications was defined as any use of statins, ezetimibe, fibrates, nicotinic acid, or bile acid sequestrants based on the verbal interview at baseline. Prescription of antihypertensive medications was defined as any usage of angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, thiazide-like diuretics, and calcium channel blockers. Antithrombotic therapy was primarily defined as any usage of aspirin. Detailed information on these drugs is presented in Table S4.1.

4.2.2.3 Identification of plasma proteins associated with T1D

To systematically assess associations between plasma protein levels and T1D in the UKB-PPP dataset, I firstly performed a basic logistic regression model, adjusting for age at assessment, sex, and ethnicity. Before including additional covariates, I assessed their individual association with T1D using the same basic logistic regression model. This step ensured that the differences in these variables between T1D cases and controls were not merely due to age, sex, or ethnicity. Variables such as eGFR, ALP, direct bilirubin, albumin, and the prescription of lipid-lowering, antihypertensive, and antithrombotic medications showed associations with T1D ($p < 0.05$) and were subsequently included as covariates in a fully adjusted logistic regression model. In both models, effect sizes were reported as log odds ratios (log OR) per standard deviation (SD) increase in protein levels. A stringent p -value threshold (10^{-6}) was applied to identify proteins significantly associated with T1D.

In addition, I further explored the association of proteins with HbA1c. The reasoning here was that if a protein shows strong association with HbA1c among those with T1D, it is more likely a consequence of glycaemia or some other aspect of T1D rather than a causal factor. To test this, I used a linear regression model (with the same full set of covariates) to assess the association between protein levels and HbA1c within T1D cases. I also explored the association of proteins with T1D duration among individuals with T1D. Proteins strongly associated with T1D duration are likewise more likely to be secondary to the disease process rather than causal. Note that for any attained age, the duration of diabetes and the age of onset of diabetes are collinear, and their effects cannot be separated. Thus, I employed a linear regression model where protein levels served as the dependent variable, duration was the primary independent variable, and age at onset, sex, ethnicity, eGFR, ALP, direct bilirubin, albumin, and medication use were included as covariates. Since the sample size is relatively small in these two case-only analyses, a less stringent p -value threshold (0.001) was used to identify proteins associated with HbA1c or T1D duration. While proteins associated with HbA1c or T1D duration in these two steps were not excluded from further causal analysis, their potential role as consequences of disease or glycaemia was explicitly noted and considered in interpretation.

4.2.2.4 Genetic assessment of causality of proteins associated with T1D

To evaluate the causality of proteins associated with T1D, I used GATE scores for these proteins as proxies for genetically predicted protein levels and tested their association with T1D in the UK Biobank and Scottish dataset. Details of the GATE analysis are described in Chapter 2. Proteins were considered as potential core proteins if they were associated with T1D in the fully adjusted model at a p -value less than 10^{-6} and were also supported by GATE score associations with a p -value less than 0.001 in either dataset with consistent direction of effect between protein and GATE score associations and a larger effect size for the protein association. Following the approach outlined in Chapter 2, I performed MR analysis for each potential core protein in both datasets, cross-referenced them with monogenic causes for autoimmune diabetes, tested the association of *cis*-scores with T1D in both datasets, annotated their encoding genes with known T1D GWAS hits, and carried out systematic literature and database reviews to identify published evidence in experimental animal models and drug effects in humans. The flowchart of the study design is shown in Figure 4.1.

4.3 Results

4.3.1 Characteristics of the UKB-PPP dataset

A total of 190 T1D cases and 36861 controls without diabetes were included to identify proteins associated with T1D. Baseline characteristics for both groups are summarised in Table 4.1. The mean age at recruitment was 54 years for T1D cases and 57 years for controls without diabetes. Among participants with T1D, the mean age at diagnosis was 29 years (SD = 12) and the mean duration of T1D was 25 years (SD = 14), 45% were female, and 88% were of European ancestry. In the control group, 56% were female, and 94% were of European ancestry. The mean HbA1c at baseline was 62.55 mmol/mol in T1D cases compared with 34.93 mmol/mol in controls. After adjustment for age, sex, and ethnicity, T1D cases demonstrated slightly higher kidney function, as indicated by a higher mean eGFR (92.37 vs. 87.08 mL/min/1.73m²). Liver function markers were within normal ranges in both groups, despite T1D cases having higher levels of ALP (90.93 vs. 83.1 U/L) and direct bilirubin (2.04 vs. 1.83 µmol/L) but lower levels of albumin (44.03 vs. 45.2 g/L). Compared with controls, participants with T1D had higher prescription rates for lipid-lowering (72% vs. 15%), antihypertensive (54% vs. 16%) or antithrombotic medications (48% vs. 12%).

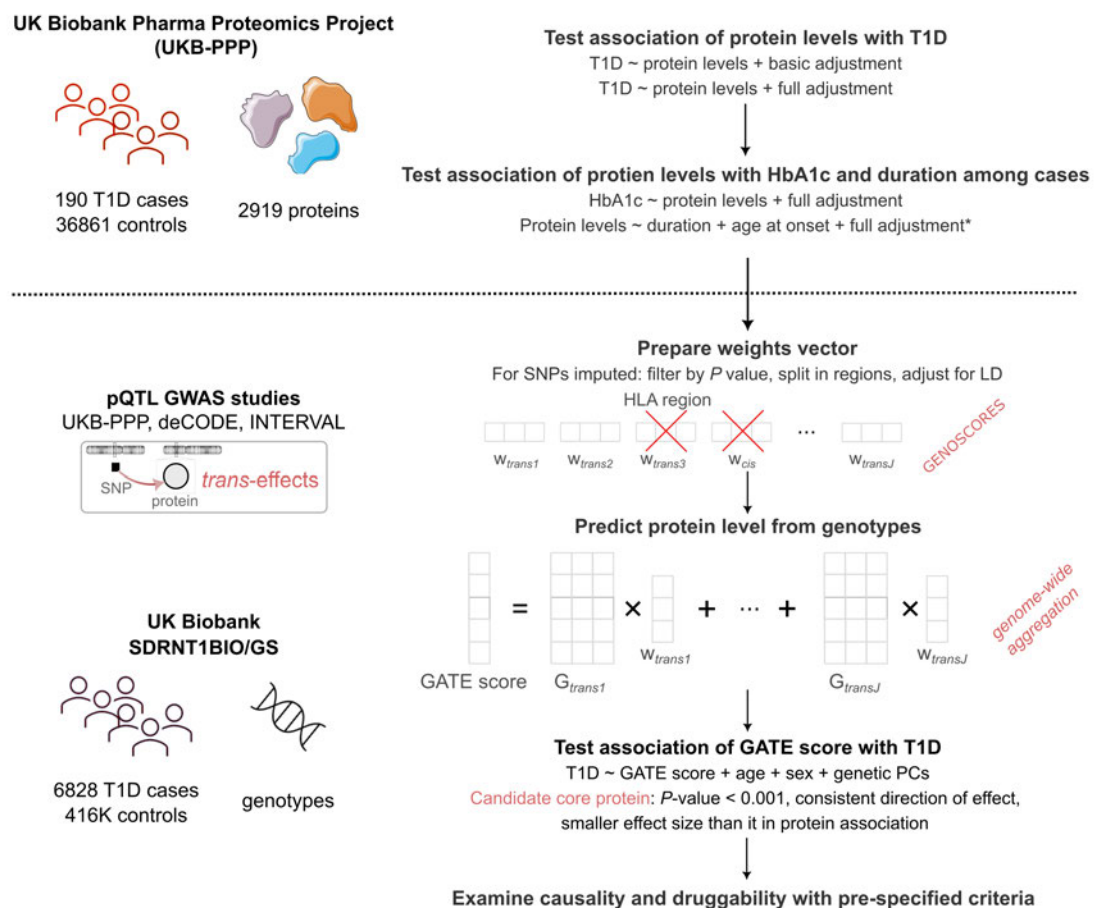


Figure 4.1: Flowchart of the study design. The upper panel outlines the association analyses between protein levels and T1D case–control status, HbA1c, and T1D duration in the UKB-PPP dataset. Models with basic adjustment included age at assessment, sex, and ethnicity as covariates, and fully adjusted models additionally accounted for kidney function, liver function, and medication use. For the analysis involving T1D duration, protein levels were specified as the dependent variable, duration as the primary independent variable, and age at onset, sex, ethnicity, kidney function, liver function, and medication use as covariates. The lower panel shows the genetic validation of protein associations using GATE scores in UK Biobank and the Scottish dataset, followed by evaluation of causal relationships and druggability.

Table 4.1: Characteristics of T1D cases and controls without diabetes in UKB-PPP dataset at study recruitment.

Characteristic	T1D cases		Controls without diabetes		<i>p</i> -value
	Mean/N	SD/%	Mean/N	SD/%	
Age (Study recruitment)	54	8	57	8	< 0.001
Age (T1D onset)	29	12			
T1D duration	25	14			
Sex (Female)	85	45%	20529	56%	0.003
Ethnicity (European)	168	88%	34633	94%	0.002
HbA1c (mmol/mol)	62.55	15.11	34.93	3.82	< 0.001
eGFR (mL/min/1.73 m ²)	92.37	25.86	87.08	17.17	0.001
ALT (U/L)	25.11	14.73	22.6	13.22	0.095
AST (U/L)	27.18	10.67	26	10.33	0.267
ALP (U/L)	90.93	29.42	83.1	25.94	< 0.001
GGT (U/L)	33.52	28.29	35.49	37.24	0.211
Direct bilirubin (μmol/L)	2.04	0.91	1.83	0.83	0.011
Total bilirubin (μmol/L)	9.18	4.52	9.15	4.42	0.496
Albumin (g/L)	44.03	2.74	45.2	2.62	< 0.001
Lipids-lowering therapy	137	72%	5645	15%	< 0.001
Antihypertensive therapy	103	54%	6002	16%	< 0.001
Antithrombotic therapy	92	48%	4525	12%	< 0.001

Continuous variables are presented with mean and standard deviation (SD), and categorical variables are presented with number (N) and percentage (%). The *p*-value for age and HbA1c was derived from a *t*-test; for sex and ethnicity, from a Chi-square test; and for all other variables, from a logistic regression model on T1D, adjusted for age, sex, and ethnicity.

4.3.2 Proteins associated with T1D

Using the basic logistic regression model adjusting for age at assessment, sex, and ethnicity, I identified 464 proteins significantly associated with T1D (*p*-value < 10⁻⁶). Due to missing data in the covariates, the fully adjusted model included 144 T1D cases and 27222 controls. With this model, 257 proteins remained to be significantly associated with T1D (*p*-value < 10⁻⁶). Of these, 26 were not identified in the basic model but emerged as significant in the fully adjusted model, indicating the presence of negative confounding in the basic model.

In the fully adjusted model, the strongest association was observed for *PLA2G1B* (phospholipase A2, an enzyme in fatty acid metabolism), with a log OR of -1.32 and a *p*-value of 2×10^{-59} . Similarly, approximately half of the top 30 proteins were involved in nutrient digestion and metabolism, including digestive enzymes (e.g., proteases and lipases) and metabolic hormones and their receptors. These proteins were inversely associated with T1D, and many of them were highly correlated with each other (Figure 4.2), particularly *CPB1*, *CPA1*, *PRSS2*, and *PNLIPRP1*, as well as *PNLIP*, *CELA3A*, *CRTC*, and *CTRL* (Pearson $r \geq 0.6$). *CD276* demonstrated the second strongest association (log OR = 1.21, *p*-value = 2×10^{-58}). It belongs to the B7 family of immune checkpoint ligands that are expressed on activated antigen-presenting cells (APCs). Notably, four additional B7 family proteins – *NCR3LG1* (B7-H6), *PDCD1LG2* (B7-DC), *CD80* (B7-1), and

CD274 (B7-H1) – were also strongly associated with T1D. However, these B7 proteins were not correlated with each other (Pearson $r \leq 0.3$).

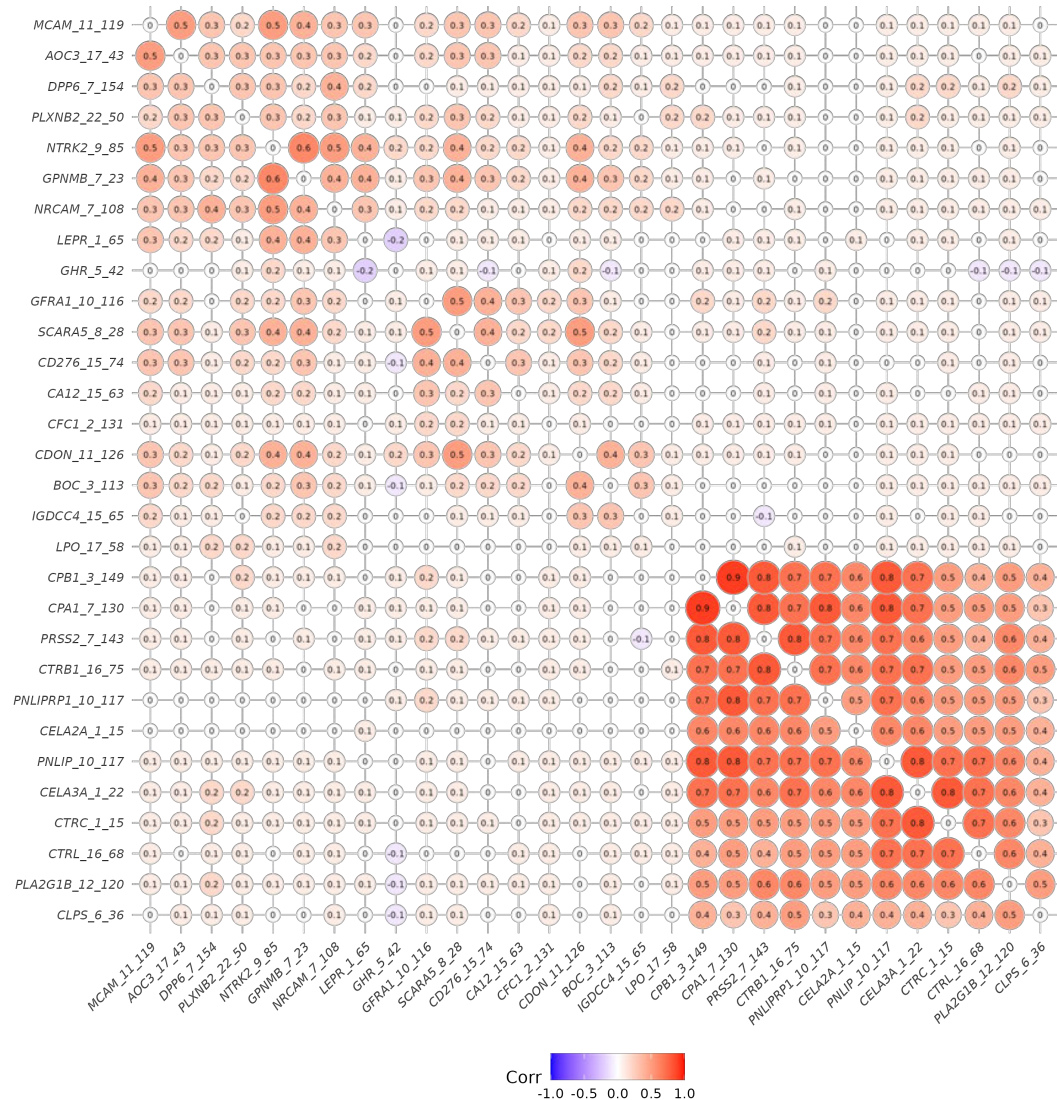


Figure 4.2: Heatmap of correlations among the top 30 T1D-associated proteins, identified in the fully adjusted model and ordered by hierarchical clustering. Proteins are labelled using gene symbol, chromosome, and transcriptional start position (in megabases), separated by underscore (“_”). Correlations were calculated using protein data from non-diabetic controls in the UKB-PPP dataset.

4.3.3 Genetic validation of protein associations using GATE scores

For proteins associated with T1D in the fully adjusted model, 255 of them had available GATE scores that could serve as genetic proxies of protein levels. In

UK Biobank, GATE scores for 14 proteins were associated with T1D (p -value < 0.001) with a consistent direction of effect but a smaller effect size compared to their protein associations. Similarly, in the Scottish dataset, GATE scores for 13 additional proteins showed significant associations with T1D (p -value < 0.001), again with consistent direction and larger magnitude in their protein associations. Combining together, 24 unique proteins were identified as potential core proteins, supported by both protein associations and GATE score associations (Table 4.2). Among these potential core proteins, *CFC1* and *KLRB1* have already been identified as putative core proteins previously in Chapter 3, while the others represent novel findings. For the metabolic proteins, only GATE scores for *PRSS2* and *PNLIPRP1* showed association with T1D, which indicated that the protein associations of metabolic proteins are more likely to be a feature of impaired exocrine functions in long-standing T1D rather than the causes of T1D. For the B7 family proteins, even though only *CD274* was supported by GATE score association, the receptor for *CD274* and *PDCD1LG2* – *PDCD1* – was the top putative core protein in Chapter 3, and the GATE score of the receptor for *CD80* – *CD28* – also revealed association with T1D.

4.3.4 Supportive evidence for causality and druggability

- For four (*CD28*, *ADA2*, *ADGRE2*, *KLRB1*) of the 24 proteins, marginalising over pleiotropic effects in the *trans*-pQTL MR analysis revealed a dose-response relationship between protein levels and T1D risk (p -value < 0.01) in either dataset (Table S4.2, Figure S4.1, and Figure S4.2). Besides, a suggestive dose-response effect (p -value < 0.05) was discovered for seven proteins (*CFC1*, *NOTCH1*, *TNFRSF4*, *GCNT1*, *PCDH17*, *SIGLEC6*, *SEMA7A*) in UK Biobank and two proteins (*CFC1*, *PNLIPRP1*) in the Scottish dataset.
- A loss-of-function mutation in *CD274* has been reported as a monogenic cause of autoimmune diabetes [Johnson et al., 2024].
- None of these 24 proteins were supported by *cis*-score association at a Bonferroni-corrected p -value threshold (0.002, 0.05/22, since *CFC1* and *DDR1* did not have *cis*-scores in either dataset). However, the *cis*-score for *ITGB2* showed an association with T1D in the Scottish dataset just below the threshold (Table S4.3, p -value = 0.003).

Table 4.2: Potential core proteins supported by both protein associations and GATE score associations.

Gene symbol	Transcriptional site		Protein association ¹		GATE score association (UK Biobank) ²				GATE score association (Scottish) ²			
	Chr	Start position (Mb)	Log OR	p-value	pQTL study	Effective number of pQTLs	Log OR	p-value	pQTL study	Effective number of pQTLs	Log OR	p-value
<i>CFC1³</i>	2	130.59	-1.08	1×10^{-43}	UKB	15.3	-0.12	4×10^{-6}	UKB	12.8	-0.10	5×10^{-7}
<i>NOTCH1</i>	9	136.49	0.60	1×10^{-12}	UKB	20.7	0.11	9×10^{-6}	UKB	20.2	0.04	0.05
<i>CD28</i>	2	203.71	0.41	2×10^{-7}	UKB	17.2	0.11	1×10^{-5}	UKB	17.5	0.06	0.003
<i>ADA2</i>	22	17.18	0.92	2×10^{-27}	UKB	13.1	0.11	2×10^{-5}	UKB	21.8	0.08	1×10^{-4}
<i>ADGRE2</i>	19	14.73	0.81	3×10^{-20}	UKB	13.1	0.11	2×10^{-5}	DeC	6.1	0.05	0.02
<i>TNFRSF4</i>	1	1.21	0.44	5×10^{-7}	UKB	7.4	0.10	6×10^{-5}	DeC	5.6	0.06	0.001
<i>GCNT1</i>	9	76.42	0.45	3×10^{-7}	UKB	9.3	0.10	7×10^{-5}	UKB	15.1	0.05	0.02
<i>PCDH17</i>	13	57.63	0.79	8×10^{-20}	UKB	4.0	0.10	7×10^{-5}	UKB	4.0	-0.01	0.5
<i>SIGLEC6</i>	19	51.52	0.62	6×10^{-19}	UKB	25.1	0.09	1×10^{-4}	UKB	25.7	0.05	0.01
<i>DLL1</i>	6	170.28	0.45	2×10^{-7}	UKB	9.0	0.09	2×10^{-4}	UKB	11.5	0.05	0.02
<i>SEMA7A</i>	15	74.41	0.80	1×10^{-22}	UKB	16.3	0.09	3×10^{-4}	DeC	9.3	0.08	6×10^{-5}
<i>ROBO1</i>	3	78.60	0.70	2×10^{-16}	DeC	13.5	0.08	4×10^{-4}	DeC	13.4	0.03	0.2
<i>ITGB2</i>	21	44.89	0.77	4×10^{-23}	UKB	15.8	0.08	8×10^{-4}	UKB	20.6	0.06	0.001
<i>OXT</i>	20	3.07	-0.50	1×10^{-8}	UKB	22.3	-0.08	9×10^{-4}	UKB	23.1	-0.04	0.03
<i>DDR1</i>	6	30.88	0.50	5×10^{-10}	UKB	14.0	0.08	0.002	UKB	16.8	0.09	3×10^{-5}
<i>KLRB1³</i>	12	9.59	0.70	6×10^{-15}	UKB	18.3	0.06	0.01	UKB	20.3	0.11	1×10^{-7}
<i>CHAD</i>	17	50.46	-0.50	4×10^{-8}	DeC	13.4	-0.06	0.01	UKB	15.2	-0.07	9×10^{-4}
<i>ALCAM</i>	3	105.37	0.44	7×10^{-7}	DeC	8.4	0.05	0.03	DeC	8.2	0.08	9×10^{-5}
<i>AMIGO2</i>	12	47.08	0.56	2×10^{-12}	UKB	20.7	0.05	0.03	UKB	20.0	0.07	5×10^{-4}
<i>PRSS2</i>	7	142.76	-1.11	2×10^{-53}	INT	10.6	-0.05	0.04	DeC	20.2	-0.07	5×10^{-4}
<i>CD274</i>	9	5.45	0.41	4×10^{-7}	UKB	8.3	0.05	0.04	DeC	2.9	0.08	6×10^{-5}
<i>MUC2</i>	11	1.07	0.73	6×10^{-17}	UKB	3.7	0.05	0.07	UKB	3.0	0.07	7×10^{-4}
<i>PNLIPRP1</i>	10	116.59	-1.02	6×10^{-36}	DeC	8.0	-0.03	0.2	UKB	15.2	-0.08	2×10^{-4}
<i>PLAT</i>	8	42.17	-0.44	6×10^{-7}	UKB	8.6	0.02	0.4	DeC	8.1	-0.07	6×10^{-4}

Chr, chromosome; Mb, megabase; Log OR, log (odds ratio) of T1D per standard deviation increase in the protein level or GATE score. UKB, UK Biobank Pharma Proteomics Project; DeC, deCODE genetics; INT, INTERVAL.

¹Protein associations were tested in the UKB-PPP dataset among 190 T1D cases and 36861 controls.

²GATE score associations were tested in the UK Biobank with 1638 cases and 365K controls for the GATE scores computed based on UKB-PPP pQTL GWAS summary statistics and 1864 cases and 408K controls for those computed based on other pQTL GWAS studies, and the Scottish dataset with 4964 cases and 7474 controls.

³Putative core proteins identified in Chapter 3.

- *CD28* is located within 200 kb of the T1D GWAS hit *CTLA4*, and *KLRB1* is located within 200 kb of the T1D GWAS hit *CD69*.
- Of these 24 proteins, 14 have not been investigated in animal models of T1D (e.g., NOD mice). Among the remaining ten proteins, six are supported by experimental evidence in animal models. For instance, *Notch1* antisense transgenic mice displayed decreased insulin secretion and reduced islet and β cell mass [Eom et al., 2021]. Short-term (4 weeks) administration of anti-CD28 antibody combined with rapamycin effectively delayed the onset of diabetes in NOD mice, while prolonged treatment (10 weeks) accelerated diabetes development [Besançon et al., 2018]. For *TNFRSF4*, knockout of its ligand *OX40L* prevented diabetes development in NOD mice by 100% and significantly reduced the incidence of diabetes in a T cell receptor transgenic mouse model [Martin-Orozco et al., 2003]. Genetic deficiency of *ITGB2* (*Itgb2*^{-/-}) protected against the development of autoimmune diabetes in NOD mice [Glawe et al., 2009]. *ALCAM* binds to *CD6*, and a combination of anti-CD6 antibody and oral insulin reversed

recent-onset diabetes in NOD mice, although this effect did not last after drug cessation [von Herrath, 2013]. Bioengineered platelets and extracellular vesicles overexpressing PD-L1 (encoded by *CD274*) and other immune checkpoint molecules were shown to ameliorate T1D in NOD mice [Ma et al., 2025, Yang et al., 2024].

- Two proteins are supported by drug effects in humans. Inhibitors targeting PD-L1 (encoded by *CD274*) have been shown to induce T1D in cancer patients [Cho and Jung, 2023]. Abatacept, a CTLA4-Ig fusion protein targeting the CD28 pathway, has been shown to slow reduction in β cell function by continued administration for over two years in recent-onset T1D patients [Orban et al., 2011], whereas no significant effect was identified in delaying the progression to stage 2 or stage 3 T1D among individuals at stage 1 [Russell et al., 2023].

Table 4.3 summarises for each potential core protein the results of the six lines of evidence that support causality and druggability of the protein for T1D.

4.3.5 Association of potential core proteins with HbA1c and T1D duration

Among the 24 potential core proteins, significant associations with HbA1c in individuals with T1D were observed for only two proteins: *ITGB2* (integrin β -2) and *CHAD* (chondroadherin). Higher plasma levels of *ITGB2* were positively associated with HbA1c, whereas higher levels of *CHAD* were negatively associated with HbA1c (Table S4.4). In addition, only one protein, *PNLIPRP1* (inactive pancreatic lipase-related protein 1), demonstrated an inverse association with T1D duration.

4.4 Discussion

In this chapter, I found 257 proteins cross-sectionally associated with T1D with a fully adjusted model. Using the GATE scores as genetic proxies for proteins, 24 were supported by associations at not only the measured level but also the genetically predicted level. Two of these 24 proteins have been identified as core proteins in Chapter 3, while the others represent novel findings, complementing the results of Chapter 3. Among the novel hits, *CD28*, *ADA2*, and *ADGRE2* are

4. Plasma proteins associated with T1D and their potential as drug targets 100

Table 4.3: Summary of supportive evidence for the potential core proteins for T1D.

Gene symbol	Mendelian randomisation ¹	Monogenic cause ²	Cis-score association ³	GWAS hit ⁴	Experimental validation ⁵	Drug effect ⁶
<i>CFC1</i> ⁷	.	.	.	.	.	.
<i>NOTCH1</i>	.	.	.	.	+	.
<i>CD28</i>	+	.	.	+	+	+
<i>ADA2</i>	+	.	.	.	.	.
<i>ADGRE2</i>	+	.	.	.	.	.
<i>TNFRSF4</i>	.	.	.	.	+	.
<i>GCNT1</i>	.	.	.	.	.	.
<i>PCDH17</i>	.	.	.	.	.	.
<i>SIGLEC6</i>	.	.	.	.	.	.
<i>DLL1</i>	.	.	.	.	.	.
<i>SEMA7A</i>	.	.	.	.	.	.
<i>ROBO1</i>	.	.	.	.	.	.
<i>ITGB2</i>	.	.	.	.	+	.
<i>OXT</i>	.	.	.	.	.	.
<i>DDR1</i>	.	.	.	.	.	.
<i>KLRB1</i> ⁷	+	.	.	.	.	.
<i>CHAD</i>	.	.	.	.	.	.
<i>ALCAM</i>	.	.	.	.	+	.
<i>AMIGO2</i>	.	.	.	.	.	.
<i>PRSS2</i>	.	.	.	.	.	.
<i>CD274</i>	.	+	.	.	+	+
<i>MUC2</i>	.	.	.	.	.	.
<i>PNLIPRP1</i>	.	.	.	.	.	.
<i>PLAT</i>	.	.	.	.	.	.

‘+’ indicates that the criterion is met; ‘.’ indicates that it is not. Criteria are defined as follows:

¹Mendelian randomisation based on *trans*-pQTLs and marginalizing over the distribution of pleiotropic effects supports a causal effect for the protein on type 1 diabetes at *p*-value < 0.01.

²There is a rare variant in the gene encoding the protein that causes monogenic autoimmune diabetes.

³A *cis*-score for the protein is associated with type 1 diabetes at *p*-value < 0.05/22.

⁴There is a known GWAS hit within 200 kb of the gene encoding the protein.

⁵There is evidence of causality through perturbation of gene encoding the protein or its product in an animal model of type 1 diabetes.

⁶An existing drug (either investigational or approved) that targets the protein or its ligand/receptor cause or ameliorate type 1 diabetes.

⁷Putative core proteins identified in Chapter 3.

supported with a dose-response effect by MR analysis; *CD274* has been reported as a monogenic cause of autoimmune diabetes; *CD28* is located proximate to known GWAS hits; *NOTCH1*, *CD28*, *TNFRSF4*, *ITGB2*, *ALCAM*, and *CD274* are supported by experimental evidence in animal models; and *CD274* and *CD28* are supported by drug effects in humans. Gene name and protein name synonyms and biological functions of the novel hits are listed in Table S4.5. Highlighted pathways are discussed below.

4.4.1 Immune checkpoint receptor and ligand

Of the novel hits, *CD28* and *TNFRSF4* (OX40) are two co-stimulatory immune checkpoint receptors on T cells, and *CD274* is an immune checkpoint ligand on APCs. *CD28* binds with its ligands *CD80* (B7-1) and *CD86* (B7-2), and *OX40* binds with its ligand *OX40L*. Both interactions induce T cell activation, differentiation, and cytokine production. These pathways appear to act at different stages

of diabetes development in NOD mice. CD28 blockade prevented diabetes in a window of up to 4 weeks, while prolonged administration accelerated diabetes development [Besançon et al., 2018]. On the contrary, only blocking OX40L at 12 weeks after birth or continuous treatment from 12 weeks onwards reduced the incidence of diabetes in NOD mice [Pakala et al., 2004], suggesting it functions as a later-stage checkpoint. It is worth noting that the effect of these pathways on Tregs can shift disease trajectory. Deficiency in CD28 or its ligands B7-1/B7-2 exacerbated spontaneous diabetes in NOD mice due to the elimination of CD4⁺CD25⁺CTLA4⁺ Tregs [Salomon et al., 2000]. Conversely, treatment with OX40 agnostic antibody (OX86), especially in combination with antigen-specific therapy (i.e., insulin B9:23 peptide), expanded CD4⁺CTLA4⁺Foxp3⁺ Tregs and reduced T1D incidence in NOD mice [Bresson et al., 2011].

Because of the higher affinity to the shared ligands and the effect of physically removing these ligands from APCs, CTLA-4 acts as a competitive inhibitor of CD28 [Edner et al., 2020]. Clinically, abatacept (CTLA4-Ig fusion protein) has shown efficacy in preserving β cell function in recent-onset T1D [Orban et al., 2011], but it failed to delay progression in individuals at stage 1 [Russell et al., 2023]. A dual antibody targeting TNF and OX40L is being trialled in phase 2.

CD274 encodes PD-L1 (B7-H1), a ligand for the inhibitory receptor PD-1. A loss-of-function mutation in *CD274* has been implicated in monogenic autoimmune diabetes [Johnson et al., 2024]. Therapies using PD-L1 overexpressing platelets or extracellular vesicles ameliorated T1D in NOD mice [Ma et al., 2025, Yang et al., 2024]. However, in this study, both the protein association and the GATE score association suggest that higher circulating levels of PD-L1 are associated with a higher risk of T1D. This may be explained by the role as decoy of soluble PD-L1, which is released by proteolytic cleavage of its membrane-bound form. While membrane-bound PD-L1 suppresses T cell response via binding with PD-1, soluble PD-L1 can act as a decoy, inhibiting membrane-bound PD-L1:PD-1 interaction, and thus promoting T cell responses and precipitating T1D [Dai et al., 2014].

4.4.2 Proteins involved in immune cell adhesion and trafficking

Several proteins play a key role in immune cell adhesion and trafficking. *ITGB2* encodes integrin β -2, a leukocyte-specific adhesion molecule. In NOD mice,

genetic deficiency inhibits the development of diabetes by impairing T cell recruitment and adhesion to islet microvascular endothelial cells [Glawe et al., 2009]. Similarly, *ALCAM* (activated leukocyte cell adhesion molecule, CD166) facilitates leukocyte adhesion and migration through its interaction with CD6. *ADGRE2* (adhesion G protein-coupled receptor E2) is another adhesion molecule predominantly expressed on activated lymphocytes and myeloid cells [Kwakkenbos et al., 2005]. *AMIGO2* (amphoterin-induced protein 2) deficiency impairs peripheral T cell migration to secondary lymphoid organs and ameliorates experimental autoimmune encephalomyelitis [Li et al., 2017]. Additionally, *GCNT1* (core2-GlcNAc-transferase) synthesises O-glycans, which in turn influences the synthesis of selectin ligands and regulates immune cell trafficking [Hobbs and Nolz, 2017].

4.4.3 Neuroimmune proteins

A few proteins primarily involved in neuronal development also modulate immune responses. *SEMA7A* (semaphorin 7A) is an axon guidance factor that helps growing axons find targets and form synapses. It also regulates immune responses, including initiating T-cell-mediated inflammation via $\alpha 1\beta 1$ integrin [Suzuki et al., 2007], promoting the differentiation and functional maintenance of cytokine-induced memory-like natural killer cells, and stimulating monocyte chemotaxis and cytokine production [Huang et al., 2025]. *ROBO1* (roundabout homolog 1), another axon guidance protein, modulates T cell recruitment, enhances T cell chemorepulsion, and delays the onset of diabetes in NOD mice [Glawe et al., 2013]. *PCDH17* (protocadherin-17) regulates presynaptic assembly [Hoshina et al., 2013] and also suppresses tumour cell growth and enhances apoptosis and autophagy of tumour cells [Hu et al., 2013]. *OXT* (oxytocin), a neuropeptide hormone associated with labour and lactation, enhances glucose uptake, stimulates insulin secretion, improves insulin sensitivity, and also exerts anti-inflammatory effects [Elabd and Sabry, 2015].

4.4.4 Other immune proteins

ADA2 encodes adenosine deaminase 2, a degrading enzyme for adenosine. It is of great importance in the maturation and maintenance of the immune system. Deficiency of *ADA2* causes a monogenic disease with a broad spectrum of inflammatory, vascular, immune dysregulatory, and haematologic phenotypes

[Barron et al., 2022]. *SIGLEC6* (sialic acid-binding Ig-like lectin 6, Siglec-6) is an inhibitory receptor expressed in age-associated B cells, a memory B cell subset that expands in chronic infection and autoimmune diseases like systemic lupus erythematosus [Tsubata, 2023]. *PLAT* (tissue-type plasminogen activator), a fibrinolytic enzyme, modulates immune responses by regulating macrophage and DC activation and enhancing T cell activity [Seillier et al., 2022]. *DDR1* (epithelial discoidin domain-containing receptor 1) is a collagen-binding receptor with tyrosine kinase activity that promotes immune exclusion in tumours by collagen fibre alignment [Sun et al., 2021].

4.4.5 Gut and pancreatic proteins

PRSS2 (trypsin-2, serine protease 2) is a pancreatic serine protease, while *PNLIPRP1* (inactive pancreatic lipase-related protein 1) is an inactive lipase that may inhibit dietary triglyceride digestion. *MUC2* (mucin-2), a gel-forming mucin produced and secreted by intestinal goblet cells, is critical for the formation of a mucosal barrier that prevents bacterial penetration [Han et al., 2022].

4.4.6 Notch signalling pathway

Both *NOTCH1* and its ligand *DLL1* were supported by associations at the measured protein level and the genetically predicted level. The Notch signalling pathway plays a critical role in embryonic development and the maintenance of tissue homeostasis. In adult murine models, *NOTCH1* is broadly expressed in the whole pancreas, while the ligands *DLL1* and *DLL4* are specifically expressed in β cells. Loss of *DLL1* and *DLL4* led to improved glucose tolerance, enhanced glucose-stimulated insulin secretion, and hyperglucagonemia. In contrast, overexpression of the intracellular domain of *DLL1* impaired insulin secretion and glucose regulation [Rubey et al., 2020]. Notch 1 antisense transgenic mice are a model for partial inhibition of Notch 1. In these mice, inhibition of Notch 1 decreased insulin secretion without affecting insulin resistance and reduced islet and β cell mass without impacting the α/β cell ratio [Eom et al., 2021]. Additionally, Notch signalling also contributes to the lineage decision of immune cells. In NOD mice, *DLL1* stimulates the development of invariant natural killer T cells (iNKT, a subset of regulatory T cells that express NK cell markers) in the foetal thymus, and decreased expression of *DLL1* leads to iNKT cell deficiency, potentially increasing T1D susceptibility [Wagner et al., 2005].

In summary, this approach identified 22 potential core proteins that did not pass the criteria of putative core proteins in Chapter 3 but have supportive evidence for causality and druggability for T1D. The strongest evidence is for *CD28* and *CD274*, highlighting the importance of immune checkpoint receptors and ligands in T1D pathogenesis. Given that CTLA-4 has higher affinity to the shared ligands and acts as a competitive inhibitor of CD28, more efforts should be directed towards the potential of CTLA-4 in T1D prevention and treatment. Moreover, another immune checkpoint pathway, OX40/OX40L, also deserves attention, as clinical trials of drugs targeting this pathway are ongoing. Despite that *ADA2* and *ADGRE2* are supported by a dose-response effect in the MR analysis, further studies are warranted to validate their roles in the pathogenesis of T1D.

4.5 Supplementary materials

4. Plasma proteins associated with T1D and their potential as drug targets 105

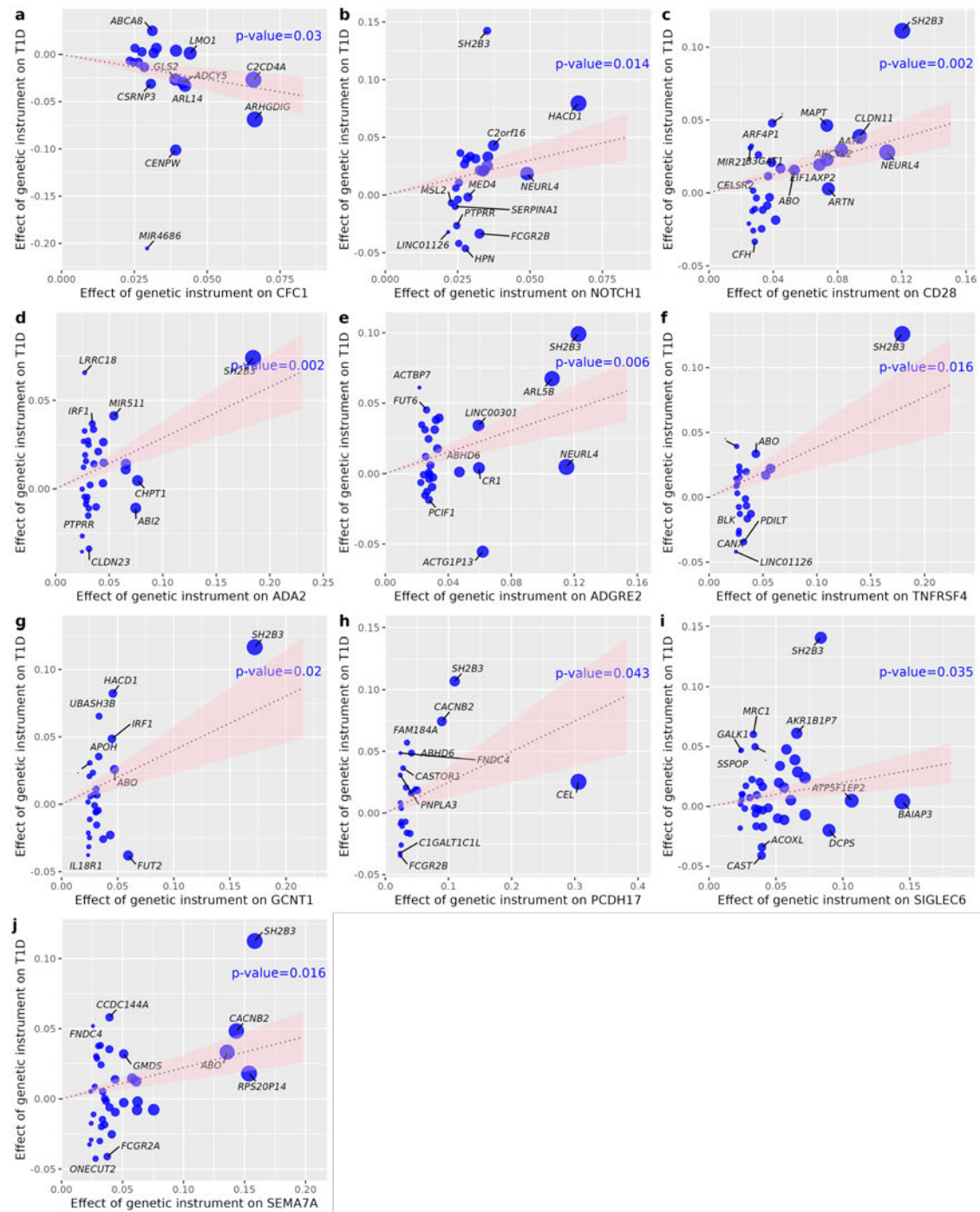


Figure S4.1: Suggestive dose-response effects of protein levels on T1D supported by the Mendelian randomisation analysis in the UK Biobank. All *trans*-pQTLs contributing to the GATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and T1D are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate.

4. Plasma proteins associated with T1D and their potential as drug targets 106

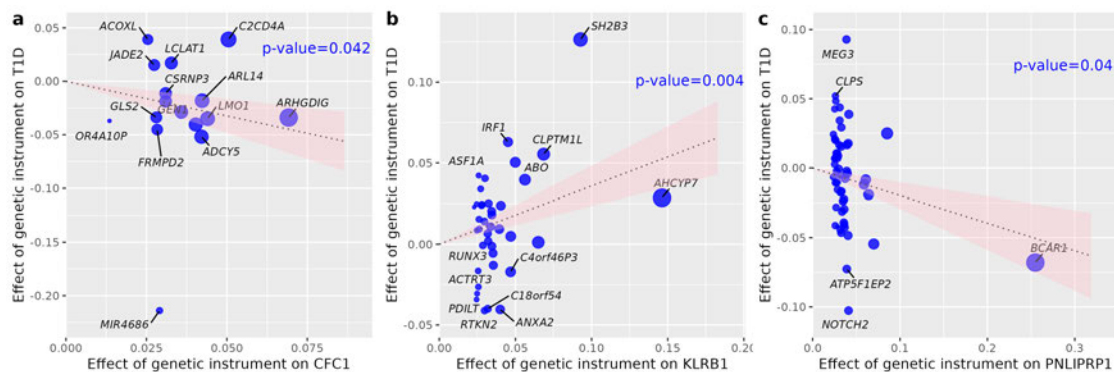


Figure S4.2: Suggestive dose-response effects of protein levels on T1D supported by the Mendelian randomisation analysis in the Scottish dataset. All *trans*-pQTLs contributing to the GATE score of the putative core proteins are used as genetic instruments. The effects of the genetic instrument on protein level and T1D are shown on the x-axis and y-axis, respectively. The size of each point is inversely proportional to the standard error of the ratio estimator. The slope of the straight line represents the maximum likelihood estimate of the causal effect parameter, while the shaded ribbon reflects one standard error above and below this maximum likelihood estimate.

Table S4.1: Code and detailed name for lipid-lowering, antihypertensive, and antithrombotic medications in UK Biobank.

Category	Code	Name
Statins	1141146234	atorvastatin
	1141192414	crestor 10mg tablet
	1140910632	eptastatin
	1140888594	fluvastatin
	1140864592	lescol 20mg capsule
	1141146138	lipitor 10mg tablet
	1140861970	lipostat 10mg tablet
	1140888648	pravastatin
	1141195196	ranzolonit 10mg tablet
	1141192410	rosuvastatin
	1141188146	simvador 10mg tablet
	1140861958	simvastatin
	1140910652	synvinolin
	1140910654	velastatin
	1140881748	zocor 10mg tablet
	1141200040	zocor heart-pro 10mg tablet
Ezetimibe	1141192736	ezetimibe
	1141192740	ezetrol 10mg tablet
Fibrates	1140861924	bezafibrate
	1140861926	bezalip 200mg tablet
	1140861928	bezalip-mono 400mg m/r tablet
	1140861954	fenofibrate
	1141172214	supralip 160mg m/r tablet
	1141162544	lipantil micro 67mg capsule
	1140862026	ciprofibrate
	1140861944	clofibrate
	1140861946	atromid-s 500mg capsule
	1140861856	gemfibrozil
	1140861858	lipid 300 capsule

4. Plasma proteins associated with T1D and their potential as drug targets 107

Nicotinic acid	1141157260	bezafibrate product
	1141157262	gemfibrozil product
	1140861868	nicotinic acid product
	1140910670	niacin
Bile acid sequestrants	1141188546	niaspan 500mg m/r tablet
	1140861848	colestid 5g/sachet granules
	1140888590	colestipol
	1140909780	colestyramine
	1141180722	colestyramine+aspartame 4g/sachet powder
	1141180734	colestyramine product
	1140861936	questran 4g/sachet powder
Angiotensin-converting enzyme inhibitors	1140861942	cholestyramine+aspartame 4g/sachet powder
	1140860750	captopril
	1140860764	captopril+hydrochlorothiazide 25mg/12.5mg tablet
	1140860752	acepril 12.5mg tablet
	1140851690	acezide 50mg tablets x56
	1140881716	acezide tablet
	1140860758	capoten 12.5mg tablet
	1140851692	capozide 50mg tablets x28
	1140881714	capozide tablet
	1140860882	cilazapril
	1140860790	enalapril maleate+hydrochlorothiazide 20mg/12.5mg tablet
	1140888552	enalapril
	1140860776	innovace 2.5mg tablet
	1140860784	innozide tablet
	1140888556	fosinopril
	1140860878	staril 10mg tablet
	1141164148	imidapril hydrochloride
	1141164154	tanatril 5mg tablet
	1140864952	lisinopril+hydrochlorothiazide 10mg/12.5mg tablet
	1140860696	lisinopril
	1140860706	carace 2.5mg tablet
	1140864910	carace 10 plus tablet
	1140864618	zestoretic 10 tablet
	1140860714	zestril 2.5mg tablet
	1140923712	moexipril
	1140923718	perdix 7.5mg tablet
	1140888560	perindopril
	1141180592	perindopril+indapamide
	1140860802	coversyl 2mg tablet
	1141180598	coversyl plus 4mg/1.25mg tablet
	1140860728	quinapril
	1140881706	accupro 5mg tablet
	1140860736	accuretic tablet
	1140860806	ramipril
	1141165470	felodipine+ramipril
	1141199940	lopace 2.5mg capsule
	1141165476	triapin mite 2.5mg/2.5mg tablet
	1141188408	tritace 1.25mg tablet
	1141153328	trandolapril+verapamil hydrochloride
	1140860904	trandolapril
	1140860912	gopten 500micrograms capsule
	1141153316	tarka 2mg/180mg m/r capsule
Angiotensin receptor blockers	1140916356	losartan
	1141145660	valsartan
	1141151016	losartan potassium+hydrochlorothiazide 50mg/12.5mg tablet
	1141152998	irbesartan
	1141156836	candesartan cilexetil

4. Plasma proteins associated with T1D and their potential as drug targets 108

	1141166006	telmisartan
	1141171336	eprosartan
	1141172682	irbesartan+hydrochlorothiazide 150mg/12.5mg tablet
	1141187788	telmisartan+hydrochlorothiazide 40mg/12.5mg tablet
	1141193282	olmesartan
	1141201038	valsartan+hydrochlorothiazide 80mg/12.5mg tablet
	1141171344	teveten 300mg tablet
	1140916362	cozaar half strength 25mg tablet
	1141151018	cozaar-comp 50mg/12.5mg tablet
	1141179974	cozaar 25mg tablet
	1141172492	micardis 20mg tablet
	1141187790	micardisplus 40mg/12.5mg tablet
	1141145668	diovan 40mg capsule
	1141201040	co-diovan 80mg/12.5mg tablet
	1141156846	amias 2mg tablet
	1141153006	aprovel 75mg tablet
	1141172686	coaprovel 150mg/12.5mg tablet
	1141193346	olmetec 10mg tablet
Thiazide-like diuretics	1140866078	indapamide
	1140909706	chlortalidone
	1141180772	triamterene+chlortalidone 50mg/50mg tablet
	1141180778	atenolol+chlortalidone
	1140866108	xipamide
	1140866092	metolazone
	1141146378	natrilix sr 1.5mg m/r tablet
	1140851364	hygroton k tablet combination pack
	1140866146	hygroton 50mg tablet
Calcium channel blockers	140860426	atenolol+nifedipine 50mg/20mg m/r capsule
	1140861088	nifedipine
	1140861190	isradipine
	1140861276	lacidipine
	1140872568	nimodipine
	1140879802	amlodipine
	1140879810	nicardipine
	1140888646	felodipine
	1140923572	adipine mr 10 m/r tablet
	1140928226	nisoldipine
	1141150500	slofedipine 20mg m/r tablet
	1141153026	lercanidipine
	1141165470	felodipine+ramipril
	1140861202	istin 5mg tablet
	1141200400	amlostin 5mg tablet
	1141153032	zanidip 10mg tablet
	1140879806	diltiazem
	1140926778	diltiazem hcl+hydrochlorothiazide 150mg/12.5mg m/r capsule
	1140861128	tildiem 60mg m/r tablet
	1140861138	adizem-60 m/r tablet
	1140926780	adizem-xl plus m/r capsule
	1140888510	verapamil
	1141153328	trandolapril+verapamil hydrochloride
	1140866460	half securon sr 120mg m/r tablet
	1140866466	securon 40mg tablet
	1141150926	verapress mr 240 m/r tablet
Aspirin	1140861806	aspirin 75mg tablet
	1140864860	nu-seals aspirin 75mg e/c tablet
	1140868226	aspirin
	1140868282	aspirin+methocarbamol 325mg/400mg tablet
	1140872040	aspirin+metoclopramide 325mg/5mg effervescent tablet
	1140882108	aspirin+cyclizine hydrochloride 500mg/25mg tablet

4. Plasma proteins associated with T1D and their potential as drug targets 109

1140882190	aspirin+glycine 500mg/133mg dispersible tablet
1140882268	aspirin+codeine 300mg/8mg tablet
1140882392	aspirin+codeine
1141163138	aspirin+papaveretum 500mg/7.71mg dispersible tablet
1141164044	isosorbide mononitrate+aspirin
1141167844	dipyridamole+aspirin
1140909772	acetylsalicylic acid
1140861804	angettes 75mg tablet
1141167848	asasantin retard m/r capsule
1140909888	benorilate
1140871080	benorylate
1140925942	caprin 75mg e/c tablet
1140923344	co-codaprin
1140856336	codis dispersible tablet
1140861808	disprin cv 100mg m/r tablet
1140882192	disprin direct dispersible tablet
1141151924	enprin 75mg e/c tablet
1140868264	equagesic tablet
1140882106	femigraine tablet
1140856394	hypon tablet
1141164050	imazin xl 60mg/75mg m/r tablet
1140856314	laboprin 300mg tablet
1140863514	laboprin dl 900mg sachet
1140856220	levius 500mg m/r tablet
1141177826	micropirin 75mg e/c tablet
1140872032	migravess effervescent tablet
1140856212	paynocil 600mg tablet
1141188536	phensic tablet
1140917408	postmi 75mg dispersible tablet
1140868294	robaxisal forte tablet
1140856440	safapryn-co tablet
1140856214	solprin 300mg dispersible tablet
1140856344	trancoprin tablet

4. Plasma proteins associated with T1D and their potential as drug targets 110

Table S4.2: Mendelian randomisation analysis of the effect of the potential core proteins on T1D.

Gene symbol	UK Biobank				Scottish dataset			
	Raw number of pQTLs	Estimate	Standard error	<i>p</i> -value	Raw number of pQTLs	Estimate	Standard error	<i>p</i> -value
<i>CFC1</i>	19	-0.529	0.244	0.03	16	-0.648	0.319	0.04
<i>NOTCH1</i>	25	0.601	0.244	0.01	23	0.112	0.250	0.7
<i>CD28</i>	29	0.316	0.103	0.002	28	0.128	0.146	0.4
<i>ADA2</i>	33	0.288	0.092	0.002	32	0.319	0.218	0.1
<i>ADGRE2</i>	28	0.382	0.140	0.006	7	0.420	0.567	0.5
<i>TNFRSF4</i>	24	0.386	0.161	0.02	6	0.892	0.851	0.3
<i>GCNT1</i>	27	0.401	0.173	0.02	24	0.310	0.285	0.3
<i>PCDH17</i>	24	0.249	0.123	0.04	23	-0.073	0.098	0.5
<i>SIGLEC6</i>	41	0.200	0.095	0.04	40	0.172	0.126	0.2
<i>DLL1</i>	19	0.283	0.162	0.08	18	0.157	0.155	0.3
<i>SEMA7A</i>	38	0.222	0.092	0.02	14	0.273	0.223	0.2
<i>ROBO1</i>	16	0.490	0.306	0.1	15	0.121	0.259	0.6
<i>ITGB2</i>	37	0.206	0.119	0.08	35	0.201	0.162	0.2
<i>OXT</i>	29	-0.348	0.183	0.06	26	-0.272	0.147	0.06
<i>DDR1</i>	25	0.241	0.134	0.07	26	0.294	0.176	0.1
<i>KLRB1</i>	42	0.164	0.119	0.2	39	0.358	0.125	0.004
<i>CHAD</i>	17	-0.124	0.188	0.5	28	-0.158	0.156	0.3
<i>ALCAM</i>	10	0.280	0.262	0.3	10	0.446	0.328	0.2
<i>AMIGO2</i>	35	0.067	0.092	0.5	35	0.174	0.105	0.1
<i>PRSS2</i>	11	-0.155	0.097	0.1	29	-0.254	0.180	0.2
<i>CD274</i>	14	0.277	0.274	0.3	3	2.493	2.510	0.3
<i>MUC2</i>	33	0.066	0.040	0.1	31	0.129	0.106	0.2
<i>PNLIPRP1</i>	36	-0.094	0.106	0.4	57	-0.198	0.097	0.04
<i>PLAT</i>	11	0.094	0.287	0.7	10	-0.487	0.292	0.1

All *trans*-pQTLs contributing to the GATE score of a potential core protein are used as genetic instruments for that protein. The maximum likelihood estimate, standard error, and *p*-value for the causal effect of each protein on T1D is reported. Estimation was performed by marginalising over the posterior distribution of pleiotropic effects of the *trans*-pQTLs.

4. Plasma proteins associated with T1D and their potential as drug targets 111

Table S4.3: *Cis*-score association for the potential core proteins with T1D.

Gene symbol	UK Biobank			Scottish dataset		
	pQTL study	Log OR	<i>p</i> -value	pQTL study	Log OR	<i>p</i> -value
<i>NOTCH1</i>	UKB	0.04	0.1	UKB	0.01	0.6
<i>CD28</i>	UKB	-0.06	0.01	UKB	-0.03	0.09
<i>ADA2</i>	UKB	0.01	0.6	UKB	0.01	0.6
<i>ADGRE2</i>	UKB	0.05	0.04	DeC	0.05	0.01
<i>TNFRSF4</i>	UKB	0.03	0.3	.	.	.
<i>GCNT1</i>	UKB	0.03	0.2	UKB	-0.02	0.2
<i>PCDH17</i>	UKB	0.00	1	UKB	0.01	0.6
<i>SIGLEC6</i>	UKB	-0.05	0.05	UKB	-0.01	0.7
<i>DLL1</i>	UKB	-0.03	0.2	UKB	0.01	0.6
<i>SEMA7A</i>	UKB	0.05	0.05	DeC	0.04	0.04
<i>ROBO1</i>	DeC	-0.02	0.4	DeC	-0.01	0.5
<i>ITGB2</i>	UKB	-0.01	0.6	UKB	-0.06	0.003
<i>OXT</i>	UKB	-0.03	0.2	UKB	-0.01	0.5
<i>KLRB1</i>	UKB	0.03	0.2	UKB	-0.02	0.3
<i>CHAD</i>	DeC	-0.01	0.8	UKB	0.00	1
<i>ALCAM</i>	DeC	-0.05	0.02	DeC	0.01	0.5
<i>AMIGO2</i>	UKB	0.00	1	UKB	-0.02	0.4
<i>PRSS2</i>	INT	0.01	0.7	DeC	-0.01	0.5
<i>CD274</i>	UKB	0.00	0.9	DeC	0.01	0.8
<i>MUC2</i>	UKB	0.02	0.4	UKB	0.00	0.8
<i>PNLIPRP1</i>	DeC	0.02	0.4	UKB	-0.02	0.4
<i>PLAT</i>	UKB	0.02	0.4	DeC	0.00	0.8

Log OR, log (odds ratio) of T1D risk per standard deviation increase in the *cis*-score. UKB, UK Biobank Pharma Proteomics Project; DeC, deCODE genetics; INT, INTERVAL.

Cis-scores for potential core proteins were computed when at least one *cis*-SNP was associated with the protein level at *p*-value < 10⁻⁶ in the pQTL GWAS.

4. Plasma proteins associated with T1D and their potential as drug targets 112

Table S4.4: Associations with T1D, HbA1c and duration for the potential core proteins in the UKB-PPP dataset.

Gene symbol	Association with T1D ¹		Association with HbA1c ²		Association with duration ³	
	Log OR ⁴	p-value	Coeff ⁵	p-value	Coeff ⁶	p-value
<i>CFC1</i>	-1.08	1×10^{-43}	0.48	0.7	-0.002	0.9
<i>NOTCH1</i>	0.60	1×10^{-12}	3.20	0.01	0.001	0.9
<i>CD28</i>	0.41	2×10^{-7}	-1.24	0.4	0.019	0.06
<i>ADA2</i>	0.92	2×10^{-27}	3.82	0.001	0.006	0.7
<i>ADGRE2</i>	0.81	3×10^{-20}	3.17	0.02	0.010	0.4
<i>TNFRSF4</i>	0.44	5×10^{-7}	-0.56	0.7	0.012	0.3
<i>GCNT1</i>	0.45	3×10^{-7}	-2.02	0.1	0.018	0.08
<i>PCDH17</i>	0.79	8×10^{-20}	2.10	0.1	0.000	1.0
<i>SIGLEC6</i>	0.62	6×10^{-19}	2.79	0.04	0.005	0.7
<i>DLL1</i>	0.45	2×10^{-7}	3.06	0.07	0.005	0.7
<i>SEMA7A</i>	0.80	1×10^{-22}	2.77	0.03	0.006	0.6
<i>ROBO1</i>	0.70	2×10^{-16}	1.95	0.2	0.002	0.9
<i>ITGB2</i>	0.77	4×10^{-23}	4.86	2×10^{-4}	0.004	0.8
<i>OXT</i>	-0.50	1×10^{-8}	3.66	0.01	-0.023	0.04
<i>DDR1</i>	0.50	5×10^{-10}	1.55	0.3	0.026	0.01
<i>KLRB1</i>	0.70	6×10^{-15}	0.68	0.6	-0.001	0.9
<i>CHAD</i>	-0.50	4×10^{-8}	-5.07	5×10^{-4}	0.018	0.1
<i>ALCAM</i>	0.44	7×10^{-7}	0.99	0.5	0.015	0.1
<i>AMIGO2</i>	0.56	2×10^{-12}	1.67	0.2	0.008	0.6
<i>PRSS2</i>	-1.11	2×10^{-53}	1.10	0.4	-0.038	0.1
<i>CD274</i>	0.41	4×10^{-7}	1.58	0.3	0.014	0.2
<i>MUC2</i>	0.73	6×10^{-17}	2.91	0.02	0.021	0.1
<i>PNLIPRP1</i>	-1.02	6×10^{-36}	0.61	0.7	-0.058	6×10^{-4}
<i>PLAT</i>	-0.44	6×10^{-7}	4.74	0.002	-0.001	0.9

¹Association with T1D was tested in a logistic regression model with T1D case-control status as the dependent variable, protein levels as the independent variable, and age at assessment, sex, ethnicity, kidney filtration, liver function and medication usage as covariates in the UKB-PPP dataset.

²Association with HbA1c was tested in a linear regression model with HbA1c as the dependent variable, protein levels as the independent variable, and age at assessment, sex, ethnicity, kidney filtration, liver function and medication usage as covariates among individuals with T1D in the UKB-PPP dataset.

³Association with duration was tested in a linear regression model with protein levels as the dependent variable, T1D duration as the independent variable, and age at onset, sex, ethnicity, kidney filtration, liver function and medication usage as covariates among individuals with T1D in the UKB-PPP dataset.

⁴Log OR, log (odds ratio) of T1D per standard deviation increase in the protein level.

⁵Coeff, coefficient representing HbA1c change per standard deviation increase in the protein level.

⁶Coeff, coefficient representing protein level change per year increase in T1D duration.

4. Plasma proteins associated with T1D and their potential as drug targets 113

Table S4.5: Gene name, protein name synonyms and biological functions of the potential core proteins for T1D.

UniProt ID	Gene names	Protein names	Biological functions
P07478	PRSS2 TRY2 TRYP2	Trypsin-2 (EC 3.4.21.4) (Anionic trypsinogen) (Serine protease 2) (Trypsin II)	A pancreatic enzyme, belongs to the trypsin family of serine proteases.
P54315	PNLIPRP1 PLRP1	Inactive pancreatic lipase-related protein 1 (PL-RP1)	May function as an inhibitor of dietary triglyceride digestion.
Q9NZK5	ADA2 ADGF CECR1 IDGFL	Adenosine deaminase 2 (EC 3.5.4.4) (Cat eye syndrome critical region protein 1)	A degrading enzyme for adenosine, which regulates immune homeostasis, is involved in multiple autoimmune diseases Gao et al. [2021] and modulates the bioactivity of insulin Hoshino et al. [1994].
P05107	ITGB2 CD18 MFI7	Integrin beta-2 (Cell surface adhesion glycoproteins LFA-1/CR3/p150,95 subunit beta) (Complement receptor C3 subunit beta) (CD antigen CD18)	A leukocyte-specific adhesion molecule that involves in phagocytosis, regulation of T cell activation, target cell killing, leukocyte recruitment, regulation of Toll-like receptor signalling, restriction of dendritic cell migration and B cell receptor signalling Fagerholm et al. [2019].
O75326	SEMA7A CD108 SEMAL	Semaphorin-7A (CDw108) (JMH blood group antigen) (John-Milton-Hargen human blood group Ag) (Semaphorin-K1) (Sema K1) (Semaphorin-L) (Sema L) (CD antigen CD108)	An immune semaphorin that initiates T-cell-mediated immune responses and has been shown to involves in various autoimmune diseases, inflammatory diseases, and tumours Suzuki et al. [2007], Song et al. [2021].
Q9UHX3	ADGRE2 EMR2	Adhesion G protein-coupled receptor E2 (EGF-like module receptor 2) (EGF-like module-containing mucin-like hormone receptor-like 2) (CD antigen CD312)	Expressed on activated lymphocytes and myeloid cells, binds to chondroitin sulfate on B cells Kwakkenbos et al. [2005].
O14917	PCDH17 PCDH68 PCH68	Protocadherin-17 (Protocadherin-68)	Regulates presynaptic assembly Hoshina et al. [2013], also acts as a tumour suppressor Hu et al. [2013].
O43699	SIGLEC6 CD33L CD33L1 OBBP1	Sialic acid-binding Ig-like lectin 6 (Siglec-6) (CD33 antigen-like 1) (CDw327) (Obesity-binding protein 1) (OB-BP1) (CD antigen CD327)	An inhibitory receptor that is primarily expressed on mast cells and represents a therapeutic target for allergic diseases and leukemia Schanin et al. [2022], Nunes et al. [2024].
Q02817	MUC2 SMUC	Mucin-2 (MUC-2) (Intestinal mucin-2)	A protein that forms the protective mucus barrier in gut.
Q9Y6N7	ROBO1 DUTT1	Roundabout homolog 1 (Deleted in U twenty twenty) (H-Robo-1)	Receptor for SLIT proteins, modulates cell migration during neuronal development.
P46531	NOTCH1 TAN1	Neurogenic locus notch homolog protein 1 (Notch 1) (hN1) (Translocation-associated notch protein TAN-1) [Cleaved into: Notch 1 extracellular truncation (NEXT); Notch 1 intracellular domain (NICD)]	Activates the Notch signalling pathway, regulates pancreatic endocrine and exocrine cell fate Eom et al. [2021].

4. Plasma proteins associated with T1D and their potential as drug targets 114

Q86SJ2	AMIGO2 <i>ALI1</i>	Amphoterin-induced protein 2 (AMIGO-2) (Alivin-1) (Differentially expressed in gastric adenocarcinomas) (DEGA)	A cell adhesion molecule that involves in cell-cell interactions, signal transduction, neuronal development, and potentially cancer metastasis.
Q08345	DDR1 <i>CAK EDDR1 NEP NTRK4 PTK3A RTK6 TRKE</i>	Epithelial discoidin domain-containing receptor 1 (Epithelial discoidin domain receptor 1) (EC 2.7.10.1) (CD167 antigen-like family member A) (Cell adhesion kinase) (Discoidin receptor tyrosine kinase) (HGK2) (Mammary carcinoma kinase 10) (MCK-10) (Protein-tyrosine kinase 3A) (Protein-tyrosine kinase RTK-6) (TRK E) (Tyrosine kinase DDR) (Tyrosine-protein kinase CAK) (CD antigen CD167a)	Receptor tyrosine kinase that regulates cell growth, differentiation and metabolism.
P01178	OXT <i>OT</i>	Oxytocin-neurophysin 1 (OT-NPI) [Cleaved into: Oxytocin (Ocytocin); Neurophysin 1]	Precursor for oxytocin and neurophysin I.
O15335	CHAD <i>SLRR4A</i>	Chondroadherin (Cartilage leucine-rich protein)	A cartilage matrix protein that mediates adhesion of isolated chondrocytes.
P10747	CD28	T-cell-specific surface glycoprotein CD28 (TP44) (CD antigen CD28)	A stimulatory immune checkpoint on T cells, binds with CD80, CD86, ICOSL.
O00548	DLL1 <i>UNQ146/PRO172</i>	Delta-like protein 1 (Drosophila Delta homolog 1) (Delta1) (H-Delta-1)	Ligand of NOTCH1, mediates cell fate decisions.
Q02742	GCNT1 <i>NACGT2</i>	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase (EC 2.4.1.102) (Core 2-branching enzyme) (Core2-GlcNAc-transferase) (C2GNT) (Core 2 GNT)	An enzyme for O-glycosylation, is required for selectin ligand synthesis.
Q9NZQ7	CD274 <i>B7H1 PDCD1L1 PDCD1LG1 PDL1</i>	Programmed cell death 1 ligand 1 (PD-L1) (PDCD1 ligand 1) (Programmed death ligand 1) (hPD-L1) (B7 homolog 1) (B7-H1) (CD antigen CD274)	Ligand of inhibitory immune checkpoint PDCD1.
P43489	TNFRSF4 <i>TXGP1L</i>	Tumor necrosis factor receptor superfamily member 4 (ACT35 antigen) (OX40L receptor) (TAX transcriptionally-activated glycoprotein 1 receptor) (CD antigen CD134)	A stimulatory immune checkpoint on T cells, binds with OX40L.
P00750	PLAT	Tissue-type plasminogen activator (t-PA) (t-plasminogen activator) (tPA) (EC 3.4.21.68) (Alteplase) (Retepase) [Cleaved into: Tissue-type plasminogen activator chain A; Tissue-type plasminogen activator chain B]	A secreted serine protease that converts the proenzyme plasminogen to plasmin, a fibrinolytic enzyme.

4. Plasma proteins associated with T1D and their potential as drug targets 115

Q13740	ALCAM <i>MEMD</i>	CD166 antigen (Activated leukocyte cell adhesion molecule) (CD antigen CD166)	A cell adhesion molecule that involves in inflammation and cancer, binds to CD6 Yang et al. [2021].
--------	--------------------------	--	---

The Hugo Gene names for the genes encoding the potential core proteins along with the corresponding UniProt protein names are shown in bold.

5

Prospective validation of proteins
associated with T1D in high-risk
individuals.

5.1 Introduction

In Chapter 3, I identified 27 putative core proteins for T1D based on GATE score associations. In Chapter 4, I used an alternative strategy, starting with a cross-sectional analysis of the associations between measured protein levels and T1D in a smaller dataset (UKB-PPP) and then following up the identified proteins by testing for the association of their GATE scores with T1D in two large datasets (the Scottish dataset and UK Biobank). Twenty-four proteins were supported by both the protein associations and GATE score associations with consistent direction of effect in this design. Combining together, there are 49 unique proteins.

In this chapter, I aimed to provide further evidence for a causal role of these proteins in T1D using data from the Environmental Determinants of Diabetes in the Young (TEDDY) study. TEDDY is a prospective cohort of newborns at high risk for T1D that were followed up longitudinally to detect initiation of islet autoimmunity (IA) and progression to T1D. Transcriptomic and proteomic measurements are available based on blood samples collected at recruitment and follow-up time-points. The hypothesis of this chapter is that if the proteins identified in the earlier Chapters are causal for T1D, their transcripts or protein levels would likely be dysregulated in individuals who progressed to IA or to T1D. Specifically, the objective of this chapter was to test the association between transcripts and protein levels for these proteins measured before the age of one and the onset of IA and of T1D within the next six years of follow-up.

In this chapter, phenotype, gene expression, and protein data were pre-processed by the TEDDY Study Group and obtained by download. I reviewed data availability, performed data cleaning, wrote the initial analysis plan which I then revised together with my supervisors and our collaborator (Prof Helen Colhoun), conducted all the statistical analyses, and wrote the original draft of the chapter. The chapter was reviewed by my supervisors and Prof Colhoun and I made all revisions based on their feedback.

5.2 Methods

5.2.1 Study population

The TEDDY study is a multicentre prospective cohort aimed at identifying environmental risk factors associated with the onset of IA and T1D in high-risk newborns

Table 5.1: HLA haplotype eligibility for inclusion in the TEDDY study.

Code	Haplotype genotypes	Abbreviation	First-degree relatives	General population
A	DRB1*04:0X~DQA1*03:0X~DQB1*03:02/DRB1*03:0X~DQA1*05:01~DQB1*02:01	DR3/4	Y	Y
B	DRB1*04:0X~DQA1*03:0X~DQB1*03:02/DRB1*04:0X~DQA1*03:0X~DQB1*03:02	DR4/4	Y	Y
C	DRB1*04:0X~DQA1*03:0X~DQB1*03:02/DRB1*08:0X~DQA1*04:01~DQB1*04:02	DR4/8	Y	Y
D	DRB1*03:0X~DQA1*05:01~DQB1*02:01/DRB1*03:0X~DQA1*05:01~DQB1*02:01	DR3/3	Y	Y
E	DRB1*04:0X~DQA1*03:0X~DQB1*03:02/DRB1*04:0X~DQA1*03:0X~DQB1*02:0X	DR4/4b	Y	N
F	DRB1*04:0X~DQA1*03:0X~DQB1*03:02/DRB1*01:0X~DQA1*01:01~DQB1*05:01	DR4/1	Y	N
G	DRB1*04:0X~DQA1*03:0X~DQB1*03:02/DRB1*13:0X~DQA1*01:02~DQB1*06:04	DR4/13	Y	N
H	DRB1*04:0X~DQA1*03:0X~DQB1*03:02/DRB1*09:0X~DQA1*03:0X~DQB1*03:03	DR4/9	Y	N
I	DRB1*03:0X~DQA1*05:01~DQB1*02:01/DRB1*09:0X~DQA1*03:0X~DQB1*03:03	DR3/9	Y	N

Source: adapted from Hagopian et al., *Pediatric Diabetes* 2011; 12: 733–743. Y, eligible for inclusion; N, not eligible for inclusion. Although DQB1*03:02 is shown above, DQB1*03:04 is acceptable in its place for inclusion. Subtyping was not required for further characterization of DQB1*02:0X and DQA1*03:0X genotypes. DR4 subtyping was required to exclude GP newborns with DRB1*04:03, but no other DRB1 subtyping was required.

[Kiviniemi et al., 2007, Hagopian et al., 2011]. Infants from the general population and families with first-degree relatives with T1D were screened for HLA class II genotypes across six clinical centres in Europe (Germany, Sweden, and Finland) and the United States (Colorado, Georgia/Florida, and Washington). A two-stage screening strategy was adopted for HLA genotyping. Nine high-risk haplotypes were identified with an estimated relative risk larger than 3 in all six datasets (Table 5.1). General population newborns with the top four HLA haplotypes (A–D) and first-degree relative newborns with all nine haplotypes (A–I) were included for the TEDDY cohort.

In total, 424,788 newborns were screened, 21,589 were eligible for inclusion, and 8676 with consent were enrolled for longitudinal follow-up at a 3-month interval from the age of 3 months to 4 years [Vehik et al., 2013]. Follow-ups included health and environmental questionnaires, anthropometric measurements, biospecimen collection, and testing for three diabetes-associated autoantibodies (GADA, IA-2A, and IAA). Subjects who seroconverted to autoantibody positivity continued follow-up at the 3-month interval until the age of 15 years, whereas the others switched to a 6-month follow-up interval.

Two primary endpoints were defined in the TEDDY study.

1. Persistent IA, characterised by the presence of one confirmed autoantibody (GADA, IA-2A or IAA) in two or more consecutive samples. All samples were double-tested in two laboratories, and only samples with concordant positive results were considered positive. If the samples were drawn before the age of 18 months, a prior negative sample or a negative cord blood/mother sample was required.
2. Clinical T1D, diagnosed using the American Diabetes Association Expert Committee classification. Two criteria must be met for the diagnosis, unless criterion (4) is met.

- (1) Casual plasma glucose ≥ 200 mg/dL with unequivocal symptoms (i.e., polyuria, polydipsia, polyphagia, and/or weight loss).
- (2) Fasting plasma glucose ≥ 126 mg/dL.
- (3) 2-hour plasma glucose ≥ 200 mg/dL during an oral glucose tolerance test (OGTT). The glucose dose is determined by body weight to a maximum of 75 g.
- (4) Unequivocal hyperglycaemia with acute metabolic decompensation (i.e., ketoacidosis).

Within the prospective cohort, a nested case-control study was established as an alternative to reduce the cost of biomarker measurements while maintaining statistical power and controlling for potential bias [Lee et al., 2014]. In the nested case-control study, subjects who developed either of the two primary endpoints were classified as cases. Matched controls were randomly selected from the remaining individuals who were event-free within ± 45 days of the case's event time (i.e., incidence density sampling). Thus, the censored time for controls was recorded as the same as the time-to-event for their matched cases. Cases and controls were matched by clinical centre, sex, and family history of T1D at a 1:1 ratio for gene expression and protein measurements.

5.2.1.1 Gene expression

Gene expression was measured in peripheral blood samples using both microarray and RNA sequencing (RNA-seq) techniques.

1. Microarray data Purified RNA was amplified, labelled, and hybridised to Illumina Human HT-12 Expression BeadChips covering over 47,000 human transcripts and splice variants [Xhonneux et al., 2021]. Arrays were scanned to generate gene expression data, which were background-corrected and normalised using the robust spline normalisation (RSN) method.

2. RNA sequencing data RNA sequences were obtained using the Illumina HiSeq2500 platform, and aligned to the GRCh38 genome and transcriptome [Lin et al., 2023]. Gene expression levels were quantified and normalised using the Total Ubiquitous (TU) algorithm with the R package NormExpression, which scales each sample to equalise the total counts for ubiquitous genes.

5.2.1.2 Serum proteins

Serum cytokines and inflammatory protein biomarkers were measured using the Olink Target 96 Inflammation panel, which covers 92 proteins. The protein levels were measured using the Olink Proximity Extension Assay (PEA) technology, and quantified as the Normalised Protein eXpression (NPX) values.

5.2.2 Data cleaning and statistical analyses

Prior to the investigation of associations with the primary endpoints, I did four exploratory analyses to understand the data structure and assess measurement quality and consistency.

- Data availability: I tabulated the availability of the microarray, RNA-seq and protein data across all timepoints.
- Within-platform consistency: Given that multiple probes were designed to target different transcripts of the same gene in the microarray data, I calculated correlations between these probes for the same gene in the same blood sample to assess whether different transcripts for the same gene have similar abundances.
- Cross-platform consistency: For genes with both microarray and RNA-seq measurements from the same blood sample and timepoint, I calculated correlations between these two platforms to evaluate the consistency of these two techniques.
- Longitudinal consistency: I calculated correlations for microarray, RNA-seq, and protein measurements at different timepoints within blood samples collected from the same individual.

Eight housekeeping genes (*ACTB*, *B2M*, *GAPDH*, *GUSB*, *PPIB*, *RPLP0*, *RPL13*, and *YWHAZ*) [Khanna et al., 2017, Hieronymus et al., 2021], expected to be constitutively and ubiquitously expressed, were used as reference genes for the assessment of within-platform consistency, cross-platform consistency, and longitudinal consistency.

Due to the missingness of the transcriptomic and proteomic data, I calculated the average transcript and protein levels before the age of one using the mean values of the 3, 6, 9, and 12-month measurements and then tested their associations with incident IA and T1D. Any IA and T1D cases developed before the age of

one were omitted from the analysis along with their controls. Since the potential confounding variables have been matched for cases and controls, a univariate conditional logistic regression model was fitted. The average transcript and protein level measurements before the age of one were scaled to a standard normal distribution, i.e., mean = 0 and standard deviation (SD) = 1. Effect estimates were presented as log conditional odds ratios (log cOR) for the risk of incident IA or T1D per SD increase of the average measurements before the age of one. The Bonferroni approach was used to account for multiple testing.

In addition, to reflect the incidence density sampling design, I stratified case-control pairs by the age at onset. Specifically, I grouped cases who developed IA or T1D at different ages: those who developed it between 12 and 24 months (age one), between 24 and 36 months (age two), between 36 and 48 months (age three), and between 48 and 60 months (age four), along with their corresponding controls, into four strata for stratification analyses. Strata with fewer than 20 case-control pairs were excluded from analysis to reduce instability due to low sample size.

5.3 Results

5.3.1 Characteristics of the TEDDY nested case-control study

The nested case-control study included 418 pairs of IA and 114 pairs of T1D (Table 5.2). Among these, 44% of IA pairs and 46% of T1D pairs were female. 23% of IA cases and 36% of T1D cases were identified through first-degree relatives of T1D patients, while 77% of IA cases and 64% of T1D cases were identified through general population screening. Geographically, 30% of IA pairs and 26% of T1D pairs were enrolled in the United States, and 70% of IA pairs and 74% of T1D pairs were enrolled in Europe. The age at onset ranged from 2.6 to 72.4 months for IA cases (mean = 23.8 months) and from 8.2 to 75.1 months for T1D cases (mean = 32.1 months). When stratified by age at onset, 105 IA cases and 5 T1D cases were identified before the age of one (< 12 months). At the age of one (12-24 months), 137 IA and 38 T1D cases developed; at the age of two (24-36 months), 96 IA and 30 T1D cases; at the age of three (36-48 months), 47 IA and 18 T1D cases; at the age of four (48-60 months), 22 IA and 16 T1D cases. Another 11 IA and 7 T1D cases developed after the age of four (> 60 months), but these were only included in the main analyses but not stratification

Table 5.2: Characteristics of persistent islet autoimmunity (IA) and type 1 diabetes (T1D) case-control pairs in the TEDDY nested case-control study.

Variable		IA case-control pairs		T1D case-control pairs	
		N	%	N	%
Sex	Female	184	44	61	54
	Male	234	56	53	46
Centre	Colorado	57	14	16	14
	Georgia/Florida	29	7	6	5
	Washington	38	9	8	7
	Finland	114	27	36	32
	Germany	37	9	18	16
	Sweden	143	34	30	26
Source	First-degree relatives	95	23	41	36
	General population	323	77	73	64
Age at onset	< 12 months	105	25	5	4
	12-24 months	137	33	38	33
	24-36 months	96	23	30	26
	36-48 months	47	11	18	16
	48-60 months	22	5	16	14
	> 60 months	11	3	7	6

analyses. Figure 5.1 visually illustrates the development of IA and T1D in the nested case-control study. Among the 114 T1D cases, 95 progressed from IA cases.

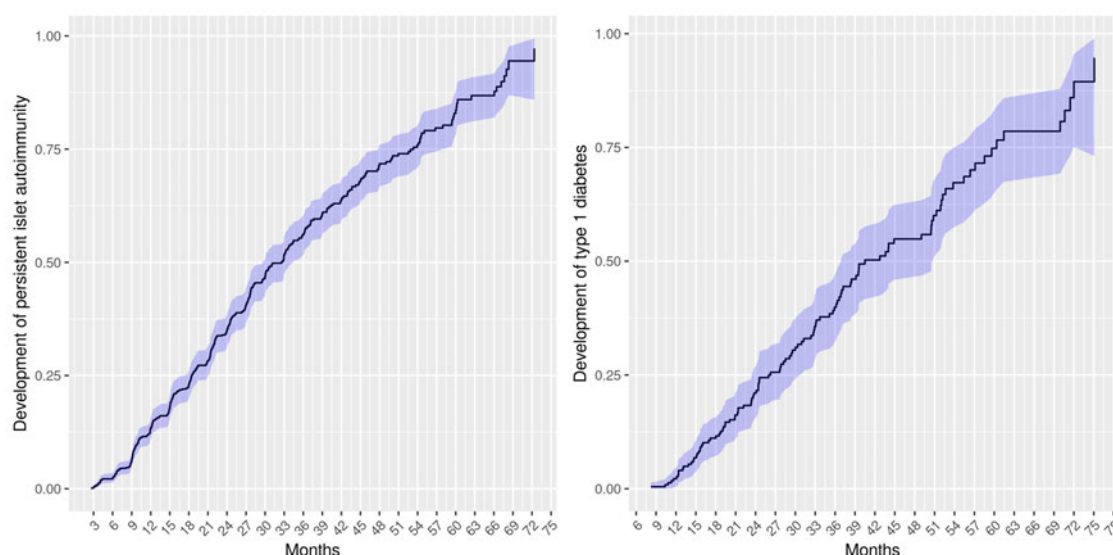


Figure 5.1: The development of persistent islet autoimmunity (IA) and type 1 diabetes (T1D) in the TEDDY nested case-control study. The age at onset ranged from 2.6 to 72.4 months for IA cases (mean = 23.8 months) and from 8.2 to 75.1 months for T1D cases (mean = 32.1 months).

Microarray gene expression data were obtained from 2013 blood samples of 401 subjects. Samples were collected between 3 and 72 months of age, with data availability shown in Figure S5.1. 169 IA case-control pairs and 62 T1D case-control pairs had microarray data available for at least one timepoint. RNA-seq

data were obtained from 5034 blood samples of 899 subjects, collected from the age of 3 to 78 months. 390 IA case-control pairs and 112 T1D case-control pairs had RNA-seq data available for at least one timepoint. Serum protein data was available for 6652 blood samples from 934 subjects, collected from the age of 3 months to 75 months. 416 IA case-control pairs and 114 T1D case-control pairs had serum protein data available for at least one timepoint. Among the 49 candidate core proteins, only eight (*CXCL11*, *CXCL9*, *CCL19*, *CD5*, *CXCL10*, *FGF19*, *LTA*, and *CD274*) were measured by the Olink Target 96 Inflammation panel.

The microarray panel included 64 probes measuring transcripts corresponding to the 49 candidate core proteins identified from the previous two chapters, along with 15 probes targeting eight housekeeping genes. The probe information is summarised in Table S5.1. Among them, 12 genes corresponding to the candidate proteins (*ADA2*, *AGR2*, *CFC1*, *CXCL11*, *DDR1*, *GCTN1*, *ITGB2*, *LGALS9*, *ROBO1*, *SIGLEC6*, *TIGIT*, and *TNFRSF4*) and six housekeeping genes (*ACTB*, *B2M*, *GAPDH*, *RPL13*, *RPLP0*, and *YWHAZ*) had multiple probes. *ADA2*, *DDR1*, *ROBO1*, and *RPLP0* each had three probes; two probes target different segments of the same isoform, while the other targets a different isoform. For *AGR2*, *GCTN1*, *ITGB2* and *SIGLEC6*, their two probes target distinct isoforms. For *CFC1*, *CXCL11*, *LGALS9*, *TIGIT*, *TNFRSF4*, *ACTB*, *B2M*, *GAPDH*, *RPL13*, and *YWHAZ*, their two probes target different segments of the same isoform. As shown in Table S5.2, for probes targeting genes of the candidate core proteins, only the probes for *ITGB2* and *LGALS9* exhibited high within-platform consistency (Pearson $r = 0.965$, 0.732), while probes for the other genes were not correlated with each other (absolute Pearson $r \leq 0.3$). Among housekeeping genes, two probes for *GAPDH* were strongly correlated (Pearson $r = 0.813$), while probes for *RPLP0*, *B2M*, *YWHAZ*, and *ACTB* showed moderate correlations (Pearson $r = 0.653$, 0.515 , 0.508 , and 0.457 , respectively).

For the 37 genes with only one probe in the microarray data, 15 (*PDCD1*, *CXCL10*, *CD5*, *FASLG*, *IDO1*, *KLRB1*, *TCN1*, *FCRL6*, *LAG3*, *ADGRE2*, *ALCAM*, *CD274*, *CD28*, *DLL1*, *NOTCH1*) showed good consistency between microarray and RNA-seq measurements (Pearson $r \geq 0.6$, Table S5.3). For the 12 genes with multiple probes, RNA-seq measurements of *LGALS9* were consistent with both of its microarray probes, while RNA-seq measurements of *TIGIT*, *ADA2*, and *GCTN1* were consistent with only one of the microarray probes. For housekeeping genes, *GAPDH* showed the strongest cross-platform correlation (Pearson $r = 0.535$), while the others showed poor agreement.

Longitudinal consistency was generally low across microarray, RNA-seq, and protein data (Table S5.4, Table S5.5, Table S5.6). For microarray data, the median between-timepoint correlation coefficient was 0.297 for housekeeping genes and 0.269 for genes corresponding to candidate core proteins. For RNA-seq data, the median between-timepoint correlation coefficient was 0.257 for housekeeping genes and 0.263 for genes corresponding to candidate core proteins. RNA-seq measurements for *PNLIPRP1* contained substantial missing data; therefore, this gene was excluded from the association analyses. Though proteins for the housekeeping genes were not included in the Olink Target 96 Inflammation panel, the eight measured proteins had a median between-timepoint correlation coefficient of 0.315.

5.3.2 Association of transcript and protein levels with incident IA and T1D

5.3.2.1 Average transcript levels measured by microarray before the age of one

The analysis included 104 IA pairs and 54 T1D pairs to examine associations between average transcript levels measured by microarray before the age of one and the development of IA and T1D within the next six years of follow-up. A Bonferroni-corrected p -value threshold was set at 8×10^{-4} (0.05/64). The study had 80% statistical power to detect effects with the magnitude of log cOR exceeding 0.58 for IA and 0.81 for T1D.

For IA, two probes showed nominal association (Table 5.3). These probes corresponded to *PDCD1* (log cOR = -0.53, p -value = 0.006) and *CD48* (log cOR = -0.5, p -value = 0.017). For T1D, two probes demonstrated nominal association: *PCDH17* (log cOR = -0.82, p -value = 0.017) and *AGR2* (log cOR = -0.8, p -value = 0.032). None of these associations survived multiple testing corrections. Full results for each probe are presented in Table S5.7 for IA and Table S5.8 for T1D.

Stratified by the age of onset, 50, 25, 17, and eight pairs of IA were in each stratum. IA pairs with the case developing during age one or age two were analysed (Table S5.7). Two probes (*CCL19* and *PRSS27*) were nominally associated with IA developed during the age of two. *CCL19* was associated with a lower risk of IA (log cOR = -2.41, p -value = 0.014), while *PRSS27* was associated with a higher risk of IA (log cOR = 0.93, p -value = 0.032). Neither association met the Bonferroni-corrected p -value threshold. As for T1D, 19, 15, eight, and eight

Table 5.3: Nominal associations of average transcripts and protein levels before the age of one with the risk of incident IA and T1D within the next six years of follow-up.

Gene symbol	Probe	Log cOR	p-value
Associations of average transcripts levels before the age of one measured by microarray with the risk of incident IA (N = 104 pairs)			
<i>PDCD1</i>	ILMN_1806725	-0.53	0.006
<i>CD48</i>	ILMN_2061043	-0.50	0.02
Associations of average transcripts levels before the age of one measured by microarray with the risk of incident T1D (N = 54 pairs)			
<i>PCDH17</i>	ILMN_1781514	-0.82	0.02
<i>AGR2</i>	ILMN_1814151	-0.80	0.03
Associations of average transcripts levels before the age of one measured by RNA-seq with the risk of incident IA (N = 274 pairs)			
<i>SIGLEC6</i>	.	-0.49	3×10^{-5}
<i>TNFRSF4</i>	.	0.30	0.007
<i>DDR1</i>	.	-0.24	0.01
<i>SELL</i>	.	0.23	0.02
<i>OXT</i>	.	0.25	0.03
<i>CD28</i>	.	-0.18	0.05
Associations of average proteins levels before the age of one with the risk of incident IA (N = 307 pairs)			
<i>LTA</i>	.	0.28	0.001
<i>CCL19</i>	.	0.20	0.02

Log cOR: log (conditional odds ratio) of IA or T1D risk per standard deviation increase in the average transcripts or protein levels before the age of one.

pairs of T1D were identified for each year. Due to the small sample size, no stratification analysis was performed.

5.3.2.2 Average transcript levels measured by RNA-seq before the age of one

Using RNA-seq data on gene expression, 274 pairs of IA and 97 pairs of T1D were included in the association analysis of mean transcript levels before the age of one with incident IA and with incident T1D. The Bonferroni-corrected *p*-value threshold was set to 0.001 (0.05/48, since *PNLIPRP1* was excluded for these analyses due to missingness). The study had 80% statistical power to detect effects with absolute log cOR exceeding 0.35 for IA and 0.59 for T1D.

Average transcript levels for *SIGLEC6*, *DDR1*, and *CD28* were associated with lower risk of IA, while *TNFRSF4*, *SELL*, and *OXT* were associated with higher risk of IA (Table 5.3). The *SIGLEC6* transcript levels displayed the strongest association with IA (log cOR = -0.49, *p*-value = 3×10^{-5}), which remained significant after Bonferroni correction. The other associations did not meet this stringent *p*-value threshold. No association was observed in the analysis of incident T1D, mainly due to limited statistical power.

Stratified by the age at onset, 123, 85, 40, and 19 pairs of IA were identified in each stratum. Subgroup analysis was conducted on incident IA that developed during the ages of one, two, and three (Table S5.9). Average transcript levels for two genes showed nominal associations with IA developed during the age of one. Average transcript levels for *OXT* were nominally associated with higher risk of incident IA developed during the age of one (log cOR = 0.33, *p*-value = 0.023), while average transcript levels for *SIGLEC6* were nominally associated with decreased risks of incident IA developed during the age of one (log cOR = -0.42, *p*-value = 0.011). Average transcript levels for nine genes showed associations with IA developed during the age of two; six (*CD28*, *DDR1*, *DLL1*, *GCNT1*, *NCR1*, *SIGLEC6*) were associated with lower risks of IA, whereas three (*CD48*, *ITGB2*, *TNFRSF4*) were associated with higher risks. Only the association between *SIGLEC6* transcripts and IA developed during the age of two met the stringent *p*-value threshold. No association was observed for IA developed during the age of three. In terms of T1D, 37, 26, 15 and 13 pairs were identified for each stratum, and stratification analysis was performed for those developed during the ages of one and two. As shown in Table S5.10, average transcript levels for one gene (*LTA*) was nominally associated with a lower risk of incident T1D developed during the age of one (log cOR = -0.73, *p*-value = 0.039). No association was observed for T1D developed during the age of two.

5.3.2.3 Average protein levels before the age of one

The analysis investigating the relationship between average protein levels before the age of one and subsequent disease development included 307 pairs of IA and 107 pairs of T1D. The Bonferroni-corrected *p*-value threshold was set to 0.006 (0.05/8, since only eight proteins were measured). The analysis had 80% statistical power to detect effects with log cOR exceeding 0.29 for IA and 0.49 for T1D.

The average levels for two proteins (*CCL19* and *LTA*) were associated with higher risk of incident IA (log cOR = 0.20 and 0.28; *p*-value = 0.016 and 0.001) (Table 5.3). The association of average protein levels for *LTA* remained statistically significant after multiple testing corrections. No association was observed in T1D.

In the subgroup analysis by age at onset, 135, 94, 45, and 22 pairs of IA were included. In Table S5.11, average protein levels for *CCL19* were associated with a higher risk of incident IA developed during the age of one (log cOR = 0.29,

p -value = 0.022). Average protein levels for *LTA* were associated with a higher risk of incident IA developed during the age of two (log cOR = 0.41 p -value = 0.012). Neither of these associations passed the stringent p -value threshold after Bonferroni correction. No association was observed for incident IA that developed during the ages of three and four. As for T1D, 38, 30, 16, and 16 pairs of T1D were identified in each year, and subgroup analysis was performed for the pairs with incident T1D developing during the age of one and during the age of two (Table S5.12). No association was observed.

5.4 Discussion

In this chapter, I presented a study of the associations of transcript and protein levels for 49 unique proteins identified in Chapter 3 and Chapter 4 with incident IA and incident T1D in the TEDDY nested case-control study.

Average transcript levels of *PDCD1*, *CD48*, *SIGLEC6*, *DDR1*, *CD28*, *TNFRSF4*, *SELL*, and *OXT* before the age of one were associated with incident IA, with *SIGLEC6* remaining significant after multiple testing corrections. Average transcript levels for *PCDH17* and *AGR2* before the age of one were nominally associated with incident T1D. At the protein level, *LTA* and *CCL19* levels before the age of one revealed associations with incident IA, with *LTA* surviving multiple testing corrections. In addition, transcript levels for six more genes (*PRSS27*, *CCL19*, *DLL1*, *GCNT1*, *NCR1*, and *ITGB2*) demonstrated nominal associations with incident IA in the stratification analysis on age at onset.

5.4.1 Methodological limitations

Overall, most associations were not significant or only nominally significant, and findings were more frequently linked to IA rather than T1D. This is likely due to the small sample size of the study. A second important factor is the source of variance (technical or biological) in the transcript and protein measurements. The low within-platform consistency, modest cross-platform consistency and poor longitudinal consistency demonstrated in this chapter suggest that the measurements have some degree of technical noise. Combined with the low sample size, the signal-to-noise ratio may be too low to distinguish inherent biological differences.

Notably, the majority of suggestive findings involved immune-related proteins, such as immune checkpoints (*PDCD1*, *CD48*, *CD28* and *TNFRSF4*), as well as cytokines (*CCL19* and *LTA*). Only *AGR2* and *PRSS27* were related to gut and exocrine pancreas functions. This could be attributed to two potential reasons. Firstly, the proteomics data in the TEDDY study was measured by the Olink Target 96 Inflammation panel. Among the 49 candidate core proteins, 38 are immune-related proteins, of which seven were included in this panel, while of the 11 pancreatic and gut proteins, only one (*FGF19*) was covered in this panel. Secondly, the participants of the TEDDY study were very young, and gut and exocrine pancreas involvement in the pathogenesis of T1D may become apparent at later disease stages.

It should also be noted that the direction of effect for transcripts in this chapter is not always consistent with the GATE score associations and protein associations in previous chapters. Both GATE scores and measured protein levels in previous chapters reflect circulating soluble protein levels in the blood. Due to the different sources of soluble proteins (i.e., alternative splicing and proteolytic cleavage from the membrane-bound form), transcripts in this chapter might represent distinct isoforms from those detected as soluble proteins in earlier chapters. Although mRNA and corresponding protein information for each probe in the microarray was provided in the Illumina Human HT-12 Expression BeadChip manifest file (Table S5.1), the specific transcript or protein isoforms captured by RNA-seq and proteomic platforms were not available. Consequently, it was not possible to determine which specific isoforms are most relevant to the pathogenesis of T1D. Future studies leveraging isoform-resolved datasets, such as those available through the Genotype-Tissue Expression (GTEx) project, could help address this limitation.

Gene expression data (from both microarray and RNA-seq) in this nested case-control study were obtained from whole blood samples [Xhonneux et al., 2021, Lin et al., 2023]. As such, measured transcripts reflect RNAs in all circulating blood cells, including white blood cell and residual red blood cell related transcripts, such as globin transcripts from reticulocytes, which pose a recognised technical challenge in whole-blood transcriptomic studies. In addition, an inherent limitation of whole-blood transcriptomics is that observed differences in gene expression may arise not only from changes in transcriptional activity within specific immune cell types, but also from differences in the relative proportions of circulating white blood cell subsets. Previous TEDDY transcriptomic analyses

have demonstrated that immune cell composition, particularly monocytes and B cells, differs between individuals who later progressed to islet autoimmunity and matched controls [Lin et al., 2023]. This study also applied a deconvolution approach to estimate white blood cell proportion and infer regression coefficients for leukocyte subpopulations. Taken together, transcriptomic associations observed in this chapter likely reflect a combination of altered immune cell composition and gene regulation within immune cells.

Finally, although only newborns carrying specific high-risk HLA haplotypes were enrolled in the TEDDY cohort (shown in Table 5.1), HLA haplotypes were not matched between cases and controls in the nested case-control study. While HLA genotyping data are available and could, in principle, be incorporated as covariates in the conditional logistic regression model, the limited sample size and restricted HLA diversity within the study raise concerns about reduced statistical power and potential model instability. Therefore, HLA haplotypes were not included as covariates in the analyses, and residual confounding by HLA genotype is acknowledged as a limitation.

With respect to the nominal association between *LTA* and incident IA (Table 5.3), although *LTA* is located within the extended HLA region, it was independently identified as a putative core protein in Chapter 3 based on GATE scores computed with *trans*-pQTLs outside the HLA region. Furthermore, the association observed in this chapter was based on average circulating *LTA* protein levels measured before the age of one using the Olink Target 96 Inflammation panel. Whereas residual confounding by HLA genotype cannot be fully excluded, the consistency between GATE score association and prospective protein level association may reflect biologically meaningful changes of *LTA* preceding the development of IA.

5.4.2 Prospective datasets for the development of T1D

Putative core genes/proteins are required to be validated in prospective cohorts to establish temporality. Given the limited sample size and statistical power of the TEDDY nested case-control study, I conducted a review to summarise additional prospective datasets for the development of T1D with existing proteomics data or biosamples, which can be used for future proteomics measurements. Brief descriptions of these resources are given below, with information on sample sizes and proteomics data available in Table 5.4.

- The Global Platform for the Prevention of Autoimmune Diabetes (GPPAD) study has screened 559,844 newborns since 2015 across five European countries (Germany, UK, Poland, Belgium, and Sweden), enrolling eligible participants into interventional trials on T1D disease-modifying therapy. Within the Primary Oral Insulin Trial (POInT), proteomic profiling was performed based on dried blood spot samples from 695 individuals with the Olink Explore 3072 panel [Tutino et al., 2025].
- The All Babies in Southeast Sweden (ABIS) study established a population-based birth cohort of 16428 individuals born between 1997 and 1999 in Southeast Sweden. By the end of 2023, 168 participants had developed T1D [Bélteki et al., 2025]. Proteomic profiling of cord blood was conducted using the Olink Explore 384 panel in 1334 individuals, including 132 T1D cases and 1202 controls [McGinnis, 2025].
- Innovative approaches to understanding and arresting type 1 diabetes (INN-ODIA) is a collaborative effort across academia, industry, and patient organisations to “fight type 1 diabetes”. While initially aiming to recruit newly diagnosed T1D patients for clinical trials, it has expanded to unaffected first-degree relatives of individuals with T1D and the general population with positive islet autoantibodies across Europe. By the end of 2023, INNODIA had screened 5300 unaffected first-degree relatives, among whom 361 consented to be intensively followed up. MS-based proteomics data is available for 460 unaffected first-degree relatives.
- The Environmental Determinants of Islet Autoimmunity (ENDIA) study has followed 1500 infants with a first-degree relative with T1D in Australia since 2013. Among them, 58 participants have seroconverted to IA. A protocol for multiomics profiling, including proteomics, has been proposed with a nested case-control design [Oakey et al., 2023]. According to the description of a Breakthrough T1D-funded research project, MS-based proteomics data has been generated from 931 plasma samples from 190 unique mother-infant pairs and is currently undergoing analysis [Penno, 2024].
- The Finnish Type 1 Diabetes Prediction and Prevention (DIPP) project has screened more than 20,000 newborns from the general population in Finland since 1994. To date, over 1100 individuals have developed IA, and more than 500 have progressed to T1D. Mass spectrometry (MS)-based proteomics have been measured in two small nested case-control studies [Moulder et al., 2015, Hirvonen et al., 2023].

- The Type 1 Diabetes TrialNet (TrialNet), which primarily serves as a platform for clinical trials of disease-modifying interventions in T1D, had screened 250,000 individuals with a first- or second-degree relative with T1D worldwide up to the end of 2024. While up-to-date cohort statistics are lacking, an earlier report indicated that by the end of 2016, 5506 individuals had been enrolled, of whom 3207 developed IA and 835 progressed to T1D [Battaglia et al., 2017]. A proteomics investigation based on MS was reported in a nested case-control study with a tiny sample size [Alcazar et al., 2021].
- The Diabetes Autoimmunity Study in the Young (DAISY) cohort recruited 2547 children between 1993 and 2004 in Denver, USA, including both general population participants and first-degree relatives of T1D patients. Proteomic profiling has been carried out in a subset of 172 individuals, focusing on complement system proteins using MS and immunoassays [Webb-Robertson et al., 2023, Hu et al., 2025].

In addition, there are more than a dozen programmes focusing primarily on the early detection and intervention of T1D (Table 5.5) [Sims et al., 2022, Leichter et al., 2025]. Biosamples were collected at recruitment and follow-ups, which might be a valuable resource for future omics investigations.

Table 5.4: Characteristics of prospective datasets for the development of T1D with proteomics data.

Study	Total sample size	Source	Outcome	Proteomics data
GPPAD	559844 screened	General population and first-degree relatives newborns		Olink Explore 3072 panel (N=695, POInT)
ABIS	16428 enrolled	General population newborns	168 T1D (by 2023)	Olink Explore 384 panel (N=1334)
INNODIA	5300 screened	General population and first-degree relatives		Mass spectrometry (N=460)
ENDIA	1500 enrolled	First-degree relatives newborns	58 IA	Mass spectrometry (N=190 mother-infant pairs)
DIPP	>20000 enrolled	General population newborns	>1100 IA, >500 T1D	Mass spectrometry
TrialNet	250000 screened (by 2024), 5506 enrolled (by 2016)	First- or second-degree relatives (age 2-45 years)	3207 IA, 835 T1D (by 2016)	Mass spectrometry
DAISY	2547 enrolled	General population and first-degree relatives (age < 8 years)	>300 IA, ~100 T1D	Mass spectrometry (N=172)

DIPP, Finnish Type 1 Diabetes Prediction and Prevention; TrialNet, Type 1 Diabetes TrialNet; ENDIA, Environmental Determinants of Islet Autoimmunity; DAISY, Diabetes Autoimmunity Study in the Young; GPPAD, Global Platform for the Prevention of Autoimmune Diabetes; POInT, Primary Oral Insulin Trial; ABIS, All Babies in Southeast Sweden; INNODIA, An Innovative approach towards understanding and arresting Type 1 diabetes.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

132

Table 5.5: Characteristics of T1D screening and intervention programmes with available samples for future omics investigations.

Study	Country	Year	Status	Sample size	Source
T1DRA	UK	Since 2023	Recruiting	20000 (target)	General population (age 18-70 years)
ELSA	UK	Since 2022	Recruiting	20000 (target)	General population (age 3-13 years)
ARAD	UK	Since 2022	Recruiting		Individuals with positive autoantibodies from T1DRA and ELSA
BOX family	UK	Since 2015	Recruiting	6000 (screened)	First- or second-degree relatives
Fr1da	Germany	Since 2015	Recruiting	200597 (screened by July 2024)	General population (age 2-10 years)
BYBYDIAB	Germany	1989-2000	Following up	1650 (recruited)	First-degree relative newborns
BABYDIET	Germany	2000-2006	Following up	791 (recruited)	First-degree relative newborns
TEENDIAB	Germany	2009-2015	Following up	610 (recruited)	First-degree relative newborns
DiPiS	Sweden	2000-2004	Following up	38683 (screened)	General population newborns
BABYSCREEN	Finland	2018-2021	Following up	9779 (screened)	General population newborns
EDENT1FI	Europe	Since 2023	Recruiting	200000 (target)	General population (age 1-17 years)
DEW-IT	US	1995-2001, 2008-2012	Following up	32500 (screened)	General population newborns
CASCADE	US	Since 2021	Recruiting	60000 (target)	General population (age < 8 months, or 4-8 years)
ASK	US	Since 2017	Recruiting	33229 (screened by October 2023)	General population (age 1-17 years)
PrIMeD	US	2018-2021	Following up	3818(recruited)	General population (age 2-16 years)
PLEDGE	US	Since 2020	Recruiting	33000 (target)	General population (age < 6, or 9-16 years)
ADIR	Israel	Since 2021	Recruiting	50000 (target)	General population (age 9 months - 5 years)
Type1Screen	Australia	Since 2019	Recruiting	3353 (screened by September 2025)	First-degree relatives (age > 2 years)

T1DRA, Type 1 diabetes risk in adults; ELSA, Early Surveillance for Autoimmune diabetes; ARAD, At Risk of Autoimmune Diabetes; BOX family, Bart's Oxford family; Fr1da, Früherkennung des Typ 1 Diabetes; DiPiS, Diabetes Prediction in Skåne; EDENT1FI, European action for the Diagnosis of Early Non-clinical Type 1 diabetes For disease Interception; DEW-IT, Diabetes Autoimmunity Research in Washington State; CASCADE, Combined Antibody Screening for Celiac and Diabetes Evaluation; ASK, Autoimmunity Screening for Kids; PrIMED, Precision Individualized Medicine for Diabetes; PLEDGE, Sanford Population Level Estimation of T1D Risk GENes in Children; ADIR, Screening for Islet Autoantibodies in the Israeli Pediatric General Population for Detection of Presymptomatic Type 1 Diabetes.

5.5 Supplementary materials

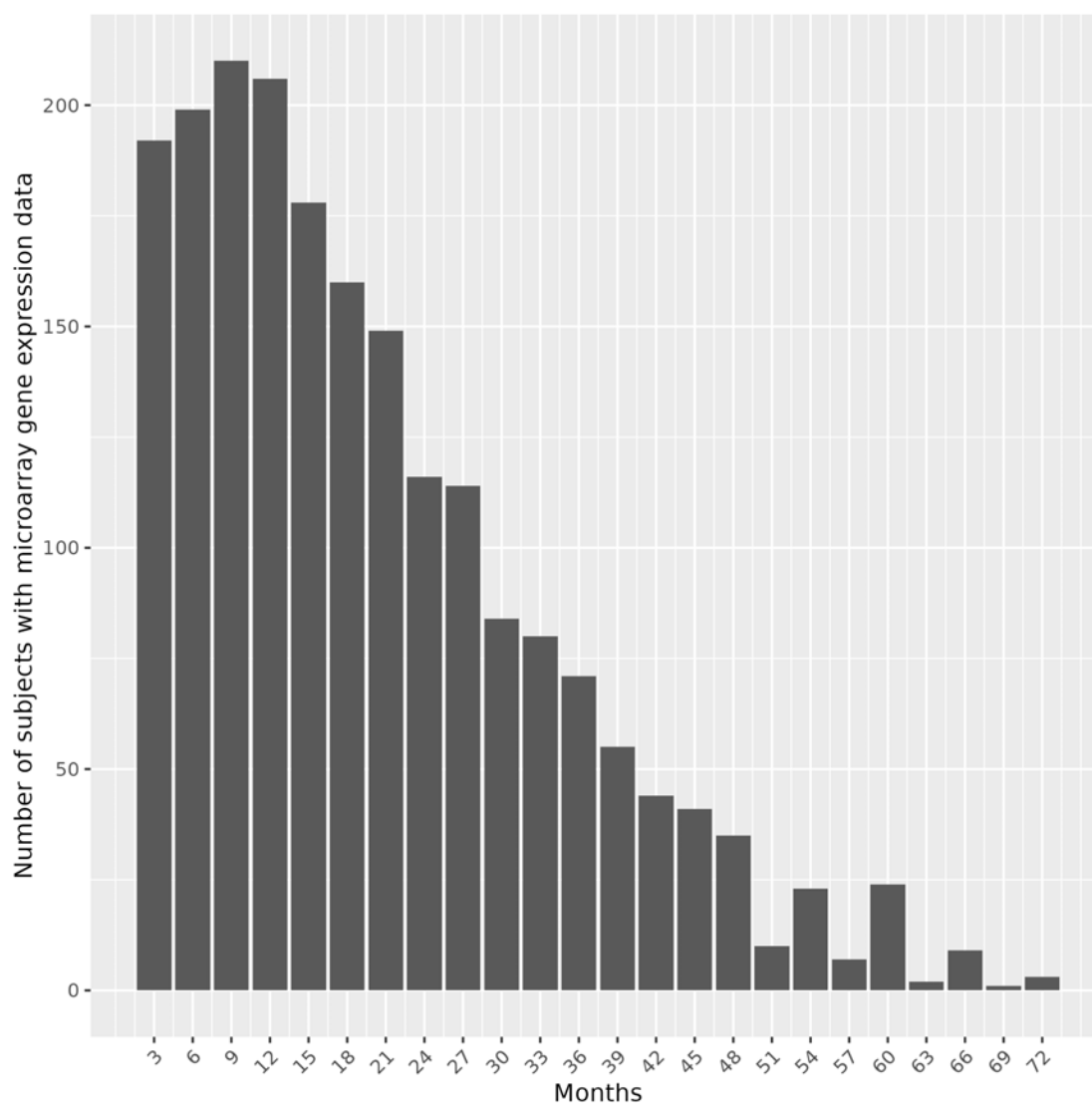


Figure S5.1: The availability of transcripts measured by microarray across time points (3 - 72 months) in 2013 blood samples from 401 subjects.

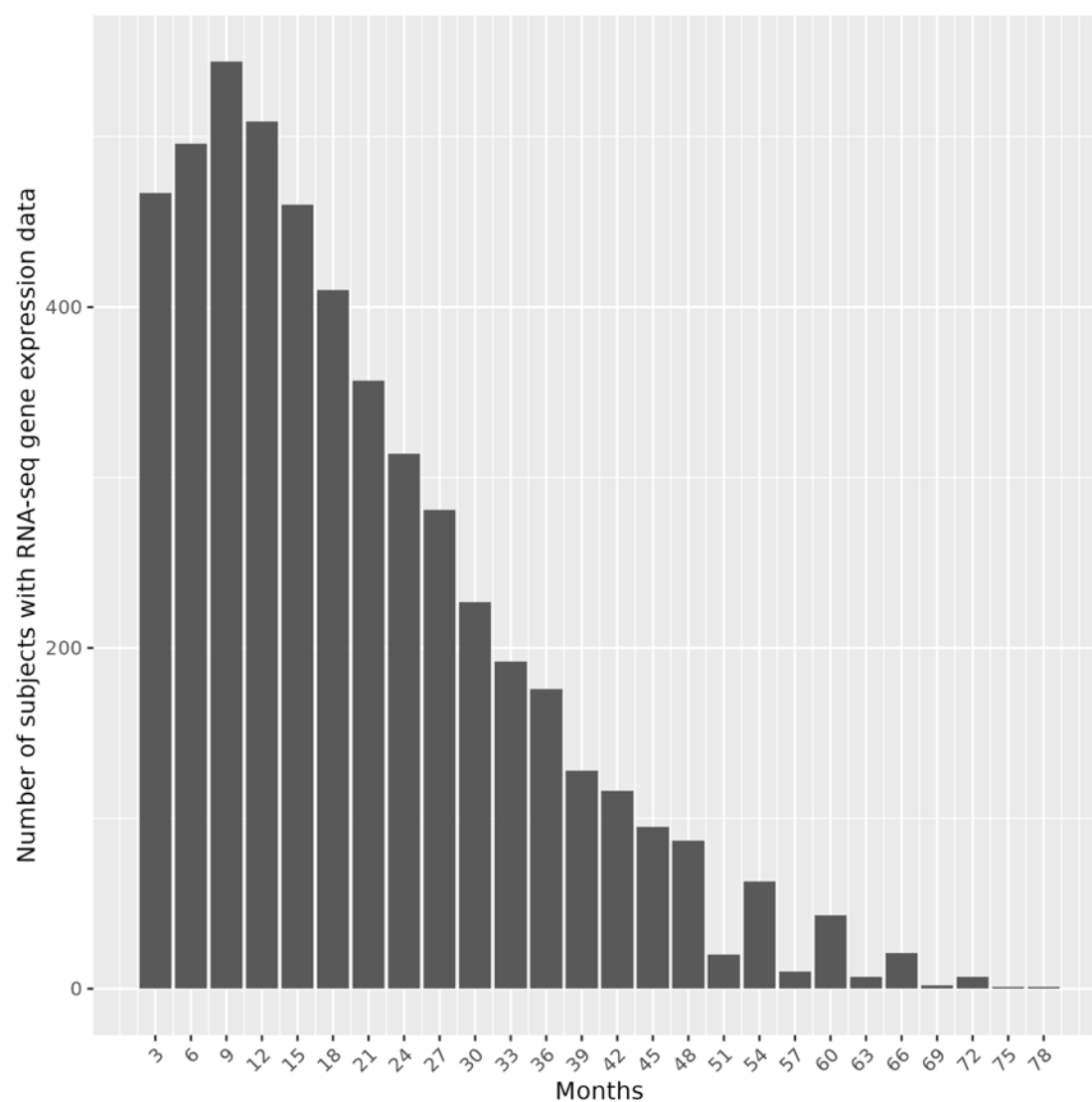


Figure S5.2: The availability of transcripts measured by RNA-seq across time points (3 - 78 months) in 5034 blood samples from 899 subjects.

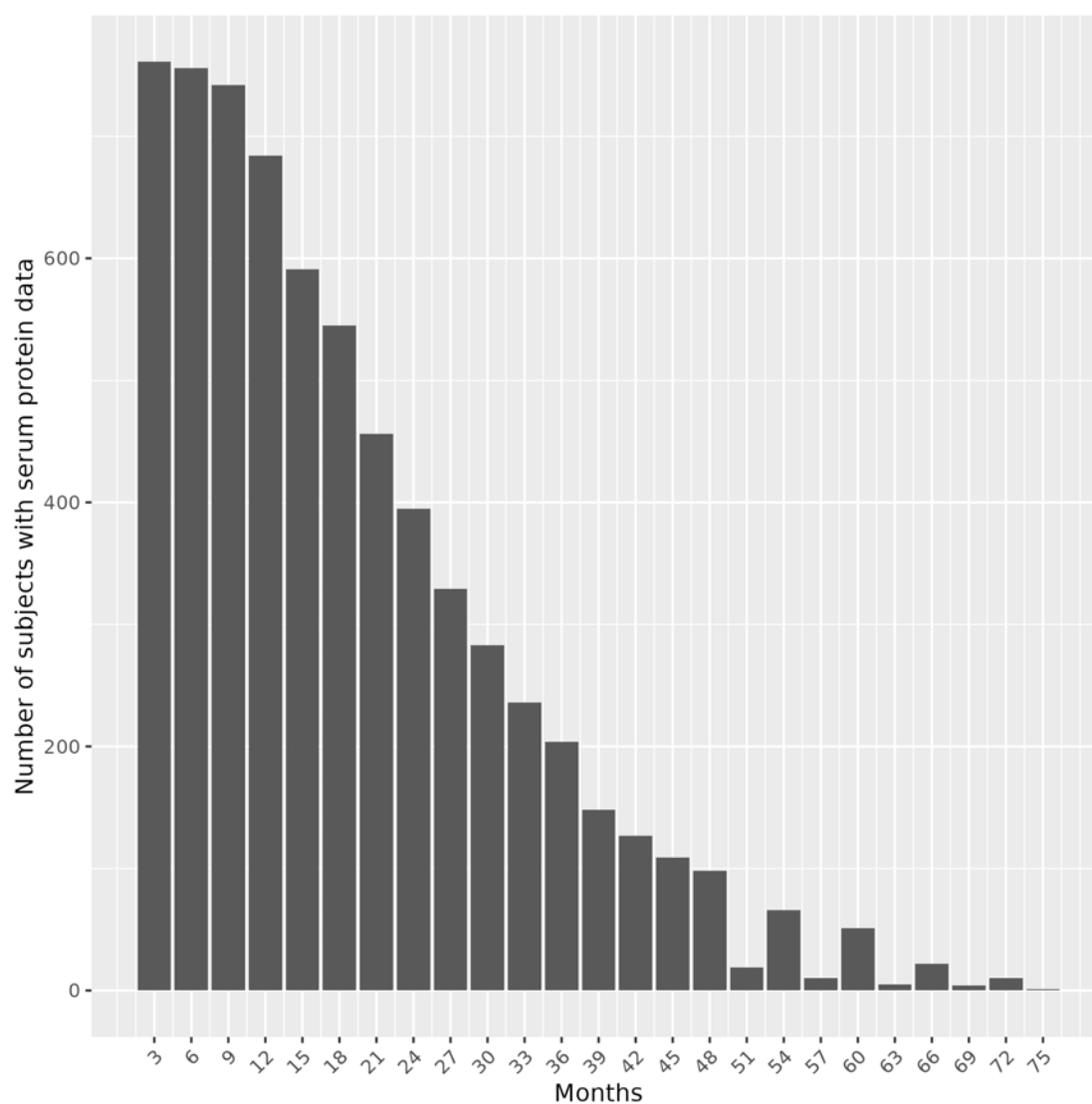


Figure S5.3: The availability of serum protein data measured by Olink Target 96 Inflammation panel across time points (3 - 75 months) in 6652 blood samples from 934 subjects.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

136

Table S5.1: Detailed probe information for the transcripts of genes corresponding to the candidate core proteins and housekeeping genes.

Gene symbol	Probe	mRNA	Protein
Genes corresponding to the candidate core proteins			
ADA2	ILMN_1666174	NM_177405.1	NP_803124.1
ADA2	ILMN_1686283	NM_017424.2	NP_059120.2
ADA2	ILMN_1751851	NM_177405.1	NP_803124.1
ADGRE2	ILMN_2353633	NM_152916.1	NP_690880.1
AGR2	ILMN_1814151	NM_006408.2	NP_006399.1
AGR2	ILMN_3307841	NM_006408.3	NP_006399.1
ALCAM	ILMN_1670870	NM_001627.2	NP_001618.2
AMIGO2	ILMN_1725896	NM_181847.3	NP_862830.1
CCK	ILMN_1712943	NM_000729.3	NP_000720.1
CCL19	ILMN_1769129	NM_006274.2	NP_006265.1
CD274	ILMN_1701914	NM_014143.2	NP_054862.1
CD28	ILMN_1749362	NM_006139.1	NP_006130.1
CD48	ILMN_2061043	NM_001778.2	NP_001769.2
CD5	ILMN_1753112	NM_014207.2	NP_055022.1
CFC1	ILMN_1668099	NM_032545.2	NP_115934.1
CFC1	ILMN_2110094	NM_032545.2	NP_115934.1
CHAD	ILMN_1700652	NM_001267.2	NP_001258.2
CRTAM	ILMN_1796247	NM_019604.2	NP_062550.2
CXCL10	ILMN_1791759	NM_001565.2	NP_001556.2
CXCL11	ILMN_2067890	NM_005409.3	NP_005400.1
CXCL11	ILMN_2067895	NM_005409.3	NP_005400.1
CXCL9	ILMN_1745356	NM_002416.1	NP_002407.1
DDR1	ILMN_1812262	NM_013993.2	NP_054699.2
DDR1	ILMN_2290547	NM_001954.4	NP_001945.3
DDR1	ILMN_2360054	NM_013994.2	NP_054700.2
DLL1	ILMN_1743373	NM_005618.3	NP_005609.3
FASLG	ILMN_1781824	NM_000639.1	NP_000630.1
FCRL6	ILMN_2074762	NM_001004310.1	NP_001004310.1
FGF19	ILMN_1760909	NM_005117.2	NP_005108.1
GCG	ILMN_1724396	NM_002054.2	NP_002045.1
GCNT1	ILMN_1666545	NM_001097635.1	NP_001091104.1
GCNT1	ILMN_2225135	NM_001490.3	NP_001481.2
IDO1	ILMN_3239965	NM_002164.4	NP_002155.1
ITGB2	ILMN_1654396	NM_000211.1	NP_000202.1
ITGB2	ILMN_2175912	NM_000211.2	NP_000202.2
KLRB1	ILMN_2079655	NM_002258.2	NP_002249.1
LAG3	ILMN_1813338	NM_002286.4	NP_002277.3
LGALS9	ILMN_1715760	NM_009587.2	NP_033665.1
LGALS9	ILMN_2412214	NM_009587.2	NP_033665.1
LTA	ILMN_1795464	NM_000595.2	NP_000586.2
MUC2	ILMN_2205622	NM_002457.2	NP_002448.2
NCR1	ILMN_1750761	NM_004829.4	NP_004820.1
NOTCH1	ILMN_1729161	NM_017617.3	NP_060087.3
OXT	ILMN_1735124	NM_000915.2	NP_000906.1
PCDH17	ILMN_1781514	NM_014459.2	NP_055274.2
PDCD1	ILMN_1806725	NM_005018.1	NP_005009.1
PLAT	ILMN_1738742	NM_000930.2	NP_000921.1
PNLIPRP1	ILMN_1775930	NM_006229.2	NP_006220.1
PRSS2	ILMN_1750234	NM_002770.2	NP_002761.1
PRSS27	ILMN_1775549	NM_031948.3	NP_114154.1
REG1B	ILMN_1681462	NM_006507.2	NP_006498.1
ROBO1	ILMN_1666468	NM_002941.2	NP_002932.1

5. Prospective validation of proteins associated with T1D in high-risk individuals.

137

<i>ROBO1</i>	ILMN_1710723	NM_133631.1	NP_598334.1
<i>ROBO1</i>	ILMN_1806790	NM_133631.1	NP_598334.1
<i>SDK2</i>	ILMN_1676709	NM_019064.3	NP_061937.3
<i>SELL</i>	ILMN_1724422	NM_000655.3	NP_000646.1
<i>SEMA7A</i>	ILMN_1756312	NM_003612.2	NP_003603.1
<i>SIGLEC6</i>	ILMN_1685630	NM_198845.1	NP_942142.1
<i>SIGLEC6</i>	ILMN_2299095	NM_198845.2	NP_942142.2
<i>TCN1</i>	ILMN_1768469	NM_001062.3	NP_001053.2
<i>TIGIT</i>	ILMN_2125017	NM_173799.2	NP_776160.1
<i>TIGIT</i>	ILMN_3236289	NM_173799.2	NP_776160.1
<i>TNFRSF4</i>	ILMN_1710204	NM_003327.2	NP_003318.1
<i>TNFRSF4</i>	ILMN_2112256	NM_003327.2	NP_003318.1
Housekeeping genes			
<i>ACTB</i>	ILMN_1777296	NM_001101.2	NP_001092.1
<i>ACTB</i>	ILMN_2152131	NM_001101.2	NP_001092.1
<i>B2M</i>	ILMN_1725427	NM_004048.2	NP_004039.1
<i>B2M</i>	ILMN_2148459	NM_004048.2	NP_004039.1
<i>GAPDH</i>	ILMN_1802252	NM_002046.3	NP_002037.2
<i>GAPDH</i>	ILMN_2038778	NM_002046.3	NP_002037.2
<i>GUSB</i>	ILMN_1669878	NM_000181.2	NP_000172.1
<i>PIIB</i>	ILMN_1703622	NM_000942.4	NP_000933.1
<i>RPL13</i>	ILMN_1709039	NM_033251.1	NP_150254.1
<i>RPL13</i>	ILMN_2413278	NM_033251.1	NP_150254.1
<i>RPLP0</i>	ILMN_1709880	NM_001002.3	NP_000993.1
<i>RPLP0</i>	ILMN_1745075	NM_053275.3	NP_444505.1
<i>RPLP0</i>	ILMN_2402090	NM_001002.3	NP_000993.1
<i>YWHAZ</i>	ILMN_1669286	NM_003406.2	NP_003397.1
<i>YWHAZ</i>	ILMN_1801928	NM_003406.2	NP_003397.1

Probe and the corresponding mRNA and protein information were obtained from the Illumina Human HT-12 Expression BeadChip manifest file. The mRNA and protein information are presented in the format of NCBI Reference Sequences.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

138

Table S5.2: Correlations between multiple probes for the same gene in the TEDDY microarray gene expression data.

Gene symbol	Probe 1	Probe 2	Correlation coefficient
Genes corresponding to the candidate core proteins with multiple probes			
ADA2	ILMN_1666174	ILMN_1751851	0.023
ADA2	ILMN_1686283	ILMN_1666174	-0.050
ADA2	ILMN_1686283	ILMN_1751851	0.307
AGR2	ILMN_3307841	ILMN_1814151	0.128
CFC1	ILMN_2110094	ILMN_1668099	0.043
CXCL11	ILMN_2067890	ILMN_2067895	-0.062
DDR1	ILMN_1812262	ILMN_2290547	0.190
DDR1	ILMN_1812262	ILMN_2360054	-0.133
DDR1	ILMN_2290547	ILMN_2360054	-0.117
GCNT1	ILMN_2225135	ILMN_1666545	0.169
ITGB2	ILMN_1654396	ILMN_2175912	0.965
LGALS9	ILMN_2412214	ILMN_1715760	0.732
ROBO1	ILMN_1666468	ILMN_1710723	0.025
ROBO1	ILMN_1666468	ILMN_1806790	0.000
ROBO1	ILMN_1710723	ILMN_1806790	-0.069
SIGLEC6	ILMN_1685630	ILMN_2299095	0.009
TIGIT	ILMN_2125017	ILMN_3236289	0.124
TNFRSF4	ILMN_1710204	ILMN_2112256	0.168
Housekeeping genes with multiple probes			
ACTB	ILMN_1777296	ILMN_2152131	0.457
B2M	ILMN_1725427	ILMN_2148459	0.515
GAPDH	ILMN_1802252	ILMN_2038778	0.813
RPL13	ILMN_2413278	ILMN_1709039	-0.109
RPLP0	ILMN_1709880	ILMN_1745075	0.653
RPLP0	ILMN_2402090	ILMN_1709880	-0.070
RPLP0	ILMN_2402090	ILMN_1745075	-0.037
YWHAZ	ILMN_1669286	ILMN_1801928	0.508

Correlation coefficients were calculated between probes for the same gene in the same blood samples using pairwise Pearson correlation.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

139

Table S5.3: Correlations between RNA-seq and microarray measurements for the same gene in the TEDDY study.

Gene symbol	Microarray probe	Correlation coefficient
Genes corresponding to the candidate core proteins		
ADA2	ILMN_1686283	0.294
ADA2	ILMN_1666174	0.009
ADA2	ILMN_1751851	0.782
ADGRE2	ILMN_2353633	0.868
AGR2	ILMN_3307841	-0.015
AGR2	ILMN_1814151	0.028
ALCAM	ILMN_1670870	0.630
AMIGO2	ILMN_1725896	0.033
CCK	ILMN_1712943	-0.013
CCL19	ILMN_1769129	0.043
CD274	ILMN_1701914	0.904
CD28	ILMN_1749362	0.637
CD48	ILMN_2061043	0.476
CD5	ILMN_1753112	0.701
CFC1	ILMN_2110094	-0.011
CFC1	ILMN_1668099	0.008
CHAD	ILMN_1700652	0.539
CRTAM	ILMN_1796247	0.167
CXCL10	ILMN_1791759	0.964
CXCL11	ILMN_2067890	0.232
CXCL11	ILMN_2067895	0.026
CXCL9	ILMN_1745356	0.404
DDR1	ILMN_1812262	0.246
DDR1	ILMN_2290547	-0.041
DDR1	ILMN_2360054	-0.024
DLL1	ILMN_1743373	0.738
FASLG	ILMN_1781824	0.884
FCRL6	ILMN_2074762	0.833
FGF19	ILMN_1760909	0.012
GCG	ILMN_1724396	-0.030
GCNT1	ILMN_2225135	0.128
GCNT1	ILMN_1666545	0.664
IDO1	ILMN_3239965	0.960
ITGB2	ILMN_1654396	0.563
ITGB2	ILMN_2175912	0.574
KLRB1	ILMN_2079655	0.784
LAG3	ILMN_1813338	0.899
LGALS9	ILMN_2412214	0.660
LGALS9	ILMN_1715760	0.659
LTA	ILMN_1795464	0.142
MUC2	ILMN_2205622	0.025
NCR1	ILMN_1750761	0.095
NOTCH1	ILMN_1729161	0.614
OXT	ILMN_1735124	0.081
PCDH17	ILMN_1781514	0.077
PDCD1	ILMN_1806725	0.603
PLAT	ILMN_1738742	0.314
PNLIPRP1	ILMN_1775930	0.023
PRSS2	ILMN_1750234	0.019
PRSS27	ILMN_1775549	0.072
REG1B	ILMN_1681462	-0.009
ROBO1	ILMN_1666468	-0.031

5. Prospective validation of proteins associated with T1D in high-risk individuals.

140

<i>ROBO1</i>	ILMN_1710723	-0.029
<i>ROBO1</i>	ILMN_1806790	0.452
<i>SDK2</i>	ILMN_1676709	0.325
<i>SELL</i>	ILMN_1724422	0.580
<i>SEMA7A</i>	ILMN_1756312	0.095
<i>SIGLEC6</i>	ILMN_1685630	-0.065
<i>SIGLEC6</i>	ILMN_2299095	0.360
<i>TCN1</i>	ILMN_1768469	0.914
<i>TIGIT</i>	ILMN_2125017	0.741
<i>TIGIT</i>	ILMN_3236289	0.044
<i>TNFRSF4</i>	ILMN_1710204	0.090
<i>TNFRSF4</i>	ILMN_2112256	0.340
Housekeeping genes		
<i>ACTB</i>	ILMN_1777296	-0.081
<i>ACTB</i>	ILMN_2152131	0.230
<i>B2M</i>	ILMN_1725427	0.174
<i>B2M</i>	ILMN_2148459	0.233
<i>GAPDH</i>	ILMN_1802252	0.302
<i>GAPDH</i>	ILMN_2038778	0.535
<i>GUSB</i>	ILMN_1669878	0.309
<i>PPIB</i>	ILMN_1703622	0.106
<i>RPL13</i>	ILMN_2413278	0.321
<i>RPL13</i>	ILMN_1709039	-0.056
<i>RPLP0</i>	ILMN_2402090	0.039
<i>RPLP0</i>	ILMN_1709880	0.010
<i>RPLP0</i>	ILMN_1745075	0.056
<i>YWHAZ</i>	ILMN_1669286	0.167
<i>YWHAZ</i>	ILMN_1801928	0.240

Correlation coefficients were calculated between RNA-seq and microarray measurements for the same gene in the same blood samples using pairwise Pearson correlation.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

141

Table S5.4: Correlations of microarray measurements between different timepoints in the TEDDY study.

Gene symbol	Probe	Corr (3m, 6m)	Corr (3m, 9m)	Corr (3m, 12m)	Corr (6m, 9m)	Corr (6m, 12m)	Corr (9m, 12m)
Genes corresponding to the candidate core proteins							
ADA2	ILMN_1666174	0.423	0.358	0.378	0.276	0.398	0.167
ADA2	ILMN_1686283	0.341	0.186	0.199	0.278	0.107	0.254
ADA2	ILMN_1751851	0.367	0.332	0.137	0.342	0.023	0.110
ADGRE2	ILMN_2353633	0.430	0.437	0.278	0.426	0.294	0.287
AGR2	ILMN_1814151	0.553	0.254	0.573	0.316	0.588	0.450
AGR2	ILMN_3307841	0.192	-0.058	0.044	0.031	-0.018	0.093
ALCAM	ILMN_1670870	0.432	0.532	0.451	0.419	0.279	0.475
AMIGO2	ILMN_1725896	0.255	0.201	0.249	0.041	0.389	0.290
CCK	ILMN_1712943	0.126	0.289	0.211	0.010	0.060	-0.052
CCL19	ILMN_1769129	0.150	0.375	0.496	0.334	0.454	0.341
CD274	ILMN_1701914	0.104	0.229	0.064	0.003	0.001	0.128
CD28	ILMN_1749362	0.494	0.439	0.404	0.403	0.339	0.313
CD48	ILMN_2061043	0.482	0.484	0.288	0.412	0.112	0.323
CD5	ILMN_1753112	0.305	0.405	0.206	0.436	0.192	0.275
CFC1	ILMN_1668099	0.141	0.300	0.147	0.005	0.164	-0.045
CFC1	ILMN_2110094	-0.022	0.066	0.073	0.005	0.150	-0.039
CHAD	ILMN_1700652	0.201	0.253	0.090	0.419	0.117	0.188
CRTAM	ILMN_1796247	0.610	0.474	0.437	0.590	0.605	0.532
CXCL10	ILMN_1791759	0.022	0.297	-0.105	-0.023	-0.092	-0.074
CXCL11	ILMN_2067890	0.173	0.199	0.111	0.226	0.095	0.190
CXCL11	ILMN_2067895	0.255	-0.049	0.137	0.030	0.236	0.091
CXCL9	ILMN_1745356	0.291	0.098	0.410	0.276	0.445	0.304
DDR1	ILMN_1812262	0.444	0.451	0.398	0.454	0.332	0.199
DDR1	ILMN_2290547	0.308	0.057	-0.019	0.061	0.040	0.238
DDR1	ILMN_2360054	0.495	0.411	0.657	0.437	0.634	0.417
DLL1	ILMN_1743373	0.621	0.601	0.547	0.642	0.517	0.598
FASLG	ILMN_1781824	0.423	0.391	0.341	0.467	0.271	0.369
FCRL6	ILMN_2074762	0.527	0.573	0.584	0.724	0.625	0.590
FGF19	ILMN_1760909	0.310	0.284	0.294	0.383	0.355	0.158
GCG	ILMN_1724396	0.093	0.084	0.001	0.177	0.095	0.141
GCNT1	ILMN_1666545	0.464	0.490	0.372	0.455	0.256	0.338
GCNT1	ILMN_2225135	0.557	0.503	0.466	0.413	0.502	0.542
IDO1	ILMN_3239965	0.169	0.216	0.026	0.243	-0.029	0.010
ITGB2	ILMN_1654396	0.533	0.504	0.411	0.487	0.160	0.320
ITGB2	ILMN_2175912	0.486	0.500	0.462	0.474	0.188	0.343
KLRB1	ILMN_2079655	0.610	0.361	0.269	0.462	0.238	0.527
LAG3	ILMN_1813338	0.532	0.235	0.125	0.472	0.179	0.330
LGALS9	ILMN_1715760	0.146	0.176	-0.047	0.341	0.096	0.144
LGALS9	ILMN_2412214	0.078	0.305	0.017	0.333	0.115	0.128
LTA	ILMN_1795464	0.256	0.279	0.296	0.434	0.442	0.377
MUC2	ILMN_2205622	0.111	0.126	0.264	0.163	0.263	0.117
NCR1	ILMN_1750761	0.270	0.071	0.057	0.137	0.183	0.168
NOTCH1	ILMN_1729161	0.485	0.264	0.285	0.229	0.293	0.271
OXT	ILMN_1735124	-0.007	0.079	-0.023	0.286	0.068	0.092
PCDH17	ILMN_1781514	0.159	0.314	0.267	0.346	0.537	0.405
PDCD1	ILMN_1806725	0.349	0.246	0.065	0.439	0.292	0.152
PLAT	ILMN_1738742	0.070	-0.029	0.127	0.346	0.372	0.515
PNLIPRP1	ILMN_1775930	0.164	0.244	0.396	0.431	0.347	0.322
PRSS2	ILMN_1750234	0.182	0.095	-0.008	0.208	-0.052	0.110
PRSS27	ILMN_1775549	0.263	0.121	0.084	0.165	-0.198	0.198
REG1B	ILMN_1681462	0.076	0.087	0.087	0.094	0.188	0.035

5. Prospective validation of proteins associated with T1D in high-risk individuals.

142

<i>ROBO1</i>	ILMN_1666468	-0.029	0.009	0.177	0.142	-0.040	-0.108
<i>ROBO1</i>	ILMN_1710723	0.101	0.096	0.060	0.199	0.044	0.081
<i>ROBO1</i>	ILMN_1806790	0.454	0.366	0.545	0.407	0.427	0.418
<i>SDK2</i>	ILMN_1676709	0.401	0.443	0.321	0.472	0.218	0.441
<i>SELL</i>	ILMN_1724422	0.319	0.412	0.351	0.343	0.248	0.250
<i>SEMA7A</i>	ILMN_1756312	0.285	0.114	0.263	0.109	-0.043	0.173
<i>SIGLEC6</i>	ILMN_1685630	0.266	0.414	0.484	0.410	0.333	0.182
<i>SIGLEC6</i>	ILMN_2299095	0.315	0.195	0.247	0.177	0.204	0.361
<i>TCN1</i>	ILMN_1768469	0.456	0.377	0.257	0.525	0.107	0.187
<i>TIGIT</i>	ILMN_2125017	0.373	0.108	0.359	0.387	0.330	0.263
<i>TIGIT</i>	ILMN_3236289	0.324	0.134	0.320	0.190	0.231	0.188
<i>TNFRSF4</i>	ILMN_1710204	0.030	0.026	0.074	0.042	0.016	0.142
<i>TNFRSF4</i>	ILMN_2112256	0.428	0.412	0.384	0.418	0.467	0.399
Housekeeping genes							
<i>ACTB</i>	ILMN_1777296	0.407	0.282	0.267	0.296	0.199	0.481
<i>ACTB</i>	ILMN_2152131	0.315	0.332	0.260	0.387	0.246	0.279
<i>B2M</i>	ILMN_1725427	0.446	0.298	0.124	0.463	0.121	0.197
<i>B2M</i>	ILMN_2148459	0.458	0.529	0.208	0.505	0.392	0.255
<i>GAPDH</i>	ILMN_1802252	0.328	0.203	0.164	0.402	0.149	0.431
<i>GAPDH</i>	ILMN_2038778	0.408	0.171	0.318	0.515	0.227	0.353
<i>GUSB</i>	ILMN_1669878	0.269	0.441	0.172	0.507	0.313	0.404
<i>PPIB</i>	ILMN_1703622	0.449	0.404	0.352	0.526	0.600	0.417
<i>RPL13</i>	ILMN_1709039	0.174	0.042	0.116	0.128	0.204	0.235
<i>RPL13</i>	ILMN_2413278	0.389	0.469	0.457	0.361	0.535	0.350
<i>RPLP0</i>	ILMN_1709880	0.260	0.163	0.196	0.340	0.334	0.288
<i>RPLP0</i>	ILMN_1745075	0.389	0.263	0.092	0.357	0.168	0.175
<i>RPLP0</i>	ILMN_2402090	0.259	-0.022	0.139	0.051	0.044	0.093
<i>YWHAZ</i>	ILMN_1669286	0.246	0.237	0.232	0.295	0.028	0.367
<i>YWHAZ</i>	ILMN_1801928	0.446	0.350	0.411	0.391	0.255	0.496

Correlation coefficients were calculated between microarray measurements for the same probe in the blood samples from the same individual at different timepoints using pairwise Pearson correlation.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

143

Table S5.5: Correlations of RNA-seq measurements between different timepoints in the TEDDY study.

Gene symbol	Corr (3m, 6m)	Corr (3m, 9m)	Corr (3m, 12m)	Corr (6m, 9m)	Corr (6m, 12m)	Corr (9m, 12m)
Genes corresponding to the candidate core proteins						
ADA2	0.252	0.260	0.141	0.156	0.077	0.180
ADGRE2	0.238	0.349	0.293	0.291	0.098	0.235
AGR2	-0.017	-0.052	-0.058	0.005	-0.033	0.042
ALCAM	0.434	0.517	0.426	0.489	0.471	0.454
AMIGO2	0.501	0.535	0.466	0.523	0.436	0.506
CCK	-0.005	-0.005	-0.005	-0.003	.	-0.004
CCL19	-0.005	0.081	0.002	0.094	0.020	0.109
CD274	0.014	-0.029	-0.035	0.087	-0.014	0.054
CD28	0.312	0.352	0.264	0.347	0.260	0.356
CD48	0.242	0.231	0.077	0.263	0.131	0.260
CD5	0.464	0.460	0.355	0.491	0.328	0.397
CFC1	-0.005	-0.005	-0.010	-0.005	-0.009	-0.011
CHAD	0.421	0.339	0.219	0.310	0.243	0.273
CRTAM	0.536	0.544	0.435	0.487	0.341	0.523
CXCL10	0.002	-0.008	-0.030	-0.005	0.078	0.003
CXCL11	0.062	-0.018	-0.041	0.001	0.025	0.031
CXCL9	0.261	0.073	0.111	0.079	0.025	0.158
DDR1	0.412	0.285	0.253	0.346	0.163	0.331
DLL1	0.408	0.475	0.383	0.431	0.394	0.467
FASLG	0.432	0.266	0.265	0.356	0.376	0.426
FCRL6	0.582	0.440	0.474	0.530	0.574	0.516
FGF19	-0.005	-0.006	-0.006	-0.005	-0.005	0.073
GCG	0.065	-0.073	-0.025	-0.029	0.094	0.061
GCNT1	0.549	0.438	0.436	0.496	0.373	0.354
IDO1	0.099	0.022	0.090	0.208	0.036	0.026
ITGB2	0.370	0.414	0.330	0.364	0.155	0.280
KLRB1	0.371	0.353	0.412	0.422	0.352	0.443
LAG3	0.342	0.131	0.095	0.205	0.181	0.120
LGALS9	0.112	0.093	0.015	0.180	0.070	0.190
LTA	0.808	0.791	0.755	0.759	0.724	0.772
MUC2	0.004	0.095	-0.001	0.176	0.065	0.011
NCR1	0.298	0.274	0.295	0.302	0.194	0.341
NOTCH1	0.309	0.250	0.343	0.341	0.189	0.300
OXT	0.064	0.150	0.062	0.099	0.078	0.091
PCDH17	0.338	0.314	0.304	0.735	0.628	0.622
PDCD1	0.437	0.184	0.213	0.328	0.318	0.359
PLAT	0.030	0.058	0.067	0.876	0.950	0.874
PNLIPRP1	.	.	.	-0.003	.	.
PRSS2	0.475	0.366	0.457	0.358	0.365	0.340
PRSS27	0.159	0.124	0.175	0.168	0.095	0.069
REG1B	0.045	0.106	0.006	0.110	0.064	0.104
ROBO1	0.546	0.525	0.604	0.605	0.435	0.628
SDK2	0.697	0.667	0.663	0.711	0.612	0.661
SELL	0.155	0.129	0.187	0.179	0.042	0.107
SEMA7A	0.495	0.407	0.319	0.425	0.284	0.444
SIGLEC6	0.532	0.505	0.514	0.487	0.471	0.481
TCN1	0.495	0.419	0.319	0.419	0.102	0.334
TIGIT	0.522	0.260	0.308	0.475	0.475	0.436
TNFRSF4	0.187	0.190	0.162	0.186	0.145	0.251
Housekeeping genes						
ACTB	0.375	0.288	0.251	0.338	0.226	0.357
B2M	0.491	0.409	0.333	0.450	0.433	0.307
GAPDH	0.246	0.157	0.149	0.254	0.154	0.252
GUSB	0.235	0.268	0.149	0.337	0.268	0.287
PPIB	0.320	0.242	0.143	0.339	0.188	0.248
RPL13	0.255	0.202	0.292	0.245	0.194	0.320
RPLP0	0.237	0.187	0.305	0.239	0.223	0.248
YWHAZ	0.338	0.408	0.294	0.365	0.258	0.221

Correlation coefficients were calculated between RNA-seq measurements for the same gene in the blood samples from the same individual at different timepoints using pairwise Pearson correlation.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

144

Table S5.6: Correlations of protein measurements between different timepoints in the TEDDY study.

Gene symbol	Corr (3m, 6m)	Corr (3m, 9m)	Corr (3m, 12m)	Corr (6m, 9m)	Corr (6m, 12m)	Corr (9m, 12m)
<i>CCL19</i>	0.565	0.389	0.288	0.537	0.411	0.451
<i>CD274</i>	0.313	0.227	0.148	0.315	0.229	0.331
<i>CD5</i>	0.570	0.492	0.456	0.521	0.538	0.533
<i>CXCL10</i>	0.357	0.269	0.205	0.314	0.224	0.304
<i>CXCL11</i>	0.359	0.181	0.104	0.214	0.209	0.270
<i>CXCL9</i>	0.379	0.253	0.149	0.356	0.242	0.276
<i>FGF19</i>	0.394	0.129	0.083	0.263	0.200	0.231
<i>LTA</i>	0.600	0.561	0.473	0.567	0.519	0.537

Correlation coefficients were calculated between measurements for the same protein in the blood samples from the same individual at different timepoints using pairwise Pearson correlation.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

145

Table S5.7: Associations of average transcripts levels measured by microarray before the age of one with the risk of incident IA.

Gene symbol	Probe	Overall (104 pairs)		Year 1 (50 pairs)		Year 2 (25 pairs)	
		Log cOR	p-value	Log cOR	p-value	Log cOR	p-value
ADA2	ILMN_1666174	0.05	0.8	-0.01	1	-0.08	0.8
ADA2	ILMN_1686283	-0.30	0.07	-0.22	0.3	-0.01	1
ADA2	ILMN_1751851	-0.15	0.3	-0.18	0.4	-0.03	0.9
ADGRE2	ILMN_2353633	0.22	0.1	0.11	0.6	0.46	0.1
AGR2	ILMN_1814151	-0.08	0.7	0.08	0.8	-0.22	0.6
AGR2	ILMN_3307841	0.01	1	-0.10	0.7	-0.73	0.1
ALCAM	ILMN_1670870	-0.02	0.9	-0.39	0.1	0.08	0.8
AMIGO2	ILMN_1725896	-0.08	0.6	-0.26	0.3	0.11	0.7
CCK	ILMN_1712943	0.10	0.5	0.04	0.8	0.39	0.3
CCL19	ILMN_1769129	-0.32	0.1	-0.05	0.9	-2.41	0.01
CD274	ILMN_1701914	0.10	0.5	0.19	0.4	0.34	0.2
CD28	ILMN_1749362	-0.02	0.9	0.07	0.8	-0.14	0.6
CD48	ILMN_2061043	-0.50	0.02	-0.34	0.3	-0.51	0.2
CD5	ILMN_1753112	-0.26	0.1	-0.25	0.3	-0.34	0.3
CFC1	ILMN_1668099	-0.29	0.08	-0.14	0.5	-0.84	0.08
CFC1	ILMN_2110094	0.09	0.5	-0.01	1	-0.07	0.8
CHAD	ILMN_1700652	0.00	1	-0.11	0.6	-0.03	0.9
CRTAM	ILMN_1796247	-0.11	0.5	-0.27	0.3	0.69	0.1
CXCL10	ILMN_1791759	-0.15	0.3	0.10	0.6	-0.04	0.9
CXCL11	ILMN_2067890	-0.09	0.5	-0.30	0.2	-0.07	0.8
CXCL11	ILMN_2067895	0.08	0.6	0.08	0.7	0.64	0.1
CXCL9	ILMN_1745356	-0.19	0.3	0.08	0.8	-0.54	0.1
DDR1	ILMN_1812262	0.18	0.4	0.06	0.8	0.78	0.3
DDR1	ILMN_2290547	-0.09	0.5	-0.09	0.7	0.13	0.7
DDR1	ILMN_2360054	-0.06	0.8	0.18	0.6	0.73	0.2
DLL1	ILMN_1743373	-0.02	0.9	-0.08	0.7	0.01	1
FASLG	ILMN_1781824	-0.07	0.6	-0.16	0.4	-0.25	0.4
FCRL6	ILMN_2074762	-0.04	0.8	-0.24	0.3	-0.01	1
FGF19	ILMN_1760909	-0.25	0.2	-0.12	0.6	-0.89	0.05
GCG	ILMN_1724396	0.17	0.3	0.16	0.4	-0.14	0.7
GCNT1	ILMN_1666545	-0.12	0.4	-0.26	0.3	-0.08	0.8
GCNT1	ILMN_2225135	0.15	0.6	-0.20	0.6	0.99	0.2
IDO1	ILMN_3239965	-0.02	0.9	-0.09	0.8	0.24	0.4
ITGB2	ILMN_1654396	0.06	0.7	-0.17	0.4	0.12	0.7
ITGB2	ILMN_2175912	0.04	0.8	-0.14	0.5	0.05	0.9
KLRB1	ILMN_2079655	-0.02	0.9	-0.10	0.6	-0.60	0.2
LAG3	ILMN_1813338	-0.26	0.1	-0.18	0.4	-0.35	0.4
LGALS9	ILMN_1715760	-0.10	0.6	-0.12	0.6	0.05	0.9
LGALS9	ILMN_2412214	-0.14	0.4	-0.11	0.7	-0.12	0.7
LTA	ILMN_1795464	-0.11	0.5	-0.15	0.5	-0.16	0.5
MUC2	ILMN_2205622	0.13	0.4	0.48	0.1	-0.13	0.6
NCR1	ILMN_1750761	0.19	0.3	0.15	0.5	0.29	0.4
NOTCH1	ILMN_1729161	0.24	0.1	0.16	0.4	0.30	0.3
OXT	ILMN_1735124	-0.17	0.2	-0.03	0.9	-0.72	0.08
PCDH17	ILMN_1781514	-0.18	0.4	0.23	0.6	-0.30	0.4
PDCD1	ILMN_1806725	-0.53	0.006	-0.32	0.2	-0.67	0.1
PLAT	ILMN_1738742	0.04	0.8	-0.14	0.5	0.28	0.4
PNLIPRP1	ILMN_1775930	-0.13	0.5	0.27	0.4	-0.31	0.5
PRSS2	ILMN_1750234	0.14	0.4	-0.01	1	-0.18	0.5
PRSS27	ILMN_1775549	0.22	0.1	-0.20	0.3	0.93	0.03
REG1B	ILMN_1681462	0.08	0.6	-0.55	0.06	0.57	0.2
ROBO1	ILMN_1666468	0.11	0.5	0.09	0.7	0.23	0.5
ROBO1	ILMN_1710723	-0.03	0.9	0.10	0.7	0.19	0.6
ROBO1	ILMN_1806790	0.15	0.4	0.12	0.6	0.43	0.3
SDK2	ILMN_1676709	0.05	0.7	0.15	0.6	0.29	0.4
SELL	ILMN_1724422	0.01	0.9	0.06	0.8	-0.32	0.4
SEMA7A	ILMN_1756312	-0.08	0.6	0.17	0.5	-0.41	0.2
SIGLEC6	ILMN_1685630	0.24	0.2	0.26	0.4	0.48	0.3
SIGLEC6	ILMN_2299095	0.15	0.4	-0.01	1	0.04	0.9
TCN1	ILMN_1768469	0.00	1	0.03	0.9	-0.08	0.8
TIGIT	ILMN_2125017	-0.07	0.6	-0.18	0.4	0.10	0.7
TIGIT	ILMN_3236289	-0.07	0.6	0.27	0.3	0.01	1
TNFRSF4	ILMN_1710204	0.01	0.9	0.17	0.5	0.01	1
TNFRSF4	ILMN_2112256	-0.09	0.6	-0.23	0.3	-0.09	0.8

Log cOR: log (conditional odds ratio) of incident IA risk per standard deviation increase in the average transcripts levels measured by microarray before the age of one.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

146

Table S5.8: Associations of average transcripts levels measured by microarray before the age of one with the risk of incident T1D.

Gene symbol	Probe	Overall (54 pairs)	
		Log cOR	p-value
ADA2	ILMN_1666174	-0.18	0.6
ADA2	ILMN_1686283	-0.15	0.5
ADA2	ILMN_1751851	0.07	0.7
ADGRE2	ILMN_2353633	-0.01	0.9
AGR2	ILMN_1814151	-0.80	0.03
AGR2	ILMN_3307841	0.21	0.3
ALCAM	ILMN_1670870	0.07	0.8
AMIGO2	ILMN_1725896	-0.46	0.09
CCK	ILMN_1712943	0.09	0.7
CCL19	ILMN_1769129	0.02	0.9
CD274	ILMN_1701914	-0.12	0.5
CD28	ILMN_1749362	-0.07	0.8
CD48	ILMN_2061043	-0.28	0.4
CD5	ILMN_1753112	0.05	0.8
CFC1	ILMN_1668099	0.17	0.4
CFC1	ILMN_2110094	0.28	0.3
CHAD	ILMN_1700652	-0.24	0.3
CRTAM	ILMN_1796247	0.05	0.8
CXCL10	ILMN_1791759	-0.11	0.6
CXCL11	ILMN_2067890	-0.43	0.1
CXCL11	ILMN_2067895	-0.13	0.5
CXCL9	ILMN_1745356	-0.27	0.2
DDR1	ILMN_1812262	0.15	0.5
DDR1	ILMN_2290547	-0.22	0.3
DDR1	ILMN_2360054	0.24	0.5
DLL1	ILMN_1743373	0.01	1
FASLG	ILMN_1781824	-0.18	0.4
FCRL6	ILMN_2074762	0.08	0.7
FGF19	ILMN_1760909	-0.32	0.2
GCG	ILMN_1724396	0.31	0.2
GCNT1	ILMN_1666545	0.34	0.1
GCNT1	ILMN_2225135	-0.02	0.9
IDO1	ILMN_3239965	-0.01	1
ITGB2	ILMN_1654396	-0.33	0.1
ITGB2	ILMN_2175912	-0.35	0.1
KLRB1	ILMN_2079655	-0.18	0.3
LAG3	ILMN_1813338	-0.36	0.2
LGALS9	ILMN_1715760	-0.04	0.9
LGALS9	ILMN_2412214	-0.11	0.6
LTA	ILMN_1795464	-0.23	0.4
MUC2	ILMN_2205622	-0.17	0.4
NCR1	ILMN_1750761	0.37	0.1
NOTCH1	ILMN_1729161	-0.20	0.4
OXT	ILMN_1735124	-0.10	0.6
PCDH17	ILMN_1781514	-0.82	0.02
PDCD1	ILMN_1806725	-0.04	0.8
PLAT	ILMN_1738742	0.16	0.4
PNLIPRP1	ILMN_1775930	-0.27	0.3
PRSS2	ILMN_1750234	0.01	1
PRSS27	ILMN_1775549	-0.36	0.1
REG1B	ILMN_1681462	0.15	0.4
ROBO1	ILMN_1666468	-0.32	0.2
ROBO1	ILMN_1710723	0.12	0.6
ROBO1	ILMN_1806790	0.42	0.09
SDK2	ILMN_1676709	0.15	0.5
SELL	ILMN_1724422	-0.18	0.5
SEMA7A	ILMN_1756312	0.04	0.8
SIGLEC6	ILMN_1685630	0.01	1
SIGLEC6	ILMN_2299095	0.10	0.6
TCN1	ILMN_1768469	-0.18	0.4
TIGIT	ILMN_2125017	-0.02	0.9
TIGIT	ILMN_3236289	0.15	0.5
TNFRSF4	ILMN_1710204	0.16	0.5
TNFRSF4	ILMN_2112256	-0.24	0.3

Log cOR: log (conditional odds ratio) of incident T1D risk per standard deviation increase in the average transcripts levels measured by microarray before the age of one.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

147

Table S5.9: Associations of average transcripts levels measured by RNA-seq before the age of one with the risk of incident IA.

Gene symbol	Overall (274 pairs)		Year 1 (123 pairs)		Year 2 (85 pairs)		Year 3 (40 pairs)	
	Log cOR	p-value	Log cOR	p-value	Log cOR	p-value	Log cOR	p-value
ADA2	-0.05	0.6	-0.13	0.3	-0.08	0.6	-0.01	1
ADGRE2	-0.05	0.6	-0.09	0.5	-0.26	0.1	0.26	0.2
AGR2	-0.01	0.9	0.03	0.8	-0.08	0.6	-0.11	0.6
ALCAM	-0.14	0.2	-0.22	0.1	-0.36	0.07	0.10	0.6
AMIGO2	-0.05	0.6	-0.08	0.6	-0.06	0.7	-0.14	0.5
CCK	-0.11	0.3	0.01	1	-0.67	0.3	.	.
CCL19	-0.09	0.3	-0.06	0.7	-0.02	0.9	-0.08	0.8
CD274	0.06	0.5	0.12	0.4	-0.08	0.6	0.16	0.6
CD28	-0.18	0.05	-0.07	0.6	-0.50	0.006	-0.20	0.4
CD48	0.14	0.1	0.12	0.3	0.49	0.009	-0.18	0.5
CD5	0.07	0.4	0.02	0.9	0.20	0.3	0.05	0.8
CFC1	0.01	0.9	0.05	0.7	-0.01	1	-2.04	1
CHAD	0.12	0.2	0.10	0.5	0.37	0.06	-0.21	0.4
CRTAM	-0.03	0.7	-0.05	0.7	-0.01	1	-0.06	0.8
CXCL10	-0.06	0.5	0.07	0.6	-0.14	0.4	-0.37	0.4
CXCL11	-0.06	0.5	0.12	0.4	-0.03	0.8	-0.75	0.2
CXCL9	0.05	0.6	0.11	0.4	0.18	0.2	-0.07	0.8
DDR1	-0.24	0.01	-0.23	0.1	-0.65	0.003	-0.15	0.5
DLL1	-0.11	0.2	0.00	1	-0.40	0.02	0.01	1
FASLG	0.00	1	0.06	0.6	-0.09	0.6	-0.06	0.8
FCRL6	-0.07	0.4	-0.12	0.3	-0.01	1	0.04	0.9
FGF19	0.02	0.8	0.04	0.8	-0.04	0.8	.	.
GCG	0.06	0.5	0.07	0.6	0.02	0.9	0.03	0.9
GCNT1	-0.09	0.3	-0.08	0.5	-0.46	0.02	0.05	0.8
IDO1	0.01	0.9	0.01	1	0.00	1	0.01	1
ITGB2	0.13	0.1	0.05	0.7	0.35	0.05	0.15	0.5
KLRB1	0.08	0.4	0.08	0.5	0.36	0.06	-0.04	0.9
LAG3	-0.13	0.4	-0.02	0.9	0.26	0.2	-4.67	0.09
LGALS9	0.13	0.1	0.21	0.1	0.29	0.08	-0.29	0.3
LTA	-0.08	0.4	-0.01	0.9	-0.18	0.3	0.26	0.3
MUC2	-0.05	0.6	-0.14	0.3	-0.04	0.8	0.02	0.9
NCR1	-0.08	0.3	-0.09	0.5	-0.57	0.007	0.44	0.09
NOTCH1	-0.03	0.8	-0.05	0.7	-0.17	0.4	0.23	0.3
OXT	0.25	0.03	0.33	0.02	0.48	0.1	0.56	0.08
PCDH17	0.07	0.4	-0.07	0.6	0.12	0.5	0.26	0.2
PDCD1	-0.17	0.09	-0.13	0.4	-0.15	0.4	-0.28	0.3
PLAT	-0.11	0.5	-0.24	0.5	0.11	0.5	-0.14	0.6
PRSS2	0.14	0.1	0.12	0.3	0.33	0.07	0.11	0.6
PRSS27	0.07	0.5	0.08	0.6	0.15	0.4	0.04	0.9
REG1B	0.06	0.5	0.18	0.2	0.00	1	-0.17	0.5
ROBO1	-0.07	0.4	-0.14	0.3	-0.29	0.1	0.34	0.1
SDK2	-0.13	0.1	-0.26	0.06	-0.05	0.8	-0.02	0.9
SELL	0.23	0.02	0.16	0.2	0.31	0.08	0.22	0.4
SEMA7A	-0.13	0.1	-0.23	0.08	0.00	1	0.01	1
SIGLEC6	-0.49	3×10^{-5}	-0.42	0.01	-0.90	6×10^{-4}	-0.08	0.7
TCN1	0.04	0.7	-0.06	0.7	0.14	0.4	0.02	0.9
TIGIT	-0.09	0.3	-0.10	0.4	-0.16	0.3	0.11	0.7
TNFRSF4	0.30	0.007	0.28	0.05	0.99	0.005	0.09	0.7

Log cOR: log (conditional odds ratio) of incident IA risk per standard deviation increase in the average transcripts levels measured by RNA-seq before the age of one.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

148

Table S5.10: Associations of average transcripts levels measured by RNA-seq before the age of one with the risk of incident T1D.

Gene symbol	Overall (97 pairs)		Year 1 (37 pairs)		Year 2 (26 pairs)	
	Log cOR	p-value	Log cOR	p-value	Log cOR	p-value
ADA2	-0.11	0.5	-0.38	0.2	-0.10	0.7
ADGRE2	-0.03	0.9	-0.05	0.8	-0.10	0.7
AGR2	0.05	0.7	0.00	1	0.06	0.8
ALCAM	0.07	0.7	0.32	0.3	0.08	0.8
AMIGO2	0.01	0.9	0.02	0.9	0.18	0.6
CCK	-1.24	1	.	.	.	.
CCL19	0.05	0.7	0.18	0.4	-0.12	0.7
CD274	-0.19	0.3	-0.47	0.2	0.01	1
CD28	0.04	0.8	-0.03	0.9	0.05	0.9
CD48	-0.24	0.1	-0.20	0.5	0.03	0.9
CD5	0.02	0.9	-0.17	0.6	-0.03	0.9
CFC1	-0.03	0.8	0.09	0.7	0.10	0.7
CHAD	-0.04	0.8	-0.06	0.8	-0.26	0.4
CRTAM	0.13	0.4	-0.02	0.9	0.40	0.2
CXCL10	-0.26	0.2	-0.31	0.4	0.18	0.6
CXCL11	-0.37	0.2	-0.80	0.2	0.25	0.5
CXCL9	0.04	0.8	-0.18	0.5	0.86	0.1
DDR1	0.13	0.5	0.13	0.7	0.28	0.4
DLL1	0.11	0.5	0.12	0.6	0.23	0.4
FASLG	0.11	0.5	0.13	0.6	0.50	0.2
FCRL6	0.03	0.8	0.17	0.5	0.07	0.8
FGF19	-2.73	1	.	.	-2.66	1
GCG	0.05	0.8	-0.28	0.3	1.07	0.09
GCNT1	0.17	0.3	0.28	0.3	0.15	0.6
IDO1	-0.24	0.3	-0.62	0.3	0.55	0.3
ITGB2	-0.23	0.2	-0.58	0.07	0.17	0.6
KLRB1	-0.06	0.7	-0.08	0.7	0.11	0.7
LAG3	-0.15	0.3	-0.26	0.3	0.03	0.9
LGALS9	0.01	0.9	-0.13	0.5	0.17	0.6
LTA	-0.11	0.5	-0.73	0.04	-0.16	0.6
MUC2	0.03	0.9	0.05	0.8	0.26	0.5
NCR1	0.02	0.9	0.07	0.8	-0.02	0.9
NOTCH1	0.02	0.9	-0.08	0.8	-0.14	0.7
OXT	0.13	0.4	-0.06	0.8	0.31	0.4
PCDH17	-0.25	0.1	-0.18	0.5	-0.42	0.3
PDCD1	-0.14	0.3	-0.09	0.7	0.05	0.9
PLAT	0.07	0.6	-0.02	0.9	0.19	0.5
PRSS2	-0.09	0.5	-0.34	0.2	0.15	0.6
PRSS27	-0.33	0.06	-0.06	0.8	-0.70	0.1
REG1B	-0.02	0.9	0.78	0.1	-0.24	0.4
ROBO1	0.21	0.2	0.22	0.4	0.01	1
SDK2	0.03	0.9	0.32	0.2	-0.19	0.5
SELL	0.02	0.9	-0.36	0.2	0.20	0.5
SEMA7A	-0.01	0.9	0.05	0.8	-0.03	0.9
SIGLEC6	0.04	0.8	0.50	0.09	-0.39	0.2
TCN1	0.03	0.8	-0.17	0.5	0.19	0.6
TIGIT	0.19	0.2	0.50	0.09	0.26	0.4
TNFRSF4	-0.01	1	-0.23	0.4	-0.07	0.8

Log cOR: log (conditional odds ratio) of incident T1D risk per standard deviation increase in the average transcripts levels measured by RNA-seq before the age of one.

5. Prospective validation of proteins associated with T1D in high-risk individuals.

149

Table S5.11: Associations of average proteins levels before the age of one with the risk of incident IA.

Gene symbol	Overall (307 pairs)		Year 1 (135 pairs)		Year 2 (94 pairs)		Year 3 (45 pairs)	
	Log cOR	p-value	Log cOR	p-value	Log cOR	p-value	Log cOR	p-value
<i>CCL19</i>	0.20	0.02	0.29	0.02	0.29	0.07	-0.31	0.2
<i>CD274</i>	0.13	0.1	0.24	0.06	0.15	0.3	-0.27	0.2
<i>CD5</i>	0.17	0.05	0.20	0.1	0.30	0.07	-0.18	0.4
<i>CXCL10</i>	0.01	0.9	-0.08	0.5	0.19	0.2	-0.28	0.2
<i>CXCL11</i>	0.07	0.4	0.08	0.5	0.08	0.6	-0.01	1
<i>CXCL9</i>	0.09	0.3	0.07	0.6	0.16	0.3	-0.11	0.6
<i>FGF19</i>	0.07	0.4	-0.01	0.9	0.20	0.2	0.31	0.2
<i>LTA</i>	0.28	0.001	0.23	0.08	0.41	0.01	0.11	0.6

Log cOR: log (conditional odds ratio) of incident IA risk per standard deviation increase in the average proteins levels before the age of one.

Table S5.12: Associations of average protein levels before the age of one with the risk of incident T1D.

Gene symbol	Overall (107 pairs)		Year 1 (38 pairs)		Year 2 (30 pairs)	
	Log cOR	p-value	Log cOR	p-value	Log cOR	p-value
<i>CCL19</i>	0.01	0.9	-0.34	0.2	0.00	1
<i>CD274</i>	-0.08	0.6	-0.40	0.2	0.44	0.1
<i>CD5</i>	0.16	0.3	0.35	0.3	-0.01	1
<i>CXCL10</i>	0.00	1	-0.21	0.4	0.40	0.2
<i>CXCL11</i>	0.00	1	-0.18	0.5	0.72	0.08
<i>CXCL9</i>	-0.05	0.7	-0.35	0.1	0.68	0.08
<i>FGF19</i>	0.31	0.07	0.26	0.5	0.20	0.6
<i>LTA</i>	0.25	0.08	0.11	0.6	0.25	0.3

Log cOR: log (conditional odds ratio) of incident T1D risk per standard deviation increase in the average proteins levels before the age of one.

6

Discussion

In this chapter, I first restate the overall aim of this thesis and the key findings from each results chapter (Chapter 3, Chapter 4, Chapter 5) and then discuss the implications of the core genes and proteins in the context of T1D pathogenesis and their potential as drug targets for T1D prevention. To systematically evaluate drug targets, I apply the Guidelines On Target Assessment for Innovative Therapeutics (GOT-IT) framework, which considers five key domains: (1) target-disease linkage (causal relationship), (2) safety (on-target or target-related toxicity), (3) microbial targets (non-human targets), (4) strategic issues (clinical needs and commercial potential), and (5) technical feasibility (druggability, assayability, and biomarker availability) [Emmerich et al., 2021]. The putative core genes detected by GATE analysis are based on human genetic evidence. Hence, the microbial target domain is not applicable. Furthermore, in T1D there is a widely recognised unmet clinical need for disease-modifying therapy to improve metabolic control and reduce or eliminate the need for insulin therapy [Mazzotta et al., 2024, Starr et al., 2025]. Therefore, this discussion focuses on the other three aspects: causal relationship, safety, and technical feasibility. Integrating main findings from this thesis with published literature and information from pharmaceutical pipelines, I classify the identified core genes and proteins into four categories: (1) strong causal evidence with existing drugs in clinical pipelines; (2) strong causal evidence with agents at preclinical stages; (3) strong causal evidence but requiring further evaluation of technical feasibility; and (4) requiring further validation of causal relationships. Finally, I discuss the strengths and limitations of this thesis and future research directions.

6.1 Statement of overall aim

As outlined in Chapter 1, T1D is an autoimmune disease with a strong genetic component. Despite the efforts in disease-modifying therapeutic strategies for T1D, only one drug, teplizumab, has been licensed for T1D prevention in high-risk individuals, and its effect is limited to delaying disease onset by a median of 32.5 months [Mondal and Pappachan, 2025, O'Donovan et al., 2024, Herold et al., 2019, Sims et al., 2021]. Given that genetic evidence for a drug target-indication pair can substantially increase the probability of success in clinical trials, the overarching aim of this thesis was to provide human genetic evidence for prioritising drug targets for T1D prevention by identifying putative core genes and proteins for T1D. To achieve this, I leveraged a novel genetic analysis method (genome-wide aggregated *trans*-effects, GATE), together with data on genetic associations

with proteins from two large-scale pQTL GWAS studies, and genotypic data from two large studies with T1D cases and controls. I first computed GATE scores for proteins by adding the effects of *trans*-acting SNPs on protein levels and then tested these scores for association with T1D. This approach allowed me to detect genes upon which the effects of common variants across the genome come together to cause disease. The omnigenic model predicts that these genes will be in the effector layer of the regulatory pathway leading to disease [Boyle et al., 2017, Liu et al., 2019] and are good candidates as drug targets since they are closer to disease impact. I also conducted a follow-up study in the TEDDY dataset, a prospective cohort of high-risk children for the development of islet autoimmunity and T1D, to evaluate if measured transcripts and protein levels for core genes and proteins differ between individuals who progressed to islet autoimmunity or clinical T1D and individuals who did not.

6.2 Statement of principal findings

In Chapter 3, I identified 27 putative core proteins based on GATE analysis in two T1D case-control datasets (the Scottish dataset and UK Biobank). Of these, 11 were supported by a dose-response relationship (i.e., one of Hill's criteria for causation, a biological gradient between exposure and outcome) on T1D using the MR approach, one has been previously reported as a monogenic cause of autoimmune diabetes, one was supported by a *cis*-score association with T1D, four are located within 200 kb of known GWAS hits for T1D, 11 were supported by an association of their measured protein level with T1D, nine were supported by experimental evidence in animal models of T1D, and one was supported by reports that drugs targeting the protein or its ligand cause T1D.

In Chapter 4, I performed a cross-sectional analysis of the association of plasma protein levels with T1D using T1D cases and controls in the subset of UK Biobank participants with available proteomics measurements to identify circulating proteins with differing levels between cases and controls. To retain proteins with a likely causal role in disease, I additionally assessed the association between the GATE scores for the identified proteins and T1D in the remaining UK Biobank participants and in the Scottish dataset. This strategy yielded 24 proteins that were associated with T1D both at the measured level and the genetically predicted level, with consistent direction of effect and larger magnitude of effect size at the measured level. Of these, two were also identified as core proteins in Chapter 3,

while the other 22 represent novel findings. Among the novel proteins, three were supported by a dose-response effect from the MR analysis, one has been reported as a monogenic cause of autoimmune diabetes, one is located within 200 kb of a known GWAS hit, five were supported by experimental evidence in animal models of T1D, and two were supported by drug effects in humans.

Combining the findings from both chapters, I identified 49 unique proteins with putative “core gene” status for T1D. In Chapter 5, I conducted follow-up analyses using data from the TEDDY nested case-control study to examine whether the measured transcript or protein levels of these 49 proteins differ between individuals who progressed to IA or T1D and those who did not. Despite having limited statistical power and limited proteomics coverage (only eight proteins had available measurements), this analysis highlighted suggestive prospective associations of immune-related proteins, such as immune checkpoints and cytokines, with IA and T1D.

6.3 Implications of key pathways in T1D pathogenesis and their therapeutic potential

In this section, I assessed the 49 identified proteins as putative drug targets based on three domains from the GOT-IT recommendation (causal relationship, safety, and technical feasibility) and categorised them into four categories with increasing requirements for further assessment.

6.3.1 Strong causal evidence with existing drugs in clinical pipelines

Of the 49 proteins, five (PD1, PD-L1, LAG-3, CD28 and OX40) had strong evidence for causality, and there are drugs in clinical pipelines targeting them in the context of T1D or of other autoimmune or immune-mediated inflammatory diseases.

PD-1/PD-L1 pathway

By integrating the GATE association, a series of pre-specified criteria for the assessment of causality, and prospective validation in the TEDDY study, this thesis underscores the central role of the PD-1/PD-L1 pathway in T1D pathogenesis. The causal relationship between PD-1 (encoded by *PDCD1*) and T1D is

established by the strong genetic association of the PD-1 GATE score with T1D, supported also by a dose-response relationship in the MR analysis, reports of a rare loss-of-function variant in this gene causing autoimmune diabetes, T1D cases induced by PD-1 inhibitors in cancer immunotherapy, experimental validation in NOD mice, the strong cross-sectional association of the PD-1 measured protein level with T1D, and a nominal association of the *PDCD1* transcript level with progression to IA in high-risk individuals. Its ligand PD-L1 (encoded by *CD274*) was detected through the cross-sectional analysis of plasma proteins in the UKB proteomics substudy, which mainly comprised prevalent cases of T1D. Its causal role is further corroborated by the GATE score association with T1D, a report of monogenic autoimmune diabetes caused by a loss-of-mutation in the *CD274* gene, experimental validation in NOD mice, and T1D cases triggered by PD-L1 inhibitor therapy.

PD-1 is broadly expressed on immune cells such as activated T cells, B cells, NK cells, and natural killer T cells, as well as myeloid cells, antigen-presenting cells, and innate lymphoid cells; PD-L1 is primarily expressed on antigen-presenting cells and cancer cells, but also on non-haematopoietic cells, including pancreatic β cells [Khan et al., 2020]. Engagement of PD-1 with PD-L1 exerts an inhibitory signal that downregulates T cell receptor and CD28 signalling, reduces downstream transcription factor induction and cytokine production, and ultimately decreases T cell proliferation and survival. In NOD mice, PD-1 knockout (*Pdcd1*^{-/-}) has been shown to accelerate the onset and frequency of T1D with remarkable Th1 polarisation of islet infiltrates [Wang et al., 2005]. Mechanistically, PD-1 deficiency leads to the proliferation and pancreatic infiltration of islet-reactive CD4⁺ T cells in a cell-intrinsic manner [Pauken et al., 2013]. Transgenic overexpression of PD-L1 in pancreatic β cells in NOD mice revealed a protective effect against T1D [Wang et al., 2008], whereas blockade of PD-1 or PD-L1 rapidly precipitates T1D [Ansari et al., 2003].

In humans, rare loss-of-function mutations in *PDCD1* and *CD274* genes have been reported to cause monogenic autoimmune diabetes [Ogishi et al., 2021, Johnson et al., 2024]. Moreover, PD-1 and PD-L1 inhibitors that are widely used for cancer treatment can cause T1D as a rare adverse effect. Antibodies blocking PD-1 (nivolumab and pembrolizumab) and PD-L1 (avelumab, durvalumab and atezolizumab) have been reported to induce T1D, with the incidence ranging from 0.3% to 3.5% [Tachibana and Imagawa, 2022]. Such a form of T1D is termed “immune checkpoint inhibitor-induced type 1 diabetes” (ICIT1D), which is

characterised by rapid onset with diabetic ketoacidosis, severe hyperglycaemia, low to undetectable endogenous insulin secretion, and a low to medium positive rate (5% - 57%) for autoantibodies to glutamic acid decarboxylase (GADA) [Cho and Jung, 2023]. Two studies have investigated genetic factors associated with ICIT1D, reporting for instance, that germline mutations in *NLRC5*, a transactivator of HLA class I genes, were implicated in ICIT1D [Caulfield et al., 2023]. Of interest, a T1D polygenic risk score (PRS) was associated with ICIT1D with an OR of 4.4 for the individuals in the top decile of T1D PRS [Ruiz-Esteves et al., 2024], suggesting that those with a high genetic predisposition to T1D are more likely to get this adverse effect of PD-1/PD-L1 inhibitors.

The above lines of evidence point to a protective role for PD-1 and PD-L1 in T1D. Similarly, the prospective association analysis of the TEDDY study in Chapter 5 showed that whole blood *PDCD1* transcripts were negatively associated with the development of IA. On the contrary, the GATE scores and the circulating protein levels for PD-1 and PD-L1 were positively associated with T1D in Chapter 3 and Chapter 4. It is worth noting that the GATE scores were based on plasma protein pQTLs and reflect the genetically predicted circulating soluble protein level in plasma, while whole blood transcripts measure mRNAs in blood cells (e.g., leukocytes).

Soluble forms of immune checkpoint receptors and ligands – produced by either mRNA alternative splicing or shedding from the membrane-bound forms – are hypothesised to function in three potential mechanisms: (1) direct activation, directly binding to the corresponding ligands or receptors and activating downstream signalling pathways; (2) competitive inhibition, competing with the membrane-bound forms for ligand or receptor binding, thereby reducing signalling through the cell-surface pathway; (3) signal blockade, incorporating into the membrane-bound complexes as monomers and interfering with signalling transduction [Yuan et al., 2024].

PDCD1 has five exons, and four alternatively spliced PD-1 mRNA transcripts have been discovered in addition to the full-length isoform [Dai et al., 2014]. These include isoforms lacking exon 2 (PD-1 Δ ex2), exon 3 (PD-1 Δ ex3), exons 2 and 3 (PD-1 Δ ex2,3), and exons 2, 3 and 4 (PD-1 Δ ex2,3,4). The soluble form of PD-1 (sPD-1) is encoded by PD-1 Δ ex3, which retains the IgV-type extracellular domain of the membrane-bound form and therefore has the ability to bind to ligands. The other three alternatively spliced mRNA transcripts that lack exon 2 lose this ligand-binding capacity. Soluble PD-L1 (sPD-L1) can be produced

through both proteolytic cleavage of the membrane-bound form and alternative splicing [Bailly et al., 2021]. Three alternatively spliced forms of sPD-L1 have been discovered, including an isoform generated by the acquisition of an intronic element (LINE-2A) in the *CD274* gene which leads to the lack of T cell inhibitory activity; an isoform generated by a 66-bp deletion in exon 4 which lacks the transmembrane and intracellular domains; and an isoform containing the first four exons but not the fifth exon and leading to a lack of a transmembrane domain at its 3' end with a unique C-terminal sequence.

Mounting studies support that sPD-1 can act as a decoy and inhibit the interaction of membrane PD-1 with PD-L1, PD-L2 and CD80, and thus promote T cell responses, and this effect may differ depending on the strength of T cell receptor signalling and degree of dendritic cell maturation [Dai et al., 2014, Gu et al., 2018, Khan et al., 2020, Tzaban et al., 2025]. In addition, the concentration of sPD-1 was reported to be positively associated with progression to T1D in islet autoantibody-positive children in a longitudinal study [Bruzzaniti et al., 2024]. This is in agreement with the direction of effect of the sPD-1 (GATE score and measured levels) on T1D reported in this thesis and supports a possible function for sPD-1 as a decoy for the cellular receptor. Levels of sPD-L1 are considered as a pool of heterogeneous isoforms instead of a single species, given the various biochemical origins, and distinct functions have been identified depending on the context [Bailly et al., 2021]. So far, no study has specifically investigated the role of sPD-L1 in T1D pathogenesis.

The formation of islet autoantibodies depends on cognate interactions between T follicular helper cells (Tfh, a subset of CD4⁺ T cells) and B cells within the germinal centres [Falcone and Fousteroi, 2020]. This process is modulated by T follicular regulatory (Tfr) cells, a specialised population of FOXP3⁺ regulatory T cells. Both Tfh and Tfr cells express *PDCD1*; its expression preserves the balance between Tfh-driven autoantibody generation and Tfr-mediated regulation, whereas its blockade increases the Tfh/Tfr cell ratio, thus leading to the production of autoantibodies to insulin (IAA). This mechanistic evidence aligns with the nominal protective association of *PDCD1* transcripts against IA observed in the TEDDY analysis.

Clinically, several PD-1 agonists that aim to stimulate the inhibitory PD-1 signal are in pipelines for several autoimmune and chronic inflammatory conditions.

- Peresolimab (Eli Lilly) is a humanised IgG1 mAb designed to stimulate the endogenous PD-1 pathway. In a phase 2a clinical trial involving 98 patients with moderate-to-severe rheumatoid arthritis, intravenous administration of 700 mg every four weeks demonstrated clinical efficacy [Tuttle et al., 2023]. However, the subsequent phase 2b trial in rheumatoid arthritis was discontinued by the manufacturer following an overall benefit/risk assessment.
- Rosnilimab (AnaptysBio) is an IgG1 mAb that binds to the proximal epitope of membrane-bound PD-1, selectively targeting and depleting pathogenic T cells in the peripheral and inflammatory tissues to restore immune homeostasis. In a phase 2b trial involving 424 patients with moderate-to-severe rheumatoid arthritis, rosnilimab achieved efficacy comparable to JAK inhibitors, with progressive improvement through six months, particularly in the 400 mg every four weeks and 600 mg every two weeks dosing groups [AnaptysBio, 2025]. The treatment was well tolerated, with no severe adverse events reported. Rosnilimab is also under phase 2 evaluation for moderate-to-severe ulcerative colitis, and relevant results are expected to be released by Q4 2025.
- JNJ-67484703 (Johnson & Johnson) is also an IgG1 mAb with dual agonistic and depleting functions. In a phase 1b clinical trial of 92 active rheumatoid arthritis patients, administration of 2 mg/kg and 3 mg/kg through 10 weeks was safe and well tolerated, and the 3 mg/kg group showed efficacy [Ling et al., 2025]. A separate phase 2 open-label proof-of-biology trial in 37 patients with rheumatoid arthritis, ulcerative colitis or primary Sjogren's syndrome demonstrated reductions in the ratio of PD-1^{hi} CD4⁺ T follicular and peripheral helper cells in disease-relevant tissues [Fisher et al., 2025]. This antibody is also being explored in a phase 2a proof-of-concept trial of atopic dermatitis (EUCTR2022-001528-14-DE).
- GenSci120 (GenSci) is an agonistic mAb generated using Beacon single plasma cell cloning and murine antibody humanisation techniques [Jiang et al., 2024]. It is currently in phase 1 clinical trials in healthy individuals (NCT06827457, NCT07040930) and will subsequently be investigated for systemic lupus erythematosus, primary Sjogren's syndrome, inflammatory bowel disease, and rheumatoid arthritis.
- S-4321 (Seismic Therapeutic) is a dual-cell bidirectional antibody that agonises the PD-1 receptor on T cells (Fab domain) and selectively engages the Fc γ receptor on antigen-presenting cells (Fc domain) [Manasson et al., 2024]. It is being assessed in phase 1 among healthy volunteers (NCT06877611).

- IMC-S118AI and IMC-U120AI (Immunocore) are two PD-1 agonists developed based on the immune modulating monoclonal T cell receptor against autoimmunity (ImmTAAI) technology. These agents specifically target cell-bound PD-1 and mimic the natural PD-1/PD-L1 interaction, ensuring PD-1 activation occurs in *cis* with T cell receptor engagement on the attacking autoreactive T cells [Curnock et al., 2021]. IMC-S118AI selectively targets pancreatic β cells through the recognition of a peptide from pre-proinsulin presented by HLA-A*02:01, and a clinical trial application is in preparation. IMC-U120AI is a non-HLA-restricted, CD1a-tethered PD1 agonist intended for clinical evaluation in atopic dermatitis.

In summary, the key findings presented in this thesis strengthen the rationale for exploring PD-1 agonists in disease-modifying therapies for T1D. Advances in their clinical trials in other autoimmune conditions may accelerate T1D trials through drug repurposing strategies. Furthermore, the emergence of β cell-specific agonists offers a promising avenue for achieving targeted and potentially safer immune modulation for T1D prevention or treatment.

LAG-3 pathway

Among the putative core proteins identified in Chapter 3, *LAG3* is also an inhibitory immune checkpoint, and its causal role in T1D pathogenesis was supported by the strong association of the GATE score for the circulating protein level with T1D, further supported by the MR analysis, the cross-sectional association of the measured circulating protein levels with T1D and by evidence from experimental animal models. Widely expressed on activated effector and regulatory T cells, NK cells, plasmacytoid dendritic cells, B cells, and even neurones, LAG-3 can selectively bind to stable peptide-MHC II complexes and also non-MHC II ligands such as fibrinogen-like protein 1, liver sinusoidal endothelial cell lectin, galectin-3, and α -synuclein fibrils [Maruhashi et al., 2020]. In NOD mice, both genetic deficiency of *LAG3* (*Lag3*^{-/-}) and blockade using anti-LAG-3 antibody accelerated the development of T1D [Bettini et al., 2011]. Also in NOD mice, LAG-3 was shown to maintain an exhaustion programme in autoreactive CD8⁺ T cells by limiting their effector-like functions and their trafficking to the islets [Grebinoski et al., 2022]. On the other hand, Treg-specific deletion of *Lag3* delayed the development of diabetes in NOD mice, indicating that LAG-3 may promote autoimmunity in a chronic autoimmune-prone environment by suppressing the proliferation and function of regulatory T cells [Zhang et al., 2017].

Soluble forms of LAG-3 can be generated either by shedding from membrane-bound LAG-3 or alternative splicing with a stop codon and additional eight amino acids from an intron downstream of exon 5 [Yuan et al., 2024]. The function of soluble LAG-3 remains to be elusive. The positive association of soluble LAG-3 with T1D in this thesis may be explained by the competitive inhibition and signal blockade hypotheses of soluble immune checkpoint proteins.

A monoclonal antibody that depletes LAG3-expressing cells, GSK2831781, has been confirmed to be safe and effective in reducing LAG3⁺ T cells in psoriasis patients in a phase 1 trial [Ellis et al., 2021], but it failed to control the inflammation in a subsequent phase 2 trial in ulcerative colitis [D'Haens et al., 2023]. Nonetheless, with a similar idea to PD-1 agonists, a LAG-3 agonist, IMP761 (Immutep), has been developed to treat T cell-mediated autoimmune diseases [Angin et al., 2020]. It is now undergoing a phase 1 trial in healthy volunteers (NCT06637865), of which preliminary data showed encouraging safety and efficacy, and further data will be released later this year [Read, 2025].

CD28 pathway

This thesis reinforces the pivotal role of the CD28 pathway in T1D pathogenesis and highlights its therapeutic potential in T1D prevention. CD28 emerged as a potential core protein in Chapter 4: its circulating levels were higher in individuals with T1D compared with non-diabetic controls in the small UKB-PPP dataset; genetically predicted levels (GATE score) were associated with T1D in a consistent direction of effect among the remaining UK Biobank participants; and MR analysis provided further support.

CD28 is a co-stimulatory immune checkpoint on T cells; its interaction with the ligands CD80 and CD86 on antigen-presenting cells is essential for T cell activation. CTLA-4 binds to the same ligands with much higher affinity, thereby functioning as a natural inhibitor of CD28. While soluble CD28 can arise from both alternative splicing and shedding of the membrane-bound form, the latter mechanism has been demonstrated in the context of autoimmune diseases such as systemic lupus erythematosus and Grave's disease [Aprilia et al., 2024]. Thus, the observed positive association between soluble CD28 and T1D risk may reflect increased CD28 expression on T cells and heightened T cell activation.

In contrast, higher CD28 transcript levels were nominally associated with a reduced risk of IA among individuals with high-risk HLA haplotypes and/or a family history of T1D in Chapter 5. This apparent discrepancy with the soluble CD28

findings may be explained by the context dependency of CD28 signalling. For instance, blocking CD28 signalling with CTLA-4 Ig or an anti-CD86 antibody at the onset of insulinitis (2-4 weeks of age) markedly reduced the incidence of diabetes in NOD mice, but the same interventions initiated later (~ 10 weeks of age) had no effect on preventing diabetes [Lenschow et al., 1995]. CD28 deficiency from birth – either via *CD28* knockout or transgenic CTLA4 Ig expression – paradoxically accelerated diabetes in the NOD mice, with increased T helper 1 cytokines (e.g., IL-2 and IFN – γ) and reduced T helper 2 cytokines (e.g., IL-4) [Lenschow et al., 1996]. Additional work using NOD mice deficient in the ligands CD80 and CD86 showed a more severe diabetes, decreased T helper 1 and T helper 2 cytokines, and a prominent reduction in CD4⁺CD25⁺ regulatory T cells, suggesting that regulatory T cells, rather than the balance between T helper 1 and T helper 2 cytokines, are the key mediators of this pathway [Salomon et al., 2000]. This conclusion was confirmed when transfer of CD4⁺CD25⁺ regulatory T cells into CD28-deficient NOD mice delayed the development of diabetes. Moreover, a selective antagonist CD28 antibody had no impact on disease onset when administered for four weeks to NOD mice aged 4, 6, or 10 weeks, but prolonged 10-week administration to six-week-old mice accelerated diabetes [Besançon et al., 2018]. In contrast, combining short-term CD28 antagonism with the mTOR inhibitor rapamycin in 10-week-old mice delayed diabetes onset by synergistically downregulating effector T cell activation and migration into pancreatic islets. Collectively, the direction of effect and timing for the CD28 signalling pathway in the development of T1D remains to be disentangled in mechanistic studies.

Clinically, given the higher affinity to the same ligands and more consistent findings of CTLA-4 over CD28, much more effort has been put into clinical trials of the CTLA-4 Ig, abatacept. As described in Section 1.4, abatacept succeeded in preserving β cell function in stage 3 T1D patients but failed to delay disease progression in stage 1 individuals. It is also evaluated together with insulin (NCT05742243) or rituximab (NCT03929601) in people with stage 3 T1D. One CD28 antagonist – FR104 – was shown to be safe and well tolerated in healthy participants [Poirier et al., 2016], but it has not been tested in any autoimmune conditions in clinical trials. ALPN-101, a dual antagonist against the CD28 and the inducible T cell costimulator (ICOS) pathway, is being investigated in a phase 2 trial of systemic lupus erythematosus (NCT04835441).

TNFRSF4 (OX40) pathway

TNFRSF4 was identified as a potential core protein in Chapter 4, where its circulating levels were higher among T1D cases as compared to non-diabetic controls, and its genetically predicted circulating levels were also associated with a higher risk of T1D. Although the MR analysis did not meet the pre-specified p -value threshold (0.01) for declaring a significant causal effect of the protein level on T1D, the result was nominally significant (p -value = 0.02). Furthermore, the *TNFRSF4* transcript levels were found to be nominally associated with an increased risk of progression to IA in the TEDDY study in Chapter 5.

TNFRSF4 belongs to the TNF receptor superfamily and encodes OX40, which is also a co-stimulatory immune checkpoint primarily expressed on activated T cells. It binds to the ligand OX40L (encoded by *TNFSF4*) on activated antigen-presenting cells. Soluble OX40 is believed to be shed from the membrane-bound form, as the concentration of the soluble form was elevated but the membrane-bound form was decreased in T1D patients compared with healthy controls [An et al., 2019].

Notably, animal experiments implicated the role of OX40/OX40L interaction on both effector and regulatory T cells in an age- and context-dependent pattern. For effector T cells, OX40L deficiency from birth (knockout) accelerated the activation and migration into islets but markedly prevented the incidence of diabetes in NOD mice and T cell receptor transgenic mice, implying that OX40L is dispensable for initial priming but important for the sustained expansion/maintenance of pathogenic effector responses [Martin-Orozco et al., 2003]. Another study discovered that the expression of OX40 and OX40L coincides in a late stage of NOD mice lifespan (~ 11-13 weeks of age, just before the initiation of β cell destruction), and blockade of OX40/OX40L interaction at 12 weeks of age and onwards (but not earlier) inhibits progression to overt diabetes [Pakala et al., 2004]. By contrast, OX40 agonism was shown to induce regulatory T cells and thereby prevent the incidence of diabetes in 6-week-old NOD mice, and this protection was synergistically increased in the combination of insulin B9:23 peptide immunisation [Bresson et al., 2011]. Administration of OX40L to 6-week-old NOD mice induced stable regulatory T cell expansion in the thymus and the periphery and thus prevented the development of diabetes [Haddad et al., 2016]. However, the same study showed that administration of OX40L in 12-week-old NOD mice induced labile (unstable) regulatory T cells in the thymus, which can gain helper-like function and promote the differentiation of naive T cells into effector T cells under pro-inflammatory conditions and lead to rapid onset of diabetes.

So far, several OX40/OX40L antagonist antibodies, including amlitelimab (Sanofi), rocatinlimab (Amgen) and telazorlimab (Ichnos Glenmark Innovation), have progressed to phase 2 clinical trials in atopic dermatitis. Amlitelimab is also being trialled in phase 2 for autoimmune diseases such as coeliac disease, systemic sclerosis, and alopecia areata. Brivekimig (Sanofi), a dual antagonist targeting both OX40L and TNF, is being trialled in T1D in phase 2.

6.3.2 Strong causal evidence with agents at preclinical stages

In this thesis, the GATE analysis and downstream assessment of causality highlighted three other proteins (galectin-9, IDO-1, and TIGIT) with strong causal evidence and promising therapeutic potential in T1D treatment and prevention; however, agents targeting these proteins and corresponding pathways are still at preclinical stages in animal models of T1D and other autoimmune conditions and haven't been progressed into clinical trials yet.

Galectin-9 pathway

LGALS9 (galectin-9, Gal-9) was identified as a putative core protein in Chapter 3, with supportive evidence from a dose-response relationship in the MR analysis, measured protein level association, and experimental validation in animal models.

Gal-9 is a ligand for an inhibitory immune checkpoint receptor TIM-3 (T cell immunoglobulin and mucin-domain-containing molecule 3). Long-term injection of Gal-9 and an anti-TIM-3 antibody were both shown to reduce development of diabetes in NOD mice by downregulating T helper 1 immunity (which promotes autoimmunity) [Kanzaki et al., 2012]. Another study revealed the effect of Gal-9 in regulating the B cell activation threshold, with loss of this regulatory network leading to spontaneous autoimmunity [Smith et al., 2021]. Bioengineered extracellular vesicles and platelets presenting high levels of immune checkpoint ligands, including Gal-9, PD-L1, PD-L2, and B and T lymphocyte attenuator (BTLA), were reported to stimulate macrophage polarisation to the immunosuppressive M2 phenotype and ameliorate new-onset T1D in NOD mice [Yang et al., 2024, Ma et al., 2025]. Taken together, the Gal-9 pathway has demonstrated promising therapeutic potential for the treatment and prevention of T1D.

IDO-1 pathway

Identified as a putative core protein in Chapter 3, the essential role of *IDO1* was supported by a dose-response relationship in the MR analysis and perturbation in experimental animal models.

IDO1 encodes indoleamine 2,3-dioxygenase 1 (IDO-1), a rate-limiting enzyme for the degradation of tryptophan into kynurenines. It is widely expressed across tissues, including antigen-presenting cells and pancreatic β cells, and its expression is inducible by inflammatory stimuli such as interferons (INF), TGF- β , lipopolysaccharide, IL-1, IL-6, IL-10 and TNF [Zhai et al., 2018, Pallotta et al., 2022]. Regulatory T cells expressing CTLA-4 can upregulate IDO-1 expression in dendritic cells via IFN- γ signalling following CTLA-4-CD80/CD86 engagement. Binding of the soluble glucocorticoid-inducible TNF receptor to its ligand activates reverse signalling in plasmacytoid dendritic cells and thus increases IDO-1 expression. In human pancreatic tissue, IDO-1 is constitutively expressed in β cells under healthy conditions, while its expression diminishes prior to β cell destruction in individuals with multiple autoantibodies [Anquetil et al., 2018]. Functionally, defective IDO-1 activity has been observed in dendritic cells in NOD mice and in peripheral blood mononuclear cells of paediatric T1D patients [Pallotta et al., 2022]. Notably, this defect is driven by IL-6 and can be corrected by the blockade of the IL-6 receptor using the recombinant antibody tocilizumab (TCZ). In parallel, a proteasome inhibitor, bortezomib (BTZ), was also shown to prevent the onset of diabetes in NOD mice through mechanisms dependent on IDO-1 and plasmacytoid dendritic cells [Mondanelli et al., 2017]. In combination with a low-dose anti-CD3 antibody, BTZ demonstrated efficacy in reversing diabetes in overtly diabetic mice. Moreover, mRNA delivery of membrane-anchored IDO-1 has shown suppressive effects in three animal models of T cell-mediated autoimmunity: experimental autoimmune encephalomyelitis (an animal model for multiple sclerosis), collagen-induced arthritis (an animal model for arthritis), and acute graft-versus-host disease [Kenney et al., 2024].

TIGIT pathway

TIGIT was identified as a putative core protein in Chapter 3 with additional supportive evidence from experimental studies. It encodes the T cell immunoreceptor with Ig and ITIM domains and is a co-inhibitory immune checkpoint receptor predominately expressed on activated T cells and NK cells. TIGIT has multiple ligands, including CD155, CD112, CD113, and CD114, as well as the Fap2 protein expressed by *F. nucleatum* [Yue et al., 2022]. Compared to a co-stimulatory immune checkpoint receptor, CD226, TIGIT has greater affinity to the

shared ligand CD155. In addition, TIGIT can even bind to CD226 directly and destroy its homodimer, thereby acting as an inhibitory competitor of CD226. It has been discovered that the TIGIT pathway is downstream to the PD1/PDL1 pathway in the hierarchy of immune checkpoints in the context of autoimmune diabetes in NOD mice [Trivedi et al., 2024]. Several studies have indicated that TIGIT regulates T1D autoimmunity through CD8⁺ T cell exhaustion, regulatory T cell affinity to T cell receptor, and balance of co-stimulatory and co-inhibitory receptors in NK cells [Yue et al., 2022]. Different strategies, including TIGIT overexpression, agonistic antibodies, recombinant CD155 protein, and TIGIT-Ig fusion protein, have been explored in mouse models for the treatment of various autoimmune diseases, such as arthritis, lupus, aplastic anaemia, and multiple sclerosis.

6.3.3 Strong causal evidence but requiring further evaluation of technical feasibility

Despite being supported by strong causal evidence in this thesis, some proteins and corresponding pathways have not been extensively studied. Further evaluation of technical feasibility is required prior to the investigation of therapeutic potentials.

CD5

CD5 was identified as a putative core protein in Chapter 3 with additional supportive evidence from the MR analysis, measured protein level association, proximity to a known T1D GWAS hit, and nominal protein level association with incident IA.

Primarily expressed on T cells and some B cells with scavenger receptor cysteine-rich domains in its extracellular region, CD5 has been recognised as a negative regulator of T cell receptor and B cell receptor signalling [Dalloul, 2009, Brown and Lacey, 2010]. In the context of cancer, anti-CD5 mAb demonstrated effects in enhancing T cell activation and suppressing tumour growth [Alotaibi et al., 2024]. In addition, immune checkpoint blockade therapy (e.g., anti-PD-1) can induce the expression of CD5 on a subset of dendritic cells, which in turn mediates the priming of antitumour T cells [He et al., 2023]. However, CD5 displays heterogeneous roles in different autoimmune disorders: while being pro-inflammatory in rheumatoid arthritis and primary Sjogren's syndrome, it presents

a protective role in Grave's disease and a dual role in systemic lupus erythematosus [Dong et al., 2025]. Compared with controls, children with T1D had lower proportions of CD8⁺ T cells that express low to medium CD5 but higher proportions of CD5-expressing B cells [Wadenpohl et al., 2021, Tompa and Faresjö, 2024]. In the 1990s, anti-CD5 immunoconjugate was investigated in a pilot phase 1 trial in subjects with recent-onset T1D and showed safety and preliminary efficacy in preserving β cell function [Skyler et al., 1993], but no follow-up trials have been registered, which might be due to the insufficient characterisation and understanding of CD5 at the molecular level [Voisinne et al., 2018]. Extensive mechanistic studies are needed to decode its role in T1D prior to the design of alternative therapeutic strategies.

CD48

CD48 turned up as a putative core protein through the GATE analyses in both the Scottish and UK Biobank datasets in Chapter 3. It was further supported by a dose-response effect in the MR analysis, a cross-sectional measured protein level association, and a nominal transcript level association with progression to IA.

CD48 (SLAMF2), a member of the signalling lymphocyte activation marker (SLAM) family, is broadly expressed on most haematopoietic cells and plays a key role in immune cell adhesion and activation [McArdel et al., 2016b]. CD48 has multiple ligands, including CD2, CD244, and also the lectin FimH on *E. coli*. Among these, CD48 exhibits a higher affinity to CD244 than to CD2, whereas CD58 – a homolog of CD48 – shows a stronger affinity to CD2 than CD48. Functionally, CD48-CD2 interaction enhances T cell signalling, while CD48-CD244 interaction can be either stimulatory or inhibitory, depending on the context. Additionally, CD48 has been implicated in the activation of the PD-1 pathway by facilitating crosslinking and phosphorylation [White et al., 2023]. CD2-derived peptides that block CD2-CD48/CD58 interactions have been shown to delay the progression of collagen-induced arthritis (an animal model for arthritis) [Gokhale et al., 2013]. Anti-CD48 mAb can attenuate experimental autoimmune encephalomyelitis (an animal model for multiple sclerosis) by limiting CD4⁺ T cell proliferation and depleting CD48-expressing CD4⁺ T cells [McArdel et al., 2016a]. Although a CD58-Ig fusion protein has shown some effectiveness in preserving β cell function and an anti-CD2 antibody is being examined in a phase 1b clinical trial in individuals with newly diagnosed T1D (see details in Section 1.4), direct targeting of CD48 itself has not yet been investigated, even in preclinical models of T1D.

NK cells

Two receptors on NK cells – *NCR1* (natural cytotoxicity triggering receptor 1, NKp46) and *KLRB1* (killer cell lectin-like receptor subfamily B member 1) – were identified as putative core proteins through the GATE analysis in Chapter 3 and were supported by other lines of evidence, including dose-response effects in the MR analyses, measured protein level association, proximity to a known T1D GWAS hit, and experimental animal models. These highlighted the important role of NK cells in the pathogenesis of T1D and the therapeutic potential of targeting NK cells in the treatment and prevention of T1D.

Innate immune cells, including the NK cells, are thought to be engaged at the earliest stages of T1D pathogenesis [Herold et al., 2024]. While the mechanism of NK cells in the development of T1D remains unknown, the NK cell receptor, NKp46 (encoded by *NCR1*), was found with the ability to recognise unknown ligand(s) on pancreatic β cells in mouse and human [Gur et al., 2010]. NKp46-deficient mice displayed a preventative effect against streptozotocin-induced diabetes, and injections of the soluble form of NKp46 at an early stage acted like a vaccine to protect NOD mice from diabetes. However, studies on NK cells and T1D are inconclusive in humans, due to the difficulty of investigating the pancreas directly and the limitations of peripheral blood samples in reflecting pancreatic pathogenesis [Kucuksezer et al., 2021]. EndoC- β H5 cells are storable and ready-to-use β cell lines with a robust and reproducible insulin secretion feature [Blanchi et al., 2023]. Our research group is currently working with biologists to validate the expression of NKp46 ligand(s) on EndoC- β H5 cells and the interaction with NKp46 fusion protein. Pull-down experiments are planned subsequently to identify the pancreatic β cell ligand(s) for NKp46.

Chemokines

Four chemokines – *CXCL9*, *CXCL10*, *CXCL11*, and *CCL19* – emerged as putative core proteins for T1D. The role of *CXCL9* in T1D was also supported by a *cis*-score association; that of *CXCL10* was supported by the MR analysis and evidence in experimental models, and that of *CCL19* was supported by a suggestive prospective association of protein level with progression to IA. *CXCL9*, *CXCL10* and *CXCL11* encode ligands for C-X-C motif chemokine receptor 3 (CXCR3), and *CCL19* encodes a ligand for chemokine receptor 7 (CCR7).

Over the past years, drugs targeting chemokines failed in clinical trials for inflammatory diseases, which has been attributed mainly to the redundancy of

the chemokine system, i.e., numerous chemokine ligands bind to a wide range of receptors and vice versa [Dyer, 2020]. However, recent studies have illustrated a different scenario, i.e., chemokine biology is in fact based on a high degree of specificity via spatiotemporally differential expression patterns, biased signalling from different ligands binding to the same receptor, and interactions with glycosaminoglycans in vascular and tissue extracellular matrix. For instance, even though the expression levels of *CXCL9*, *CXCL10*, and *CXCL11* are highly correlated with each other during inflammation, their expression is not related to the expression of the receptor *CXCR3*, suggesting they are likely produced at different sites [Ridley et al., 2023]. While *CXCL9* has the ability to recruit T cells *in vivo*, the other two ligands do not show similar effects. Moreover, they are distributed distinctly within a tissue or on cell surfaces and have different affinities to glycosaminoglycans. Therefore, in-depth mechanistic studies would be required to localise and understand the effects and time points of chemokines in the pathogenesis of T1D before considering for therapeutic targeting.

Fas-mediated apoptosis

FASLG was identified as a putative core protein in Chapter 3 with additional support from the MR analysis, proximity to a known T1D GWAS hit and evidence in experimental studies. It encodes Fas ligand (FasL), which is primarily expressed on cytotoxic T cells and NK cells. The soluble form of FasL is shed from the membrane-bound form [Trivedi et al., 2019].

The interaction between the Fas receptor and FasL is central for regulating programmed cell death (i.e., apoptosis). In T1D, Fas/FasL interaction has been found to be involved in insulin secretion and immune homeostasis modulation physiologically and to mediate β cell apoptosis and islet injury in inflammatory circumstances [Yolcu et al., 2017]. However, the therapeutic implications of Fas/FasL interaction in the prevention of T1D remain undetermined. Although NOD mice that are deficient in both the membrane-bound form and the soluble form of FasL were protected from diabetes and neutralisation with anti-FasL antibody showed preventative efficacy [Nakayama et al., 2002], another study reported that the soluble form of FasL is not required in the pathogenesis of diabetes [Trivedi et al., 2019]. Furthermore, multiple mechanisms engage in the regulation of β cell death in a redundant and interrelated pattern; it is thus difficult to designate the exact role of Fas-mediated apoptosis in the pathogenesis of T1D. Therefore, more studies are needed to understand the exact mechanism of the Fas pathway in T1D before considering possible therapeutic investigations.

Notch pathway

Both *NOTCH1* and one of its ligands, *DLL1*, were identified as potential core proteins using the alternative strategy in Chapter 4. Experimental studies have revealed the dual roles of the Notch pathway in regulating pancreatic homeostasis [Rubey et al., 2020, Eom et al., 2021] and deciding immune cell lineage in T1D [Wagner et al., 2005]. In addition to *NOTCH1* and *DLL1*, another three Notch receptors (encoded by *NOTCH2*, *NOTCH3*, and *NOTCH4*) and four Notch ligands (encoded by *DLL3*, *DLL4*, *JAG1*, and *JAG2*) have been discovered in mammals [Allen and Maillard, 2021, Christopoulos et al., 2021]. Over the past decades, much effort has been put into the investigation of inhibiting the Notch pathway with γ -secretase inhibitors and antibodies targeting Notch receptors and ligands in the therapy of cancers as well as inflammatory diseases, while the outcomes of these investigations have not led to any effective treatments. One key issue is that Notch receptors and ligands are ubiquitously expressed on numerous cell types while their roles are highly context-dependent. Instead of prolonged systematic pan-Notch inhibition which might induce intolerable side effects, pulses of Notch inhibition that target specific receptors/ligands at specific times in the disease course may offer more promising therapeutic potentials.

6.3.4 Requiring further validation of causal relationships

Pancreatic and gut proteins

In Chapter 3, eight proteins involved in exocrine pancreas and gut functions (*PRSS27*, *AGR2*, *FGF19*, *GCG*, *CCK*, *CFC1*, *TCN1*, *REG1B*) emerged as putative core proteins, and in Chapter 4, three other pancreatic and gut proteins (*PRSS2*, *MUC2*, *PNLIPRP1*) were identified as potential core proteins. Among these 11 proteins, only *PRSS27* was supported by a dose-response effect in the MR analysis, *REG1B* was supported by experimental evidence, and *GCG* is located within 200 kb from a T1D GWAS hit, *IFIH1*. *FUT2*, a known T1D GWAS hit that is involved in gut mucus homeostasis, is a peripheral master regulator for all pancreatic and gut proteins detected, except *CFC1*. Thus, these pancreatic and gut proteins are less likely to be causal, since the direct effect of *FUT2* might confound the GATE score associations.

Though T1D is primarily characterised by the autoimmune destruction of endocrine β cells, exocrine pancreas abnormalities are also a prominent feature of T1D [Atkinson and Mirmira, 2023]. Compared to healthy individuals, people

with T1D tend to have a smaller pancreas (in both weight and volume) with immune cell infiltration, acinar atrophy, ductal/vascular fibrosis, angiopathy, and dysfunction [Wright et al., 2023, Foster et al., 2020, Nigi et al., 2020, Vecchio et al., 2019]. The timing of exocrine pancreas abnormalities in the development of T1D remains debatable: while some studies showed a reduction of exocrine pancreas mass and function prior to the appearance of the first autoantibody, others suggested this reduction occurs between stage 1 and stage 3 [Bruggeman and Schatz, 2023]. The underlying mechanism is not well understood, and several hypotheses, including developmental defects, insulinopenia, endocrine deficiencies induced exocrine atrophy, and autoimmune or inflammatory destruction of exocrine cells, have been proposed. Moreover, the insufficient sources of pancreatic tissues and the inability to investigate the trajectory of pancreatic changes *in vivo* make it more difficult to understand the biology. With the development of pancreatic databases, such as Pancreatlas <https://pancreatlas.org/>, PANC-DB <https://hpap.pmacs.upenn.edu/>, and PanKbase <https://pankbase.org/>, more resources will become available for the investigation of the exocrine pancreas in the pathogenesis of T1D.

Other immune proteins

In Chapter 3 and Chapter 4, more than a dozen proteins with general immune regulatory functions were identified, but most have only limited or no additional supporting evidence for a causal role in T1D. Among the 27 putative core proteins, *CRTAM* and *SDK2* were only supported with measured protein level association, while *SELL*, *LTA*, and *FCRL6* lacked additional supporting evidence. Among the 22 novel potential core proteins, *ADA2* and *ADGRE2* showed dose-response effects in the MR analyses; *ITGB2* and *ALCAM* were supported by experimental evidence, and the remaining proteins lacked corroborating evidence. Although some of these proteins (*PDCH17*, *SIGLEC6*, *DDR1*, *SELL*, *OXT*) were supported by prospective associations between their transcript levels and incident IA and T1D in Chapter 5, little is known about their exact roles in the pathogenesis of T1D. Additional analyses of genetic and biomarker data and mechanistic studies are warranted to further validate the associations with the remaining 19 proteins reported in this thesis and clarify their biological relevance in T1D.

6.4 Strengths and limitations

In this thesis, one of the key strengths is that I applied a novel genetic approach inspired by the omnigenic hypothesis, which posits that most of the genetic effects on a typical complex trait are mediated through weak *trans*-effects of common variants that coalesce on the expression of a relatively sparse set of “core genes” in relevant tissues. While the omnigenic hypothesis was first proposed in 2017, it was only after deCODE and UKB-PPP released summary statistics for *trans*-effects of pQTLs on thousands of proteins that it became possible to compute aggregated *trans*-scores and test for association with disease. As stated in Chapter 1, conventional post-GWAS approaches generally assume that effector genes for a disease phenotype are located within or near GWAS loci and focus on *cis*-effects of genetic variants on molecular traits (*cis*-molQTL). However, approximately 70% of heritability of gene expression is attributed to weak *trans*-effects [Ouwens et al., 2020], and disease-relevant genes tend to be enriched with redundant enhancer domains and depleted of *cis*-QTLs of large effect due to selective constraints [Wang and Goldstein, 2020, Mostafavi et al., 2023]. On the other hand, *trans*-effects are typically weak and pleiotropic and can thus “escape” selective pressure. By aggregating these weak *trans*-effects, the GATE analysis is capable of identifying genes that are closer to disease in the regulatory pathway, offering more plausible therapeutic targets than those identified through conventional GWAS approaches, which often yield upstream genetic determinants.

The second strength is the use of a series of pre-specified criteria to further assess causality of the putative core proteins in T1D. Firstly, MR was employed to test if the effects of *trans*-pQTL instruments on disease risk are proportional to their effects on protein level, i.e., the causal effect of the protein on disease is supported by multiple independent loci. Then, six other sources of evidence were evaluated: (1) rare variants in the gene causing monogenic diabetes, (2) association of *cis*-score with T1D, (3) T1D GWAS hit within 200 kb, (4) association of measured protein level with T1D, (5) support from perturbation in experimental models of T1D, and (6) drugs targeting the protein are effective or cause the disease in humans.

The third strength is that I employed an alternative strategy to identify potential core proteins by directly analysing measured protein levels in a small, prevalent case-control study and validating cross-sectional associations with genetically predicted protein levels (i.e., GATE scores) in two large T1D case-control

datasets. This strategy complemented the original GATE analysis that uses a stringent p -value threshold of 10^{-6} for declaring putative core proteins based on the strength of association between the GATE score and T1D. Using a stringent threshold limits the detection of false positive findings, but it also misses true effects that are smaller or noisier. By employing the alternative strategy, I was able to detect several additional putative core genes that did not meet the stringent p -value threshold in the original GATE study but still had strong associations with T1D at $p < 0.001$. Using this more liberal threshold from the start would have resulted in over two hundred genes being detected, many of which would have been false positive associations. By filtering first based on the cross-sectional association with protein levels, I was able to reduce this number to 24 proteins with the most consistent evidence of association with T1D across genetic and proteomic data.

The fourth strength is that I followed up my putative core genes using data from a prospective nested case-control study (TEDDY). The prospective study design established temporality for the associations by showing that dysregulations of transcripts or proteins happen prior to the onset of disease, which provided strong evidence for a possible causal relationship.

In addition, a major contribution of my work and a strength of my thesis is the thorough consideration of the issues of reverse causation and confounding when attributing causality to GATE-disease associations and how these can be evaluated and addressed (Table 6.1). Associations of GATE scores with diseases can reflect reverse causation if (1) there is a high prevalence of the disease in the learning set where the SNP-protein coefficients are derived; (2) the protein is pathognomonic of disease even if the prevalence of that disease is low in the learning set, e.g., an elevated thyroid-stimulating hormone is pathognomonic of hypothyroidism. Regarding point 1, T1D is relatively rare in the general population (prevalence is 0.5%). In the UKB-PPP dataset, there are 232 individuals with T1D and of European ancestry. These individuals would have been included in the learning set for the SNP-protein coefficients made available by [Sun et al., 2023], which I used to construct GATE scores. The SNP-protein coefficients computed after excluding these 232 individuals were highly correlated (Pearson $r > 0.999$) with the published summary statistics for all putative core proteins. This confirmed that the inclusion of the small proportion of individuals with T1D in the learning set did not affect the SNP-protein coefficients. Regarding point 2,

Table 6.1: Potential issues and solutions for differentiating causality from reverse causation and confounding.

Potential issue	Solution
High prevalence of disease in pQTL studies	T1D is rare (prevalence is 0.5%)
Protein is pathognomonic of disease	Check distribution among cases and controls
Overfitting	Independent training set and test set
Direct effect of <i>trans</i> -pQTLs on disease	Filtering on the number of <i>trans</i> -pQTLs, GATE score composition, MR analysis

all candidate core proteins identified in Chapter 3 and Chapter 4 have overlapping distributions between T1D cases and non-diabetic controls, and none are pathognomonic for T1D.

Associations of GATE scores with diseases may be confounded by (1) overfitting due to sample overlap (i.e., the learning set in which the SNP-protein coefficients are derived includes participants that are also part of the disease case-control “test” dataset); (2) effects of *trans*-pQTLs on disease that are not mediated through the protein under study (i.e., direct effects). In this thesis, the Scottish dataset was not included in any pQTL GWAS studies, and participants of the UKB-PPP dataset were excluded when testing the associations between T1D and GATE scores computed from summary statistics from the UKB-PPP pQTL dataset. Therefore, the learning sets for SNP-protein associations and the test sets for evaluating GATE score-T1D associations were fully independent without overlapping samples. To address potential confounding due to direct effects of *trans*-pQTLs on T1D, firstly, I excluded the HLA region that has a large direct effect on T1D from GATE score computation; secondly, I filtered out GATE scores with less than five effective *trans*-pQTLs when declaring putative core proteins in Chapter 3. This criterion ensures that a GATE score is not dominated by one or more *trans*-pQTLs with a direct effect on T1D not mediated by the protein, and thus minimises potential confounding. This also helps increase the signal-to-noise ratio for GATE score associations with T1D. Furthermore, I examined the GATE score composition by computing the correlations between GATE scores (which indicate sharing of *trans*-pQTLs) and by grouping overlapping *trans*-pQTLs into clumps. Finally, I employed a novel MR approach that explicitly models the direct effects of genetic instruments on the outcome and integrates these effects to calculate the marginal likelihood of the causal effect. This approach controls the type I error rate (i.e., false positive reports for causality) even in the presence of large direct (pleiotropic) effects.

Nonetheless, this thesis also has limitations. Firstly, GATE scores in this thesis reflect genetically predicted levels of soluble proteins in the blood. Because of the complexity in functional mechanisms of soluble proteins and the presence of multiple isoforms, for some proteins, the directions of effects observed in this thesis were inconsistent with what would be expected from known biology and evidence reported in experimental studies. Mechanistic studies are warranted to understand the mechanism by which candidate core proteins lead to the development of T1D at the tissue-specific level. Secondly, the validation study of the putative core proteins in the prospective cohort for progression to IA or T1D presented in Chapter 5 was constrained by insufficient statistical power due to limited sample size and by the limited coverage of the proteome, with only eight of the 49 candidate core proteins having been measured in the Olink Inflammation panel in the TEDDY study. Although transcript levels for all these 49 proteins were available from RNA-seq and microarray datasets, these measurements reflect mRNAs in blood cells, which may differ functionally from circulating proteins and can also be noisy. Large-scale proteomics studies, with broad coverage of circulating proteins in prospective studies with larger sample sizes, are needed to further validate the effects of candidate core proteins in the pathogenesis of T1D.

6.5 Future directions

From the methodological perspective, future QTL studies with larger sample sizes will significantly increase the number of effective *trans*-loci for each GATE score and improve the proportion of molecular trait variance explained by GATE scores. This, in turn, will enhance the statistical power of both GATE and *trans*-QTL-based MR analyses to detect proteins with a causal effect on disease and to better distinguish causality from confounding in the presence of direct (pleiotropic) effects. In the coming years, three large QTL studies will become available: (1) the eQTLGen Phase 2, analysing whole blood transcripts of 43,301 individuals to map gene expression QTLs; (2) the Trans-Omics for Precision Medicine programme, integrating RNA-sequencing and whole genome sequencing from 14,324 samples of diverse ancestries across six tissues and cell types to identify gene expression and RNA splicing QTLs [Orchard et al., 2025]; (3) the UKB-PPP Phase 2, profiling over 5400 proteins with the Olink Explore HT panel in 500,000 UK Biobank participants to map protein QTLs. In addition to using summary

data from larger QTL studies, future GATE analyses should also incorporate *trans*-pQTLs located on sex chromosomes to increase the number of effective *trans*-loci for each GATE score. The GATE analyses presented in this thesis evaluated GATE scores for proteins encoded by genes located on sex chromosomes, however, only autosomal *trans*-pQTLs were aggregated when computing GATE scores. For *trans*-pQTLs on sex chromosomes, sex-specific genotype dosages – 0, 1, 2 for females and either 0, 2 or 0, 1 for males depending on dosage compensation model – should be applied for the computation of GATE scores.

In parallel to the new QTL studies, large-scale disease case-control datasets with extensive phenotypic and genetic data will further improve the statistical power to detect core genes and core proteins and provide additional supportive evidence for existing findings through replication. For example, the Our Future Health cohort aims to recruit five million participants, with health record linkage and genotyping. By June 2025, 1.8 million participants had been enrolled, including 5880 with reports of T1D diagnosis. As summarised in Section 5.4.2, prospective cohorts for IA and T1D with proteomics data or available biosamples in which to generate proteomics measurements will provide valuable resources for validating core genes and core proteins. Additionally, burden tests of rare variants may serve as a complementary approach to strengthen causal assessment.

From the disease phenotype perspective, T1D is a heterogeneous condition, with recent research suggesting it can be classified into six distinct endotypes based on age at onset, genetic and immunological profiles, C-peptide levels, and body weight [Weston et al., 2024]. Applying GATE analysis within each endotype could help disentangle the complex pathophysiology and inform endotype-specific disease-modifying therapeutic strategies. People with T1D are at a higher risk for a series of microvascular and macrovascular complications. In my initial analysis of diabetic kidney disease in the SDRNT1BIO cohort (not included in this thesis), I did not detect putative core genes using the established criteria. However, this analysis had a limited sample size (577 cases and 4809 controls). Applying GATE analysis in large-scale genetic consortiums of diabetes complications will improve power and may identify actionable drug targets for slowing disease progression and preventing complications. Beyond T1D, our research group has applied the GATE approach to several autoimmune diseases, including rheumatoid arthritis [Spiliopoulou et al., 2025], systemic lupus erythematosus [Iakovliev et al., 2025], giant cell arthritis, and autoimmune thyroid disease. Expanding this pipeline to additional autoimmune conditions will help dissect both shared and

disease-specific molecular pathways in autoimmunity, thus ultimately promoting novel drug discovery and repurposing.

From the results interpretation perspective, it remains challenging to elucidate the underlying biology. The main difficulties arise from the following two aspects: (1) the processes by which soluble proteins are produced and released into the bloodstream (e.g., alternative splicing, proteolytic cleavage from membrane-bound form, secretion from other tissues); (2) the tissue-specific roles of soluble proteins in T1D pathogenesis. Addressing these questions will require integrative bioinformatic analyses to understand the spatial context of protein production and circulation, as well as experimental studies to bridge statistical associations with mechanistic insights. Despite showing mainly cross-sectional measured protein level associations, the key findings of this thesis will also be useful for informing prospective evaluations of proteins and constructing prediction models for progression from stage 1 to stage 3 of T1D.

6.6 Conclusion

In conclusion, the novel human genetic evidence provided in this thesis for the pivotal roles of certain proteins in the pathogenesis of T1D is useful in i) confirming the primary importance of the immune checkpoint pathway, ii) highlighting the importance of the innate immune pathway, iii) reinforcing the already accepted central role for T cells in T1D development and progression, iv) supporting the hypothesis of a role for the exocrine pancreas in the primary pathogenesis of T1D, and v) helping prioritise potential drug targets for T1D disease-modifying therapy.

7

Bibliography

- Tahira Akbar, Stuart McGurnaghan, Colin NA Palmer, Shona J Livingstone, John Petrie, John Chalmers, Robert S Lindsay, John A McKnight, Donald WM Pearson, Alan W Patrick, James Walker, Helen C Looker, and Helen M Colhoun. Cohort Profile: Scottish Diabetes Research Network Type 1 Bioresource Study (SDRNT1BIO). *Int J Epidemiol*, 46(3):796–796i, June 2017. ISSN 0300-5771. doi: 10.1093/ije/dyw152.
- Oscar Alcazar, Luis F. Hernandez, Ernesto S. Nakayasu, Carrie D. Nicora, Charles Ansong, Michael J. Muehlbauer, James R. Bain, Ciara J. Myer, Sanjoy K. Bhattacharya, Peter Buchwald, and Midhat H. Abdulreda. Parallel Multi-Omics in High-Risk Subjects for the Identification of Integrated Biomarker Signatures of Type 1 Diabetes. *Biomolecules*, 11(3):383, March 2021. ISSN 2218-273X. doi: 10.3390/biom11030383.
- Mohammad Alhadj Ali, Yuk-Fun Liu, Sefina Arif, Danijela Tatovic, Hina Shariff, Vivienne B. Gibson, Norkhairin Yusuf, Roman Baptista, Martin Eichmann, Nedyalko Petrov, Susanne Heck, Jennie H. M. Yang, Timothy I. M. Tree, Irma Pujol-Autonell, Lorraine Yeo, Lucas Baumard, Rachel Stenson, Alex Howell, Alison Clark, Zoe Boulton, Jake Powrie, Laura Adams, Florence S. Wong, Stephen Luzio, Gareth Dunseath, Kate Green, Alison O’Keefe, Graham Bayly, Natasha Thorogood, Robert Andrews, Nicola Leech, Frank Joseph, Sunil Nair, Susan Seal, HoYee Cheung, Craig Beam, Robert Hills, Mark Peakman, and Colin M. Dayan. Metabolic and immune effects of immunotherapy with proinsulin peptide in human new-onset type 1 diabetes. *Science Translational Medicine*, 9(402):eaaf7779, August 2017. doi: 10.1126/scitranslmed.aaf7779.
- Frederick Allen and Ivan Maillard. Therapeutic Targeting of Notch Signaling: From Cancer to Inflammatory Disorders. *Front. Cell Dev. Biol.*, 9, May 2021. ISSN 2296-634X. doi: 10.3389/fcell.2021.649205.
- Faizah M. Alotaibi, Wei-Ping Min, and James Koropatnick. CD5 blockade, a novel immune checkpoint inhibitor, enhances T cell anti-tumour immunity and delays tumour growth in mice harbouring poorly immunogenic 4T1 breast tumour homografts. *Front Immunol*, 15, February 2024. ISSN 1664-3224. doi: 10.3389/fimmu.2024.1256766.
- P. Ambery, T. W. Donner, N. Biswas, J. Donaldson, J. Parkin, and C. M. Dayan. Efficacy and safety of low-dose oteelixizumab anti-CD3 monoclonal antibody in preserving C-peptide secretion in adolescent type 1 diabetes: DEFEND-2, a randomized, placebo-controlled, double-blind, multi-centre study. *Diabetic Medicine*, 31(4):399–402, 2014. ISSN 1464-5491. doi: 10.1111/dme.12361.

- American Diabetes Association Professional Practice Committee. 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes—2025. *Diabetes Care*, 48(Supplement_1):S27–S49, December 2024. ISSN 0149-5992. doi: 10.2337/dc25-S002.
- Jingnan An, Sisi Ding, Sicheng Li, Lili Sun, Xin Chang, Ziyi Huang, Bin Zhou, Chen Fang, Cuiping Liu, and Xueguang Zhang. Enhancement of the Soluble Form of OX40 and OX40L Costimulatory Molecules but Reduction of the Membrane Form in Type 1 Diabetes (T1D). *J Immunol Res*, 2019:1780567, August 2019. ISSN 2314-8861. doi: 10.1155/2019/1780567.
- AnaptysBio. Anaptys Announces Positive Rosnilimab Data Updated Through Six Months in Robust Phase 2b Trial in RA | AnaptysBio, Inc. <https://ir.anaptysbio.com/news-releases/news-release-details/anaptys-announces-positive-rosnilimab-data-updated-through-six/>, June 2025.
- Warren Anderson, Fariba Barahmand-pour-Whitman, Peter S Linsley, Karen Cerosaletti, Jane H Buckner, and David J Rawlings. PTPN22 R620W gene editing in T cells enhances low-avidity TCR responses. *eLife*, 12:e81577, March 2023. ISSN 2050-084X. doi: 10.7554/eLife.81577.
- Mathieu Angin, Chrystelle Brignone, and Frédéric Triebel. A LAG-3–Specific Agonist Antibody for the Treatment of T Cell–Induced Autoimmune Diseases. *J Immunol*, 204(4):810–818, February 2020. ISSN 0022-1767. doi: 10.4049/jimmunol.1900823.
- Florence Anquetil, Giada Mondanelli, Nathaly Gonzalez, Teresa Rodriguez Calvo, Jose Zapardiel Gonzalo, Lars Krogvold, Knut Dahl-Jørgensen, Benoit Van den Eynde, Ciriana Orabona, Ursula Grohmann, and Matthias G. von Herrath. Loss of IDO1 Expression From Human Pancreatic β -Cells Precedes Their Destruction During the Development of Type 1 Diabetes. *Diabetes*, 67(9):1858–1866, June 2018. ISSN 0012-1797. doi: 10.2337/db17-1281.
- Mohammed Javeed I. Ansari, Alan D. Salama, Tanuja Chitnis, R. Neal Smith, Hideo Yagita, Hisaya Akiba, Tomohide Yamazaki, Miyuki Azuma, Hideyuki Iwai, Samia J. Khoury, Hugh Auchincloss, Jr., and Mohamed H. Sayegh. The Programmed Death-1 (PD-1) Pathway Regulates Autoimmune Diabetes in Nonobese Diabetic (NOD) Mice. *Journal of Experimental Medicine*, 198(1): 63–69, July 2003. ISSN 0022-1007. doi: 10.1084/jem.20022125.

- Andrea Aprilia, Kusworini Handono, Hidayat Sujuti, Akhmad Sabarudin, and Nuning Winaris. sCD163, sCD28, sCD80, and sCTLA-4 as soluble marker candidates for detecting immunosenescence. *Immunity & Ageing*, 21(1):9, January 2024. ISSN 1742-4933. doi: 10.1186/s12979-023-00405-0.
- Ronnie Aronson, Peter A. Gottlieb, Jens S. Christiansen, Thomas W. Donner, Emanuele Bosi, Bruce W. Bode, Paolo Pozzilli, and the DEFEND Investigator Group. Low-Dose Otelixizumab Anti-CD3 Monoclonal Antibody DEFEND-1 Study: Results of the Randomized Phase III Study in Recent-Onset Human Type 1 Diabetes. *Diabetes Care*, 37(10):2746–2754, September 2014. ISSN 0149-5992. doi: 10.2337/dc13-0327.
- Robin Assfalg, Jan Knoop, Kristi L. Hoffman, Markus Pfirrmann, Jose Maria Zapardiel-Gonzalo, Anna Hofelich, Anne Eugster, Marc Weigelt, Claudia Matzke, Julia Reinhardt, Yannick Fuchs, Melanie Bunk, Andreas Weiss, Markus Hippich, Kathrin Halfter, Stefanie M. Hauck, Jörg Hasford, Joseph F. Petrosino, Peter Achenbach, Ezio Bonifacio, and Anette-Gabriele Ziegler. Oral insulin immunotherapy in children at risk for type 1 diabetes in a randomised controlled trial. *Diabetologia*, 64(5):1079–1092, May 2021. ISSN 1432-0428. doi: 10.1007/s00125-020-05376-1.
- Mark A. Atkinson and Raghavendra G. Mirmira. The pathogenic “symphony” in type 1 diabetes: A disorder of the immune system, β cells, and exocrine pancreas. *Cell Metab*, 35(9):1500–1518, September 2023. ISSN 1550-4131. doi: 10.1016/j.cmet.2023.06.018.
- Nicolas Aubert, Simon Brunel, Daniel Olive, and Gilles Marodon. Blockade of HVEM for Prostate Cancer Immunotherapy in Humanized Mice. *Cancers*, 13(12):3009, January 2021. ISSN 2072-6694. doi: 10.3390/cancers13123009.
- Christian Bailly, Xavier Thuru, and Bruno Quesnel. Soluble Programmed Death Ligand-1 (sPD-L1): A Pool of Circulating Proteins Implicated in Health and Diseases. *Cancers (Basel)*, 13(12):3034, June 2021. ISSN 2072-6694. doi: 10.3390/cancers13123034.
- Jeffrey C. Barrett, David G. Clayton, Patrick Concannon, Beena Akolkar, Jason D. Cooper, Henry A. Erlich, Cécile Julier, Grant Morahan, Jørn Nerup, Concepcion Nierras, Vincent Plagnol, Flemming Pociot, Helen Schuilenburg, Deborah J. Smyth, Helen Stevens, John A. Todd, Neil M. Walker, and Stephen S. Rich. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*, 41(6):703–707, June 2009. ISSN 1546-1718. doi: 10.1038/ng.381.

- Karyl S. Barron, Ivona Aksentijevich, Natalie T. Deutch, Deborah L. Stone, Patrycja Hoffmann, Ryan Videgar-Laird, Ariane Soldatos, Jenna Bergerson, Camilo Toro, Cornelia Cudrici, Michele Nehrebecky, Tina Romeo, Anne Jones, Manfred Boehm, Jennifer A. Kanakry, Dimana Dimitrova, Katherine R. Calvo, Hawwa Alao, Devika Kapuria, Gil Ben-Yakov, Dominique C. Pichard, Londa Hathaway, Alessandra Brofferio, Elisa McRae, Natalia Sampaio Moura, Oskar Schnappauf, Sofia Rosenzweig, Theo Heller, Edward W. Cowen, Daniel L. Kastner, and Amanda K. Ombrello. The Spectrum of the Deficiency of Adenosine Deaminase 2: An Observational Analysis of a 60 Patient Cohort. *Front. Immunol.*, 12, January 2022. ISSN 1664-3224. doi: 10.3389/fimmu.2021.811473.
- Manuela Battaglia, Mark S. Anderson, Jane H. Buckner, Susan M. Geyer, Peter A. Gottlieb, Thomas W. H. Kay, Åke Lernmark, Sarah Muller, Alberto Pugliese, Bart O. Roep, Carla J. Greenbaum, and Mark Peakman. Understanding and preventing type 1 diabetes through the unique working model of TrialNet. *Diabetologia*, 60(11):2139–2147, November 2017. ISSN 1432-0428. doi: 10.1007/s00125-017-4384-2.
- Craig A. Beam, Colleen MacCallum, Kevan C. Herold, Diane K. Wherrett, Jerry Palmer, and Johnny Ludvigsson. GAD vaccine reduces insulin loss in recently diagnosed type 1 diabetes: Findings from a Bayesian meta-analysis. *Diabetologia*, 60(1):43–49, January 2017. ISSN 0012-186X. doi: 10.1007/s00125-016-4122-1.
- G. I. Bell, S. Horita, and J. H. Karam. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes*, 33(2):176–183, February 1984. ISSN 0012-1797, 0012-1797. doi: 10.2337/diabetes.33.2.176.
- Malin Béteky, Jeanette Wahlberg, and Johnny Ludvigsson. Infections and antibiotic use in early childhood have limited importance in developing manifest type 1 diabetes – The ABIS cohort study. *Front Endocrinol (Lausanne)*, 16:1529447, February 2025. ISSN 1664-2392. doi: 10.3389/fendo.2025.1529447.
- S. T. Bennett, A. M. Lucassen, S. C. L. Gough, E. E. Powell, D. E. Undlien, L. E. Pritchard, M. E. Merriman, Y. Kawaguchi, M. J. Dronsfield, F. Pociot, J. Nerup, N. Bouzekri, A. Cambon-Thomsen, K. S. Rønningen, A. H. Barnett, S. C. Bain, and J. A. Todd. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat Genet*, 9(3):284–292, March 1995. ISSN 1546-1718. doi: 10.1038/ng0395-284.

- Alix Besançon, Tania Goncalves, Fabrice Valette, Caroline Mary, Bernard Vanhove, Lucienne Chatenoud, and Sylvaine You. A selective CD28 antagonist and rapamycin synergise to protect against spontaneous autoimmune diabetes in NOD mice. *Diabetologia*, 61(8):1811–1816, August 2018. ISSN 1432-0428. doi: 10.1007/s00125-018-4638-7.
- Maria Bettini, Andrea L. Szymczak-Workman, Karen Forbes, Ashley H. Castellaw, Mark Selby, Xiaoyu Pan, Charles G. Drake, Alan J. Korman, and Dario AA Vignali. Accelerated autoimmune diabetes in the absence of LAG-3. *J Immunol*, 187(7):3493, August 2011. doi: 10.4049/jimmunol.1100714.
- Vikash J. Bhagwandin, Leola W. T. Hau, Jon Mallen-St. Clair, Paul J. Wolters, and George H. Caughey. Structure and Activity of Human Pancreasin, a Novel Tryptic Serine Peptidase Expressed Primarily by the Pancreas*. *J Biol Chem*, 278(5):3363–3371, January 2003. ISSN 0021-9258. doi: 10.1074/jbc.M209353200.
- Jan Bláha, Tereza Skálová, Barbora Kalousková, Ondřej Skořepa, Denis Cmunť, Valéria Grobárová, Samuel Pazicky, Edit Poláchová, Celeste Abreu, Jan Stránský, Tomáš Koval', Jarmila Dušková, Yuguang Zhao, Karl Harlos, Jindřich Hašek, Jan Dohnálek, and Ondřej Vaněk. Structure of the human NK cell NKR-P1:LLT1 receptor:ligand complex reveals clustering in the immune synapse. *Nat Commun*, 13(1):5022, August 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-32577-6.
- Bruno Blanchi, Marion Taurand, Claire Colace, Sofia Thomaidou, Charlotte Audeoud, Federica Fantuzzi, Toshiaki Sawatani, Sevda Gheibi, Joan Sabadell-Basallote, Fransje W.J. Boot, Thibault Chantier, Aline Piet, Charlotte Cavanihac, Marion Pilette, Adélie Balguerie, Hamza Olleik, Françoise Carlotti, Miriam Ejarque, Malin Fex, Hindrik Mulder, Miriam Cnop, Decio L. Eizirik, Ouardane Jouannot, Anne-Lise Gaffuri, Paul Czernichow, Arnaud Zaldumbide, Raphaël Scharfmann, and Philippe Ravassard. EndoC-βH5 cells are storable and ready-to-use human pancreatic beta cells with physiological insulin secretion. *Mol Metab*, 76:101772, July 2023. ISSN 2212-8778. doi: 10.1016/j.molmet.2023.101772.
- Jeffrey A. Bluestone, Jane H. Buckner, Mark Fitch, Stephen E. Gitelman, Shipra Gupta, Marc K. Hellerstein, Kevan C. Herold, Angela Lares, Michael R. Lee, Kelvin Li, Weihong Liu, S. Alice Long, Lisa M. Masiello, Vinh Nguyen,

- Amy L. Putnam, Mary Rieck, Peter H. Sayre, and Qizhi Tang. Type 1 diabetes immunotherapy using polyclonal regulatory T cells. *Science Translational Medicine*, 7(315):315ra189–315ra189, November 2015. doi: 10.1126/scitranslmed.aad4134.
- Ezio Bonifacio, Anette-G. Ziegler, Georgeanna Klingensmith, Edith Schober, Polly J. Bingley, Marietta Rottenkolber, Anke Theil, Anne Eugster, Ramona Puff, Claudia Peplow, Florian Buettner, Karin Lange, Jörg Hasford, Peter Achenbach, and for the Pre-POINT Study Group. Effects of High-Dose Oral Insulin on Immune Responses in Children at High Risk for Type 1 Diabetes: The Pre-POINT Randomized Clinical Trial. *JAMA*, 313(15):1541–1549, April 2015. ISSN 0098-7484. doi: 10.1001/jama.2015.2928.
- Amélie Bonnefond, Ranjit Unnikrishnan, Alessandro Doria, Martine Vaxillaire, Rohit N. Kulkarni, Viswanathan Mohan, Vincenzo Trischitta, and Philippe Froguel. Monogenic diabetes. *Nat Rev Dis Primers*, 9(1):12, March 2023. ISSN 2056-676X. doi: 10.1038/s41572-023-00421-w.
- Nunzio Bottini, Torkel Vang, Francesco Cucca, and Tomas Mustelin. Role of *PTPN22* in type 1 diabetes and other autoimmune diseases. *Seminars in Immunology*, 18(4):207–213, August 2006. ISSN 1044-5323. doi: 10.1016/j.smim.2006.03.008.
- Evan A. Boyle, Yang I. Li, and Jonathan K. Pritchard. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, 169(7):1177–1186, June 2017. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2017.05.038.
- Jonathan P. Bradfield, Hui-Qi Qu, Kai Wang, Haitao Zhang, Patrick M. Sleiman, Cecilia E. Kim, Frank D. Mentch, Haijun Qiu, Joseph T. Glessner, Kelly A. Thomas, Edward C. Frackelton, Rosetta M. Chiavacci, Marcin Imielinski, Dimitri S. Monos, Rahul Pandey, Marina Bakay, Struan F. A. Grant, Constantin Polychronakos, and Hakon Hakonarson. A Genome-Wide Meta-Analysis of Six Type 1 Diabetes Cohorts Identifies Multiple Associated Loci. *PLOS Genetics*, 7(9):e1002293, September 2011. ISSN 1553-7404. doi: 10.1371/journal.pgen.1002293.
- Damien Bresson, Georgia Fouteri, Yulia Manenkova, Michael Croft, and Matthias von Herrath. Antigen-specific prevention of type 1 diabetes in NOD mice is ameliorated by OX40 agonist treatment. *Journal of Autoimmunity*, 37(4):342–351, December 2011. ISSN 0896-8411. doi: 10.1016/j.jaut.2011.10.001.

- Alberto Brini and Edwin R. van den Heuvel. Missing Data Imputation with High-Dimensional Data. *The American Statistician*, April 2024. ISSN 0003-1305.
- Marion H. Brown and Erica Lacey. A Ligand for CD5 is CD5. *J Immunol*, 185 (10):6068–6074, November 2010. ISSN 0022-1767. doi: 10.4049/jimmunol.0903823.
- Brittany S. Bruggeman and Desmond A. Schatz. Type 1 Diabetes: A Disorder of the Exocrine and Endocrine Pancreas. *J Cell Immunol*, Volume 5(Issue 4): 120–126, November 2023. ISSN 2689-2812. doi: 10.33696/immunology.5.177.
- Sara Bruzzaniti, Erica Piemonte, Dario Bruzzese, Maria Teresa Lepore, Rocky Strollo, Lavinia Izzo, Francesca Di Candia, Adriana Franzese, Maurizio Bifulco, Enza Mozzillo, Johnny Ludvigsson, Giuseppe Matarese, and Mario Galgani. Progression of type 1 diabetes is associated with high levels of soluble PD-1 in islet autoantibody-positive children. *Diabetologia*, 67(4):714–723, April 2024. ISSN 1432-0428. doi: 10.1007/s00125-023-06075-3.
- Stephen Burgess, George Davey Smith, Neil M. Davies, Frank Dudbridge, Dipender Gill, M. Maria Glymour, Fernando P. Hartwig, Zoltán Kutalik, Michael V. Holmes, Cosetta Minelli, Jean V. Morrison, Wei Pan, Caroline L. Relton, and Evropi Theodoratou. Guidelines for performing Mendelian randomization investigations: Update for summer 2023. *Wellcome Open Research*, 4:186, August 2023. doi: 10.12688/wellcomeopenres.15555.3.
- Stephen Burgess, Benjamin Woolf, Amy M. Mason, Mika Ala-Korpela, and Dipender Gill. Addressing the credibility crisis in Mendelian randomization. *BMC Medicine*, 22(1):374, September 2024. ISSN 1741-7015. doi: 10.1186/s12916-024-03607-5.
- Paul R. Burton, David G. Clayton, Lon R. Cardon, Nick Craddock, Panos Deloukas, Audrey Duncanson, Dominic P. Kwiatkowski, Mark I. McCarthy, Willem H. Ouwehand, Nilesh J. Samani, John A. Todd, Peter Donnelly, Jeffrey C. Barrett, Paul R. Burton, Dan Davison, Peter Donnelly, Doug Easton, David Evans, Hin-Tak Leung, Jonathan L. Marchini, Andrew P. Morris, Chris C. A. Spencer, Martin D. Tobin, Lon R. Cardon, David G. Clayton, Antony P. Attwood, James P. Boorman, Barbara Cant, Ursula Everson, Judith M. Hussey, Jennifer D. Jolley, Alexandra S. Knight, Kerstin Koch, Elizabeth Meech, Sarah Nutland, Christopher V. Prowse, Helen E. Stevens, Niall C. Taylor, Graham R. Walters, Neil M. Walker, Nicholas A. Watkins, Thilo Winzer,

John A. Todd, Willem H. Ouwehand, Richard W. Jones, Wendy L. McArdle, Susan M. Ring, David P. Strachan, Marcus Pembrey, Gerome Breen, David St Clair, Sian Caesar, Katherine Gordon-Smith, Lisa Jones, Christine Fraser, Elaine K. Green, Detelina Grozeva, Marian L. Hamshire, Peter A. Holmans, Ian R. Jones, George Kirov, Valentina Moskvina, Ivan Nikolov, Michael C. O'Donovan, Michael J. Owen, Nick Craddock, David A. Collier, Amanda Elkin, Anne Farmer, Richard Williamson, Peter McGuffin, Allan H. Young, I. Nicol Ferrier, Stephen G. Ball, Anthony J. Balmforth, Jennifer H. Barrett, D. Timothy Bishop, Mark M. Iles, Azhar Maqbool, Nadira Yuldasheva, Alistair S. Hall, Peter S. Braund, Paul R. Burton, Richard J. Dixon, Massimo Mangino, Suzanne Stevens, Martin D. Tobin, John R. Thompson, Nilesch J. Samani, Francesca Bredin, Mark Tremelling, Miles Parkes, Hazel Drummond, Charles W. Lees, Elaine R. Nimmo, Jack Satsangi, Sheila A. Fisher, Alastair Forbes, Cathryn M. Lewis, Clive M. Onnie, Natalie J. Prescott, Jeremy Sanderson, Christopher G. Mathew, Jamie Barbour, M. Khalid Mohiuddin, Catherine E. Todhunter, John C. Mansfield, Tariq Ahmad, Fraser R. Cummings, Derek P. Jewell, John Webster, Morris J. Brown, David G. Clayton, G. Mark Lathrop, John Connell, Anna Dominiczak, Nilesch J. Samani, Carolina A. Braga Marcano, Beverley Burke, Richard Dobson, Johannie Gungadoo, Kate L. Lee, Patricia B. Munroe, Stephen J. Newhouse, Abiodun Onipinla, Chris Wallace, Mingzhan Xue, Mark Caulfield, Martin Farrall, Anne Barton, The Biologics in RA Genetics and Genomics (BRAGGS), Ian N. Bruce, Hannah Donovan, Steve Eyre, Paul D. Gilbert, Samantha L. Hider, Anne M. Hinks, Sally L. John, Catherine Potter, Alan J. Silman, Deborah P. M. Symmons, Wendy Thomson, Jane Worthington, David G. Clayton, David B. Dunger, Sarah Nutland, Helen E. Stevens, Neil M. Walker, Barry Widmer, John A. Todd, Timothy M. Frayling, Rachel M. Freathy, Hana Lango, John R. B. Perry, Beverley M. Shields, Michael N. Weedon, Andrew T. Hattersley, Graham A. Hitman, Mark Walker, Kate S. Elliott, Christopher J. Groves, Cecilia M. Lindgren, Nigel W. Rayner, Nicholas J. Timpson, Eleftheria Zeggini, Mark I. McCarthy, Melanie Newport, Giorgio Sirugo, Emily Lyons, Fredrik Vannberg, Adrian V. S. Hill, Linda A. Bradbury, Claire Farrar, Jennifer J. Pointon, Paul Wordsworth, Matthew A. Brown, Jayne A. Franklyn, Joanne M. Heward, Matthew J. Simmonds, Stephen C. L. Gough, Sheila Seal, Breast Cancer Susceptibility Collaboration (UK), Michael R. Stratton, Nazneen Rahman, Maria Ban, An Goris, Stephen J. Sawcer, Alastair Compston, David Conway, Muminatou Jallow, Melanie Newport, Giorgio Sirugo, Kirk A. Rockett, Dominic P. Kwiatkowski, Suzannah J. Bumpstead, Amy Chaney, Kate

- Downes, Mohammed J. R. Ghor, Rhian Gwilliam, Sarah E. Hunt, Michael Inouye, Andrew Keniry, Emma King, Ralph McGinnis, Simon Potter, Rathi Ravindrarajah, Pamela Whittaker, Claire Widden, David Withers, Panos Deloukas, Hin-Tak Leung, Sarah Nutland, Helen E. Stevens, Neil M. Walker, John A. Todd, Doug Easton, David G. Clayton, Paul R. Burton, Martin D. Tobin, Jeffrey C. Barrett, David Evans, Andrew P. Morris, Lon R. Cardon, Niall J. Cardin, Dan Davison, Teresa Ferreira, Joanne Pereira-Gale, Ingileif B. Hallgrimsdóttir, Bryan N. Howie, Jonathan L. Marchini, Chris C. A. Spencer, Zhan Su, Yik Ying Teo, Damjan Vukcevic, Peter Donnelly, David Bentley, Matthew A. Brown, Lon R. Cardon, Mark Caulfield, David G. Clayton, Alistair Compston, Nick Craddock, Panos Deloukas, Peter Donnelly, Martin Farrall, Stephen C. L. Gough, Alistair S. Hall, Andrew T. Hattersley, Adrian V. S. Hill, Dominic P. Kwiatkowski, Christopher G. Mathew, Mark I. McCarthy, Willem H. Ouwehand, Miles Parkes, Marcus Pembrey, Nazneen Rahman, Nilesh J. Samani, Michael R. Stratton, John A. Todd, Jane Worthington, The Wellcome Trust Case Control Consortium, Management Committee, Data and Analysis Committee, UK Blood Services and University of Cambridge Controls, 1958 Birth Cohort Controls, Bipolar Disorder, Coronary Artery Disease, Crohn's Disease, Hypertension, Rheumatoid Arthritis, Type 1 Diabetes, Type 2 Diabetes, Tuberculosis, Ankylosing Spondylitis, Autoimmune Thyroid Disease, Breast Cancer, Multiple Sclerosis, Gambian Controls, Genotyping DNA, Data QC and Informatics, Statistics, and Primary Investigators. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447(7145):661–678, June 2007. ISSN 1476-4687. doi: 10.1038/nature05911.
- Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T. Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O'Connell, Adrian Cortes, Samantha Welsh, Alan Young, Mark Effingham, Gil McVean, Stephen Leslie, Naomi Allen, Peter Donnelly, and Jonathan Marchini. The UK Biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726):203–209, October 2018. ISSN 1476-4687. doi: 10.1038/s41586-018-0579-z.
- Per-Ola Carlsson, Xiaomeng Hu, Hanne Scholz, Sofie Ingvas, Torbjörn Lundgren, Tim Scholz, Olof Eriksson, Per Liss, Di Yu, Tobias Deuse, Olle Korsgren, and Sonja Schrepfer. Survival of Transplanted Allogeneic Beta Cells with No Immunosuppression. *New England Journal of Medicine*, 0(0), August 2025. ISSN 0028-4793. doi: 10.1056/NEJMoa2503822.

- Bendix Carstensen, Pernille Falberg Rønn, and Marit Eika Jørgensen. Prevalence, incidence and mortality of type 1 and type 2 diabetes in Denmark 1996–2016. *BMJ Open Diab Res Care*, 8(1), May 2020. ISSN 2052-4897. doi: 10.1136/bmjdr-2019-001071.
- Rosaura Casas, Fabrícia Dietrich, Hugo Barcenilla, Beatriz Tavira, Jeanette Wahlberg, Peter Achenbach, and Johnny Ludvigsson. Glutamic Acid Decarboxylase Injection Into Lymph Nodes: Beta Cell Function and Immune Responses in Recent Onset Type 1 Diabetes Patients. *Front. Immunol.*, 11, October 2020. ISSN 1664-3224. doi: 10.3389/fimmu.2020.564921.
- Jasmine I. Caulfield, Lilach Aizenbud, Ana Luisa Perdigoto, Eric Meffre, Lucia Jilaveanu, Dominika A. Michalek, Stephen S. Rich, Yariv Aizenbud, Adebowale Adeniran, Kevan C. Herold, Matthew R. Austin, and Harriet Kluger. Germline genetic variants are associated with development of insulin-dependent diabetes in cancer patients treated with immune checkpoint inhibitors. *J Immunother Cancer*, 11(3), March 2023. ISSN 2051-1426. doi: 10.1136/jitc-2022-006570.
- Zijing Chen, Shawna Downing, and Emmanuel S. Tzanakakis. Four Decades After the Discovery of Regenerating Islet-Derived (Reg) Proteins: Current Understanding and Challenges. *Frontiers in Cell and Developmental Biology*, 7, October 2019. ISSN 2296-634X. doi: 10.3389/fcell.2019.00235.
- Joshua Chiou, Ryan J. Geusz, Mei-Lin Okino, Jee Yun Han, Michael Miller, Rebecca Melton, Elisha Beebe, Paola Benaglio, Serina Huang, Katha Korgaonkar, Sandra Heller, Alexander Kleger, Sebastian Preissl, David U. Gorkin, Maïke Sander, and Kyle J. Gaulton. Interpreting type 1 diabetes risk with genetics and single-cell epigenomics. *Nature*, 594(7863):398–402, June 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-03552-w.
- Yun Kyung Cho and Chang Hee Jung. Immune-Checkpoint Inhibitors-Induced Type 1 Diabetes Mellitus: From Its Molecular Mechanisms to Clinical Practice. *Diabetes Metab J*, 47(6):757–766, November 2023. ISSN 2233-6079, 2233-6087. doi: 10.4093/dmj.2023.0072.
- Feng-Cheng Chou, Shing-Jia Shieh, and Huey-Kang Sytwu. Attenuation of Th1 response through galectin-9 and T-cell Ig mucin 3 interaction inhibits autoimmune diabetes in NOD mice. *Eur J Immunol*, 39(9):2403–2411, 2009. ISSN 1521-4141. doi: 10.1002/eji.200839177.

- Panagiotis F. Christopoulos, Torleif T. Gjølborg, Stig Krüger, Guttorm Haraldsen, Jan Terje Andersen, and Eirik Sundlisæter. Targeting the Notch Signaling Pathway in Chronic Inflammatory Diseases. *Front Immunol*, 12:668207, April 2021. ISSN 1664-3224. doi: 10.3389/fimmu.2021.668207.
- Maria C. Costanzo, Laura W. Harris, Yue Ji, Aoife McMahon, Noël P. Burt, and Jason Flannick. Realizing the promise of genome-wide association studies for effector gene prediction. *Nat Genet*, 57(7):1578–1587, July 2025. ISSN 1546-1718. doi: 10.1038/s41588-025-02210-5.
- Adam P. Curnock, Giovanna Bossi, Jyothi Kumaran, Lindsay J. Bawden, Rita Figueiredo, Rajeevkumar Tawar, Katherine Wiseman, Emma Henderson, Sec Julie Hoong, Veronica Gonzalez, Hemza Ghadbane, David E. O. Knight, Ronan O'Dwyer, David X. Overton, Christina M. Lucato, Nicola M. G. Smith, Carlos R. Reis, Keith Page, Lorraine M. Whaley, Michelle L. McCully, Stephen Hearty, Tara M. Mahon, and Peter Weber. Cell-targeted PD-1 agonists that mimic PD-L1 are potent T cell inhibitors. *JCI Insight*, 6(20), October 2021. ISSN 0021-9738. doi: 10.1172/jci.insight.152468.
- Suya Dai, Ru Jia, Xiao Zhang, Qiwen Fang, and Lijuan Huang. The PD-1/PD-Ls pathway and autoimmune diseases. *Cellular Immunology*, 290(1):72–79, July 2014. ISSN 00088749. doi: 10.1016/j.cellimm.2014.05.006.
- Ali Dalloul. CD5: A safeguard against autoimmunity and a shield for cancer cells. *Autoimmunity Reviews*, 8(4):349–353, February 2009. ISSN 1568-9972. doi: 10.1016/j.autrev.2008.11.007.
- Randall S. Davis. Roles for the FCRL6 Immunoreceptor in Tumor Immunology. *Front Immunol*, 11, October 2020. ISSN 1664-3224. doi: 10.3389/fimmu.2020.575175.
- Geert D'Haens, Laurent Peyrin-Biroulet, Daniel J. B. Marks, Edoardo Lisi, Lia Liefwaard, Andrew Beaton, Naren Srinivasan, Gerben Bouma, Naveen Prasad, Raymond Cameron, Zeid Kayali, Ruth Tarzi, Stephen Hanauer, and William J. Sandborn. A randomised, double-blind, placebo-controlled study of the LAG-3-depleting monoclonal antibody GSK2831781 in patients with active ulcerative colitis. *Aliment Pharmacol Ther*, 58(3):283–296, 2023. ISSN 1365-2036. doi: 10.1111/apt.17557.

- Jiawen Dong, Kaibo Zhang, Jinhui Hong, Jialu Ye, Leping Wang, Qiang Ke, Yuyang Ye, Keting Jin, Yalun Chen, and Jing Wu. The multiple functions of CD5 in diseases related to immune disorders. *Annals of Medicine*, 57(1):2519682, December 2025. ISSN 0785-3890. doi: 10.1080/07853890.2025.2519682.
- Shen Dong, Kamir J. Hiam-Galvez, Cody T. Mowery, Kevan C. Herold, Stephen E. Gitelman, Jonathan H. Esensten, Weihong Liu, Angela P. Lares, Ashley S. Leinbach, Michael Lee, Vinh Nguyen, Stanley J. Tamaki, Whitney Tamaki, Courtney M. Tamaki, Morvarid Mehdizadeh, Amy L. Putnam, Matthew H. Spitzer, Chun Jimmie Ye, Qizhi Tang, and Jeffrey A. Bluestone. The effect of low-dose IL-2 and Treg adoptive cell therapy in patients with type 1 diabetes. *JCI Insight*, 6(18), September 2021. ISSN 0021-9738. doi: 10.1172/jci.insight.147474.
- I. Durinovic-Belló, R. P. Wu, V. H. Gersuk, S. Sanda, H. G. Shilling, and G. T. Nepom. Insulin gene VNTR genotype associates with frequency and phenotype of the autoimmune response to proinsulin. *Genes Immun*, 11(2):188–193, March 2010. ISSN 1476-5470. doi: 10.1038/gene.2009.108.
- Douglas P. Dyer. Understanding the mechanisms that facilitate specificity, not redundancy, of chemokine-mediated leukocyte recruitment. *Immunology*, 160(4):336–344, 2020. ISSN 1365-2567. doi: 10.1111/imm.13200.
- Natalie M. Edner, Gianluca Carlesso, James S. Rush, and Lucy S. K. Walker. Targeting co-stimulatory molecules in autoimmune disease. *Nat Rev Drug Discov*, 19(12):860–883, December 2020. ISSN 1474-1784. doi: 10.1038/s41573-020-0081-9.
- Seham Elabd and Ismail Sabry. Two Birds with One Stone: Possible Dual-Role of Oxytocin in the Treatment of Diabetes and Osteoporosis. *Front. Endocrinol.*, 6, August 2015. ISSN 1664-2392. doi: 10.3389/fendo.2015.00121.
- Grimur Hjorleifsson Eldjarn, Egil Ferkingstad, Sigrun H. Lund, Hannes Helgason, Olafur Th Magnusson, Kristbjorg Gunnarsdottir, Thorunn A. Olafsdottir, Bjarni V. Halldorsson, Pall I. Olason, Florian Zink, Sigurjon A. Gudjonsson, Gardar Sveinbjornsson, Magnus I. Magnusson, Agnar Helgason, Asmundur Oddsson, Gisli H. Halldorsson, Magnus K. Magnusson, Saedis Saevarsdottir, Thjodbjorg Eiriksdoottir, Gisli Masson, Hreinn Stefansson, Ingileif Jonsdottir, Hilma Holm, Thorunn Rafnar, Pall Melsted, Jona Saemundsdottir, Gudmundur L. Norddahl, Gudmar Thorleifsson, Magnus O. Ulfarsson, Daniel F. Gudbjartsson, Unnur Thorsteinsdottir, Patrick Sulem, and Kari Stefansson.

Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature*, 622(7982):348–358, October 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06563-x.

Joanne Ellis, Daniel J.B. Marks, Naren Srinivasan, Christine Barrett, Thomas G. Hopkins, Anna Richards, Rainard Fuhr, Muna Albayaty, Martin Coenen, Lia Liefwaard, Karen Leavens, Katherine L. Nevin, Shuo Tang, Stephen A. Hughes, Léa Fortunato, Ken Edwards, Yi Cui, Rabia Anselm, Christopher J. Delves, Emilie Charles, Maria Feeney, Thomas M. Webb, Sara J. Brett, Tim S. Schmidt, John Stone, Caroline O.S. Savage, Nicolas Wisniacki, and Ruth M. Tarzi. Depletion of LAG-3+ T Cells Translated to Pharmacology and Improvement in Psoriasis Disease Activity: A Phase I Randomized Study of mAb GSK2831781. *Clinical Pharmacology & Therapeutics*, 109(5):1293–1303, 2021. ISSN 1532-6535. doi: 10.1002/cpt.2091.

Christoph H. Emmerich, Lorena Martinez Gamboa, Martine C. J. Hofmann, Marc Bonin-Andresen, Olga Arbach, Pascal Schendel, Björn Gerlach, Katja Hempel, Anton Bespalov, Ulrich Dirnagl, and Michael J. Parnham. Improving target assessment in biomedical research: The GOT-IT recommendations. *Nat Rev Drug Discov*, 20(1):64–81, January 2021. ISSN 1474-1784. doi: 10.1038/s41573-020-0087-3.

Young Sil Eom, A.-Ryeong Gwon, Kyung Min Kwak, Jin-Young Youn, Heekyoung Park, Kwang-Won Kim, and Byung-Joon Kim. Notch1 Has an Important Role in β -Cell Mass Determination and Development of Diabetes. *Diabetes Metab J*, 45(1):86–96, January 2021. ISSN 2233-6087. doi: 10.4093/dmj.2019.0160.

Henry Erlich, Ana Maria Valdes, Janelle Noble, Joyce A. Carlson, Mike Varney, Pat Concannon, Josyf C. Mychaleckyj, John A. Todd, Persia Bonella, Anna Lisa Fear, Eva Lavant, Anthony Louey, Priscilla Moonsamy, and for the Type 1 Diabetes Genetics Consortium. HLA DR-DQ Haplotypes and Genotypes and Type 1 Diabetes Risk: Analysis of the Type 1 Diabetes Genetics Consortium Families. *Diabetes*, 57(4):1084–1092, April 2008. ISSN 0012-1797. doi: 10.2337/db07-1331.

Susanna C. Fagerholm, Carla Guenther, Marc Lloret Asens, Terhi Savinko, and Liisa M. Uotila. Beta2-Integrins and Interacting Proteins in Leukocyte Trafficking, Immune Suppression, and Immunodeficiency Disease. *Front. Immunol.*, 10, February 2019. ISSN 1664-3224. doi: 10.3389/fimmu.2019.00254.

- Marika Falcone and Georgia Fousteri. Role of the PD-1/PD-L1 Dyad in the Maintenance of Pancreatic Immune Tolerance for Prevention of Type 1 Diabetes. *Front Endocrinol (Lausanne)*, 11:569, August 2020. ISSN 1664-2392. doi: 10.3389/fendo.2020.00569.
- Rudolf S. N. Fehrmann, Ritsert C. Jansen, Jan H. Veldink, Harm-Jan Westra, Danny Arends, Marc Jan Bonder, Jingyuan Fu, Patrick Deelen, Harry J. M. Groen, Asia Smolonska, Rinse K. Weersma, Robert M. W. Hofstra, Wim A. Buurman, Sander Rensen, Marcel G. M. Wolfs, Mathieu Platteel, Alexandra Zhernakova, Clara C. Elbers, Eleanora M. Festen, Gosia Trynka, Marten H. Hofker, Christiaan G. J. Saris, Roel A. Ophoff, Leonard H. van den Berg, David A. van Heel, Cisca Wijmenga, Gerard J. te Meerman, and Lude Franke. Trans-eQTLs Reveal That Independent Genetic Variants Associated with a Complex Phenotype Converge on Intermediate Genes, with a Major Role for the HLA. *PLoS Genet*, 7(8):e1002197, August 2011. ISSN 1553-7404. doi: 10.1371/journal.pgen.1002197.
- Egil Ferkingstad, Patrick Sulem, Bjarni A. Atlason, Gardar Sveinbjornsson, Magnus I. Magnusson, Edda L. Styrnisdottir, Kristbjorg Gunnarsdottir, Agnar Helgason, Asmundur Oddsson, Bjarni V. Halldorsson, Brynjar O. Jensson, Florian Zink, Gisli H. Halldorsson, Gisli Masson, Gudny A. Arnadottir, Hildigunnur Katrinardottir, Kristinn Juliusson, Magnus K. Magnusson, Olafur Th Magnusson, Run Fridriksdottir, Saedis Saevarsdottir, Sigurjon A. Gudjonsson, Simon N. Stacey, Solvi Rognvaldsson, Thjodbjorg Eiriksdottir, Thorunn A. Olafsdottir, Valgerdur Steinthorsdottir, Vinicius Tragante, Magnus O. Ulfarsson, Hreinn Stefansson, Ingileif Jonsdottir, Hilma Holm, Thorunn Rafnar, Pall Melsted, Jona Saemundsdottir, Gudmundur L. Norddahl, Sigrun H. Lund, Daniel F. Gudbjartsson, Unnur Thorsteinsdottir, and Kari Stefansson. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*, 53(12):1712–1721, December 2021. ISSN 1546-1718. doi: 10.1038/s41588-021-00978-w.
- G Feutren, R Assan, G Karsenty, H Du Rostu, J Sirmai, L Papoz, B Vialettes, P Vexiau, M Rodier, A Lallemand, and J.-F. Bach. CYCLOSPORIN INCREASES THE RATE AND LENGTH OF REMISSIONS IN INSULIN-DEPENDENT DIABETES OF RECENT ONSET. *The Lancet*, 328(8499):119–124, July 1986. ISSN 01406736. doi: 10.1016/S0140-6736(86)91943-4.

- B. A. Fisher, J. van de Wiel, J. D. Turner, Y. S. Liu, T. Hosack, J. Young, S. Al-Taei, A. Baranskaya, F. Humby, J. McMurray, S. Gilbert, C. Shave, D. H. Gardner, S. Nayar, T. Cox, R. Muthinti, A. Recchioni, M. C. Maybury, P. de Pablo, S. Bowman, A. Rowe, C. D. Buckley, A. Richter, M. J. Loza, S. Marciniak, W. F. Ng, A. Pratt, S. Travis, V. Homer, and A. Filer. POS0094 PHARMACODYNAMIC ACTIVITY OF JNJ-67484703 IN RHEUMATOID ARTHRITIS, ULCERATIVE COLITIS AND SJOGREN'S DISEASE (PARIS): RESULTS OF A PHASE II PROOF OF BIOLOGY TRIAL. *Ann Rheum Dis*, 84(Supplement 1):389–390, June 2025. ISSN 0003-4967. doi: 10.1016/j.ard.2025.05.482.
- Vincenzo Forgetta, Despoina Manousaki, Roman Istomine, Stephanie Ross, Marie-Catherine Tessier, Luc Marchand, Min Li, Hui-Qi Qu, Jonathan P. Bradfield, Struan F.A. Grant, Hakon Hakonarson, the DCCT/EDIC Research Group, Andrew D. Paterson, Ciriaco Piccirillo, Constantin Polychronakos, and J. Brent Richards. Rare Genetic Variants of Large Effect Influence Risk of Type 1 Diabetes. *Diabetes*, 69(4):784–795, January 2020. ISSN 0012-1797. doi: 10.2337/db19-0831.
- Gregory P. Forlenza, Jennifer McVean, Roy W. Beck, Colleen Bauza, Ryan Bailey, Bruce Buckingham, Linda A. DiMeglio, Jennifer L. Sherr, Mark Clements, Anna Neyman, Carmella Evans-Molina, Emily K. Sims, Laurel H. Messer, Laya Ekhlaspour, Ryan McDonough, Michelle Van Name, Diana Rojas, Shannon Beasley, Stephanie DuBose, Craig Kollman, Antoinette Moran, and CLVer Study Group. Effect of Verapamil on Pancreatic Beta Cell Function in Newly Diagnosed Pediatric Type 1 Diabetes: A Randomized Clinical Trial. *JAMA*, 329(12):990–999, March 2023. ISSN 0098-7484. doi: 10.1001/jama.2023.2064.
- Timothy P. Foster, Brittany Bruggeman, Martha Campbell-Thompson, Mark A. Atkinson, Michael J. Haller, and Desmond A. Schatz. EXOCRINE PANCREAS DYSFUNCTION IN TYPE 1 DIABETES. *Endocr Pract*, 26(12):1505–1513, December 2020. ISSN 1530-891X. doi: 10.4158/EP-2020-0295.
- Yannick F. Fuchs, Virag Sharma, Anne Eugster, Gloria Kraus, Robert Morgenstern, Andreas Dahl, Susanne Reinhardt, Andreas Petzold, Annett Lindner, Doreen Löbel, and Ezio Bonifacio. Gene Expression-Based Identification of Antigen-Responsive CD8+ T Cells on a Single-Cell Level. *Front Immunol*, 10, November 2019. ISSN 1664-3224. doi: 10.3389/fimmu.2019.02568.

- Jason Gaglia and Stephan Kissler. Anti-CD3 antibody for the prevention of type 1 diabetes - a story of perseverance. *Biochemistry*, 58(40):4107, September 2019. doi: 10.1021/acs.biochem.9b00707.
- Katarzyna Anna Gajewska, Regien Biesma, Seamus Sreenan, and Kathleen Bennett. Prevalence and incidence of type 1 diabetes in Ireland: A retrospective cross-sectional study using a national pharmacy claims data from 2016. *BMJ Open*, 10(4):e032916, April 2020. ISSN 2044-6055, 2044-6055. doi: 10.1136/bmjopen-2019-032916.
- Eric R. Gamazon, Heather E. Wheeler, Kaanan P. Shah, Sahar V. Mozaffari, Keston Aquino-Michaels, Robert J. Carroll, Anne E. Eyler, Joshua C. Denny, Dan L. Nicolae, Nancy J. Cox, and Hae Kyung Im. A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet*, 47(9):1091–1098, September 2015. ISSN 1546-1718. doi: 10.1038/ng.3367.
- Zhao-wei Gao, Xi Wang, Hui-zhong Zhang, Fang Lin, Chong Liu, and Ke Dong. The roles of adenosine deaminase in autoimmune diseases. *Autoimmunity Reviews*, 20(1):102709, January 2021. ISSN 1568-9972. doi: 10.1016/j.autrev.2020.102709.
- Graeme Gardner and Christopher A. Fraker. Natural Killer Cells as Key Mediators in Type I Diabetes Immunopathology. *Front Immunol*, 12, August 2021. ISSN 1664-3224. doi: 10.3389/fimmu.2021.722979.
- Yan Ge, Suna Onengut-Gumuscu, Aaron R. Quinlan, Aaron J. Mackey, Jocyn-dra A. Wright, Jane H. Buckner, Tania Habib, Stephen S. Rich, and Patrick Concannon. Targeted Deep Sequencing in Multiple-Affected Sibships of European Ancestry Identifies Rare Deleterious Variants in PTPN22 That Confer Risk for Type 1 Diabetes. *Diabetes*, 65(3):794–802, December 2015. ISSN 0012-1797. doi: 10.2337/db15-0322.
- Stephen E. Gitelman and Jeffrey A. Bluestone. Regulatory T cell therapy for type 1 diabetes: May the force be with you. *Journal of Autoimmunity*, 71:78–87, July 2016. ISSN 0896-8411. doi: 10.1016/j.jaut.2016.03.011.
- Stephen E. Gitelman, Brian N. Bundy, Ele Ferrannini, Noha Lim, J. Lori Blanchfield, Linda A. DiMeglio, Eric I. Felner, Jason L. Gaglia, Peter A. Gottlieb, S. Alice Long, Andrea Mari, Raghavendra G. Mirmira, Philip Raskin, Srinath Sanda, Eva Tsalikian, John M. Wentworth, Steven M. Willi, Jeffrey P. Krischer, Jeffrey A. Bluestone, Mayalin Barr, J. Lori Blanchfield, Jeffrey A. Bluestone,

- Jeanne Buchanan, Brian N. Bundy, Joanne Cabbage, Peter Coleman, Monica De La Vega, Linda A. DiMeglio, Carmella Evans-Molina, Eric I. Felner, Ele Ferrannini, Christine Ferrara, Jason L. Gaglia, Stephen E. Gitelman, Peter A. Gottlieb, Felicity Healy, Laurie Higgins, Megan Hildinger, Margaret Jenkins, Nora Kayton Bryant, Amanda Kinderman, Nisha Koshy, Brianne Kost, Jeffrey P. Krischer, Suzanne Krishfield, Olena Kucheruk, Noha Lim, Karen Lindley, S. Alice Long, Manasa Mantravadi, Andrea Mari, Shelley Mesfin, Aaron Michels, Mary Ellen Migre, Pantea Minnock, Raghavendra G. Mirmira, Elham Mohammed-Nur, Jennifer Nelson, Ashvin Nursing, Ryan O'Donnell, Diana Olivos, Melissa Parker, Philip Raskin, Leanne Redl, Nicole Reed, Brittany Resnick, Srinath Sanda, Peter Sayre, Elisavet Serti, Emily Sims, Karen Smith, Carol Soppe, Fiona Stuart, Sarah Szubowicz, Michel Tansey, Jennifer Terrell, Sarah Tersey, Christine Torok, Eva Tsalikian, Kelly Watson, John M. Wentworth, Rebecca Wesch, Steven Willi, and Stephanie Woerner. Imatinib therapy for patients with recent-onset type 1 diabetes: A multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *The Lancet Diabetes & Endocrinology*, 9(8):502–514, August 2021. ISSN 2213-8587, 2213-8595. doi: 10.1016/S2213-8587(21)00139-X.
- Adebola Matthew Giwa, Rizwan Ahmed, Zahra Omidian, Neha Majety, Kagan Ege Karakus, Sarah M. Omer, Thomas Donner, and Abdel Rahim A. Hamad. Current understandings of the pathogenesis of type 1 diabetes: Genetics to environment. *World Journal of Diabetes*, 11(1):13–25, January 2020. doi: 10.4239/wjd.v11.i1.13.
- John D. Glawe, D. Ross Patrick, Meng Huang, Christopher D. Sharp, Shayne C. Barlow, and Christopher G. Kevil. Genetic Deficiency of Itgb2 or ItgaL Prevents Autoimmune Diabetes Through Distinctly Different Mechanisms in NOD/LtJ Mice. *Diabetes*, 58(6):1292–1301, February 2009. ISSN 0012-1797. doi: 10.2337/db08-0804.
- John D. Glawe, Eleni M. Mijalis, William C. Davis, Shayne C. Barlow, Neslihan Gungor, Robert McVie, and Christopher G. Kevil. SDF-1–CXCR4 differentially regulates autoimmune diabetogenic T cell adhesion through ROBO1–SLIT2 interactions in mice. *Diabetologia*, 56(10):2222–2230, October 2013. ISSN 1432-0428. doi: 10.1007/s00125-013-2978-x.

- Ameya Gokhale, Shanthi Kanthala, John Latendresse, Veena Taneja, and Seetharama Satyanarayanajois. Immunosuppression by Co-stimulatory Molecules: Inhibition of CD2-CD48/CD58 Interaction by Peptides from CD2 to Suppress Progression of Collagen-induced Arthritis in Mice. *Chem Biol Drug Des*, 82(1):106–118, July 2013. ISSN 1747-0277. doi: 10.1111/cbdd.12138.
- Andrew J. Grant and Stephen Burgess. A Bayesian approach to Mendelian randomization using summary statistics in the univariable and multivariable settings with correlated pleiotropy. *Am J Hum Genet*, 111(1):165–180, January 2024. ISSN 0002-9297. doi: 10.1016/j.ajhg.2023.12.002.
- Stephanie Grebinoski, Qianxia Zhang, Anthony R. Cillo, Sasikanth Manne, Hanxi Xiao, Erin A. Brunazzi, Tracy Tabib, Carly Cardello, Christine G. Lian, George F. Murphy, Robert Lafyatis, E. John Wherry, Jishnu Das, Creg J. Workman, and Dario A. A. Vignali. Autoreactive CD8⁺ T cells are restrained by an exhaustion-like program that is maintained by LAG3. *Nat Immunol*, 23(6):868–877, June 2022. ISSN 1529-2916. doi: 10.1038/s41590-022-01210-5.
- Carla J. Greenbaum, Elisavet Serti, Katharina Lambert, Lia J. Weiner, Sai Kanaparthi, Sandra Lord, Stephen E. Gitelman, Darrell M. Wilson, Jason L. Gaglia, Kurt J. Griffin, William E. Russell, Philip Raskin, Antoinette Moran, Steven M. Willi, Eva Tsalikian, Linda A. DiMeglio, Kevan C. Herold, Wayne V. Moore, Robin Goland, Mark Harris, Maria E. Craig, Desmond A. Schatz, David A. Baidal, Henry Rodriguez, Kristina M. Utzschneider, Hendrik J. Nel, Carol L. Soppe, Karen D. Boyle, Karen Cerosaletti, Lynette Keyes-Elstein, S. Alice Long, Ranjeny Thomas, James G. McNamara, Jane H. Buckner, and Srinath Sanda. IL-6 receptor blockade does not slow β cell loss in new-onset type 1 diabetes. *JCI Insight*, 6(21), November 2021. ISSN 0021-9738. doi: 10.1172/jci.insight.150074.
- Daqian Gu, Xiang Ao, Yu Yang, Zhuo Chen, and Xiang Xu. Soluble immune checkpoints in cancer: Production, function and biological significance. *J Immunother Cancer*, 6(1):132, December 2018. ISSN 2051-1426. doi: 10.1186/s40425-018-0449-0.
- Chamutal Gur, Angel Porgador, Moran Elboim, Roi Gazit, Saar Mizrahi, Noam Stern-Ginossar, Hagit Achdout, Hormas Ghadially, Yuval Dor, Tomer Nir, Victoria Doviner, Oren HersHKovitz, Michal Mendelson, Yaakov Naparstek, and Ofer Mandelboim. The activating receptor NKp46 is essential for the development of type 1 diabetes. *Nat Immunol*, 11(2):121–128, February 2010. ISSN 1529-2916. doi: 10.1038/ni.1834.

- Alexander Gusev, Arthur Ko, Huwenbo Shi, Gaurav Bhatia, Wonil Chung, Brenda W. J. H. Penninx, Rick Jansen, Eco J. C. de Geus, Dorret I. Boomsma, Fred A. Wright, Patrick F. Sullivan, Elina Nikkola, Marcus Alvarez, Mete Civelek, Aldons J. Lusi, Terho Lehtimäki, Emma Raitoharju, Mika Kähönen, Ilkka Seppälä, Olli T. Raitakari, Johanna Kuusisto, Markku Laakso, Alkes L. Price, Päivi Pajukanta, and Bogdan Pasaniuc. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*, 48(3):245–252, March 2016. ISSN 1546-1718. doi: 10.1038/ng.3506.
- Christine S. Haddad, Palash Bhattacharya, Khaled Alharshaw, Alejandra Marinelarena, Prabhakaran Kumar, Osama El-Sayed, Hatem A. Elshabrawy, Alan L. Epstein, and Bellur S. Prabhakar. Age-dependent divergent effects of OX40L treatment on the development of diabetes in NOD mice. *Autoimmunity*, 49(5):298–311, July 2016. ISSN 0891-6934. doi: 10.1080/08916934.2016.1183657.
- William A Hagopian, Henry Erlich, Åke Lernmark, Marian Rewers, Anette G Ziegler, Olli Simell, Beena Akolkar, Robert Vogt Jr, Alan Blair, Jorma Ilonen, Jeffrey Krischer, JinXiong She, and the TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY): Genetic criteria and international diabetes risk screening of 421 000 infants. *Pediatric Diabetes*, 12(8):733–743, 2011. ISSN 1399-5448. doi: 10.1111/j.1399-5448.2011.00774.x.
- Michael J. Haller, Stephen E. Gitelman, Peter A. Gottlieb, Aaron W. Michels, Stephen M. Rosenthal, Jonathan J. Shuster, Baiming Zou, Todd M. Brusko, Maigan A. Hulme, Clive H. Wasserfall, Clayton E. Mathews, Mark A. Atkinson, and Desmond A. Schatz. Anti-thymocyte globulin/G-CSF treatment preserves β cell function in patients with established type 1 diabetes. *J. Clin. Invest.*, 125(1):448–455, January 2015. ISSN 0021-9738. doi: 10.1172/JCI78492.
- Michael J. Haller, Desmond A. Schatz, Jay S. Skyler, Jeffrey P. Krischer, Brian N. Bundy, Jessica L. Miller, Mark A. Atkinson, Dorothy J. Becker, David Baidal, Linda A. DiMeglio, Stephen E. Gitelman, Robin Goland, Peter A. Gottlieb, Kevan C. Herold, Jennifer B. Marks, Antoinette Moran, Henry Rodriguez, William Russell, Darrell M. Wilson, Carla J. Greenbaum, and Type 1 Diabetes TrialNet ATG-GCSF Study Group. Low-Dose Anti-Thymocyte Globulin (ATG) Preserves β -Cell Function and Improves HbA1c in New-Onset Type 1 Diabetes. *Diabetes Care*, 41(9):1917–1925, July 2018. ISSN 0149-5992. doi: 10.2337/dc18-0494.

- Changhao Han, Qinghua Geng, Jingjing Qin, Yulin Li, and Huarong Yu. Activation of 5-Hydroxytryptamine 4 Receptor Improves Colonic Barrier Function by Triggering Mucin 2 Production in a Mouse Model of Type 1 Diabetes. *The American Journal of Pathology*, 192(6):876–886, June 2022. ISSN 0002-9440, 1525-2191. doi: 10.1016/j.ajpath.2022.03.002.
- Ulf Hannelius, Craig A. Beam, and Johnny Ludvigsson. Efficacy of GAD-alum immunotherapy associated with HLA-DR3-DQ2 in recently diagnosed type 1 diabetes. *Diabetologia*, 63(10):2177–2181, October 2020. ISSN 1432-0428. doi: 10.1007/s00125-020-05227-z.
- Agnès Hartemann, Gilbert Bensimon, Christine A. Payan, Sophie Jacqueminet, Olivier Bourron, Nathalie Nicolas, Michèle Fonfrede, Michelle Rosenzweig, Claude Bernard, and David Klatzmann. Low-dose interleukin 2 in patients with type 1 diabetes: A phase 1/2 randomised, double-blind, placebo-controlled trial. *The Lancet Diabetes & Endocrinology*, 1(4):295–305, December 2013. ISSN 2213-8587, 2213-8595. doi: 10.1016/S2213-8587(13)70113-X.
- Mingyu He, Kate Roussak, Feiyang Ma, Nicholas Borchering, Vince Garin, Mike White, Charles Schutt, Trine I. Jensen, Yun Zhao, Courtney A. Iberg, Kairav Shah, Himanshi Bhatia, Daniel Korenfeld, Sabrina Dinkel, Judah Gray, Alina Ulezko Antonova, Stephen Ferris, David Donermeyer, Cecilia Lindestam Arlehamn, Matthew M. Gubin, Jingqin Luo, Laurent Gorvel, Matteo Pellegrini, Alessandro Sette, Thomas Tung, Rasmus Bak, Robert L. Modlin, Ryan C. Fields, Robert D. Schreiber, Paul M. Allen, and Eynav Klechevsky. CD5 expression by dendritic cells directs T cell immunity and sustains immunotherapy responses. *Science*, 379(6633):eabg2752, February 2023. doi: 10.1126/science.abg2752.
- S. Taylor Head, Felipe Dezem, Andrei Todor, Jingjing Yang, Jasmine Plummer, Simon Gayther, Siddhartha Kar, Joellen Schildkraut, and Michael P. Epstein. Cis- and trans-eQTL TWASs of breast and ovarian cancer identify more than 100 susceptibility genes in the BCAC and OCAC consortia. *The American Journal of Human Genetics*, 111(6):1084–1099, June 2024. ISSN 0002-9297, 1537-6605. doi: 10.1016/j.ajhg.2024.04.012.
- K. C. Herold, S. E. Gitelman, S. M. Willi, P. A. Gottlieb, F. Waldron-Lynch, L. Devine, J. Sherr, S. M. Rosenthal, S. Adi, M. Y. Jalaludin, A. W. Michels, J. Dziura, and J. A. Bluestone. Teplizumab treatment may improve C-peptide responses in participants with type 1 diabetes after the new-onset period: A

- randomised controlled trial. *Diabetologia*, 56(2):391–400, February 2013a. ISSN 1432-0428. doi: 10.1007/s00125-012-2753-4.
- Kevan C. Herold, William Hagopian, Julie A. Auger, Ena Poumian-Ruiz, Lesley Taylor, David Donaldson, Stephen E. Gitelman, David M. Harlan, Danlin Xu, Robert A. Zivin, and Jeffrey A. Bluestone. Anti-CD3 Monoclonal Antibody in New-Onset Type 1 Diabetes Mellitus. *New England Journal of Medicine*, 346(22):1692–1698, May 2002. ISSN 0028-4793. doi: 10.1056/NEJMoa012864.
- Kevan C. Herold, Stephen E. Gitelman, Umesh Masharani, William Hagopian, Brygida Bisikirska, David Donaldson, Kristina Rother, Beverly Diamond, David M. Harlan, and Jeffrey A. Bluestone. A Single Course of Anti-CD3 Monoclonal Antibody hOKT3 γ 1(Ala-Ala) Results in Improvement in C-Peptide Responses and Clinical Parameters for at Least 2 Years after Onset of Type 1 Diabetes. *Diabetes*, 54(6):1763–1769, June 2005. ISSN 0012-1797. doi: 10.2337/diabetes.54.6.1763.
- Kevan C. Herold, Stephen E. Gitelman, Mario R. Ehlers, Peter A. Gottlieb, Carla J. Greenbaum, William Hagopian, Karen D. Boyle, Lynette Keyes-Elstein, Sudeepta Aggarwal, Deborah Phippard, Peter H. Sayre, James McNamara, Jeffrey A. Bluestone, and the AbATE Study Team. Teplizumab (Anti-CD3 mAb) Treatment Preserves C-Peptide Responses in Patients With New-Onset Type 1 Diabetes in a Randomized Controlled Trial: Metabolic and Immunologic Features at Baseline Identify a Subgroup of Responders. *Diabetes*, 62(11):3766–3774, October 2013b. ISSN 0012-1797. doi: 10.2337/db13-0345.
- Kevan C. Herold, Brian N. Bundy, S. Alice Long, Jeffrey A. Bluestone, Linda A. DiMeglio, Matthew J. Dufort, Stephen E. Gitelman, Peter A. Gottlieb, Jeffrey P. Krischer, Peter S. Linsley, Jennifer B. Marks, Wayne Moore, Antoinette Moran, Henry Rodriguez, William E. Russell, Desmond Schatz, Jay S. Skyler, Eva Tsalikian, Diane K. Wherrett, Anette-Gabriele Ziegler, and Carla J. Greenbaum. An Anti-CD3 Antibody, Teplizumab, in Relatives at Risk for Type 1 Diabetes. *New England Journal of Medicine*, 381(7):603–613, August 2019. ISSN 0028-4793. doi: 10.1056/NEJMoa1902226.
- Kevan C. Herold, Stephen E. Gitelman, Peter A. Gottlieb, Laura A. Knecht, Ralph Raymond, and Eleanor L. Ramos. Teplizumab: A Disease-Modifying Therapy for Type 1 Diabetes That Preserves β -Cell Function. *Diabetes Care*, 46(10):1848–1856, August 2023. ISSN 0149-5992. doi: 10.2337/dc23-0675.

- Kevan C. Herold, Thomas Delong, Ana Luisa Perdigoto, Noah Biru, Todd M. Brusko, and Lucy S. K. Walker. The immunology of type 1 diabetes. *Nat Rev Immunol*, 24(6):435–451, June 2024. ISSN 1474-1741. doi: 10.1038/s41577-023-00985-4.
- Kristin Hieronymus, Benjamin Dorschner, Felix Schulze, Neeta L. Vora, Joel S. Parker, Jennifer Lucia Winkler, Angela Rösen-Wolff, and Stefan Winkler. Validation of reference genes for whole blood gene expression analysis in cord blood of preterm and full-term neonates and peripheral blood of healthy adults. *BMC Genomics*, 22(1):489, June 2021. ISSN 1471-2164. doi: 10.1186/s12864-021-07801-0.
- M. Karoliina Hirvonen, Niina Lietzén, Robert Moulder, Santosh D. Bhosale, Jaakko Koskenniemi, Mari Vähä-Mäkilä, Mirja Nurmio, Matej Orešič, Jorma Ilonen, Jorma Toppari, Riitta Veijola, Heikki Hyöty, Harri Lähdesmäki, Mikael Knip, Lu Cheng, and Riitta Lahesmaa. Serum APOC1 levels are decreased in young autoantibody positive children who rapidly progress to type 1 diabetes. *Scientific Reports*, 13:15941, September 2023. doi: 10.1038/s41598-023-43039-4.
- Samuel J. Hobbs and Jeffrey C. Nolz. Regulation of T Cell Trafficking by Enzymatic Synthesis of O-Glycans. *Front. Immunol.*, 8, May 2017. ISSN 1664-3224. doi: 10.3389/fimmu.2017.00600.
- Naosuke Hoshina, Asami Tanimura, Miwako Yamasaki, Takeshi Inoue, Ry-oji Fukabori, Teiko Kuroda, Kazumasa Yokoyama, Tohru Tezuka, Hiroshi Sagara, Shinji Hirano, Hiroshi Kiyonari, Masahiko Takada, Kazuto Kobayashi, Masahiko Watanabe, Masanobu Kano, Takanobu Nakazawa, and Tadashi Yamamoto. Protocadherin 17 Regulates Presynaptic Assembly in Topographic Corticobasal Ganglia Circuits. *Neuron*, 78(5):839–854, June 2013. ISSN 0896-6273. doi: 10.1016/j.neuron.2013.03.031.
- Tomoaki Hoshino, Kentaro Yamada, Kazuhiro Masuoka, Kyogo Itoh, Kyohei Nonaka, and Kotaro Oizumi. Elevated adenosine deaminase activity in the serum of patients with diabetes mellitus. *Diabetes Research and Clinical Practice*, 25(2):97–102, September 1994. ISSN 0168-8227. doi: 10.1016/0168-8227(94)90034-5.

- Xiaotong Hu, Xinbing Sui, Lili Li, Xuefeng Huang, Rong Rong, Xianwei Su, Qinglan Shi, Lijuan Mo, Xingsheng Shu, Yeye Kuang, Qian Tao, and Chao He. Protocadherin 17 acts as a tumour suppressor inducing tumour cell apoptosis and autophagy, and is frequently methylated in gastric and colorectal cancers. *The Journal of Pathology*, 229(1):62–73, 2013. ISSN 1096-9896. doi: 10.1002/path.4093.
- Xiaowei Hu, Bobbie-Jo M. Webb-Robertson, Hemang M. Parikh, Ernesto S. Nakayasu, Suna Onengut-Gumuscu, Wei-Min Chen, Ashley Frazer-Abel, Thomas O. Metz, Stephen S. Rich, Marian J. Rewers, and Ani Manichaik. Genome-wide mapping of complement system proteins for islet autoimmunity in the DAISY and TEDDY children, February 2025. ISSN 2693-5015.
- Qing Huang, Yali Wang, Zhiruo Zhang, Yueying Yang, Hua Li, and Lixia Chen. SEMA7A: A glycoprotein with therapeutic potential in inflammatory diseases and tumor development. *International Journal of Biological Macromolecules*, 318:144907, July 2025. ISSN 0141-8130. doi: 10.1016/j.ijbiomac.2025.144907.
- Jin-Ling Huo, Ya-Tao Wang, Wen-Jia Fu, Nan Lu, and Zhang-Suo Liu. The promising immune checkpoint LAG-3 in cancer immunotherapy: From basic research to clinical application. *Front Immunol*, 13, July 2022. ISSN 1664-3224. doi: 10.3389/fimmu.2022.956090.
- Andrii Iakovliev, Stuart J. McGurnaghan, Caroline Hayward, Marco Colombo, Debby Lipschutz, Athina Spiliopoulou, Helen M. Colhoun, and Paul M. McKeigue. Genome-wide aggregated trans-effects on risk of type 1 diabetes: A test of the “omnigenic” sparse effector hypothesis of complex trait genetics. *Am J Hum Genet*, 110(6):913–926, June 2023. ISSN 0002-9297. doi: 10.1016/j.ajhg.2023.04.003.
- Andrii Iakovliev, Olivia Castellini-Pérez, Buddhiprabha Erabadda, Javier Martín, Guillermo Barturen, Paul M. McKeigue, Elena Carnero-Montoro, Marta E. Alarcón-Riquelme, and Athina Spiliopoulou. Discovery of core genes for systemic lupus erythematosus via genome-wide aggregated trans-effects analysis. *Genes Immun*, pages 1–12, September 2025. ISSN 1476-5470. doi: 10.1038/s41435-025-00352-4.

- Richard A. Insel, Jessica L. Dunne, Mark A. Atkinson, Jane L. Chiang, Dana Dabelea, Peter A. Gottlieb, Carla J. Greenbaum, Kevan C. Herold, Jeffrey P. Krischer, Åke Lernmark, Robert E. Ratner, Marian J. Rewers, Desmond A. Schatz, Jay S. Skyler, Jay M. Sosenko, and Anette-G. Ziegler. Staging Presymptomatic Type 1 Diabetes: A Scientific Statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care*, 38(10):1964–1974, September 2015. ISSN 0149-5992. doi: 10.2337/dc15-1419.
- Aleksandar Ivetic, Hannah Louise Hoskins Green, and Samuel James Hart. L-selectin: A Major Regulator of Leukocyte Adhesion, Migration and Signaling. *Front Immunol*, 10, May 2019. ISSN 1664-3224. doi: 10.3389/fimmu.2019.01068.
- Wenbo Jiang, Xuebin Chu, Lei Song, Xue Li, Weili Xue, Lingyun Li, Ranran Zhao, Fei Gu, and John Xu. GenSci120, a humanized PD-1 agonistic monoclonal antibody, enhances the binding of both PD-L1 and PD-L2 to PD-1, and mitigates the symptoms of GVHD in a mouse model. *J Immunol*, 212(1_Supplement): 0718_6289, May 2024. ISSN 0022-1767. doi: 10.4049/jimmunol.212.suppl.0718.6289.
- Matthew B. Johnson, Masato Ogishi, Clara Domingo-Vila, Elisa De Franco, Matthew N. Wakeling, Zineb Imane, Brittany Resnick, Evangelia Williams, Rui Pedro Galão, Richard Caswell, James Russ-Silsby, Yoann Seeleuthner, Darawan Rinchai, Iris Fagniez, Basilin Benson, Matthew J. Dufort, Cate Speake, Megan E. Smithmyer, Michelle Hudson, Rebecca Dobbs, EXE-T1D consortium, Zoe Quandt, Andrew T. Hattersley, Peng Zhang, Stephanie Boisson-Dupuis, Mark S. Anderson, Jean-Laurent Casanova, Timothy I. Tree, and Richard A. Oram. Human inherited PD-L1 deficiency is clinically and immunologically less severe than PD-1 deficiency. *J Exp Med*, 221(6): e20231704, April 2024. ISSN 0022-1007. doi: 10.1084/jem.20231704.
- Julie Josse and François Husson. missMDA: A Package for Handling Missing Values in Multivariate Data Analysis. *Journal of Statistical Software*, 70:1–31, April 2016. ISSN 1548-7660. doi: 10.18637/jss.v070.i01.
- Motoko Kanzaki, Jun Wada, Koichi Sugiyama, Atsuko Nakatsuka, Sanae Teshigawara, Kazutoshi Murakami, Kentaro Inoue, Takahiro Terami, Akihiro Katayama, Jun Eguchi, Hisaya Akiba, Hideo Yagita, and Hirofumi Makino. Galectin-9 and T Cell Immunoglobulin Mucin-3 Pathway Is a Therapeutic Target for Type 1 Diabetes. *Endocrinology*, 153(2):612–620, February 2012. ISSN 0013-7227. doi: 10.1210/en.2011-1579.

- Atsuko Kataoka, Kazuhiko Yamada, Teruki Hagiwara, Masayoshi Terayama, Takayuki Sugimoto, Kyoko Nohara, Toru Igari, Chizu Yokoi, and Yuki I. Kawamura. Expression Status of Serine Protease 27: A Prognostic Marker for Esophageal Squamous Cell Carcinoma Treated with Preoperative Chemotherapy/Chemoradiotherapy. *Ann Surg Oncol*, 28(9):5373–5381, September 2021. ISSN 1534-4681. doi: 10.1245/s10434-020-09550-y.
- Daniel H. Katz, Jeremy M. Robbins, Shuliang Deng, Usman A. Tahir, Alexander G. Bick, Akhil Pampana, Zhi Yu, Debby Ngo, Mark D. Benson, Zsu-Zsu Chen, Daniel E. Cruz, Dongxiao Shen, Yan Gao, Claude Bouchard, Mark A. Sarzynski, Adolfo Correa, Pradeep Natarajan, James G. Wilson, and Robert E. Gerszten. Proteomic profiling platforms head to head: Leveraging genetics and clinical traits to compare aptamer- and antibody-based methods. *Science Advances*, 8(33):eabm5164, August 2022. doi: 10.1126/sciadv.abm5164.
- Eiji Kawasaki, Takahiro Fukuyama, Aira Uchida, Yoko Sagara, Yuko Nakano, Hidekazu Tamai, Masayuki Tojikubo, Yuji Hiromatsu, and Nobuhiko Koga. Development of type 1 diabetes in a patient treated with anti-interleukin-6 receptor antibody for rheumatoid arthritis. *Journal of Diabetes Investigation*, 13(4):738–740, 2022. ISSN 2040-1124. doi: 10.1111/jdi.13706.
- Laurie L. Kenney, Rebecca Suet-Yan Chiu, Michelle N. Dutra, Alexandra Wactor, Chris Honan, Lukas Shelerud, Joshua J. Corrigan, Kelly Yu, Joseph D. Ferrari, Kate L. Jeffrey, Eric Huang, and Paul L. Stein. mRNA-delivery of IDO1 suppresses T cell-mediated autoimmunity. *CR Med*, 5(9), September 2024. ISSN 2666-3791. doi: 10.1016/j.xcrm.2024.101717.
- Bart Keymeulen, Evy Vandemeulebroucke, Anette G. Ziegler, Chantal Mathieu, Leonard Kaufman, Geoff Hale, Frans Gorus, Michel Goldman, Markus Walter, Sophie Candon, Liliane Schandene, Laurent Crenier, Christophe De Block, Jean-Marie Seigneurin, Pieter De Pauw, Denis Pierard, Ilse Weets, Peppy Rebello, Pru Bird, Eleanor Berrie, Mark Frewin, Herman Waldmann, Jean-François Bach, Daniel Pipeleers, and Lucienne Chatenoud. Insulin Needs after CD3-Antibody Therapy in New-Onset Type 1 Diabetes. *New England Journal of Medicine*, 352(25):2598–2608, June 2005. ISSN 0028-4793. doi: 10.1056/NEJMoa043980.
- Bart Keymeulen, André van Maurik, Dave Inman, João Oliveira, Rene McLaughlin, Rachel M. Gittelman, Bart O. Roep, Pieter Gillard, Robert Hilbrands, Frans Gorus, Chantal Mathieu, Ursule Van de Velde, Nicolas Wisniacki, and

- Antonella Napolitano. A randomised, single-blind, placebo-controlled, dose-finding safety and tolerability study of the anti-CD3 monoclonal antibody oteelixumab in new-onset type 1 diabetes. *Diabetologia*, 64(2):313–324, February 2021. ISSN 1432-0428. doi: 10.1007/s00125-020-05317-y.
- Muhammad Khan, Zhihong Zhao, Sumbal Arooj, Yuxiang Fu, and Guixiang Liao. Soluble PD-1: Predictive, Prognostic, and Therapeutic Value for Cancer Immunotherapy. *Front Immunol*, 11, November 2020. ISSN 1664-3224. doi: 10.3389/fimmu.2020.587460.
- P Khanna, KL Johnson, and JL Maron. Optimal reference genes for RT-qPCR normalization in the newborn. *Biotechnic & Histochemistry*, 92(7):459–466, October 2017. ISSN 1052-0295. doi: 10.1080/10520295.2017.1362474.
- Minna Kiviniemi, Robert Hermann, Jussi Nurmi, Anette G. Ziegler, Mikael Knip, Olli Simell, Riitta Veijola, Timo Lövgren, Jorma Ilonen, and The TEDDY Study Group. A High-Throughput Population Screening System for the Estimation of Genetic Risk for Type 1 Diabetes: An Application for the TEDDY (The Environmental Determinants of Diabetes in the Young) Study. *Diabetes Technology & Therapeutics*, 9(5):460–472, October 2007. ISSN 1520-9156, 1557-8593. doi: 10.1089/dia.2007.0229.
- Craig Knox, Mike Wilson, Christen M. Klinger, Mark Franklin, Eponine Oler, Alex Wilson, Allison Pon, Jordan Cox, Na Eun Lucy Chin, Seth A. Strawbridge, Marysol Garcia-Patino, Ray Kruger, Aadhavya Sivakumaran, Selena Sanford, Rahil Doshi, Nitya Khetarpal, Omolola Fatokun, Daphnee Doucet, Ashley Zubkowski, Dorsa Yahya Rayat, Hayley Jackson, Karxena Harford, Afia Anjum, Mahi Zakir, Fei Wang, Siyang Tian, Brian Lee, Jaanus Liigand, Harrison Peters, Ruo Qi Rachel Wang, Tue Nguyen, Denise So, Matthew Sharp, Rodolfo da Silva, Cyrella Gabriel, Joshua Scantlebury, Marissa Jasin-ski, David Ackerman, Timothy Jewison, Tanvir Sajed, Vasuk Gautam, and David S. Wishart. DrugBank 6.0: The DrugBank Knowledgebase for 2024. *Nucleic Acids Res*, 52(D1):D1265–D1275, January 2024. ISSN 1362-4962. doi: 10.1093/nar/gkad976.
- Dana Koishybayeva, Adilet Amirashov, Saule Balmagambetova, Nadiar M. Mussin, and Amin Tamadon. Mesenchymal stem cell therapy for diabetes: An umbrella review. *Tissue and Cell*, 96:103043, October 2025. ISSN 0040-8166. doi: 10.1016/j.tice.2025.103043.

- Umut Can Kucuksezer, Esin Aktas Cetin, Fehim Esen, Ilhan Tahrali, Nilgun Akdeniz, Metin Yusuf Gelmez, and Gunnur Deniz. The Role of Natural Killer Cells in Autoimmune Diseases. *Front. Immunol.*, 12, February 2021. ISSN 1664-3224. doi: 10.3389/fimmu.2021.622306.
- Mark J Kwakkenbos, Walter Pouwels, Mourad Matmati, Martin Stacey, Hsi-Hsien Lin, Siamon Gordon, René A W Van Lier, and Jörg Hamann. Expression of the largest CD97 and EMR2 isoforms on leukocytes facilitates a specific interaction with chondroitin sulfate on B cells. *Journal of Leukocyte Biology*, 77(1):112–119, January 2005. ISSN 0741-5400. doi: 10.1189/jlb.0704402.
- Stanley Lasch, Peter Müller, Monika Bayer, Josef M. Pfeilschifter, Andrew D. Luster, Edith Hintermann, and Urs Christen. Anti-CD3/Anti-CXCL10 Antibody Combination Therapy Induces a Persistent Remission of Type 1 Diabetes in Two Mouse Models. *Diabetes*, 64(12):4198–4211, August 2015. ISSN 0012-1797. doi: 10.2337/db15-0479.
- Hye-Seung Lee, Brant R. Burkhardt, Wendy McLeod, Susan Smith, Chris Eberhard, Kristian Lynch, David Hadley, Marian Rewers, Olli Simell, Jin-Xiong She, Bill Hagopian, Ake Lernmark, Beena Akolkar, Anette G. Ziegler, Jeffrey P. Krischer, and The TEDDY study Group. Biomarker discovery study design for type 1 diabetes in The Environmental Determinants of Diabetes in the Young (TEDDY) study. *Diabetes/Metabolism Research and Reviews*, 30(5):424–434, 2014. ISSN 1520-7560. doi: 10.1002/dmrr.2510.
- Steven B Leichter, Jamie L Felton, Cristy Geno Rasmussen, Patrick Rizzuto, Natalie Bellini, Osagie Ebekozien, and Rifka Schulman-Rosenbaum. Establishing Screening Programs for Presymptomatic Type 1 Diabetes: Practical Guidance for Diabetes Care Providers. *J Clin Endocrinol Metab*, 110(8):2371–2382, August 2025. ISSN 0021-972X. doi: 10.1210/clinem/dgaf194.
- D J Lenschow, S C Ho, H Sattar, L Rhee, G Gray, N Nabavi, K C Herold, and J A Bluestone. Differential effects of anti-B7-1 and anti-B7-2 monoclonal antibody treatment on the development of diabetes in the nonobese diabetic mouse. *J Exp Med*, 181(3):1145–1155, March 1995. ISSN 0022-1007. doi: 10.1084/jem.181.3.1145.
- Deborah J. Lenschow, Kevan C. Herold, Lesley Rhee, Bina Patel, Ann Koons, Hui-Yu Qin, Elaine Fuchs, Bhagarith Singh, Craig B. Thompson, and Jeffrey A. Bluestone. CD28/B7 Regulation of Th1 and Th2 Subsets in the Development of Autoimmune Diabetes. *Immunity*, 5(3):285–293, September 1996. ISSN 1074-7613. doi: 10.1016/S1074-7613(00)80323-4.

- Zhilin Li, Mohd Moin Khan, Juha Kuja-Panula, Hongyun Wang, Yu Chen, Deyin Guo, Zhi Jane Chen, Riitta Lahesmaa, Heikki Rauvala, and Li Tian. AMIGO2 modulates T cell functions and its deficiency in mice ameliorates experimental autoimmune encephalomyelitis. *Brain, Behavior, and Immunity*, 62:110–123, May 2017. ISSN 0889-1591. doi: 10.1016/j.bbi.2017.01.009.
- Jake Lin, Elaheh Moradi, Karoliina Salenius, Suvi Lehtipuro, Tomi Häkkinen, Jutta E. Laiho, Sami Oikarinen, Sofia Randelin, Hemang M. Parikh, Jeffrey P. Krischer, Jorma Toppari, Åke Lernmark, Joseph F. Petrosino, Nadim J. Ajami, Jin-Xiong She, William A. Hagopian, Marian J. Rewers, Richard E. Lloyd, Kirsi J. Rautajoki, Heikki Hyöty, Matti Nykter, and TEDDY Study Group. Distinct transcriptomic profiles in children prior to the appearance of type 1 diabetes-linked islet autoantibodies and following enterovirus infection. *Nat Commun*, 14(1):7630, November 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-42763-9.
- Irving Ling, Stanley Marciniak, Stephen Clarke, Vani Lakshminarayanan, Matthew J Loza, Sophia Liva, Tong Wang, Ashley Orillion, Ilia Tikhonov, and Erika Noss. POS0597 SAFETY, TOLERABILITY, AND ACTIVITY OF JNJ-67484703 IN PARTICIPANTS WITH ACTIVE RHEUMATOID ARTHRITIS: RESULTS OF A MULTICENTER, DOUBLE-BLIND, PLACEBO-CONTROLLED, RANDOMIZED, MULTIPLE-DOSE PHASE 1b STUDY. *Ann Rheum Dis*, 84 (Supplement 1):794–795, June 2025. ISSN 0003-4967. doi: 10.1016/j.ard.2025.05.977.
- Xuanyao Liu, Yang I. Li, and Jonathan K. Pritchard. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell*, 177(4):1022–1034.e6, May 2019. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2019.04.014.
- Yuk-Fun Liu, Jake Powrie, Sefina Arif, Jennie H.M. Yang, Evangelia Williams, Leena Khatri, Mamta Joshi, Loic Lhuillier, Nikolaos Fountoulakis, Emma Smith, Craig Beam, Anna Lorenc, Mark Peakman, and Timothy Tree. Immune and Metabolic Effects of Antigen-Specific Immunotherapy Using Multiple β -Cell Peptides in Type 1 Diabetes. *Diabetes*, 71(4):722–732, January 2022. ISSN 0012-1797. doi: 10.2337/db21-0728.
- Jingli Lu, Jiyun Liu, Lulu Li, Yan Lan, and Yan Liang. Cytokines in type 1 diabetes: Mechanisms of action and immunotherapeutic targets. *Clin Transl Immunology*, 9(3):e1122, March 2020. ISSN 2050-0068. doi: 10.1002/cti2.1122.

- Johnny Ludvigsson, Maria Faresjö, Maria Hjorth, Stina Axelsson, Mikael Chéramy, Mikael Pihl, Outi Vaarala, Gun Forsander, Sten Ivarsson, Calle Johansson, Agne Lindh, Nils-Östen Nilsson, Jan Åman, Eva Örtqvist, Peter Zerhouni, and Rosaura Casas. GAD Treatment and Insulin Secretion in Recent-Onset Type 1 Diabetes. *New England Journal of Medicine*, 359(18): 1909–1920, October 2008. ISSN 0028-4793. doi: 10.1056/NEJMoa0804328.
- Johnny Ludvigsson, David Krisky, Rosaura Casas, Tadej Battelino, Luis Castaño, James Greening, Olga Kordonouri, Timo Otonkoski, Paolo Pozzilli, Jean-Jacques Robert, Henk J. Veeze, and Jerry Palmer. GAD65 Antigen Therapy in Recently Diagnosed Type 1 Diabetes Mellitus. *New England Journal of Medicine*, 366(5):433–442, February 2012. ISSN 0028-4793. doi: 10.1056/NEJMoa1107096.
- Johnny Ludvigsson, Jeanette Wahlberg, and Rosaura Casas. Intralymphatic Injection of Autoantigen in Type 1 Diabetes. *New England Journal of Medicine*, 376(7):697–699, February 2017. ISSN 0028-4793. doi: 10.1056/NEJMc1616343.
- Johnny Ludvigsson, Zdenek Sumnik, Terezie Pelikanova, Lia Nattero Chavez, Elena Lundberg, Itxaso Rica, Maria A. Martínez-Brocca, Marisol Ruiz de Adana, Jeanette Wahlberg, Anastasia Katsarou, Ragnar Hanas, Cristina Hernandez, Maria Clemente León, Ana Gómez-Gila, Marcus Lind, Marta Ferrer Lozano, Theo Sas, Ulf Samuelsson, Stepanka Pruhova, Fabricia Dietrich, Sara Puente Marin, Anders Nordlund, Ulf Hannelius, and Rosaura Casas. Intralymphatic Glutamic Acid Decarboxylase With Vitamin D Supplementation in Recent-Onset Type 1 Diabetes: A Double-Blind, Randomized, Placebo-Controlled Phase IIb Trial. *Diabetes Care*, 44(7):1604–1612, May 2021. ISSN 0149-5992. doi: 10.2337/dc21-0318.
- Johnny Ludvigsson, Linnea Eriksson, Christoph Nowak, Pedro F. Teixeira, Martina Widman, Anton Lindqvist, Rosaura Casas, Marcus Lind, and Ulf Hannelius. Phase III, randomised, double-blind, placebo-controlled, multicentre trial to evaluate the efficacy and safety of rhGAD65 to preserve endogenous beta cell function in adolescents and adults with recently diagnosed type 1 diabetes, carrying the genetic HLA DR3-DQ2 haplotype: The DIAGNODE-3 study protocol. *BMJ Open*, 12(10):e061776, October 2022. ISSN 2044-6055, 2044-6055. doi: 10.1136/bmjopen-2022-061776.

- Justin M. Luningham, Junyu Chen, Shizhen Tang, Philip L. De Jager, David A. Bennett, Aron S. Buchman, and Jingjing Yang. Bayesian Genome-wide TWAS Method to Leverage both cis- and trans-eQTL Information through Summary Statistics. *The American Journal of Human Genetics*, 107(4):714–726, October 2020. ISSN 0002-9297. doi: 10.1016/j.ajhg.2020.08.022.
- Yan Lv, Xiao Ma, Yuxin Ma, Yuxin Du, and Jifeng Feng. A new emerging target in cancer immunotherapy: Galectin-9 (LGALS9). *Genes & Diseases*, 10(6): 2366–2382, November 2023. ISSN 2352-3042. doi: 10.1016/j.gendis.2022.05.020.
- Yumeng Ma, Fanqiang Meng, Zhongda Lin, Yanjun Chen, Tianyu Lan, Zhaoxin Yang, Rui Diao, Xiaozhou Zhang, Qi Chen, Chi Zhang, Yishi Tian, Chanjuan Li, Wenli Fang, Xin Liang, and Xudong Zhang. Bioengineering Platelets Presenting PD-L1, Galectin-9 and BTLA to Ameliorate Type 1 Diabetes. *Advanced Science*, 12(16):2501139, 2025. ISSN 2198-3844. doi: 10.1002/advs.202501139.
- Dianna J Magliano and Edward J Boyko. IDF Atlas 11th Edition. Technical report, International Diabetes Federation, 2025.
- Julia Manasson, Marisella Panduro, Michael Ciani, Minasri Borah, Stephanie Grebinoski, Joshua Vitlip, Stephen Lutz, Ishan Sharma, Elliott Wittenberg, Allison Colthart, Samuel Perry, Maria Cecilia Ramello, Chelsea R. Parker Harp, Jyothsna Visweswaraiah, Ryan Peckner, Alex Pellerin, Heather Vital, John Sundy, Nathan Higginson-Scott, Kevin L. Otipoby, and Daniela Cipolletta. S-4321, a Novel Dual-cell Bidirectional PD-1:FcyRIIb Selective Agonist Antibody for the Treatment of Autoimmune Disease. In *ACR Meeting Abstracts*, November 2024.
- Camila Manrique-Acevedo, Irl B. Hirsch, and Robert H. Eckel. Prevention of Cardiovascular Disease in Type 1 Diabetes. *New England Journal of Medicine*, 390(13):1207–1217, April 2024. ISSN 0028-4793. doi: 10.1056/NEJMra2311526.
- Federica M. Marelli-Berg, Klaus Okkenhaug, and Vincenzo Mirenda. A two-signal model for T cell trafficking. *Trends in Immunology*, 28(6):267–273, June 2007. ISSN 14714906. doi: 10.1016/j.it.2007.04.008.

Fergal J Martin, M Ridwan Amode, Alisha Aneja, Olanrewaju Austine-Orimoloye, Andrey G Azov, If Barnes, Arne Becker, Ruth Bennett, Andrew Berry, Jyothish Bhai, Simarpreet Kaur Bhurji, Alexandra Bignell, Sanjay Boddu, Paulo R Branco Lins, Lucy Brooks, Shashank Budhanuru Ramaraju, Mehrnaz Charkhchi, Alexander Cockburn, Luca Da Rin Fiorretto, Claire Davidson, Kamalkumar Dodiya, Sarah Donaldson, Bilal El Houdaigui, Tamara El Naboulsi, Reham Fatima, Carlos Garcia Giron, Thiago Genez, Gurpreet S Ghattaoraya, Jose Gonzalez Martinez, Cristi Guijarro, Matthew Hardy, Zoe Hollis, Thibaut Hourlier, Toby Hunt, Mike Kay, Vinay Kaykala, Tuan Le, Diana Lemos, Diego Marques-Coelho, José Carlos Marugán, Gabriela Alejandra Merino, Louisse Paola Mirabueno, Aleena Mushtaq, Syed Nakib Hossain, Denye N Ogeh, Manoj Pandian Sakthivel, Anne Parker, Malcolm Perry, Ivana Pil-izota, Irina Prosovetkskaia, José G Pérez-Silva, Ahamed Imran Abdul Salam, Nuno Saraiva-Agostinho, Helen Schuilenburg, Dan Sheppard, Swati Sinha, Botond Sipos, William Stark, Emily Steed, Ranjit Sukumaran, Dulika Sumathipala, Marie-Marthe Suner, Likhitha Surapaneni, Kyösti Sutinen, Michal Szpak, Francesca Floriana Tricomi, David Urbina-Gómez, Andres Veidenberg, Thomas A Walsh, Brandon Walts, Elizabeth Wass, Natalie Willhoft, Jamie Allen, Jorge Alvarez-Jarreta, Marc Chakiachvili, Bethany Flint, Stefano Giorggetti, Leanne Haggerty, Garth R Ilsley, Jane E Loveland, Benjamin Moore, Jonathan M Mudge, John Tate, David Thybert, Stephen J Trevanion, Andrea Winterbottom, Adam Frankish, Sarah E Hunt, Magali Ruffier, Fiona Cunningham, Sarah Dyer, Robert D Finn, Kevin L Howe, Peter W Harrison, Andrew D Yates, and Paul Flicek. Ensembl 2023. *Nucleic Acids Res*, 51(D1):D933–D941, January 2023. ISSN 0305-1048. doi: 10.1093/nar/gkac958.

S. Martin, G. Schernthaner, J. Nerup, F. A. Gries, V. A. Koivisto, J. Dupré, E. Standl, P. Hamet, R. McArthur, M. H. Tan, K. Dawson, A. E. Mehta, S. Van Vliet, B. von Graffenried, C. Stiller, and H. Kolb. Follow-up of cyclosporin A treatment in Type 1 (insulin-dependent) diabetes mellitus: Lack of long-term effects. *Diabetologia*, 34(6):429–434, June 1991. ISSN 1432-0428. doi: 10.1007/BF00403182.

Natalia Martin-Orozco, Zhibin Chen, Laurent Poirot, Elzbieta Hyatt, Andy Chen, Osami Kanagawa, Arlene Sharpe, Diane Mathis, and Christophe Benoist. Paradoxical Dampening of Anti-Islet Self-Reactivity but Promotion of Diabetes by OX40 Ligand. *The Journal of Immunology*, 171(12):6954–6960, December 2003. ISSN 0022-1767, 1550-6606. doi: 10.4049/jimmunol.171.12.6954.

- Takumi Maruhashi, Daisuke Sugiura, Il-mi Okazaki, and Taku Okazaki. LAG-3: From molecular functions to clinical applications. *J Immunother Cancer*, 8(2), September 2020. ISSN 2051-1426. doi: 10.1136/jitc-2020-001014.
- Ashish K Marwaha, Samuel Chow, Anne M Pesenacker, Laura Cook, Annika Sun, S Alice Long, Jennie H M Yang, Kirsten A Ward-Hartstonge, Evangelia Williams, Clara Domingo-Vila, Khalif Halani, Kristina M Harris, Timothy I M Tree, Megan K Levings, Thomas Elliott, Rusung Tan, and Jan P Dutz. A phase 1b open-label dose-finding study of ustekinumab in young adults with type 1 diabetes. *Immunotherapy Advances*, 2(1):ltab022, January 2022. ISSN 2732-4303. doi: 10.1093/immadv/ltab022.
- Lucy Mastrandrea, Jihneeh Yu, Torsten Behrens, John Buchlis, Christine Albini, Shannon Fournier, and Teresa Quattrin. Etanercept Treatment in Children With New-Onset Type 1 Diabetes: Pilot randomized, placebo-controlled, double-blind study. *Diabetes Care*, 32(7):1244–1249, April 2009. ISSN 0149-5992. doi: 10.2337/dc09-0054.
- Francesco Antonio Mazzotta, Lorenzo Lucaccini Paoli, Alessandro Rizzi, Linda Tartaglione, Maria Laura Leo, Valentina Popolla, Annarita Barberio, Luca Viti, Mauro Di Leo, Alfredo Pontecorvi, and Dario Pitocco. Unmet needs in the treatment of type 1 diabetes: Why is it so difficult to achieve an improvement in metabolic control? *Nutr. Diabetes*, 14(1):58, August 2024. ISSN 2044-4052. doi: 10.1038/s41387-024-00319-w.
- Shannon L. McArdel, Daniel R. Brown, Raymond A. Sobel, and Arlene H. Sharpe. Anti-CD48 mAb attenuates experimental autoimmune encephalomyelitis by limiting the number of pathogenic CD4+ T cells. *J Immunol*, 197(8):3038–3048, October 2016a. ISSN 0022-1767. doi: 10.4049/jimmunol.1600706.
- Shannon L. McArdel, Cox Terhorst, and Arlene H. Sharpe. Roles of CD48 in regulating immunity and tolerance. *Clin Immunol*, 164:10–20, March 2016b. ISSN 1521-6616. doi: 10.1016/j.clim.2016.01.008.
- Adam McGinnis. Genes and Microbes—Old Foes and New Revelations on the Road to Type 1 Diabetes. <https://www.adameetingnews.org/genes-and-microbes-old-foes-and-new-revelations-on-the-road-to-type-1-diabetes/>, June 2025.

- Paul M. McKeigue, Athina Spiliopoulou, Stuart McGurnaghan, Marco Colombo, Luke Blackbourn, Timothy J. McDonald, Suna Onengut-Gomuscu, Stephen S. Rich, Colin N. A. Palmer, John A. McKnight, Mark W. J. Strachan, Alan W. Patrick, John Chalmers, Robert S. Lindsay, John R. Petrie, Sandeep Thekkepat, Andrew Collier, Sandra MacRury, and Helen M. Colhoun. Persistent C-peptide secretion in Type 1 diabetes and its relationship to the genetic architecture of diabetes. *BMC Med*, 17(1):165, August 2019. ISSN 1741-7015. doi: 10.1186/s12916-019-1392-8.
- Paul M. McKeigue, Andrii Iakovliev, Athina Spiliopoulou, Buddhiprabha Erabadda, and Helen M. Colhoun. Inference of causal and pleiotropic effects with multiple weak genetic instruments: Application to effect of adiponectin on type 2 diabetes, April 2024.
- Jonathan Mertens, Christophe De Block, Maarten Spinhoven, Ann Driessen, Sven M. Francque, and Wilhelmus J. Kwanten. Hepatopathy Associated With Type 1 Diabetes: Distinguishing Non-alcoholic Fatty Liver Disease From Glycogenic Hepatopathy. *Front. Pharmacol.*, 12, October 2021. ISSN 1663-9812. doi: 10.3389/fphar.2021.768576.
- Dominika A Michalek, Courtney Tern, Wei Zhou, Catherine C Robertson, Emily Farber, Paul Campolieto, Wei-Min Chen, Suna Onengut-Gomuscu, and Stephen S Rich. A multi-ancestry genome-wide association study in type 1 diabetes. *Human Molecular Genetics*, 33(11):958–968, June 2024. ISSN 0964-6906. doi: 10.1093/hmg/ddae024.
- Eric Vallabh Minikel, Jeffery L. Painter, Coco Chengliang Dong, and Matthew R. Nelson. Refining the impact of genetic evidence on clinical success. *Nature*, 629(8012):624–629, May 2024. ISSN 1476-4687. doi: 10.1038/s41586-024-07316-0.
- Sunetra Mondal and Joseph M Pappachan. Current perspectives and the future of disease-modifying therapies in type 1 diabetes. *World J Diabetes*, 16(1), January 2025. ISSN 1948-9358. doi: 10.4239/wjd.v16.i1.99496.
- Giada Mondanelli, Elisa Albin, Maria T. Pallotta, Claudia Volpi, Lucienne Chate-noud, Chantal Kuhn, Francesca Fallarino, Davide Matino, Maria L. Belladonna, Roberta Bianchi, Carmine Vacca, Silvio Bicciato, Louis Boon, Giovanni Ricci, Ursula Grohmann, Paolo Puccetti, and Ciriana Orabona. The Proteasome Inhibitor Bortezomib Controls Indoleamine 2,3-Dioxygenase 1 Breakdown and Restores Immune Regulation in Autoimmune Diabetes. *Front Immunol*, 8, April 2017. ISSN 1664-3224. doi: 10.3389/fimmu.2017.00428.

- Antoinette Moran, Brian Bundy, Dorothy J Becker, Linda A DiMeglio, Stephen E Gitelman, Robin Goland, Carla J Greenbaum, Kevan C Herold, Jennifer B Marks, Philip Raskin, Srinath Sanda, Desmond Schatz, Diane K Wherrett, Darrell M Wilson, Jeffrey P Krischer, Jay S Skyler, Linda Pickersgill, Eelco De Koning, Anette-G Ziegler, Bernhard Böehm, Klaus Badenhoop, Nanette Schloot, Jens Friis Bak, Paolo Pozzilli, Didac Mauricio, Marc Y Donath, Luis Castaño, Ana Wägner, Hans Henrik Lervang, Hans Perrild, and Thomas Mandrup-Poulsen. Interleukin-1 antagonism in type 1 diabetes of recent onset: Two multicentre, randomised, double-blind, placebo-controlled trials. *The Lancet*, 381(9881):1905–1915, June 2013. ISSN 0140-6736. doi: 10.1016/s0140-6736(13)60023-9.
- Rhiannon Morris, Liesl Butler, Andrew Perkins, Nadia J. Kershaw, and Jeffrey J. Babon. The Role of LNK (SH2B3) in the Regulation of JAK-STAT Signalling in Haematopoiesis. *Pharmaceuticals*, 15(1), December 2021. ISSN 1424-8247. doi: 10.3390/ph15010024.
- Hakhamanesh Mostafavi, Jeffrey P. Spence, Sahin Naqvi, and Jonathan K. Pritchard. Systematic differences in discovery of genetic effects on gene expression and complex traits. *Nat Genet*, 55(11):1866–1875, November 2023. ISSN 1546-1718. doi: 10.1038/s41588-023-01529-1.
- Robert Moulder, Santosh D. Bhosale, Timo Erkkilä, Essi Laajala, Jussi Salmi, Elizabeth V. Nguyen, Henna Kallionpää, Juha Mykkänen, Mari Vähä-Mäkilä, Heikki Hyöty, Riitta Veijola, Jorma Ilonen, Tuula Simell, Jorma Toppari, Mikael Knip, David R. Goodlett, Harri Lähdesmäki, Olli Simell, and Riitta Lahesmaa. Serum Proteomes Distinguish Children Developing Type 1 Diabetes in a Cohort With HLA-Conferred Susceptibility. *Diabetes*, 64(6):2265–2278, January 2015. ISSN 0012-1797. doi: 10.2337/db14-0983.
- Maki Nakayama, Masao Nagata, Hisafumi Yasuda, Kenji Arisawa, Reiko Kotani, Katsumi Yamada, Shahead Ali Chowdhury, Sagarika Chakrabarty, Zhen Zi Jin, Hideo Yagita, Koichi Yokono, and Masato Kasuga. Fas/Fas Ligand Interactions Play an Essential Role in the Initiation of Murine Autoimmune Diabetes. *Diabetes*, 51(5):1391–1397, May 2002. ISSN 0012-1797. doi: 10.2337/diabetes.51.5.1391.
- Kirsti Näntö-Salonen, Antti Kupila, Satu Simell, Heli Siljander, Tiina Salonsaari, Anne Hekkala, Sari Korhonen, Risto Erkkola, Jukka I. Sipilä, Lotta Haavisto, Marja Siltala, Juhani Tuominen, Jari Hakalax, Heikki Hyöty, Jorma Ilonen,

- Riitta Veijola, Tuula Simell, Mikael Knip, and Olli Simell. Nasal insulin to prevent type 1 diabetes in children with HLA genotypes and autoantibodies conferring increased risk of disease: A double-blind, randomised controlled trial. *The Lancet*, 372(9651):1746–1755, November 2008. ISSN 0140-6736, 1474-547X. doi: 10.1016/S0140-6736(08)61309-4.
- C Nederstigt, B S Uitbeijerse, L G M Janssen, E P M Corssmit, E J P de Koning, and O M Dekkers. Associated auto-immune disease in type 1 diabetes patients: A systematic review and meta-analysis. *European Journal of Endocrinology*, 180(2):135–144, February 2019. ISSN 0804-4643. doi: 10.1530/EJE-18-0515.
- Sergey Nejentsev, Neil Walker, David Riches, Michael Egholm, and John A. Todd. Rare Variants of IFIH1, a Gene Implicated in Antiviral Responses, Protect Against Type 1 Diabetes. *Science*, 324(5925):387–389, April 2009. doi: 10.1126/science.1167728.
- Laura Nigi, Carla Maccora, Francesco Dotta, and Guido Sebastiani. From immunohistological to anatomical alterations of human pancreas in type 1 diabetes: New concepts on the stage. *Diabetes/Metabolism Research and Reviews*, 36(4):e3264, 2020. ISSN 1520-7560. doi: 10.1002/dmrr.3264.
- Lorenza Nisticò, Raffaella Buzzetti, Lynn E. Pritchard, Bart Van der Auwera, Claudio Giovannini, Emanuele Bosi, Maria Teresa Martinez Larrad, Manuel Serrano Rios, C. C. Chow, Clive S. Cockram, Karen Jacobs, Catherine Mijovic, Stephen C. Bain, Anthony H. Barnett, Christina L. Vandewalle, Frans Schuit, Frans K. Gorus, Belgian Diabetes Registry, Roberto Tosi, Paolo Pozzilli, and John A. Todd. The CTLA-4 Gene Region of Chromosome 2q33 Is Linked to, and Associated with, Type 1 Diabetes. *Human Molecular Genetics*, 5(7):1075–1080, July 1996. ISSN 0964-6906. doi: 10.1093/hmg/5.7.1075.
- Janelle A. Noble. Fifty years of HLA-associated type 1 diabetes risk: History, current knowledge, and future directions. *Front. Immunol.*, 15, September 2024. ISSN 1664-3224. doi: 10.3389/fimmu.2024.1457213.
- Janelle A. Noble and Ana M. Valdes. Genetics of the HLA Region in the Prediction of Type 1 Diabetes. *Curr Diab Rep*, 11(6):533–542, December 2011. ISSN 1534-4827. doi: 10.1007/s11892-011-0223-x.

Jessica Nunes, Rakeb Tafesse, Charlene Mao, Matthew Purcell, Xiaokui Mo, Liwen Zhang, Meixiao Long, Matthew G. Cyr, Christoph Rader, and Natarajan Muthusamy. Siglec-6 as a therapeutic target for cell migration and adhesion in chronic lymphocytic leukemia. *Nat Commun*, 15(1):5180, June 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-48678-3.

Helena Oakey, Lynne C. Giles, Rebecca L. Thomson, Kim-Anh Lê Cao, Pat Ashwood, James D. Brown, Emma J. Knight, Simon C. Barry, Maria E. Craig, Peter G. Colman, Elizabeth A. Davis, Emma E. Hamilton-Williams, Leonard C. Harrison, Aveni Haynes, Ki Wook Kim, Kylie-Ann Mallitt, Kelly McGorm, Grant Morahan, William D. Rawlinson, Richard O. Sinnott, Georgia Soldatos, John M. Wentworth, Jennifer J. Couper, and Megan A. S. Penno. Protocol for a nested case-control study design for omics investigations in the Environmental Determinants of Islet Autoimmunity cohort. *Annals of Medicine*, 55(1):2198255, December 2023. ISSN 0785-3890. doi: 10.1080/07853890.2023.2198255.

Alexander J. O'Donovan, Seth Gorelik, and Laura M. Nally. Shifting the paradigm of type 1 diabetes: A narrative review of disease-modifying therapies. *Front. Endocrinol.*, 15, November 2024. ISSN 1664-2392. doi: 10.3389/fendo.2024.1477101.

Masato Ogishi, Rui Yang, Caner Aytekin, David Langlais, Mathieu Bourgey, Taushif Khan, Fatima Al Ali, Mahbuba Rahman, Ottavia M. Delmonte, Maya Chrabieh, Peng Zhang, Conor Gruber, Simon J. Pelham, András N. Spaan, Jérémie Rosain, Wei-Te Lei, Scott Drutman, Matthew D. Hellmann, Margaret K. Callahan, Matthew Adamow, Phillip Wong, Jedd D. Wolchok, Geetha Rao, Cindy S. Ma, Yuka Nakajima, Tomonori Yaguchi, Kenji Chamoto, Samuel C. Williams, Jean-Francois Emile, Flore Rozenberg, Michael S. Glickman, Franck Rapaport, Gaspard Kerner, Garrett Allington, Ilhan Tezcan, Deniz Cagdas, Ferda O. Hosnut, Figen Dogu, Aydan Ikinciogullari, V. Koneti Rao, Leena Kainulainen, Vivien Béziat, Jacinta Bustamante, Silvia Vilarinho, Richard P. Lifton, Bertrand Boisson, Laurent Abel, Dusan Bogunovic, Nico Marr, Luigi D. Notarangelo, Stuart G. Tangye, Tasuku Honjo, Philippe Gros, Stéphanie Boisson-Dupuis, and Jean-Laurent Casanova. Inherited PD-1 deficiency underlies tuberculosis and autoimmunity in a child. *Nat Med*, 27(9):1646–1654, September 2021. ISSN 1546-170X. doi: 10.1038/s41591-021-01388-5.

- Graham D. Ogle, Fei Wang, Aveni Haynes, Gabriel A. Gregory, Thomas W. King, Kylie Deng, Dana Dabelea, Steven James, Alicia J. Jenkins, Xia Li, Ronald C.W. Ma, David M. Maahs, Richard A. Oram, Catherine Pihoker, Jannet Svensson, Zhiguang Zhou, Dianna J. Magliano, and Jayanthi Maniam. Global type 1 diabetes prevalence, incidence, and mortality estimates 2025: Results from the International diabetes Federation Atlas, 11th Edition, and the T1D Index Version 3.0. *Diabetes Research and Clinical Practice*, 225:112277, July 2025. ISSN 0168-8227. doi: 10.1016/j.diabres.2025.112277.
- Suna Onengut-Gumuscu, Wei-Min Chen, Oliver Burren, Nick J. Cooper, Aaron R. Quinlan, Josyf C. Mychaleckyj, Emily Farber, Jessica K. Bonnie, Michal Szpak, Ellen Schofield, Premanand Achuthan, Hui Guo, Mary D. Fortune, Helen Stevens, Neil M. Walker, Lucas D. Ward, Anshul Kundaje, Manolis Kellis, Mark J. Daly, Jeffrey C. Barrett, Jason D. Cooper, Panos Deloukas, John A. Todd, Chris Wallace, Patrick Concannon, and Stephen S. Rich. Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet*, 47(4):381–386, April 2015. ISSN 1546-1718. doi: 10.1038/ng.3245.
- Suna Onengut-Gumuscu, Patrick Concannon, Beena Akolkar, Henry A Erlich, Cécile Julier, Grant Morahan, Concepcion R Nierras, Flemming Pociot, John A Todd, and Stephen S Rich. Type 1 Diabetes Genetics Consortium. *The Journal of Clinical Endocrinology & Metabolism*, 110(6):1505–1513, June 2025. ISSN 0021-972X. doi: 10.1210/clinem/dgaf181.
- Tihamer Orban, Brian Bundy, Dorothy J. Becker, Linda A. DiMeglio, Stephen E. Gitelman, Robin Goland, Peter A. Gottlieb, Carla J. Greenbaum, Jennifer B. Marks, Roshanak Monzavi, Antoinette Moran, Philip Raskin, Henry Rodriguez, William E. Russell, Desmond Schatz, Diane Wherrett, Darrell M. Wilson, Jeffrey P. Krischer, and Jay S. Skyler. Co-stimulation modulation with abatacept in patients with recent-onset type 1 diabetes: A randomised, double-blind, placebo-controlled trial. *The Lancet*, 378(9789):412–419, July 2011. ISSN 0140-6736, 1474-547X. doi: 10.1016/S0140-6736(11)60886-6.
- Tihamer Orban, Brian Bundy, Dorothy J. Becker, Linda A. DiMeglio, Stephen E. Gitelman, Robin Goland, Peter A. Gottlieb, Carla J. Greenbaum, Jennifer B. Marks, Roshanak Monzavi, Antoinette Moran, Mark Peakman, Philip Raskin, William E. Russell, Desmond Schatz, Diane K. Wherrett, Darrell M. Wilson, Jeffrey P. Krischer, Jay S. Skyler, and the Type 1 Diabetes TrialNet Abatacept Study Group. Costimulation Modulation With Abatacept in Patients

With Recent-Onset Type 1 Diabetes: Follow-up 1 Year After Cessation of Treatment. *Diabetes Care*, 37(4):1069–1075, March 2014. ISSN 0149-5992. doi: 10.2337/dc13-0604.

Peter Orchard, Thomas W. Blackwell, Linda Kachuri, Peter J. Castaldi, Michael H. Cho, Stephanie A. Christenson, Peter Durda, Stacey Gabriel, Craig P. Hersh, Scott Huntsman, Seungyong Hwang, Roby Joehanes, Mari Johnson, Xingnan Li, Honghuang Lin, Ching-Ti Liu, Yongmei Liu, Angel C. Y. Mak, Ani W. Manichaikul, David Paik, Aabida Saferali, Joshua D. Smith, Kent D. Taylor, Russell P. Tracy, Jiongming Wang, Mingqiang Wang, Joshua S. Weinstock, Jeffrey Weiss, Heather E. Wheeler, Ying Zhou, Sebastian Zoellner, Joseph C. Wu, Luisa Mestroni, Sharon Graw, Matthew R. G. Taylor, Victor E. Ortega, Craig W. Johnson, Weiniu Gan, Goncalo Abecasis, Deborah A. Nickerson, Namrata Gupta, Kristin Ardlie, Prescott G. Woodruff, Yinan Zheng, Russell P. Bowler, Deborah A. Meyers, Alex Reiner, Charles Kooperberg, Elad Ziv, Vasani S. Ramachandran, Martin G. Larson, L. Adrienne Cupples, Esteban G. Burchard, Edwin K. Silverman, Stephen S. Rich, Nancy Heard-Costa, Hua Tang, Jerome I. Rotter, Albert V. Smith, Daniel Levy, NHLBI TOPMed Consortium Multi-Omics Working Group, NHLBI TOPMed Consortium, François Aguet, Laura Scott, Laura M. Raffield, and Stephen C. J. Parker. Cross-cohort analysis of expression and splicing quantitative trait loci in TOPMed, February 2025.

Klaasjan G. Ouwers, Rick Jansen, Michel G. Nivard, Jenny van Dongen, Maia J. Frieser, Jouke-Jan Hottenga, Wibowo Arindrarto, Annique Claringbould, Maarten van Iterson, Hailiang Mei, Lude Franke, Bastiaan T. Heijmans, Peter A. C. 't Hoen, Joyce van Meurs, Andrew I. Brooks, Brenda W. J. H. Penninx, and Dorret I. Boomsma. A characterization of cis- and trans-heritability of RNA-Seq-based gene expression. *Eur J Hum Genet*, 28(2):253–263, February 2020. ISSN 1476-5438. doi: 10.1038/s41431-019-0511-5.

Fernando Ovalle, Tiffany Grimes, Guanlan Xu, Anish J. Patel, Truman B. Grayson, Lance A. Thielen, Peng Li, and Anath Shalev. Verapamil and beta cell function in adults with recent-onset type 1 diabetes. *Nat Med*, 24(8):1108–1112, August 2018. ISSN 1546-170X. doi: 10.1038/s41591-018-0089-4.

Syamasundar V. Pakala, Pratima Bansal-Pakala, Beth S. Halteman, and Michael Croft. Prevention of diabetes in NOD mice at a late stage by targeting OX40/OX40 ligand interactions. *European Journal of Immunology*, 34(11): 3039–3046, 2004. ISSN 1521-4141. doi: 10.1002/eji.200425141.

- Maria Teresa Pallotta, Sofia Rossini, Chiara Suvieri, Alice Coletti, Ciriana Orabona, Antonio Macchiarulo, Claudia Volpi, and Ursula Grohmann. Indoleamine 2,3-dioxygenase 1 (IDO1): An up-to-date overview of an eclectic immunoregulatory enzyme. *FEBS J*, 289(20):6099–6118, 2022. ISSN 1742-4658. doi: 10.1111/febs.16086.
- Xiongfeng Pan, Atipatsa C. Kaminga, Sanjay Kinra, Shi Wu Wen, Hongying Liu, Xinrui Tan, and Aizhong Liu. Chemokines in Type 1 Diabetes Mellitus. *Front Immunol*, 12, February 2022. ISSN 1664-3224. doi: 10.3389/fimmu.2021.690082.
- Haipeng Pang, Ying Xia, Shuoming Luo, Gan Huang, Xia Li, Zhiguo Xie, and Zhiguang Zhou. Emerging roles of rare and low-frequency genetic variants in type 1 diabetes mellitus. *Journal of Medical Genetics*, 58(5):289–296, May 2021. ISSN 0022-2593, 1468-6244. doi: 10.1136/jmedgenet-2020-107350.
- Haipeng Pang, Jian Lin, Shuoming Luo, Gan Huang, Xia Li, Zhiguo Xie, and Zhiguang Zhou. The missing heritability in type 1 diabetes. *Diabetes, Obesity and Metabolism*, 24(10):1901–1911, October 2022. ISSN 1462-8902. doi: 10.1111/dom.14777.
- Huiwen Pang, Zhuo Chen, and Felicity Y. Han. Nano-based therapy for type 1 diabetes: From immuno-intervention to insulin delivery. *Acta Biomaterialia*, 201:101–120, July 2025. ISSN 1742-7061. doi: 10.1016/j.actbio.2025.06.016.
- Kristen E. Pauken, Marc K. Jenkins, Miyuki Azuma, and Brian T. Fife. PD-1, but Not PD-L1, Expressed by Islet-Reactive CD4+ T Cells Suppresses Infiltration of the Pancreas During Type 1 Diabetes. *Diabetes*, 62(8):2859–2869, July 2013. ISSN 0012-1797. doi: 10.2337/db12-1475.
- Megan Penno. Expanded data mining of accumulated functional genomics data for identifying mechanistic biomarkers informative of T1D risk and prevention. <https://www.breakthrough1d.org/grants/australia/sa/expanded-data-mining-of-accumulated-functional-genomics-data-for-identifying-mechanistic-biomarkers-informative-of-t1d-risk-and-prevention/>, June 2024.
- Judith Pérez-Granado, Janet Piñero, and Laura I. Furlong. Benchmarking post-GWAS analysis tools in major depression: Challenges and implications. *Front Genet*, 13:1006903, October 2022. ISSN 1664-8021. doi: 10.3389/fgene.2022.1006903.

- Mark D. Pescovitz, Carla J. Greenbaum, Heidi Krause-Steinrauf, Dorothy J. Becker, Stephen E. Gitelman, Robin Goland, Peter A. Gottlieb, Jennifer B. Marks, Paula F. McGee, Antoinette M. Moran, Philip Raskin, Henry Rodriguez, Desmond A. Schatz, Diane Wherrett, Darrell M. Wilson, John M. Lachin, and Jay S. Skyler. Rituximab, B-Lymphocyte Depletion, and Preservation of Beta-Cell Function. *New England Journal of Medicine*, 361(22):2143–2152, November 2009. ISSN 0028-4793. doi: 10.1056/NEJMoa0904452.
- Mark D. Pescovitz, Carla J. Greenbaum, Brian Bundy, Dorothy J. Becker, Stephen E. Gitelman, Robin Goland, Peter A. Gottlieb, Jennifer B. Marks, Antoinette Moran, Philip Raskin, Henry Rodriguez, Desmond A. Schatz, Diane K. Wherrett, Darrell M. Wilson, Jeffrey P. Krischer, Jay S. Skyler, and the Type 1 Diabetes TrialNet Anti-CD20 Study Group. B-Lymphocyte Depletion With Rituximab and β -Cell Function: Two-Year Results. *Diabetes Care*, 37(2):453–459, January 2014. ISSN 0149-5992. doi: 10.2337/dc13-0626.
- Maik Pietzner, Eleanor Wheeler, Julia Carrasco-Zanini, Johannes Raffler, Nicola D. Kerrison, Erin Oerton, Victoria P. W. Auyeung, Jian'an Luan, Chris Finan, Juan P. Casas, Rachel Ostroff, Steve A. Williams, Gabi Kastenmüller, Markus Ralser, Eric R. Gamazon, Nicholas J. Wareham, Aroon D. Hingorani, and Claudia Langenberg. Genetic architecture of host proteins involved in SARS-CoV-2 infection. *Nat Commun*, 11(1):6397, December 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-19996-z.
- Maik Pietzner, Eleanor Wheeler, Julia Carrasco-Zanini, Nicola D. Kerrison, Erin Oerton, Mine Koprulu, Jian'an Luan, Aroon D. Hingorani, Steve A. Williams, Nicholas J. Wareham, and Claudia Langenberg. Synergistic insights into human health from aptamer- and antibody-based proteomic profiling. *Nat Commun*, 12(1):6822, November 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-27164-0.
- Juho Piironen and Aki Vehtari. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*, 11(2): 5018–5051, January 2017. ISSN 1935-7524, 1935-7524. doi: 10.1214/17-EJS1337SI.
- Nicolas Poirier, Gilles Blancho, Maryvonne Hiance, Caroline Mary, Tim Van Assche, Jos Lempoels, Steven Ramael, Weirong Wang, Virginie Thepenier, Cecile Braudeau, Nina Salabert, Regis Josien, Ian Anderson, Ian Gourley, Jean-Paul Soulillou, Didier Coquoz, and Bernard Vanhove. First-in-Human Study

- in Healthy Subjects with FR104, a Pegylated Monoclonal Antibody Fragment Antagonist of CD28. *J Immunol*, 197(12):4593–4602, December 2016. ISSN 1550-6606. doi: 10.4049/jimmunol.1601538.
- Alberto Pugliese, Markus Zeller, Alarico Fernandez, Laura J. Zalcberg, Richard J. Bartlett, Camillo Ricordi, Massimo Pietropaolo, George S. Eisenbarth, Simon T. Bennett, and Dhavalkumar D. Patel. The insulin gene is transcribed in the human thymus and transcription levels correlate with allelic variation at the INS VNTR-IDD2 susceptibility locus for type 1 diabetes. *Nat Genet*, 15(3): 293–297, March 1997. ISSN 1546-1718. doi: 10.1038/ng0397-293.
- H-Q Qu, JP Bradfield, SFA Grant, H Hakonarson, and C Polychronakos. Remapping the type 1 diabetes association of the CTLA4 locus. *Genes Immun*, 10 (Suppl 1):S27–S32, December 2009. ISSN 1466-4879. doi: 10.1038/gene.2009.88.
- Teresa Quattrin, Michael J. Haller, Andrea K. Steck, Eric I. Felner, Yinglei Li, Yichuan Xia, Jocelyn H. Leu, Ramin Zoka, Joseph A. Hedrick, Mark R. Rigby, and Frank Vercauteren. Golumab and Beta-Cell Function in Youth with New-Onset Type 1 Diabetes. *New England Journal of Medicine*, 383(21):2007–2017, November 2020. ISSN 0028-4793. doi: 10.1056/NEJMoa2006136.
- 5 Min Read. Immutep Announces Positive Update from Phase I Study of IMP761, a First-in-Class LAG-3 Agonist Antibody for Autoimmune Diseases. <https://www.biospace.com/press-releases/immutep-announces-positive-update-from-phase-i-study-of-imp761-a-first-in-class-lag-3-agonist-antibody-for-autoimmune-diseases>, June 2025.
- Maria J. Redondo, Suna Onengut-Gumuscu, and Kyle J. Gaulton. Genetics of Type 1 Diabetes. In Jean M. Lawrence, Sarah Stark Casagrande, William H. Herman, Deborah J. Wexler, and William T. Cefalu, editors, *Diabetes in America*. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), Bethesda (MD), 2023.
- Trevor W. Reichman, James F. Markmann, Jon Odorico, Piotr Witkowski, John J. Fung, Martin Wijkstrom, Fouad Kandeel, Eelco J. P. de Koning, Anne L. Peters, Chantal Mathieu, Leslie S. Kean, Bote G. Bruinsma, Chenkun Wang, Molly Mascia, Bastiano Sanna, Gautham Marigowda, Felicia Pagliuca, Doug Melton, Camillo Ricordi, and Michael R. Rickels. Stem Cell-Derived, Fully Differentiated Islets for Type 1 Diabetes. *New England Journal of Medicine*, 0 (0), June 2025. ISSN 0028-4793. doi: 10.1056/NEJMoa2506549.

- Shengxiang Ren, Qinrui Tian, Nadav Amar, Hui Yu, Christopher J. Rivard, Charles Caldwell, Terry L. Ng, Megan Tu, Yiwei Liu, Dexiang Gao, Kim Ellison, Kenichi Suda, Leslie Rozeboom, Gareth Rivalland, Paul Mitchell, Caicun Zhou, and Fred R. Hirsch. The immune checkpoint, HVEM may contribute to immune escape in non-small cell lung cancer lacking PD-L1 expression. *Lung Cancer*, 125:115–120, November 2018. ISSN 0169-5002. doi: 10.1016/j.lungcan.2018.09.004.
- Stephen S Rich. Mapping Genes in Diabetes: Genetic Epidemiological Perspective. *Diabetes*, 39(11):1315–1319, November 1990. ISSN 0012-1797. doi: 10.2337/diab.39.11.1315.
- Amanda J. L. Ridley, Yaqing Ou, Richard Karlsson, Nabina Pun, Holly L. Birchenough, Iashia Z. Mulholland, Mary L. Birch, Andrew S. MacDonald, Thomas A. Jowitt, Craig Lawless, Rebecca L. Miller, and Douglas P. Dyer. Chemokines form complex signals during inflammation and disease that can be decoded by extracellular matrix proteoglycans. *Science Signaling*, 16(810): eadf2537, November 2023. doi: 10.1126/scisignal.adf2537.
- Mark R Rigby, Linda A DiMeglio, Marc S Rendell, Eric I Felner, Jean M Dostou, Stephen E Gitelman, Chetanbabu M Patel, Kurt J Griffin, Eva Tsalikian, Peter A Gottlieb, Carla J Greenbaum, Nicole A Sherry, Wayne V Moore, Roshanak Monzavi, Steven M Willi, Philip Raskin, Antoinette Moran, William E Russell, Ashley Pinckney, Lynette Keyes-Elstein, Michael Howell, Sudeepta Aggarwal, Noha Lim, Deborah Phippard, Gerald T Nepom, James McNamara, and Mario R Ehlers. Targeting of memory T cells with alefacept in new-onset type 1 diabetes (T1DAL study): 12 month results of a randomised, double-blind, placebo-controlled phase 2 trial. *The Lancet Diabetes & Endocrinology*, 1(4): 284–294, December 2013. ISSN 2213-8587. doi: 10.1016/S2213-8587(13)70111-6.
- Mark R. Rigby, Kristina M. Harris, Ashley Pinckney, Linda A. DiMeglio, Marc S. Rendell, Eric I. Felner, Jean M. Dostou, Stephen E. Gitelman, Kurt J. Griffin, Eva Tsalikian, Peter A. Gottlieb, Carla J. Greenbaum, Nicole A. Sherry, Wayne V. Moore, Roshanak Monzavi, Steven M. Willi, Philip Raskin, Lynette Keyes-Elstein, S. Alice Long, Sai Kanaparthi, Noha Lim, Deborah Phippard, Carol L. Soppe, Margret L. Fitzgibbon, James McNamara, Gerald T. Nepom, and Mario R. Ehlers. Alefacept provides sustained clinical and immunological effects in new-onset type 1 diabetes patients. *J Clin Invest*, 125(8):3285–3296, September 2016. ISSN 0021-9738. doi: 10.1172/JCI81722.

- N. Risch. Assessing the role of HLA-linked and unlinked determinants of disease. *American Journal of Human Genetics*, 40(1):1, January 1987.
- Catherine C. Robertson, Jamie R. J. Inshaw, Suna Onengut-Gumuscu, Wei-Min Chen, David Flores Santa Cruz, Hanzhi Yang, Antony J. Cutler, Daniel J. M. Crouch, Emily Farber, S. Louis Bridges, Jeffrey C. Edberg, Robert P. Kimberly, Jane H. Buckner, Panos Deloukas, Jasmin Divers, Dana Dabelea, Jean M. Lawrence, Santica Marcovina, Amy S. Shah, Carla J. Greenbaum, Mark A. Atkinson, Peter K. Gregersen, Jorge R. Oksenberg, Flemming Pociot, Marian J. Rewers, Andrea K. Steck, David B. Dunger, Type 1 Diabetes Genetics Consortium, Linda S. Wicker, Patrick Concannon, John A. Todd, and Stephen S. Rich. Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. *Nat Genet*, 53(7):962–971, July 2021. ISSN 1546-1718. doi: 10.1038/s41588-021-00880-5.
- Bart O. Roep, Nanette Solvason, Peter A. Gottlieb, Joana R. F. Abreu, Leonard C. Harrison, George S. Eisenbarth, Liping Yu, Michael Leviten, William A. Hagopian, John B. Buse, Matthias von Herrath, Joanne Quan, Robert S. King, William H. Robinson, Paul J. Utz, Hideki Garren, The BHT-3021 Investigators, and Lawrence Steinman. Plasmid-Encoded Proinsulin Preserves C-Peptide While Specifically Reducing Proinsulin-Specific CD8+ T Cells in Type 1 Diabetes. *Science Translational Medicine*, 5(191):191ra82–191ra82, June 2013. doi: 10.1126/scitranslmed.3006103.
- Mary R. Rooney, Jingsha Chen, Christie M. Ballantyne, Ron C. Hoogeveen, Eric Boerwinkle, Bing Yu, Keenan A. Walker, Pascal Schlosser, Elizabeth Selvin, Nilanjan Chatterjee, David Couper, Morgan E. Grams, and Josef Coresh. Plasma proteomic comparisons change as coverage expands for SomaLogic and Olink, July 2024.
- Michelle Rosenzweig, Randa Salet, Roberta Lorenzon, Nicolas Tchitchek, Alexandra Roux, Claude Bernard, Jean-Claude Carel, Caroline Storey, Michel Polak, Jacques Beltrand, Chloé Amouyal, Agnès Hartemann, Pierre Corbeau, Eric Vicaut, Cecile Bibal, Pierre Bougnères, Tu-Anh Tran, and David Klatzmann. Low-dose IL-2 in children with recently diagnosed type 1 diabetes: A Phase I/II randomised, double-blind, placebo-controlled, dose-finding study. *Diabetologia*, 63(9):1808–1821, September 2020. ISSN 1432-0428. doi: 10.1007/s00125-020-05200-w.

- Sarah H. Ross and Doreen A. Cantrell. Signaling and Function of Interleukin-2 in T Lymphocytes. *Annu Rev Immunol*, 36:411–433, April 2018. ISSN 0732-0582. doi: 10.1146/annurev-immunol-042617-053352.
- Marina Rubey, Nirav Florian Chhabra, Daniel Gradinger, Adrián Sanz-Moreno, Heiko Lickert, Gerhard K.H. Przemeck, and Martin Hrabě de Angelis. DLL1- and DLL4-Mediated Notch Signaling Is Essential for Adult Pancreatic Islet Homeostasis. *Diabetes*, 69(5):915–926, February 2020. ISSN 0012-1797. doi: 10.2337/db19-0795.
- Karina N. Ruiz-Esteves, Kaitlyn R. Shank, Aaron J. Deutsch, Alekhya Gunturi, Natalia Chamorro-Pareja, Caitlin A. Colling, Leyre Zubiri, Katherine Perlman, Tianqi Ouyang, Alexandra-Chloé Villani, Jose C. Florez, Alexander Gusev, Kerry L. Reynolds, Karen K. Miller, Miriam S. Udler, Meghan E. Sise, and Michelle Rengarajan. Identification of Immune Checkpoint Inhibitor–Induced Diabetes. *JAMA Oncol*, 10(10):1409–1416, October 2024. ISSN 2374-2437. doi: 10.1001/jamaoncol.2024.3104.
- William E. Russell, Brian N. Bundy, Mark S. Anderson, Laura A. Cooney, Stephen E. Gitelman, Robin S. Goland, Peter A. Gottlieb, Carla J. Greenbaum, Michael J. Haller, Jeffrey P. Krischer, Ingrid M. Libman, Peter S. Linsley, S. Alice Long, Sandra M. Lord, Daniel J. Moore, Wayne V. Moore, Antoinette M. Moran, Andrew B. Muir, Philip Raskin, Jay S. Skyler, John M. Wentworth, Diane K. Wherrett, Darrell M. Wilson, Anette-Gabriele Ziegler, Kevan C. Herold, and Type 1 Diabetes TrialNet Study Group. Abatacept for Delay of Type 1 Diabetes Progression in Stage 1 Relatives at Risk: A Randomized, Double-Masked, Controlled Trial. *Diabetes Care*, 46(5):1005–1013, March 2023. ISSN 0149-5992. doi: 10.2337/dc22-2200.
- Benoît Salomon, Deborah J. Lenschow, Lesley Rhee, Neda Ashourian, Bhargarith Singh, Arlene Sharpe, and Jeffrey A. Bluestone. B7/CD28 Costimulation Is Essential for the Homeostasis of the CD4+CD25+ Immunoregulatory T Cells that Control Autoimmune Diabetes. *Immunity*, 12(4):431–440, April 2000. ISSN 1074-7613. doi: 10.1016/S1074-7613(00)80195-8.
- Philip Salu, Daniel Tuvin, and Katie M. Reindl. AGR2 knockdown induces ER stress and mitochondria fission to facilitate pancreatic cancer cell death. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1872(1):119854, January 2025. ISSN 0167-4889. doi: 10.1016/j.bbamcr.2024.119854.

- Frantisek Saudek, Tereza Havrdova, Petr Boucek, Ludmila Karasova, Peter Novota, and Jelena Skibova. Polyclonal Anti-T-Cell Therapy for Type 1 Diabetes Mellitus of Recent Onset. *Rev Diabet Stud*, 1(2):80–88, 2004. ISSN 1613-6071. doi: 10.1900/RDS.2004.1.80.
- Julia Schanin, Wouter Korver, Emily C. Brock, John Leung, Zachary Benet, Thuy Luu, Katherine Chang, Alan Xu, Naomi De Freitas, Kenneth Luehrsen, Michael A. Brehm, Alan Wong, and Bradford A. Youngblood. Discovery of an agonistic Siglec-6 antibody that inhibits and reduces human mast cells. *Commun Biol*, 5(1):1226, November 2022. ISSN 2399-3642. doi: 10.1038/s42003-022-04207-w.
- Amand F. Schmidt, Chris Finan, Maria Gordillo-Marañón, Folkert W. Asselbergs, Daniel F. Freitag, Riyaz S. Patel, Benoît Tyl, Sandesh Chopade, Rupert Faraway, Magdalena Zwierzyna, and Aroon D. Hingorani. Genetic drug target validation using Mendelian randomisation. *Nat Commun*, 11(1):3255, June 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-16969-0.
- Célia Seillier, Pauline Hélie, Gautier Petit, Denis Vivien, Diego Clemente, Brigitte Le Mauff, Fabian Docagne, and Olivier Toutirais. Roles of the tissue-type plasminogen activator in immune response. *Cellular Immunology*, 371:104451, January 2022. ISSN 0008-8749. doi: 10.1016/j.cellimm.2021.104451.
- Nigel Sharfe, Harjit K. Dadi, Michal Shahr, and Chaim M. Roifman. Human immune disorder arising from mutation of the α chain of the interleukin-2 receptor. *Proceedings of the National Academy of Sciences*, 94(7):3168–3171, April 1997. doi: 10.1073/pnas.94.7.3168.
- Ashok Sharma, Xiang Liu, David Hadley, William Hagopian, Wei-Min Chen, Suna Onengut-Gumuscu, Carina Törn, Andrea K. Steck, Brigitte I. Frohnert, Marian Rewers, Anette-G. Ziegler, Åke Lernmark, Jorma Toppari, Jeffrey P. Krischer, Beena Akolkar, Stephen S. Rich, and Jin-Xiong She. Identification of non-HLA genes associated with development of islet autoimmunity and type 1 diabetes in the prospective TEDDY cohort. *Journal of Autoimmunity*, 89:90–100, May 2018. ISSN 0896-8411. doi: 10.1016/j.jaut.2017.12.008.
- Nicole Sherry, William Hagopian, Johnny Ludvigsson, Sunil M. Jain, Jack Wahlen, Robert J. Ferry, Bruce Bode, Stephen Aronoff, Christopher Holland, David Carlin, Karen L. King, Ronald L. Wilder, Stanley Pillemer, Ezio Bonvini, Syd Johnson, Kathryn E. Stein, Scott Koenig, Kevan C. Herold, and

- Anastasia G. Daifotis. Teplizumab for treatment of type 1 diabetes (Protégé study): 1-year results from a randomised, placebo-controlled trial. *The Lancet*, 378(9790):487–497, August 2011. ISSN 0140-6736, 1474-547X. doi: 10.1016/S0140-6736(11)60931-8.
- Beverley M. Shields, Timothy J. McDonald, Richard Oram, Anita Hill, Michelle Hudson, Pia Leete, Ewan R. Pearson, Sarah J. Richardson, Noel G. Morgan, Andrew T. Hattersley, and The T1G Consortium. C-Peptide Decline in Type 1 Diabetes Has Two Phases: An Initial Exponential Fall and a Subsequent Stable Phase. *Diabetes Care*, 41(7):1486–1492, June 2018. ISSN 0149-5992. doi: 10.2337/dc18-0465.
- Janet Silverstein, Noel Maclaren, William Riley, Rebecca Spillar, Doreen Radjenovic, and Suzanne Johnson. Immunosuppression with Azathioprine and Prednisone in Recent-Onset Insulin-Dependent Diabetes Mellitus. *New England Journal of Medicine*, 319(10):599–604, September 1988. ISSN 0028-4793. doi: 10.1056/NEJM198809083191002.
- Emily K. Sims, Brian N. Bundy, Kenneth Stier, Elisavet Serti, Noha Lim, S. Alice Long, Susan M. Geyer, Antoinette Moran, Carla J. Greenbaum, Carmella Evans-Molina, Kevan C. Herold, and Type 1 Diabetes TrialNet Study Group. Teplizumab improves and stabilizes beta cell function in antibody-positive high-risk individuals. *Sci Transl Med*, 13(583):eabc8980, March 2021. doi: 10.1126/scitranslmed.abc8980.
- Emily K. Sims, Rachel E.J. Besser, Colin Dayan, Cristy Geno Rasmussen, Carla Greenbaum, Kurt J. Griffin, William Hagopian, Mikael Knip, Anna E. Long, Frank Martin, Chantal Mathieu, Marian Rewers, Andrea K. Steck, John M. Wentworth, Stephen S. Rich, Olga Kordonouri, Anette-Gabriele Ziegler, Kevan C. Herold, and for the NIDDK Type 1 Diabetes TrialNet Study Group. Screening for Type 1 Diabetes in the General Population: A Status Report and Perspective. *Diabetes*, 71(4):610–623, March 2022. ISSN 0012-1797. doi: 10.2337/dbi20-0054.
- Jay S. Skyler, Todd J. Lorenz, Sherwyn Schwartz, George S. Eisenbarth, Daniel Einhorn, Jerry P. Palmer, Jennifer B. Marks, Carla Greenbaum, Elizabeth A. Saria, and Vera Byers. Effects of an anti-CD5 immunoconjugate (CD5-Plus) in recent onset type I diabetes mellitus: A preliminary investigation. *J Diabetes Complications*, 7(4):224–232, October 1993. ISSN 1056-8727. doi: 10.1016/S0002-9610(05)80249-1.

Blair H. Smith, Harry Campbell, Douglas Blackwood, John Connell, Mike Connor, Ian J. Deary, Anna F. Dominiczak, Bridie Fitzpatrick, Ian Ford, Cathy Jackson, Gillian Haddow, Shona Kerr, Robert Lindsay, Mark McGilchrist, Robin Morton, Graeme Murray, Colin NA Palmer, Jill P. Pell, Stuart H. Ralston, David St Clair, Frank Sullivan, Graham Watt, Roland Wolf, Alan Wright, David Porteous, and Andrew D. Morris. Generation Scotland: The Scottish Family Health Study; a new resource for researching genes and heritability. *BMC Medical Genetics*, 7(1):74, October 2006. ISSN 1471-2350. doi: 10.1186/1471-2350-7-74.

Blair H Smith, Archie Campbell, Pamela Linksted, Bridie Fitzpatrick, Cathy Jackson, Shona M Kerr, Ian J Deary, Donald J MacIntyre, Harry Campbell, Mark McGilchrist, Lynne J Hocking, Lucy Wisely, Ian Ford, Robert S Lindsay, Robin Morton, Colin N A Palmer, Anna F Dominiczak, David J Porteous, and Andrew D Morris. Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *International Journal of Epidemiology*, 42(3): 689–700, June 2013. ISSN 0300-5771. doi: 10.1093/ije/dys084.

Logan K Smith, Kareem Fawaz, and Bebhinn Treanor. Galectin-9 regulates the threshold of B cell activation and autoimmunity. *eLife*, 10:e64557, August 2021. ISSN 2050-084X. doi: 10.7554/eLife.64557.

Deborah Smyth, Jason D. Cooper, Joanne E. Collins, Joanne M. Heward, Jayne A. Franklyn, Joanna M.M. Howson, Adrian Vella, Sarah Nutland, Helen E. Rance, Lisa Maier, Bryan J. Barratt, Cristian Guja, Constantin Ionescu-Tîrgoviște, David A. Savage, David B. Dunger, Barry Widmer, David P. Strachan, Susan M. Ring, Neil Walker, David G. Clayton, Rebecca C.J. Twells, Stephen C.L. Gough, and John A. Todd. Replication of an Association Between the Lymphoid Tyrosine Phosphatase Locus (LYP/PTPN22) With Type 1 Diabetes, and Evidence for Its Role as a General Autoimmunity Locus. *Diabetes*, 53(11):3020–3023, November 2004. ISSN 0012-1797. doi: 10.2337/diabetes.53.11.3020.

Deborah J. Smyth, Jason D. Cooper, Rebecca Bailey, Sarah Field, Oliver Burren, Luc J. Smink, Cristian Guja, Constantin Ionescu-Tirgoviste, Barry Widmer, David B. Dunger, David A. Savage, Neil M. Walker, David G. Clayton, and John A. Todd. A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. *Nat Genet*, 38(6):617–619, June 2006. ISSN 1546-1718. doi: 10.1038/ng1800.

- Deborah J. Smyth, Jason D. Cooper, Joanna M.M. Howson, Pamela Clarke, Kate Downes, Trupti Mistry, Helen Stevens, Neil M. Walker, and John A. Todd. FUT2 Nonsecretor Status Links Type 1 Diabetes Susceptibility and Resistance to Infection. *Diabetes*, 60(11):3081–3084, October 2011. ISSN 0012-1797. doi: 10.2337/db11-0638.
- Yao Song, Liu Wang, Lan Zhang, and Dingming Huang. The involvement of semaphorin 7A in tumorigenic and immunoinflammatory regulation. *Journal of Cellular Physiology*, 236(9):6235–6248, 2021. ISSN 1097-4652. doi: 10.1002/jcp.30340.
- Athina Spiliopoulou, Marco Colombo, Darren Plant, Nisha Nair, Jing Cui, Marieke JH Coenen, Katsunori Ikari, Hisashi Yamanaka, Saedis Saevarsdottir, Leonid Padyukov, S. Louis Bridges Jr, Robert P. Kimberly, Yukinori Okada, Piet L. CM van Riel, Gertjan Wolbink, Irene E. van der Horst-Bruinsma, Niek de Vries, Paul P. Tak, Koichiro Ohmura, Helena Canhã, Henk-Jan Guchelaar, Tom WJ Huizinga, Lindsey A. Criswell, Soumya Raychaudhuri, Michael E. Weinblatt, Anthony G. Wilson, Xavier Mariette, John D. Isaacs, Ann W. Morgan, Costantino Pitzalis, Anne Barton, and Paul McKeigue. Association of response to TNF inhibitors in rheumatoid arthritis with quantitative trait loci for CD40 and CD39. *Ann Rheum Dis*, 78(8):1055–1061, August 2019. ISSN 0003-4967, 1468-2060. doi: 10.1136/annrheumdis-2018-214877.
- Athina Spiliopoulou, Andrii Iakovliev, Darren Plant, Megan Sutcliffe, Seema Sharma, Cankut Cubuk, Myles Lewis, Costantino Pitzalis, Anne Barton, and Paul M. McKeigue. Genome-Wide Aggregated Trans Effects Analysis Identifies Genes Encoding Immune Checkpoints as Core Genes for Rheumatoid Arthritis. *Arthritis Rheumatol*, n/a(n/a), 2025. ISSN 2326-5205. doi: 10.1002/art.43125.
- Lynn Starr, Sanjoy Dutta, Thomas Danne, Stephen R. Karpen, Campbell Hutton, and Aaron Kowalski. The Urgent Need for Breakthrough Therapies and a World Without Type 1 Diabetes. *Diabetes Ther*, 16(6):1063–1076, June 2025. ISSN 1869-6953. doi: 10.1007/s13300-025-01735-6.
- Andrea K Steck, Ping Xu, Susan Geyer, Maria J Redondo, Peter Antinozzi, John M Wentworth, Jay Sosenko, Suna Onengut-Gumuscu, Wei-Min Chen, Stephen S Rich, Alberto Pugliese, and the Type 1 Diabetes TrialNet Study Group. Can Non-HLA Single Nucleotide Polymorphisms Help Stratify Risk in TrialNet Relatives at Risk for Type 1 Diabetes? *The Journal of Clinical*

Endocrinology & Metabolism, 102(8):2873–2880, August 2017. ISSN 0021-972X. doi: 10.1210/jc.2016-4003.

Andreas Strasser, Philipp J. Jost, and Shigekazu Nagata. The Many Roles of FAS Receptor Signaling in the Immune System. *Immunity*, 30(2):180–192, February 2009. ISSN 1074-7613. doi: 10.1016/j.immuni.2009.01.001.

Benjamin B. Sun, Joseph C. Maranville, James E. Peters, David Stacey, James R. Staley, James Blackshaw, Stephen Burgess, Tao Jiang, Ellie Paige, Praveen Surendran, Clare Oliver-Williams, Mihir A. Kamat, Bram P. Prins, Sheri K. Wilcox, Erik S. Zimmerman, An Chi, Narinder Bansal, Sarah L. Spain, Angela M. Wood, Nicholas W. Morrell, John R. Bradley, Nebojsa Janjic, David J. Roberts, Willem H. Ouwehand, John A. Todd, Nicole Soranzo, Karsten Suhre, Dirk S. Paul, Caroline S. Fox, Robert M. Plenge, John Danesh, Heiko Runz, and Adam S. Butterworth. Genomic atlas of the human plasma proteome. *Nature*, 558(7708):73–79, June 2018. ISSN 1476-4687. doi: 10.1038/s41586-018-0175-2.

Benjamin B. Sun, Joshua Chiou, Matthew Traylor, Christian Benner, Yi-Hsiang Hsu, Tom G. Richardson, Praveen Surendran, Anubha Mahajan, Chloe Robins, Steven G. Vasquez-Grinnell, Liping Hou, Erika M. Kvikstad, Oliver S. Burren, Jonathan Davitte, Kyle L. Ferber, Christopher E. Gillies, Åsa K. Hedman, Sile Hu, Tinchu Lin, Rajesh Mikkilineni, Rion K. Pendergrass, Corran Pickering, Bram Prins, Denis Baird, Chia-Yen Chen, Lucas D. Ward, Aimee M. Deaton, Samantha Welsh, Carissa M. Willis, Nick Lehner, Matthias Arnold, Maria A. Wörheide, Karsten Suhre, Gabi Kastenmüller, Anurag Sethi, Madeleine Cule, Anil Raj, Hyun Ming Kang, Lucy Burkitt-Gray, Eugene Melamed, Mary Helen Black, Eric B. Fauman, Joanna M. M. Howson, Hyun Min Kang, Mark I. McCarthy, Paul Nioi, Slavé Petrovski, Robert A. Scott, Erin N. Smith, Sándor Szalma, Dawn M. Waterworth, Lyndon J. Mitnau, Joseph D. Szustakowski, Bradford W. Gibson, Melissa R. Miller, and Christopher D. Whelan. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*, 622(7982):329–338, October 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06592-6.

Xiujie Sun, Bogang Wu, Huai-Chin Chiang, Hui Deng, Xiaowen Zhang, Wei Xiong, Junquan Liu, Aaron M. Rozeboom, Brent T. Harris, Eline Blommaert, Antonio Gomez, Roderic Espin Garcia, Yufan Zhou, Payal Mitra, Madeleine Prevost, Deyi Zhang, Debarati Banik, Claudine Isaacs, Deborah Berry, Catherine Lai, Krysta Chaldekis, Patricia S. Latham, Christine A. Brantner, Anastas

- Popratiloff, Victor X. Jin, Ningyan Zhang, Yanfen Hu, Miguel Angel Pujana, Tyler J. Curiel, Zhiqiang An, and Rong Li. Tumour DDR1 promotes collagen fibre alignment to instigate immune exclusion. *Nature*, 599(7886):673–678, November 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-04057-2.
- Kazuhiro Suzuki, Tatsusada Okuno, Midori Yamamoto, R. Jeroen Pasterkamp, Noriko Takegahara, Hyota Takamatsu, Tomoe Kitao, Junichi Takagi, Paul D. Rennert, Alex L. Kolodkin, Atsushi Kumanogoh, and Hitoshi Kikutani. Semaphorin 7A initiates T-cell-mediated inflammatory responses through $\alpha 1\beta 1$ integrin. *Nature*, 446(7136):680–684, April 2007. ISSN 1476-4687. doi: 10.1038/nature05652.
- Megumi Tachibana and Akihisa Imagawa. Type 1 diabetes related to immune checkpoint inhibitors. *Best Practice & Research Clinical Endocrinology & Metabolism*, 36(3):101657, May 2022. ISSN 1521-690X. doi: 10.1016/j.beem.2022.101657.
- Ralf Tambets, Anastassia Kolde, Peep Kolberg, Michael I. Love, and Kaur Alasoo. Extensive co-regulation of neighboring genes complicates the use of eQTLs in target gene prioritization. *HGGADVANCE*, 5(4), October 2024. ISSN 2666-2477. doi: 10.1016/j.xhgg.2024.100348.
- Wei Tang, Jun Chen, Tianlong Ji, and Xiufeng Cong. TIGIT, a novel immune checkpoint therapy for melanoma. *Cell Death Dis*, 14(7):1–9, July 2023. ISSN 2041-4889. doi: 10.1038/s41419-023-05961-3.
- Danijela Tatovic, Ashish Marwaha, Peter Taylor, Stephanie J. Hanna, Kym Carter, W. Y. Cheung, Steve Luzio, Gareth Dunseath, Hayley A. Hutchings, Gail Holland, Steve Hiles, Greg Fegan, Evangelia Williams, Jennie H. M. Yang, Clara Domingo-Vila, Emily Pollock, Muntaha Wadud, Kirsten Ward-Hartstonge, Susie Marques-Jones, Jane Bowen-Morris, Rachel Stenson, Megan K. Levings, John W. Gregory, Timothy I. M. Tree, and Colin Dayan. Ustekinumab for type 1 diabetes in adolescents: A multicenter, double-blind, randomized phase 2 trial. *Nat Med*, 30(9):2657–2666, September 2024. ISSN 1546-170X. doi: 10.1038/s41591-024-03115-2.
- The Diabetes Prevention Trial-Type 1 Study Group. Effects of Oral Insulin in Relatives of Patients With Type 1 Diabetes: The Diabetes Prevention Trial–Type 1. *Diabetes Care*, 28(5):1068–1076, May 2005. ISSN 0149-5992. doi: 10.2337/diacare.28.5.1068.

- The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res*, 51(D1):D523–D531, January 2023. ISSN 0305-1048. doi: 10.1093/nar/gkac1052.
- Puchong Thirawatananond, Matthew E. Brown, Lindsey K. Sachs, Juan M. Arnoletti, Wen-I Yeh, Amanda L. Posgai, Melanie R. Shapiro, Yi-Guang Chen, and Todd M. Brusko. Treg-Specific CD226 Deletion Reduces Diabetes Incidence in NOD Mice by Improving Regulatory T-Cell Stability. *Diabetes*, 72(11): 1629–1640, August 2023. ISSN 0012-1797. doi: 10.2337/db23-0307.
- David Thomas, Daniel Chancellor, Amanda Micklus, Sara LaFever, Michael Hay, Shomesh Chaudhuri, Robert Bowden, and Andrew W. Lo. Clinical Development Success Rates and Contributing Factors 2011–2020. Technical report, Biotechnology Innovation Organization; Informa Pharma Intelligence; Quantitative Life Sciences, February 2021.
- James Tobin. Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26(1):24–36, 1958. ISSN 0012-9682. doi: 10.2307/1907382.
- John A. Todd, Neil M. Walker, Jason D. Cooper, Deborah J. Smyth, Kate Downes, Vincent Plagnol, Rebecca Bailey, Sergey Nejentsev, Sarah F. Field, Felicity Payne, Christopher E. Lowe, Jeffrey S. Szeszko, Jason P. Hafler, Lauren Zeitels, Jennie H. M. Yang, Adrian Vella, Sarah Nutland, Helen E. Stevens, Helen Schuilenburg, Gillian Coleman, Meeta Maisuria, William Meadows, Luc J. Smink, Barry Healy, Oliver S. Burren, Alex A. C. Lam, Nigel R. Ovington, James Allen, Ellen Adlem, Hin-Tak Leung, Chris Wallace, Joanna M. M. Howson, Cristian Guja, Constantin Ionescu-Tîrgoviște, Matthew J. Simmonds, Joanne M. Heward, Stephen C. L. Gough, David B. Dunger, Linda S. Wicker, and David G. Clayton. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet*, 39(7):857–864, July 2007. ISSN 1546-1718. doi: 10.1038/ng2068.
- Andrea Tompa and Maria Faresjö. Shift in the B cell subsets between children with type 1 diabetes and/or celiac disease. *Clin Exp Immunol*, 216(1):36–44, April 2024. ISSN 1365-2249. doi: 10.1093/cei/uxad136.
- Maomeng Tong, Ian McHardy, Paul Ruegger, Maryam Goudarzi, Purna C. Kashyap, Talin Haritunians, Xiaoxiao Li, Thomas G. Graeber, Emma Schwager, Curtis Huttenhower, Albert J. Fornace, Justin L. Sonnenburg, Dermot PB McGovern, James Borneman, and Jonathan Braun. Reprogramming of gut

- microbiome energy metabolism by the FUT2 Crohn's disease risk polymorphism. *ISME J*, 8(11):2193–2206, November 2014. ISSN 1751-7370. doi: 10.1038/ismej.2014.64.
- Prerak Trivedi, Gaurang Jhala, David J. De George, Chris Chiu, Claudia Selck, Tingting Ge, Tara Catterall, Lorraine Elkerbout, Louis Boon, Nicole Joller, Thomas W. Kay, Helen E. Thomas, and Balasubramanian Krishnamurthy. TIGIT acts as an immune checkpoint upon inhibition of PD1 signaling in autoimmune diabetes. *Front Immunol*, 15, March 2024. ISSN 1664-3224. doi: 10.3389/fimmu.2024.1370907.
- Prerak M. Trivedi, Stacey Fynch, Lucy M. Kennedy, Jonathan Chee, Balasubramanian Krishnamurthy, Lorraine A. O'Reilly, Andreas Strasser, Thomas W. H. Kay, and Helen E. Thomas. Soluble FAS ligand is not required for pancreatic islet inflammation or beta-cell destruction in non-obese diabetic mice. *Cell Death Discovery*, 5(1):1–5, September 2019. ISSN 2058-7716. doi: 10.1038/s41420-019-0217-z.
- Takeshi Tsubata. The ligand interactions of B cell Siglecs are involved in the prevention of autoimmunity to sialylated self-antigens and in the quality control of signaling-competent B cells. *International Immunology*, 35(10):461–473, October 2023. ISSN 1460-2377. doi: 10.1093/intimm/dxad030.
- Mauro Tutino, Nancy Yiu-Lin Yu, Konstantinos Hatzikotoulas, Young-Chan Park, Peter Kreitmaier, Georgia Katsoula, Reinhard Berner, Kristina Casteels, Helena Elding Larsson, Olga Kordonouri, Mariusz Ołtarzewski, Agnieszka Szybowska, Raffael Ott, Andreas Weiss, Christiane Winkler, Jose Zapardiel-Gonzalo, Agnese Petrera, Stefanie M. Hauck, Ezio Bonifacio, Anette-Gabriele Ziegler, and Eleftheria Zeggini. Genetics of circulating proteins in newborn babies at high risk of type 1 diabetes. *Nat Commun*, 16(1):3750, April 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-58972-3.
- Jay Tuttle, Edit Drescher, Jesus Abraham Simón-Campos, Paul Emery, Maria Greenwald, Alan Kivitz, Hyungmin Rha, Pia Yachi, Christina Kiley, and Ajay Nirula. A Phase 2 Trial of Peresolimab for Adults with Rheumatoid Arthritis. *New England Journal of Medicine*, 388(20):1853–1862, May 2023. ISSN 0028-4793. doi: 10.1056/NEJMoa2209856.

Shay Tzaban, Ori Stern, Elad Zisman, Galit Eisenberg, Shiri Klein, Shoshana Frankenburg, and Michal Lotem. Alternative splicing of modulatory immune receptors in T lymphocytes: A newly identified and targetable mechanism for anticancer immunotherapy. *Front Immunol*, 15:1490035, January 2025. ISSN 1664-3224. doi: 10.3389/fimmu.2024.1490035.

Hironori Ueda, Joanna M. M. Howson, Laura Esposito, Joanne Heward, Snook, Giselle Chamberlain, Daniel B. Rainbow, Kara M. D. Hunter, Annabel N. Smith, Gianfranco Di Genova, Mathias H. Herr, Ingrid Dahlman, Felicity Payne, Deborah Smyth, Christopher Lowe, Rebecca C. J. Twells, Sarah Howlett, Barry Healy, Sarah Nutland, Helen E. Rance, Vin Everett, Luc J. Smink, Alex C. Lam, Heather J. Cordell, Neil M. Walker, Cristina Bordin, John Hulme, Costantino Motzo, Francesco Cucca, J. Fred Hess, Michael L. Metzker, Jane Rogers, Simon Gregory, Amit Allahabadia, Ratnasingam Nithiyananthan, Eva Tuomilehto-Wolf, Jaakko Tuomilehto, Polly Bingley, Kathleen M. Gillespie, Dag E. Undlien, Kjersti S. Rønningen, Cristian Guja, Constantin Ionescu-Tîrgoviște, David A. Savage, A. Peter Maxwell, Dennis J. Carson, Chris C. Patterson, Jayne A. Franklyn, David G. Clayton, Laurence B. Peterson, Linda S. Wicker, John A. Todd, and Stephen C. L. Gough. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. *Nature*, 423(6939):506–511, May 2003. ISSN 1476-4687. doi: 10.1038/nature01621.

Emil Uffelmann, Qin Qin Huang, Nchangwi Syntia Munung, Jantina de Vries, Yukinori Okada, Alicia R. Martin, Hilary C. Martin, Tuuli Lappalainen, and Danielle Posthuma. Genome-wide association studies. *Nat Rev Methods Primers*, 1(1):59, August 2021. ISSN 2662-8449. doi: 10.1038/s43586-021-00056-9.

Michiaki Unno, Koji Nata, Naoya Noguchi, Yoichi Narushima, Takako Akiyama, Takayuki Ikeda, Kei Nakagawa, Shin Takasawa, and Hiroshi Okamoto. Production and Characterization of Reg Knockout Mice: Reduced Proliferation of Pancreatic β -Cells in Reg Knockout Mice. *Diabetes*, 51(suppl_3):S478–S483, December 2002. ISSN 0012-1797. doi: 10.2337/diabetes.51.2007.S478.

Petros Vafiadis, Simon T. Bennett, John A. Todd, Joseph Nadeau, Rosemarie Grabs, Cynthia G. Goodyer, Saman Wickramasinghe, Eleanor Colle, and Constantin Polychronakos. Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. *Nat Genet*, 15(3):289–292, March 1997. ISSN 1546-1718. doi: 10.1038/ng0397-289.

- Jean Van Rampelbergh, Peter Achenbach, Richard David Leslie, Mohammad Al-hadj Ali, Colin Dayan, Bart Keymeulen, Katharine R. Owen, Martin Kindermans, Frédéric Parmentier, Vincent Carlier, Roxana R. Ahangarani, Evelien Gebruers, Nicolas Bovy, Luc Vanderelst, Marcelle Van Mechelen, Pierre Vandepapelière, and Christian Boitard. First-in-human, double-blind, randomized phase 1b study of peptide immunotherapy IMCY-0098 in new-onset type 1 diabetes. *BMC Medicine*, 21(1):190, May 2023. ISSN 1741-7015. doi: 10.1186/s12916-023-02900-z.
- E. Vandemeulebroucke, F. K. Gorus, K. Decochez, I. Weets, B. Keymeulen, C. De Block, J. Tits, D. G. Pipeleers, and C. Mathieu. Insulin treatment in IA-2A-positive relatives of type 1 diabetic patients. *Diabetes & Metabolism*, 35(4):319–327, September 2009. ISSN 1262-3636. doi: 10.1016/j.diabet.2009.02.005.
- Julio Angel Vázquez-Martínez, Miguel Angel Gómez-Lim, Edgar Morales-Ríos, Jorge Alberto Gonzalez-y-Merchand, and Vianney Ortiz-Navarrete. Short Disordered Epitope of CRTAM Ig-Like V Domain as a Potential Target for Blocking Antibodies. *Int J Mol Sci*, 21(22):8798, November 2020. doi: 10.3390/ijms21228798.
- Federica Vecchio, Gloria Messina, Anna Giovenzana, and Alessandra Petrelli. New Evidence of Exocrine Pancreatopathy in Pre-symptomatic and Symptomatic Type 1 Diabetes. *Curr Diab Rep*, 19(10):92, August 2019. ISSN 1539-0829. doi: 10.1007/s11892-019-1223-5.
- Kendra Vehik, Steven W. Fiske, Chad A. Logan, Daniel Agardh, Corrado M. Cilio, William Hagopian, Olli Simell, Merja Roivainen, Jin-Xiong She, Thomas Briese, Sami Oikarinen, Heikki Hyoty, Anette-G. Ziegler, Marian Rewers, Ake Lernmark, Beena Akolkar, Jeffrey P. Krischer, Brant R. Burkhardt, and The TEDDY Study Group. Methods, quality control and specimen management in an international multicentre investigation of type 1 diabetes: TEDDY. *Diabetes/Metabolism Research and Reviews*, 29(7):557–567, 2013. ISSN 1520-7560. doi: 10.1002/dmrr.2427.
- Adrian Vella, Jason D. Cooper, Christopher E. Lowe, Neil Walker, Sarah Nutland, Barry Widmer, Richard Jones, Susan M. Ring, Wendy McArdle, Marcus E. Pembrey, David P. Strachan, David B. Dunger, C. J. Rebecca Twells, David G. Clayton, and John A. Todd. Localization of a Type 1 Diabetes Locus in the IL2RA/CD25 Region by Use of Tag Single-Nucleotide Polymorphisms. *The*

American Journal of Human Genetics, 76(5):773–779, May 2005. ISSN 0002-9297, 1537-6605. doi: 10.1086/429843.

Guillaume Voisinne, Anne Gonzalez de Peredo, and Romain Roncagalli. CD5, an Undercover Regulator of TCR Signaling. *Front. Immunol.*, 9, December 2018. ISSN 1664-3224. doi: 10.3389/fimmu.2018.02900.

Matthias von Herrath, Stephen C. Bain, Bruce Bode, Jesper Ole Clausen, Ken Coppieters, Leylya Gaysina, Janusz Gumprecht, Troels Krarup Hansen, Chantal Mathieu, Cristobal Morales, Ofri Mosenzon, Stine Segel, George Tsoukas, Thomas R. Pieber, Bernhard Ludvik, Rudolf Prager, Bernhard Paulweber, Christoph F. Ebenbichler, B. Keymeulen, C. De Block, Loren Grossman, Robyn Houlden, Patrice Perron, Thomas Ransom, Peter Senior, S. John Weisnagel, Vincent Woo, Richard Dumas, David Thompson, Tina Vilsbøll, Jeppe Gram, Claus Bogh Juhl, Janne Hukkanen, Jorma Lahtela, Leo Niskanen, Donal O'Shea, Timothy O'Brien, Seamus Sreenan, Julio Wainstein, Moshe Phillip, Hilla Knobler, Francesco Dotta, Pier Marco Piatti, Trevisan Roberto, Agostino Gnasso, Hanne Gulseth, John Cooper, Ewa Pankowska, Monika Lukaszewicz, Bogumił Wolnik, Isabel Manita, Olinda Marques, Cristina Roque, Rosa Maria Príncipe, Celestino Neves, Susana Heitor, L. Ruyatkina, Irina Dvoryashina, Gulnar Vagapova, Lidiya Belousova, Marina Sergeeva-Kondrachenko, Andrey Peskov, Elena Frolova, Albina Golovach, Marina Kunit-syna, Natalia Krasnopeevea (Kabachkova), Irina Ipatko, Carmen De la Cuesta, Francisco José Tinahones, Mercedes Rigla, Juan Francisco Merino, Luis Alberto Gómez, Mercè Fernández, Rafael Simó, Mikael Rydén, Johan Jendle, Karin Filipsson, Björn Eliasson, Borys Mankovsky, Iurii Lymar, Liubov Sokolova, Galyna Myshanych, Tetiana Zlova, Maryna Vlasenko, Petro Kuskalo -, Hamish Courtney, Colin Dayan, Patrick English, Simon Heller, Andrew B. Johnson, Sunil Nair, R. D. Leslie, P. Narendran, Nick Oliver, Shenaz Ramtoola, Jim Shaw, Adie Viljoen -, Amer Al-Karadsheh, Jean Marie Dostou, Sumana Gangi, Peter Gottlieb, Terri Jerkins, Michael Magnotti, Jennifer Marks, Samer Nakhle, Gholamreza Bonabi, Lyle Myers, Richard Pratley, William Hagopian, Jeremy Pettus, Bernt Johan von Scholten, Johnna D. Wesley, and Frederik F. Kreiner. Anti-interleukin-21 antibody and liraglutide for the preservation of β -cell function in adults with recent-onset type 1 diabetes: A randomised, double-blind, placebo-controlled, phase 2 trial. *The Lancet Diabetes & Endocrinology*, 9(4):212–224, April 2021. ISSN 2213-8587, 2213-8595. doi: 10.1016/S2213-8587(21)00019-X.

- Matthias G. von Herrath. Combination therapy with anti-CD6 and oral insulin immunization reverses recent onset diabetes in non obese diabetic mice but fails to induce lasting tolerance. *Clinical Immunology*, 149(3, Part B):440–441, December 2013. ISSN 1521-6616. doi: 10.1016/j.clim.2013.08.004.
- Linda von Hoff, Eva Kärgel, Vedran Franke, Erik McShane, Kathrin W. Schulz-Beiss, Giannino Patone, Nikolai Schleussner, Marina Kolesnichenko, Norbert Hübner, Oliver Daumke, Matthias Selbach, Altuna Akalin, Stephan Mathas, and Claus Scheidereit. Autocrine LTA signaling drives NF- κ B and JAK-STAT activity and myeloid gene expression in Hodgkin lymphoma. *Blood*, 133(13): 1489–1494, March 2019. ISSN 0006-4971. doi: 10.1182/blood-2018-08-871293.
- Urmo Vösa, Annique Claringbould, Harm-Jan Westra, Marc Jan Bonder, Patrick Deelen, Biao Zeng, Holger Kirsten, Ashis Saha, Roman Kreuzhuber, Seyhan Yazar, Harm Brugge, Roy Oelen, Dylan H. de Vries, Monique GP van der Wijst, Silva Kasela, Natalia Pervjakova, Isabel Alves, Marie-Julie Favé, Mawussé Agbessi, Mark W. Christiansen, Rick Jansen, Ilkka Seppälä, Lin Tong, Alexander Teumer, Katharina Schramm, Gibran Hemani, Joost Verlouw, Hanieh Yaghootkar, Reyhan Sönmez, Andrew Brown, Viktorija Kukushkina, Anette Kalnapienkis, Sina Rüeger, Eleonora Porcu, Jaanika Kronberg, Johannes Kettunen, Bernett Lee, Futao Zhang, Ting Qi, Jose Alquicira Hernandez, Wibowo Arindrarto, Frank Beutner, Bios Consortium, i2QTL Consortium, Julia Dmitrieva, Mahmoud Elansary, Benjamin P. Fairfax, Michel Georges, Bastiaan T. Heijmans, Alex W. Hewitt, Mika Kähönen, Yungil Kim, Julian C. Knight, Peter Kovacs, Knut Krohn, Shuang Li, Markus Loeffler, Urko M. Marigorta, Hailang Mei, Yukihide Momozawa, Martina Müller-Nurasyid, Matthias Nauck, Michel G. Nivard, Brenda WJH Penninx, Jonathan K. Pritchard, Olli T. Raitakari, Olaf Rotzschke, Eline P. Slagboom, Coen DA Stehouwer, Michael Stumvoll, Patrick Sullivan, Peter AC 't Hoen, Joachim Thiery, Anke Tönjes, Jenny van Dongen, Maarten van Iterson, Jan H. Veldink, Uwe Völker, Robert Warmerdam, Cisca Wijmenga, Morris Swertz, Anand Andiappan, Grant W. Montgomery, Samuli Ripatti, Markus Perola, Zoltan Kutalik, Emmanouil Dermitzakis, Sven Bergmann, Timothy Frayling, Joyce van Meurs, Holger Prokisch, Habibul Ahsan, Brandon L. Pierce, Terho Lehtimäki, Dorret I. Boomsma, Bruce M. Psaty, Sina A. Gharib, Philip Awadalla, Lili Milani, and Willem H. Ouwehand. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet*, 53(9):1300, September 2021. doi: 10.1038/s41588-021-00913-z.

- Josefine Wadenpohl, Julia Seyfarth, Paul Hehenkamp, Maximilian Hoffmann, Sebastian Kummer, Christina Reinauer, Carsten Döing, Katharina Förtsch, Ertan Mayatepek, Thomas Meissner, and Marc Jacobsen. CD5-expressing CD8+ T-cell subsets differ between children with type 1 diabetes and controls. *Immunology & Cell Biology*, 99(10):1077–1084, 2021. ISSN 1440-1711. doi: 10.1111/imcb.12488.
- Melany J D Wagner, Shabbir Hussain, Mala Mehan, Joseph M Verdi, and Terry L Delovitch. A Defect in Lineage Fate Decision during Fetal Thymic Invariant NKT Cell Development May Regulate Susceptibility to Type 1 Diabetes1. *The Journal of Immunology*, 174(11):6764–6771, June 2005. ISSN 0022-1767. doi: 10.4049/jimmunol.174.11.6764.
- Michaela Waibel, John M. Wentworth, Michelle So, Jennifer J. Couper, Fergus J. Cameron, Richard J. MacIsaac, Gabby Atlas, Alexandra Gorelik, Sara Litwak, Laura Sanz-Villanueva, Prerak Trivedi, Simi Ahmed, Francis J. Martin, Madeleine E. Doyle, Jessica E. Harbison, Candice Hall, Balasubramanian Krishnamurthy, Peter G. Colman, Leonard C. Harrison, Helen E. Thomas, and Thomas W. H. Kay. Baricitinib and β -Cell Function in Patients with New-Onset Type 1 Diabetes. *New England Journal of Medicine*, 389(23):2140–2150, December 2023. ISSN 0028-4793. doi: 10.1056/NEJMoa2306691.
- Markus Walter, Areti Philotheou, François Bonnici, Anette-G. Ziegler, Roland Jimenez, and on behalf of the NBI-6024 Study Group. No Effect of the Altered Peptide Ligand NBI-6024 on β -Cell Residual Function and Insulin Needs in New-Onset Type 1 Diabetes. *Diabetes Care*, 32(11):2036–2040, August 2009. ISSN 0149-5992. doi: 10.2337/dc09-0449.
- Baihan Wang, Alfred Pozarickij, Mohsen Mazidi, Neil Wright, Pang Yao, Saredo Said, Andri Iona, Christiana Kartsonaki, Hannah Fry, Kuang Lin, Yiping Chen, Huaidong Du, Daniel Avery, Dan Schmidt-Valle, Canqing Yu, Dianjianyi Sun, Jun Lv, Michael Hill, Liming Li, Derrick A. Bennett, Rory Collins, Robin G. Walters, Robert Clarke, Iona Y. Millwood, and Zhengming Chen. Comparative studies of 2168 plasma proteins measured by two affinity-based platforms in 4000 Chinese adults. *Nat Commun*, 16(1):1869, February 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-56935-2.
- Chia-Jen Wang, Feng-Cheng Chou, Chi-Hong Chu, Jen-Chine Wu, Shih-Hua Lin, Deh-Ming Chang, and Huey-Kang Sytwu. Protective Role of Programmed

- Death 1 Ligand 1 (PD-L1) in Nonobese Diabetic Mice : The Paradox in Transgenic Models. *Diabetes*, 57(7):1861–1869, July 2008. ISSN 0012-1797. doi: 10.2337/db07-1260.
- Jian Wang, Taku Yoshida, Fumio Nakaki, Hiroshi Hiai, Taku Okazaki, and Tasuku Honjo. Establishment of NOD-Pdcd1-/- mice as an efficient animal model of type I diabetes. *Proc Natl Acad Sci U S A*, 102(33):11823–11828, August 2005. doi: 10.1073/pnas.0505497102.
- Shusen Wang, Yuanyuan Du, Boya Zhang, Gaofan Meng, Zewen Liu, Soon Yi Liew, Rui Liang, Zhengyuan Zhang, Xiangheng Cai, Shuangshuang Wu, Wei Gao, Dewei Zhuang, Jiaqi Zou, Hui Huang, Mingyang Wang, Xiaofeng Wang, Xuelian Wang, Ting Liang, Tengli Liu, Jiabin Gu, Na Liu, Yanling Wei, Xuejie Ding, Yue Pu, Yixiang Zhan, Yu Luo, Peng Sun, Shuangshuang Xie, Jiuxia Yang, Yiqi Weng, Chunlei Zhou, Zhenglu Wang, Shuang Wang, Hongkui Deng, and Zhongyang Shen. Transplantation of chemically induced pluripotent stem-cell-derived islets under abdominal anterior rectus sheath in a type 1 diabetes patient. *Cell*, 187(22):6152–6164.e18, October 2024. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2024.09.004.
- Xinchen Wang and David B. Goldstein. Enhancer Domains Predict Gene Pathogenicity and Inform Gene Discovery in Complex Disease. *The American Journal of Human Genetics*, 106(2):215–233, February 2020. ISSN 0002-9297. doi: 10.1016/j.ajhg.2020.01.012.
- Jeremy T. Warshauer, Jeffrey A. Bluestone, and Mark S. Anderson. New Frontiers in the Treatment of Type 1 Diabetes. *Cell Metabolism*, 31(1):46–61, January 2020. ISSN 1550-4131. doi: 10.1016/j.cmet.2019.11.017.
- Taylor K. Watson, Aaron B.I. Rosen, Travis Drow, Jacob A. Medjo, Matthew A. MacQuivey, Yan Ge, H. Denny Liggitt, Dane A. Grosvenor, Kimberly A. Dill-McFarland, Matthew C. Altman, Patrick J. Concannon, Jane H. Buckner, David J. Rawlings, and Eric J. Allenspach. Reduced Function of the Adaptor SH2B3 Promotes T1D via Altered Cytokine-Regulated, T-Cell-Intrinsic Immune Tolerance. *Diabetes*, 74(6):943–955, March 2025. ISSN 0012-1797. doi: 10.2337/db24-0655.
- Bobbie-Jo M. Webb-Robertson, Ernesto S. Nakayasu, Fran Dong, Kathy C. Waugh, Javier E. Flores, Lisa M. Bramer, Athena A. Schepmoes, Yuqian Gao, Thomas L. Fillmore, Suna Onengut-Gumuscu, Ashley Frazer-Abel, Stephen S. Rich, V. Michael Holers, Thomas O. Metz, and Marian J. Rewers. Decrease

in multiple complement proteins associated with development of islet autoimmunity and type 1 diabetes. *iScience*, 27(2):108769, December 2023. ISSN 2589-0042. doi: 10.1016/j.isci.2023.108769.

Yuxia Wei, Shengxin Liu, Tomas Andersson, Maria Feychting, Ralf Kuja-Halkola, and Sofia Carlsson. Familial aggregation and heritability of childhood-onset and adult-onset type 1 diabetes: A Swedish register-based cohort study. *The Lancet Diabetes & Endocrinology*, 12(5):320–329, May 2024. ISSN 22138587. doi: 10.1016/S2213-8587(24)00068-8.

Craig Sinclair Weston, Bernhard O. Boehm, and Paolo Pozzilli. Type 1 diabetes: A new vision of the disease based on endotypes. *Diabetes/Metabolism Research and Reviews*, 40(2):e3770, 2024. ISSN 1520-7560. doi: 10.1002/dmrr.3770.

Diane K Wherrett, Brian Bundy, Dorothy J Becker, Linda A DiMeglio, Stephen E Gitelman, Robin Goland, Peter A Gottlieb, Carla J Greenbaum, Kevan C Herold, Jennifer B Marks, Roshanak Monzavi, Antoinette Moran, Tihamer Orban, Jerry P Palmer, Philip Raskin, Henry Rodriguez, Desmond Schatz, Darrell M Wilson, Jeffrey P Krischer, and Jay S Skyler. Antigen-based therapy with glutamic acid decarboxylase (GAD) vaccine in patients with recent-onset type 1 diabetes: A randomised double-blind trial. *The Lancet*, 378(9788):319–327, July 2011. ISSN 0140-6736. doi: 10.1016/S0140-6736(11)60895-7.

Della White, Alexandra Cote-Martin, Marina Bleck, Nicole Garaffa, Abdulsalam Shaaban, Helen Wu, Dongmei Liu, David Young, Justin Scheer, Ivo C. Lorenz, Andrew Nixon, Jay S. Fine, Fergus R. Byrne, M.Lamine Mbow, and Miguel E. Moreno-Garcia. Programmed Cell Death-1 (PD-1) anchoring to the GPI-linked co-receptor CD48 reveals a novel mechanism to modulate PD-1-dependent inhibition of human T cells. *Molecular Immunology*, 156:31–38, April 2023. ISSN 01615890. doi: 10.1016/j.molimm.2023.02.007.

Jordan J. Wright, Adel Eskaros, Annika Windon, Rita Bottino, Regina Jenkins, Amber M. Bradley, Radhika Aramandla, Sharon Philips, Hakmook Kang, Diane C. Saunders, Marcela Brissova, and Alvin C. Powers. Exocrine Pancreas in Type 1 and Type 2 Diabetes: Different Patterns of Fibrosis, Metaplasia, Angiopathy, and Adiposity. *Diabetes*, 73(7):1140–1152, October 2023. ISSN 0012-1797. doi: 10.2337/db23-0009.

- Writing Committee for the Type 1 Diabetes TrialNet Oral Insulin Study Group. Effect of Oral Insulin on Prevention of Diabetes in Relatives of Patients With Type 1 Diabetes: A Randomized Clinical Trial. *JAMA*, 318(19):1891–1902, November 2017. ISSN 0098-7484. doi: 10.1001/jama.2017.17070.
- Louis-Pascal Xhonneux, Oliver Knight, Åke Lernmark, Ezio Bonifacio, William A. Hagopian, Marian J. Rewers, Jin-Xiong She, Jorma Toppari, Hemang Parikh, Kenneth G.C. Smith, Anette-G. Ziegler, Beena Akolkar, Jeffrey P. Krischer, and Eoin F. McKinney. Transcriptional networks in at-risk individuals identify signatures of type 1 diabetes progression. *Sci Transl Med*, 13(587):eabd5666, March 2021. ISSN 1946-6234. doi: 10.1126/scitranslmed.abd5666.
- Zhiguo Xie, Christopher Chang, and Zhiguang Zhou. Molecular Mechanisms in Autoimmune Type 1 Diabetes: A Critical Review. *Clinic Rev Allerg Immunol*, 47(2):174–192, October 2014. ISSN 1559-0267. doi: 10.1007/s12016-014-8422-2.
- Guanlan Xu, Junqin Chen, Gu Jing, and Anath Shalev. Preventing β -Cell Loss and Diabetes With Calcium Channel Blockers. *Diabetes*, 61(4):848–856, March 2012. ISSN 0012-1797. doi: 10.2337/db11-0955.
- Guifeng Xu, Buyun Liu, Yangbo Sun, Yang Du, Linda G. Snetselaar, Frank B. Hu, and Wei Bao. Prevalence of diagnosed type 1 and type 2 diabetes among US adults in 2016 and 2017: Population based study. *BMJ*, 362:k1497, September 2018. ISSN 0959-8138, 1756-1833. doi: 10.1136/bmj.k1497.
- Masahito Yamagata. Structure and Functions of Sidekicks. *Frontiers in Molecular Neuroscience*, 13, August 2020. ISSN 1662-5099. doi: 10.3389/fnmol.2020.00139.
- Jing Yan, Di Chen, Zi Ye, Xuqiang Zhu, Xueyuan Li, Henan Jiao, Mengjiao Duan, Chaoli Zhang, Jingliang Cheng, Lixia Xu, Hongjiang Li, and Dongming Yan. Molecular mechanisms and therapeutic significance of Tryptophan Metabolism and signaling in cancer. *Mol Cancer*, 23(1):241, October 2024. ISSN 1476-4598. doi: 10.1186/s12943-024-02164-y.
- Yiming Yang, Andrew J. Sanders, Q. Ping Dou, David G. Jiang, Amber Xinyu Li, and Wen G. Jiang. The Clinical and Theranostic Values of Activated Leukocyte Cell Adhesion Molecule (ALCAM)/CD166 in Human Solid Cancers. *Cancers (Basel)*, 13(20):5187, October 2021. ISSN 2072-6694. doi: 10.3390/cancers13205187.

- Zhaoxin Yang, Zhirang Zhang, Liyan Li, Zhangyan Jing, Yumeng Ma, Tianyu Lan, Yuan Li, Zhongda Lin, Wenli Fang, Jinjie Zhang, Jinling Zhang, Xin Liang, Benqing Wu, Yi Zheng, and Xudong Zhang. Bioengineered Artificial Extracellular Vesicles Presenting PD-L1 and Gal-9 Ameliorate New-Onset Type 1 Diabetes. *Diabetes*, 73(8):1325–1335, May 2024. ISSN 0012-1797. doi: 10.2337/db23-0987.
- Esma S. Yolcu, Haval Shirwan, and Nadir Askenasy. Fas/Fas-Ligand Interaction As a Mechanism of Immune Homeostasis and β -Cell Cytotoxicity: Enforcement Rather Than Neutralization for Treatment of Type 1 Diabetes. *Front Immunol*, 8, March 2017. ISSN 1664-3224. doi: 10.3389/fimmu.2017.00342.
- Li Yuan, Yuxia Wang, Xuxia Shen, Fujun Ma, Jun Wang, and Fang Yan. Soluble form of immune checkpoints in autoimmune diseases. *Journal of Autoimmunity*, 147:103278, July 2024. ISSN 0896-8411. doi: 10.1016/j.jaut.2024.103278.
- Chenran Yue, Sheng Gao, Shuting Li, Zhouhang Xing, Hengrong Qian, Ying Hu, Wenqian Wang, and Chunyan Hua. TIGIT as a Promising Therapeutic Target in Autoimmune Diseases. *Front Immunol*, 13:911919, June 2022. ISSN 1664-3224. doi: 10.3389/fimmu.2022.911919.
- Barbara Zdrazil, Eloy Felix, Fiona Hunter, Emma J Manners, James Blackshaw, Sybilla Corbett, Marleen de Veij, Harris Ioannidis, David Mendez Lopez, Juan F Mosquera, Maria Paula Magarinos, Nicolas Bosc, Ricardo Arcila, Tevfik Kizilören, Anna Gaulton, A Patrícia Bento, Melissa F Adasme, Peter Monecke, Gregory A Landrum, and Andrew R Leach. The ChEMBL Database in 2023: A drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res*, 52(D1):D1180–D1192, January 2024. ISSN 0305-1048. doi: 10.1093/nar/gkad1004.
- Lijie Zhai, Erik Ladomersky, Alicia Lenzen, Brenda Nguyen, Ricky Patel, Kristen L. Lauing, Meijing Wu, and Derek A. Wainwright. IDO1 in cancer: A Gemini of immune checkpoints. *Cell Mol Immunol*, 15(5):447–457, May 2018. ISSN 2042-0226. doi: 10.1038/cmi.2017.143.
- Qianxia Zhang, Maria Chikina, Andrea L. Szymczak-Workman, William Horne, Jay K. Kolls, Kate M. Vignali, Daniel Normolle, Maria Bettini, Creg J. Workman, and Dario A. A. Vignali. LAG3 limits regulatory T cell proliferation and function in autoimmune diabetes. *Sci Immunol*, 2(9):eaah4569, March 2017. doi: 10.1126/sciimmunol.aah4569.

- Jie Zheng, Valeriia Haberland, Denis Baird, Venexia Walker, Philip C. Haycock, Mark R. Hurle, Alex Gutteridge, Pau Erola, Yi Liu, Shan Luo, Jamie Robinson, Tom G. Richardson, James R. Staley, Benjamin Elsworth, Stephen Burgess, Benjamin B. Sun, John Danesh, Heiko Runz, Joseph C. Maranville, Hannah M. Martin, James Yarmolinsky, Charles Laurin, Michael V. Holmes, Jimmy Z. Liu, Karol Estrada, Rita Santos, Linda McCarthy, Dawn Waterworth, Matthew R. Nelson, George Davey Smith, Adam S. Butterworth, Gibran Hemani, Robert A. Scott, and Tom R. Gaunt. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet*, 52 (10):1122–1131, October 2020. ISSN 1546-1718. doi: 10.1038/s41588-020-0682-6.
- Qian Zhong, Xina Xiao, Yijie Qiu, Zhiqiang Xu, Chunyu Chen, Baochen Chong, Xinjun Zhao, Shan Hai, Shuangqing Li, Zhenmei An, and Lunzhi Dai. Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications. *MedComm (2020)*, 4(3):e261, May 2023. ISSN 2688-2663. doi: 10.1002/mco2.261.
- Chenjia Zhou, Ying Gao, Peilun Ding, Tao Wu, and Guang Ji. The role of CXCL family members in different diseases. *Cell Death Discovery*, 9(1):1–12, July 2023. ISSN 2058-7716. doi: 10.1038/s41420-023-01524-9.
- Anette-Gabriele Ziegler, Peter Achenbach, Reinhard Berner, Kristina Casteels, Thomas Danne, Melanie Gündert, Joerg Hasford, Verena Sophia Hoffmann, Olga Kordonouri, Karin Lange, Helena Elding Larsson, Markus Lundgren, Matthew D. Snape, Agnieszka Szypowska, John A. Todd, and Ezio Bonifacio. Oral insulin therapy for primary prevention of type 1 diabetes in infants with high genetic risk: The GPPAD-POInT (global platform for the prevention of autoimmune diabetes primary oral insulin trial) study protocol. *BMJ Open*, 9 (6):e028578, June 2019. ISSN 2044-6055, 2044-6055. doi: 10.1136/bmjopen-2018-028578.
- Maciej Zieliński, Magdalena Żalińska, Dorota Iwaszkiewicz-Grześ, Mateusz Gliwiński, Matylda Hennig, Anna Jaźwińska-Curyłło, Halla Kamińska, Justyna Sakowska, Anna Wołoszyn-Durkiewicz, Radosław Owczuk, Wojciech Młynarski, Przemysław Jarosz-Chobot, Artur Bossowski, Agnieszka Szadkowska, Janusz Siebert, Małgorzata Myśliwiec, Natalia Marek-Trzonkowska, and Piotr Trzonkowski. Combined therapy with CD4+CD25highCD127- T regulatory cells and anti-CD20 antibody in recent-onset type 1 diabetes is

superior to monotherapy: Randomized phase I/II trial. *Diabetes, Obesity and Metabolism*, 24(8):1534–1543, August 2022. ISSN 1462-8902. doi: 10.1111/dom.14723.