Ecole doctorale Biologie-Sante de Lille (EDBSL)

Academic year: 2024-2025

A multi-omics approach to identify key genes in the endocrine and exocrine pancreas and their role in T2D

Presented by: Lucas Maurin

Under the supervision of **Dr Amna Khamis** and **Professor Philippe Froguel**

Defended on the 11th of December 2024 in front of the following jury:

Rapporteur: **Professor Michael Weedon**, University of Exeter, UK

Rapporteur: **Dr Lorenzo Pasquali**, Pompeu Fabra University, Spain

Thesis supervisor: **Dr Amna Khamis**, University of Lille, France

Thesis supervisor: **Professor Philippe Froguel**, University of Lille, France

Ecole doctorale Biologie-Sante de Lille (EDBSL)

Année universitaire : 2024-2025

Une approche multi-omiques pour identifier les gènes clés du pancréas endocrinien et exocrine et leur rôle dans le développement du DT2

Presenté par : Lucas Maurin

Sous la direction du **Dr Amna Khamis** et **Professeur Philippe Froguel**

Soutenu le 11 décembre 2024 devant le jury suivant :

Rapporteur : **Professeur Michael Weedon**, University of Exeter, RU

Rapporteur : **Dr Lorenzo Pasquali**, Pompeu Fabra University, Espagne

Directrice de thèse : **Dr Amna Khamis**, Université de Lille, France

Directeur de thèse : **Professeur Philippe Froguel**, Université de Lille, France

Table of contents

1 Summary of thesis in french (Résumé de la thèse en français)

Le diabète de type 2 (DT2) est une maladie multifactorielle caractérisée par une hyperglycémie chronique, et causée par des facteurs génétiques et environnementaux, tels que le vieillissement. Alors que les études d'association pangénomique (GWAS) ont commencé à identifier les causes génétiques du DT2, les études d'association épigénomique (EWAS) ont rencontré un succès limité dans la caractérisation de l'impact de l'environnement en raison de la taille très réduite des échantillons étudiés et du manque d'études fonctionnelles. De plus, l'ampleur de l'interaction entre la variation génétique et épigénétique reste mal comprise. L'objectif de cette thèse était de contribuer à notre compréhension de la façon dont les facteurs environnementaux influencent la pathogenèse du DT2 et sa progression vers des complications associées, notamment l'adénocarcinome canalaire pancréatique (PDAC).

Dans le premier projet, nous avons étudié l'interaction entre les modifications épigénétiques associées au DT2 et l'âge, et la variation génétique dans les îlots pancréatiques de 124 individus, dont 16 atteints de DT2. Nous avons développé une approche intégrative combinant la méthylation de l'ADN, l'expression génique et le génotypage pour identifier des associations en triade, en examinant si les variations génétiques et épigénétiques s'influencent mutuellement. Nous avons identifié 301 et 743 CpG associés à l'âge et au DT2, qui influencent l'expression des gènes voisins. Parmi ceux-ci, moins de 10 % étaient influencés par des variants génétiques, suggérant que les modifications épigénétiques induites par l'environnement sont indépendantes de la variation génétique. Notamment, seuls trois gènes, *SIX3*, *ST6GAL1* et *TIPIN*, se colocalisaient avec des variants de risque du GWAS du DT2 et étaient également sous régulation épigénétique. La caractérisation des gènes régulés épigénétiquement a mis en évidence des candidats clés du DT2, notamment *OPRD1* et *MEG3*. Finalement, nous montrons que l'ajout de scores de risque épigénétique aux scores de risque polygénique a amélioré la prédiction du risque de DT2. Nos

résultats suggèrent que la plupart des gènes sont régulés soit par des facteurs génétiques, soit par des facteurs épigénétiques, mais rarement par les deux.

Dans le deuxième projet, nous avons exploré l'influence épigénétique du DT2 sur le pancréas exocrine afin de comprendre pourquoi les individus atteints de DT2 sont plus à risque de développer le PDAC, l'un des cancers les plus mortels. Nous avons réalisé une EWAS pour le DT2 (25 individus atteints de DT2 et 116 non-diabétiques) et identifié une seule hyperméthylation dans cg15549216, située dans le gène *Pancreatic Lipase Related Protein 1* (*PNLIPRP1*), qui était corrélée à une diminution de l'expression du gène. L'inhibition de *Pnliprp1* dans la lignée cellulaire acinaire de rat AR42J a augmenté les niveaux de cholestérol, réduit la prolifération et induit une métaplasie acino-canalaire (ADM), caractéristiques des premiers stades du PDAC. Il est à noter que cet effet a été inversé par un traitement avec des statines, mettant en évidence le potentiel translationnel de ces résultats. De plus, une analyse de variants rares à partir de la UK Biobank a lié *PNLIPRP1* au cholestérol LDL, confirmant les résultats fonctionnels. Nous proposons un modèle où les mécanismes épigénétiques et génétiques agissent indépendamment mais synergisent pour favoriser les lésions du pancréas et la progression de la maladie.

Cette thèse souligne l'importance d'étudier la méthylation de l’ADN pour identifier les facteurs environnementaux qui contribuent à la maladie. Nos résultats révèlent que ces altérations épigénétiques sont en grande partie indépendantes des facteurs génétiques, soulignant leur rôle complémentaire dans la pathogenèse du DT2. De plus, *PNLIPRP1* illustre comment les études épigénomiques peuvent identifier de nouveaux biomarqueurs à pertinence translationnelle, offrant de nouvelles perspectives sur le DT2.

2 Résumé de thèse vulgarisé pour le grand public en français

Le diabète de type 2 (DT2) est une maladie caractérisée par un excès de sucre sanguin, causée par des facteurs génétiques et environnementaux comme le vieillissement, une alimentation riche, l'obésité et la sédentarité. Cette thèse vise à comprendre comment l'environnement contribue au DT2 et ses complications. Dans le premier projet, nous avons étudié comment le vieillissement et le diabète affectent l'ADN des cellules productrices d'insuline. Nous avons découvert que le diabète et l'âge entraînent des modifications de l'ADN, altérant le fonctionnement des gènes régulant le sucre sanguin. Ensuite, nous avons exploré l'impact du diabète sur le pancréas exocrine. Le gène PNLIPRP1 est sous-exprimé chez les diabétiques. En laboratoire, sa modification induit des changements similaires aux débuts du cancer du pancréas, mais ces effets sont inversés avec des médicaments comme les statines. Comprendre ces changements pourrait aider à prévenir ou traiter le DT2 et ses complications.

3 Thesis summary

Type 2 diabetes (T2D) is a multifactorial, complex disease characterised by chronic elevated blood glucose, and caused by genetic and environmental factors, such as ageing. While genome-wide association studies (GWAS) have successfully identified the genetic causes of T2D, epigenome-wide association studies (EWAS) have had limited success in capturing the environmental impact due to the tissue-specificity of epigenetic changes, very small sample sizes, and the lack of functional studies. Furthermore, the extent of the interaction between genetic and epigenetic variation remains poorly understood. The objective of this thesis was to contribute to our understanding of how environmental factors contribute to T2D pathogenesis, and its progression towards related complications, notably pancreatic ductal adenocarcinoma (PDAC).

In the first project, we investigated the interplay between age and T2D-associated epigenetic changes and genetic variation in pancreatic islets of 124 individuals, of which 16 had T2D. We developed a novel integrative approach combining DNA methylation, gene expression, and genotyping to identify triad associations, examining whether genetic and epigenetic influence each other. We identified 301 and 743 CpGs associated with age and T2D, respectively, which impacted nearby gene expression (within a 2 Mb window). Of these, less than 10 % were influenced by nearby genetic variants, suggesting that environmentally-driven epigenetic changes operate largely independently of genetic variation. Notably, only three genes, *SIX3*, *ST6GAL1*, and *TIPIN*, were found to co-localise with T2D GWAS risk variants, and were also under epigenetic regulation. Characterisation of the epigenetically-regulated genes highlighted key T2D candidates, including *OPRD1* and *MEG3*. Importantly, adding methylation risk scores (MRS) to polygenic risk scores (PGS) improved T2D risk prediction, underscoring the additive value of epigenetic studies. Our findings suggest that most genes are regulated either by genetic or epigenetic factors, but rarely both.

In the second project, we explored the epigenetic influence of T2D in the exocrine pancreas, to explore why T2D individuals are at a higher risk of developing pancreatic disease, notably PDAC, one of the deadliest cancers. We performed an EWAS for T2D (25 T2D individuals and 116 non-diabetic) and identified a single hypermethylation in cg15549216, located in the Pancreatic Lipase Related Protein 1 (*PNLIPRP1*) gene, which was corelated with a decreased expression of the gene. Knockdown of *Pnliprp1* in the rat acinar cell line AR42J increased cholesterol levels, reduced proliferation, and induced acinar-to-ductal metaplasia (ADM), hallmarks of the early stages of PDAC. Notably, this effect was reversed by treatment of statin, highlighting the translational potential of these findings. Additionally, a rare variant analysis using the UKBiobank linked *PNLIPRP1* to LDL-cholesterol, confirming the functional results. We propose a model where epigenetic and genetic mechanisms act independently but synergise to promote pancreas injury and disease progression.

This thesis underscores the importance of studying DNA methylation as an unbiased approach for identifying environmental factors that contribute to disease. Our findings reveal that these epigenetic alterations are largely independent of genetic factors, underscoring their complementary role T2D pathogenesis. Additionally, *PNLIPRP1* serves as an example of how epigenomic studies can indeed identify novel biomarkers with a translational relevance, offering new insights into disease mechanisms and progression.

4 Thesis summary for the general public

Type 2 diabetes (T2D) is a disease characterised by high blood sugar levels, caused by genetic and environmental factors like aging, high-calorie diets, obesity, and lack of physical activity. This thesis explores how the environment contributes to diabetes and its complications. The first project investigated how ageing and diabetes interact with DNA in insulin-producing pancreatic cells. We found that both diabetes and ageing lead to specific DNA modifications that alter blood sugar regulating genes. In the second project, we explored how T2D impacts other parts of the pancreas. We identified a gene (*PNLIPRP1*) modified in diabetics. Lab experiments showed that modifying this gene caused changes similar to early pancreatic cancer, but these effects can be reversed with drugs like statins. This research shows that environmental factors can alter DNA in ways that influence diseases. Understanding these changes may lead to new ways to prevent or treat T2D and its complications.

5 Acknowledgements

Without any exaggeration, these past three years and some have transformed me at every level beyond what can be attributed to time alone. While I don't have data to support this, I believe it is largely due to the incredible people I was surrounded by.

First, I'd like to thank Professor Michael Weedon and Dr Lorenzo Pasquali for agreeing to share their knowledge and time by assessing the work I have performed during my PhD.

I would also like to thank Dr Nicolas Venteclef and Dr Mark Ibberson, who took the time to read my end of year reports and followed my project these past three years. Their advice not only guided me in identifying areas for improvement but also inspired me to recognise what could be achieved.

Dr Amna Khamis, I hope that if I ever happen to supervise a student, I will show the same kindness, patience and understanding you did. Thank you for being so generous with your time, which led to many important and interesting discussions about science itself, but also everything else around science, despite myself not being the ideal student. Thank you also for being a model of determination and resolve, which inspired me to push through the various challenges which inevitably rear themselves through the years. Finally, thank you for taking a chance on me – I arrived here with no vision or plan, but I leave with purpose, which as it turns out is quite important as you go through life. Most students complain about their supervisors (this is in fact scientifically proven), so I guess I got really lucky.

Un grand merci au Professeur Philippe Froguel, qui m'a offert l'opportunité de rejoindre son équipe suite à un stage au laboratoire. Nos réunions et discussions ont toujours été une source d'inspiration, apportant de nouvelles idées et perspectives scientifiques qui ont enrichi et affiné mes réflexions. Ces échanges ont considérablement amélioré les projets sur lesquels j'ai travaillé, ainsi que ceux à venir. Merci aussi au Dr Amélie Bonnefond, qui m'a prodigué de nombreux conseils et suggestions qui m'ont permis d'améliorer mon travail, en particulier pour l'article sur *PNLIPRP1*, qui ont largement contribué à l'amélioration du papier.

Un énorme merci à toute l'équipe du laboratoire, qui m'a accueilli chaleureusement. Merci particulièrement à Justine pour son courage et sa bonne humeur, malgré le fait qu'elle devait partager un bureau avec Maxine, Vincent et moi-même, mais aussi pour sa précieuse contribution au reviewing de l'article sur *PNLIPRP1*, qui a été

nettement amélioré grâce à son travail. Merci également à Sarah, qui m'a partagé sa sagesse et ses conseils de doctorante, ainsi qu'à Hélène et Mélanie, qui m'ont énormément aidée pour toutes les démarches administratives, surtout au début de ma thèse pendant la période du Covid. Un grand merci à Constance et Nicolas pour leur bonne humeur et leur aide avec les réservations de trains en première classe et d'hôtels cinq étoiles lors des déplacements en congrès. Merci à Vincent Massy pour son assistance quand mon ordinateur refusait de fonctionner. Je remercie également toute l'équipe Biostat, passée et présente, qui a contribué à mes projets et m'a permis de me former en tant que biostatisticien. Un grand merci à Audrey pour son aide (et sa patience) lors des manipulations d'immunofluorescence, et à Mehdi, la force tranquille, qui savait comment réagir face à toutes les situations imprévues pendant la thèse.

None of this work would have been possible without the invaluable contributions of collaborators from across Europe, who provided the samples that made our data collection possible. I extend my heartfelt thanks to Professor Pierro Marchetti, Lorella Marselli, and their entire team for the islet and exocrine tissue samples they generously supplied. Special thanks also to Dr. Mark Ibberson and Miriam Cnop for providing the islet tissue. Finally, I would like to express my gratitude to Dr. Julie Kerr-Conte and François Pattou for the whole pancreas tissue slides they provided.

Un énorme merci à Vincent, avec qui j'ai partagé la majorité de ma thèse, aussi bien au laboratoire qu'en dehors. Si nos vies étaient des chapitres, tu serais certainement l'un des personnages principaux de celui-ci, et avec un peu de chance, des chapitres à venir aussi. Tu m'as apporté une toute nouvelle perspective sur la vie et, de manière générale, tu as grandement contribué à la personne que je suis aujourd'hui.

Merci aussi à Alexandre, qui, avec Vincent, complète le trio de Lille. Tout comme Vincent, tu représentes une partie importante de ce chapitre, et tu as contribué à faire de moi la personne que je suis aujourd'hui. Je sais déjà que nos chemins continueront à se croiser.

Un immense merci à Amélie, qui a été un véritable pilier tout au long de ma thèse. Ta gentillesse et ton incroyable capacité à prendre soin des autres font de toi la personne la plus douce et bienveillante que j'aie jamais rencontrée. Tu as marqué ma vie d'une façon qui va bien au-delà des mots.

Merci aussi à Nawel, avec qui j’ai partagé de nombreux moment joyeux et rempli de rire. J’ai rarement rencontré une personne aussi bienveillante et aussi couramment de bonne humeur.

Merci à Maxime, qui m’a toujours aidé, que ce soit pour des questions administratives ou des manipulations, et avec qui nous avons formé une équipe de choc pour l’évacuation des déchets du laboratoire.
Steven et Louise pour leur constante (ou presque) bonne humeur et gentillesse qui ont nettement améliorer mon temps sur Lille.

Thanks also to Matt, Adem, Rupert, Jez, Rayan, and Andy with whom I shared many moments, unfortunately online because of the distance, that made this whole experience lighter and more enjoyable.

Finalement, à mes parents, ma sœur et mon frère, ainsi qu'au reste de la famille, qui m'ont soutenu (et supporté) inconditionnellement pendant la thèse, ainsi que toutes les années auparavant.

6 Abbreviations

ADA	American Diabetes Association
ADM	Acinar-to-Ductal Metaplasia
AI	Artificial Intelligence
AUC	Area Under the Curve
BMI	Body Mass Index
BP	Blood Pressure
CpG	Cytosine Preceding a Guanine in a DNA sequence
DMR	Differentially Methylated Regions
DNMT	DNA Methyltransferases
eQTL	Expression Quantitative Trait Loci
ER	Endoplasmic reticulum
EWAS	Epigenome-Wide Association Studies
FDR	False Discovery Rate
G6P	Glucose-6-Phosphate
GSIS	Glucose-Stimulated Insulin Secretion
GTEx	a large eQTL database
GWAS	Genome-Wide Association Study
HDL	High density Lipoprotein
IAPP	Islet Amyloid Polypeptide
INSR	Insulin Receptor
IQR	Interquartile Range
IR	Insulin Resistance
ISI	Insulin Secretion Index
IVW	Inverse Variance Weighted
KD	Knockdown
LDL-C	Low Density Lipoprotein
LINE1	Long Interspersed Nuclear Element-1
MAF	Minor Allele Frequency
MARD	Mild Age-Related Diabetes
MEGA	Methylation-Expression adjusted for Genotype Analysis
ML	Machine Learning
MOD	Mild Obesity-Related Diabetes
MODY	Maturity onset diabetes of the young
mQTL	Methylation quantitative trait loci
MR	Mendelian Randomisation
MRS	Methylation risk scores
PC	Principal Component
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PDAC	Pancreatic Ductal Adenocarcinoma
PGS	Polygenic Risk Scores
PP	Pancreatic Polypeptide
QC	Quality Control
RE	Repetitive Elements
RF	Risk Factors

ROC	Receiver Operating Characteristic
RR	Relative Risk
RT	Room temperature
SCDNAM	Single cell DNA methylation
SIDD	Severe Insulin Deficient Diabetes
SIRD	Severe Insulin Resistant Diabetes
SNPs	Single Nucleotide Polymorpisms
T1D	Type 1 Diabetes
T2D	Type 2 Diabetes
TCA	Tricarboxylic Acid
TE	Transposable Elements
TFs	Transcription Factors
TPM	Transcript Per Million
WGBS	Whole-Genome Bisulfite Sequencing
WHR	Waist-Hip Ratio

7 Introduction

7.1 Diabetes

Diabetes refers to a group of metabolic diseases characterised by chronic elevated blood glucose levels, or hyperglycaemia (Karamanou et al. 2016), as a result of insufficient insulin production by insulin-secreting pancreatic β-cells (Abel et al. 2024). In 2021, an estimated 529 million individuals were afflicted with diabetes (Home et al., n.d.), and projections estimate 642 million and 1.31 billion cases in 2035 and 2050, respectively (Abel et al. 2024; Home et al., IDF diabetes atlas). Diabetes poses a significant burden, leading to reduced quality of life and a shorter life expectancy for those affected (Ahmad et al. 2022). Economically, the impact is substantial; in France, healthcare expenditure for diabetes has been sharply increasing, from 7.69 billion euros in 2015, to 8.59 billion, or 5.1% of total healthcare expenditures in 2019 (Rachas et al. 2022). Similar trends are observed worldwide (**Figure 1**). This growing challenge is further amplified by the increasing prevalence of an aging population, which contributes to rising healthcare demands. As a result, diabetes and related conditions represent a major healthcare challenge of the 21st century (United Nations, World Population Prospects Report, 2015).

Diabetes is a heterogeneous disease, which is broadly classified into several subtypes: type 1 diabetes (T1D), monogenic diabetes, and type 2 diabetes (T2D).T1D is characterised by an autoimmune-mediated destruction of pancreatic β-cells, leading to an absolute deficiency in insulin production (Saberzadeh-Ardestani et al. 2018). Additionally, there are over 40 currently identified rare monogenic forms of diabetes, including maturity-onset diabetes of the young (MODY) and neonatal diabetes, which result from highly penetrant heterozygous mutations in genes crucial for β-cell function, leading to impaired insulin secretion (Skoczek, Dulak, and Kachamakova-Trojanowska 2021; Bonnefond et al. 2023). Finally, T2D, the most prevalent type of diabetes, which is characterised by progressive loss of insulin secretion by pancreatic β-cells, combined with insulin resistance (IR) in peripheral tissues (Abel et al. 2024).

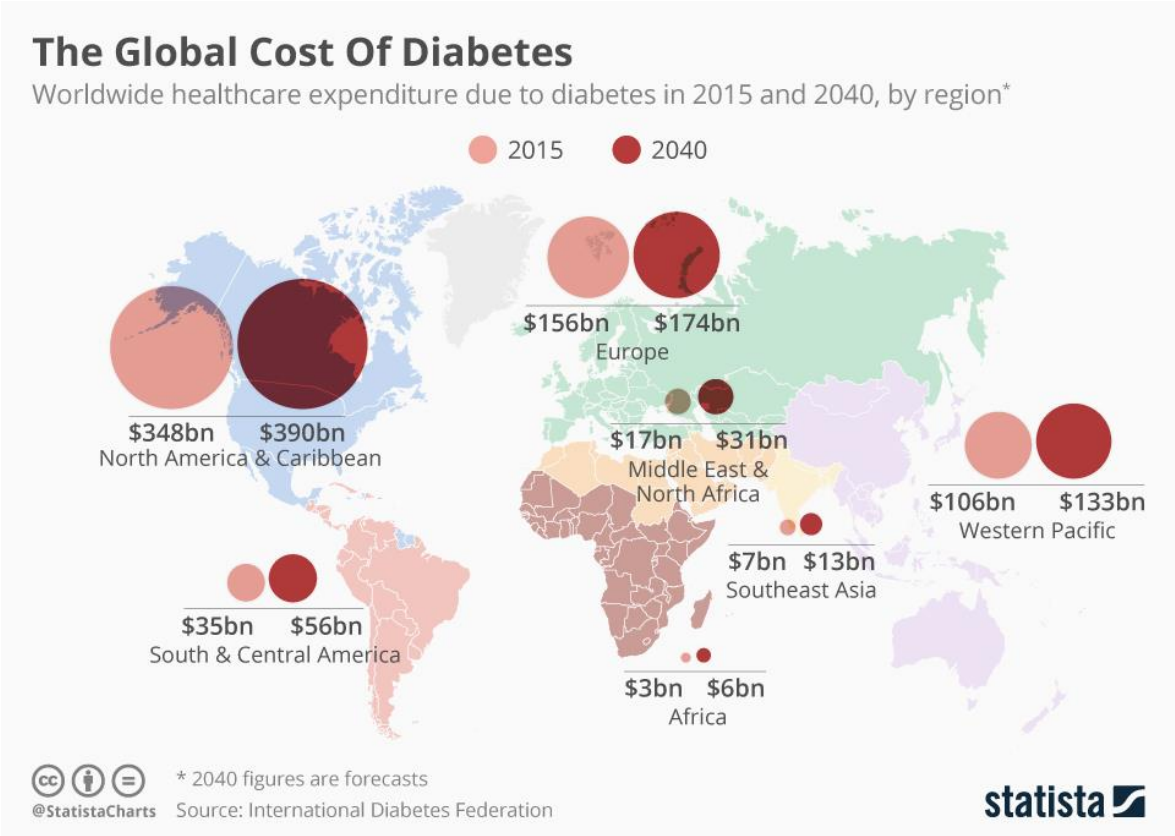


Figure 1: Estimated worldwide healthcare spending on diabetes in 2015 and 2040. Light red indicates 2015 and dark red indicates 2040. Size of the circles represent nominal healthcare spending. Adapated from ("Infographic: The Global Cost Of Diabetes" 2016).

7.1.1 Type 2 diabetes

T2D currently accounts for an estimated 96% of cases (Abel et al. 2024), and is the most rapidly increasing type of diabetes in terms of prevalence (**Figure 2**) (Ong et al. 2023). T2D is a multifactorial disease driven by both environmental and genetic factors. Key environmental contributors include ageing, obesity, physical inactivity, and sedentary lifestyles—all of which have played a significant role in the rising global incidence of T2D (Chatterjee, Khunti, and Davies 2017). Moreover, emerging evidence suggests that mental health and psychological well-being are also crucial risk factors, with chronic stress, depression, and anxiety linked to the development and progression of the disease (Kelly and Ismail 2015).

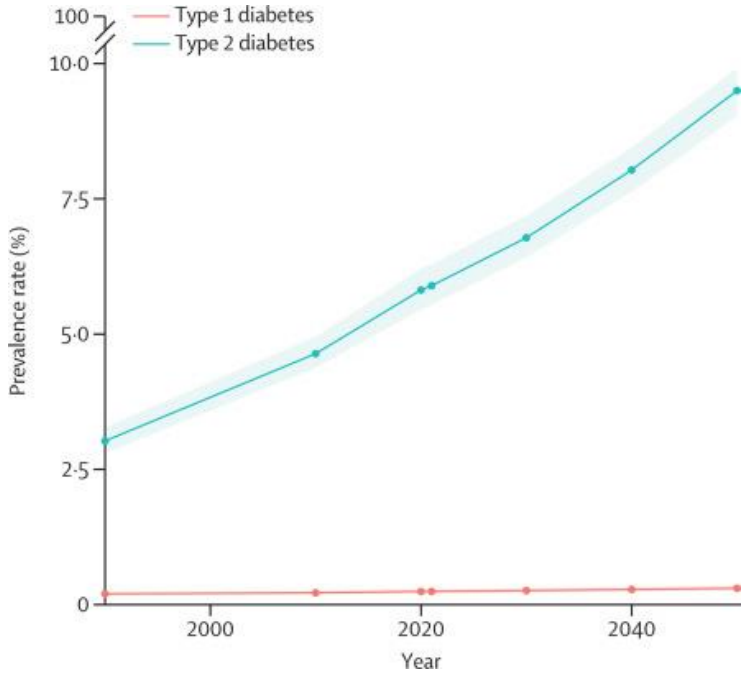


Figure 2: Type 1 and type 2 diabetes prevalence up to 2050. Type 1 diabetes incidence has remained largely stable through the years whilst type 2 diabetes incidence is rapidly increasing. Adapted from (Ong et al. 2023).

7.1.1.1 T2D pathophysiology

T2D is heterogenous and results from varying degrees of deficient insulin secretion and insulin sensitivity (Galicia-Garcia et al. 2020). The most prevalent model for the development of T2D involves a gradual increase in IR, often driven by obesity or elevated BMI which occurs in conjunction with loss of β-cell function (Reed, Bain, and Kanamarlapudi 2021) (**Figure 3**). In the context of decreased insulin sensitivity, a temporary glycaemic balance can be maintained as β-cells compensate by increasing in number and size—obesity can increase β-cell size by an estimated 50%—and consequently secrete greater volumes of insulin (hyperinsulinemia) to maintain normoglycaemia (Fonseca 2009; Linnemann, Baan, and Davis 2014). However, elevated insulin secretion and increased β-cell activity are not sustainable and can lead to various forms of stress, including oxidative stress, endoplasmic reticulum stress, dyslipidaemia, amyloid accumulation, and inflammation (Corkey, Deeney, and Merrins 2021). These stressors collectively contribute to the dysfunction and

eventual death of β-cells, culminating in the onset of T2D (Boland, Rhodes, and Grimsby 2017). Importantly, T2D

In practice, T2D manifestation is a combination of relative pancreatic islet function and peripheral IR owing from an interplay of environmental exposures and genetic backgrounds. For instance, obesity promotes T2D through IR, whereas ageing contributes to T2D primarily via β-cell decline (Reed, Bain, and Kanamarlapudi 2021; Tudurí et al. 2022). This heterogeneity has been defined into four distinct clusters. Severe Insulin Deficient Diabetes (SIDD), Severe Insulin Resistant Diabetes (SIRD), Mild Obesity-Related Diabetes (MOD), and Mild Age-Related Diabetes (MARD), all of which present with varying degrees of insulin deficiency and insulin resistance (Ahlqvist, Prasad, and Groop 2020). Notably, the heterogeneity of the disease highlights the need of diverse cohorts which capture each of the subtypes to properly identify the underlying physiological alterations driving each of these groups.

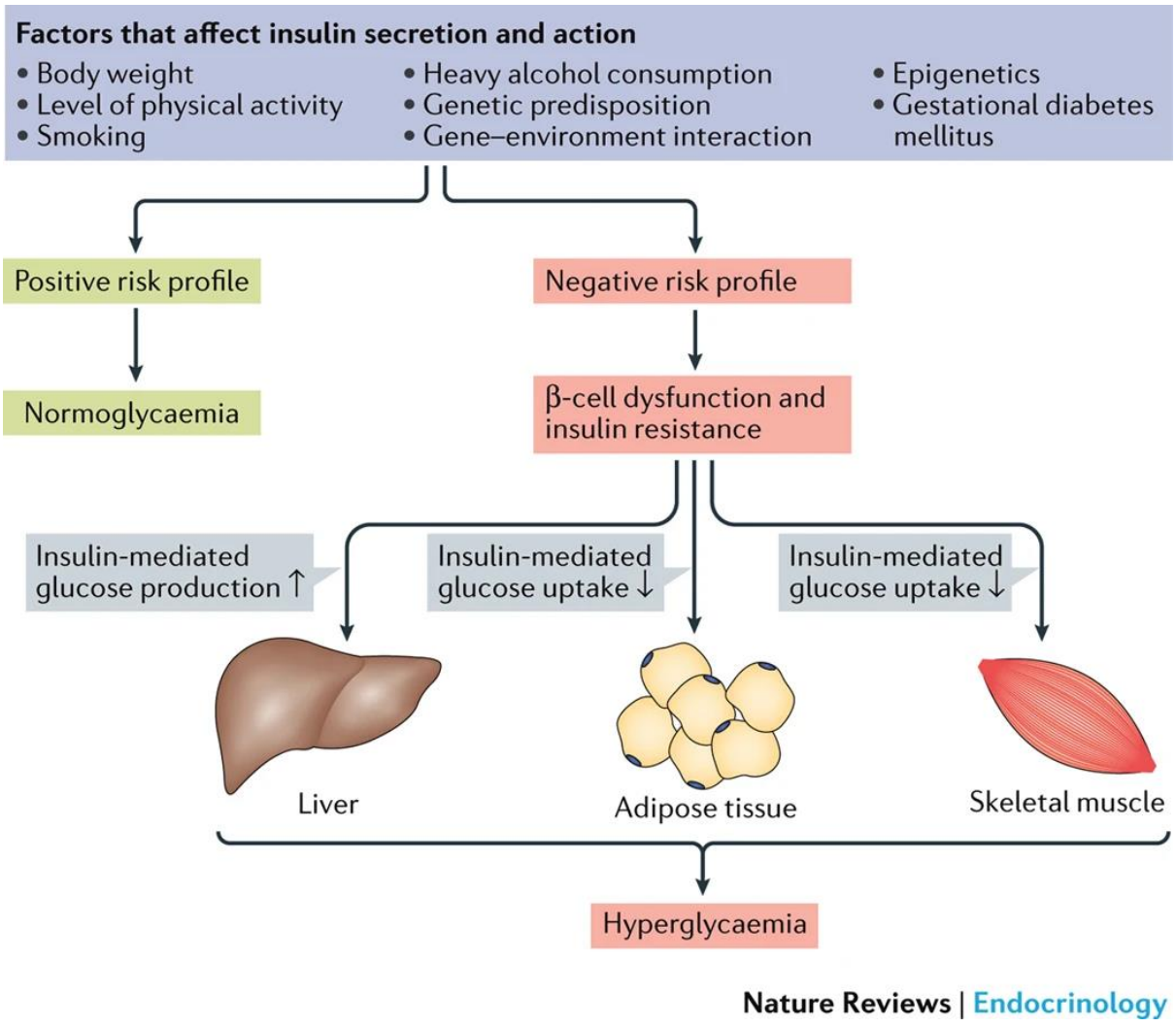


Figure 3 : Overview of the aetiology and pathophysiology of type 2 diabetes. Several environmental factors impact both insulin secretion, and its action, including BMI, physical activity and smoking. These environmental factors can be exacerbated by disease-favourable genetic profiles, leading to the development of T2D. (Zheng, Ley, and Hu 2018).

7.1.1.2 Complications of T2D

T2D is associated with numerous chronic complications stemming from deficient insulin secretion and insulin resistance, significantly contributing to disease mortality. These complications are broadly categorised into microvascular and macrovascular complications, which represent the majority of T2D-associated morbidity and mortality.

7.1.1.3 Microvascular and macrovascular complications

Microvascular and macrovascular complications represent the majority of T2D-associated complications and deaths. Indeed, T2D presents with hyperglycaemia, which damages large blood vessels (i.e., macrovascular), and small blood vessels (i.e., microvascular). Macrovascular complications attributed to diabetes amounted to 179 million deaths in the US, or 53% of all diabetes-related deaths (Parker et al. 2024). Meanwhile, microvascular complications represented an estimated 26.8% of diabetes-related deaths (W. Ling et al. 2020). HbA1C, glycated haemoglobin, which is a measure of average blood glucose, is heavily correlated to diabetes complication incidence and mortality. Indeed, each 1% decrease in HbA1C levels results in a 34% reduction in microvascular complication incidence (Stratton et al. 2000). Similarly, 1% decreases in HbA1C lowers the risk of myocardial infarctions, strokes, and heart failure by 14%, 12%, and 16% respectively (Stratton et al. 2000). Finally, a 1-point reduction in HbA1C lowers diabetes-related deaths by 21% and all-cause mortality by 17%, highlighting the importance of preventing and managing T2D to limit the occurrence of complications (Stratton et al. 2000).

7.1.1.4 Cancers

T2D is also a known risk factor for certain cancers (Cannata et al. 2010). A pooled analysis indicated that the relative risk (RR) for 19 different cancers in individuals with T2D is 1.15 (S. Ling et al. 2020). These included the liver, which features a RR of 2.23, and the kidney, which features a RR of 1.32, both of which are insulin target tissues (S. Ling et al. 2020). Notably, the RR for pancreatic ductal adenocarcinoma (PDAC) is substantially higher, at 2.09, highlighting a strong association between T2D and PDAC development (S. Ling et al. 2020; Yan Li et al. 2019). This association was recently reinforced by a mendelian randomization (MR) which showed a causal role for T2D in PDAC initiation (Maina et al. 2023). PDAC is a rapidly lethal and one of the most difficult to treat form of cancer, due to its late diagnosis. While the exact mechanisms linking T2D to PDAC remain unclear, factors such as hyperinsulinemia,

hyperglycaemia, and chronic inflammation are correlated with an increased risk of PDAC (Yan Li et al. 2019). Acinar cells, which undergo acinar-to-ductal metaplasia (ADM) in response to stress, are exposed to high levels of insulin in states of hyperinsulinemia (Egozi et al. 2020). These cells express the insulin receptor (*INSR*) and are responsive to insulin signalling, which can influence their function and morphology (Pandol 2011). Indeed, exposure of acinar cells to excessive insulin triggers increases in cell sizes and transcriptional alterations, which may induce stress, thus promoting ADM (Egozi et al. 2020). Excessive ADM is one of the hallmarks of PDAC initiation (Shu Li and Xie 2022).

7.2 The pancreas

The pancreas plays a central role in both insulin production and glucose homeostasis, as well as in the secretion of digestive enzymes, underscoring its critical importance in the pathophysiology of T2D and overall metabolic regulation (Leung 2010). Structurally, the pancreas consists of two anatomically and functionally distinct compartments: the endocrine pancreas, which comprises the islets of Langerhans responsible for hormone secretion, and the exocrine pancreas, which produces digestive enzymes (**Figure 4**). Despite their distinct roles, there is significant evidence supporting a bidirectional relationship between these compartments. Indeed, dysfunction in the endocrine pancreas can induce exocrine dysfunction, and vice versa. As described above, T2D increases PDAC risk (Tan et al. 2017). Conversely, PDAC onset often triggers new-onset T2D (Tan et al. 2017).

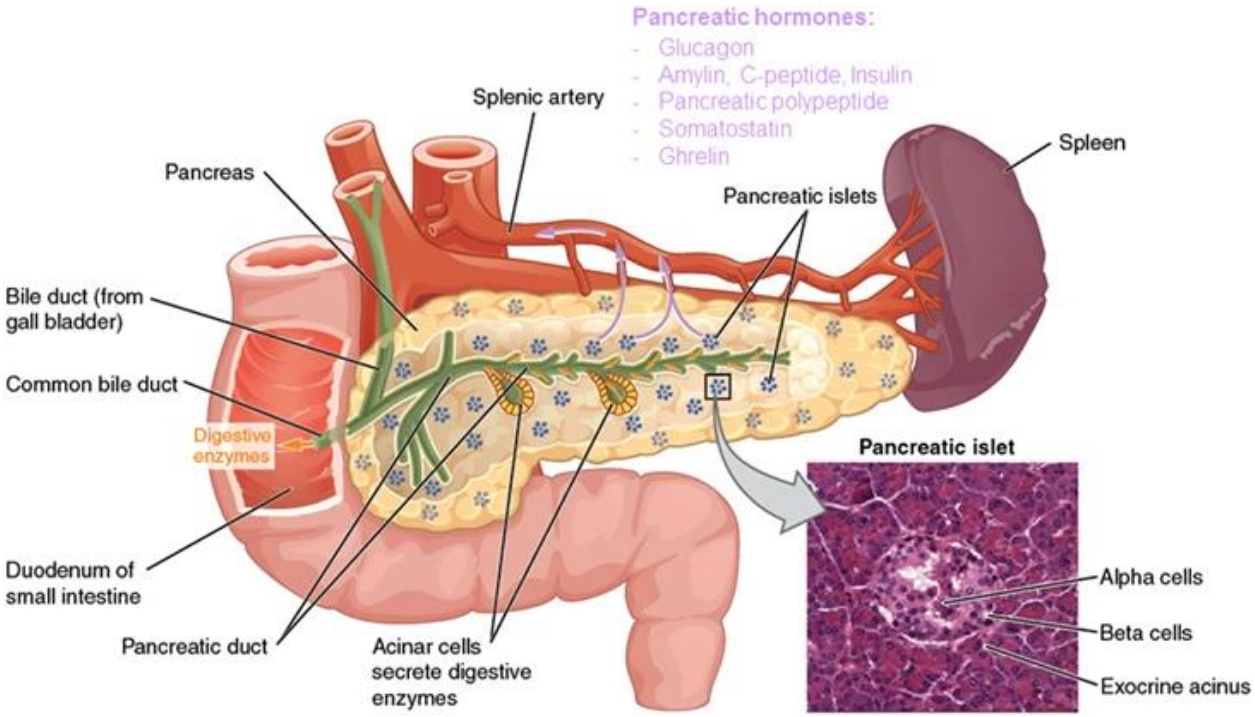


Figure 4: Overview of the pancreas. The exocrine compartment forms the majority of the organ and is composed of enzyme-secreting acinar cells and ductal cells. Pancreatic islets are dispersed throughout the exocrine pancreas and are composed primarily of insulin and glucagon secreting β- and α-cells respectively. Adapted from (Röder et al. 2016).

7.2.1 The endocrine pancreas

Pancreatic islets of the endocrine pancreas constitute only 1–2% of the pancreas and are crucial for maintaining glucose homeostasis throughout the body. This is achieved by orchestrating the production and secretion of two antagonistic hormones: insulin, produced by β-cells, and glucagon, produced by α-cells (Walker et al. 2021). In addition to these primary hormones, the islets produce somatostatin, secreted by δ-cells; pancreatic polypeptide (PP), secreted by PP cells (also known as γ-cells); and ghrelin, secreted by ε-cells. These additional hormones modulate the activity of α- and β-cells, contributing to the fine-tuning of glucose homeostasis (Walker et al. 2021). Pancreatic islets are predominantly composed of β-cells, which account for 55-75% of the islet volume, followed by α-cells, which make up 30-45% of the volume (Cabrera et al. 2006). In contrast, δ-cells, γ-cells, and ε-cells together comprise less than 10% of the total islet cell population (Cabrera et al. 2006).

7.2.2 Pancreatic β-cells and insulin secretion

Pancreatic β-cells are extensively studied due to their crucial role in regulating glucose homeostasis through insulin secretion. Insulin, encoded by the *INS* gene, is initially produced as preproinsulin, a precursor molecule. The Golgi apparatus and endoplasmic reticulum process preproinsulin into insulin and C-peptide, which are stored in vesicles ready for secretion by β-cells. Insulin secretion is initiated through glucose-stimulated insulin secretion (GSIS), a tightly regulated process which is enabled by the ability of β-cells to detect changes in plasma glucose levels and respond according to the body's metabolic needs (**Figure 5**) (Suckale and Solimena 2008). GSIS is triggered by an increase in extracellular glucose levels, resulting in the entrance of glucose into β-cells through glucose transporter 2 (GLUT2). Next, glucose is phosphorylated into glucose-6-phosphate (G6P) by glucokinase. Glucokinase acts as the key censor for glucose in β-cells, indeed, its activity is continuous and directly proportional to intracellular glucose levels (Suckale and Solimena 2008). G6P is utilised as fuel and converted into pyruvate by the glycolytic pathway for use in the tricarboxylic acid cycle pathway (TCA) (Suckale and Solimena 2008). TCA pathway activity increases ATP generation in the mitochondria. An increase in intra-cellular ATP concentrations causes the Kir6.2 channel ATP-dependant potassium channels, encoded by *KCNJ11* and *ABCC8*, to close (Seino et al. 2000). The importance of this channel is evidenced in monogenic diabetes caused by mutations in *KCNJ11* and *ABCC8* (Babenko et al. 2006; L. Liu et al. 2013). Kir6.2 closing induces a depolarisation of the plasma membrane, which results in the opening of voltage-dependent calcium (Ca^{2+}) channels. Calcium influx causes the insulin and C-peptide containing vesicles to fuse with the plasma membrane, releasing the peptides into the bloodstream. Insulin secretion is biphasic, whereby initial GSIS induces the secretion of 1% of insulin-containing vesicles present in β-cells and triggers a secondary, longer-lasting, phase of insulin secretion (Campbell and Newgard 2021). This secondary phase of insulin

secretion allows for the steady secretion of insulin for hours following GSIS (Campbell and Newgard 2021).

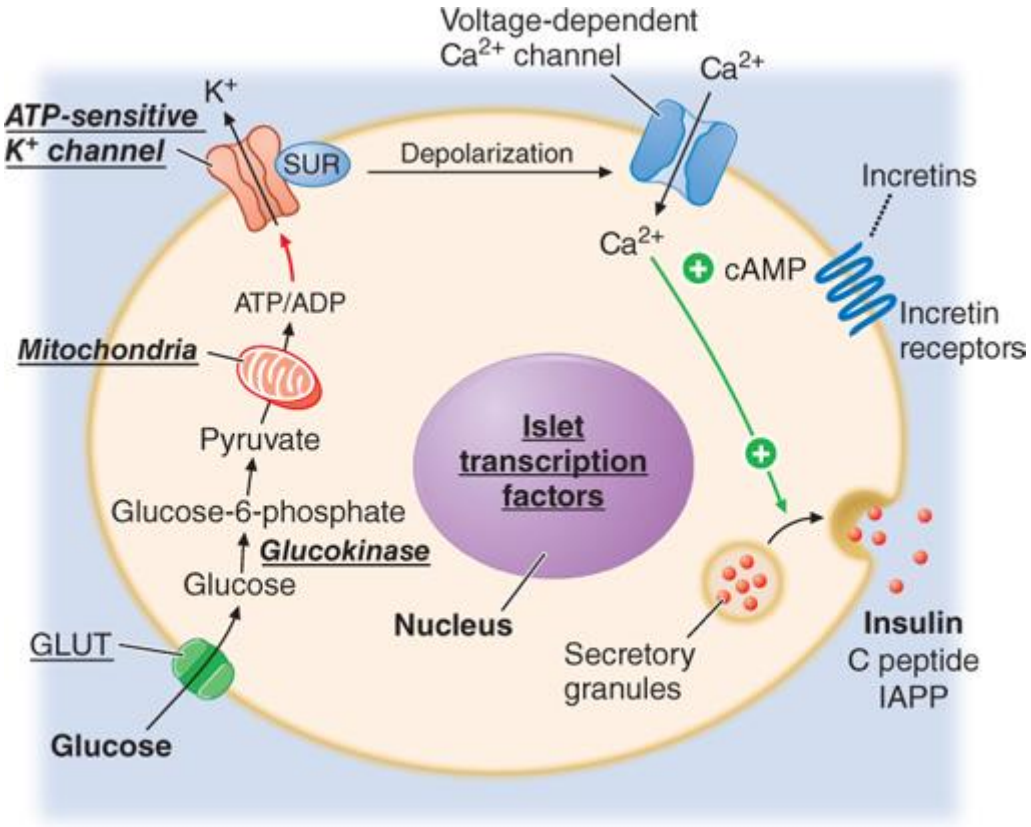


Figure 5: Overview of the glucose stimulated insulin secretion pathway. Glucose enters the cell via the glucose transporter GLUT2. Glucose is then processed the enzyme glucokinase to glucose-6-phosphate (G6P). G6P is used in glycolysis for ATP generation. An increase in ATP tiggers the closing of ATP sensitive K^+, depolarising the membrane, leading to the opening of Ca^{2+} channels, triggering insulin vesicle exocytosis (Kasper et al. 2018).

Post release, insulin enacts its glucose-lowering action by binding to the INSR in the liver, adipose tissue, and skeletal muscle. Insulin binding to INSR induces a molecular cascade that their promotes glucose uptake in the skeletal muscle adipose tissue, or a

reduction in glucose production by the liver (Dong et al. 2006; Petersen and Shulman 2018).

Insulin production is tightly regulated and highly dependent upon proper β-cell and overall pancreatic islet function. Mutations in a number of key islet transcription factors (TF), such as *HNF1A,* or K^+ channels result in MODY (Anık et al. 2015). Mutations in *HNF1A* lead to MODY3, characterised by progressive dysfunction of pancreatic β-cells and loss of insulin production, ultimately causing hyperglycaemia (Anık et al. 2015). These individuals can be treated with insulin therapy. Mutations in *KCNJ11* and *ABCC8* result in deficient insulin secretion owing to permanent K^+ opening. These are better treated by sulphonylureas, which close K^+ channels and restore proper insulin secretion in response to glucose levels (Bowman et al. 2020) This is a powerful example of personalised medicine and the importance of understanding the underlying causes of diabetes.

7.2.3 The exocrine pancreas

The exocrine pancreas constitutes approximately 95% to 98% of the entire pancreas and is composed mainly of acinar cells (about 82%) along with ductal cells. The primary function of acinar cells is to produce and secrete digestive enzymes: amylase for carbohydrates, lipase for lipids, and proteases for proteins. Ductal cells support acinar cells in two primary ways: (1) by forming ducts through which acinar secretions flow into the gastrointestinal tract, and (2) by producing and secreting sodium bicarbonate (HCO_3^-) to neutralise stomach acid in the duodenum. The ducts are connected to clusters of acinar cells via centroacinar cells, which possess both acinar and ductal characteristics (Leung 2010). Together, the acinar and centroacinar cells form the functional unit of the exocrine pancreas, known as the acinus. Each acinus is connected to the ductal system, which converges into the main pancreatic duct, ultimately transporting pancreatic secretions into the gut (Leung 2010).

7.3 The genetics of type 2 diabetes

7.3.1 The heritability of type 2 diabetes and early genetic studies

Early family studies demonstrated a strong genetic component to the development of T2D (Turner et al. 1995). This was later confirmed by two twin-studies, with cohorts sizes of 5,810 and 34,166 individuals, which showed a heritability of the disease ranging from 31% to 72%, respectively (Almgren et al. 2011; Willemsen et al. 2015). Therefore, huge efforts were aimed at unravelling the genetic causes of T2D. Early genetic investigations using family pedigrees identified highly penetrant single variants that co-occurred with disease. This approach identified several genes, including glucokinase, encoded by the *GCK* gene, and the TF *HNF4α* (Fajans, Bell, and Polonsky 2001; Froguel et al. 1993). However, T2D is a complex polygenic disease, caused by several genetic variants with modest effect sizes. To detect these variants, association studies, composed of many individuals to ensure sufficient power, are required. Early targeted association studies were successful in identifying candidate genes for T2D including *KCNJ11* and *PPARG,* but were limited in scope because of the technology available at the time (Altshuler et al. 2000; Gloyn et al. 2003).

7.3.2 Genome wide association studies

Genome wide association studies (GWAS) revolutionised the genetics of T2D. GWAS are very large genetic case-control studies that utilise microarrays to identify single nucleotide polymorphisms (SNPs) associated with a given trait or disease. The first GWAS in T2D was published in 2007, and utilised a cohort of 1275 individuals, identifying 5 SNPs associated with the disease, including in the previously known *TCF7L2* gene, but also in *SLC30A8*, now an established T2D gene (Sladek et al. 2007; Flannick et al. 2014). Since then, a further 158 GWAS for T2D have been published, with gradually increasing sample sizes and number of SNPs being tested. This increase was due to the advancements in array technologies as well as reference panels, such as HapMap or the 1000 genome project, which enable the imputation of

variants not present in genotyping arrays (International HapMap Consortium 2003; Auton et al. 2015). Today, very large multi-ethnic cohorts have been assembled and continue to grow, with the most recent cohort including 428,452 cases and 2,107,149 million non-diabetic controls, which identified 1,289 genetic variants associated with T2D (Suzuki et al. 2024). Indeed, GWAS has been extremely successful in uncovering genetic variation involved in T2D risk (**Figure 6**).

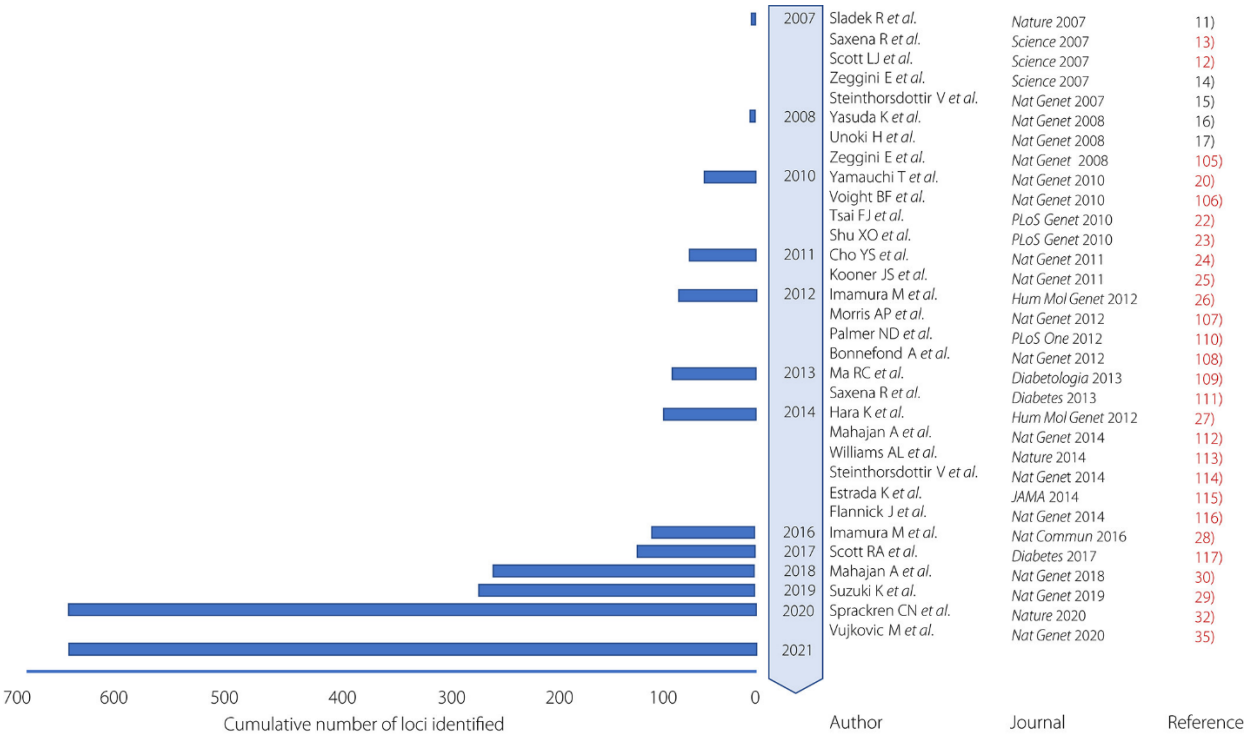


Figure 6: Cumulative number of loci associated with type 2 diabetes per year by genome-wide association studies. The number of loci associated with the disease has increased exponentially through the years, primarily due to very large sample sizes. Adapted from: (Shojima and Yamauchi 2023).

7.3.3 GWAS: limitations and drawing mechanistic insights from associations

Although GWAS have been extremely successful in identifying genetic variants associated with T2D, they feature several important limitations. Most SNPs identified by GWAS lie in non-coding regions of the genome, making it difficult to link a variant to a specific gene and, thereby, to understand the biological mechanisms associated with the T2D risk variant. When a SNP is located within a gene and impacts the amino acid sequence, it is relatively straightforward to identify the gene driving the association, as is exemplified by rs17887074 and rs1801282, which are located within the coding regions of the *SLC30A8* gene and the *PPARG* gene, respectively (Kreienkamp et al. 2023), both of which are implicated in insulin secretion. However, the vast majority of SNPs (> 90 %) lie in non-coding region of the genome (Tak and Farnham 2015). These SNPs do not directly impact protein sequences, rather, they may alter the expression pattern of target genes. In T2D GWAS (and more generally, GWAS for polygenic traits), many SNPs lie in promoter or enhancer regions (Miguel-Escalada et al. 2019; Suzuki et al. 2024).

To gain biological insight from GWAS findings, a powerful tool known as expression quantitative trait *loci* (eQTL) was developed. eQTL integrates genotyping data with gene expression data, *e.g.,* RNA sequencing, to ascertain associations between SNPs and gene expression levels to identify genes targeted by genetic variants. Typically, cis-eQTL analyses focus on nearby genes—within a 500 kb or 1 Mb window—whilst trans-eQTL analyses examine genes located more than 5 Mb away from the genetic variant (Bryois et al. 2014). Since gene expression data is tissue-specific, it is essential that eQTL studies are performed in the tissue relevant to the disease being studied (Arvanitis et al. 2022).

The largest eQTL study to date in human islets was conducted in 514 organ donors, which identified over 1 million associations in 21,115 genes (Alonso et al. 2021). To focus on T2D GWAS signals, colocalisation analyses, which rely on Bayesian inference to identify shared causal signals between GWAS and eQTL SNPs, were performed. Using this approach, this study identified 53 distinct genes target genes of previously

reported T2D-associated SNPs, including known T2D candidate genes such as *CCND2* and *SIX3*, as well as novel genes such as *RMST* (Bevacqua, Dai, et al. 2021; Yaghootkar et al. 2015; Alonso et al. 2021). Furthermore, by integrating their findings with the Genotype-Tissue Expression (GTEx) project, a large eQTL database spanning of 54 tissues, they demonstrated that pancreatic islet specific eQTLs were overrepresented relative to other tissues (Alonso et al. 2021), and were particularly enriched in previously identified islet-specific regulatory regions (Miguel-Escalada et al. 2019). These results underscore that a significant portion of T2D heritability arises from genetically driven physiological changes in pancreatic islets and highlight the importance of investigating biologically relevant tissues.

Another approach is functional validation, where genes of interest are investigated and characterised in pancreatic β-cell or animal models. For instance, an eQTL study found a novel gene *ZMIZ1,* and found that this gene was indeed involved in insulin secretion (Bunt et al. 2015). Another eQTL study found that the *FCHSD2* gene was the target genes of the *ARAP1* locus (Khamis et al. 2019). Further functional validation in the EndoC-βH1, a β-cell model, revealed that the deletion of the enhancer region resulted in the decreased expression of the *FCHSD2* gene and the reduction in insulin secretion (Hu et al. 2021; Pasquali et al. 2014). Similarly, another recent study validated functionally validated the rs12712929 SNP, which had been previously linked to the *SIX3* gene in eQTL studies (Spracklen et al. 2018). Crispr-Cas9 editing in human islets of the genetic variant impaired gene expression of *SIX3* and nearby *SIX3-AS1,* and reduced insulin secretion (Bevacqua, Dai, et al. 2021). Overall, these studies have identified novel contributors of insulin secretion and are gradually translating GWAS signals into concrete mechanistic insights into the pathophysiology of T2D.

7.3.4 Leveraging GWAS to predict disease risk, cause, and consequence

In addition to identifying the genetic contribution of common disease, GWAS associations have been leveraged to predict disease risk and predisposition through polygenic risk scores (PGS) (Padilla-Martínez et al. 2020). PGS aggregate the presence of alleles from predefined loci (T2D GWAS loci for instance) to generate scores for individuals being tested, with higher scores indicating greater genetic predisposition to the disease being assessed. Several PGS models, incorporating different numbers of SNPs, have been developed T2D. Weedon et al. (2006) generated the first PGS for T2D using three SNPs associated to the key T2D genes: *KCNJ11*, *PPARG*, and *TCF7L2*. The score featured a predictive power of an area under the curve (AUC) value of 0.58. Importantly, individuals carrying all six risk alleles of these SNPs had an odds ratio of 5.71 for developing the disease (Weedon et al. 2006). Subsequent studies incorporated risk factors alongside their PGS to enhance predictive power. Khera et al. (2018) utilised 7 million SNPs, and the resulting PGS improved the AUC of clinical variables from 0.66 to 0.73. Similarly, Läll et al. (2017) leveraged 1000 SNPs, and improved the AUC of clinical variables alone from 0.718 to 0.767. Although PGS improve risk prediction for T2D only modestly, they hold significant clinical value. PGS enable patient stratification into low- and high-risk groups, facilitating targeted preventive interventions and personalised healthcare strategies (Khera et al. 2018).

Finally, GWAS have also been leveraged to better understand the epidemiological links between several risk factors, including BMI and smoking, and T2D, through MR. MR is a powerful tool used to determine the causality of one trait, the exposure, or another trait, the outcome. MR assesses causality based on the principle that genetic variation is determined at birth and remains constant, thus serving as an unconfounded proxy for environmental exposures. MR exploits SNPs associated with an exposure (typically identified through GWAS), such as BMI, to determine whether these same SNPs contribute to the risk of an outcome—in this case, T2D. To date, MR studies have demonstrated a causal role of obesity in T2D incidence, as well as in glycaemic traits, such as impaired insulin sensitivity and insulin secretion (T. Wang et

al. 2018). Similarly, MR has been used to confirm the causality of T2D in the development of certain cancers, such as pancreatic and kidney cancer, but not liver or breast cancer (Yuan et al. 2020). Appropriately designed MR isolate the exposure as the sole contributor to the outcome, indicating that T2D-specific alterations are driving certain outcomes, providing important information for the management of T2D complications.

Overall, the development of SNP arrays and GWAS has significantly advanced our understanding of T2D pathophysiology by identifying candidate genes, improving disease prediction, and uncovering factors contributing to both the incidence and progression of the disease.

7.4 Epigenetics

Whilst GWAS have been successful in characterizing the genetic component of T2D, they currently explain only about 20% of the overall genetic risk (Suzuki et al. 2024). This falls short of the heritability estimates, which range from approximately 31% to 71% (Almgren et al. 2011; Willemsen et al. 2015). Furthermore, GWAS do not capture the contribution of environmental factors. To address these issues, and continue characterising T2D, studies have now begun exploring the contribution of epigenetics to disease susceptibility. Epigenetic marks, particularly DNA methylation, are altered by environmental factors, such as ageing, hence epigenetic may mediate the contribution of environmental risk factors to T2D susceptibility. These modifications can affect gene expression and contribute to disease development (Smith, Hetzel, and Meissner 2024). Indeed, risk factors for T2D identified by epidemiological studies including obesity and ageing have been associated with epigenetic modifications which subsequently impact gene expression (Zheng, Ley, and Hu 2018; Charlotte Ling and Rönn 2019). Therefore, these factors and their contribution to T2D can be explored using epigenetic variation as a proxy for environmental factors that

contribute to disease, offering a complementary approach to GWAS to understand the mechanisms underlying T2D pathophysiology.

7.4.1 Overview of epigenetics

Epigenetics is defined by Cavalli and Heard (2019) as "the study of molecules and mechanisms that can perpetuate alternative gene activity states in the context of the same DNA sequence", *i.e.,* the modification of gene expression without modifying the genome. These molecules and mechanisms, such as DNA methylation, histone modifications, and non-coding and small RNAs, are stable over time and through cellular division, and can be passed down to future generations. In multicellular organisms, this stability allows for the existence of distinct cellular states (Cavalli and Heard 2019). The epigenome, however, can be influenced by environmental stimuli. For instance, infections, heat shock and undernourishment have all been shown to induce epigenetic alterations in C.elegans with functional consequences on physiological outcomes (Klosin et al. 2017; Rechavi, Minevich, and Hobert 2011; Rechavi et al. 2014). Consequently, studying the epigenome offers valuable insight into how environmental exposures influence gene expression and contribute to disease susceptibility. Recent technological advances—particularly in high-throughput DNA methylation arrays—have revolutionised the field of epigenetics. These tools enable comprehensive analysis of epigenetic modifications across the genome, providing new opportunities to explore their roles in gene regulation and disease.

7.4.2 Fundamentals of DNA methylation

DNA methylation is an epigenetic modification characterised by the addition of a methyl group to the 5th carbon on a cytosine residue to form 5-methylcytosine (Moore, Le, and Fan 2013). While all cytosines—approximately 25% of nucleotides—can be methylated, the vast majority (about 98%) of DNA methylation occurs at

cytosines followed by guanine nucleotides, known as CpG dinucleotides (Jin, Li, and Robertson 2011). DNA methylation patterns are established during early development, particularly at implantation, following a global demethylation event in the zygote (Smith, Hetzel, and Meissner 2024). This process sets the foundational methylome, which is then dynamically modified throughout the organism's lifespan. DNA methylation is crucial for proper development, contributing to genomic imprinting, X-chromosome inactivation, regulation of repetitive element transcription, and other essential processes (Jin, Li, and Robertson 2011).

The process of DNA methylation is governed by the DNA methyltransferases (DNMT) family of enzymes. DNMT3A and DNMT3B are de-novo methylators which target unmethylated CpG dinucleotides (Figure 7). Mice with *Dnmt3a* knockout (KO) die at 4 weeks of age, whilst *Dnmt3b* KO are embryonic lethal (E14.5-E18.5), highlighting the importance of DNA methylation to development (E. Li 2002; Okano et al. 1999). DNMT1 is primarily responsible for maintaining existing methylation patterns by methylating hemimethylated CpGs—sites where only one DNA strand is methylated—during DNA replication (Jin, Li, and Robertson 2011). Hemi-methylated CpGs occur during DNA replication, as the daughter strand is replicated without the DNA methylation pattern of the father strand. Dnmt1 KO mice are embryonic lethal (Jin, Li, and Robertson 2011). Despite the high variability in DNA methylation patterns among species, DNMT enzymes are highly conserved throughout evolution (Law and Jacobsen 2010). This conservation suggests strong selective pressure on these genes, highlighting the critical role of DNA methylation in the organismal survival.

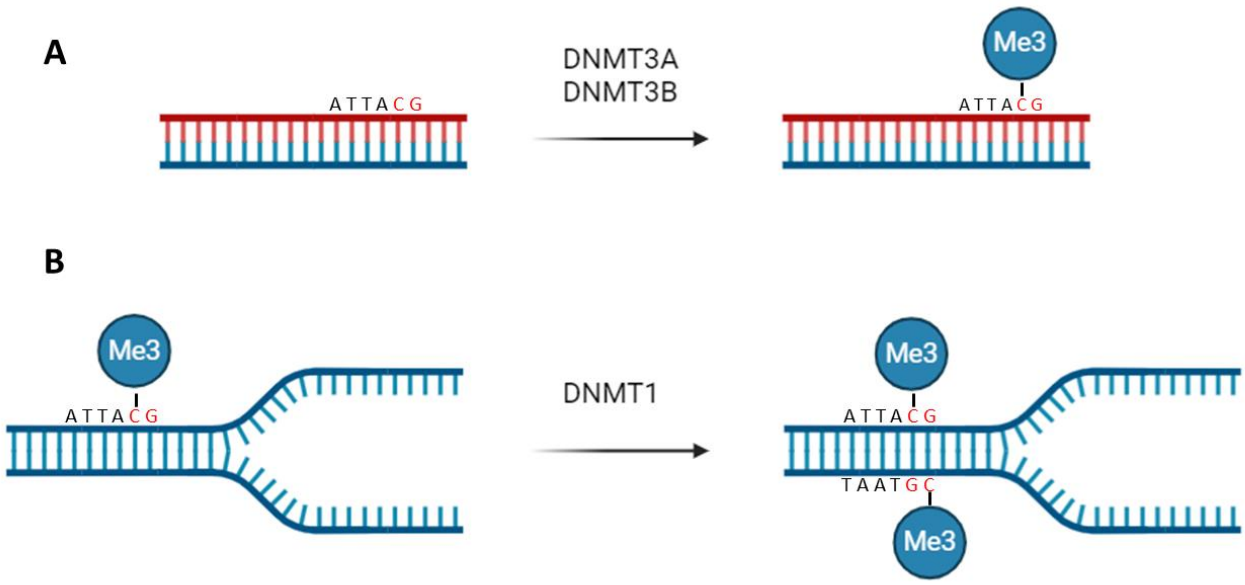


Figure 7: Application and maintenance of DNA methylation marks. A) De novo methylation is applied to cytosines preceding a guanine residue by DNMT3A or DNMT3B. B) DNMT1 replicates DNA methylation marks from the parental strand onto the newly synthesised daughter strand during DNA replication. Figure generated in BioRender.

7.4.3 Distribution of CpGs throughout the genome

CpG dinucleotides are not evenly distributed throughout the genome. Assuming a random distribution of nucleotides, CpG dinucleotides should occur at a frequency of 1 in 16 dinucleotides, given the four possible nucleotides (A, T, C, G) and the 16 possible dinucleotide combinations. Early on, it was observed that CpG dinucleotides occur only at 21% of the expected frequency or roughly every 1/100 dinucleotide pair in the genome (Illingworth and Bird 2009; McClelland and Ivarie 1982). Despite this overall depletion, there are regions known as CpG islands—typically ranging from 200 to 2,000 base pairs—that are rich in CpG dinucleotides (Illingworth and Bird 2009). These CpG islands are mostly unmethylated and are overrepresented in the 5' untranslated region (5' UTR) of genes. Approximately 70% of all human gene promoters contain CpG islands, underscoring their significance in gene regulation (Illingworth and Bird 2009). The remaining CpGs, located outside of islands, are isolated and found in repetitive elements (REs) and transposable elements (TEs)

which make up the bulk (70%) of the genome and are typically packaged in heterochromatin (Koning et al. 2011). These CpGs are most often methylated, unlike CpG islands, and contribute to the silencing of REs and TEs (Pappalardo and Barra 2021).

7.4.4 The contribution of DNA methylation to transcriptional regulation

As CpG dinucleotides are overrepresented in regions critical for regulating gene expression (i.e., promoters), several studies have investigated the role of DNA methylation in this process. Early models suggested that hypermethylation in regulatory regions, particularly promoters and enhancers, prevents TF binding, therefore, repressing gene expression (**Figure 8A**). In this context, changes in methylation levels of CpG islands within promoters can either activate or repress the expression of tumour suppressors implicated in cancer (Zhu, Wang, and Qian 2016). This model, which suggests that hypermethylation leads to downregulation and hypomethylation results in upregulation, often affecting the gene closest to the CpG, has remained prevalent in the field, perhaps because of its mechanistic simplicity. However, it has become increasingly clear that the regulation of gene expression by DNA methylation is more complex. For instance, a genome-wide methylation study in cancer demonstrated a positive correlation between the methylation of many CpG sites and gene expression (Irizarry et al. 2009). Certain TF are able to bind methylated DNA, offering a mechanistic explanation for the observed positive correlations between DNA methylation and gene expression (Zhu, Wang, and Qian 2016). A classic example is MECP2, which contains a methyl-CpG binding domain (MBD). MECP2 acts as both a transcriptional activator or repressor, depending on the context (Figure 8B) (Chahrour et al. 2008). Importantly, certain TFs, such as KLF4, SMYM3, AP2α, exhibit a greater DNA-binding affinity when their target motif features methylated CpGs. This provides a mechanism through which DNA hypermethylation can enhance TF activity, whether they function as repressors or activators (Spruijt et al. 2013). To date, over 100 TFs have been identified that bind methylated DNA *in vitro*, and several have

been validated *in vivo* (Zhu, Wang, and Qian 2016). Beyond regulatory regions, DNA methylation regulates heterochromatin conformation (Grewal 2023). Loss of methylation results in chromatin decompensation, exposing previously inactive DNA, and in particular REs, which are overly represented in heterochromatin (Figure 8C) (Pappalardo and Barra 2021). Heterochromatin disruption impacts normal expression patterns, genomic stability and chromosomal organisation (Bodega and Orlando 2014; Pappalardo and Barra 2021; Shapiro and von Sternberg 2005). Genes encoding all types of RNAs are present in REs and TEs and are expressed upon heterochromatin decompensation (Toubiana et al. 2018). In cancer, global hypomethylations in REs and TEs are observed, particularly in the TEs LINE1s (Long interspersed nuclear element-1). The degree of LINE1 hypomethylation is correlated to tumour progression (Igarashi et al. 2010). Beyond disrupting transcriptional patterns, RE exposure increases the occurrence of recombination events, leading to deletions, insertions, translocations, and inversions, all of which are linked to disease and promote genomic instability (George and Alani 2012).

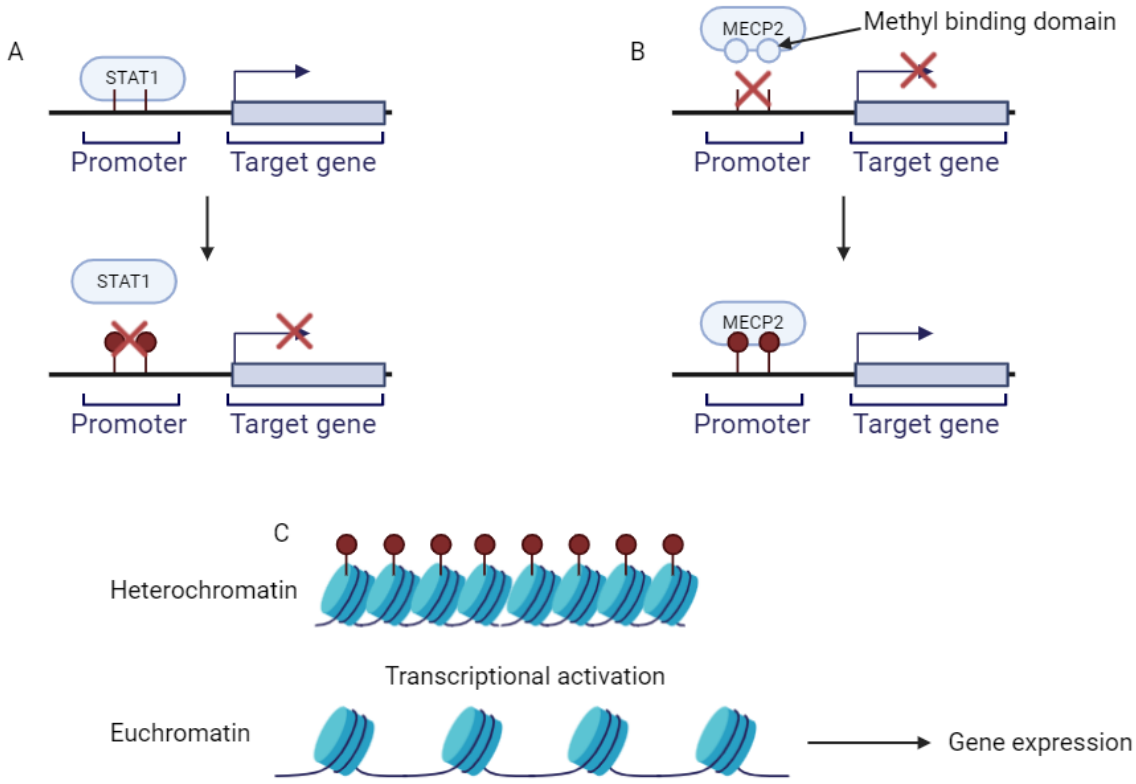


Figure 8: Overview of the mechanisms by which DNA methylation contributes to gene expression. A) STAT1 binds a matching promoter, activating gene expression. Subsequent DNA methylation of the promoter prevents STAT1 binding and therefore represses expression. **B)** MECP2 features a methylation binding domain and cannot bind unmethylated promoters. Methylation of the promoter enables MECP2 binding and initiates transcription. **C)** DNA methylation maintains heterochromatin structure. Hypomethylation of chromatin results in transcriptionally active euchromatin. Generated in BioRender.

7.4.5 Modification of DNA methylation patterns

Changes in DNA methylation are categorised into hypermethylations and hypomethylations, owing to the distinct mechanisms which drive these processes. DNA hypermethylation, which refers to an increase in the proportion of the methylation of a given CpG dinucleotide in a given sample, is an active process mediated by the DNMT enzyme family (Jeltsch and Jurkowska 2014). DNA hypermethylation occurs primarily in tissue specific CpG rich promoters, actively

transcribed gene bodies, and intragenic and intergenic enhancers (Ehrlich, 2019). Paradoxically, hypermethylations are not targeted at exact CpGs, but specific CpGs are consistently hypermethylated in different tissues and physiological contexts (Jeltsch and Jurkowska 2014). This may be explained by the transcriptional alterations driven by hypermethylations, which could enhance cellular fitness and therefore be propagated as cellular replication occurs (Jeltsch and Jurkowska 2014). Indeed, in adrenocortical carcinoma, hypermethylation of specific promoters is associated with cancer progression and escape from the immune system (Rauluseviciute, Drabløs, and Rye 2020). Overall, a combination of DNMT activity and histone marks define hypermethylation patterns (Yinglu Li, Chen, and Lu 2021). In adrenocortical carcinoma, DNMT1 and DNMT3A are both overexpressed and their expression levels are correlated to the hypermethylation of promoters (Rauluseviciute, Drabløs, and Rye 2020). In conjunction, DNMTs show greater affinity for regions marked by H3K9me2/3 histones than regions marked by H3K27me3 histones (Yinglu Li, Chen, and Lu 2021). Importantly, histone marks are not randomly distributed and, therefore, partly direct DNA methylation patterns (Yinglu Li, Chen, and Lu 2021). Indeed, the H3K27me3 and H3K4me3 marks, for which DNMTs have low affinity, are predominantly located in CpG islands and contribute to their maintenance in a hypomethylated state (Yinglu Li, Chen, and Lu 2021). Meanwhile, H3K9me3 marks are enriched in heterochromatin regions to preserve their hypermethylated state (Yinglu Li, Chen, and Lu 2021). DNA hypomethylations are largely stochastic, indeed they are thought to result from a lack of DNA methylation maintenance in a context of cellular stress or elevated cellular proliferation without adequate compensation in the expression of DNMTs. Consequently, hypomethylated CpG sites are less replicable than hypomethylated sites (Tarkhov et al. 2024).

7.4.6 Measuring and studying DNA methylation

DNA methylation, has only recently begun receiving significant attention, despite the concept of epigenetics being first been described over 80 years ago (Waddlngton

1942). Early studies utilised bisulfite sequencing, a method that involves chemically converting DNA using sodium bisulfite to detect DNA methylation (**Figure 9**) During this process, unmethylated cytosines are deaminated by sodium bisulfite and converted into uracil. However, methylated cytosines are resistant to this conversion and remain unchanged. When the DNA is then amplified through polymerase chain reaction (PCR), the uracils are copied as thymines, while the methylated cytosines remain as cytosines. This enables the possibility to distinguish between methylated and unmethylated cytosines by comparing the sequences after bisulfite treatment. Early studies performed bisulfite conversion followed by polymerase chain reaction (PCR) to measure DNA methylation at specific promoters, such as the INS promoter, but generally they were limited in scale, and could only assess a few promoters or target regions per analysis (Paz et al. 2003; B. T. Yang et al. 2011).

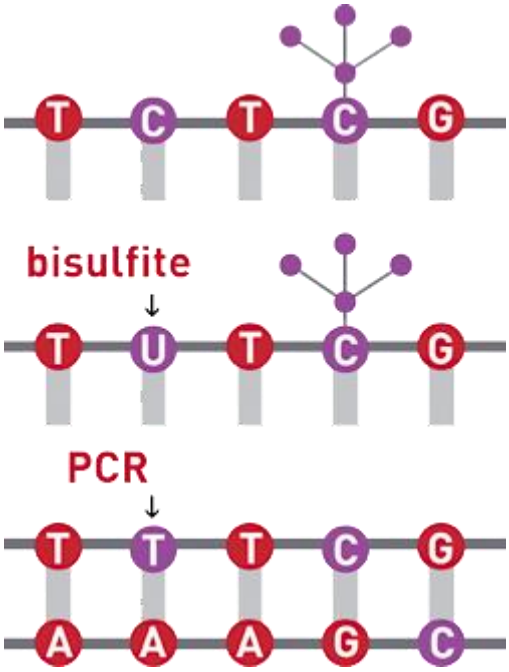


Figure 9: Bisulfite conversion of DNA. Treatment of DNA with sodium bisulfite converts cytosines with no methylation to uracils, whilst methylated cytosines are unchanged. Uracils are substituted to thymines during DNA synthesis. Adapted from: Diagenode.com

The development of DNA methylation arrays over the last decade has revolutionised the field, enabling genome-wide screening of methylation sites. The state-of-the-art

platforms for studying methylation include the Illumina 450k and EPICv1 arrays, which cover 450,000 and 850,000 CpGs, respectively, out of the approximately 28 million CpGs in the genome (Mansell et al. 2019). A more recent addition, the EPICv2 array, which covers 930,000 CpGs, has recently been released. These arrays specifically target CpG sites that are located within regulatory regions, including CpG islands, promoters, enhancers, and transcription factor binding sites.

Methylation arrays measure DNA methylation at each CpG site again by leveraging bisulfite conversion of DNA. In the Infinium MethylationEPIC array, two separate probes per CpG site—one for the methylated (C) state and one for the unmethylated (T) state are used. Each probe ends immediately before the CpG site. During single-base extension, a labelled nucleotide complementary to the target base is incorporated (**Figure 10**). Fluorescent dideoxynucleotides (green for methylated and red for unmethylated DNA) are used to quantify methylation as a beta value, ranging from 0 (unmethylated) to 1 (fully methylated). For instance, a beta value of 0 means no methylation at a given CpG site for the sample being measured, while 1 indicates complete methylation at this CpG.

Methylation arrays have facilitated a more global approach, similar to GWAS, for identifying DNA methylation changes associated with a given trait. These epigenome-wide association studies (EWAS) have been studied in various contexts, and the EWAS Catalog, a database referencing studies with more than 100 samples and at least 100,000 tested CpGs, currently contains 407 entries.

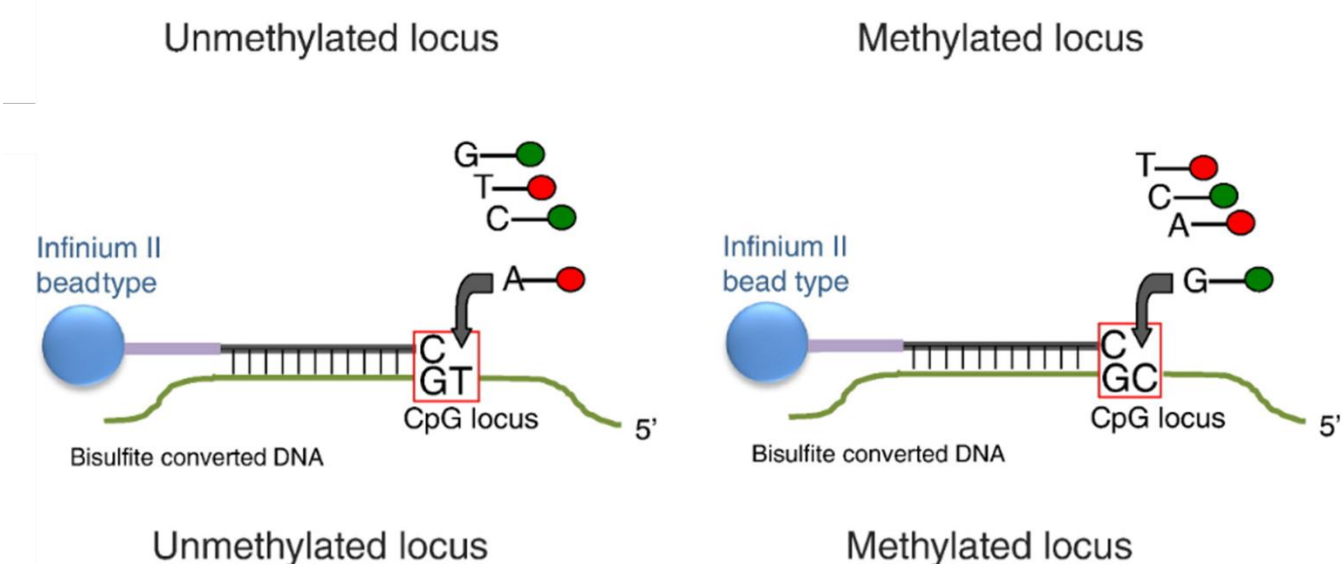


Figure 10: Overview of DNA methylation array technology. To discriminate between methylated and unmethylated CpGs, DNA is bisulfite-converted, whereby unmethylated cytosines are converted to uracil. DNA is then enzymatically fragmented and applied to the array. For each CpG quantified in the array, there are two probes, one which hybridises with the sequence containing a CG, and another with the sequence containing UG, corresponding to methylated and unmethylated DNA, respectively. DNA methylation at a given CpG is then quantified by the relative fluorescence emitted by each of the two probes. Adapted from: Bibikova et al. 2011

7.4.7 DNA methylation studies in T2D and related exposures

In recent years, EWAS have uncovered DNA methylation changes associated with T2D risk, particularly in genes related to insulin secretion, glucose metabolism, and inflammation (Rönn et al. 2023; Bacos et al. 2023). These findings suggest that DNA methylation may mediate the effects of environmental factors on T2D development. Moreover, risk factors like obesity, ageing, and hyperglycaemia have been shown to drive specific methylation changes, further linking epigenetic modifications to T2D progression and its complications (X. Wang et al. 2010; Bollati et al. 2009). These studies are crucial for identifying potential biomarkers and therapeutic targets for T2D and related metabolic disorders.

7.4.8 DNA methylation and T2D

Methylation studies for T2D are limited. Initial small-scale studies of DNA methylation in the context of T2D were focused on specific regions linked to T2D, including the promoters of INS, PDX1, GLP1R, and PPARGC1A, and their expression in pancreatic islet tissue (Hall et al. 2013; C. Ling et al. 2008; B. T. Yang et al. 2011; Beatrice T. Yang et al. 2012). Since then, several EWAS for T2D have been performed. In blood, Chambers et al. (2015), identified five CpG methylation sites associated with incident T2D, highlighting the potential value of DNA methylation to disease prediction and characterisation. However, the focus has shifted towards metabolically relevant tissues, especially pancreatic islets, as blood DNA methylation changes do not consistently replicate in these tissues, hindering the understanding of the underlying molecular mechanisms in disease (de Mello et al. 2014). Early EWAS in pancreatic islets faced challenges due to limited sample sizes. Volkmar et al., (2012) analysed 5 cases and 11 controls, and identified 276 differentially methylated CpGs linked to gene expression changes, enriched in oxidative stress pathways, a component of T2D pathophysiology (Volkmar et al. 2012). Similarly, Dayeh et al., (2014) found 1,649 differentially methylated CpGs in a cohort of 49 pancreatic islets (15 cases), with genes linked to genes with a role in T2D, notably CDKN1A, a gene later implicated in T2D development (Dayeh et al. 2014; Muhammad et al. 2021). The most recent EWAS in pancreatic islets for T2D, which encompassed 100 samples (25 T2D cases), identified 7260 differentially methylated CpGs (Rönn et al. 2023).

In addition, several studies have explored DNA methylation in other metabolic tissues. In the liver, Nilsson et al., (2015) identified 251 differentially methylated CpGs associated with T2D. These CpGs were linked to changes in the expression of 29 genes (Nilsson et al. 2015). Studies in skeletal muscle and adipose tissue are more limited. Ribel-Madsen et al., (2012) performed an EWAS in 11 skeletal muscle and 5 adipose tissue samples of monozygotic twins discordant for T2D (Ribel-Madsen et al. 2012). In skeletal muscle, only one differentially methylated CpG was identified, whilst none were found in the adipose tissue, likely due to limited sample sizes (Ribel-

Madsen et al. 2012). Generally, in the liver and pancreatic islets, a global hypomethylation was observed, with more targeted hypermethylation in promoters, coherent with methylation patterns observed in disease, notably cancer (Nilsson et al. 2015; Rönn et al. 2023).

Although EWAS have successfully identified DNA methylation changes associated to traits, in order to gain biological insight, these alterations must be linked to 1) transcriptomics (or other biological effectors, such as proteomics or metabolomics), and 2) functionally validated. Emerging studies have begun translating these EWAS signals into candidate genes for T2D. Recent studies have integrated DNA methylation with transcriptomics, and have identified OPRD1 and PAX5 as epi-genetically regulated target genes in T2D (Rönn et al. 2023; Bacos et al. 2023). Further functional validation, through siRNA silencing of these genes in healthy human islets, reduced insulin secretion (Bacos et al. 2023). These results highlight 1) the potential of epigenetics to offer insights that extend beyond traditional biomarkers, and 2) the ability of epigenetics to complement genetics studies in characterisation of the molecular architecture of T2D (Bacos et al. 2023). However, while candidates are beginning to be identified, studies which linking DNA methylation to transcriptomics remain limited in scope, as they focus solely on genes in close proximity to the differentially methylated CpG. This approach fails to capture the complexity of the impact of on transcriptomic alterations (Rönn et al. 2023).

7.4.9 DNA methylation in T2D related exposures

DNA methylation has been widely studied in the context of BMI, due to its contribution to disease susceptibility. Initial DNA methylation studies for BMI were performed in blood. The first study, composed of 7 obese and 7 lean individuals, assessed with the HumanMethylation27 chip, composed of 27,000 CpG probes, identified a hypermethylation in the *UBASH3A* gene, and a hypomethylation in the *TRIM3* gene, but drew no further biological conclusions regarding the consequences

of these alterations (X. Wang et al. 2010). A subsequent study by Demerath et al. (2015), in 2097 individuals identified 8 BMI-associated CpGs. Notably, they found BMI-associated CpGs located in proximity to the *HIF3A*, *CPT1A*, and *ABCG1* genes. These genes have since been characterised and shown to contribute to BMI or associated traits, including T2D. *HIF3A* inhibition promotes white adipocyte tissue (WAT) browning, a process which protects from both obesity but also T2D (Cuomo et al. 2022). Meanwhile, *ABCG1* expression in humans is correlated to BMI, a finding reinforced by mice *Abcg1* silencing, which exhibited reduced fat mass development (Frisdal and Le Goff 2015). Importantly, *ABCG1* deficiency also impairs insulin secretion from β-cells (Harris et al. 2018). Finally, *CPT1A* was found to promote fatty acid oxidation in the liver, and its dysregulation promotes IR (Weber et al. 2020; Sarnowski et al. 2023). Demonstrably, these DNA methylation changes appear to either promote obesity or downstream consequences of the disease. This was confirmed by Wahl et al. (2017), who performed a large EWAS (10,261 samples) for BMI in blood. They leveraged transcriptomics and mendelian randomization to 1) link DNA methylation changes to transcriptomic alterations, 2) show that DNA methylation were both cause and consequence of BMI exposure, and 3) that these methylation changes could predict later risk of T2D (Wahl et al. 2017).

7.4.10 DNA methylation and age

Age is a major risk factor for age-related disease, including T2D (Fazeli, Lee, and Steinhauser 2019). Ageing is particularly significant as it leads to a gradual, systemic decline in physiological functions across the body. Understanding the molecular mechanisms of ageing is crucial not only for developing therapies to treat age-related diseases but also for potentially slowing down the ageing process itself. Therefore, DNA methylation in the context of ageing has attracted significant interest, as "epigenetic alterations" have been identified as one of the key hallmarks of ageing (López-Otín et al. 2023).

An early methylation study for age, which targeted 1,413 CpG sites identified a several trends which have since become wide accepted: 1) DNA methylation is strongly correlated to age, 2) DNA methylation in CpG islands increases with age, and decreases outside of these regions, and 3) DNA methylation changes are highly tissue specific (Christensen et al. 2009). A subsequent EWAS identified 37,911 differentially methylated CpGs in the monocytes of 1264 individuals associated with age (Reynolds et al. 2014). This study confirmed that ageing induces a global hypomethylation, with fewer targeted increases in specific regions, including CpG islands, transcription factor binding sites, and enhancers (Reynolds et al. 2014). Additionally, they identified increases in methylation in genes, such as ELOVL2 (cg16867657) and FHL2 (cg06639320), which have since been replicated in several tissues, such as blood, pancreatic islets, liver, and kidney (Reynolds et al. 2014; Bacos et al. 2016a; Bysani et al. 2017; Slieker et al. 2018). A similar trend was observed in other tissues, for instance, in human brain with age, the majority were hypomethylated, but a hypermethylation trend was observed in CpG islands, and hypomethylation in other regions (J. Yang et al. 2015).

An EWAS for age in pancreatic islets identified 241 differentially methylated sites, all of which were hypermethylated and enriched for CpG islands (Bacos et al. 2016a). Moreover, differentially methylated CpG sites in proximity to ELOVL2 and FHL2 were differentially methylated, furthering the observation that there are specific age-associated alterations which are ubiquitous across diverse tissues (Bacos et al. 2016a). This observation led to the classification of age-associated CpGs into two categories: the first are “co-regulated”, which are consistent across individuals and species, while the second category are stochastic and poorly correlated between individuals and species (Tarkhov et al. 2024). Co-regulated CpG sites are strongly correlated to chronological age (Horvath and Topol 2024). This finding has facilitated the development of DNA methylation clocks that utilise a select number of CpGs to predict biological age (**Figure 11**) (Horvath and Raj 2018). These clocks not only predict biological age in blood, but also across several tissues, including the brain,

liver, and skin (Horvath and Raj 2018). Additionally, DNA methylation change with age were shown to correlate with clinical characteristics, such as visceral fat level and liver density (A. T. Lu et al. 2019). Accordingly, DNA methylation age can predict time to death (remaining lifespan) more accurately than chronological age (A. T. Lu et al. 2019; Horvath and Topol 2024). This is supported by observations that deviations of biological age from chronological age—referred to as age acceleration—are associated with earlier disease onset (Bell et al. 2019). For example, a longitudinal study demonstrated that individuals with an older DNA methylation age at baseline were more likely to develop cancer earlier (Bartlett et al. 2019).

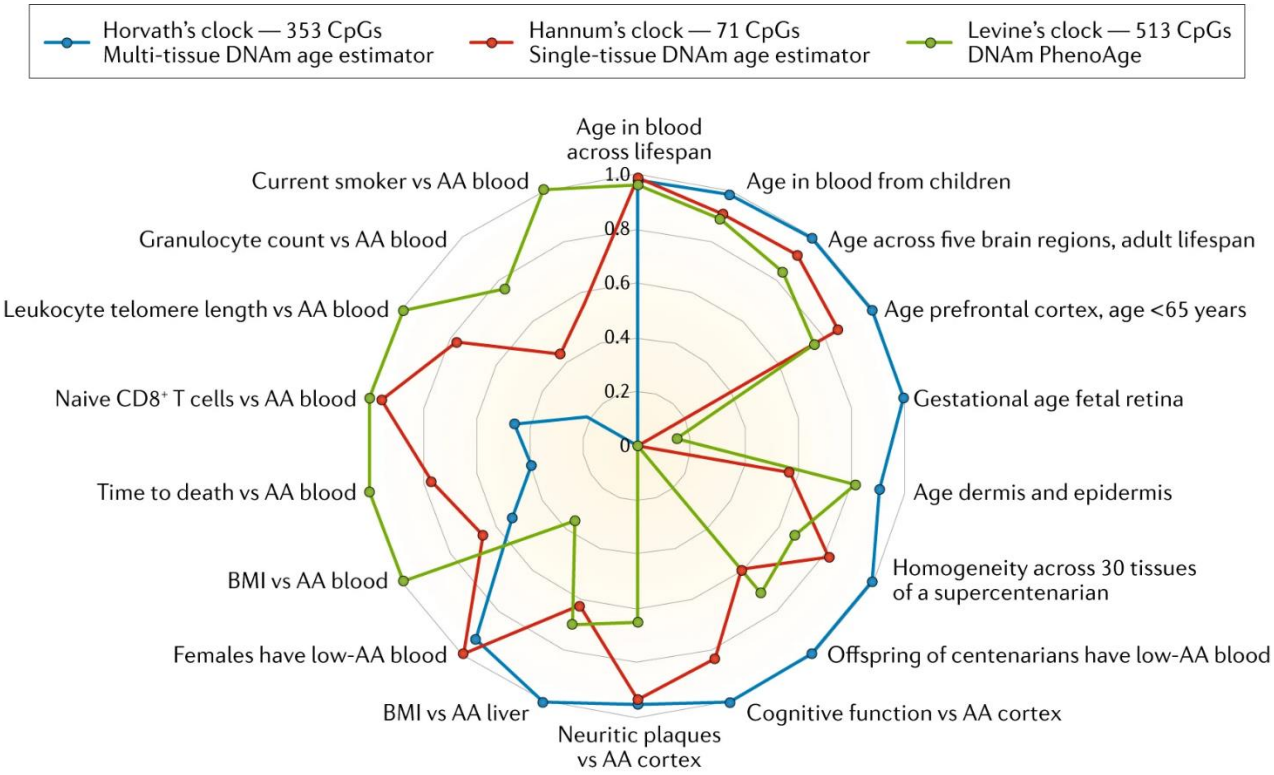


Figure 11: Overview of DNA methylation clocks. A radar plot describes the performance of three DNA methylation clocks in calculating biological age in different human tissues: Horvath's (blue line), Hannum's (red line), and Levine's (green line) DNA methylation clocks. Each clock features distinct performance owing to the CpGs selected for inclusion in the clocks. AA stands for age acceleration; therefore, BMI vs AA blood denotes age acceleration in blood in the context of BMI. Adapted from (Horvath and Raj 2018)

Given the correlation between DNA methylation and lifespan, researchers have investigated whether lifestyle interventions associated with increased lifespan and health span could slow biological ageing. A 2-year diet intervention (plant rich, low meat, low glycemic load foods) in 219 women resulted in a deceleration of biological ageing, as computed by the DNAmGrimAA clock (Fiorito et al., 2021). Further studies have found that all conventional “healthy” and “common-sense” lifestyle actionables, such as eating fish regularly, low alcohol intake, exercise (**Figure 12**), and healthy weight are all associated with biological ageing (Quach et al., 2017). The study of DNA methylation in ageing holds great promise. These results highlight that 1) we can utilise DNA methylation to identify the exposures that promote or limit ageing, 2) we can pinpoint individuals at a greater risk of developing disease. However, it remains unclear whether these age-associated alterations are functional or merely consequences of various exposures.

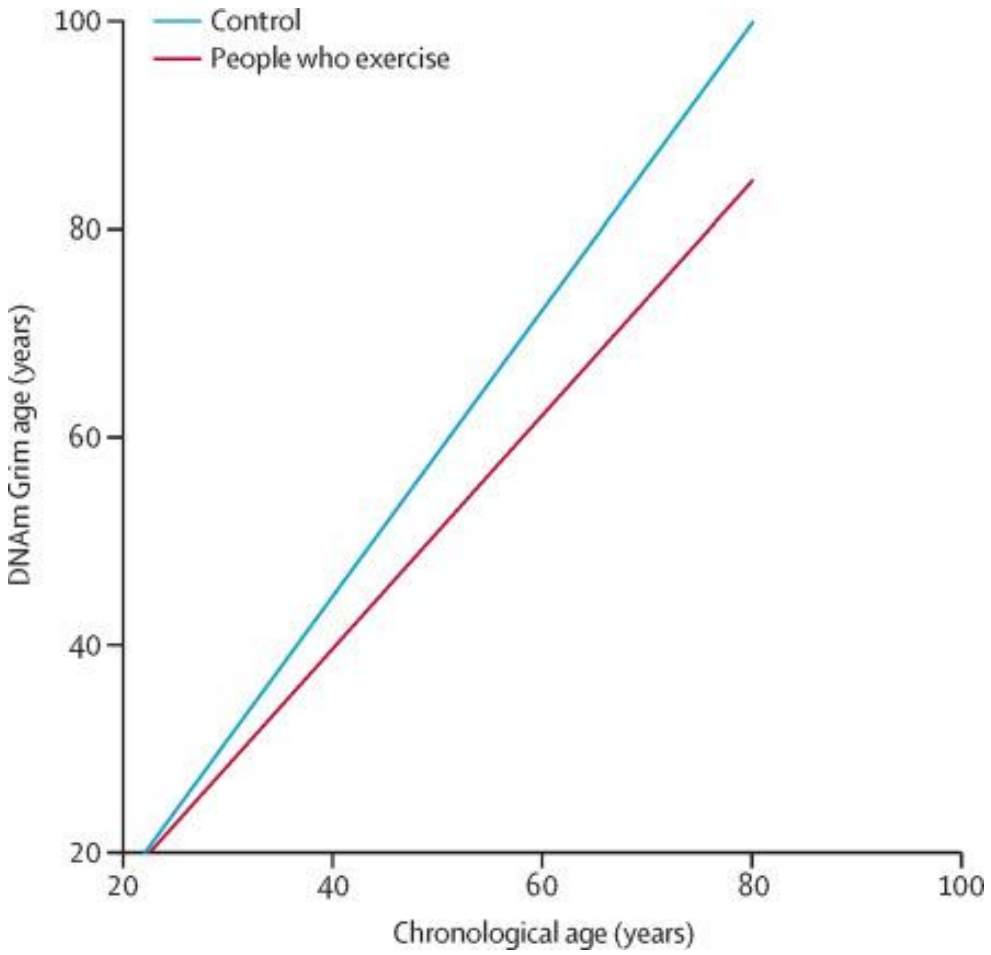


Figure 12: Comparison of biological age relative to chronological age between individuals who exercise or not. Biological age was calculated using the DNAm Grim age DNA methylation clock. The X-axis shows chronological age in years, and the y-axis shows DNA methylation age calculated with the DNAm Grim age DNA methylation clock. Adapted from (Horvath and Topol 2024).

7.4.11 <u>The interactions of genetics and DNA methylation</u>

DNA methylation is also under genetic regulation. Methylation quantitative trait loci (mQTL) link genetic variation, primarily SNPs, to changes in target CpG methylation. Current estimations propose that 10%-45% of the human methylome is under influence by nearby genetic variants (Villicaña and Bell 2021). mQTLs have been linked to various diseases, including cardiovascular disease, and T2D (R.-K. Liu et al. 2021; Huan et al. 2019). Recently, Stefansson et al. (2024) observed that CpGs under the influence of a genetic variant were more likely to impact gene expression. Furthermore, they proposed a model where SNPs impact DNA methylation which subsequently impacts gene expression, highlighting the importance of considering

DNA methylation in the context of genetic studies to fully capture the heritability of the disease (Stefansson et al. 2024). However, it is currently unclear whether the CpGs under the influence of genetic variants are the same CpGs targeted by environmental variables such as age.

7.4.12 The state of DNA methylation today: lessons learnt and limitations

To date, DNA methylation studies have successfully identified alterations associated with a wide variety of traits and diseases across several tissues. A revolution is currently underway in the field of ageing following the development of ageing clocks, which offers promise for strategies aimed at slowing ageing and preventing disease (Horvath and Topol 2024). Additionally, in the context of T2D, EWAS have led to the identification of novel candidate genes (Bacos et al. 2023; Cuomo et al. 2022).

However, the potential of EWAS for T2D has yet to be fully realised. First, only four EWAS for T2D or age have been published in pancreatic islets (**Table 1**). Second, sample sizes remain limited, and only a single EWAS with over 100 samples has been published (Rönn et al. 2023). Second, as demonstrated by BMI based EWAS, DNA methylation changes can induce alterations which promote complications, such as T2D, yet only study Bacos et al (2016), has investigated the contribution of a known T2D risk factor to DNA methylation in a biologically relevant tissue. Another limitation of EWAS, similar to GWAS, is that most DNA methylation sites are located in non-coding regions, making it difficult to pinpoint the causal gene. Additionally, most studies have not properly integrated transcriptomic data to DNA methylation studies to determine the target genes and their biological consequences. For instance, the most recent EWAS for T2D focused on genes located within 10 kb of a differentially methylated CpG (Rönn et al. 2023), despite a recent study demonstrating that many CpG-gene interactions occur beyond a 10 kb window (Kim et al. 2023). Moreover, many EWAS continue to be performed in blood, rather than in biologically relevant tissues. Of the 10 EWAS for BMI published since 2020, only two were performed in adipose tissue, while the remaining eight being performed in blood (Keller et al.

2023). Finally, genetics contributes significantly to DNA methylation variation, yet it is rarely controlled for in EWAS, which fundamentally assess environmental exposure. This oversight may lead to potentially spurious associations (Hawe et al. 2022).

Table 1: Published EWAS performed in pancreatic islets for age and T2D

Reference	Sample size	Trait	Significant CpGs
Volkmar et al., 2012	16 (5 cases)	T2D	276*
Dayeh et al., 2014	49 (15 cases)	T2D	1649
Ronn et al., 2023	100 (25 cases)	T2D	7260
Bacos et al., 2016	87	Age	241

*276 sites reaching unadjusted $p < 0.01$ and > 5% change in methylation level. T2D: type 2 diabetes

7.5 Project aims

The aim of this PhD project was to gain a biological understanding of epigenetic regulation in the context of age and T2D within the endocrine and exocrine pancreas. We sought to perform in-depth multi-omic analyses and functional characterisation to derive functional insights from epigenomic studies. In the first study, we conducted integrative multi-omic analyses in pancreatic islets assess the degree of interaction between genetic variants and epigenetic modifications and to identify the functional consequences of DNA methylation changes associated with age and T2D. In the second study, we aimed to identify epigenetic changes associated with T2D in the exocrine pancreas, that could explain why individuals with T2D have an increased risk of developing pancreatic disease. These studies build upon existing literature, addressing the limitations in the current state of epigenetic research related to T2D and contributing to a better understanding of the complex cis-regulatory architecture of DNA methylation and genetics.

8 Project 1: Environment-mediated DNA methylation changes operate independently of genetic variation to impact the expression of key pancreatic islet genes

8.1 Introduction

T2D is characterised by chronic hyperglycaemia, primarily due to the inability of pancreatic β-cells to secrete sufficient insulin to meet the body's needs (Galicia-Garcia et al. 2020). The decline in β-cell function arises from a complex interplay between genetic and environmental factors. A comprehensive understanding of how genetic and environmental factors influence these molecular changes is critical for unravelling the pathogenesis of T2D. This requires a detailed investigation into the cellular events occurring in human tissues, as these interactions shape physiological traits and disease development. Gaining insights into how these molecular features interact with each other is key to developing effective therapies, interventions, and diagnostic tools for conditions like T2D.

GWAS have been instrumental in identifying the genetic contribution to T2D, revealing over 1,000 genetic variants associated with the disease (Suzuki et al. 2024). In parallel, epigenetic studies in pancreatic islets have shed light on how ageing contributes to T2D pathophysiology. Some research has demonstrated that genetic variants associated with T2D can be linked to changes in nearby DNA methylation that affect gene expression. For instance, one study proposed that the presence of a T2D risk allele allows FOXA1/FOXA2 transcription factors to bind enhancer regions, thereby reducing DNA methylation at the *CAMK1D* promoter and increasing its expression, which may elevate the risk of T2D (Xue, Wu, Zhu, Zhang, Kemper, Zheng, Yengo, Lloyd-Jones, Sidorenko, Wu, Consortium, et al. 2018). To confirm this, a recent study using nanopore sequencing in blood highlighted that genetic variation often serves as the primary driver behind the correlation between methylation and gene expression, identifying thousands of allele-specific methylation quantitative trait loci (Stefansson et al. 2024). These indicate an intricate relationship between genetics

and DNA methylation. Conversely, certain genes, such as *CPT1A*, are 1) epigenetically dysregulated by an environmental exposure (obesity), and 2) are not under the influence of any genetic variant (Weber et al. 2020; Hirota et al. 2007). Similarly, many EWAS have linked environmental exposures to robust epigenetic modifications, particularly in ageing, where CpGs in proximity to *ELOVL2* are consistently hypermethylated (Slieker et al. 2018). Together, these studies suggest that epigenetic modifications can both be dependent and independent of genetic regulation.

While much of the existing literature has focused on the influence of general genetic variation on methylation and gene expression, we took a systematic approach by first mapping the epigenetic landscape of pancreatic islets in response to both ageing and T2D. We then integrated these epigenetic changes with genotyping and transcriptomic data from the same individuals within a single model. This approach was developed to explore the extent to which epigenetic changes associated with age and T2D are influenced by genetic factors or primarily driven by environmental exposures. Our results revealed that the epigenetic alterations linked to age and T2D are largely distinct from genetic regulation. These insights provide a deeper understanding of the cis-regulatory architecture involved in pancreatic islet function and the molecular mechanisms underpinning β-cell decline in T2D.

8.2 Materials and methods

8.2.1 Clinical characteristics of the Epi-Islet organ donor cohort

Pancreatic islets were obtained from 144 brain-dead organ donors, with an age range of 22 to 96 years. The characteristics of the donors are detailed in **Table 2**. These samples were collected from the IMIDIA consortium (Solimena et al. 2018). Of these samples, 25 had T2D, based on the American Diabetes Association (ADA) guidelines (ADA, 2019). Insulin secretion index (ISI) of the pancreatic islet was assessed from all the human islets obtained in this cohort. Differences in clinical characteristics between non-diabetic individuals and individuals were assessed using the Wilcoxon rank sum test. Ethical approval was obtained by next-of-kin's (for organ donors) or patient's (for surgical cases) informed consent, and with the approval of the local ethics committees in Pisa.

8.2.2 DNA extraction, methylation arrays and statistical analysis

DNA was extracted from pancreatic islets samples using the NucleoSpin Tissue kit (T740952.50; Mackerey-Nagel). Bisulphite conversion was performed in a total of 500 ng of DNA from our samples using the EZ DNA Methylation kit (5001; Zymo Research) and subjected to Illumina's 850K EPIC array. Methylation array data was imported using the minfi R package (version 1.50.0) (Aryee et al. 2014). QC steps removed CpG probes if they were: located on sex chromosomes cross-hybridising, non-cg or had a detection threshold p-value of less than 0.01. Probes near or in SNPs were retained. Samples with less than 99 % probes with a detection p-value lower than 0.01 were excluded. Probe-design biases and batch effects were normalised using R packages Enmix (1.40.2) (Z. Xu, Niu, and Taylor 2021) and SVA (version 3.52.0) (Leek et al. 2012) respectively. Following QC, 809,700 probes remained. Sample call rate threshold was set at 95 %. Following QC, 144 samples remained for further analysis.

8.2.3 Principal component analysis

Principal component analysis (PCA) was employed to assess the influence of clinical variables on DNA methylation variation in our dataset and variability across samples. This was performed using the flashpcaR R package (version 2.1) (Abraham, Qiu, and Inouye 2017). The 10 principal components were computed, and an ANOVA was used to test the association of each variable with each PC. Outliers were identified based on Euclidean distance from the centre. The squared sum of standardized scores across the first 10 principal components was calculated for each sample, and those exceeding the 75th percentile plus three times the interquartile range (IQR) were flagged as outliers.

8.2.4 Epigenome wide association study (EWAS)

We applied linear regression models to associate two traits, age and T2D, with CpG methylation level at a given probe. Potential confounding variables were identified from our PCA analysis, *i.e.,* any variable significantly associated with the first principal component (PC1) was flagged as a potential confounder. Based on this, T2D status, age, sex and islet purity were associated with PC1 (**Supplementary figure 1**). Therefore, for age, the best model was adjusted for: sex, islet purity, T2D status. For T2D, the best model was adjusted for: sex, islet purity and age. Additionally, cellular composition, which was estimated using the R package RefFreeEWAS (version 2.2), was also included in both models (Houseman et al. 2016), where we adjusted for J- 1 cell types (where J is the number of cell types defined), and to avoid collinearity, we remove the cell type with the lower estimated proportion . The EWAS models are highlighted below.

$$CpG\ (M\ value) \sim \boldsymbol{Age} + Sex + T2D\ status + Islet\ purity + Cellular\ components$$

$$CpG\ (M\ value) \sim \boldsymbol{T2D\ status} + Sex + Age\ + Islet\ purity + Cellular\ components$$

β-values denote methylation levels, where 0 indicates 0 % methylation and 1 indicates 100 % methylation. For the analyses, β-values were transformed to the more statistically robust -M-values (Du et al. 2010). EWAS was conducted using the limma R package (Ritchie et al. 2015). The bias- and inflation-corrected P-values were then adjusted for multiple testing using the false discovery rate (FDR) method from Benjamini-Hochberg. In order to control for false positives due to genomic inflation, we measured and handled the potential inflation and bias in our results using the bacon method (R package bacon), a Bayesian method based on estimation of the empirical null distribution (van Iterson et al. 2017). For all EWAS, a bacon-FDR < 0.05 was used to determine statistical significance. EWAS identified CpGs were overlapped with islet-specific regulatory regions generated by (Miguel-Escalada et al. 2019). The enrichment of EWAS-identified CpGs for promoters, inactive and active enhancers, and unasigned regions, relative to all CpGs present in the Infinium MethylationEPIC array was assessed using a Chi^2 test.

8.2.5 RNA-sequencing

RNA-sequencing was performed in 142 pancreatic islet samples using 200 ng of RNA using the KAPA mRNA HyperPrep kit (Roche Sequencing). The librairies were sequenced in 2x75 bp paired-end reads using the NovaSeq6000 Illumina system. The average read per sample was of 72,000,000. These were mapped to the human genome on hg37. Sequence demultiplexing was performed using the bcl2fastq Conversion Software (Illumina; version v2.20.0.422). The QC was performed using the FastQC software (version v0.11.9). The removal of adaptor sequences and low-quality

bases was performed with Trimmomatic (version v0.39) (Bolger, Lohse, and Usadel 2014). Genes were quantified using RSEM (version v1.3.0), read with tximport R package (version 1.26.1) and normalised using the method "vst" from the R package DESeq2 (version 1.44.0) (Love, Huber, and Anders 2014).

8.2.6 Genotyping

Genotyping data was generated from DNA extracted from whole pancreas samples from the same study participants. DNA extraction was performed with the DNeasy Blood & Tissue kit (Qiagen, Germany). Genotyping was done with the GWAS Illumina HumanOmni2.5 arrays on the Illumina iScan. Genotypes were called with GenomeStudio software. SNPs were excluded based on the following criteria: call rate < 0.9%, MAF < 0.01%, and a Hardy-Weinberg equilibrium p-value < 1×10^{-4}. Samples were excluded if they featured: discordant sex, a heterozygosity rate greater than four times the standard deviation from the mean heterozygosity rate, or a relatedness (determined by identity by descent) > 0.2. QC resulted in the exclusion of 20 samples and 116,241 SNPs. Imputation was performed with the Haplotype Reference Consortium Panel (McCarthy et al. 2016).

8.2.7 eQTL and colocalization analyses

We combined genotyping and transcriptomics data to identify eQTLs using the nominal pass function from the QTLtools software (version 1.2) (Delaneau et al. 2017). eQTLs were corrected for age, sex, islet purity, and sample origin. A 2 Mb cis-window was used, which includes 1 Mb on each side between the SNP and transcription start site (TSS) of the gene. To detect signals shared between eQTL and selected GWAS, colocalization analysis was focused on the following GWAS data: T2D (Suzuki et al. 2024), proinsulin (Broadaway et al. 2023), random glucose (Lagou et al. 2023), and modified Stumvoll ISI (Williamson et al. 2023). RedRibboncoloc (version 1.3),R package (Piron et al. 2024), which builds on the coloc (version 5.2.3)

(Giambartolomei et al., 2014) package, was used. RedRibbon computes the posterior probabilities for two key outcomes: (1) the likelihood that a region shares a causal variant between the two association summary statistics being tested, and (2) the likelihood that each individual variant present in both summary statistics is the causal variant. A colocalisation was considered significant when the posterior probability for a shared causal variant (PP.H4) exceeded 0.75.

8.2.8 Methylation quantitative trait loci

To identify CpGs under genetic influence, we performed genome-wide mQTL analyses by integrating genotyping and methylation data. mQTLs were performed with the QTLtools software (version 1.2) (Delaneau et al. 2017) using the nominal pass option in cis. A 2 Mb window was used (1 Mb distance between SNPs and CpGs on each side). The mQTLs were corrected for age, sex, and T2D status. All mQTLs with an FDR p-value < 0.05 were considered significant and used in further downstream analyses.

8.2.9 Methylation-Expression adjusted for Genotype Analysis (MEGA)

To address the contribution of DNA methylation changes to gene expression, we regressed all age- or T2D-associated CpGs identified in the EWAS against the expression of all genes within a 2 Mb window. The following formulae were applied:

$$CpG(M\ value) \sim Gene\ expression + Age\ (trait) + Sex + T2D\ status + Islet\ purity + Cellular\ components + Sample\ origin + Genotype$$

$$CpG\ (M\ value) \sim + Gene\ expression + T2D\ status\ (trait) + Sex + Age + Islet\ purity + Cellular\ components + Sample\ origin + Genotype$$

This resulted in trio associations (SNP-DNAm-gene expression). To determine DNA methylation sites that were not associated with nearby gene expression, we focused on 1/ significant DNA methylation and gene expression associations, and 2/ removed

any associations of significant CpGs with nearby genetic variants (genotype). To determine the overarching relationship between the trait and gene expression we computed the “effect direction”. The effect direction of the association was determined by extracting the trait and gene estimate and multiplying the estimate signs, yielded positively and negatively regulated genes with age and T2D.

8.2.10 Pathway analysis

We performed pathway enrichment analysis using the Metascape tool (metascape.org), an integrated platform designed to provide comprehensive functional analysis of gene lists (Zhou et al. 2019). This tool provides an enrichment of pathways based on the following databases: Gene Ontology, Kyoto Encyclpedia of Genes and Genomes, and Reactome.

8.2.11 Risk scores

8.2.12 Methylation risk score

To assess the ability of DNA methylation changes to predict T2D status in our cohorts, we generated methylation risk scores (MRS). For this, we selected CpGs associated (bacon FDR p-value <0.05) with age in our cohort that were 1) located in CpG islands, and 2) located in regulatory regions, to ensure high likelihood of biological relevance. The selected CpGs were intersected with CpGs previously associated with age in 1) a previously published article for pancreatic islets and age (Bacos et al. 2016a) and 2) associated with blood in several published articles obtained through EWAS Catalog (Florath et al. 2014; Reynolds et al. 2014; Tajuddin et al. 2019; McCartney et al. 2019) This criteria resulted in 11 CpGs associated in all studies (**Supplementary Table 9**). To generate MRS for each individual, we applied an additive model, whereby the M-value of each CpG was added. A binary logistic regression was fitted to evaluate the association between CpG methylation score as predictor and T2D status

(cases/controls) as outcome. The resulting MRS was associated to age, ISI, and mean glycaemia using linear models.

8.2.13 Polygenic risk score

To predict T2D status in our cohort using genetic variants, we generated PGS using the PRSice-2 tool (Choi and O'Reilly 2019). We utilised summary statistics' weights from a recent T2D GWAS of multiple ancestry (Suzuki et al., 2024, referred to as the base dataset), and built PGS in our datasets of 124 individuals of European ancestry (referred to as the target dataset). The base and target datasets were strictly independent. To maximise the PGS prediction ability, we applied a pruning and thresholding (P+T) approach with the following parameters: a clumping window of 250 kb, a linkage disequilibrium threshold of $r^2 < 0.1$, and SNP P-value thresholds ranging from 5×10^{-8} to 0.5. This filtering process resulted in the inclusion of 1,475 independent SNPs using a genome-wide p-value threshold of 5×10^{-8}. The PGS was then computed using an additive model. A binary logistic regression was fitted to evaluate the association between PRS as predictor and T2D status as outcome.

8.2.14 Risk score integration

To assess the predictive ability of PGS, MRS, and risk factors (RFs, including age, sex, BMI), we built six classifiers, namely T2D ~ RFs; T2D ~ PGS + RFs; T2D ~ MRS + RFs; T2D ~ PGS + MRS; T2D ~ PGS + MRS + RFs; and evaluated their performance using Receiver Operating Characteristic (ROC) curves from the pROC R package (version 1.18.5) (Robin et al. 2010). As our sample size was limited, we opted for k-fold cross validation to limit overfitting. In this process, the dataset was partitioned into ten approximately equal subsets. For each of the ten iterations, one subset was held out as the validation set, and the remaining nine subsets were used to train the PGS model. We reported the AUC with their 95% confidence interval for each classifier to quantify their predictive accuracy.

8.3 Results

8.3.1 Epi-islets cohort clinical characteristics

Pancreatic islet samples were obtained from 144 organ donors, including 25 with T2D, spanning an age range of 22 to 96 years (**Table 2**). BMI was not significantly associated with T2D status (P = 0.12), and individuals with T2D exhibited higher mean glycaemia (P < 0.001) and reduced ISI (P < 0.001). DNA was extracted from pancreatic islets and profiled for DNA methylation using the Infinium MethylationEPIC array. The resulting data were used to conduct an epigenome-wide association study (EWAS) to investigate associations with both age and T2D across the entire cohort.

Table 2: Clinical characteristics of the epi-islets organ donor cohort

Characteristic	Overall, N = 144[1]	ND, N = 123[1]	T2D, N = 21[1]	p-value[2]
Age (years)	69 (22, 96)	68 (22, 96)	76 (58, 92)	**0.031**
Sex				0.2
Female	60 (42%)	54 (44%)	6 (29%)	
Male	84 (58%)	69 (56%)	15 (71%)	
BMI (kg/m²)	24.5 (23.0, 27.1)	24.5 (22.9, 26.9)	25.8 (23.4, 27.7)	0.12
Mean glycaemia (mmol/l)	149 (121, 182)	144 (120, 173)	193 (165, 227)	**<0.001**
insulin secretory index (ISI)	2.34 (1.67, 3.38)	2.57 (1.75, 3.60)	1.80 (1.40, 2.14)	**<0.001**
[1] n (%); Median (IQR)				
[2] Wilcoxon rank sum test; Pearson's Chi-squared test				

8.3.2 Epigenome-wide association study for age and T2D

We identified 1,092 age-associated CpGs (bacon corrected P < 0.05), with 86% showing hypermethylations (**Figure 13A; Supplementary Table 1**). Our results were consistent with previous studies, as we replicated 147 CpGs identified in an EWAS for pancreatic islets (**Figure 13B**), and found CpGs which are ubiquitously hypermethylated with age across multiple cohorts and tissues in genes like *ELOVL2* and *FHL2* (Bacos et al. 2016; Bysani et al. 2017) (**Supplementary Table 2**) . In contrast, 3,262 CpGs were associated with T2D, the majority of which were hypomethylated

(**Figure 13C; Supplementary Table 3**). Unlike age, the overlap between our results and a recent T2D EWAS (Rönn et al., 2023) was minimal (88 CpGs, 0.3%; **Figure 13D**), suggesting the presence of significant variability in the methylation profiles of T2D. In accordance with this, we observed substantial heterogeneity in the DNA methylation profiles of our individuals with T2D (**Supplementary figure 3**).

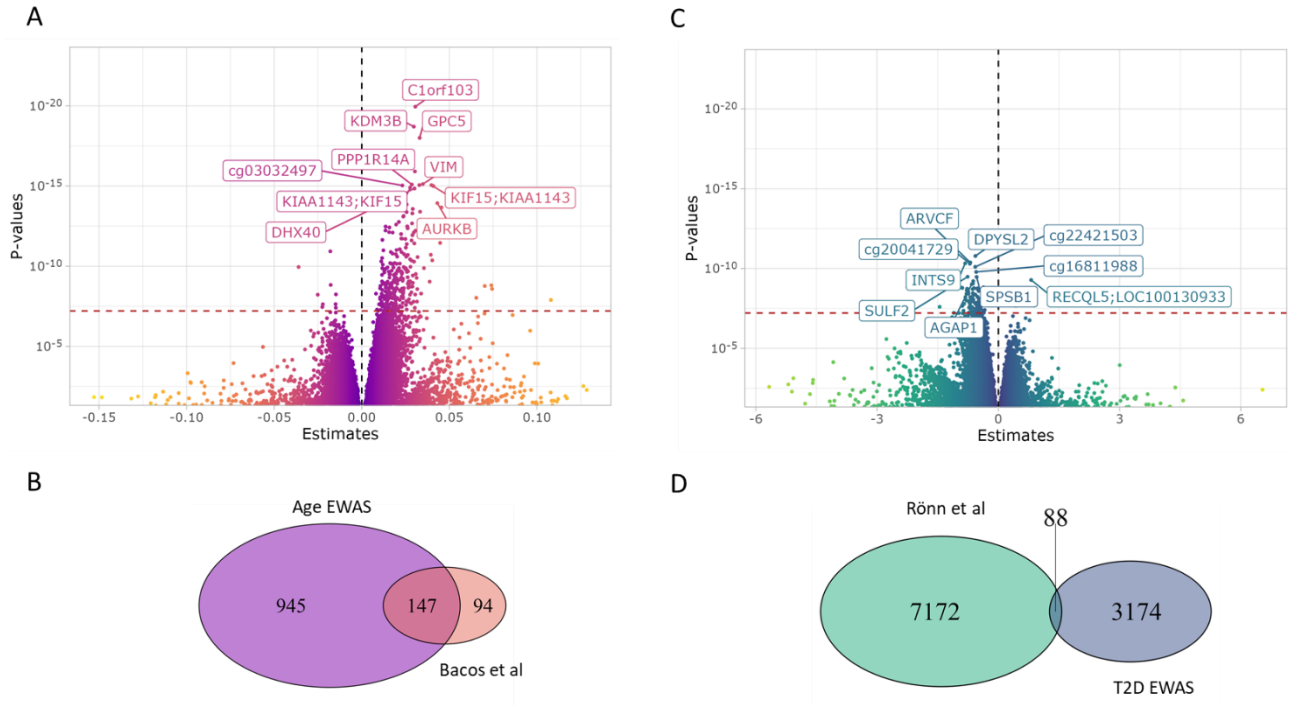


Figure 13: Epigenome wide association study reveals genes associated with age and T2D. A) Volcano plot showing DNA methylation changes associated with age, and **B)** comparison with Bacos et al., to identify shared CpGs associated with age. **C)** Volcano plot showing DNA methylation changes associated with T2D, and **D)** comparison with Ronn et al., to identify shared CpGs associated with T2D.

8.3.3 The interplay between DNA methylation, gene expression, and genetic variants

This study aimed to assess the relationship between genetic variants and environment-associated DNA methylation changes. To this end, RNA sequencing and genotyping data were generated from all individuals, of which 124 passed all QC for all generated omics. We developed a linear model to test associations between CpGs and nearby gene expression (within a 2 Mb window), and including the nearby genetic variants for each specific CpG site. We employed this approach to investigate the prevalence of two gene regulatory models. The first model involves genetic variants, alongside environmental variables, influencing DNA methylation, which in turn impacts gene expression, termed regulatory triads. The second model considers environmental factors, such as age or T2D, that directly influence CpG methylation and affect gene expression without the involvement of SNPs (**Figure 14**).

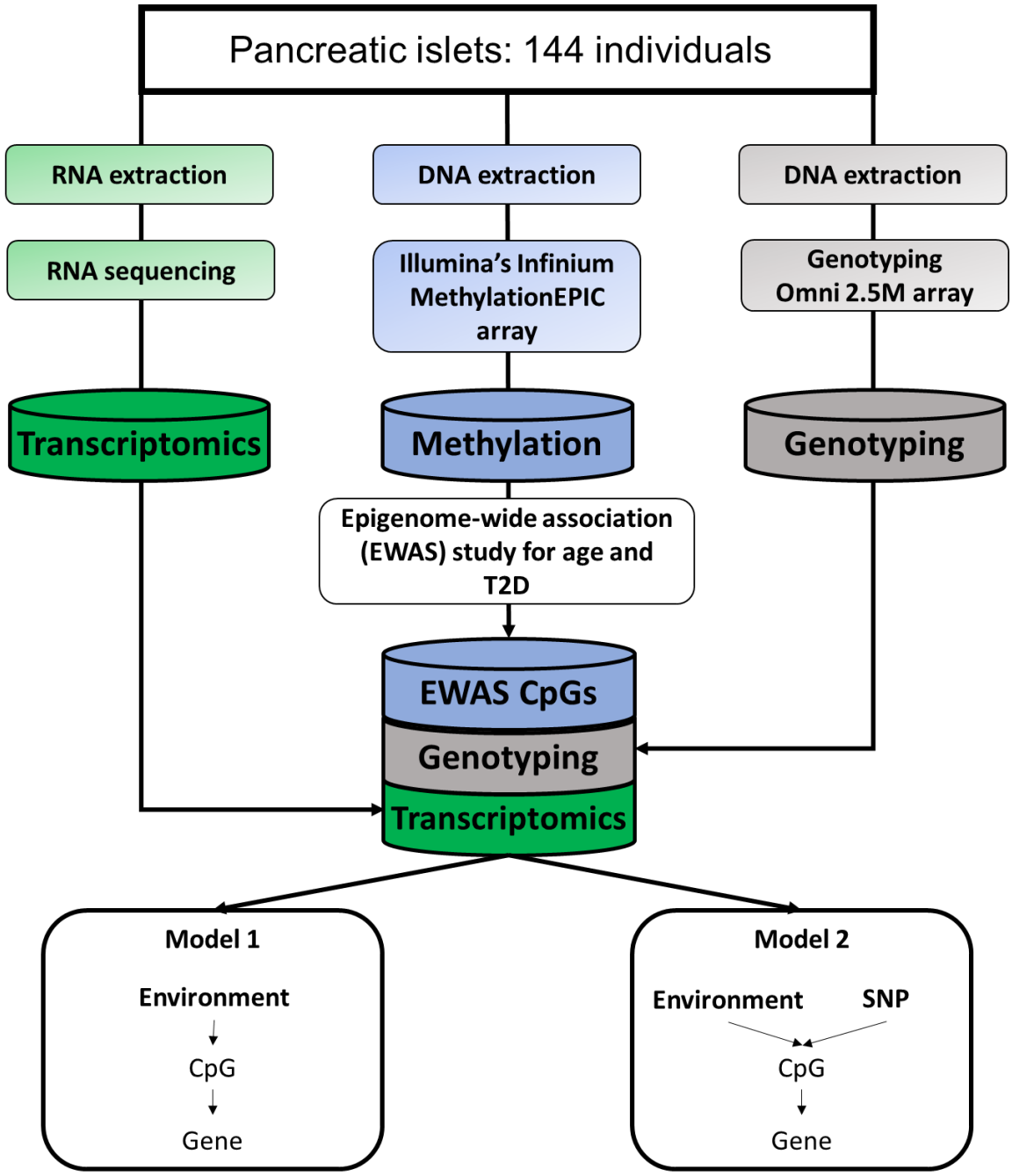


Figure 14: Overview of the project pipeline. 144 organ donors were recruited for this study. Methylation, genotyping, and transcriptomic data were generated using Illumina methylationEPIC arrays, Omni2.5M genotyping chips, and RNA sequencing, respectively. After quality control (QC), 124 individuals remained for further analysis. EWAS identified CpGs associated with both age and T2D. These significant CpGs were integrated with genotyping and transcriptomic data in a unified model. This model aimed to determine whether CpGs associated with nearby gene expression were influenced by 1) genetic variants or 2) independent environmental factors, within a 2 Mb window.

Our results revealed that only 12 triad interactions were detected for the 1092 age-associated CpGs (**Supplementary Table 4**), and 57 triads for the 3062 CpGs T2D-associated CpGs (**Supplementary Table 5**), which is less than expected, as we found 15% of all CpGs to be influenced by SNPs in our genome-wide mQTL. These findings suggest that SNP-mediated epigenetic regulation, influenced by environmental factors, is rare and likely context-specific. To further investigate genetic contributions to gene expression, we performed a colocalisation analysis to identify eQTLs (SNPs affecting gene expression) and compared these findings with our triad analysis (**Supplementary Table 6**). This approach allowed us to examine whether SNPs influence gene expression independently of CpG methylation. The colocalisation analysis specifically focused on genetic variants overlapping with T2D-related GWAS signals, highlighting key regulatory SNPs that may contribute to disease risk through their impact on gene expression.

For our T2D-associated CpGs, this resulted in two target genes associated with T2D GWAS SNPs, *SIX3* and *ST6GAL1* (**Table 3**). These genes are co-regulated by different SNPs, with a genetic variant influencing gene expression and another CpG methylation. This dual regulation underscores the complex genetic-epigenetic interplay contributing to β-cell dysfunction in T2D, positioning both *SIX3* and *ST6GAL1* as promising candidates into the molecular mechanisms underlying T2D. For instance, we found that the decreased methylation of the cg06478249 CpG, located in an inactive enhancer, is associated with T2D (FDR-storey = 2.98×10^{-8}; **Figure 15A)**, and a decreased methylation is associated with an increased expression of the *ST6GAL1* gene **(Figure 15B)** Moreover, the rs12632862-GG variant is associated with decreased methylation at cg06478249, as demonstrated in our triad analysis **(Figure 15C)**. Additionally, the colocalization analysis revealed that the T2D risk variant rs3887925-TT, located in an active enhancer, is linked to increased *ST6GAL1* expression. Interestingly, rs12632862-GG and rs3887925-TT are in strong linkage disequilibrium ($R^2 = 0.87$), suggesting that these two variants work in concert to regulate *ST6GAL1* expression and methylation **(Figure 15D)**.

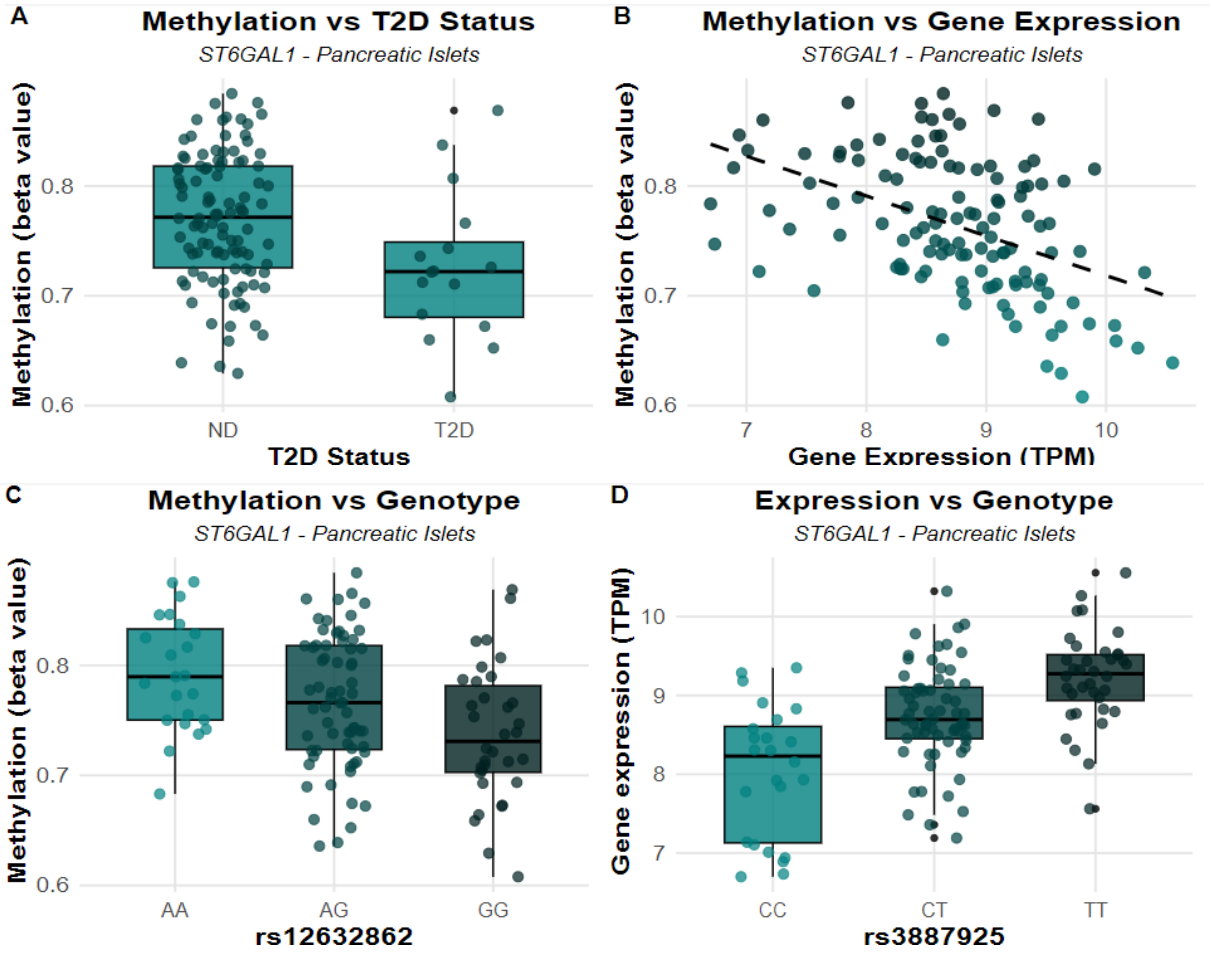


Figure 15: The relationships between DNA methylation, gene expression, and genotypic variation for the ST6GAL1 gene in pancreatic islets. A) Methylation (beta values) vs T2D status: compares the methylation of the CpG site cg06478249 between non-diabetic (ND) and type 2 diabetes (T2D) individuals. **B)** Methylation vs Gene Expression: A scatter plot illustrating the negative association between the methylation levels of cg06478249 and *ST6GAL1* gene expression (TPM). **C)** Genotype rs12632862 (SNP1) vs Methylation: Boxplot showing the effect of rs12632862 genotypes (AA, AG, GG) on the methylation levels of cg06478249. **D)** Genotype rs3887925 (SNP2) vs Gene Expression: Boxplot showing the association between rs3887925 genotypes (CC, CT, TT) and *ST6GAL1* expression.

Table 3: SNP-CpG-Gene triads with colocalisation signals for T2D or related traits

SNP	Position	Association trait	Gene Name	GWAS	H4 probability	SNP effect direction	CpG effect direction
rs3887925	3_186665645_C_T	T2D	*ST6GAL1*	T2D	0.999	Increase	Increased
rs12712928	2_45192080_G_C	T2D	*SIX3-AS1*	Random glucose	0.986	Decrease	Decreased
rs12712929	2_45192105_G_T	T2D	*SIX3-AS1*	Proinsulin	0.986	Decrease	Decreased
rs12712928	2_45192080_G_C	T2D	*SIX3*	Proinsulin	0.967	Decreased	Decreased
rs12712929	2_45192080_G_C	T2D	*SIX3*	Random glucose	0.966	Decreased	Decreased
rs2053005	15_66704449_G_A	Age	*TIPIN*	T2D	0.844	Increased	Increased

Additionally, only one age-associated target gene, *TIPIN*, was also under genetic regulation. The cg06993413, located in an active enhancer, was associated with age (**Figure 16A**), and this CpG was associated with the upregulation of the *TIPIN* gene (estimate = 0.20, FDR-storey = 0.004; **Figure 16B**)**.** While not forming a triad, increased *TIPIN* gene expression was associated with the recently identified rs2053005-AA T2D GWAS risk variant (eQTL p-value = 4.20×10^{-12}; **Figure 16C**). Additionally, *TIPIN* gene expression was significantly increased with mean glycaemia in our organ donors, further confirming our results (**Figure 16D)**. This suggests a more indirect but potentially important role in gene regulation, particularly in the context of ageing and T2D, offering further avenues for understanding its contribution to disease mechanisms.

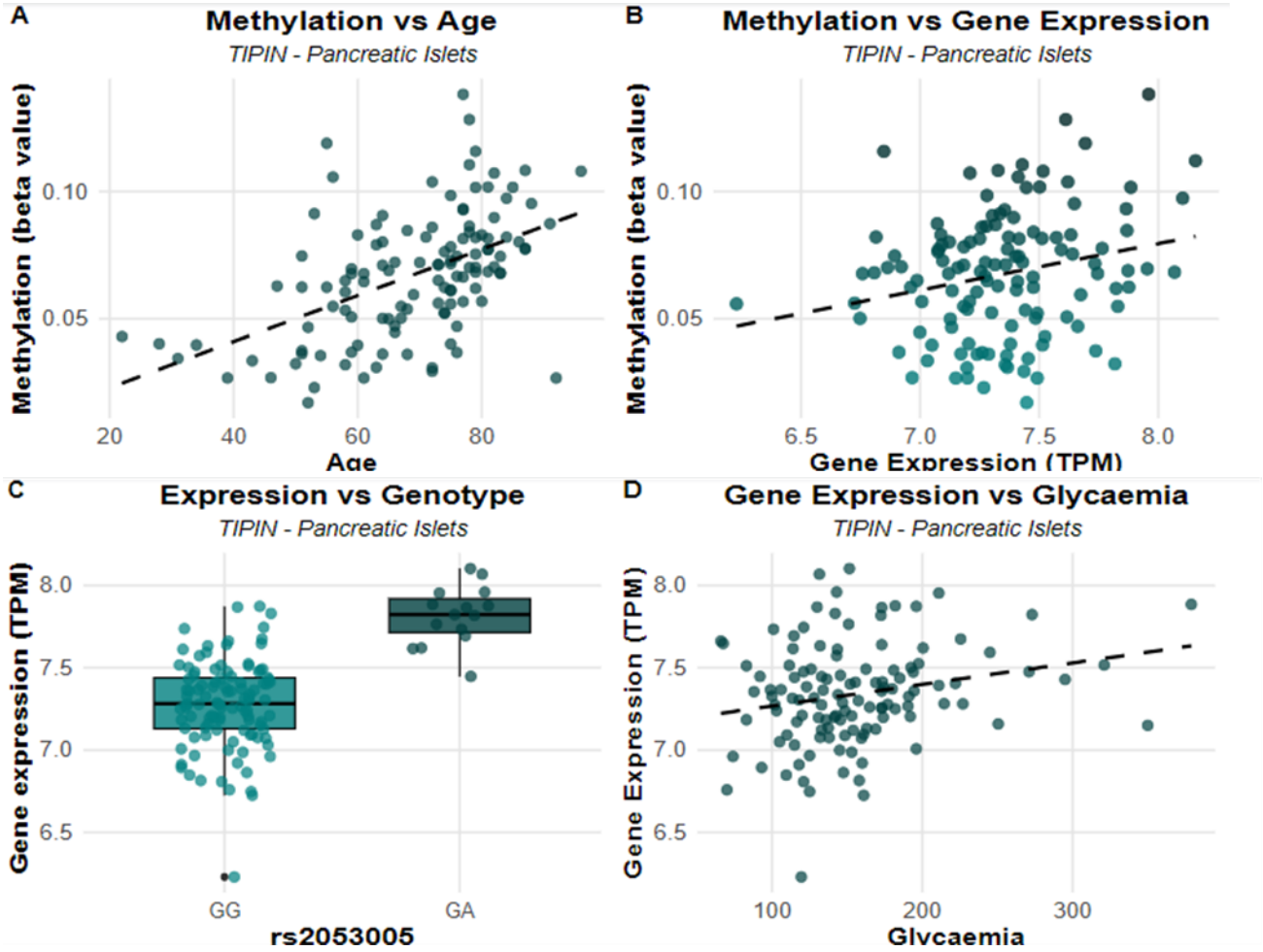


Figure 16: The relationships between DNA methylation, gene expression, and genotypic variation for the *TIPIN* gene in pancreatic islets. A) Methylation (beta values) vs age: compares the methylation of the CpG site cg06993413 with age**. B)** Methylation vs Gene Expression: A scatter plot illustrating the positive association between the methylation levels of cg06993413 and *TIPIN* gene expression (TPM). **C)** Genotype vs gene expression (eQTL): Boxplot showing the effect of rs2053005 genotypes (GG, GA) on the gene expression levels of *TIPIN*. **D)** Gene expression and mean glycaemia: Scatterplot showing the association between TIPIN expression and mean glycaemia of the organ donors.

.

8.3.4 The interplay between DNA methylation and gene expression, independent of genetic variation

Having established that environment-associated CpGs are depleted of genetic influence, we sought to assess how these impact nearby gene expression. We focused on CpGs associated with age or T2D (FDR p-bacon < 0.05) and with nearby gene expression (FDR < 0.05) and excluded CpGs associated with nearby SNPs (FDR > 0.05). Using this approach, we found that 288 unique CpGs (25%) were associated with the expression of at least one nearby gene (**Supplementary Table 4**). Similarly, we found that of the T2D-associated CpGs, 700 CpGs (21%) were associated with a nearby gene (**Supplementary Table 5**). For both T2D and age-associated CpGs, the vast majority of CpGs did not target the nearest genes (**Supplementary figure 4**), and no over 40% of CpGs targeted genes at a distance > 500kb for both age and T2D (**Figure 17A-B**). Age-associated CpGs that targeted a nearby gene were enriched for CpGs islands ($P = 2.368 \times 10^{-12}$), promoters ($P = 2.599 \times 10^{-4}$), inactive enhancers ($P = 2.714 \times 10^{-2}$), and were depleted in non-regulatory ($P = 7.836 \times 10^{-10}$) and open sea regions ($P = 2.368 \times 10^{-12}$) (**Figure 17C**). In contrast, T2D-associated CpGs were enriched for open sea regions ($P = 4.306 \times 10^{-5}$), active enhancers ($P = 7.489 \times 10^{-4}$) and depleted for CpG islands ($P = 4.306 \times 10^{-5}$) and promoters ($P = 5.392 \times 10^{-4}$) (**Figure 17D**). This suggests that while age impacts gene expression more directly through promoters, T2D may lead to broader, less targeted epigenetic dysregulation. Additionally, we observed no uniform pattern based on the genomic location of the CpGs (**Supplementary figure 5**).

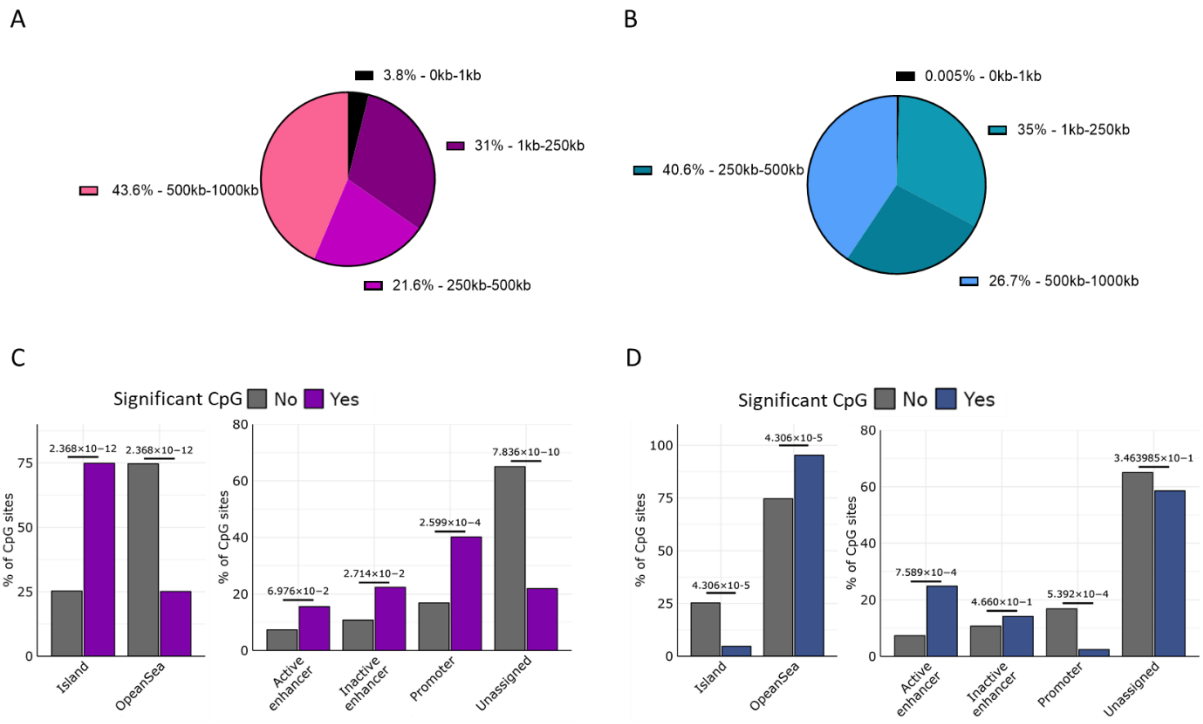


Figure 17: Target genes of age-and type 2 diabetes (T2D) associated CpGs. A-B) Pie charts demonstrating the number the number of CpG-gene associations split by CpG-gene distance for age (left), and T2D (right). **C)** Bar plots showing the count of CpGs in the context of their genomic (CpG islands, open sea) or regulatory position (active enhancer, inactive enhancer, promoter, unassigned). Purple indicates the proportion (%) of significant CpGs associated with a change in gene expression in the age analysis located in each region whilst grey indicates the proportion of all CpGs present in the Infinium MethylationEPIC array. P-values, as determined by Chi-squared analysis are provided above each bar. **D)** Bar plots showing the count of CpGs in the context of their genomic location for T2D. Blue indicates CpGs associated with a change in gene expression in the T2D analysis whilst grey indicates the proportion of all CpGs present in the Infinium MethylationEPIC array. P-values, as determined by Chi-squared analysis.

Among the 300 genes linked to age-associated CpGs, several play important roles in epigenetic regulation, including *DNMT3B* and *TET3*, which were both downregulated with age, potentially indicating a shift in methylation dynamics with age (**Figure 18A**). Furthermore, this analysis highlighted several key genes involved in pancreatic islet stress and insulin secretion, including *OPRD1*, a recently described gene implicated in insulin secretion (Meulebrouck et al. 2024), and *PRDX3*, whose decrease is associated with the inhibition of β-cell apoptosis (Wolf et al. 2010).

For T2D, we identified 552 unique dysregulated genes, which included *HES1*, *NEURO1*, *INSM1*, all of which are crucial to β-cell identity and function (Y. Bar et al. 2012; Jia et al. 2015; Romer et al. 2019; Stancill et al. 2017). (**Figure 18C**). Additionally, genes crucial for β-cell insulin secretion, such as *SLC2A2* (encoding GLUT2), were downregulated in T2D, reinforcing the link between T2D-associated methylation changes and β-cell dysfunction (Sansbury et al. 2012).

To bring mechanistic insight into DNA methylation changes in age, we performed pathway analysis using Metascape (Zhou et al. 2019) for the 271 target genes (**Figure 18A**). We found 20 significantly enriched pathways (and 92 sub-pathways) (**Supplementary table 7**). The most significant pathway was the "Intracellular signaling by second messengers", *i.e.,* of which the PI3K/AKT pathway was the most enriched sub-pathway. PI3K/AKT activation is necessary for the maintenance of β-cell mass and function, notably insulin secretion (Huang et al. 2018). Additionally, we found that downstream pathways of this PI3K/AKT activity pathways were also significant, including "regulation of cell growth" and "regulation of intrinsic apoptotic signalling pathway". We also observed an enrichment of the TBC/RABGAP pathway, a key contributor to insulin secretion in response to glucose uptake in pancreatic β-cells (C. Lu et al. 2022). Finally, "regulation of cellular response to stress", a key driver of age-associated β-cell decline (Aguayo-Mazzucato 2020). Together, this highlights an epigenetically mediated dysregulation of pathways crucial to β-cell function and insulin secretion with age. This suggests that age-associated epigenetic changes may lead to a decline in these key regulatory processes (**Figure 18C**).

In T2D, the most dysregulated pathways included pathways relevant to pancreatic β-cell function and insulin secretion, including "regulation of protein-containing complex assembly", "regulation of cell growth" and "microtubule cytoskeleton organisation", and "regulation of cellular component size". (**Figure 18; Supplementary table 8**).

Notably, only three enriched pathways overlapped between age and T2D: “regulation of protein-containing complex assembly”, “regulation of cell growth”, and “neutral lipid catabolic process”, highlighting the distinct consequences of age and T2D-mediated DNA methylation alterations. Building on this, we identified only 23 genes shared between these two analyses, which were under the regulation of distinct CpGs. These included key T2D genes, including *MEG3,* a long non-coding RNA which is involved in insulin production and β-cell apoptosis (You et al. 2016), and *CDKN1A,* which is involved in β-cell proliferation (E. E. Xu et al. 2017).

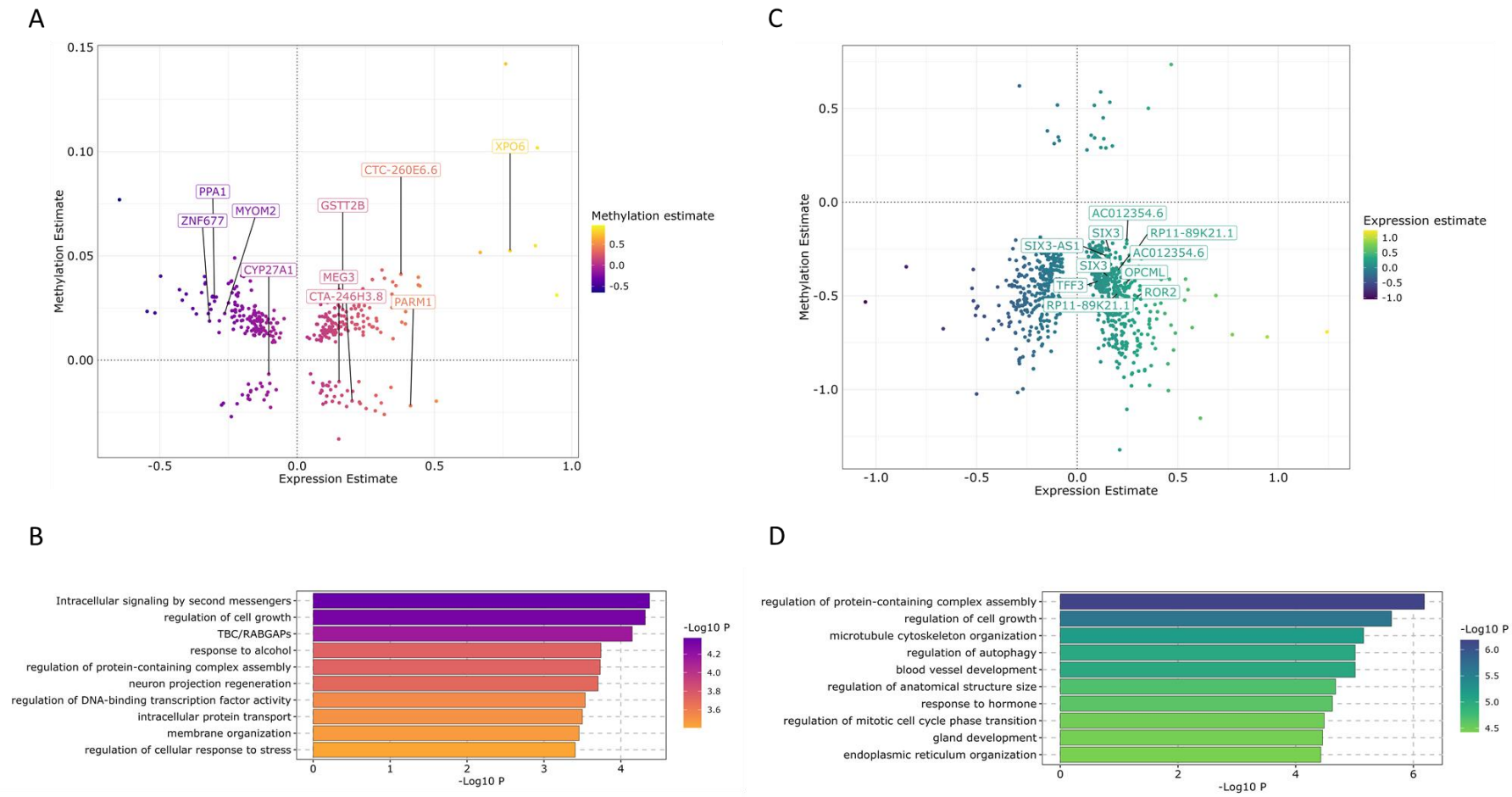


Figure 18: Analysis of target genes. **A)** Scatterplot showing the relationship between CpG methylation (age) and gene expression. Each point indicates a significant CpG-gene association. The top 10 most significant CpG-gene associations are labelled. Colours represent the methylation estimate. **B)** Pathway analysis using Metascape. The 10 most significantly dysregulated pathways are shown. The x-axis indicates the -Log10 p-value. Colours represent the p-value gradient. **C)** Scatterplot showing the relationship between CpG methylation (T2D) and gene expression. **D)** Pathway analysis for T2D-associated target genes using Metascape.

Overall, our findings for both T2D and age are consistent, demonstrating that CpGs impacted by environmental factors tend to be independent of nearby genetic influence. Furthermore, the target genes are also largely distinct, suggesting that genetics and epigenetics operate distinctly one from another (**Figure 19**)

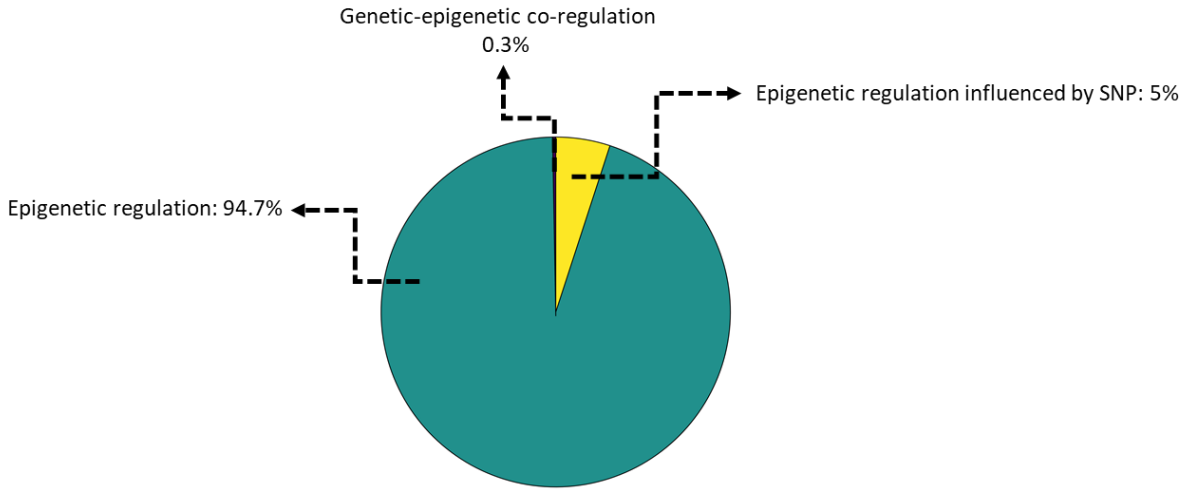


Figure 19: Summary of genes under the control of CpGs (94.7%), CpGs influenced by SNPs (5%) and CpGs and SNPs independently (0.3%). Pie chart pools the genes identified in the age and T2D analyses.

8.3.5 Linking epigentically regulated genes to pancreatic islet function

To explore the functional relevance of our findings, we correlated the identified genes with the insulin secretory index (ISI) calculated from our human donor islets, an *in vitro* measure that reflects the functional capacity of pancreatic islets to secrete insulin under both basal and glucose challenge condition. Pancreatic islets of individuals with T2D in this study exhibited weaker ISI values ($P < 0.001$; **Table 2**). We found 38 age-associated genes (**Supplementary Table 9**) significantly correlated with ISI, including *OPRD1*, a gene involved in insulin secretion, and *CYP27A1*, a key regulator of cholesterol metabolism (**Figure 20A**; Escher et al., 2003; Meulebrouck et al., 2024). For T2D, we identified 52 genes impacting ISI (**Supplementary Table 10**), several of which have previously been tied to insulin secretion, including *SYT7* and *SLC2A2* (**Figure 20B**) (Dolai et al. 2016; Sansbury et al. 2012). Interestingly, five of the genes associated with ISI were solute carriers, *SLC2A2*, *SLC5A4*, *SLC7A4*, *SLC6A6*, and *SLC8A1*. Whilst only *SLC2A2* (GLUT2) has an important role in β-cells, solute carriers are central to pancreatic islet function and these were all highly expressed in pancreatic islets (Alonso et al. 2021; Le et al. 2024). Finally, ST6GAL1, which we identified as a triad and colocalised gene was associated with ISI, providing a mechanistic explanation for its association with T2D risk variants. Generally, this analysis highlighted key players to insulin secretion, further confirming that the methylation changes we observed are functionally linked to β-cell failure in T2D (**Figure 20B**).

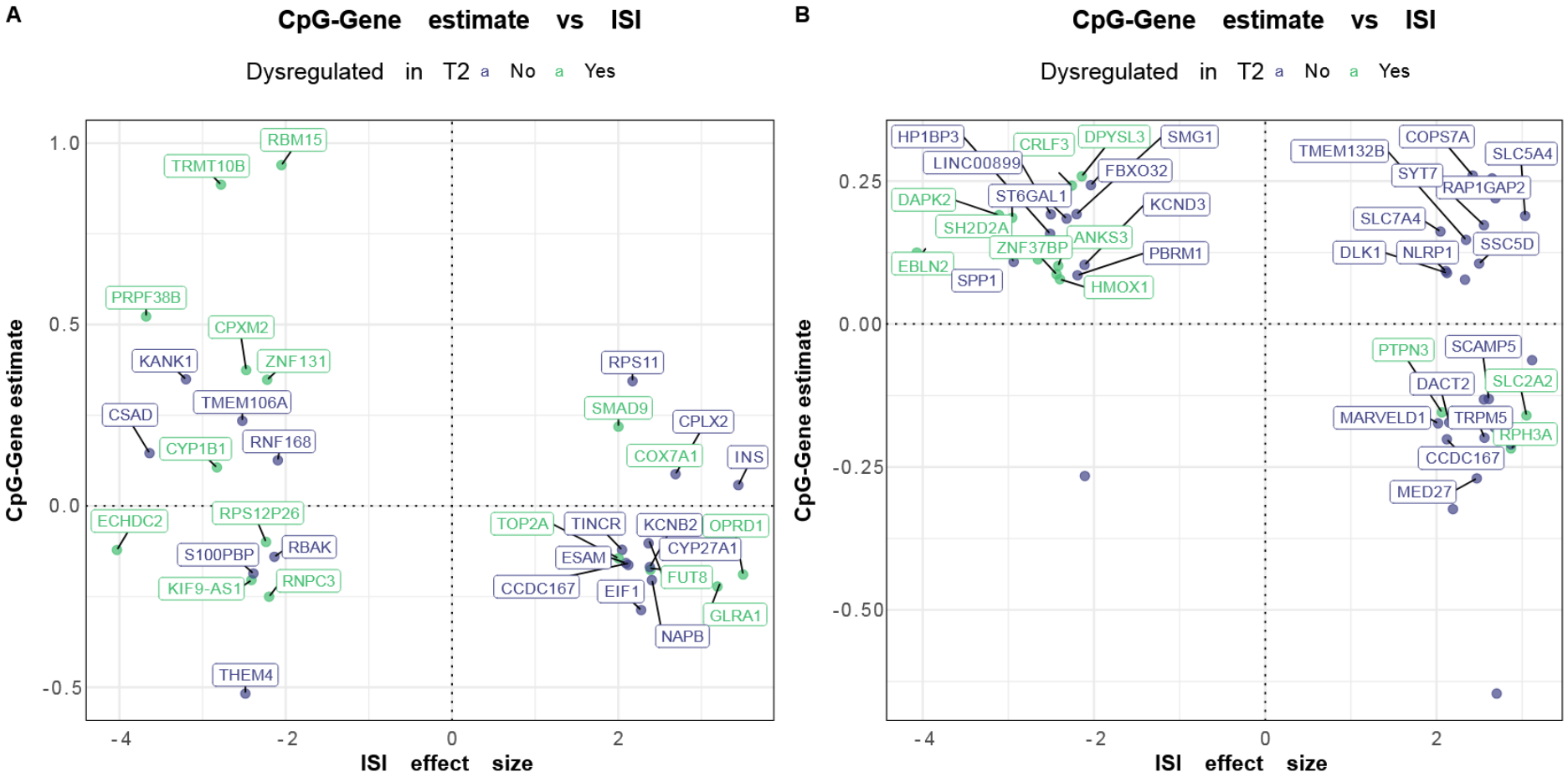


Figure 20: Epigenetic target genes that are also dysregulated with insulin secretion index (ISI) for the same islets used in this study. A-B) Scatterplot of ISI-gene expression against CpG-gene expression for age and T2D, respectively. The x-axis shows the gene expression association with ISI, and the y-axis shows the methylation change estimate with age. Genes labelled in green are also dysregulated with T2D in the same cohort.

8.3.6 Methylation risk scores predict diabetes status and enhance polygenic scores

Given our observation that epigenetic changes associated with age and T2D are largely independent of genetic variation, we hypothesised that integrating DNA methylation could improve the predictive power of existing genetic models. To test this, we developed methylation risk scores (MRS) using age-associated CpG sites identified within our cohort. We focused on CpG sites validated in two independent settings: 1) a pancreatic islet cohort of 87 individuals (Bacos et al. 2016) and 2) four blood-based cohorts (Florath et al. 2014; Reynolds et al. 2014; McCartney et al. 2019; Tajuddin et al. 2019).

CpG sites were included in the MRS were based on the following criteria: 1) localised within a CpG island, 2) positioned in regulatory regions (promoter or enhancer), 3) significant in our EWAS for age, and 4) showing a methylation estimate with age greater than 0.2. From this, we identified 11 CpGs (**Supplementary Table 11**). First, since DNA methylation is known to correlate with age, and accelerated ageing has been linked to various diseases (Horvath and Topol 2024), we confirmed two points: 1) our MRS could accurately predict chronological age (estimate = 149.90, $P = 2.2 \times 10^{-16}$ **Figure 21A**), and 2) individuals with T2D displayed accelerated ageing ($P = 1.505 \times 10^{-6}$; **Figure 21B**). Next, we evaluated the predictive power of the MRS for T2D status using receiver operating characteristic (ROC) curve analysis, yielding an AUC of 0.769, comparable to the performance of recently developed PGS for T2D (Ge et al., 2022; **Figure 21C**). Moreover, our MRS was significantly associated with clinical and metabolic outcomes, including ISI (estimate = -5.26, $P = 9.34 \times 10^{-6}$; **Figure 21D**) and mean glycemia (estimate = 152.04, P = 0.03; **Figure 21E**). These results validate the broader applicability of our MRS.

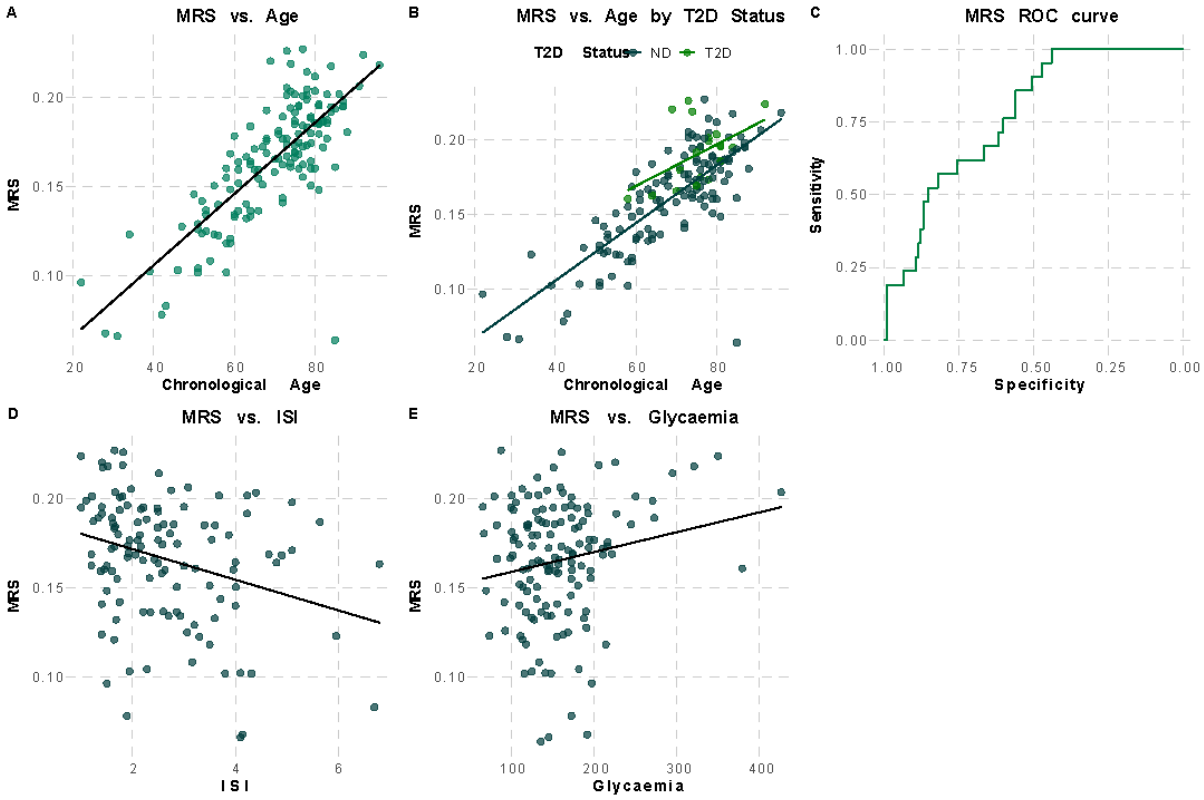


Figure 21: Predictive performance of methylation risk score (MRS), polygenic risk scores (PGS). A) ROC curve describing the predictive capabilities of MRS for T2D in the Epi-islets cohort. **B)** Scatter plot showing the relationship between chronological age and MRS, with each point representing an individual. **C)** Scatter plot showing the relationship between chronological age and MRS split by diabetes status. Purple dots indicate individuals with T2D, while black dots indicate non-diabetic individuals. Solid lines represent linear models fitted between chronological age and MRS for each group. **D-E)** Scatter plots displaying the relationships between insulin secretion index (ISI) **D)** and mean glycaemia **E**) against MRS. The solid lines in each panel represent linear models fitted between MRS and either ISI or mean glycaemia.

To explore whether integrating MRS with genetic data could improve T2D prediction, we generated a PGS using the most recent GWAS data for T2D (Suzuki et al., 2024). Our PGS achieved an AUC of 0.76 (**Figure 22**). When integrating PGS with basic clinical variables (age, sex, BMI), the model's performance improved to an AUC of 0.87. When we combined the MRS with PGS, we observed a significant additive effect, with the combined score reaching an AUC of 0.88. The highest predictive accuracy (AUC = 0.90) was achieved by combining MRS, PGS, and clinical variables (**Figure 22**).

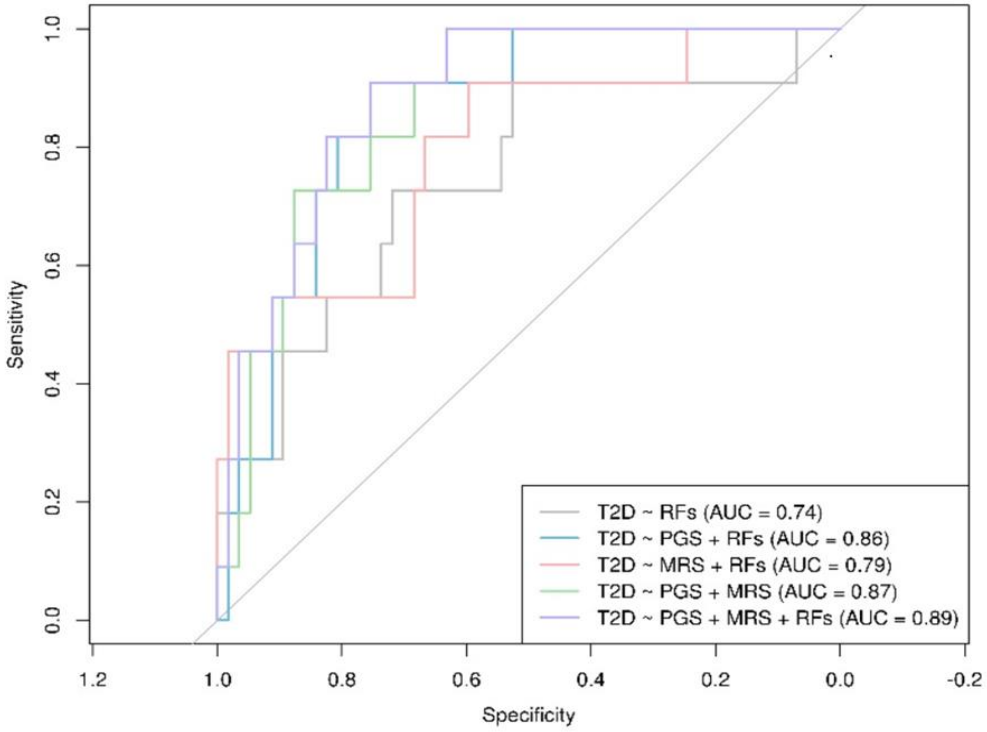


Figure 22: Predictive performance of methylation risk score (MRS) and polygenic risk scores (PGS). ROC curves comparing the predictive capabilities of MRS, PGS, risk factors (RFs), and their combinations T2D. The area under the curve (AUC) values for each model are shown in the bottom right corner of the panel.

8.4 Discussion

In this work, we integrated DNA methylation, transcriptomics, and genotyping to assess the contribution of genetic variants to environmentally-mediated DNA methylation changes in pancreatic islets. Surprisingly, only a fraction of age or T2D associated CpGs were under the influence of a nearby genetic variant. Furthermore, of the genes targeted by DNA methylation changes, only three were shared with colocalised signals for T2D. This suggests that a full understanding of disease and physiological processes requires the integrated study of both genetics and epigenetics. Indeed, we confirmed this with the observation that combining MRS with PGS enhances predictive capabilities for T2D, highlighting the utility of DNA methylation in improving genetics-based disease risk prediction. Finally, characterisation of genes under CpG regulation detected an enrichment in key pancreatic islet pathways, providing insights into the contribution of environmental variables to the mechanisms underlying T2D and pancreatic islet function. . Consequently, we propose that epigenetic profiling may serve as a powerful complementary approach to genetic studies for identifying novel candidate genes.

8.5 The contribution of genetics to epigenetically-mediated changes

Here, we assessed the extent to which environmentally mediated changes in DNA methylation which impact gene expression were under genetic control. Strikingly, only a fraction of environment-associated CpGs which impacted gene expression were associated with nearby SNPs, indicating that environment-mediated DNA methylation changes are acting independently of genetic alterations. While GWAS, eQTL, and colocalization have successfully identified a number of T2D *loci* (Alonso et al. 2021; Mandla et al. 2024), our findings suggest that epigenetic alterations, which impact the expression of distinct genes, may uncover targets that remain undetected through conventional genetic approaches. This presents a unique opportunity to characterise traits of interest using DNA methylation-based studies. In contrast to our

findings, recent work found that CpGs influencing gene expression are predominantly enriched for mQTLs, indicating strong genetic regulation (Stefansson et al. 2024). However, these correlations were assessed in a phenotype agnostic manner, which may overlook environment-specific interactions with DNA methylation. By focusing on age or T2D, our study captures these environmental influences. Additionally, it is expected that a region containing a CpG and a SNP will contribute to changes in gene expression more often than CpGs alone. Notably, despite the observed depletion of epigenetic and genetic interactions, the identified triad genes *i.e., SIX3*, *ST6GAL1*, and *TIPIN* are of interest in the context of T2D. Indeed, *SIX3* is an already well established T2D candidate gene, whose expression is key to insulin secretion (Alonso et al. 2021; Bevacqua, Dai, et al. 2021). *ST6GAL1* s implicated in N-glycosylation and has been previously been colocalised with T2D GWAS SNPs in an independent pancreatic islet cohort (Alonso et al. 2021; Bevacqua, Lam, et al. 2021; Rudman, Gornik, and Lauc 2019). N-glycosylation is associated with T2D, and *ST6GAL1* expression in β-cells is also associated with T1D risk, suggesting a likely role for the gene in T2D given its correlation with ISI in our cohort (Rudman et al. 2023). Finally, *TIPIN* is unknown to diabetes, however it is linked to both the circadian rhythm and cell cycle replication (Gotter, Suppa, and Emanuel 2007), and the circadian rhythm is associated to all facets of T2D (Parameswaran and Ray 2022).

8.6 DNA methylation and gene expression

DNA methylation alterations are well-established biomarkers for age and age-related diseases (Salameh, Bejaoui, and El Hajj 2020). However, it remains difficult to 1) discern which CpGs are actively influencing gene expression. To address this, we comprehensively integrated DNA methylation to gene expression. In addition, we investigated age, which acts upstream of our disease of interest, T2D, to identify CpGs and genes which could promote disease incidence.

Through this approach, we identified numerous differentially methylated age-associated CpGs associated with gene expression involved in: 1) β-cell proliferation, 2) β-cell apoptosis, and 3) secondary messenger signaling, particularly PI3K-pAKT signaling—all pathways crucial to pancreatic islet function and T2D pathophysiology (Galicia-Garcia et al. 2020). Disruptions in these enriched pathways suggest a plausible mechanistic link to disease, indicating that at least some DNA methylation changes may contribute to T2D initiation or early β-cell dysfunction, which is characteristic of ageing (Tudurí et al. 2022). A weakness of our approach is that our model generates mere associations which do not provide certainty whether the DNA methylation is a cause or consequence of the transcriptomic alteration. However, our model did detect several genes which have previously been implicated in T2D, such as *CDKN1A*, *MEG3*, *OPRD1*, and *INS*, improving our confidence in the biological value of our analysis (You et al. 2016; Meulebrouck et al. 2024; Muhammad et al. 2021).

On this basis, we propose that age-associated DNA methylation contribute to age-associated physiological decline, leading to T2D. Indeed, by assessing the age-associated target genes with ISI, we identified several deleterious associations. For instance, we found age was associated with decreased *KCNB2*, and *GLRA1* expresion. Both these genes positively associated with ISI in our cohort, and have previously been shown to promote insulin secretion, yet are downregulated by age (Fu et al. 2017; C. Ling 2020). Similarly, *KANK1*, which negatively correlated with ISI and is known to impact insulin secretion was upregulated with age (Yin et al. 2024). Conversely, we found *INS* to be upregulated by age and to be positively associated with ISI, demonstrating a protective effect of age, which has previously been suggested by Bacos et al. (2016). In accordance with this, certain DNA methylation alterations are thought to confer selective advantages to cancerous cells, promoting their survival and overall tumour progression (Loukas et al., 2023). Certain age-associated DNA methylation changes may therefore be protective, which is consistent with previous propositions of DNA methylation being a mechanism enabling environmental adaptation (Flores, Wolschin, and Amdam 2013).

In our target gene analysis for T2D-associated CpGs, we identified genes distinct from the age analysis. However, we again found pathways implicated in the pathophysiology of T2D, such as regulation of autophagy, blood vessel development and regulation of anatomical structure size (Brissova et al. 2015; Watada and Fujitani 2015). Importantly, these pathways are generally observed following T2D onset in pancreatic islets. For instance, in T2D, islets can temporarily feature increased size to enhance insulin secretion capabilities (Watada and Fujitani 2015). In conjunction, islets feature defective autophagy (in part resulting from increased B-cell activity promoted by increased β-cell size), which contributes to β-cell dysfunction and death (Watada & Fujitani, 2015). Finally, T2D islets feature greater vascularization, which is correlated to amyloid plaque deposition, a direct consequence of T2D exposure in pancreatic islets (Brissova et al., 2015). More importantly, key regulators of pancreatic islet cellular identity, *ie., HES1*, *INSM1*, *NEUROD1*, and *SLC2A2*, all implicated in the pathway gland development were downregulated (Romer et al. 2019; Jia et al. 2015; Stancill et al. 2017; Secco et al. 2022). Demonstrably, T2D exposure appears to, at least, alter DNA methylation patterns in a manner which is correlated to further aggravations of T2D development in pancreatic islets. Corroborating this, we found a decreased expression of *SLC2A2* and *ENO1*, both of which positively correlated with ISI in our cohort and previously validated to contribute to insulin secretion (Luo et al. 2024; Sansbury et al. 2012). Interestingly, we found five solute transporters epigenetically disrupted by T2D and associated with ISI. Solute transporters are fundamental to β-cell nutrient sensing and insulin secretion, and could be one of the levers through which these cells adapt to changing metabolic conditions (Schumann et al. 2020).

8.7 DNA methylation patterns in age and T2D

We found that age is associated with an overwhelming hypermethylation of CpG sites in pancreatic islets, predominantly located in CpG islands and regulatory elements,

especially promoters. Consistent with previous findings by Bacos et al. (2016) our results indicate a strong correlation between age and DNA hypermethylation in these regions. Furthermore, we replicated over 60% of the age-associated CpG sites identified in this earlier study (Bacos et al. 2016). These trends are consistent with age-associated DNA methylation changes observed in other tissues such as blood, liver, and kidney, suggesting a precise and systemic pattern of DNA methylation modification with age (Bysani et al. 2017; Jansen et al. 2019; Heylen et al. 2019).

In contrast to age, T2D was associated with a global hypomethylation. Hypomethylations are generally the result of deficiencies in DNA methylation machinery, largely induced by cellular stress (Jeltsch and Jurkowska 2014). Given that cellular stress is a hallmark of T2D in pancreatic islets, T2D may induce a state of cellular chaos, disrupting the normal regulatory processes that maintain cell function. (Alonso et al. 2021; Smith, Hetzel, and Meissner 2024). This state could lead to hypomethylation changes large enough to be detected, despite their stochastic nature and our limited sample size (Jeltsch and Jurkowska 2014). The stochastic nature of these hypomethylations may also explain why only 1% of our T2D-associated CpG sites were replicated in previously published work on pancreatic islets (Rönn et al. 2023). Consequently, T2D-associated DNA methylation patterns in the endocrine pancreas likely cannot currently serve as prognostic biomarkers for disease initiation or progression in the same manner as age-associated DNA methylation changes. An alternative, but likely synergistic explanation, is that T2D is a highly heterogenous disease, and the lack of consistent methylation patterns may be the result of different cohorts featuring distinct T2D subtypes (Ahlqvist, Prasad, and Groop 2020). Indeed, subtypes of colorectal cancer feature distinct methylation patterns (Weisenberger, Liang, and Lenz 2018). Hence, it may be of value to assess the methylome of the currently described subtypes of T2D to 1) generate markers and 2) assess the biological contribution of said marker to the pathophysiology of each subtype. This would allow for rapid identification of subtype as well as characterisation of the molecular alterations, potentially identifying a methylation

signature most representative to each subtype's unique pathophysiology. Indeed, a recent study analysed the methylome of T2D subtypes and identified distinct methylation patterns for each subtype (Schrader et al. 2022). A complete analysis of these subtypes could provide more replicable and better understanding into the pathophysiology of T2D.

8.8 Methylation-based risk scores

We confirmed the ability of DNA methylation to predict T2D by generating a MRS based on robust age-associated changes in DNA methylation and demonstrated its ability to predict T2D status. Importantly, integration of the MRS to PGS improved prediction in an additive manner, surpassing the performance of PGS and traditional risk factors alone. This indicates that 1) DNA methylation provides additive predictive value beyond conventional PGS approaches, and 2) DNA methylation captures environmental variables beyond age, sex, and BMI. Indeed, we observed correlations between our MRS and ISI and glycaemia in our cohort. Consistent with this, clinical variables have successfully been imputed from DNA methylation data (Kalyakulina et al. 2022). Furthermore, our score supports previous findings linking disease states with older biological age (Bell et al. 2019). It would be of interest to validate our MRS score in age in other cohorts, as this could enable earlier detection of T2D and inform preventative strategies.

8.9 Limitations

This work features several limitations which are important to consider. First, whilst we did integrate ISI in our analysis, our work remains limited to statistical associations. Further functional characterisation is necessary to validate the identified genes and their mechanistic roles. This would provide a deeper understanding of the biological processes involved and confirm the causal relationships suggested by our findings. Indeed, we are currently functionally validating four genes of our identified genes, *SACM1L*, *FBXO27*, *STK38L*, and *USP4*.

These genes have not been previously linked to T2D and were epigenetically dysregulated in both the age and T2D analysis. Furthermore, whilst we did improve on previous studies, our sample size remains limited. This may have contributed to the weak replication of T2D-associated DNA methylation changes and the absence of age-related hypomethylation, a well-documented phenomenon in other tissues (Salameh, Bejaoui, and El Hajj 2020). Additionally, clinical variables for the cohort were limited to age, sex, BMI, and ISI, restricting our ability to link alterations to T2D-relevant variables such as HOMA2B or HOMA-IR. Although we had access to mean glycaemia, it is an imperfect measurement in hospital settings for organ donors, hindering our ability to link methylation alterations to more precise T2D-relevant indicators.

8.9.1 Conclusion

Here, we conducted an extensive DNA methylation and genotyping analysis in pancreatic islets to assess the relationship between epigenetics and genetics and to identify novel genes implicated in disease, particularly in the context of ageing and T2D. This work highlights the potential of integrating epigenetic and genetic data to uncover genes and regulatory mechanisms that may not be captured by traditional genetic approaches alone. Our findings offer valuable insights into the distinct and shared pathways between ageing and T2D, emphasizing the role of environmental factors in shaping epigenetic changes. These insights contribute to a deeper understanding of disease progression and open new avenues for identifying biomarkers and therapeutic targets. Ultimately, we propose a model whereby genes are regulated either by genetic or epigenetic factors, but rarely both.

9 Project 2: DNA methylation investigation of the exocrine pancreas identifies *PNLIPRP1* as a link between type 2 diabetes cholesterol metabolism, and precancerous states

9.1 Introduction

The pancreas is viewed as two distinct organs because of the different functions of the endocrine and exocrine compartments. The endocrine pancreas secretes a variety of hormones, most notably insulin, which is central to T2D pathophysiology. On the other hand, the exocrine pancreas primarily consists of acinar cells that secrete digestive enzymes, responsible for breaking down large molecules (Pandol, 2011). Consequently, the majority of research on T2D focuses on the endocrine pancreas, whereas the exocrine pancreas, which represents 95% to 98% of the pancreas, has remained largely overlooked in the study of T2D (Pandol, 2011). However, both compartments are interdependent. Indeed, the exocrine pancreas provides structure, blood flow and contributes to the islet microenvironment (Pandol, 2011). Furthermore, a recent study demonstrated that acinar cells located in close proximity to pancreatic islets are enlarged and may contribute to islet expansion to enhance insulin secretion by releasing trypsin to break down the extra-cellular matrix that surrounds islets (Egozi et al., 2020). In the other direction, insulin is a critical regulator of exocrine homeostasis and, indeed, a lack of trophic insulin induces exocrine atrophy, whilst an excess promotes fibrosis (Czakó et al., 2009). Furthermore, epidemiological studies show a strong physiological link between the two compartments of the pancreas. Indeed, endocrine disease (T2D) causally predisposes individuals to developing PDAC, a disease of the exocrine (Yuan et al. 2020; Maina et al. 2023). Conversely, PDAC is associated with new-onset T2D (Yan Li et al. 2019). Recent evidence showed that T2D duration is associated with both ADM and PDAC in the exocrine pancreas (Wright et al. 2024). ADM is an important

initiating step in PDAC development (Marstrand-Daucé et al. 2023). Understanding the underlying mechanisms which drive T2D to promote ADM and PDAC is essential to the prevention, early detection, and treatment of this hard to detect and deadly disease. We hypothesised that T2D (and age) exposure triggers molecular alterations which could be detected by assessing the methylome of the exocrine pancreas. This project was published and is available in print format the bottom of this document.

9.2 Materials and methods

9.2.1 DNA methylation measurement of human samples and epigenome-wide association study

We obtained 155 pancreas samples, of which 32 had T2D according to the American Diabetes Association (ADA) guideline (ADA, 2019), from the IMIDIA consortium (Solimena et al. 2018b). Next of kin's consent was obtained for pancreas sample collection along with approval from the ethics committees in both Pisa and Hannover.

To extract DNA, we utilised the NucleoSpin Tissue kit (T740952.50; Mackerey-Nagel). We performed bisulphite conversion with 800 ng of DNA using the EZ DNA Methylation kit (5001; Zymo Research). Bisulphite converted DNA was applied to Illumina's Infinium methylationEPIC array, which covers a total of 930,000 CpGs. The minfi R package was used to import the methylationEPIC array data into Rstudio (Aryee et al. 2014). We performed quality control (QC) of array data by removing CpG probes which met the following conditions: located in sex chromosomes or SNPs, which cross-hybridised, were non-cg probes, or featured a detection threshold p-value < 0.01. Additionally, any sample with a probe detection threshold lower than 99% was removed (probe detection p-value = $p < 0.01$). Finally, two samples with discordant sex were also discarded. To correct for probe-design biases as well as batch effect, we utilised the packages Enmix and SVA respectively (Z. Xu et al. 2016; Leek et al. 2020). Post-QC, 141 samples and 746,912 CpGs were retained for further analysis (**Table 8**). We assessed population structure of our cohort with principal component analysis (PCA) using the 1,000 genomes reference panel (**Supplementary figure 6**).

To perform EWAS, linear regression models were applied to associate age or T2D with CpG methylation level at a given probe. Whole pancreatic tissue samples were expected to include a variety of cell types and thereby be a potential confounding effect on DNA methylation. Consequently, cell composition was estimated using the R package RefFreeEWAS and included in the linear regression to address this

confounder (**Supplementary figure 6**) (Houseman et al., 2016). Multiple testing was corrected using the Bonferroni method and CpGs with an p-value < 0.05 were considered significant. The EWAS models are described below:

$$CpG\ (M\ value) = Age \sim Sex + T2D + BMI + Cellular\ composition$$

$$CpG\ (M\ value) = T2D \sim Sex + Age + BMI + Cellular\ composition$$

9.2.2 Comparing the DNA methylation profiles of the exocrine and endocrines pancreas:

To ensure our whole pancreas samples (98% exocrine) were distinct from pancreatic islets at the DNA methylation level, we compared our PCA data with pancreatic islet data that was obtained from the IMIDIA cohort, utilising a method identical to the one described above. Pancreatic islet samples were collected from 144 organ donors, ranging in age from 22 to 96 years (average age = 69 years). No significant differences were observed in terms of clinical characteristics, notably, age, sex, BMI, or type 2 diabetes (T2D) status between the two cohorts (**Supplementary Table 12**). PCA of methylation data from both exocrine and pancreatic islets was conducted using beta values from each group, with the analysis performed using the flashPCA package in R (**Supplementary figure 7**) (Abraham, Qiu, and Inouye 2017).

9.2.3 Differentially methylated regions (DMR) and enhancers

The dmrcate package was used to identify differentially methylated regions (DMRs) (Peters et al., 2015). A CpG with p-value < 0.05 was considered significant and we defined a DMR as a region with at least 2 CpGs within the Gaussian kernel bandwidth (lambda) equals to 1 kb. To determine if specific CpGs and DMRs were situated within regulatory regions, we consulted the Genehancer (Fishilevich et al. 2017) and dbSUPER (Khan and Zhang 2016) databases.

9.2.4 UK Biobank to identify rare variant associations in *PNLIPRP1:*

We consulted exome sequencing data from 191,000 UKbiobank participants to detect null variants for a minor allele frequency (MAF) <1% in the *PNLIPRP1* gene (UKBiobank research application #67575). To assess associations between *PNLIPRP1* null variants and metabolic traits including BMI, glucose and lipids, we applied the MiST method (Sun, Zheng, and Hsu 2013), which tests rare variants in a single cluster (at the gene scale). A score π represents the average effect of the cluster, while τ denotes the effect heterogeneity within the cluster. The overall p-value assesses the association between the set of variants and the trait of interest. For each trait, we adjusted for relevant covariates. A trait was considered significant if the p-value for the direct burden effect of the cluster ($\hat{\pi}$ p-value) was less than 0.05, and if the effect direction for the variants was consistent, as indicated by a lack of heterogeneity (tau p-value or P.value.S.tau > 0.05).

9.2.5 GTEx to determine tissue expression of PNLIPRP1

The median gene-level Transcript per million (TPM) from RNA-sequencing data by tissue from Genotype-Tissue Expression (GTEx) Portal (https://gtexportal.org/home/datasets/) was used to identify the expression pattern of *PNLIPRP1*.

9.2.6 Genotyping and Mendelian Randomisation

To assess the direction of causality between traits of interest (*i.e.,* T2D, LDL-C, and CpG methylation), we performed bi-directional two-sample MR. To obtain genetic associations for both T2D and LDL-C, we consulted previously published large scale European GWAS for T2D (Xue, Wu, Zhu, Zhang, Kemper, Zheng, Yengo, Lloyd-Jones, Sidorenko, Wu, eQTLGen Consortium, et al. 2018) and LDL-C (Graham et al. 2021). All MR analyses were performed using the TwoSampleMR (version 0.5.7) R software package.

To obtain genetic associations between SNPs and DNA methylation necessary for our MR analysis, we genotyped 111 control samples from our organ donors (Illumina HumanOmni2.5 arrays) using the Illumina iScan. Genotypes were called using the Genome-studio software and single nucleotide polymorphisms (SNPs) were kept for further analysis according to the following thresholds: 1) minor allele frequency > 0.05, 2) Hardy-Weinberg equilibrium > 1×10^{-4} and 3) call rate > 0.95. Imputation was based on the Haplotype Reference Consortium Panel. Ancestry clustering was performed using the 1000 genome reference panel (**Supplementary figure 8**) (Auton et al. 2015) aFollowing QC, all individuals were suitable for downstream analyses. mQTL analysis was performed using the QTLtools software (Delaneau et al. 2017), adjusting for age, sex, and BMI.

We performed bi-directional MR was performed, first by using T2D or LDL-C as the exposure and CpG methylation as the outcome (referred to as forward MR), then, by using CpG methylation as the exposure and T2D or LDL-C as the outcome (referred to as reverse MR). The Inverse Variance Weighted (IVW) methods was reported as the main MR method. IVW requires the validity of all genetic instrument, or a balanced pleiotropy. To validate this assumption, we conducted other complementary MR methods which uses different assumptions, such as simple median, weighted median (more robust to outliers), MR Egger (sensitive to outliers, its intercept is a test to evaluate horizontal pleiotropy). The F-statistic was used to verify the strength of instrument. Linkage disequilibrium was assessed with the ld_clump() function from the R software package ieugwasr (version 0.1.5). Finally, the leave-one-out analysis was used to verify whether any variant was driving our findings and heterogeneity was assessed using the Cochran's Q test.

Forward instrument design *(T2D/LDL-C to CpG):* We extracted genome-wide significant SNPs from the aforementioned GWAS. For T2D to CpG, we obtained 118 independent signals (LD $r^2 < 0.2$) and 241 independent signals ($r^2 < 0.2$) for LDL-C to CpG. We performed a trans mQTL at SNP-CpG pairs of interest as described in the above section.

Reverse instrument design *(CpG to T2D/LDL-C):* To identify SNPs that act as proxies for *PNLIPRP1* methylation, we performed a cis-mQTL for cg15549216, cg06606475, and cg08580014 with a 50kb window. mQTL signals were FDR-corrected.and SNPs in LD were pruned using the ld_clump() function. SNPs with an r^2 of < 0.75 and an FDR < 0.75 were considered to be viable instruments.

9.2.7 RNA expression of *PNLIPRP1* in organ donors

Samples were stored in OCT blocks at -80°C prior to being utilised. OCT was thawed at room temperature (RT) and the pancreas sample were cut into chuncks and incubated in Trizol (15596-026; Thermofisher), 3% DTT and 5% B-Mercaptoethanol and vortexed. The supernatant containing RNA was collected, the RNA was isolated using Trizol (15596-026; Thermofisher). The isolated RNA underwent reverse transcription using the High-capacity cDNA reverse Transcription Kit (Applied Biosystems; 4368814). qPCRs were performed with the QuantStudio Pro 7 (Applied Biosystems). Amplication was measured using the SYBRgreen reagent mix (A25918; ThermoFisher). Two-tailed t-tests were performed using GraphPad Prism (GraphPad software Inc). Eleven samples were processed (six controls and five T2D individuals), matched for age, sex and BMI.

9.2.8 Functional characterisation in AR42J

General cellular culture protocol: To characterise the role of *PNLIPRP1*, the AR42J acinar cell line was used for all *in vitro* assays (CRL-1492; ATCC). For growth, cells were incubated at 37°C and 5% CO_2 with RPMI 1640 Glutamax medium (61870044; Gibco), supplemented with 10% FBS (26140079; Gibco), 0.01% penicillin/streptomycin (P/S) (15410-122; Life Technologies). Medium was changed every 48 hours.

Diabetogenic treatment: To mimic diabetes exposure, we treated AR42J cells to high glucose (20mM), insulin (100 nM) or both, for 72 hours along with 0.1% FBS

(26140079; Gibco) medium. The cells, 2×10^5 per well, were plated in 6-well plates. Following treatment, RNA from the cells was harvested with Trizol (15596-026; Thermofisher) as described above and was quantified by qPCR as described above.

Akt response and glucose uptake: AR42J cells were plated and serum starved overnight, washed with PBS and stimulated with or without 200 nM insulin for 1 hour. Protein was harvested in RIPA buffer supplemented with protease and phosphatase inhibitors. We used anti pAKT (S473; Cell Signaling) and anti-Akt (9272; Cell Signaling), and the secondary antibody was goat pAb to Rb igG (Ab205718; Abcam). Details of all antibodies utilised in this work are listed in **Supplementary Table 13**. To quantify glucose uptake in AR42J cells, we used the colorimetric glucose detection kit (EIAGLUC; Invitrogen), and followed the protocol provided by the manufacturer.

siRNA knockdown: All transfections were performed in AR42J and in suspension with the AR42J Transfection Reagent kit (1181; Altogen). We used *Pnliprp1* siRNA (M-099515-01-0010; Dharmacon) or non-targeting control siRNA (D-001810-10-20; Dharmacon). Following a 48- or 72-hours period (depending on the experiment or assay performed), RNA and protein were harvested as described above. Extracted protein from *Pnliprp1* knockdown (KD) were quantified with Western blotting. For Western blotting, cells were lysed and protein harvested using a RIPA buffer (89900; ThermoFisher) and protein levels were quantified using the Pierce BCA Protein Assay Kit (23225; ThermoFisher). Protein were separated using a 10% SDS-PAGE gel and transferred to a nitrocellulose membrane using the iBlot2 Gel Transfer Device (Life Technologies). Membranes non-specific sites were blocked with 5% non-fat dry milk or 5% BSA. Incubation was overnight at 4°C. Finally, membranes were incubated with secondary antibody for 1 hour at RT. The detection of the secondary antibody was carried out using the LI-COR Biosciences imaging system and the results were analysed with ImageJ.

RNA Sequencing: Library preparation from *Pnliprp1* KD RNA was performed using the KAPA mRNA HyperPrep Preparation Kit (Roche) and sequenced with Illumina's NovaSeq 6000. The mean sequencing depth was of 100 million 100 bp paired-end reads per sample. Illumina raw data were demultiplexed using bcl2fastq v2.20.0.422 (Illumina) and adapters trimming step was performed using cutadapt (version 3.2) Reads were mapped with STAR version 2.7.1a to the Rattus norvegicus.Rnor6 genome. Raw and normalised read counts were generated using RSEM v1.3.0 with a GTF file from Ensembl version 102, and gene name annotations from Ensembl v102. Differential expression analysis was performed with DeSeq2 in R. Pathway analysis was performed using Metascape or EnrichR.

Immuhistochemistry: Human pancreatic tissue sections, sourced from both control and type 2 diabetes (T2D) individuals, were provided by the Inserm UMR1190 unit at the University of Lille, France. Sections were carefully examined to ensure the absence of fibrotic or pathological changes. Paraffin was removed from the slides using 100% xylene, followed by a series of ethanol washes (100% to 50%) for rehydration. Antigen retrieval was achieved by incubating the slides in sodium citrate buffer (pH 6). To block non-specific binding sites, the slides were treated with a solution of 1x PBS, 0.01% Triton X-100, and 5% goat serum for 30 minutes at RT. Primary antibodies were applied overnight at 4°C, and secondary antibodies were applied for 1 hour at RT. Imaging was performed with a Zeiss LSM 710 NLO confocal laser scanning microscope.

MTS proliferation (197010; Acbam) and Cholesterol Ester-Glo Assay (J3190; Promega) methods were performed as per protocol instructions in AR42J following 72 hours *Pnliprp1* KD.

9.3 Results

9.3.1.1 Age-associated DNA methylation changes in the exocrine pancreas

As age is a major risk factor for pancreatic disease, we first aimed to bring insight into the epigenetic mechanisms of ageing in the exocrine pancreas (Mellenthin et al. 2022). For this purpose, we measured DNA methylation in whole pancreas samples obtained from 141 organ donors (17-89 years, median of 67 years) of European descent, using the Illumina Infinium MethylationEPIC array (**Table 4)** (**Supplementary figure 8**). To validate that the observed whole pancreas epigenetic profile was indeed mainly exocrine tissue, we compared the methylation profile to pancreatic islet methylation profiles (via Infinium MethylationEPIC arrays) from 125 individuals from the same cohort of organ donors, matched for age, sex and body mass index (BMI) (**Supplementary Table 12**). As expected, we found that the methylation profiles of pancreatic islets and whole pancreas were clearly distinct, and this was consistent with four pancreatic islet samples extracted and handled in parallel with the exocrine preparations (**Supplementary figure 7**). Additionally, we found that even pancreatic islet preparations with low islet purity (i.e., percent of islets compared to other pancreatic tissues, such as exocrine), did not overlap with exocrine preparations (**Supplementary figure 7**).

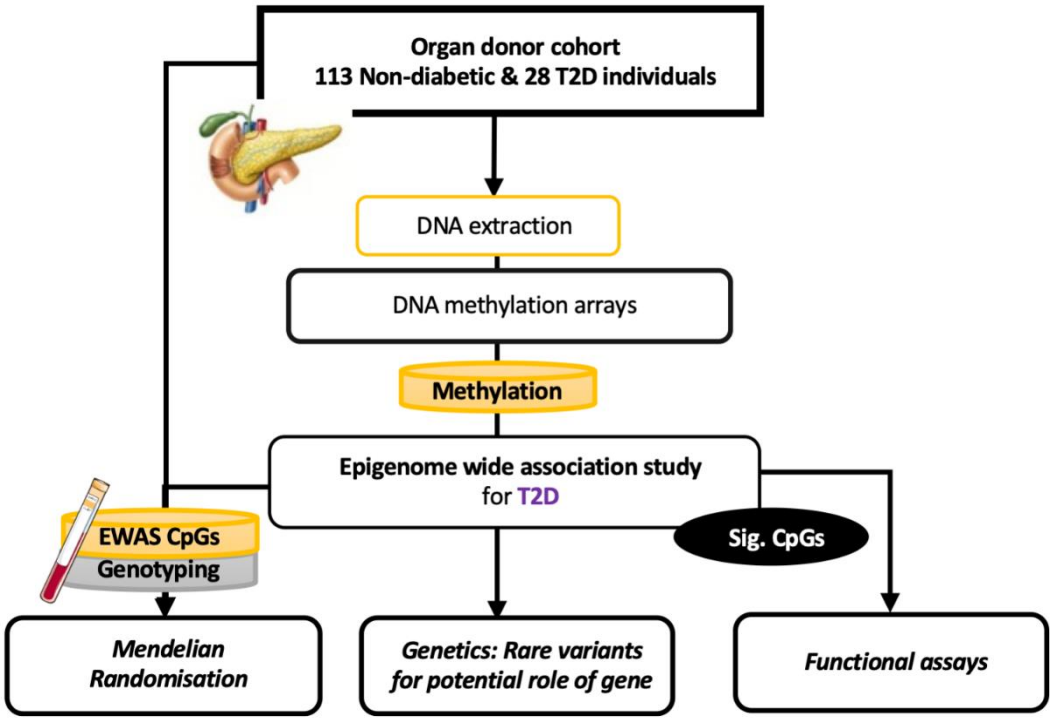


Figure 23: Overview of the exocrine project study design. DNA from **w**hole pancreas tissue was extracted from organ donors and subjected to Illumina's Infinium MethylationEPIC array to perform epigenome wide association studies (EWAS) with age and T2D in the cohort. Subsequently, genetic and functional assays were performed to assess the function of candidate genes.

Table 4: Exocrine organ donor cohort characteristics

Characteristic	**Non-diabetics N = 131**	**T2D N = 28**	**P-value***
Sex			0.016
Female	61 (54%)	8 (29%)	
Male	52 (46%)	20 (71%)	
Age (years)*	67 (53, 76)	74 (69, 79)	0.004
BMI (kg/m^2)*	25 (23.1, 27.1)	26.0 (24.2, 28.2)	0.13

Mean is represented, along with the first and third quartile for the data

We then performed an epigenome-wide association study (EWAS) to determine methylation sites associated with age. We found that a total of 718 CpGs associated with age, of which the vast majority (> 85%) were hypermethylated, which is in accordance with our data in pancreatic islets as well blood and other tissues (Bonferroni-corrected p < 0.05; **Figure 24A; Supplementary Table 14**) (Bysani et al. 2017). Compared with pancreatic islets methylation profile, only 30 % of CpGs (197 / 718) were shared between endocrine pancreas and exocrine pancreas, but they were consistent in direction of effect (**Supplementary Table 15**). Interestingly, the most significant differentially methylated CpGs with age in exocrine tissue were all shared with pancreatic islets, and were previously found in several other tissues, including liver, kidney and blood: cg23606718 (*FAM123*), cg16867657 and cg21572722 (both in *ELOVL2*), and cg06639320 (*FHL2*) (**Figure 24B**) (Hastuti and Beandrade 2022; Slieker et al. 2018; Fulea et al. 2021). To determine the biological age of our samples, and whether T2D induces accelerated ageing in the exocrine pancreas, we utilised the Horvath clock, composed of 353 CpGs (Horvath, 2013). DNA methylation age of pancreas samples correlated with the chronological age of the donors (R2 = 0.55; p = 3.7 × 10-41). Surprisingly, we did not find a difference in DNA methylation age among patients with T2D compared to controls, despite T2D being strongly associated with age (Fazeli, Lee, and Steinhauser 2020) (p = 0.18; **Supplementary figure 9**).

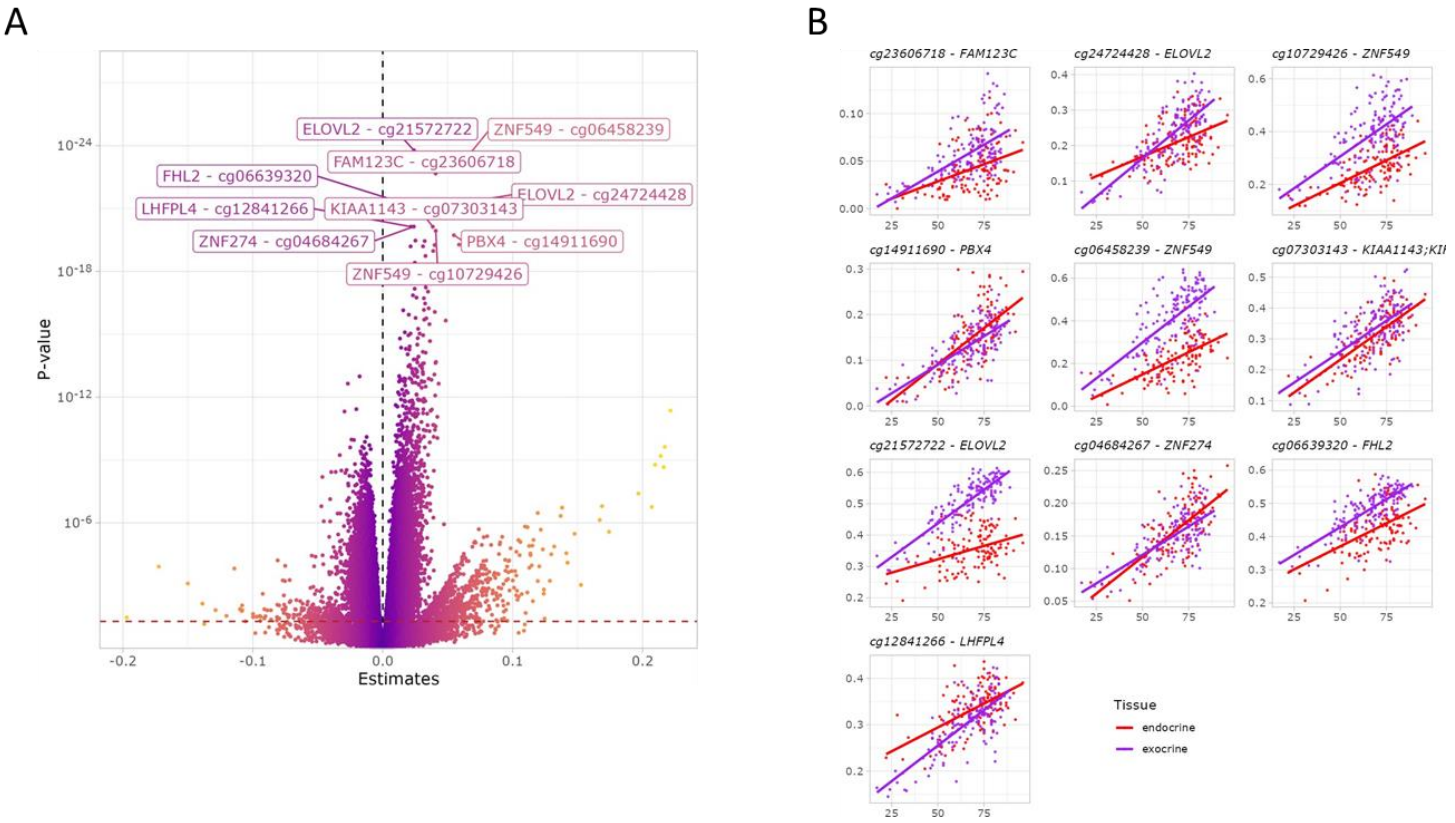


Figure 24: Epigenome-wide association study of whole pancreas samples for age. **A)** Volcano plot depicting the differentially methylated CpGs associatedwith age. The 10 most significant CpGs are labelled, with the CpG probe and the gene nearest to the probe. **B)** Boxplots showing the top 10 differentially methylated CpGs in exocrine (purple) and endocrine (red) pancreas. The X-axis represents the β-value and y-axis the age

9.3.1.2 T2D-associated DNA methylation changes in the exocrine pancreas

Next, we conducted an epigenome-wide association study (EWAS) to identify DNA methylation changes associated with type 2 diabetes (T2D). Our analysis revealed a single CpG site, cg15549216, that showed significant hypermethylation in individuals with T2D after Bonferroni correction (p=0.025; **Figure 25A**). This CpG site is located within the gene body of *PNLIPRP1*, which encodes Pancreatic Lipase Related Protein 1. It was found to be 11.4% more methylated in T2D patients compared to controls (Figure 27B; estimate=0.6; standard error=0.1). Additionally, cg15549216 was positively correlated with glucose levels (p = 1.34 × 10^-4; **Supplementary figure 10**).

We also identified a differentially methylated region (DMR) associated with an increased risk of T2D, which included cg15549216 and two nearby CpG sites, cg06606475 and cg08580014 (**Figure 25B-D**). These flanking CpGs exhibited consistent directional effects: cg06606475, located 921 base pairs (bp) upstream of cg15549216, was hypermethylated by 9.0% (p = 5.9 × 10^-5; estimate=0.38; SE=0.09), while cg08580014, situated 370 bp downstream of cg15549216, showed a 6.3% increase in methylation (p = 4.0 × 10^-4; estimate = 0.27; SE = 0.07). Analysis using the Genehancer and dbSUPER enhancer databases revealed that this region, including and surrounding cg15549216, lies within a 12-kb segment, which is the only identified super-enhancer of the *PNLIPRP1* gene in human whole pancreas tissue (**Figure 25E**) (Fishilevich et al. 2017; Khan and Zhang 2016).

On the basis that age is a significant risk factor for T2D, we investigated whether the cg15549216 CpG site was associated with age in our cohort (Fazeli, Lee, and Steinhauser 2020). We observed a nominal association between age and the hypermethylation of the cg15549216 site in all individuals (unadjusted p = 0.010). Interestingly, this association was no longer significant following adjustment for T2D, indicating an interaction of the disease with the methylation of cg15549216 (unadjusted p = 0.93). Consequently, we tested an interactive model and asked whether both T2D and age contribute to cg15549216 methylation. Our analysis revealed that, at the cg15549216 site, methylation levels increased with age among

T2D individuals compared to non-diabetic controls (M-value estimate = 0.01; p = 4 × 10^-10; **Figure 25F**). We found no association between cg15549216 methylation and BMI (p = 0.54), sex (p = 0.32), T2D duration (p = 0.39), or statin treatment (p = 0.32).

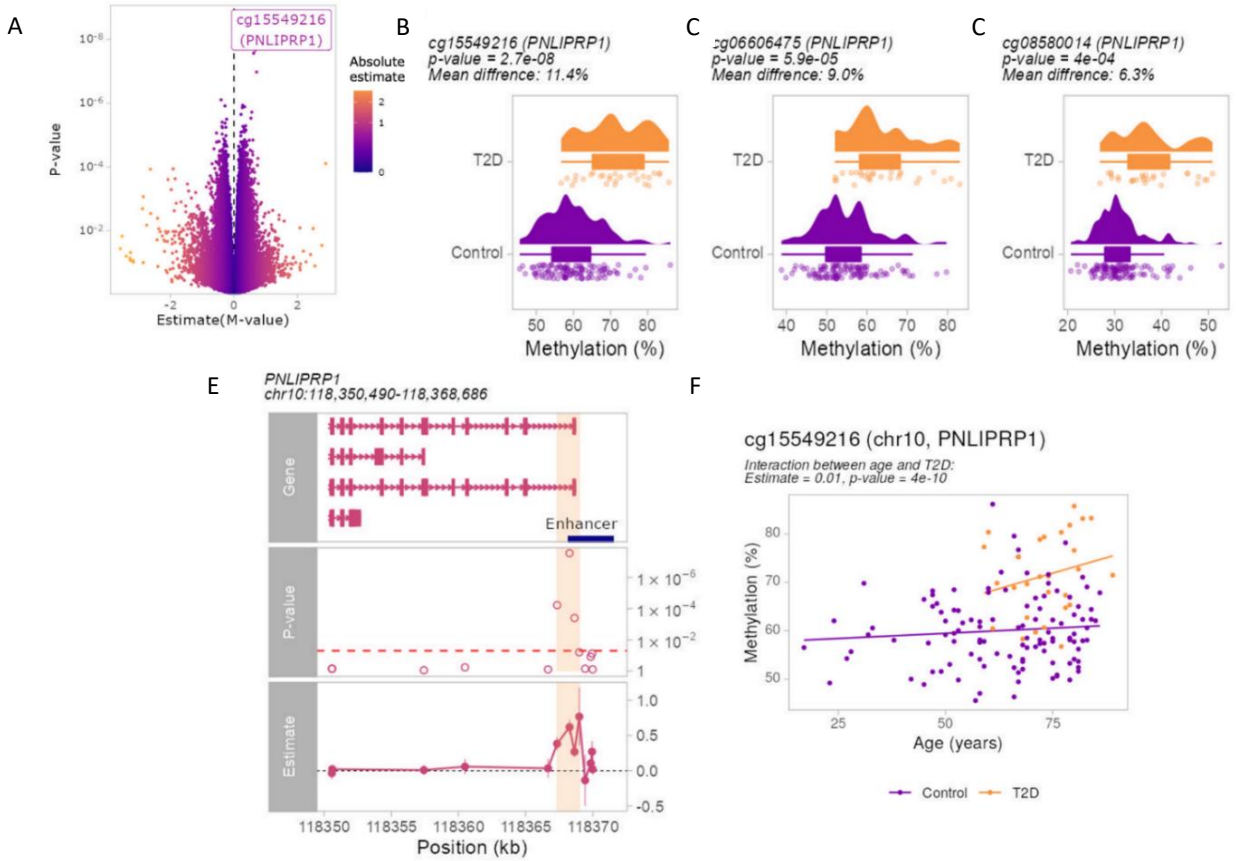


Figure 25: Epigenome-wide association study in whole pancreas tissue for Type 2 Diabetes (T2D). A) A volcano plot illustrating the differentially methylated CpG sites associated with T2D, with the only significant CpG site (Bonferroni correction < 0.05) highlighted. The colour gradient represents the absolute estimate (M-values). **B)** Boxplot showing the distribution of methylation levels at cg15549216 in individuals with T2D (orange) compared to non-diabetic controls (purple), with the x-axis displaying the percentage methylation (beta-values). **C-D)** Boxplots depicting the distribution of methylation levels at cg06606475 and cg08580014, respectively, in both T2D patients and non-diabetic controls. **E)** Visualisation of the PNLIPRP1 gene, highlighting the differentially methylated region (DMR) in orange, and the enhancer region in blue. CpG estimates and p-values for all CpG sites within the PNLIPRP1 gene are also provided. **F)** Scatterplot illustrating the interaction between cg15549216 methylation and age in organ donor individuals, with non-diabetic controls represented in purple and T2D patients in orange.

9.3.1.3 **PNLIPRP1** ***null variants link the gene to metabolic traits***

To explore the function of *PNLIPRP1*, we first examined whether rare loss of function variants (minor allele frequency [MAF] <1%) in the *PNLIPRP1* gene were associated with T2D and relevant metabolic traits, including glucose and lipid levels. This was predicated on the basis that 1) *PNLIPRP1* methylation was associated with T2D, and 2) the exocrine pancreas is central to overall metabolism (Pandol 2011). Here we utilised whole-exome sequencing data from 191,000 participants in the UKBiobank, where we identified a total of 44 null variants (i.e., nonsense, frameshift, canonical ±1 or 2 splice sites) (**Supplementary Table 16**).

Our analysis revealed that *PNLIPRP1* null variants were associated with increased glycaemia ($p\pi = 1.1 \times 10^{-3}$; effect size = 0.13; SE = 0.040; **Table 5**). Additionally, we found significant associations between *PNLIPRP1* rare variants and several metabolic traits: increased LDL-cholesterol levels (LDL-C; $p\pi = 0.034$; effect size=0.10; SE = 0.049), HDL levels ($p\pi=0.026$; effect size=0.05; SE = 0.022), waist-to-hip ratio ($p\pi = 2.7 \times 10^{-3}$; effect size = 0.23; SE = 0.006), waist circumference ($p\pi=7.6 \times 10^{-3}$; effect size = 2.19; SE = 0.820), BMI ($p\pi = 3.1 \times 10^{-3}$; effect size = 0.039; SE = 0.011), diastolic blood pressure ($p\pi = 2.02 \times 10^{-5}$; effect size = 2.89; SE = 0.677), and systolic blood pressure ($p\pi = 2.9 \times 10^{-3}$; effect size = 3.30; SE = 1.112). Interestingly, null variants were not associated with disease manifestations of disrupted glucose and lipid traits, notably T2D risk ($p\pi = 0.48$), obesity ($p\pi = 0.29$), hypertension ($p\pi = 0.94$), or triglyceride levels ($p\pi = 0.46$) (Table 5). Next, we asked if we could reproduce this result in common variants. Using the Type 2 Diabetes Knowledge Portal (https://t2d.hugeamp.org/), we discovered that common SNPs (MAF >1%) within the *PNLIPRP1* locus were strongly associated with increased LDL-C levels ($p = 2.0 \times 10^{-14}$; Supplementary Tables 5 and 6). Furthermore, common variants in *PNLIPRP1* were linked to non-high-density lipoprotein (HDL) cholesterol levels and apolipoprotein B, with significance after multiple testing correction ($p \leq 2.5 \times 10^{-6}$). Interestingly, these same variants were not associated with T2D or glucose-related traits in a cohort of up to 1.61 million participants (**Supplementary figure 11**).

Table 5: Null variants associations with *PNLIPRP1*

Trait	Number of individuals	Pi hat (π)	Standard error	P value (p π)	Number of variants
Diastolic BP*	168,374	2.889	0.677	2.0×10^{-5}	39
WHR**	190,739	0.023	0.006	2.7×10^{-3}	41
log BMI	187,727	0.039	0.011	3.1×10^{-3}	40
Glucose	159,764	0.133	0.041	1.1×10^{-3}	38
Systolic BP*	168,367	3.303	1.1119	2.9×10^{-3}	39
Waist circumference	190,751	2.188	0.820	7.6×10^{-3}	41
HDL	156,077	0.050	0.022	2.6×10^{-2}	38
LDL	169,625	0.104	0.049	3.4×10^{-2}	38
Hba1c	177,038	0.439	0.236	0.06	39
Obesity	107,219	-0.146	0.138	0.29	31
Log Triglyceride	169,815	-0.009	0.011	0.46	39
T2D	171,651	0.177	0.252	0.48	39
Hypertension	180,290	0.010	0.126	0.94	40

*BP = blood pressure

**WHR = waist-hip ratio

9.3.1.4 Causal relationship between PNLIPRP1 methylation and T2D-related traits

To identify whether *PNLIPRP1* methylation was causally associated with T2D and LDL-C, and the reverse, we performed MR. To identify SNPs acting as proxies for *PNLIPRP1* methylation, we genotyped 111 control individuals in our cohort and performed mQTLs. For T2D, we used 118 proxy SNPs associated with increased T2D risk and 6 proxy SNPs for associated with *PNLIPRP1* rmethylation (**Supplementary Tables 17 and 18**). Using the IVW method, we found evidence that increased T2D risk was causal to cg15549216 hypermethylation with an estimate of 0.23 (95% CI = 0.029-0.43; p = 0.025; **Figure 26A**). However, we did not find a significant association using MR methods following different assumptions, such as the simple median and weighted median methods, possibly pointing to the presence of horizontal pleiotropy or heterogeneity in our genetic instruments. To ensure the validity of our result, we performed another robust MR method, MR-Egger, which showed no evidence of causal association (estimate = 0.102; 95% CI = -0.353-0.557; p = 0.659), and ruled out the possibility of horizontal pleiotropy (intercept = 0.011; 95% CI = -0.024-0.045; p = 0.55). In the reverse direction, we found no evidence of a causal effect of *PNLIPRP1* methylation (**Figure 26A**).

For LDL-C, we used 9 proxy SNPs to represent *PNLIPRP1* methylation and 241 proxy SNPs for increased LDL-C levels (**Supplementary Table 19 and 20)**. We found no evidence of a causal effect of LDL-C on *PNLIPRP1* methylation, however *PNLIPRP1* methylation was associated with increased LDL-C levels, with an estimate of 0.064 using the IVW method (95% CI=0.024-0.105; *p*=0.0019) (**Figure 26B**). This result was consistent with the simple median (estimate=0.071; 95% CI=0.043-0.099; $p=5.9 \times 10^{-7}$) and the weighted median methods (estimate=0.068; 95% CI=0.044-0.091; $p=2.2 \times 10^{-8}$). The MR-Egger intercept showed no evidence of horizontal pleiotropy (estimate=0.000; 95% CI=-0.012-0.012; *p*=0.31; **Figure 26B**). In all significant MR analyses, the leave-one-out analysis found that no single SNP altered the results, suggesting that the observed association was robust (**Supplementary Tables 21 and 22**). Altogether, our data provide some evidence that *PNLIPRP1* hypermethylation

increases LDL-C levels and suggest a trend that T2D status increases cg15549216 methylation.

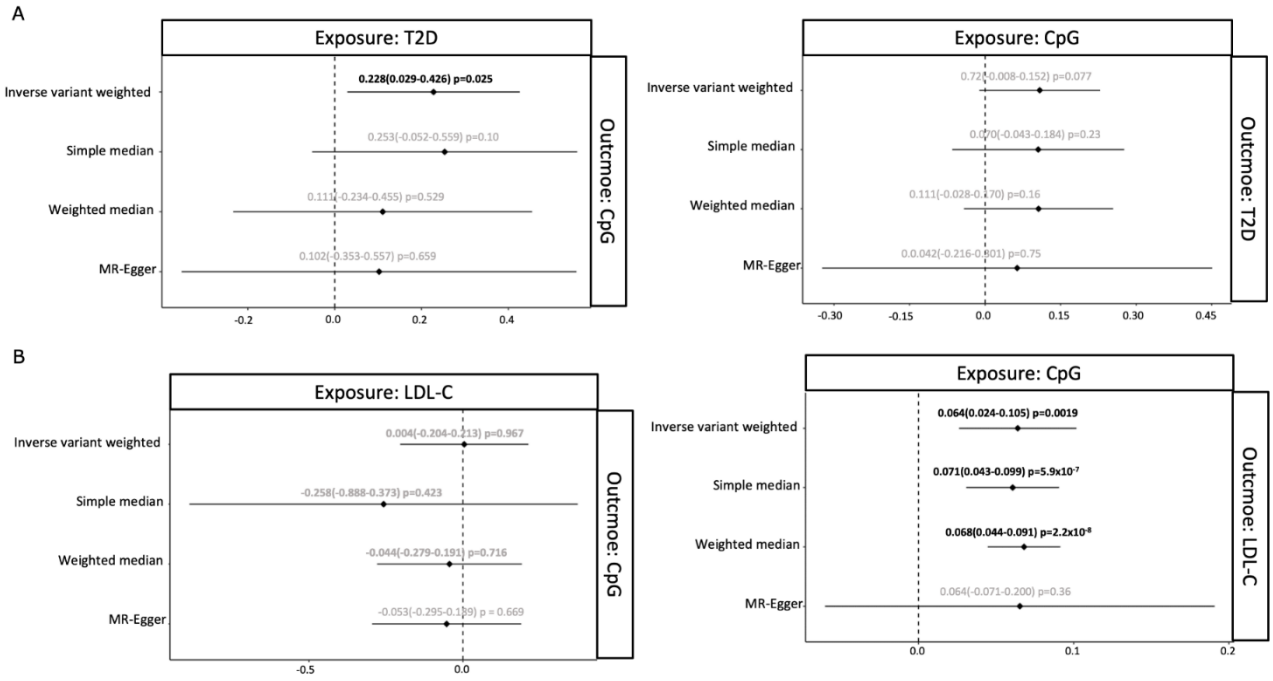


Figure 26: Mendelian Randomisation (MR) to determine causality. **A)** Bi-directional MR to assess causality between T2D and *PNLIPRP1* methylation (CpG), and **B)** between LDL-cholesterol and *PNLIPRP1* methylation, using the inverse variance weighted, simple median, weighted median and MR-Egger methods. The x-axis represents the estimates. The statistically significant associations are shown in black text, and non-significant values are shown in grey, depicting the estimate value, as well as the 95% CI values in parentheses.

9.3.1.5 Linking PNLIPRP1 expression to T2D and methylation of the gene

Next, we asked if *PNLIPRP1* expression is dysregulated in T2D. To this end, we quantified *PNLIPRP1* gene expression in RNA extracted from whole pancreas tissue from a subset of five T2D donors and six non-diabetic controls, matched for age, sex, and BMI (**Supplementary Table 23**). Our analysis revealed that *PNLIPRP1* expression was downregulated in T2D donors (53% reduction, p=0.011; **Figure 27A**). No association was found between *PNLIPRP1* expression and age (p = 0.17). To evaluate the impact of methylation on *PNLIPRP1* expression, we performed a linear regression analysis of methylation levels and PNLIPRP1 gene expression. We observed a significant correlation between increased cg15549216 methylation and decreased *PNLIPRP1* expression (p = 0.042; R^2 = 0.35; **Figure 27B**).

Given that our samples are whole pancreas, we asked where PNLIPRP1 was being expressed to pin-point further characterisation. First, we queried the GTEx database, where we found that *PNLIPRP1* is exclusively expressed in the whole pancreas and not in any of the remaining 53 tissues referenced in the GTEx database (**Supplementary figure 12**) (https://gtexportal.org/home/). The TIGER database, the largest repository of pancreatic islets RNA-sequencing data, supplemented with RNA-seq data from GTex further indicated that *PNLIPRP1* is predominantly expressed in the whole pancreas (median Transcript per Million [TPM] = 2581) with much lower expression in pancreatic islets (median TPM = 36) (**Supplementary figure 13**). To confirm the acinar-specific expression of *PNLIPRP1* at the protein level, we conducted immunofluorescence staining on human pancreatic tissue for PNLIPRP1, KRT19 (a ductal cell marker), and insulin (a marker for pancreatic islets). The results demonstrated that *PNLIPRP1* protein is exclusively expressed in acinar cells, with no expression in pancreatic islets or ductal cells (**Figure 27C**).

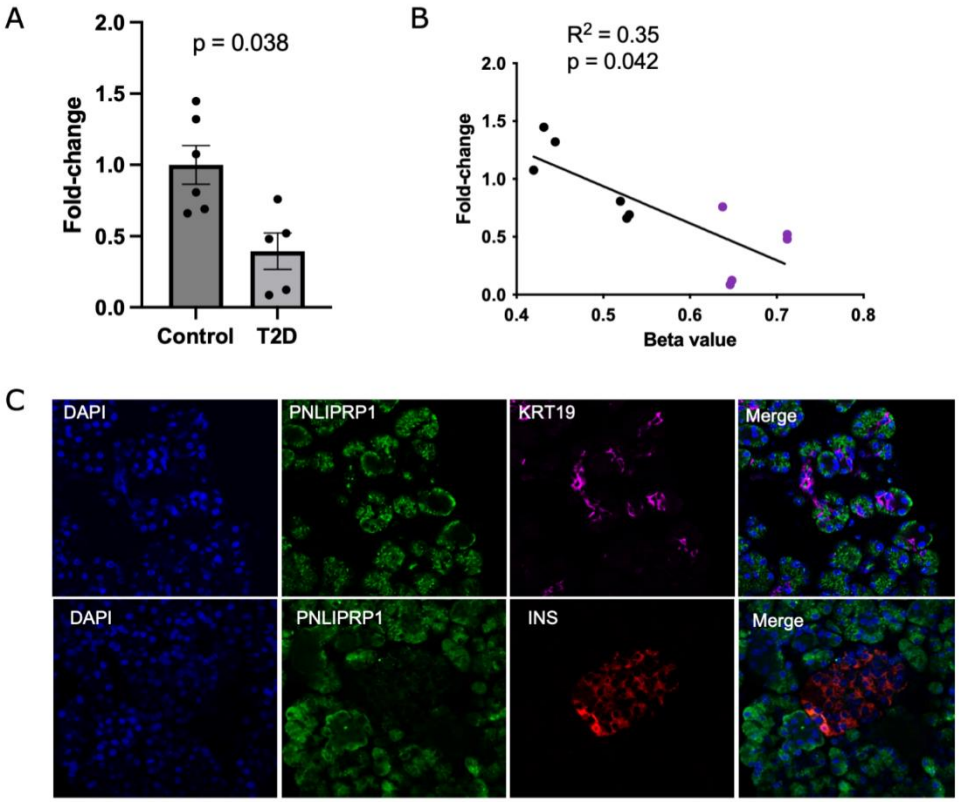


Figure 27: **Examination of *PNLIPRP1* expression and localisation. A)** The RNA expression levels of the PNLIPRP1 gene were measured in a sample of 11 individuals (5 with Type 2 Diabetes (T2D) and 6 without diabetes, matched for age, sex, and BMI). A t-test was performed to compare the gene expression levels between the T2D group and the non-diabetic controls. Error bars represent the standard error of the mean. **B)** The relationship between the cg15549216 probe and *PNLIPRP1* gene RNA expression was analysed using qPCR, with results normalised to the housekeeping gene *RPLP0*. Black dots depict non-diabetic subjects, while purple dots denote individuals with T2D. **C)** Immunofluorescence staining of pancreatic tissue samples from healthy individuals shows the localization of PNLIPRP1, KRT19 (a marker for ductal cells), and INS (insulin, a marker for pancreatic islets). DAPI was used to stain the nuclei.

9.3.1.6 A diabetogenic exposure downregulates Pnliprp1 and promotes acinar-to-ductal metaplasia in AR42J cells

Having determined that *PNLIPRP1* is associated with glucose and lipid traits, and dysregulated in T2D, we performed functional characterisation of the gene. First, we verified that T2D dysregulates the gene in the rat acinar cell line AR42J. To this end, we first verified that AR42J respond to insulin by measuring phosphorylated AKT levels following insulin exposure (**Supplementary figure 14**). Next, we treated the cells to a diabetogenic environment, namely high glucose and high insulin. High glucose or high insulin alone did not impact the expression of the gene, but the combination of the two resulted in a 35% decrease of *Pnliprp1* expression (**Figure 28B**). In the skeletal muscle and adipose tissue, insulin induces glucose uptake. We asked if this occurs in AR42J as it may explain why only high glucose and insulin treatment together induced a downregulation of the gene. We found that the medium of cells treated to high glucose and insulin had 19% less glucose, indicating uptake by the cells ($p < 0.0001$; **Figure 28B**), indicating increased glucose uptake by the cells. Given that T2D is associated with ADM in human pancreatic tissue, we asked if we could observe a similar phenomenon in our cells (Wright et al. 2024). To this end we assayed markers of ADM, namely cellular proliferation, cholesterol, and markers of exocrine identity (Grisan et al. 2021; Carrer et al. 2019). An MTS proliferation assays confirmed that high glucose and insulin led to a 25% reduction in cell proliferation ($p = 0.0089$; **Figure 28C**). However, total cholesterol content was not impacted ($p = 0.1002$; **Figure 28D**). We examined the expression of acinar and ductal markers in cells treated with high glucose and insulin. Two acinar markers, *Cpa2* ($p = 0.0281$) *and* Ctrl ($p = 0.0022$), were downregulated, while *Prss1* ($p = 0.5054$) and *Amy2* ($p = 0.7209$) remained unchanged. The ductal markers *Krt19* ($p = 0.0140$) and *Hnf1b* ($p = 0.0344$) were upregulated, indicating ADM following high glucose and insulin exposure (**Figure 28E**). To confirm this, we performed immunofluorescence analysis of whole pancreas tissue from individuals with T2D and controls. We did not observe notable differences in PNLIPRP1 protein levels, but in some cells KRT19 and

PNLIPRP1 were found to colocalise, indicating ADM (**Supplementary figure 15**). Together, these results suggest that high glucose and insulin treatment induces a downregulation of *Pnliprp1*, a decrease in proliferation, and ADM in these cells.

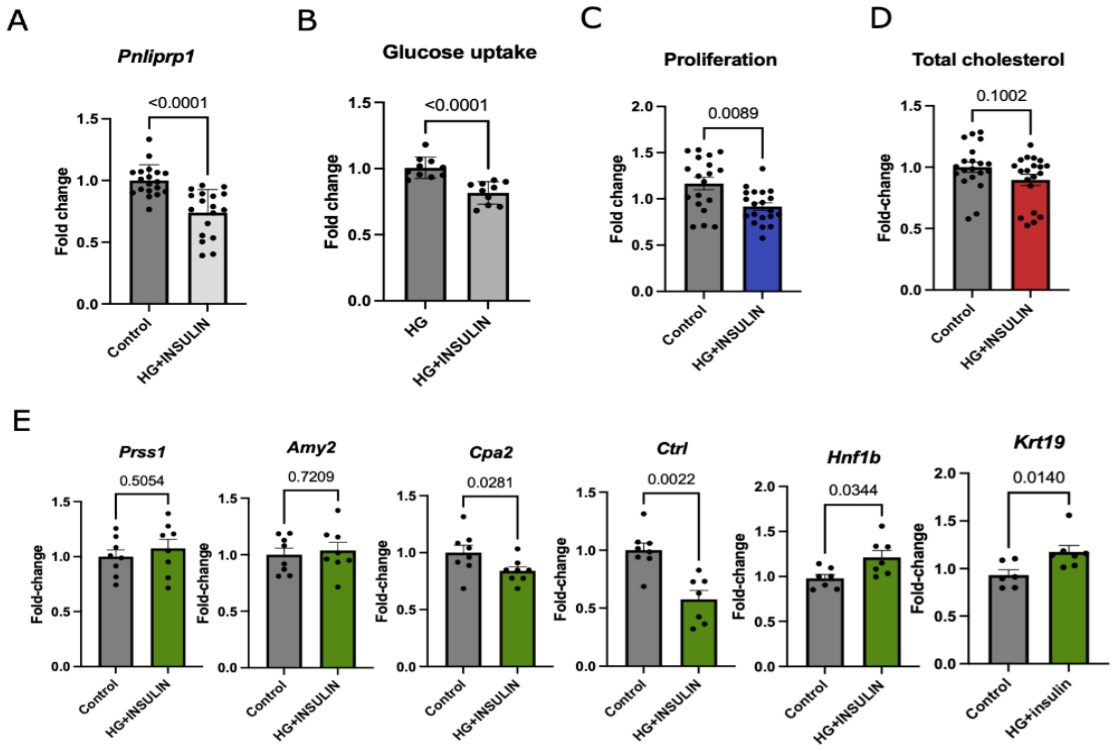


Figure 28: Exposing AR42J to a diabetogenic environment via high glucose and insulin treatment. A) *PNLIPRP1* expression following high glucose (HG; 20 mM glucose) and insulin (100 nM) treatment compared to untreated controls, as determined by qPCR. A two-tailed t-test was performed to assess statistical significance. Performed in four biological replicates. **B)** Glucose uptake by AR42J cells following treatment with high glucose (20 mmol/L) and insulin (100 nmol/L), compared to non-treated AR42J cells, measured by qPCR. A two-tailed t-test was performed to assess statistical significance. Performed in six biological replicates. **C)** MTS proliferation assay after 48 hours of HG and insulin treatment in AR42J cells, compared to controls. **D)** Total cholesterol levels measured after 48 hours of HG (20 mM glucose) and insulin (100 nM) treatment compared to controls. **E)** Gene expression of acinar and ductal markers in AR42J cells treated with HG and insulin compared to controls. A two-tailed t-test was performed to assess statistical significance. Performed in three biological replicates

9.3.1.7 *Pnliprp1* knockdown induces a dysregulation of the cell cycle, cholesterol metabolism, and ADM

Based on the acinar-specific expression pattern of *PNLIPRP1*, and the consistent pattern of *PNLIPRP1* downregulation *in vivo* and *in vitro* in the context of T2D, we further explored the functional role of *Pnliprp1* in AR42J. To this end, we performed a KD of *Pnliprp1* in AR42J to determine the downstream consequences of the dysregulation of gene. First, we validated our KD and observed a 60% reduction in the expression of the gene (**Figure 28A**) and a 34% decrease at the protein level (**Supplementary figure 16**). Next, to obtain a global overview of the consequences of knocking-down *Pnliprp1*, we performed RNA sequencing. We confirmed that *Pnliprp1* was one of the most significant down-regulated genes (FDR = 0.025). To assess the pathways dysregulated by the KD, we performed a pathway analysis with enrichR using all dysregulated genes but split by differential expression direction (up- or downregulated) (nominal $p < 0.05$) (**Supplementary Table 24**). In the downregulated genes, (587 genes), the cell cycle pathway (adjusted $p = 2.21 \times 10^{-12}$; **Supplementary Table 25**) was the most disrupted, whilst in the upregulated genes, (437 genes) the "cholesterol biosynthesis" and "SREBP control of lipid biosynthesis" were the most dysregulated pathways (adjusted p = 0.0044; **Supplementary Table 26**; **Figure 28B**). To valiate this, we assessed proliferation in AR42J following the KD of Pnliprp1 using the MTS proliferation and observed a 22% reduction in cellular proliferation (p = 0.0027; **Figure 28C**). Next, we quantified total cholesterol content in AR42J following *Pnliprp1* KD and found a 29% increase (p = 0.0017; **Figure 28D**). We asked whether we could revert this increase in cholesterol with simvastatin, a cholesterol-lowering drug and indeed found that the drug was able to rescue the increase in cholesterol biosynthesis to levels of the control (**Figure 28E**). Demonstrably, the dysregulation of *Pnliprp1* alters both the cell cycle and cholesterol metabolism.

Finally to link *Pnliprp1* expression to ADM, we assessed the expression of acinar and ductal markers. We found that the expression of three acinar markers was downregulated: *Prss1* ($p = 0.0066$), *Amy2* ($p = 0.0029$), *Cpa2* ($p = 0.035$) (also downregulated by HG + INS). Finally, both *Krt19* ($p = 0.0010$) and *Hnf1b* ($p = 0.020$) which were upregulated by a diabetogenic environment were also upregulated following the KD of *Pnliprp1* (**Figure 28F**). Taken together, these results suggest that *Pnliprp1* invalidation appears to induce ADM, which was also induced by a diabetogenic environment.

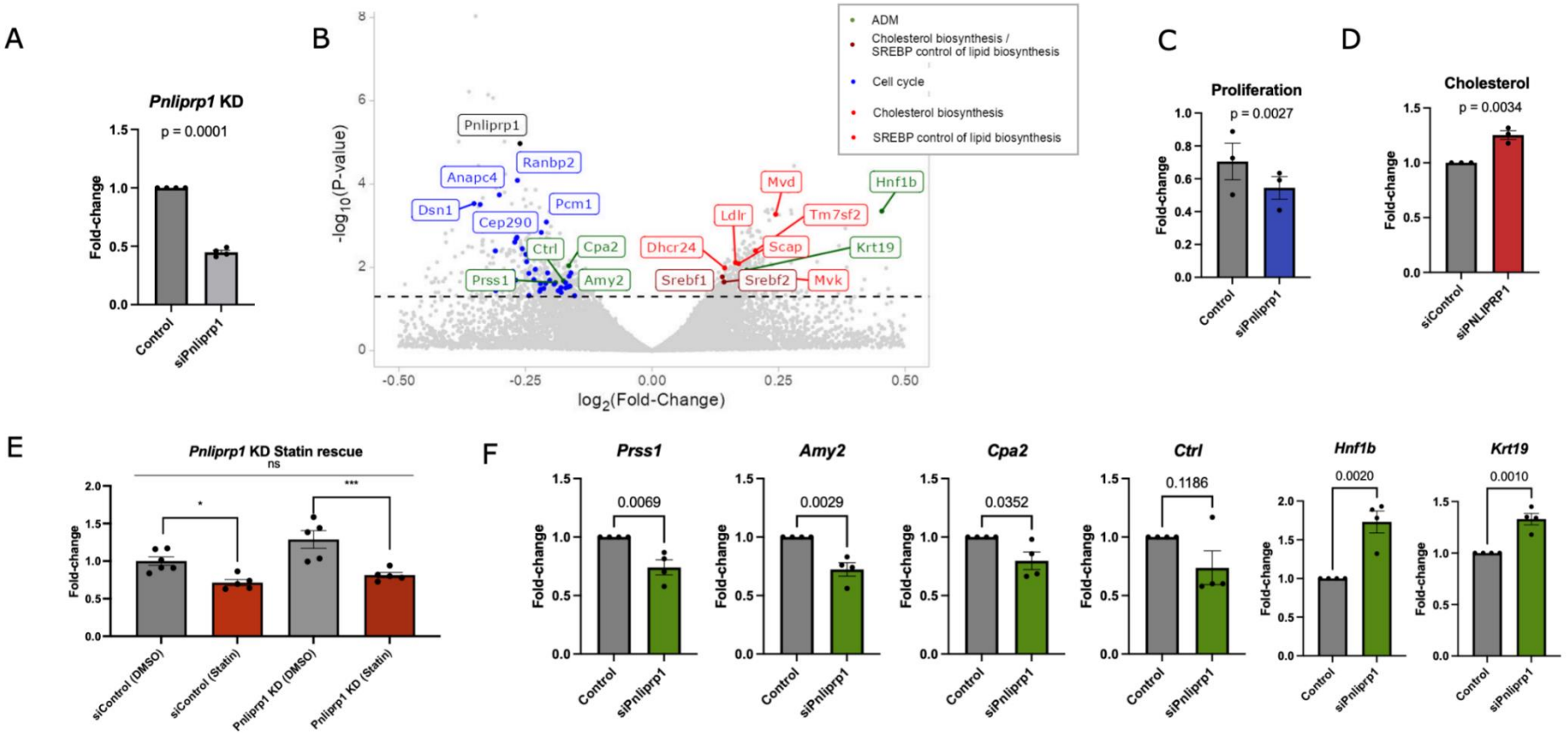


Figure 29: Characterising the downstream consequences of *Pnliprp1* downregulation in AR42J. A) Confirmation of Pnliprp1 knockdown (KD) in four biological replicates using qPCR. A two-tailed t-test was performed to assess statistical significance. **B)** Volcano plot showing the dysregulated genes following *Pnliprp1* KD in AR42J. Genes of interest are coloured and highlighted: *Pnliprp1* (black), cell cycle genes (blue), acinar-to-ductal metaplasia genes (green), and cholesterol metabolism genes (red). **C)** Quantification of AR42J proliferation following KD of *Pnliprp1* (72 hours) with the MTS proliferation assay. Performed in three biological replicates. **D)** Measurement of total cholesterol following the 48 hours KD of Pnliprp1 in AR42J. Performed in three biological replicates. **E)** Measurement of total cholesterol following *Pnliprp1* KD (48 hours) and statin treatment (24 hours) in AR42J cells. DMSO was used as a control for statin treatment. Performed in two biological replicates. Error bars represent the standard error. * = $p<0.05$, *** = $p< 0.001$. **F)** Expression level of acinar and ductal markers in AR42J cells following the 72 hours KD of *Pnliprp1* as measured by qPCR. A two-tailed t-test was performed to assess statistical significance. Performed in four biological replicates.

9.4 Discussion

Here we performed the first EWAS for T2D in the exocrine pancreas. We identified a single hypermethylation associated with T2D and 718 differentially methylated CpGs associated with age. We observed that many age-associated alterations were shared between the endocrine and exocrine pancreas. For T2D, the single hypermethylation, in cg15549216, was located in a super-enhancer predicted to target the *PNLIPRP1* gene, prompting further functional characterisation which revealed a role for the gene in cholesterol. Finally, we reproduce recent data which suggest a role for T2D in ADM, an important step in PDAC initiation.

We propose that *PNLIPRP1* may act as a mediator between T2D and various metabolic disruptions that promote downstream comorbidities such as PDAC. First, we found that the gene was downregulated in individuals with T2D, and that its expression was correlated to the methylation status of the hypermethylated cg15549216. We found that both rare and common variants of the gene were associated with LDL-cholesterol, which we confirmed by 1) MR analysis of cg15549216 and LDL-levels, and 2) by KO of *Pnliprp1* in AR42J cells. These results showed that *Pnliprp1* downregulation increases LDL-levels. This is coherent with previous studies which proposed that *PNLIPRP1* is an inhibitor of *Pancreatic lipase* (*PNLIP*) which breaks down fats during digestion (Wagner et al. 2022). Further, an evolutionary study has linked gene loss of *PNLIPRP1* to low fat diets in many species, cementing the notion that it is linked to lipid metabolism (Wagner et al. 2022; Hecker, Sharma, and Hiller 2019). The relation between *PNLIPRP1* and cholesterol is of particular interest. Indeed, increased cholesterol levels are known to support PDAC carcinogenesis, in part through acinar-to-ductal metaplasia (ADM) which is promoted by cholesterol in acinar cells (Grisan et al. 2021). Supporting this observation, we found that KD of *Pnliprp1* in AR42J cell induced both a reduction in proliferation, downregulation of acinar markers, and upregulation of ductal markers, which also

suggest ADM. ADM is traditionally reversible, however, in states of sustained stress, these cells become predisposed to progression towards PDAC, the tumour with the worst prognosis (Neuhöfer et al. 2021; Chuvin et al. 2017; J. Liu et al. 2016; Shi et al. 2013). Importantly, *PNLIPRP1* has already been linked to PDAC. Indeed, it is one of the most significantly downregulated gene in the disease (Zhang et al. 2013).

In addition to a direct impact on cholesterol, *PNLIPRP1* appears linked to multiple metabolic traits. Indeed, we found that rare and common genetic variants are associated with increased LDL-cholesterol and metabolic traits. Specifically, rare variants were associated with LDL-cholesterol, HDL-cholesterol, BMI, glucose levels, waist-to-hip ratio, waist circumference, diastolic blood pressure and systolic blood pressure. It seems unlikely that *PNLIPRP1* is driving all these alterations, rather they may be consequences of cholesterol metabolism disruption

Importantly, we were able to down-regulate *Pnliprp1* and replicate the effects of KD the gene by exposing our cells to a diabetogenic environment. Indeed, insulin or glucose alone did not trigger alterations in *Pnliprp1* expression. These results confirm previous observations that T2D alters exocrine tissue (Wright et al. 2024). This is important not only to PDAC, but other diseases of the exocrine such as pancreatitis. Our data further demonstrate the importance of proper endocrine pancreas function to exocrine pancreas health. Indeed, in T1D, where insulin is lost, the exocrine pancreas undergoes atrophy and fails to secrete sufficient digestive hormones, sometimes requiring hormone replacement therapy (Wright et al. 2020).

Finally, we show that age-associated DNA methylation changes in the exocrine pancreas are linear and representative of age. Furthermore, we find that many age-associated DNA methylation changes in the exocrine pancreas overlap with age-associated changes in both the endocrine pancreas, and other tissues such as blood. Finally, individuals with T2D featured accelerated ageing, confirming observations in other tissue whereby accelerated ageing is associated with disease (Horvath and Topol 2024; Fafián-Labora and O'Loghlen 2020). The consistence of age-associated

DNA methylation changes across tissue suggest that age is a well-controlled physiological process.

It should be noted that this work features several limitations. First, given the extent of changes associated with T2D exposure (PDAC, ADM) in the exocrine pancreas, it is unlikely that *PNLIPRP1* is the only differentially methylated gene, however, lack of statistical power likely prevented us from identifying additional alterations. Second, our functional characterisation is limited to rat acinar cells, which do not recapitulate the full complexity of tissue, and may not be representative of human physiology, however, primary human acinar cells rapidly de-differentiate in vitro and are therefore a poor model. Finally, to assess the role of cholesterol in inducing ADM, it would have been of interest to treat the AR42J cells with cholesterol.

This work further demonstrates the value of DNA methylation in studying the molecular events upstream of disease. Indeed, here we assessed the contribution of an exposure, T2D, on DNA methylation. We found that T2D induces a hypermethylation in *PNLIPRP1*, a gene which we link to ADM and may therefore contribute to PDAC initiation in the exocrine pancreas, providing important insights into the development of the disease. More generally, we reinforce the observation that T2D promotes ADM in the exocrine pancreas, likely through both high glucose and high insulin exposure. Deciphering the molecular alterations leading to disease is essential for advancing our knowledge of both conditions and improving therapeutic approaches. Here, we demonstrate that DNA methylation studies are suitable vehicles for such investigations.

10 Concluding remarks and future perspectives

10.1 Summary

In this work, we explored the (epi)genetic landscape of both the endocrine and exocrine compartments of the pancreas in the context of age and T2D. Our findings reveal significant associations between both traits and DNA methylation changes and that these epigenetic changes are largely independent of nearby genetic variants. Importantly, we linked these modifications to expression changes in previously established genes as well as probable novel candidates for T2D. This underscores the utility of DNA methylation in characterizing disease mechanisms and physiological processes. Within the endocrine pancreas, we uncovered associations that provide clear mechanistic insights into T2D and β-cell function. In the exocrine pancreas, we observed hypermethylation that led to the identification and characterization of *PNLIPRP1*, a novel gene that may play a role in T2D-related disruptions of exocrine function by modulating cholesterol metabolism. Furthermore, our data strongly support a model where genetics and epigenetics contribute independently to traits and disease. This is evidenced by the limited influence of genetics in our MEGA analyses and by the additive effect of MRS to PGS in predicting T2D. These findings highlight the importance of investigating both genetic and epigenetic factors to obtain a comprehensive understanding of T2D, and more generally, biology.

10.2 DNA methylation-based studies: a complementary approach to traditional eQTL and colocalisation approaches

Ultimately, the aim of omics studies is to uncover genes and pathways implicated in disease. While purely *in silico* omics approaches cannot definitively establish candidate genes and their biological consequences, they provide a valuable starting point. GWAS and eQTL analyses for T2D have indeed pinpointed signals that were later validated *in vitro*, such as *FCHSD2* (Hu et al. 2021). As environmental contributions to T2D are indeed crucial, we propose that DNA methylation studies

could serve a similar function to genetics-based approaches in identifying novel candidate genes. For instance, in our exocrine study, our EWAS revealed a hypermethylation of specific CpG site led to the identification of *PNLIPRP1*, a gene with a probable role in T2D-associated alterations in the exocrine pancreas. In the pancreatic islets study, we found that changes in DNA methylation were associated with alterations in key T2D genes and pathways. Furthermore, we identified genes not previously linked to T2D but with a probable role in the disease, such as *RBM15*, a regulator of *HES1*, a key β-cell identity marker (Bolinches-Amorós et al. 2014).

However, DNA methylation-based approaches feature two important limitations: (1) they detect many dysregulated genes, not all of which are relevant in a disease context, and (2) it is challenging to determine which CpG sites are causal rather than merely correlated to changes in gene expression. The first limitation can be partially addressed by linking the expression of identified genes to relevant phenotypic traits. In our work, we linked gene expression to the insulin secretion index (ISI), a direct measure of pancreatic islet function. This enabled us to identify probable candidates, such as *ECHDC2*, which regulates glycolysis, a key component of GSIS, but future studies should aim to obtain a larger number of clinical variables. The second limitation can be addressed with genome-editing strategies, such as CRISPR, which have been successfully used to validate eQTL signals, including *SIX3* (Bevacqua, Dai, et al. 2021; Hu et al. 2021). However, genome editing is an intensive process and difficult to scale, meaning that CpG-gene pairs of interest need to be prioritised before characterization. Here MR could be leveraged to link target CpGs to traits. Exposure SNPs could be generated via mQTL. However, we identified few CpGs under genetic influence in the Epi-islets cohort, suggesting that this approach may not be widely applicable.

Alternatively, CpG-gene interactions could be filtered upstream using single-cell DNA methylation (scDNAm) and single-cell transcriptomics analyses. Individual cells, of the same cell type, could be grouped by the methylation status of a target CpG. Then, case-control studies could be performed where individual cells are grouped based on

the methylation status of the CpG site of interest. The resulting impact on target gene expression could then be measured to determine whether a CpG-gene pair warrants further characterization. Characterisation of CpG-Gene pairs is important not only in the context of gene expression, but also because CpGs which demonstrably impact gene expression are likely to be more robust markers of disease.

DNA methylation also complements GWAS and eQTL analyses because of its distinct nature from genetics. SNPs are determined at conception, and their effects are stable throughout an individual's lifetime. In contrast, DNA methylation patterns vary over time and in response to environmental stimuli (Smith, Hetzel, and Meissner 2024). Consequently, other approaches are needed to identify these environmentally regulated genes. Indeed, as described in the first project, few of our identified CpG-gene pairs were under genetic control, either through the SNP-gene axis or the SNP-CpG-gene axis. This finding underscores the great potential of environmentally focused DNA methylation studies to identify novel genes only under epigenetic control.

Finally, DNA methylation alterations cell-type and tissue-specific. This is demonstrated by DNA methylation based smoking predictors and age clock, which can accurately predict whether individuals smoke and their biological age, based on the methylation status of a select few CpG site (Bollepalli et al. 2019; Horvath and Topol 2024). Additionally, DNA methylation can differentiate between distinct treatments and cell types in the adipose tissue of obese individuals (Macartney-Coxson et al. 2017). Many age-related diseases are heterogeneous and feature distinct pathophysiological clusters. In T2D, which comprises at least five distinct clusters, certain groups exhibit greater IR while others are characterised by β-cell dysfunction (Ahlqvist, Tuomi, and Groop 2019). Furthermore, each cluster features distinct disease progression, with retinopathy and neuropathy being more prevalent in insulin deficient patients whilst IR-driven diabetes is enriched in fatty liver disease (Ahlqvist, Tuomi, and Groop 2019). Each subtype results from disruptions in different genes, pathways, cell types, tissues, and organs. Recent work confirmed the

heterogeneity of each subtype at the DNA methylation levels (Schrader et al. 2022). Integration of each subtype to transcriptomics could shed light into the genes and pathways contributing to each of these. At the least, DNA methylation could likely improve currently available genetic-based methods for T2D subtypes partitioning (DiCorpo et al. 2022)

10.3 General limitations of current methylation-based studies

10.3.1 Biased and incomplete coverage of the methylome by arrays

Whilst DNA methylation has proven useful in the identification of biomarkers and candidate genes, currently used array-based methods have several limitations. First, the largest array, the Infinium MethylationEPIC v2.0, only covers an estimated 3% of the 28,000,000 CpGs present in the genome (Lander et al. 2001). Unlike SNPs from DNA arrays, it is not possible to impute missing CpGs, and the remaining CpGs remain overlooked. This issue is further exacerbated by the non-random selection of CpGs, as methylation arrays were initially designed to assess cancer-specific regions of interest, and more broadly promoters and enhancer regions (Moran, Arribas, and Esteller 2016).

Additionally, DNA methylation patterns are phenotype specific, as evidenced by smoking and age, which can accurately be predicted with a select few CpGs. This suggests that despite identifying many associations with the disease, regions relevant to T2D may be unscreened in current EWAS studies (Salameh, Bejaoui, and El Hajj 2020; Bollepalli et al. 2019). Furthermore, arrays, limited to less than a million probes are overwhelmingly focused on promoters, enhancers, and open chromatin regions identified by the ENCODE project (Moran, Arribas, and Esteller 2016). However, DNA methylation is an important regulatory component of compact chromatin regions, or heterochromatin (Pappalardo and Barra 2021). CpGs in heterochromatin are typically highly methylated and a decrease in their methylation level is associated with

chromatin deconvolution and transcriptional activation of RE and TE (Pappalardo and Barra 2021).

Activation of RE is linked to oncogene activation (Pappalardo and Barra 2021). In various cancers, hypomethylations alter the entire 3D structure of chromatin, impacting nucleus size, and DNA spatial organisation (Zeimet et al. 2011; McDonald et al. 2011). We and others have clearly demonstrated a trend for global hypomethylations in T2D, however, whether these induce a complete remodelling of DNA structure is unknown but should be assessed. This could be addressed with whole-genome bisulfite sequencing (WGBS), which enables complete and comprehensive screening of the genome. In cancer, WGBS has been successfully been used to screen heterochromatin regions, which are indeed heavily hypomethylated and associated with activation of TEs (D. Bar et al. 2022). However, it is important to note that as the number of CpG sites under testing increases, larger sample sizes will be necessary to achieve sufficient power to detect meaningful changes in DNA methylation accurately. As WGBS becomes more cost-effective, and larger cohorts begin to be assembled, WBGS will greatly improve the scope of DNA methylation studies.

10.3.2 CpG effect size and making sense of CpG-gene interactions

Similar to SNPs, EWAS identified CpGs feature modest effect sizes. Indeed, the largest methylation changes we observed (both in T2D) in exocrine tissue was of 11.2% (cg15549216), and of 4.08% (cg17240976) in pancreatic islets. As with common variants, it is likely that each CpG induces mild changes, which additively contribute to phenotypes. An exciting opportunity lies in the characterizing of CpG-Gene interactions to develop our understanding of how these modest alterations shape transcriptional patterns given the success of GWAS and downstream approaches. This is exemplified by *PNLIPRP1*, which we found to be hypermethylated with T2D. Characterisation of the gene demonstrated a role in cholesterol metabolism, with important implications in ADM and PDAC. These results demonstrate the utility of identifying these single DNA methylation changes that could contribute to disease

and further our understanding of disease. More generally, with sufficient understanding of how DNA methylation shapes expression patterns, we could begin developing models which predict, without transcriptomic data, the impact of a CpG on gene expression, and largely improve the impact of single-OMIC EWAS.

10.4 Future opportunities

As more EWAS focusing on T2D in pancreatic islets are published, meta-analyses present an important opportunity to overcome current power limitations and detect robust, reproducible associations. To maximise their effectiveness, it is essential to standardise methodologies across studies and address potential heterogeneity in sample populations and data acquisition techniques. Importantly, studies should aim to consider clinical subtypes or include key clinical measures relevant to T2D—such as HOMA2B and the ISI —to delineate the relative contributions of the various components of T2D, such as insulin secretion and IR to DNA methylation patterns. Additionally, integrating multidimensional data can enhance our understanding of T2D pathogenesis and aid in the discovery of clinically relevant biomarkers.

OMICS datasets are becoming larger and more complex. Additionally, integration of multiple dataset vastly increases this complexity, especially in biology where systems are intimately linked *ie.,* a change in expression impacts DNA methylation which impacts gene expression which impacts protein levels. To resolve this, the utilization of machine learning is rapidly expanding (in fact, the Nobel prise for physics was awarded to ML and AI on the day of writing this). ML is excellent at identifying patterns in complex data and multi-dimensional data. ML has not yet been applied in the context of T2D and DNA methylation. However, it has successfully been leveraged in acute myeloid leukaemia, where it significantly outperformed traditional hypothesis-driven association methods in the identification of drug sensitivity markers (Lee et al. 2018).

Finally, proteomics and metabolomics are rapidly becoming more available. These omics (especially transcriptomics) directly reflect biological activity and are therefore

closer to phenotype. Indeed, proteomics and metabolomics captures processes not detected by RNA-sequencing. Alzheimer's disease is driven by long-term Aβ peptide accumulation in neurons, a process not detected by transcriptomics owing to the short-term nature of RNAs in the cells (Shaomin Li and Stern 2022). Already, DNA methylation has been integrated to proteomics to identify biomarkers for Alzheimer's, Parkinson's, and autism (Mahony and O'Ryan 2021; Suhre and Zaghlool 2021). Integration of additional OMICS is certain to bring us ever closer to the characterisation of disease, physiology, and metabolism.

10.5 A proposition on the fundamental purpose of DNA methylation

Methylation patterns are predicated by both targeted and stochastic mechanisms (Jeltsch and Jurkowska 2014). At the most basic level, we generally observe targeted hypermethylations in regulatory regions, primarily promoters, and stochastic hypomethylations in non-promoter regions (Smith, Hetzel, and Meissner 2024). In addition, we are also aware that CpGs are not normally distributed throughout the genome. Indeed, they are enriched in promoters and CpG islands and depleted outside of these regions (Smith, Hetzel, and Meissner 2024). Generally, the DNA methylation ecosystem features strong evolutionary conversion, indicating an important underlying function (Smith, Hetzel, and Meissner 2024). Therefore, targeted changes in DNA methylation should provide, at least in some cases, benefits. Indeed, this is a well-established phenomenon in cancer cells, where DNA methylations enable greater cellular survival and proliferation, relative to healthy cells (Timp and Feinberg 2013). Consequently, a function of DNA methylation may be that it generates cellular heterogeneity, which is then selected for, or against, by various environmental exposures to enhance cellular survival. We observed evidence in our data that certain DNA methylation changes appear to be protective of T2D, such as a change in DNA methylation which increased *INS* expression with increasing age. Similarly, seemingly deleterious modifications, which impair β-cell function, could be a mechanism to prevent cellular damage associated with excessive β-cell activity (Cerf 2013; Eizirik, Pasquali, and Cnop 2020). Understanding the fundamental

function of DNA methylation, and why changes in methylation occur, whether these be deleterious or protective, is important. In the future, better study designs to comprehensively assess the contribution of this mechanism to physiology are important.

10.6 Conclusion

DNA methylation in the of T2D and the pancreas is an emerging field of research. Our data demonstrates promise for DNA methylation in the characterisation of the disease as well as incidence prediction. However, current studies face limitations, ranging from sample size to structural shortcomings of arrays, which fail to comprehensively assess the methylome. Recent studies have seen a significant shift in the ability to identify epigenetic changes associated with complex disease. Although methylation studies in isolation have yielded limited results, when integrated with other OMICS, they add an interesting perspective in the study of complex disease. It is important to note that this journey into investigating epigenetic modifications has only recently started and there remains a great deal to examine the untapped potential of this field. Further statistical modelling, ML and AI approaches will improve and advance this field and aid in the unravelling of molecular mechanisms that influence complex disease. This is yet a young field that has promising potential future perspectives. By overcoming the limitations described, future research can enhance our understanding of the epigenetic mechanisms underlying T2D and improve the identification of biomarkers for better diagnosis and treatment strategies.

11 References

Abel, E. Dale, Anna L. Gloyn, Carmella Evans-Molina, Joshua J. Joseph, Shivani Misra, Utpal B. Pajvani, Judith Simcox, Katalin Susztak, and Daniel J. Drucker. 2024. "Diabetes Mellitus—Progress and Opportunities in the Evolving Epidemic." *Cell* 187 (15): 3789–3820. https://doi.org/10.1016/j.cell.2024.06.029.

Abraham, Gad, Yixuan Qiu, and Michael Inouye. 2017. "FlashPCA2: Principal Component Analysis of Biobank-Scale Genotype Datasets." *Bioinformatics* 33 (17): 2776–78. https://doi.org/10.1093/bioinformatics/btx299.

Aguayo-Mazzucato, Cristina. 2020. "Functional Changes in Beta Cells during Ageing and Senescence." *Diabetologia* 63 (10): 2022–29. https://doi.org/10.1007/s00125-020-05185-6.

Ahlqvist, Emma, Rashmi B. Prasad, and Leif Groop. 2020. "Subtypes of Type 2 Diabetes Determined From Clinical Parameters." *Diabetes* 69 (10): 2086–93. https://doi.org/10.2337/dbi20-0001.

Ahlqvist, Emma, Tiinamaija Tuomi, and Leif Groop. 2019. "Clusters Provide a Better Holistic View of Type 2 Diabetes than Simple Clinical Features." *The Lancet Diabetes & Endocrinology* 7 (9): 668–69. https://doi.org/10.1016/S2213-8587(19)30257-8.

Ahmad, Ehtasham, Soo Lim, Roberta Lamptey, David R. Webb, and Melanie J. Davies. 2022. "Type 2 Diabetes." *The Lancet* 400 (10365): 1803–20. https://doi.org/10.1016/S0140-6736(22)01655-5.

Almgren, P., M. Lehtovirta, B. Isomaa, L. Sarelin, M. R. Taskinen, V. Lyssenko, T. Tuomi, L. Groop, and for the Botnia Study Group. 2011. "Heritability and Familiality of Type 2 Diabetes and Related Quantitative Traits in the Botnia Study." *Diabetologia* 54 (11): 2811–19. https://doi.org/10.1007/s00125-011-2267-5.

Alonso, Lorena, Anthony Piron, Ignasi Morán, Marta Guindo-Martínez, Sílvia Bonàs-Guarch, Goutham Atla, Irene Miguel-Escalada, et al. 2021. "TIGER: The Gene

Expression Regulatory Variation Landscape of Human Pancreatic Islets." *Cell Reports* 37 (2). https://doi.org/10.1016/j.celrep.2021.109807.

Altshuler, D, J N Hirschhorn, M Klannemark, C M Lindgren, M C Vohl, J Nemesh, C R Lane, et al. 2000. "The Common PPARgamma Pro12Ala Polymorphism Is Associated with Decreased Risk of Type 2 Diabetes." *Nature Genetics* 26 (1): 76–80. https://doi.org/10.1038/79216.

Anık, Ahmet, Gönül Çatlı, Ayhan Abacı, and Ece Böber. 2015. "Maturity-Onset Diabetes of the Young (MODY): An Update." *Journal of Pediatric Endocrinology and Metabolism* 28 (3–4): 251–63. https://doi.org/10.1515/jpem-2014-0384.

Arvanitis, Marios, Karl Tayeb, Benjamin J. Strober, and Alexis Battle. 2022. "Redefining Tissue Specificity of Genetic Regulation of Gene Expression in the Presence of Allelic Heterogeneity." *American Journal of Human Genetics* 109 (2): 223. https://doi.org/10.1016/j.ajhg.2022.01.002.

Aryee, Martin J., Andrew E. Jaffe, Hector Corrada-Bravo, Christine Ladd-Acosta, Andrew P. Feinberg, Kasper D. Hansen, and Rafael A. Irizarry. 2014a. "Minfi: A Flexible and Comprehensive Bioconductor Package for the Analysis of Infinium DNA Methylation Microarrays." *Bioinformatics* 30 (10): 1363–69. https://doi.org/10.1093/bioinformatics/btu049.

———. 2014b. "Minfi: A Flexible and Comprehensive Bioconductor Package for the Analysis of Infinium DNA Methylation Microarrays." *Bioinformatics* 30 (10): 1363–69. https://doi.org/10.1093/bioinformatics/btu049.

Auton, Adam, Gonçalo R. Abecasis, David M. Altshuler, Richard M. Durbin, Gonçalo R. Abecasis, David R. Bentley, Aravinda Chakravarti, et al. 2015. "A Global Reference for Human Genetic Variation." *Nature* 526 (7571): 68–74. https://doi.org/10.1038/nature15393.

Babenko, Andrey P., Michel Polak, Hélène Cavé, Kanetee Busiah, Paul Czernichow, Raphael Scharfmann, Joseph Bryan, Lydia Aguilar-Bryan, Martine Vaxillaire, and

Philippe Froguel. 2006. “Activating Mutations in the ABCC8 Gene in Neonatal Diabetes Mellitus.” *New England Journal of Medicine* 355 (5): 456–66. https://doi.org/10.1056/NEJMoa055068.

Bacos, Karl, Linn Gillberg, Petr Volkov, Anders H. Olsson, Torben Hansen, Oluf Pedersen, Anette Prior Gjesing, et al. 2016a. “Blood-Based Biomarkers of Age-Associated Epigenetic Changes in Human Islets Associate with Insulin Secretion and Diabetes.” *Nature Communications* 7 (1): 11089. https://doi.org/10.1038/ncomms11089.

———. 2016b. “Blood-Based Biomarkers of Age-Associated Epigenetic Changes in Human Islets Associate with Insulin Secretion and Diabetes.” *Nature Communications* 7 (1): 11089. https://doi.org/10.1038/ncomms11089.

Bacos, Karl, Alexander Perfilyev, Alexandros Karagiannopoulos, Elaine Cowan, Jones K. Ofori, Ludivine Bertonnier-Brouty, Tina Rönn, et al. 2023. “Type 2 Diabetes Candidate Genes, Including *PAX5*, Cause Impaired Insulin Secretion in Human Pancreatic Islets.” *The Journal of Clinical Investigation* 133 (4). https://doi.org/10.1172/JCI163612.

Bar, Dror, Lior Fishman, Yueyuan Zheng, Irene Unterman, Devorah Schlesinger, Amir Eden, De-Chen Lin, and Benjamin P. Berman. 2022. “A Local Sequence Signature Defines a Subset of Heterochromatin-Associated CpGs with Minimal Loss of Methylation in Healthy Tissues but Extensive Loss in Cancer.” bioRxiv. https://doi.org/10.1101/2022.08.16.504069.

Bar, Yael, Holger A. Russ, Elad Sintov, Leeat Anker-Kitai, Sarah Knoller, and Shimon Efrat. 2012. “Redifferentiation of Expanded Human Pancreatic β-Cell-Derived Cells by Inhibition of the NOTCH Pathway.” *The Journal of Biological Chemistry* 287 (21): 17269–80. https://doi.org/10.1074/jbc.M111.319152.

Bartlett, Alexandra H., Jane W. Liang, Jose Vladimir Sandoval-Sierra, Jay H. Fowke, Eleanor M. Simonsick, Karen C. Johnson, and Khyobeni Mozhui. 2019. “Longitudinal

Study of Leukocyte DNA Methylation and Biomarkers for Cancer Risk in Older Adults." *Biomarker Research* 7 (1): 10. https://doi.org/10.1186/s40364-019-0161-3.

Bell, Christopher G., Robert Lowe, Peter D. Adams, Andrea A. Baccarelli, Stephan Beck, Jordana T. Bell, Brock C. Christensen, et al. 2019. "DNA Methylation Aging Clocks: Challenges and Recommendations." *Genome Biology* 20 (1): 249. https://doi.org/10.1186/s13059-019-1824-y.

Bevacqua, Romina J., Xiaoqing Dai, Jonathan Y. Lam, Xueying Gu, Mollie S. H. Friedlander, Krissie Tellez, Irene Miguel-Escalada, et al. 2021. "CRISPR-Based Genome Editing in Primary Human Pancreatic Islet Cells." *Nature Communications* 12 (1): 2397. https://doi.org/10.1038/s41467-021-22651-w.

Bevacqua, Romina J., Jonathan Y. Lam, Heshan Peiris, Robert L. Whitener, Seokho Kim, Xueying Gu, Mollie S. H. Friedlander, and Seung K. Kim. 2021. "SIX2 and SIX3 Coordinately Regulate Functional Maturity and Fate of Human Pancreatic β Cells." *Genes & Development* 35 (3–4): 234–49. https://doi.org/10.1101/gad.342378.120.

Bibikova, Marina, Bret Barnes, Chan Tsan, Vincent Ho, Brandy Klotzle, Jennie M. Le, David Delano, et al. 2011. "High Density DNA Methylation Array with Single CpG Site Resolution." *Genomics*, New Genomic Technologies and Applications, 98 (4): 288–95. https://doi.org/10.1016/j.ygeno.2011.07.007.

Bodega, Beatrice, and Valerio Orlando. 2014. "Repetitive Elements Dynamics in Cell Identity Programming, Maintenance and Disease." *Current Opinion in Cell Biology*, Cell cycle, differentiation and disease, 31 (December):67–73. https://doi.org/10.1016/j.ceb.2014.09.002.

Boland, Brandon B., Christopher J. Rhodes, and Joseph S. Grimsby. 2017. "The Dynamic Plasticity of Insulin Production in β-Cells." *Molecular Metabolism* 6 (9): 958–73. https://doi.org/10.1016/j.molmet.2017.04.010.

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Bolinches-Amorós, Arantxa, Belén Mollá, David Pla-Martín, Francesc Palau, and Pilar González-Cabo. 2014. “Mitochondrial Dysfunction Induced by Frataxin Deficiency Is Associated with Cellular Senescence and Abnormal Calcium Metabolism.” *Frontiers in Cellular Neuroscience* 8. https://doi.org/10.3389/fncel.2014.00124.

Bollati, Valentina, Joel Schwartz, Robert Wright, Augusto Litonjua, Letizia Tarantini, Helen Suh, David Sparrow, Pantel Vokonas, and Andrea Baccarelli. 2009. “Decline in Genomic DNA Methylation through Aging in a Cohort of Elderly Subjects.” *Mechanisms of Ageing and Development* 130 (4): 234–39. https://doi.org/10.1016/j.mad.2008.12.003.

Bollepalli, Sailalitha, Tellervo Korhonen, Jaakko Kaprio, Simon Anders, and Miina Ollikainen. 2019. “EpiSmokEr: A Robust Classifier to Determine Smoking Status from DNA Methylation Data.” *Epigenomics* 11 (13): 1469–86. https://doi.org/10.2217/epi-2019-0206.

Bonnefond, Amélie, Ranjit Unnikrishnan, Alessandro Doria, Martine Vaxillaire, Rohit N. Kulkarni, Viswanathan Mohan, Vincenzo Trischitta, and Philippe Froguel. 2023. “Monogenic Diabetes.” *Nature Reviews Disease Primers* 9 (1): 1–16. https://doi.org/10.1038/s41572-023-00421-w.

Bowman, Pamela, Frances Mathews, Fabrizio Barbetti, Maggie H. Shepherd, Janine Sanchez, Barbara Piccini, Jacques Beltrand, et al. 2020. “Long-Term Follow-up of Glycemic and Neurological Outcomes in an International Series of Patients With Sulfonylurea-Treated ABCC8 Permanent Neonatal Diabetes.” *Diabetes Care* 44 (1): 35. https://doi.org/10.2337/dc20-1520.

Brissova, Marcela, Alena Shostak, Corinne L. Fligner, Frank L. Revetta, Mary K. Washington, Alvin C. Powers, and Rebecca L. Hull. 2015. “Human Islets Have Fewer

Blood Vessels than Mouse Islets and the Density of Islet Vascular Structures Is Increased in Type 2 Diabetes." *Journal of Histochemistry & Cytochemistry* 63 (8): 637–45. https://doi.org/10.1369/0022155415573324.

Broadaway, K. Alaine, Xianyong Yin, Alice Williamson, Victoria A. Parsons, Emma P. Wilson, Anne H. Moxley, Swarooparani Vadlamudi, et al. 2023. "Loci for Insulin Processing and Secretion Provide Insight into Type 2 Diabetes Risk." *American Journal of Human Genetics* 110 (2): 284–99. https://doi.org/10.1016/j.ajhg.2023.01.002.

Bryois, Julien, Alfonso Buil, David M. Evans, John P. Kemp, Stephen B. Montgomery, Donald F. Conrad, Karen M. Ho, et al. 2014. "Cis and Trans Effects of Human Genomic Variants on Gene Expression." *PLOS Genetics* 10 (7): e1004461. https://doi.org/10.1371/journal.pgen.1004461.

Bunt, Martijn van de, Jocelyn E. Manning Fox, Xiaoqing Dai, Amy Barrett, Caleb Grey, Lei Li, Amanda J. Bennett, et al. 2015. "Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors." *PLOS Genetics* 11 (12): e1005694. https://doi.org/10.1371/journal.pgen.1005694.

Bysani, Madhusudhan, Alexander Perfilyev, Vanessa D de Mello, Tina Rönn, Emma Nilsson, Jussi Pihlajamäki, and Charlotte Ling. 2017. "Epigenetic Alterations in Blood Mirror Age-Associated Dna Methylation and Gene Expression Changes in Human Liver." *Epigenomics* 9 (2): 105–22. https://doi.org/10.2217/epi-2016-0087.

Cabrera, Over, Dora M. Berman, Norma S. Kenyon, Camillo Ricordi, Per-Olof Berggren, and Alejandro Caicedo. 2006. "The Unique Cytoarchitecture of Human Pancreatic Islets Has Implications for Islet Cell Function." *Proceedings of the National Academy of Sciences* 103 (7): 2334–39. https://doi.org/10.1073/pnas.0510790103.

Campbell, Jonathan E., and Christopher B. Newgard. 2021. "Mechanisms Controlling Pancreatic Islet Cell Function in Insulin Secretion." *Nature Reviews Molecular Cell Biology* 22 (2): 142–58. https://doi.org/10.1038/s41580-020-00317-7.

Cannata, Dara, Yvonne Fierz, Archana Vijayakumar, and Derek LeRoith. 2010. “Type 2 Diabetes and Cancer: What Is the Connection?” *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine* 77 (2): 197–213. https://doi.org/10.1002/msj.20167.

Carrer, Alessandro, Sophie Trefely, Steven Zhao, Sydney L. Campbell, Robert J. Norgard, Kollin C. Schultz, Simone Sidoli, et al. 2019. “Acetyl-CoA Metabolism Supports Multistep Pancreatic Tumorigenesis.” *Cancer Discovery* 9 (3): 416–35. https://doi.org/10.1158/2159-8290.CD-18-0567.

Cavalli, Giacomo, and Edith Heard. 2019. “Advances in Epigenetics Link Genetics to the Environment and Disease.” *Nature* 571 (7766): 489–99. https://doi.org/10.1038/s41586-019-1411-0.

Cerf, Marlon E. 2013. “Beta Cell Dysfunction and Insulin Resistance.” *Frontiers in Endocrinology* 4 (March). https://doi.org/10.3389/fendo.2013.00037.

Chahrour, Maria, Sung Yun Jung, Chad Shaw, Xiaobo Zhou, Stephen T. C. Wong, Jun Qin, and Huda Y. Zoghbi. 2008. “MeCP2, a Key Contributor to Neurological Disease, Activates and Represses Transcription.” *Science* 320 (5880): 1224–29. https://doi.org/10.1126/science.1153252.

Chambers, John C, Marie Loh, Benjamin Lehne, Alexander Drong, Jennifer Kriebel, Valeria Motta, Simone Wahl, et al. 2015. “Epigenome-Wide Association of DNA Methylation Markers in Peripheral Blood from Indian Asians and Europeans with Incident Type 2 Diabetes: A Nested Case-Control Study.” *The Lancet Diabetes & Endocrinology* 3 (7): 526–34. https://doi.org/10.1016/S2213-8587(15)00127-8.

Chatterjee, Sudesna, Kamlesh Khunti, and Melanie J Davies. 2017. “Type 2 Diabetes.” *The Lancet* 389 (10085): 2239–51. https://doi.org/10.1016/S0140-6736(17)30058-2.

Choi, Shing Wan, and Paul F O’Reilly. 2019. “PRSice-2: Polygenic Risk Score Software for Biobank-Scale Data.” *GigaScience* 8 (7): giz082. https://doi.org/10.1093/gigascience/giz082.

Christensen, Brock C., E. Andres Houseman, Carmen J. Marsit, Shichun Zheng, Margaret R. Wrensch, Joseph L. Wiemels, Heather H. Nelson, et al. 2009. “Aging and Environmental Exposures Alter Tissue-Specific DNA Methylation Dependent upon CpG Island Context.” *PLoS Genetics* 5 (8): e1000602. https://doi.org/10.1371/journal.pgen.1000602.

Chuvin, Nicolas, David F. Vincent, Roxane M. Pommier, Lindsay B. Alcaraz, Johann Gout, Cassandre Caligaris, Karam Yacoub, et al. 2017. “Acinar-to-Ductal Metaplasia Induced by Transforming Growth Factor Beta Facilitates KRASG12D-Driven Pancreatic Tumorigenesis.” *Cellular and Molecular Gastroenterology and Hepatology* 4 (2): 263–82. https://doi.org/10.1016/j.jcmgh.2017.05.005.

Corkey, Barbara E., Jude T. Deeney, and Matthew J. Merrins. 2021. “What Regulates Basal Insulin Secretion and Causes Hyperinsulinemia?” *Diabetes* 70 (10): 2174–82. https://doi.org/10.2337/dbi21-0009.

Cuomo, Francesca, Carmela Dell’Aversana, Teresa Chioccarelli, Veronica Porreca, Francesco Manfrevola, Chiara Papulino, Vincenzo Carafa, Rosaria Benedetti, Lucia Altucci, and Gilda Cobellis. 2022. “HIF3A Inhibition Triggers Browning of White Adipocytes via Metabolic Rewiring.” *Frontiers in Cell and Developmental Biology* 9 (January). https://doi.org/10.3389/fcell.2021.740203.

Dayeh, Tasnim, Petr Volkov, Sofia Salö, Elin Hall, Emma Nilsson, Anders H. Olsson, Clare L. Kirkpatrick, et al. 2014. “Genome-Wide DNA Methylation Analysis of Human Pancreatic Islets from Type 2 Diabetic and Non-Diabetic Donors Identifies Candidate Genes That Influence Insulin Secretion.” *PLoS Genetics* 10 (3). https://doi.org/10.1371/journal.pgen.1004160.

Delaneau, Olivier, Halit Ongen, Andrew A. Brown, Alexandre Fort, Nikolaos I. Panousis, and Emmanouil T. Dermitzakis. 2017. “A Complete Tool Set for Molecular QTL Discovery and Analysis.” *Nature Communications* 8 (1): 15452. https://doi.org/10.1038/ncomms15452.

Demerath, Ellen W., Weihua Guan, Megan L. Grove, Stella Aslibekyan, Michael Mendelson, Yi-Hui Zhou, Åsa K. Hedman, et al. 2015. “Epigenome-Wide Association Study (EWAS) of BMI, BMI Change and Waist Circumference in African American Adults Identifies Multiple Replicated Loci.” *Human Molecular Genetics* 24 (15): 4464–79. https://doi.org/10.1093/hmg/ddv161.

DiCorpo, Daniel, Jessica LeClair, Joanne B. Cole, Chloé Sarnowski, Fariba Ahmadizar, Lawrence F. Bielak, Anneke Blokstra, et al. 2022. “Type 2 Diabetes Partitioned Polygenic Scores Associate With Disease Outcomes in 454,193 Individuals Across 13 Cohorts.” *Diabetes Care* 45 (3): 674–83. https://doi.org/10.2337/dc21-1395.

Dolai, Subhankar, Li Xie, Dan Zhu, Tao Liang, Tairan Qin, Huanli Xie, Youhou Kang, Edwin R. Chapman, and Herbert Y. Gaisano. 2016. “Synaptotagmin-7 Functions to Replenish Insulin Granules for Exocytosis in Human Islet β-Cells.” *Diabetes* 65 (7): 1962–76. https://doi.org/10.2337/db15-1436.

Dong, Xiaocheng, Sunmin Park, Xueying Lin, Kyle Copps, Xianjin Yi, and Morris F. White. 2006. “Irs1 and Irs2 Signaling Is Essential for Hepatic Glucose Homeostasis and Systemic Growth.” *The Journal of Clinical Investigation* 116 (1): 101–14. https://doi.org/10.1172/JCI25735.

Du, Pan, Xiao Zhang, Chiang-Ching Huang, Nadereh Jafari, Warren A. Kibbe, Lifang Hou, and Simon M. Lin. 2010. “Comparison of Beta-Value and M-Value Methods for Quantifying Methylation Levels by Microarray Analysis.” *BMC Bioinformatics* 11 (1): 587. https://doi.org/10.1186/1471-2105-11-587.

Egozi, Adi, Keren Bahar Halpern, Lydia Farack, Hagar Rotem, and Shalev Itzkovitz. 2020. “Zonation of Pancreatic Acinar Cells in Diabetic Mice.” *Cell Reports* 32 (7): 108043. https://doi.org/10.1016/j.celrep.2020.108043.

Eizirik, Décio L., Lorenzo Pasquali, and Miriam Cnop. 2020. “Pancreatic β-Cells in Type 1 and Type 2 Diabetes Mellitus: Different Pathways to Failure.” *Nature Reviews Endocrinology* 16 (7): 349–62. https://doi.org/10.1038/s41574-020-0355-7.

Escher, Genevieve, Zygmunt Krozowski, Kevin D. Croft, and Dmitri Sviridov. 2003. "Expression of Sterol 27-Hydroxylase (CYP27A1) Enhances Cholesterol Efflux *." *Journal of Biological Chemistry* 278 (13): 11015–19. https://doi.org/10.1074/jbc.M212780200.

Fafián-Labora, Juan Antonio, and Ana O'Loghlen. 2020. "Classical and Nonclassical Intercellular Communication in Senescence and Ageing." *Trends in Cell Biology* 30 (8): 628–39. https://doi.org/10.1016/j.tcb.2020.05.003.

Fajans, Stefan S., Graeme I. Bell, and Kenneth S. Polonsky. 2001. "Molecular Mechanisms and Clinical Pathophysiology of Maturity-Onset Diabetes of the Young." *New England Journal of Medicine* 345 (13): 971–80. https://doi.org/10.1056/NEJMra002168.

Fazeli, Pouneh K., Hang Lee, and Matthew L. Steinhauser. 2020. "Aging Is a Powerful Risk Factor for Type 2 Diabetes Mellitus Independent of Body Mass Index." *Gerontology* 66 (2): 209–10. https://doi.org/10.1159/000501745.

Fazeli, Pouneh K., Hang Lee, and Matthew L. Steinhauser. 2019. "Aging Is a Powerful Risk Factor for Type 2 Diabetes Mellitus Independent of Body Mass Index." *Gerontology* 66 (2): 209–10. https://doi.org/10.1159/000501745.

Fishilevich, Simon, Ron Nudel, Noa Rappaport, Rotem Hadar, Inbar Plaschkes, Tsippi Iny Stein, Naomi Rosen, et al. 2017. "GeneHancer: Genome-Wide Integration of Enhancers and Target Genes in GeneCards." *Database: The Journal of Biological Databases and Curation* 2017 (January):bax028. https://doi.org/10.1093/database/bax028.

Flannick, Jason, Gudmar Thorleifsson, Nicola L. Beer, Suzanne B. R. Jacobs, Niels Grarup, Noël P. Burtt, Anubha Mahajan, et al. 2014. "Loss-of-Function Mutations in SLC30A8 Protect against Type 2 Diabetes." *Nature Genetics* 46 (4): 357–63. https://doi.org/10.1038/ng.2915.

Florath, Ines, Katja Butterbach, Heiko Müller, Melanie Bewerunge-Hudler, and Hermann Brenner. 2014. “Cross-Sectional and Longitudinal Changes in DNA Methylation with Age: An Epigenome-Wide Analysis Revealing over 60 Novel Age-Associated CpG Sites.” *Human Molecular Genetics* 23 (5): 1186–1201. https://doi.org/10.1093/hmg/ddt531.

Flores, Kevin B., Florian Wolschin, and Gro V. Amdam. 2013. “The Role of Methylation of DNA in Environmental Adaptation.” *Integrative and Comparative Biology* 53 (2): 359–72. https://doi.org/10.1093/icb/ict019.

Fonseca, Vivian A. 2009. “Defining and Characterizing the Progression of Type 2 Diabetes.” *Diabetes Care* 32 (Suppl 2): S151–56. https://doi.org/10.2337/dc09-S301.

Frisdal, Eric, and Wilfried Le Goff. 2015. “Adipose ABCG1: A Potential Therapeutic Target in Obesity?” *Adipocyte* 4 (4): 315–18. https://doi.org/10.1080/21623945.2015.1023491.

Froguel, Philippe, Habib Zouali, Nathalie Vionnet, Gilberto Velho, Martine Vaxillaire, Fang Sun, Suzanne Lesage, et al. 1993. “Familial Hyperglycemia Due to Mutations in Glucokinase – Definition of a Subtype of Diabetes Mellitus.” *New England Journal of Medicine* 328 (10): 697–702. https://doi.org/10.1056/NEJM199303113281005.

Fu, Jianyang, Xiaoqing Dai, Gregory Plummer, Kunimasa Suzuki, Austin Bautista, John M. Githaka, Laura Senior, et al. 2017. “Kv2.1 Clustering Contributes to Insulin Exocytosis and Rescues Human β-Cell Dysfunction.” *Diabetes* 66 (7): 1890–1900. https://doi.org/10.2337/db16-1170.

Fulea, R. C., L. Reynard, D. Young, and G. Bou-Gharios. 2021. “FHL2 Promoter DNA Methylation Increases with Chronological Age in Joint Tissues and Impacts Target Gene Expression.” *Osteoarthritis and Cartilage* 29 (April):S310. https://doi.org/10.1016/j.joca.2021.02.409.

Galicia-Garcia, Unai, Asier Benito-Vicente, Shifa Jebari, Asier Larrea-Sebal, Haziq Siddiqi, Kepa B. Uribe, Helena Ostolaza, and César Martín. 2020. “Pathophysiology of

Type 2 Diabetes Mellitus." *International Journal of Molecular Sciences* 21 (17): 6275. https://doi.org/10.3390/ijms21176275.

Ge, Tian, Marguerite R. Irvin, Amit Patki, Vinodh Srinivasasainagendra, Yen-Feng Lin, Hemant K. Tiwari, Nicole D. Armstrong, et al. 2022. "Development and Validation of a Trans-Ancestry Polygenic Risk Score for Type 2 Diabetes in Diverse Populations." *Genome Medicine* 14 (1): 70. https://doi.org/10.1186/s13073-022-01074-2.

George, Carolyn M., and Eric Alani. 2012. "Multiple Cellular Mechanisms Prevent Chromosomal Rearrangements Involving Repetitive DNA." *Critical Reviews in Biochemistry and Molecular Biology* 47 (3): 297–313. https://doi.org/10.3109/10409238.2012.675644.

Giambartolomei, Claudia, Damjan Vukcevic, Eric E. Schadt, Lude Franke, Aroon D. Hingorani, Chris Wallace, and Vincent Plagnol. 2014. "Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics." *PLOS Genetics* 10 (5): e1004383. https://doi.org/10.1371/journal.pgen.1004383.

Gloyn, Anna L, Michael N Weedon, Katharine R Owen, Martina J Turner, Bridget A Knight, Graham Hitman, Mark Walker, et al. 2003. "Large-Scale Association Studies of Variants in Genes Encoding the Pancreatic Beta-Cell KATP Channel Subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) Confirm That the KCNJ11 E23K Variant Is Associated with Type 2 Diabetes." *Diabetes* 52 (2): 568–72. https://doi.org/10.2337/diabetes.52.2.568.

Gotter, Anthony L., Christine Suppa, and Beverly S. Emanuel. 2007. "Mammalian TIMELESS and Tipin Are Evolutionarily Conserved Replication Fork-Associated Factors." *Journal of Molecular Biology* 366 (1): 36–52. https://doi.org/10.1016/j.jmb.2006.10.097.

Graham, Sarah E., Shoa L. Clarke, Kuan-Han H. Wu, Stavroula Kanoni, Greg J. M. Zajac, Shweta Ramdas, Ida Surakka, et al. 2021. "The Power of Genetic Diversity in Genome-

Wide Association Studies of Lipids." *Nature* 600 (7890): 675–79. https://doi.org/10.1038/s41586-021-04064-3.

Grewal, Shiv I. S. 2023. "The Molecular Basis of Heterochromatin Assembly and Epigenetic Inheritance." *Molecular Cell* 83 (11): 1767–85. https://doi.org/10.1016/j.molcel.2023.04.020.

Grisan, Francesca, Martina Spacci, Carlotta Paoli, Andrea Costamagna, Marco Fantuz, Miriam Martini, Konstantinos Lefkimmiatis, and Alessandro Carrer. 2021. "Cholesterol Activates Cyclic AMP Signaling in Metaplastic Acinar Cells." *Metabolites* 11 (3): 141. https://doi.org/10.3390/metabo11030141.

Hall, Elin, Tasnim Dayeh, Clare L. Kirkpatrick, Claes B. Wollheim, Marloes Dekker Nitert, and Charlotte Ling. 2013. "DNA Methylation of the Glucagon-like Peptide 1 Receptor (GLP1R) in Human Pancreatic Islets." *BMC Medical Genetics* 14 (1): 76. https://doi.org/10.1186/1471-2350-14-76.

Harris, Megan T., Syed Saad Hussain, Candice M. Inouye, Anna M. Castle, and J. David Castle. 2018. "Reinterpretation of the Localization of the ATP Binding Cassette Transporter ABCG1 in Insulin-Secreting Cells and Insights Regarding Its Trafficking and Function." *PLOS ONE* 13 (9): e0198383. https://doi.org/10.1371/journal.pone.0198383.

Hastuti, Pramudji, and Maya Uzia Beandrade. 2022. "Epigenetic in DNA Methylation and Metabolic Syndrome."

Hawe, Johann S., Rory Wilson, Katharina T. Schmid, Li Zhou, Lakshmi Narayanan Lakshmanan, Benjamin C. Lehne, Brigitte Kühnel, et al. 2022. "Genetic Variation Influencing DNA Methylation Provides Insights into Molecular Mechanisms Regulating Genomic Function." *Nature Genetics* 54 (1): 18–29. https://doi.org/10.1038/s41588-021-00969-x.

Hecker, Nikolai, Virag Sharma, and Michael Hiller. 2019. "Convergent Gene Losses Illuminate Metabolic and Physiological Changes in Herbivores and Carnivores."

Proceedings of the National Academy of Sciences of the United States of America 116 (8): 3036–41. https://doi.org/10.1073/pnas.1818504116.

Heylen, Line, Bernard Thienpont, Pieter Busschaert, Ben Sprangers, Dirk Kuypers, Matthieu Moisse, Evelyne Lerut, Diether Lambrechts, and Maarten Naesens. 2019. "Age-Related Changes in DNA Methylation Affect Renal Histology and Post-Transplant Fibrosis." *Kidney International* 96 (5): 1195–1204. https://doi.org/10.1016/j.kint.2019.06.018.

Home, Resources, Living with diabetes, Acknowledgement, FAQs, Contact, and Privacy Policy. n.d. "IDF Diabetes Atlas." Accessed August 3, 2024. https://diabetesatlas.org/.

Horvath, Steve. 2013. "DNA Methylation Age of Human Tissues and Cell Types." *Genome Biology* 14 (10): 3156. https://doi.org/10.1186/gb-2013-14-10-r115.

Horvath, Steve, and Kenneth Raj. 2018. "DNA Methylation-Based Biomarkers and the Epigenetic Clock Theory of Ageing." *Nature Reviews Genetics* 19 (6): 371–84. https://doi.org/10.1038/s41576-018-0004-3.

Horvath, Steve, and Eric J Topol. 2024. "Digitising the Ageing Process with Epigenetic Clocks." *The Lancet* 404 (10451): 423. https://doi.org/10.1016/S0140-6736(24)01554-X.

Houseman, E. Andres, Molly L. Kile, David C. Christiani, Tan A. Ince, Karl T. Kelsey, and Carmen J. Marsit. 2016. "Reference-Free Deconvolution of DNA Methylation Data and Mediation by Cell Composition Effects." *BMC Bioinformatics* 17 (1): 259. https://doi.org/10.1186/s12859-016-1140-4.

Hu, Ming, Inês Cebola, Gaelle Carrat, Shuying Jiang, Sameena Nawaz, Amna Khamis, Mickaël Canouil, et al. 2021. "Chromatin 3D Interaction Analysis of the STARD10 Locus Unveils FCHSD2 as a Regulator of Insulin Secretion." *Cell Reports* 34 (5). https://doi.org/10.1016/j.celrep.2021.108703.

Huan, Tianxiao, Roby Joehanes, Ci Song, Fen Peng, Yichen Guo, Michael Mendelson, Chen Yao, et al. 2019. “Genome-Wide Identification of DNA Methylation QTLs in Whole Blood Highlights Pathways for Cardiovascular Disease.” *Nature Communications* 10 (1): 4267. https://doi.org/10.1038/s41467-019-12228-z.

Huang, Xingjun, Guihua Liu, Jiao Guo, and Zhengquan Su. 2018. “The PI3K/AKT Pathway in Obesity and Type 2 Diabetes.” *International Journal of Biological Sciences* 14 (11): 1483–96. https://doi.org/10.7150/ijbs.27173.

Igarashi, Shinichi, Hiromu Suzuki, Takeshi Niinuma, Haruo Shimizu, Masanori Nojima, Hiroyuki Iwaki, Takayuki Nobuoka, et al. 2010. “A Novel Correlation between LINE-1 Hypomethylation and the Malignancy of Gastrointestinal Stromal Tumors.” *Clinical Cancer Research* 16 (21): 5114–23. https://doi.org/10.1158/1078-0432.CCR-10-0581.

Illingworth, Robert S., and Adrian P. Bird. 2009. “CpG Islands – ‘A Rough Guide.’” *FEBS Letters*, Prague Special Issue: Functional Genomics and Proteomics, 583 (11): 1713–20. https://doi.org/10.1016/j.febslet.2009.04.012.

“Infographic: The Global Cost Of Diabetes.” 2016. Statista Daily Data. November 14, 2016. https://www.statista.com/chart/6700/the-global-cost-of-diabetes.

International HapMap Consortium. 2003. “The International HapMap Project.” *Nature* 426 (6968): 789–96. https://doi.org/10.1038/nature02168.

Irizarry, Rafael A., Christine Ladd-Acosta, Bo Wen, Zhijin Wu, Carolina Montano, Patrick Onyango, Hengmi Cui, et al. 2009. “Genome-Wide Methylation Analysis of Human Colon Cancer Reveals Similar Hypo- and Hypermethylation at Conserved Tissue-Specific CpG Island Shores.” *Nature Genetics* 41 (2): 178–86. https://doi.org/10.1038/ng.298.

Iterson, Maarten van, Erik W. van Zwet, Bastiaan T. Heijmans, and the BIOS Consortium. 2017. “Controlling Bias and Inflation in Epigenome- and Transcriptome-Wide Association Studies Using the Empirical Null Distribution.” *Genome Biology* 18 (1): 19. https://doi.org/10.1186/s13059-016-1131-9.

Jansen, Rick J., Lin Tong, Maria Argos, Farzana Jasmine, Muhammad Rakibuz-Zaman, Golam Sarwar, Md. Tariqul Islam, et al. 2019. “The Effect of Age on DNA Methylation in Whole Blood among Bangladeshi Men and Women.” *BMC Genomics* 20 (1): 704. https://doi.org/10.1186/s12864-019-6039-9.

Jeltsch, Albert, and Renata Z. Jurkowska. 2014. “New Concepts in DNA Methylation.” *Trends in Biochemical Sciences* 39 (7): 310–18. https://doi.org/10.1016/j.tibs.2014.05.002.

Jia, Shiqi, Andranik Ivanov, Dinko Blasevic, Thomas Müller, Bettina Purfürst, Wei Sun, Wei Chen, Matthew N Poy, Nikolaus Rajewsky, and Carmen Birchmeier. 2015. “Insm1 Cooperates with Neurod1 and Foxa2 to Maintain Mature Pancreatic B-cell Function.” *The EMBO Journal* 34 (10): 1417–33. https://doi.org/10.15252/embj.201490819.

Jin, Bilian, Yajun Li, and Keith D. Robertson. 2011. “DNA Methylation.” *Genes & Cancer* 2 (6): 607–17. https://doi.org/10.1177/1947601910393957.

Jj, Wright, Eskaros A, Windon A, Bottino R, Jenkins R, Bradley Am, Aramandla R, et al. 2024. “Exocrine Pancreas in Type 1 and Type 2 Diabetes: Different Patterns of Fibrosis, Metaplasia, Angiopathy, and Adiposity.” *Diabetes* 73 (7). https://doi.org/10.2337/db23-0009.

Kalyakulina, Alena, Igor Yusipov, Maria Giulia Bacalini, Claudio Franceschi, Maria Vedunova, and Mikhail Ivanchenko. 2022. “Disease Classification for Whole-Blood DNA Methylation: Meta-Analysis, Missing Values Imputation, and XAI.” *GigaScience* 11 (January):giac097. https://doi.org/10.1093/gigascience/giac097.

Karamanou, Marianna, Athanase Protogerou, Gregory Tsoucalas, George Androutsos, and Effie Poulakou-Rebelakou. 2016. “Milestones in the History of Diabetes Mellitus: The Main Contributors.” *World Journal of Diabetes* 7 (1): 1–7. https://doi.org/10.4239/wjd.v7.i1.1.

Kasper, Dennis L., Anthony S. Fauci, Stephen L. Hauser, Dan L. Longo, J. Larry Jameson, and Joseph Loscalzo. 2018. *Harrison's Principles of Internal Medicine 20/E (Vol.1 & Vol.2) (Ebook)*. McGraw Hill Professional.

Keller, Maria, Stina Ingrid Alice Svensson, Kerstin Rohde-Zimmermann, Peter Kovacs, and Yvonne Böttcher. 2023. "Genetics and Epigenetics in Obesity: What Do We Know so Far?" *Current Obesity Reports* 12 (4): 482–501. https://doi.org/10.1007/s13679-023-00526-z.

Kelly, Shona J., and Mubarak Ismail. 2015. "Stress and Type 2 Diabetes: A Review of How Stress Contributes to the Development of Type 2 Diabetes." *Annual Review of Public Health* 36 (Volume 36, 2015): 441–62. https://doi.org/10.1146/annurev-publhealth-031914-122921.

Khamis, Amna, Mickaël Canouil, Afshan Siddiq, Hutokshi Crouch, Mario Falchi, Manon von Bulow, Florian Ehehalt, et al. 2019. "Laser Capture Microdissection of Human Pancreatic Islets Reveals Novel eQTLs Associated with Type 2 Diabetes." *Molecular Metabolism* 24 (June):98–107. https://doi.org/10.1016/j.molmet.2019.03.004.

Khan, Aziz, and Xuegong Zhang. 2016. "dbSUPER: A Database of Super-Enhancers in Mouse and Human Genome." *Nucleic Acids Research* 44 (D1): D164-171. https://doi.org/10.1093/nar/gkv1002.

Khera, Amit V., Mark Chaffin, Krishna G. Aragam, Mary E. Haas, Carolina Roselli, Seung Hoan Choi, Pradeep Natarajan, et al. 2018. "Genome-Wide Polygenic Scores for Common Diseases Identify Individuals with Risk Equivalent to Monogenic Mutations." *Nature Genetics* 50 (9): 1219–24. https://doi.org/10.1038/s41588-018-0183-z.

Kim, Soyeon, Zhongli Xu, Erick Forno, Yidi Qin, Hyun Jung Park, Molin Yue, Qi Yan, et al. 2023. "*Cis*- and *Trans*-eQTM Analysis Reveals Novel Epigenetic and Transcriptomic Immune Markers of Atopic Asthma in Airway Epithelium." *Journal of Allergy and Clinical Immunology* 152 (4): 887–98. https://doi.org/10.1016/j.jaci.2023.05.018.

Klosin, Adam, Eduard Casas, Cristina Hidalgo-Carcedo, Tanya Vavouri, and Ben Lehner. 2017. “Transgenerational Transmission of Environmental Information in C. Elegans.” *Science* 356 (6335): 320–23. https://doi.org/10.1126/science.aah6412.

Koning, A. P. Jason de, Wanjun Gu, Todd A. Castoe, Mark A. Batzer, and David D. Pollock. 2011. “Repetitive Elements May Comprise Over Two-Thirds of the Human Genome.” *PLOS Genetics* 7 (12): e1002384. https://doi.org/10.1371/journal.pgen.1002384.

Kreienkamp, Raymond J., Benjamin F. Voight, Anna L. Gloyn, and Miriam S. Udler. 2023. “Genetics of Type 2 Diabetes.” In *Diabetes in America*, edited by Jean M. Lawrence, Sarah Stark Casagrande, William H. Herman, Deborah J. Wexler, and William T. Cefalu. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). http://www.ncbi.nlm.nih.gov/books/NBK597726/.

Lagou, Vasiliki, Longda Jiang, Anna Ulrich, Liudmila Zudina, Karla Sofia Gutiérrez González, Zhanna Balkhiyarova, Alessia Faggian, et al. 2023. “GWAS of Random Glucose in 476,326 Individuals Provide Insights into Diabetes Pathophysiology, Complications and Treatment Stratification.” *Nature Genetics* 55 (9): 1448–61. https://doi.org/10.1038/s41588-023-01462-3.

Läll, Kristi, Reedik Mägi, Andrew Morris, Andres Metspalu, and Krista Fischer. 2017. “Personalized Risk Prediction for Type 2 Diabetes: The Potential of Genetic Risk Scores.” *Genetics in Medicine* 19 (3): 322–29. https://doi.org/10.1038/gim.2016.103.

Lander, Eric S., Lauren M. Linton, Bruce Birren, Chad Nusbaum, Michael C. Zody, Jennifer Baldwin, Keri Devon, et al. 2001. “Initial Sequencing and Analysis of the Human Genome.” *Nature* 409 (6822): 860–921. https://doi.org/10.1038/35057062.

Law, Julie A., and Steven E. Jacobsen. 2010. “Establishing, Maintaining and Modifying DNA Methylation Patterns in Plants and Animals.” *Nature Reviews Genetics* 11 (3): 204–20. https://doi.org/10.1038/nrg2719.

Le, Jiamei, Yilong Chen, Wei Yang, Ligong Chen, and Jianping Ye. 2024. “Metabolic Basis of Solute Carrier Transporters in Treatment of Type 2 Diabetes Mellitus.” *Acta Pharmaceutica Sinica B* 14 (2): 437–54. https://doi.org/10.1016/j.apsb.2023.09.004.

Lee, Su-In, Safiye Celik, Benjamin A. Logsdon, Scott M. Lundberg, Timothy J. Martins, Vivian G. Oehler, Elihu H. Estey, et al. 2018. “A Machine Learning Approach to Integrate Big Data for Precision Medicine in Acute Myeloid Leukemia.” *Nature Communications* 9 (1): 42. https://doi.org/10.1038/s41467-017-02465-5.

Leek, Jeffrey T., W. Evan Johnson, Hilary S. Parker, Elana J. Fertig, Andrew E. Jaffe, Yuqing Zhang, John D. Storey, and Leonardo Collado Torres. 2020. *Sva: Surrogate Variable Analysis*.

Leek, Jeffrey T., W. Evan Johnson, Hilary S. Parker, Andrew E. Jaffe, and John D. Storey. 2012. “The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments.” *Bioinformatics* 28 (6): 882–83. https://doi.org/10.1093/bioinformatics/bts034.

Leung, Po Sing. 2010. *The Renin-Angiotensin System: Current Research Progress in The Pancreas: The RAS in the Pancreas*. Vol. 690. Advances in Experimental Medicine and Biology. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-9060-7.

Li, En. 2002. “Chromatin Modification and Epigenetic Reprogramming in Mammalian Development.” *Nature Reviews Genetics* 3 (9): 662–73. https://doi.org/10.1038/nrg887.

Li, Shaomin, and Andrew M. Stern. 2022. “Bioactive Human Alzheimer Brain Soluble Aβ: Pathophysiology and Therapeutic Opportunities.” *Molecular Psychiatry* 27 (8): 3182–91. https://doi.org/10.1038/s41380-022-01589-5.

Li, Shu, and Keping Xie. 2022. “Ductal Metaplasia in Pancreas.” *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1877 (2): 188698. https://doi.org/10.1016/j.bbcan.2022.188698.

Li, Yan, Xiaohui Bian, Shuyi Wei, Meizhi He, and Yuelian Yang. 2019. “The Relationship between Pancreatic Cancer and Type 2 Diabetes: Cause and Consequence.” *Cancer Management and Research* 11 (September):8257–68. https://doi.org/10.2147/CMAR.S211972.

Li, Yinglu, Xiao Chen, and Chao Lu. 2021. “The Interplay between DNA and Histone Methylation: Molecular Mechanisms and Disease Implications.” *EMBO Reports* 22 (5): e51803. https://doi.org/10.15252/embr.202051803.

Ling, C. 2020. “Epigenetic Regulation of Insulin Action and Secretion – Role in the Pathogenesis of Type 2 Diabetes.” *Journal of Internal Medicine* 288 (2): 158–67. https://doi.org/10.1111/joim.13049.

Ling, C., S. Del Guerra, R. Lupi, T. Rönn, C. Granhall, H. Luthman, P. Masiello, P. Marchetti, L. Groop, and S. Del Prato. 2008. “Epigenetic Regulation of PPARGC1A in Human Type 2 Diabetic Islets and Effect on Insulin Secretion.” *Diabetologia* 51 (4): 615–22. https://doi.org/10.1007/s00125-007-0916-5.

Ling, Charlotte, and Tina Rönn. 2019. “Epigenetics in Human Obesity and Type 2 Diabetes.” *Cell Metabolism* 29 (5): 1028–44. https://doi.org/10.1016/j.cmet.2019.03.009.

Ling, Suping, Karen Brown, Joanne K. Miksza, Lynne Howells, Amy Morrison, Eyad Issa, Thomas Yates, Kamlesh Khunti, Melanie J. Davies, and Francesco Zaccardi. 2020. “Association of Type 2 Diabetes With Cancer: A Meta-Analysis With Bias Analysis for Unmeasured Confounding in 151 Cohorts Comprising 32 Million People.” *Diabetes Care* 43 (9): 2313–22. https://doi.org/10.2337/dc20-0204.

Ling, Wei, Yi Huang, Yan-Mei Huang, Rong-Rong Fan, Yi Sui, and Hai-Lu Zhao. 2020. “Global Trend of Diabetes Mortality Attributed to Vascular Complications, 2000–2016.” *Cardiovascular Diabetology* 19 (1): 182. https://doi.org/10.1186/s12933-020-01159-5.

Linnemann, Amelia K., Mieke Baan, and Dawn Belt Davis. 2014. “Pancreatic β-Cell Proliferation in Obesity.” *Advances in Nutrition* 5 (3): 278–88. https://doi.org/10.3945/an.113.005488.

Liu, Jun, Naoki Akanuma, Chengyang Liu, Ali Naji, Glenn A. Halff, William K. Washburn, Luzhe Sun, and Pei Wang. 2016. “TGF-B1 Promotes Acinar to Ductal Metaplasia of Human Pancreatic Acinar Cells.” *Scientific Reports* 6 (August):30904. https://doi.org/10.1038/srep30904.

Liu, Limei, Kazuaki Nagashima, Takao Yasuda, Yanjun Liu, Hai-rong Hu, Guang He, Bo Feng, et al. 2013. “Mutations in KCNJ11 Are Associated with the Development of Autosomal Dominant, Early-Onset Type 2 Diabetes.” *Diabetologia* 56 (12): 2609–18. https://doi.org/10.1007/s00125-013-3031-9.

Liu, Rui-Ke, Xu Lin, Zun Wang, Jonathan Greenbaum, Chuan Qiu, Chun-Ping Zeng, Yong-Yao Zhu, Jie Shen, and Hong-Wen Deng. 2021. “Identification of Novel Functional CpG-SNPs Associated with Type 2 Diabetes and Birth Weight.” *Aging (Albany NY)* 13 (7): 10619. https://doi.org/10.18632/aging.202828.

López-Otín, Carlos, Maria A. Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. 2023. “Hallmarks of Aging: An Expanding Universe.” *Cell* 186 (2): 243–78. https://doi.org/10.1016/j.cell.2022.11.001.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550. https://doi.org/10.1186/s13059-014-0550-8.

Lu, Ake T., Austin Quach, James G. Wilson, Alex P. Reiner, Abraham Aviv, Kenneth Raj, Lifang Hou, et al. 2019. “DNA Methylation GrimAge Strongly Predicts Lifespan and Healthspan.” *Aging (Albany NY)* 11 (2): 303. https://doi.org/10.18632/aging.101684.

Lu, Chunting, Qingtong Zhao, Dan Wang, Yunlu Feng, Lie Feng, Zejian Li, and Qiping Shi. 2022. “Rab35 Regulates Insulin Secretion via Phogrin in Pancreatic β Cells.”

Clinical and Experimental Pharmacology and Physiology 49 (1): 104–12. https://doi.org/10.1111/1440-1681.13581.

Luo, Xiumei, Cheng Luan, Jingqi Zhou, Yingying Ye, Wei Zhang, Ruchi Jain, Enming Zhang, and Ning Chen. 2024. "Glycolytic Enzyme Enolase-1 Regulates Insulin Gene Expression in Pancreatic β-Cell." *Biochemical and Biophysical Research Communications* 706 (April):149735. https://doi.org/10.1016/j.bbrc.2024.149735.

Macartney-Coxson, Donia, Miles C. Benton, Ray Blick, Richard S. Stubbs, Ronald D. Hagan, and Michael A. Langston. 2017. "Genome-Wide DNA Methylation Analysis Reveals Loci That Distinguish Different Types of Adipose Tissue in Obese Individuals." *Clinical Epigenetics* 9 (1): 48. https://doi.org/10.1186/s13148-017-0344-4.

Mahony, Caitlyn, and Colleen O'Ryan. 2021. "Convergent Canonical Pathways in Autism Spectrum Disorder from Proteomic, Transcriptomic and DNA Methylation Data." *International Journal of Molecular Sciences* 22 (19): 10757. https://doi.org/10.3390/ijms221910757.

Maina, Jared G., Zhanna Balkhiyarova, Arie Nouwen, Igor Pupko, Anna Ulrich, Mathilde Boissel, Amélie Bonnefond, et al. 2023. "Bidirectional Mendelian Randomization and Multiphenotype GWAS Show Causality and Shared Pathophysiology Between Depression and Type 2 Diabetes." *Diabetes Care* 46 (9): 1707–14. https://doi.org/10.2337/dc22-2373.

Mandla, Ravi, Kim Lorenz, Xianyong Yin, Ozvan Bocher, Alicia Huerta-Chagoya, Ana Luiza Arruda, Anthony Piron, et al. 2024. "Multi-Omics Characterization of Type 2 Diabetes Associated Genetic Variation." *medRxiv*, July, 2024.07.15.24310282. https://doi.org/10.1101/2024.07.15.24310282.

Mansell, Georgina, Tyler J. Gorrie-Stone, Yanchun Bao, Meena Kumari, Leonard S. Schalkwyk, Jonathan Mill, and Eilis Hannon. 2019. "Guidance for DNA Methylation Studies: Statistical Insights from the Illumina EPIC Array." *BMC Genomics* 20 (1): 366. https://doi.org/10.1186/s12864-019-5761-7.

Marstrand-Daucé, Louis, Diane Lorenzo, Anaïs Chassac, Pascal Nicole, Anne Couvelard, and Cécile Haumaitre. 2023. “Acinar-to-Ductal Metaplasia (ADM): On the Road to Pancreatic Intraepithelial Neoplasia (PanIN) and Pancreatic Cancer.” *International Journal of Molecular Sciences* 24 (12): 9946. https://doi.org/10.3390/ijms24129946.

McCarthy, Shane, Sayantan Das, Warren Kretzschmar, Olivier Delaneau, Andrew R Wood, Alexander Teumer, Hyun Min Kang, et al. 2016. “A Reference Panel of 64,976 Haplotypes for Genotype Imputation.” *Nature Genetics* 48 (10): 1279–83. https://doi.org/10.1038/ng.3643.

McCartney, Daniel L., Futao Zhang, Robert F. Hillary, Qian Zhang, Anna J. Stevenson, Rosie M. Walker, Mairead L. Bermingham, et al. 2019. “An Epigenome-Wide Association Study of Sex-Specific Chronological Ageing.” *Genome Medicine* 12 (1): 1. https://doi.org/10.1186/s13073-019-0693-z.

McClelland, Michael, and Robert Ivarie. 1982. “Asymmetrical Distribution of CpG in an ‘Average’ Mammalian Gene.” *Nucleic Acids Research* 10 (23): 7865–77. https://doi.org/10.1093/nar/10.23.7865.

McDonald, Oliver G., Hao Wu, Winston Timp, Akiko Doi, and Andrew P. Feinberg. 2011. “Genome-Scale Epigenetic Reprogramming during Epithelial-to-Mesenchymal Transition.” *Nature Structural & Molecular Biology* 18 (8): 867–74. https://doi.org/10.1038/nsmb.2084.

Mellenthin, Claudia, Vasile Daniel Balaban, Ana Dugic, and Stephane Cullati. 2022. “Risk Factors for Pancreatic Cancer in Patients with New-Onset Diabetes: A Systematic Review and Meta-Analysis.” *Cancers* 14 (19): 4684. https://doi.org/10.3390/cancers14194684.

Mello, Vanessa Derenji Ferreira de, Leena Pulkkinen, Marianne Lalli, Marjukka Kolehmainen, Jussi Pihlajamäki, and Matti Uusitupa. 2014. “DNA Methylation in

Obesity and Type 2 Diabetes." *Annals of Medicine* 46 (3): 103–13. https://doi.org/10.3109/07853890.2013.857259.

Meulebrouck, Sarah, Judith Merrheim, Gurvan Queniat, Cyril Bourouh, Mehdi Derhourhi, Mathilde Boissel, Xiaoyan Yi, et al. 2024. "Functional Genetics Reveals the Contribution of Delta Opioid Receptor to Type 2 Diabetes and Beta-Cell Function." *Nature Communications* 15 (1): 6627. https://doi.org/10.1038/s41467-024-51004-6.

Miguel-Escalada, Irene, Silvia Bonàs-Guarch, Inês Cebola, Joan Ponsa-Cobas, Julen Mendieta-Esteban, Goutham Atla, Biola M. Javierre, et al. 2019. "Human Pancreatic Islet Three-Dimensional Chromatin Architecture Provides Insights into the Genetics of Type 2 Diabetes." *Nature Genetics* 51 (7): 1137–48. https://doi.org/10.1038/s41588-019-0457-0.

Moore, Lisa D., Thuc Le, and Guoping Fan. 2013. "DNA Methylation and Its Basic Function." *Neuropsychopharmacology* 38 (1): 23–38. https://doi.org/10.1038/npp.2012.112.

Moran, Sebastian, Carles Arribas, and Manel Esteller. 2016. "Validation of a DNA Methylation Microarray for 850,000 CpG Sites of the Human Genome Enriched in Enhancer Sequences." *Epigenomics* 8 (3): 389–99. https://doi.org/10.2217/epi.15.114.

Muhammad, Syed Aun, Syeda Tahira Qousain Naqvi, Thanh Nguyen, Xiaogang Wu, Fahad Munir, Muhammad Babar Jamshed, and QiYu Zhang. 2021. "Cisplatin's Potential for Type 2 Diabetes Repositioning by Inhibiting *CDKN1A, FAS*, and *SESN1*." *Computers in Biology and Medicine* 135 (August):104640. https://doi.org/10.1016/j.compbiomed.2021.104640.

Nations Unies, ed. 2015. *World Population Ageing, 2015: Highlights*. ST-ESA-SER.A 368. New York: United Nations.

Neuhöfer, Patrick, Caitlin M. Roake, Stewart J. Kim, Ryan J. Lu, Robert B. West, Gregory W. Charville, and Steven E. Artandi. 2021. "Acinar Cell Clonal Expansion in

Pancreas Homeostasis and Carcinogenesis." *Nature* 597 (7878): 715–19. https://doi.org/10.1038/s41586-021-03916-2.

Nilsson, Emma, Ashok Matte, Alexander Perfilyev, Vanessa D. de Mello, Pirjo Käkelä, Jussi Pihlajamäki, and Charlotte Ling. 2015. "Epigenetic Alterations in Human Liver From Subjects With Type 2 Diabetes in Parallel With Reduced Folate Levels." *The Journal of Clinical Endocrinology & Metabolism* 100 (11): E1491–1501. https://doi.org/10.1210/jc.2015-3204.

Okano, Masaki, Daphne W. Bell, Daniel A. Haber, and En Li. 1999. "DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development." *Cell* 99 (3): 247–57. https://doi.org/10.1016/S0092-8674(00)81656-6.

Ong, Kanyin Liane, Lauryn K. Stafford, Susan A. McLaughlin, Edward J. Boyko, Stein Emil Vollset, Amanda E. Smith, Bronte E. Dalton, et al. 2023. "Global, Regional, and National Burden of Diabetes from 1990 to 2021, with Projections of Prevalence to 2050: A Systematic Analysis for the Global Burden of Disease Study 2021." *The Lancet* 402 (10397): 203–34. https://doi.org/10.1016/S0140-6736(23)01301-6.

Padilla-Martínez, Felipe, Francois Collin, Miroslaw Kwasniewski, and Adam Kretowski. 2020. "Systematic Review of Polygenic Risk Scores for Type 1 and Type 2 Diabetes." *International Journal of Molecular Sciences* 21 (5): 1703. https://doi.org/10.3390/ijms21051703.

Pandol, Stephen. 2011. *The Exocrine Pancreas*. Morgan & Claypool Publishers.

Pappalardo, Xena Giada, and Viviana Barra. 2021. "Losing DNA Methylation at Repetitive Elements and Breaking Bad." *Epigenetics & Chromatin* 14 (1): 25. https://doi.org/10.1186/s13072-021-00400-z.

Parameswaran, Gokul, and David W. Ray. 2022. "Sleep, Circadian Rhythms, and Type 2 Diabetes Mellitus." *Clinical Endocrinology* 96 (1): 12–20. https://doi.org/10.1111/cen.14607.

Parker, Emily D., Janice Lin, Troy Mahoney, Nwanneamaka Ume, Grace Yang, Robert A. Gabbay, Nuha A. ElSayed, and Raveendhara R. Bannuru. 2024. “Economic Costs of Diabetes in the U.S. in 2022.” *Diabetes Care* 47 (1): 26–43. https://doi.org/10.2337/dci23-0085.

Pasquali, Lorenzo, Kyle J. Gaulton, Santiago A. Rodríguez-Seguí, Loris Mularoni, Irene Miguel-Escalada, İldem Akerman, Juan J. Tena, et al. 2014. “Pancreatic Islet Enhancer Clusters Enriched in Type 2 Diabetes Risk-Associated Variants.” *Nature Genetics* 46 (2): 136–43. https://doi.org/10.1038/ng.2870.

Paz, Maria F., Susan Wei, Juan C. Cigudosa, Sandra Rodriguez-Perales, Miguel A. Peinado, Tim Hui-Ming Huang, and Manel Esteller. 2003. “Genetic Unmasking of Epigenetically Silenced Tumor Suppressor Genes in Colon Cancer Cells Deficient in DNA Methyltransferases.” *Human Molecular Genetics* 12 (17): 2209–19. https://doi.org/10.1093/hmg/ddg226.

Petersen, Max C., and Gerald I. Shulman. 2018. “Mechanisms of Insulin Action and Insulin Resistance.” *Physiological Reviews* 98 (4): 2133–2223. https://doi.org/10.1152/physrev.00063.2017.

Piron, Anthony, Florian Szymczak, Theodora Papadopoulou, Maria Inês Alvelos, Matthieu Defrance, Tom Lenaerts, Décio L. Eizirik, and Miriam Cnop. 2024. “RedRibbon: A New Rank–Rank Hypergeometric Overlap for Gene and Transcript Expression Signatures.” *Life Science Alliance* 7 (2). https://doi.org/10.26508/lsa.202302203.

Rachas, Antoine, Christelle Gastaldi-Ménager, Pierre Denis, Pauline Barthélémy, Panayotis Constantinou, Jérôme Drouin, Dimitri Lastier, et al. 2022. “The Economic Burden of Disease in France From the National Health Insurance Perspective.” *Medical Care* 60 (9): 655–64. https://doi.org/10.1097/MLR.0000000000001745.

Rauluseviciute, Ieva, Finn Drabløs, and Morten Beck Rye. 2020. “DNA Hypermethylation Associated with Upregulated Gene Expression in Prostate Cancer

Demonstrates the Diversity of Epigenetic Regulation." *BMC Medical Genomics* 13 (1): 6. https://doi.org/10.1186/s12920-020-0657-6.

Rechavi, Oded, Leah Houri-Ze'evi, Sarit Anava, Wee Siong Sho Goh, Sze Yen Kerk, Gregory J. Hannon, and Oliver Hobert. 2014. "Starvation-Induced Transgenerational Inheritance of Small RNAs in C. Elegans." *Cell* 158 (2): 277–87. https://doi.org/10.1016/j.cell.2014.06.020.

Rechavi, Oded, Gregory Minevich, and Oliver Hobert. 2011. "Transgenerational Inheritance of an Acquired Small RNA-Based Antiviral Response in C. Elegans." *Cell* 147 (6): 1248–56. https://doi.org/10.1016/j.cell.2011.10.042.

Reed, Josh, Stephen Bain, and Venkateswarlu Kanamarlapudi. 2021. "A Review of Current Trends with Type 2 Diabetes Epidemiology, Aetiology, Pathogenesis, Treatments and Future Perspectives." *Diabetes, Metabolic Syndrome and Obesity*, August. https://www.tandfonline.com/doi/abs/10.2147/DMSO.S319895.

Reynolds, Lindsay M., Jackson R. Taylor, Jingzhong Ding, Kurt Lohman, Craig Johnson, David Siscovick, Gregory Burke, et al. 2014. "Age-Related Variations in the Methylome Associated with Gene Expression in Human Monocytes and T Cells." *Nature Communications* 5 (November):5366. https://doi.org/10.1038/ncomms6366.

Ribel-Madsen, Rasmus, Mario F. Fraga, Stine Jacobsen, Jette Bork-Jensen, Ester Lara, Vincenzo Calvanese, Agustin F. Fernandez, et al. 2012. "Genome-Wide Analysis of DNA Methylation Differences in Muscle and Fat from Monozygotic Twins Discordant for Type 2 Diabetes." *PLOS ONE* 7 (12): e51302. https://doi.org/10.1371/journal.pone.0051302.

Ritchie, Matthew E., Belinda Phipson, Di Wu, Yifang Hu, Charity W. Law, Wei Shi, and Gordon K. Smyth. 2015. "Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies." *Nucleic Acids Research* 43 (7): e47. https://doi.org/10.1093/nar/gkv007.

Robin, Xavier, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. 2010. “pROC: Display and Analyze ROC Curves.” https://doi.org/10.32614/CRAN.package.pROC.

Röder, Pia V, Bingbing Wu, Yixian Liu, and Weiping Han. 2016. “Pancreatic Regulation of Glucose Homeostasis.” *Experimental & Molecular Medicine* 48 (3): e219. https://doi.org/10.1038/emm.2016.6.

Romer, Anthony I., Ruth A. Singer, Lina Sui, Dieter Egli, and Lori Sussel. 2019. “Murine Perinatal β-Cell Proliferation and the Differentiation of Human Stem Cell–Derived Insulin-Expressing Cells Require NEUROD1.” *Diabetes* 68 (12): 2259–71. https://doi.org/10.2337/db19-0117.

Rönn, Tina, Jones K. Ofori, Alexander Perfilyev, Alexander Hamilton, Karolina Pircs, Fabian Eichelmann, Sonia Garcia-Calzon, et al. 2023. “Genes with Epigenetic Alterations in Human Pancreatic Islets Impact Mitochondrial Function, Insulin Secretion, and Type 2 Diabetes.” *Nature Communications* 14 (1): 8040. https://doi.org/10.1038/s41467-023-43719-9.

Rudman, Najda, Olga Gornik, and Gordan Lauc. 2019. “Altered N-Glycosylation Profiles as Potential Biomarkers and Drug Targets in Diabetes.” *FEBS Letters* 593 (13): 1598–1615. https://doi.org/10.1002/1873-3468.13495.

Rudman, Najda, Simranjeet Kaur, Vesna Simunović, Domagoj Kifer, Dinko Šoić, Toma Keser, Tamara Štambuk, et al. 2023. “Integrated Glycomics and Genetics Analyses Reveal a Potential Role for N-Glycosylation of Plasma Proteins and IgGs, as Well as the Complement System, in the Development of Type 1 Diabetes.” *Diabetologia* 66 (6): 1071–83. https://doi.org/10.1007/s00125-023-05881-z.

Saberzadeh-Ardestani, Bahar, Razieh Karamzadeh, Mohsen Basiri, Ensiyeh Hajizadeh-Saffar, Aisan Farhadi, A.M. James Shapiro, Yaser Tahamtani, and Hossein Baharvand. 2018. “Type 1 Diabetes Mellitus: Cellular and Molecular Pathophysiology at A

Glance.” *Cell Journal (Yakhteh)* 20 (3): 294–301. https://doi.org/10.22074/cellj.2018.5513.

Salameh, Yasmeen, Yosra Bejaoui, and Nady El Hajj. 2020. “DNA Methylation Biomarkers in Aging and Age-Related Diseases.” *Frontiers in Genetics* 11 (March). https://doi.org/10.3389/fgene.2020.00171.

Sansbury, F. H., S. E. Flanagan, J. A. L. Houghton, F. L. Shuixian Shen, A. M. S. Al-Senani, A. M. Habeb, M. Abdullah, A. Kariminejad, S. Ellard, and A. T. Hattersley. 2012. “SLC2A2 Mutations Can Cause Neonatal Diabetes, Suggesting GLUT2 May Have a Role in Human Insulin Secretion.” *Diabetologia* 55 (9): 2381–85. https://doi.org/10.1007/s00125-012-2595-0.

Sarnowski, Chloé, Tianxiao Huan, Yiyi Ma, Roby Joehanes, Alexa Beiser, Charles S. DeCarli, Nancy L. Heard-Costa, et al. 2023. “Multi-Tissue Epigenetic Analysis Identifies Distinct Associations Underlying Insulin Resistance and Alzheimer’s Disease at CPT1A Locus.” *Clinical Epigenetics* 15 (1): 173. https://doi.org/10.1186/s13148-023-01589-4.

Schrader, Silja, Alexander Perfilyev, Emma Ahlqvist, Leif Groop, Allan Vaag, Mats Martinell, Sonia García-Calzón, and Charlotte Ling. 2022. “Novel Subgroups of Type 2 Diabetes Display Different Epigenetic Patterns That Associate With Future Diabetic Complications.” *Diabetes Care* 45 (7): 1621–30. https://doi.org/10.2337/dc21-2489.

Schumann, Tina, Jörg König, Christine Henke, Diana M. Willmes, Stefan R. Bornstein, Jens Jordan, Martin F. Fromm, and Andreas L. Birkenfeld. 2020. “Solute Carrier Transporters as Potential Targets for the Treatment of Metabolic Disease.” Edited by Martin C. Michel. *Pharmacological Reviews* 72 (1): 343–79. https://doi.org/10.1124/pr.118.015735.

Secco, Blandine, Kevin Saitoski, Karima Drareni, Antoine Soprani, Severine Pechberty, Latif Rachdi, Nicolas Venteclef, and Raphaël Scharfmann. 2022. “Loss of Human Beta Cell Identity in a Reconstructed Omental Stromal Cell Environment.” *Cells* 11 (6): 924. https://doi.org/10.3390/cells11060924.

Seino, S, T Iwanaga, K Nagashima, and T Miki. 2000. “Diverse Roles of K(ATP) Channels Learned from Kir6.2 Genetically Engineered Mice.” *Diabetes* 49 (3): 311–18. https://doi.org/10.2337/diabetes.49.3.311.

Shapiro, James A., and Richard von Sternberg. 2005. “Why Repetitive DNA Is Essential to Genome Function.” *Biological Reviews* 80 (2): 227–50. https://doi.org/10.1017/S1464793104006657.

Shi, G., D. DiRenzo, C. Qu, D. Barney, D. Miley, and S. F. Konieczny. 2013. “Maintenance of Acinar Cell Organization Is Critical to Preventing Kras-Induced Acinar-Ductal Metaplasia.” *Oncogene* 32 (15): 1950–58. https://doi.org/10.1038/onc.2012.210.

Shojima, Nobuhiro, and Toshimasa Yamauchi. 2023. “Progress in Genetics of Type 2 Diabetes and Diabetic Complications.” *Journal of Diabetes Investigation* 14 (4): 503–15. https://doi.org/10.1111/jdi.13970.

Skoczek, Dawid, Józef Dulak, and Neli Kachamakova-Trojanowska. 2021. “Maturity Onset Diabetes of the Young—New Approaches for Disease Modelling.” *International Journal of Molecular Sciences* 22 (14): 7553. https://doi.org/10.3390/ijms22147553.

Sladek, Robert, Ghislain Rocheleau, Johan Rung, Christian Dina, Lishuang Shen, David Serre, Philippe Boutin, et al. 2007. “A Genome-Wide Association Study Identifies Novel Risk Loci for Type 2 Diabetes.” *Nature* 445 (7130): 881–85. https://doi.org/10.1038/nature05616.

Slieker, Roderick C., Caroline L. Relton, Tom R. Gaunt, P. Eline Slagboom, and Bastiaan T. Heijmans. 2018. “Age-Related DNA Methylation Changes Are Tissue-Specific with ELOVL2 Promoter Methylation as Exception.” *Epigenetics & Chromatin* 11 (1): 25. https://doi.org/10.1186/s13072-018-0191-3.

Smith, Zachary D., Sara Hetzel, and Alexander Meissner. 2024. “DNA Methylation in Mammalian Development and Disease.” *Nature Reviews Genetics*, August, 1–24. https://doi.org/10.1038/s41576-024-00760-8.

Solimena, Michele, Anke M. Schulte, Lorella Marselli, Florian Ehehalt, Daniela Richter, Manuela Kleeberg, Hassan Mziaut, et al. 2018a. “Systems Biology of the IMIDIA Biobank from Organ Donors and Pancreatectomised Patients Defines a Novel Transcriptomic Signature of Islets from Individuals with Type 2 Diabetes.” *Diabetologia* 61 (3): 641–57. https://doi.org/10.1007/s00125-017-4500-3.

———. 2018b. “Systems Biology of the IMIDIA Biobank from Organ Donors and Pancreatectomised Patients Defines a Novel Transcriptomic Signature of Islets from Individuals with Type 2 Diabetes.” *Diabetologia* 61 (3): 641–57. https://doi.org/10.1007/s00125-017-4500-3.

Spracklen, Cassandra N., Jinxiu Shi, Swarooparani Vadlamudi, Ying Wu, Meng Zou, Chelsea K. Raulerson, James P. Davis, et al. 2018. “Identification and Functional Analysis of Glycemic Trait Loci in the China Health and Nutrition Survey.” *PLOS Genetics* 14 (4): e1007275. https://doi.org/10.1371/journal.pgen.1007275.

Spruijt, Cornelia G., Felix Gnerlich, Arne H. Smits, Toni Pfaffeneder, Pascal W. T. C. Jansen, Christina Bauer, Martin Münzel, et al. 2013. “Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives.” *Cell* 152 (5): 1146–59. https://doi.org/10.1016/j.cell.2013.02.004.

Stancill, Jennifer S., Jean-Philippe Cartailler, Hannah W. Clayton, James T. O’Connor, Matthew T. Dickerson, Prasanna K. Dadi, Anna B. Osipovich, David A. Jacobson, and Mark A. Magnuson. 2017. “Chronic β-Cell Depolarization Impairs β-Cell Identity by Disrupting a Network of Ca2+-Regulated Genes.” *Diabetes* 66 (8): 2175–87. https://doi.org/10.2337/db16-1355.

Stefansson, Olafur Andri, Brynja Dogg Sigurpalsdottir, Solvi Rognvaldsson, Gisli Hreinn Halldorsson, Kristinn Juliusson, Gardar Sveinbjornsson, Bjarni Gunnarsson, et al. 2024. “The Correlation between CpG Methylation and Gene Expression Is Driven by Sequence Variants.” *Nature Genetics* 56 (8): 1624–31. https://doi.org/10.1038/s41588-024-01851-2.

Stratton, Irene M., Amanda I. Adler, H. Andrew W. Neil, David R. Matthews, Susan E. Manley, Carole A. Cull, David Hadden, Robert C. Turner, and Rury R. Holman. 2000. "Association of Glycaemia with Macrovascular and Microvascular Complications of Type 2 Diabetes (UKPDS 35): Prospective Observational Study." *BMJ* 321 (7258): 405–12. https://doi.org/10.1136/bmj.321.7258.405.

Suckale, Jakob, and Michele Solimena. 2008. "Pancreas Islets in Metabolic Signaling - Focus on the β-Cell." *Nature Precedings*, June, 1–1. https://doi.org/10.1038/npre.2008.1724.2.

Suhre, K., and S. Zaghlool. 2021. "Connecting the Epigenome, Metabolome and Proteome for a Deeper Understanding of Disease." *Journal of Internal Medicine* 290 (3): 527–48. https://doi.org/10.1111/joim.13306.

Sun, Jianping, Yingye Zheng, and Li Hsu. 2013. "A Unified Mixed-Effects Model for Rare-Variant Association in Sequencing Studies." *Genetic Epidemiology* 37 (4): 334–44. https://doi.org/10.1002/gepi.21717.

Suzuki, Ken, Konstantinos Hatzikotoulas, Lorraine Southam, Henry J. Taylor, Xianyong Yin, Kim M. Lorenz, Ravi Mandla, et al. 2024. "Genetic Drivers of Heterogeneity in Type 2 Diabetes Pathophysiology." *Nature* 627 (8003): 347–57. https://doi.org/10.1038/s41586-024-07019-6.

Tajuddin, Salman M., Dena G. Hernandez, Brian H. Chen, Nicole Noren Hooten, Nicolle A. Mode, Mike A. Nalls, Andrew B. Singleton, et al. 2019. "Novel Age-Associated DNA Methylation Changes and Epigenetic Age Acceleration in Middle-Aged African Americans and Whites." *Clinical Epigenetics* 11 (1): 119. https://doi.org/10.1186/s13148-019-0722-1.

Tak, Yu Gyoung, and Peggy J. Farnham. 2015. "Making Sense of GWAS: Using Epigenomics and Genome Engineering to Understand the Functional Relevance of SNPs in Non-Coding Regions of the Human Genome." *Epigenetics & Chromatin* 8 (1): 57. https://doi.org/10.1186/s13072-015-0050-4.

Tan, Jiaxin, Yu You, Fei Guo, Jianhua Xu, Haisu Dai, and Ping Bie. 2017. “Association of Elevated Risk of Pancreatic Cancer in Diabetic Patients: A Systematic Review and Meta-Analysis.” *Oncology Letters* 13 (3): 1247–55. https://doi.org/10.3892/ol.2017.5586.

Tarkhov, Andrei E., Thomas Lindstrom-Vautrin, Sirui Zhang, Kejun Ying, Mahdi Moqri, Bohan Zhang, Alexander Tyshkovskiy, Orr Levy, and Vadim N. Gladyshev. 2024. “Nature of Epigenetic Aging from a Single-Cell Perspective.” *Nature Aging* 4 (6): 854–70. https://doi.org/10.1038/s43587-024-00616-0.

Timp, Winston, and Andrew P. Feinberg. 2013. “Cancer as a Dysregulated Epigenome Allowing Cellular Growth Advantage at the Expense of the Host.” *Nature Reviews Cancer* 13 (7): 497–510. https://doi.org/10.1038/nrc3486.

Toubiana, Shir, Guillaume Velasco, Adi Chityat, Angela M Kaindl, Noam Hershtig, Aya Tzur-Gilat, Claire Francastel, and Sara Selig. 2018. “Subtelomeric Methylation Distinguishes between Subtypes of Immunodeficiency, Centromeric Instability and Facial Anomalies Syndrome.” *Human Molecular Genetics* 27 (20): 3568–81. https://doi.org/10.1093/hmg/ddy265.

Tudurí, Eva, Sergi Soriano, Lucía Almagro, Eduard Montanya, Paloma Alonso-Magdalena, Ángel Nadal, and Ivan Quesada. 2022. “The Pancreatic β-Cell in Ageing: Implications in Age-Related Diabetes.” *Ageing Research Reviews* 80 (September):101674. https://doi.org/10.1016/j.arr.2022.101674.

Turner, Robert C, Andrew T Hattersley, Joanne T Shaw, and Jonathan C Levy. 1995. “Type II Diabetes: Clinical Aspects of Molecular Biological Studies.” *Diabetes* 44 (1): 1–10. https://doi.org/10.2337/diab.44.1.1.

Villicaña, Sergio, and Jordana T. Bell. 2021. “Genetic Impacts on DNA Methylation: Research Findings and Future Perspectives.” *Genome Biology* 22 (1): 127. https://doi.org/10.1186/s13059-021-02347-6.

Volkmar, Michael, Sarah Dedeurwaerder, Daniel A Cunha, Matladi N Ndlovu, Matthieu Defrance, Rachel Deplus, Emilie Calonne, et al. 2012. “DNA Methylation Profiling Identifies Epigenetic Dysregulation in Pancreatic Islets from Type 2 Diabetic Patients.” *The EMBO Journal* 31 (6): 1405–26. https://doi.org/10.1038/emboj.2011.503.

Waddlngton, C. H. 1942. “The Pupal Contraction as an Epigenetic Crisis in Drosophila.” *Proceedings of the Zoological Society of London* A111 (3–4): 181–88. https://doi.org/10.1111/j.1469-7998.1942.tb08480.x.

Wagner, Franziska, Irina Ruf, Thomas Lehmann, Rebecca Hofmann, Sylvia Ortmann, Christian Schiffmann, Michael Hiller, Clara Stefen, and Heiko Stuckas. 2022. “Reconstruction of Evolutionary Changes in Fat and Toxin Consumption Reveals Associations with Gene Losses in Mammals: A Case Study for the Lipase Inhibitor PNLIPRP1 and the Xenobiotic Receptor NR1I3.” *Journal of Evolutionary Biology* 35 (2): 225–39. https://doi.org/10.1111/jeb.13970.

Wahl, Simone, Alexander Drong, Benjamin Lehne, Marie Loh, William R. Scott, Sonja Kunze, Pei-Chien Tsai, et al. 2017. “Epigenome-Wide Association Study of Body Mass Index, and the Adverse Outcomes of Adiposity.” *Nature* 541 (7635): 81–86. https://doi.org/10.1038/nature20784.

Walker, John T, Diane C Saunders, Marcela Brissova, and Alvin C Powers. 2021. “The Human Islet: Mini-Organ With Mega-Impact.” *Endocrine Reviews* 42 (5): 605–57. https://doi.org/10.1210/endrev/bnab010.

Wang, Tao, Rong Zhang, Xiaojing Ma, Shiyun Wang, Zhen He, Yeping Huang, Bo Xu, et al. 2018. “Causal Association of Overall Obesity and Abdominal Obesity with Type 2 Diabetes: A Mendelian Randomization Analysis.” *Obesity* 26 (5): 934–42. https://doi.org/10.1002/oby.22167.

Wang, Xiaoling, Haidong Zhu, Harold Snieder, Shaoyong Su, David Munn, Gregory Harshfield, Bernard L. Maria, et al. 2010. “Obesity Related Methylation Changes in

DNA of Peripheral Blood Leukocytes." *BMC Medicine* 8 (1): 87. https://doi.org/10.1186/1741-7015-8-87.

Watada, Hirotaka, and Yoshio Fujitani. 2015. "Minireview: Autophagy in Pancreatic β-Cells and Its Implication in Diabetes." *Molecular Endocrinology* 29 (3): 338–48. https://doi.org/10.1210/me.2014-1367.

Weber, Minéia, Paula Mera, Josefina Casas, Javier Salvador, Amaia Rodríguez, Sergio Alonso, David Sebastián, et al. 2020. "Liver CPT1A Gene Therapy Reduces Diet-induced Hepatic Steatosis in Mice and Highlights Potential Lipid Biomarkers for Human NAFLD." https://doi.org/10.1096/fj.202000678R.

Weedon, Michael N., Mark I. McCarthy, Graham Hitman, Mark Walker, Christopher J. Groves, Eleftheria Zeggini, N. William Rayner, et al. 2006. "Combining Information from Common Type 2 Diabetes Risk Polymorphisms Improves Disease Prediction." *PLOS Medicine* 3 (10): e374. https://doi.org/10.1371/journal.pmed.0030374.

Weisenberger, D. J., G. Liang, and H.-J. Lenz. 2018. "DNA Methylation Aberrancies Delineate Clinically Distinct Subsets of Colorectal Cancer and Provide Novel Targets for Epigenetic Therapies." *Oncogene* 37 (5): 566–77. https://doi.org/10.1038/onc.2017.374.

Willemsen, Gonneke, Kirsten J. Ward, Christopher G. Bell, Kaare Christensen, Jocelyn Bowden, Christine Dalgård, Jennifer R. Harris, et al. 2015. "The Concordance and Heritability of Type 2 Diabetes in 34,166 Twin Pairs From International Twin Registers: The Discordant Twin (DISCOTWIN) Consortium." *Twin Research and Human Genetics* 18 (6): 762–71. https://doi.org/10.1017/thg.2015.83.

Williamson, Alice, Dougall M. Norris, Xianyong Yin, K. Alaine Broadaway, Anne H. Moxley, Swarooparani Vadlamudi, Emma P. Wilson, et al. 2023. "Genome-Wide Association Study and Functional Characterization Identifies Candidate Genes for Insulin-Stimulated Glucose Uptake." *Nature Genetics* 55 (6): 973–83. https://doi.org/10.1038/s41588-023-01408-9.

Wolf, Gabriele, Nicole Aumann, Marta Michalska, Antje Bast, Jürgen Sonnemann, James F Beck, Uwe Lendeckel, Philip Newsholme, and Reinhard Walther. 2010. "Peroxiredoxin III Protects Pancreatic β Cells from Apoptosis." *Journal of Endocrinology* 207 (2): 163–75. https://doi.org/10.1677/JOE-09-0455.

Wright, Jordan J., Adel Eskaros, Annika Windon, Rita Bottino, Regina Jenkins, Amber M. Bradley, Radhika Aramandla, et al. 2024. "Exocrine Pancreas in Type 1 and Type 2 Diabetes: Different Patterns of Fibrosis, Metaplasia, Angiopathy, and Adiposity." *Diabetes* 73 (7): 1140–52. https://doi.org/10.2337/db23-0009.

Wright, Jordan J., Diane C. Saunders, Chunhua Dai, Greg Poffenberger, Brynn Cairns, David V. Serreze, David M. Harlan, Rita Bottino, Marcela Brissova, and Alvin C. Powers. 2020. "Decreased Pancreatic Acinar Cell Number in Type 1 Diabetes." *Diabetologia* 63 (7): 1418–23. https://doi.org/10.1007/s00125-020-05155-y.

Xu, Eric E., Shugo Sasaki, Thilo Speckmann, Cuilan Nian, and Francis C. Lynn. 2017. "SOX4 Allows Facultative β-Cell Proliferation Through Repression of Cdkn1a." *Diabetes* 66 (8): 2213–19. https://doi.org/10.2337/db16-1074.

Xu, Zongli, Liang Niu, Leping Li, and Jack A. Taylor. 2016. "ENmix: A Novel Background Correction Method for Illumina HumanMethylation450 BeadChip." *Nucleic Acids Research* 44 (3): e20. https://doi.org/10.1093/nar/gkv907.

Xu, Zongli, Liang Niu, and Jack A. Taylor. 2021. "The ENmix DNA Methylation Analysis Pipeline for Illumina BeadChip and Comparisons with Seven Other Preprocessing Pipelines." *Clinical Epigenetics* 13 (1): 216. https://doi.org/10.1186/s13148-021-01207-1.

Xue, Angli, Yang Wu, Zhihong Zhu, Futao Zhang, Kathryn E. Kemper, Zhili Zheng, Loic Yengo, Luke R. Lloyd-Jones, Julia Sidorenko, Yeda Wu, eQTLGen Consortium, et al. 2018. "Genome-Wide Association Analyses Identify 143 Risk Variants and Putative Regulatory Mechanisms for Type 2 Diabetes." *Nature Communications* 9 (1): 2941. https://doi.org/10.1038/s41467-018-04951-w.

Xue, Angli, Yang Wu, Zhihong Zhu, Futao Zhang, Kathryn E. Kemper, Zhili Zheng, Loic Yengo, Luke R. Lloyd-Jones, Julia Sidorenko, Yeda Wu, eQTLGen Consortium, et al. 2018. “Novel Susceptibility Loci and Genetic Regulation Mechanisms for Type 2 Diabetes.” bioRxiv. https://doi.org/10.1101/284570.

Yaghootkar, Hanieh, Alena Stancáková, Rachel M. Freathy, Jagadish Vangipurapu, Michael N. Weedon, Weijia Xie, Andrew R. Wood, et al. 2015. “Association Analysis of 29,956 Individuals Confirms That a Low-Frequency Variant at CCND2 Halves the Risk of Type 2 Diabetes by Enhancing Insulin Secretion.” *Diabetes* 64 (6): 2279–85. https://doi.org/10.2337/db14-1456.

Yang, B. T., T. A. Dayeh, C. L. Kirkpatrick, J. Taneera, R. Kumar, L. Groop, C. B. Wollheim, M. D. Nitert, and C. Ling. 2011. “Insulin Promoter DNA Methylation Correlates Negatively with Insulin Gene Expression and Positively with HbA1c Levels in Human Pancreatic Islets.” *Diabetologia* 54 (2): 360–67. https://doi.org/10.1007/s00125-010-1967-6.

Yang, Beatrice T., Tasnim A. Dayeh, Petr A. Volkov, Clare L. Kirkpatrick, Siri Malmgren, Xingjun Jing, Erik Renström, Claes B. Wollheim, Marloes Dekker Nitert, and Charlotte Ling. 2012. “Increased DNA Methylation and Decreased Expression of PDX-1 in Pancreatic Islets from Patients with Type 2 Diabetes.” *Molecular Endocrinology (Baltimore, Md.)* 26 (7): 1203–12. https://doi.org/10.1210/me.2012-1004.

Yang, Jingyun, Lei Yu, Christopher Gaiteri, Gyan P. Srivastava, Lori B. Chibnik, Sue E. Leurgans, Julie A. Schneider, Alexander Meissner, Philip L. De Jager, and David A. Bennett. 2015. “Association of DNA Methylation in the Brain with Age in Older Persons Is Confounded by Common Neuropathologies.” *The International Journal of Biochemistry & Cell Biology* 67 (October):58–64. https://doi.org/10.1016/j.biocel.2015.05.009.

Yin, Yan-Bin, Wei Ji, Ying-Lan Liu, Qian-Hao Gao, Dong-Dong He, Shi-Lin Xu, Jing-Xin Fan, and Li-Hai Zhang. 2024. “cNPAS2 Induced β Cell Dysfunction by Regulating

KANK1 Expression in Type 2 Diabetes." *World Journal of Diabetes* 15 (9): 1932–41. https://doi.org/10.4239/wjd.v15.i9.1932.

You, LiangHui, Ning Wang, DanDan Yin, LinTao Wang, FeiYan Jin, YaNan Zhu, QingXin Yuan, and Wei De. 2016. "Downregulation of Long Noncoding RNA Meg3 Affects Insulin Synthesis and Secretion in Mouse Pancreatic Beta Cells." *Journal of Cellular Physiology* 231 (4): 852–62. https://doi.org/10.1002/jcp.25175.

Yuan, Shuai, Siddhartha Kar, Paul Carter, Mathew Vithayathil, Amy M. Mason, Stephen Burgess, and Susanna C. Larsson. 2020. "Is Type 2 Diabetes Causally Associated With Cancer Risk? Evidence From a Two-Sample Mendelian Randomization Study." *Diabetes* 69 (7): 1588–96. https://doi.org/10.2337/db20-0084.

Zeimet, Alain G., Heidi Fiegl, Georg Goebel, Francis Kopp, Claude Allasia, Daniel Reimer, Ilona Steppan, Elisabeth Mueller-Holzner, Melanie Ehrlich, and Christian Marth. 2011. "DNA Ploidy, Nuclear Size, Proliferation Index and DNA-Hypomethylation in Ovarian Cancer." *Gynecologic Oncology*, New Biomarkers and Risk Predictors for Prevention and Early Detection of Gynecologic Cancers, 121 (1): 24–31. https://doi.org/10.1016/j.ygyno.2010.12.332.

Zhang, Geng, Peijun He, Hanson Tan, Anuradha Budhu, Jochen Gaedcke, B. Michael Ghadimi, Thomas Ried, et al. 2013. "Integration of Metabolomics and Transcriptomics Revealed a Fatty Acid Network Exerting Growth Inhibitory Effects in Human Pancreatic Cancer." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 19 (18): 4983–93. https://doi.org/10.1158/1078-0432.CCR-13-0209.

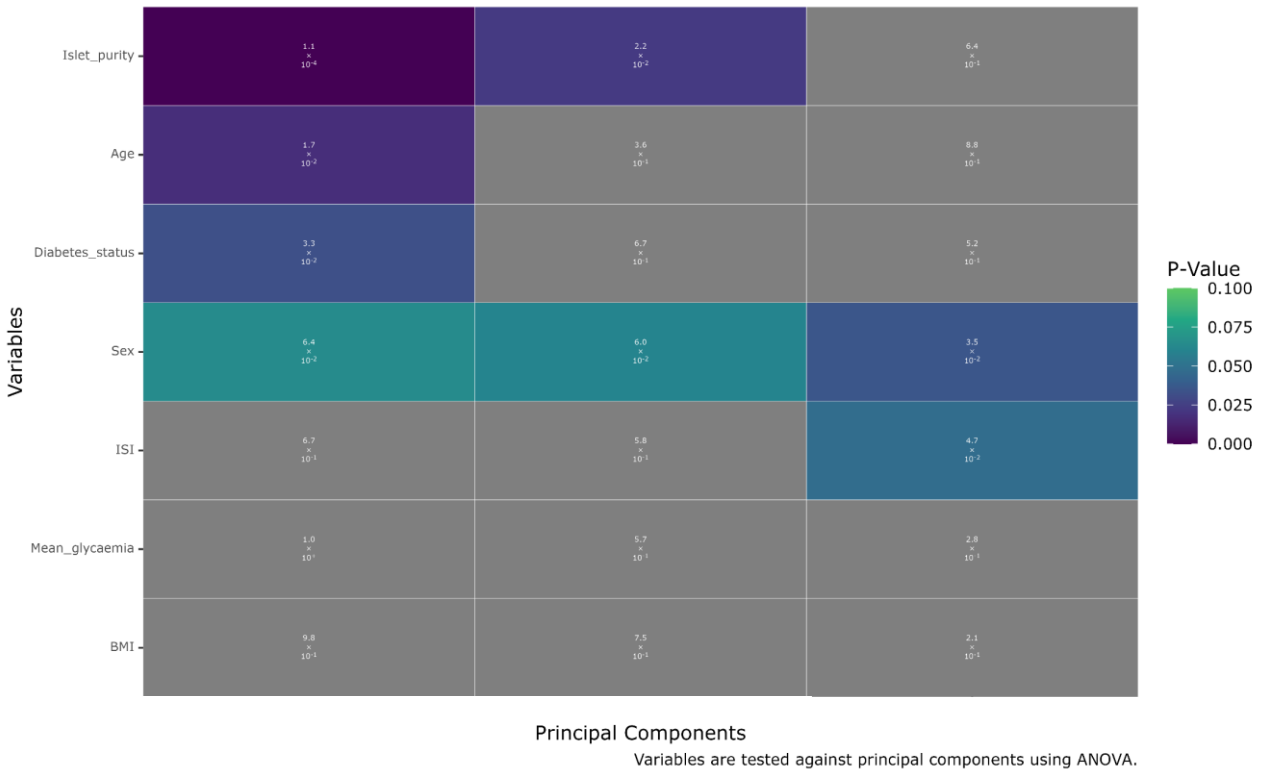
Zheng, Yan, Sylvia H. Ley, and Frank B. Hu. 2018. "Global Aetiology and Epidemiology of Type 2 Diabetes Mellitus and Its Complications." *Nature Reviews Endocrinology* 14 (2): 88–98. https://doi.org/10.1038/nrendo.2017.151.

Zhou, Yingyao, Bin Zhou, Lars Pache, Max Chang, Alireza Hadj Khodabakhshi, Olga Tanaseichuk, Christopher Benner, and Sumit K. Chanda. 2019. “Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets.” *Nature Communications* 10 (1): 1523. https://doi.org/10.1038/s41467-019-09234-6.

Zhu, Heng, Guohua Wang, and Jiang Qian. 2016. “Transcription Factors as Readers and Effectors of DNA Methylation.” *Nature Reviews Genetics* 17 (9): 551–65. https://doi.org/10.1038/nrg.2016.83.

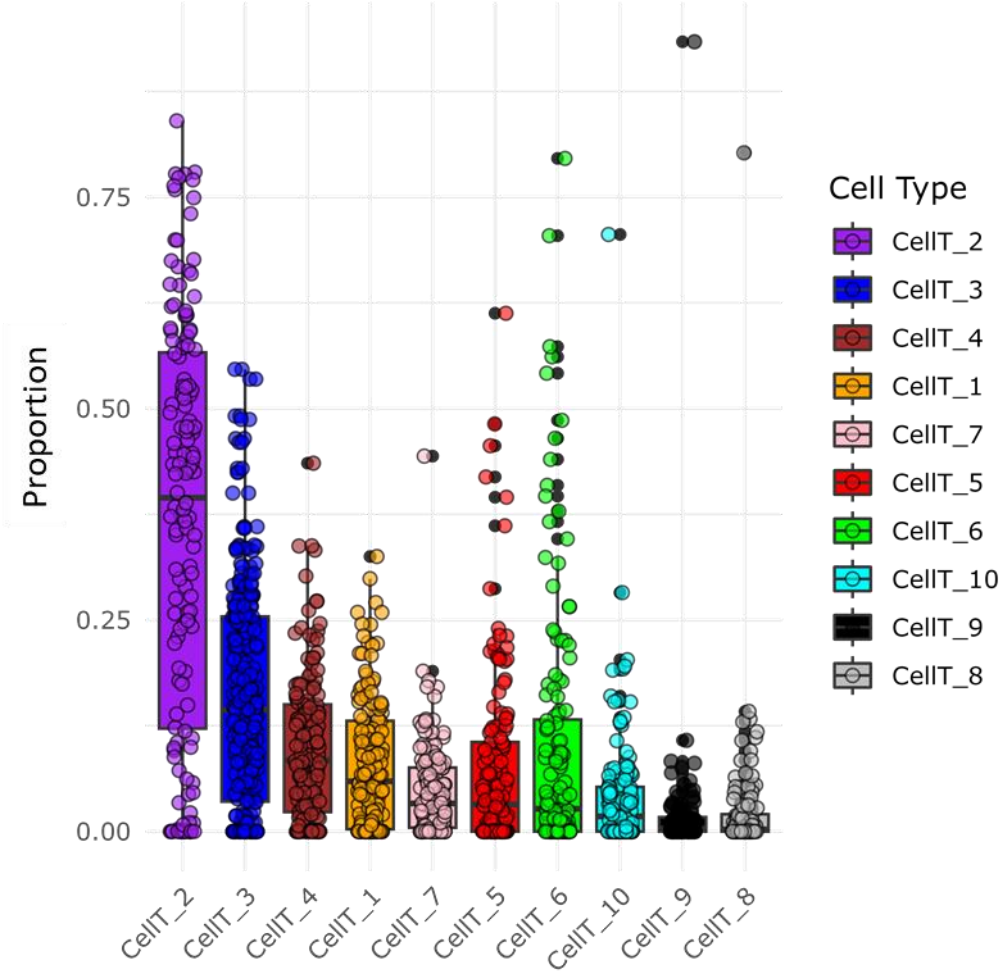
12 Supplementary figures

12.1 Project 1

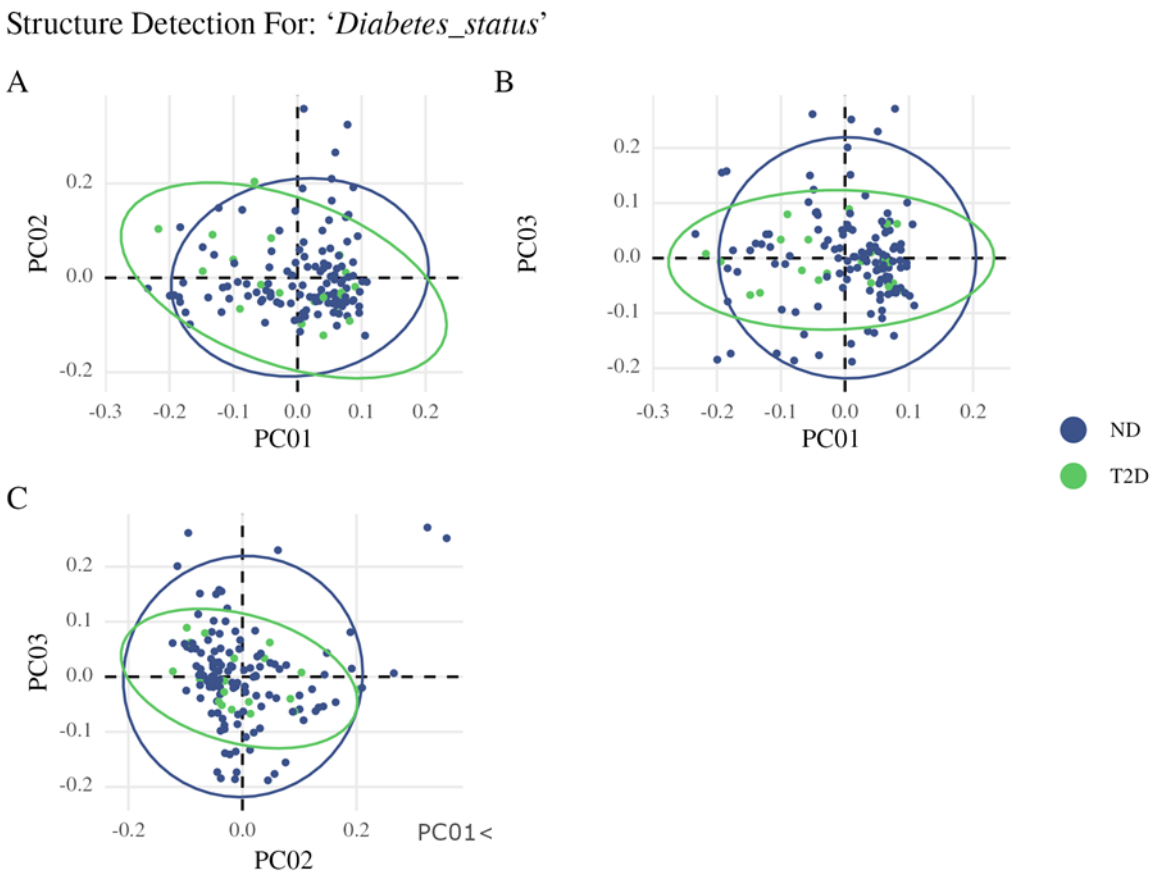


Supplementary figure 1: Principal component analysis of cohort characteristics against DNA methylation β-values. The first three PCs are represented and ordered by the proportion of variance in the data they explain. P-values were computed with

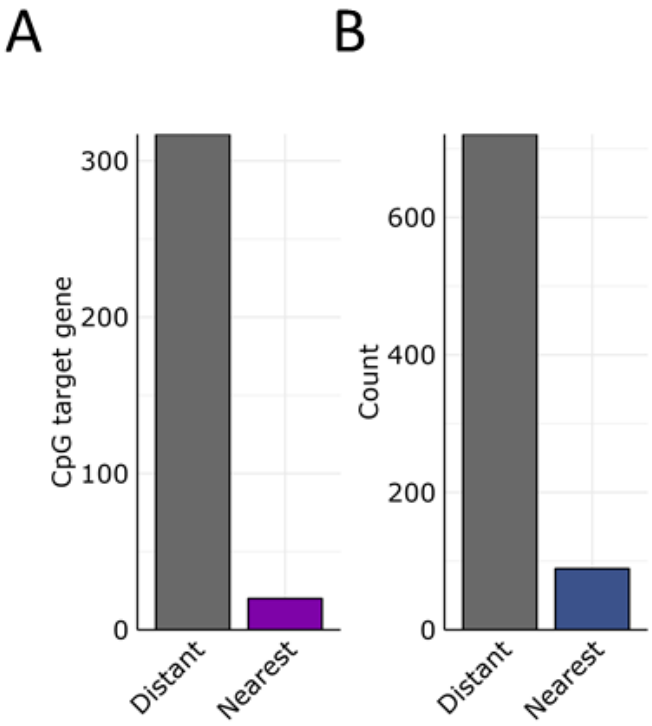
an ANOVA test. Colours gradient represents P-values, with darker colours indicating lowers values.



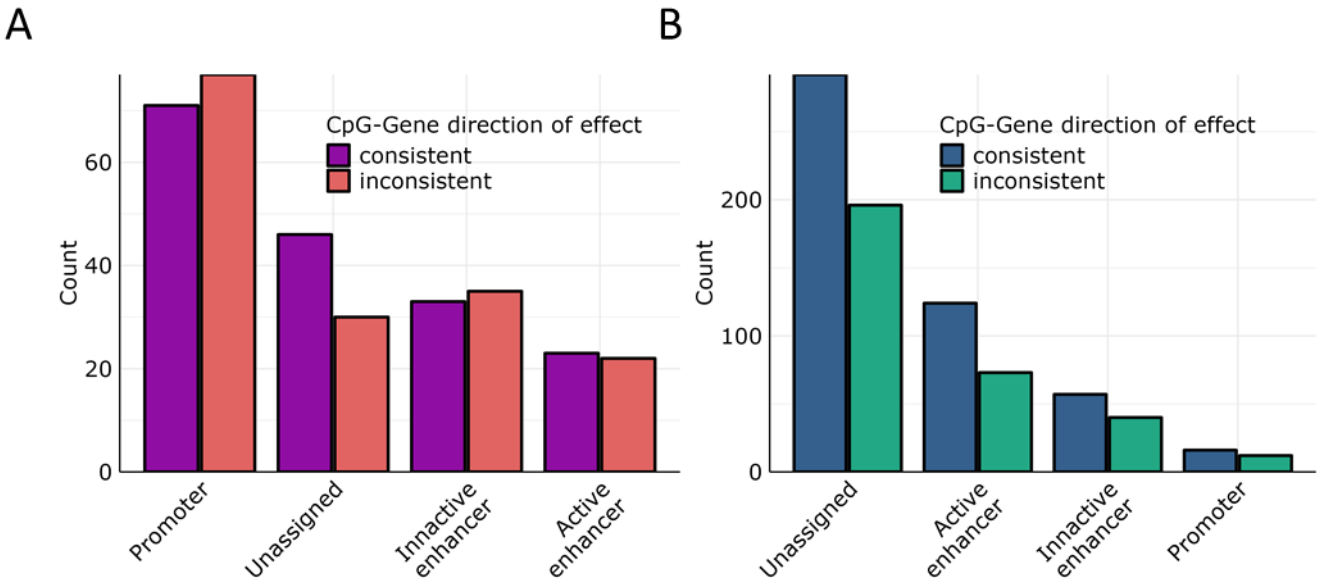
Supplementary figure 2: Pancreatic islet tissue samples estimated cell type composition. Cell types were estimated using DNA methylation data with the RefFreeEWAS r package (Houseman et al. 2016). Colours are arbitrary and indicate different cell types. Each point represents an individual sample.



Supplementary figure 3: Principal component analysis (PCA) of DNA methylation data from pancreatic islets. A-C) PC 1-3 of the pancreatic islet samples. Pancreatic islet samples coloured blue originated from control individuals, whilst green indicates samples originating from individuals with type 2 diabetes.

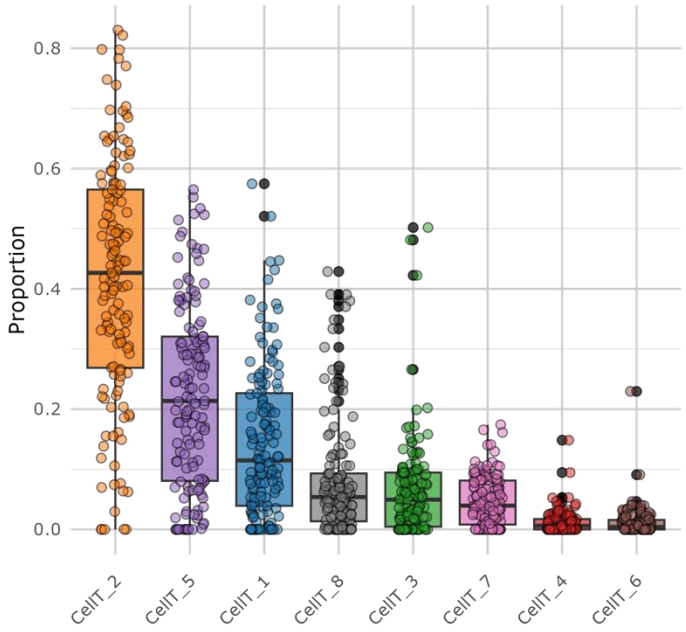


Supplementary figure 4: Number of CpGs targeting their nearest gene, or not. A-B) Bar plots showing the nominal number of CpGs which target the nearest gene or not for **A)** age (purple) and for **B)** T2D (blue). Grey indicates CpGs targeting distant genes for both **A)** and **B)**.

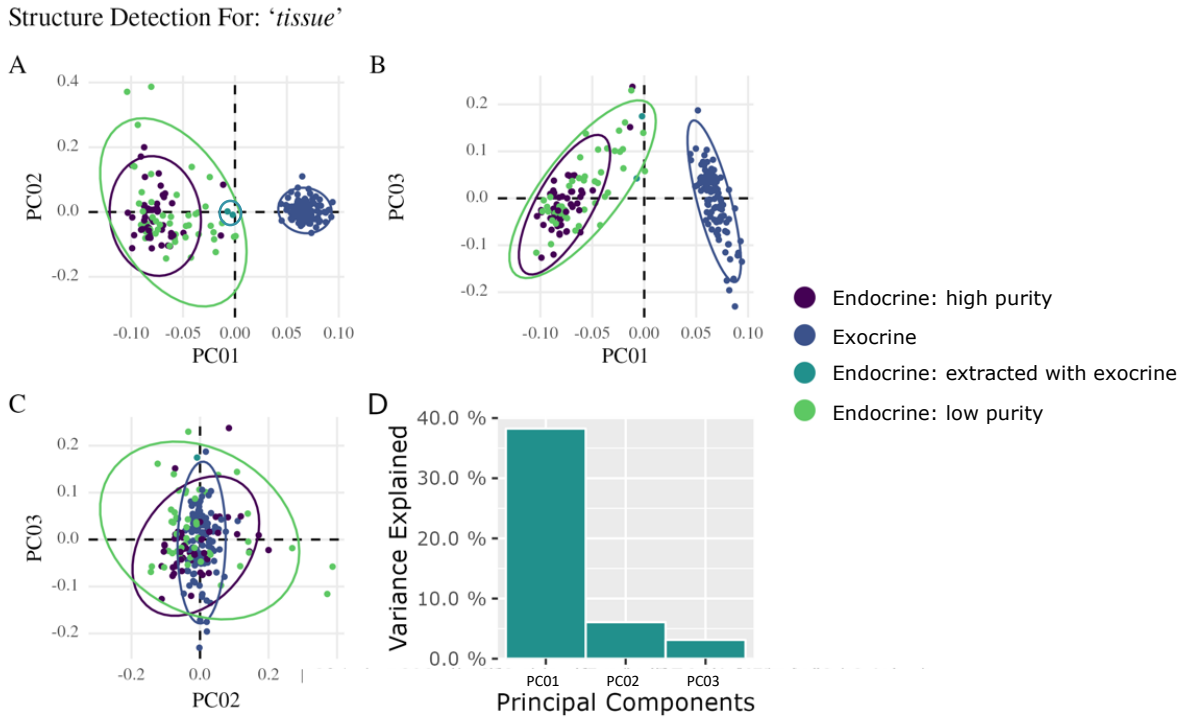


Supplementary figure 5: Bar plots showing the absolute number of CpGs associated with gene expression and the direction in which these CpGs influence gene expression, split by regulatory region. A) Bar plot for genes dysregulated in the age analysis. Purple indicates consistent effect direction *ie.,* an increase in methylation increases gene expression, and purple indicates the opposite. **B)** Bar plot for genes dysregulated in the T2D analysis. Blue indicates consistent effect direction, while green indicates inconsistent effect direction.

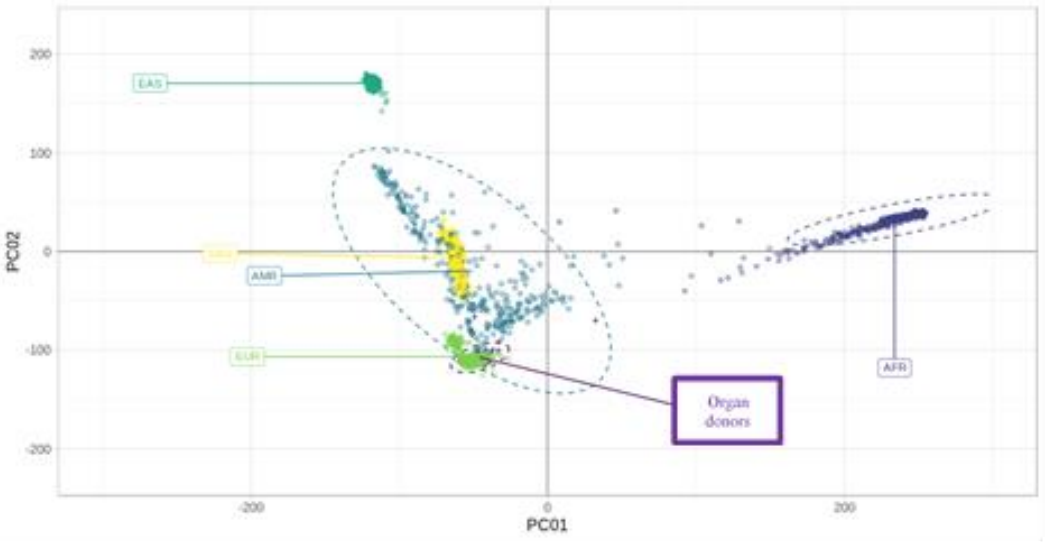
13 Project 2



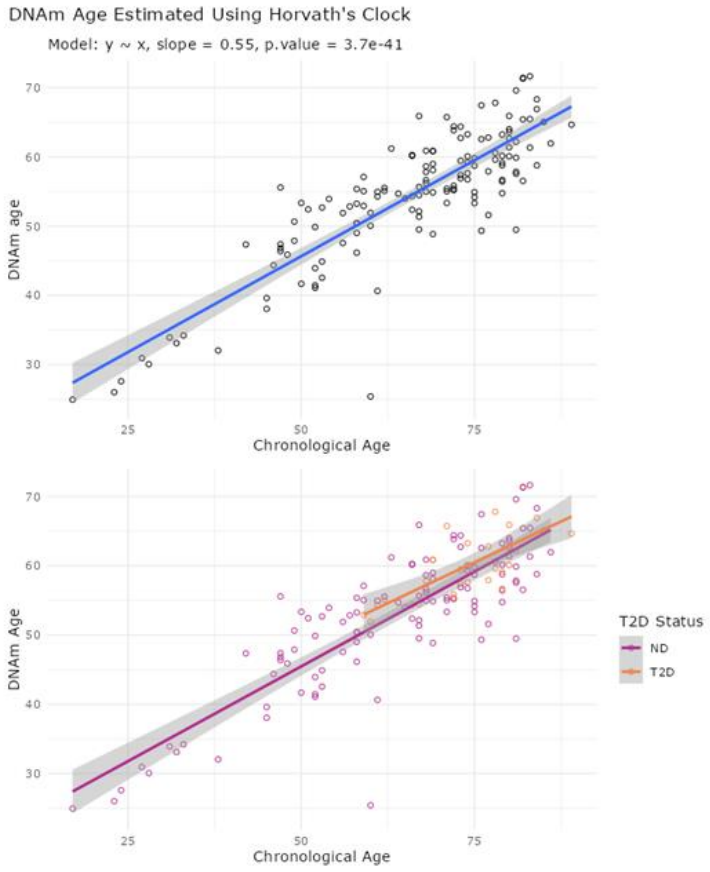
Supplementary figure 6: Exocrine tissue estimated cell type composition. Cell types were estimated using DNA methylation data with the RefFreeEWAS r package (Houseman et al. 2016).



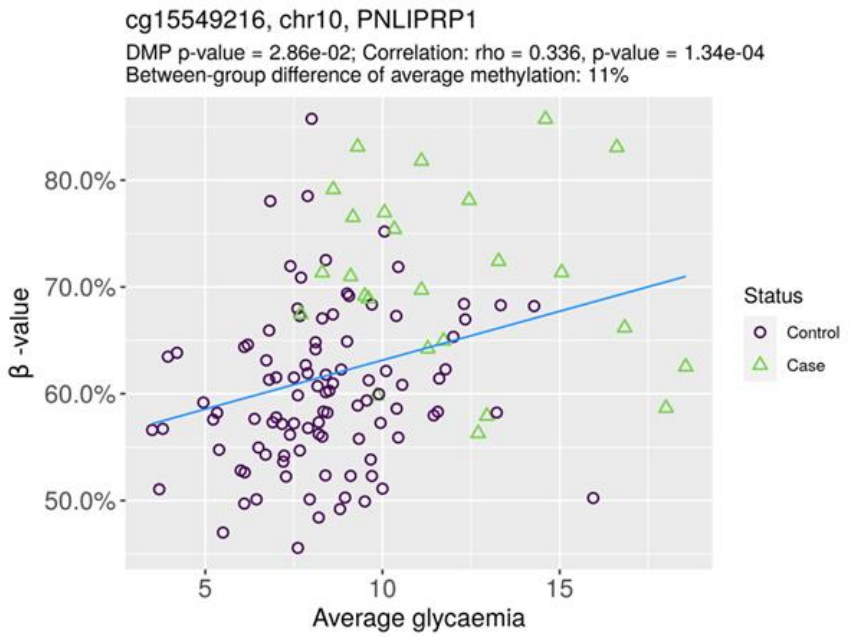
Supplementary figure 7: Principal component analysis (PCA) of DNA methylation data from exocrine and endocrine preparations. A-C) PC 1-3 of the pancreas samples. Purple indicates high purity endocrine preparations. Dark blue indicates exocrine preparations. Teal indicates endocrine preparations that were extracted along with exocrine preparations. Light green indicates endocrine preparations of low purity (*ie.,* with some exocrine content).



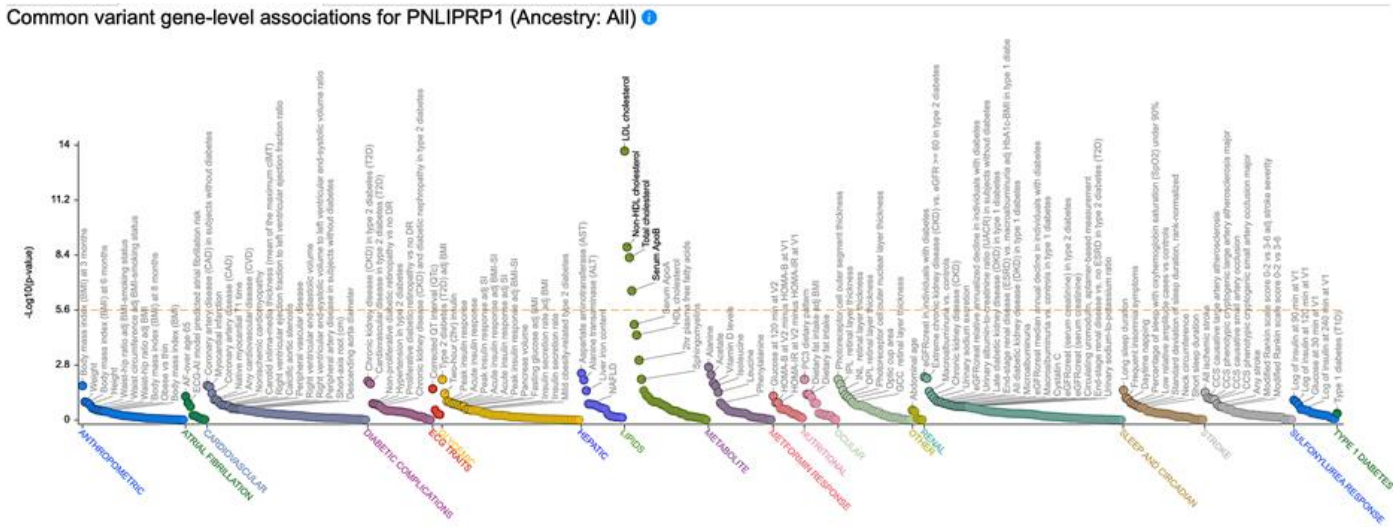
Supplementary figure 8: Ancestry clustering of exocrine organ donors using the 1,000-genome reference panel. Organ donors are in purple. Ad Mixed americans are in dark green, East Asians are in teal. South Asian are in yellow. Europeans are in green. Africans are in dark blue.



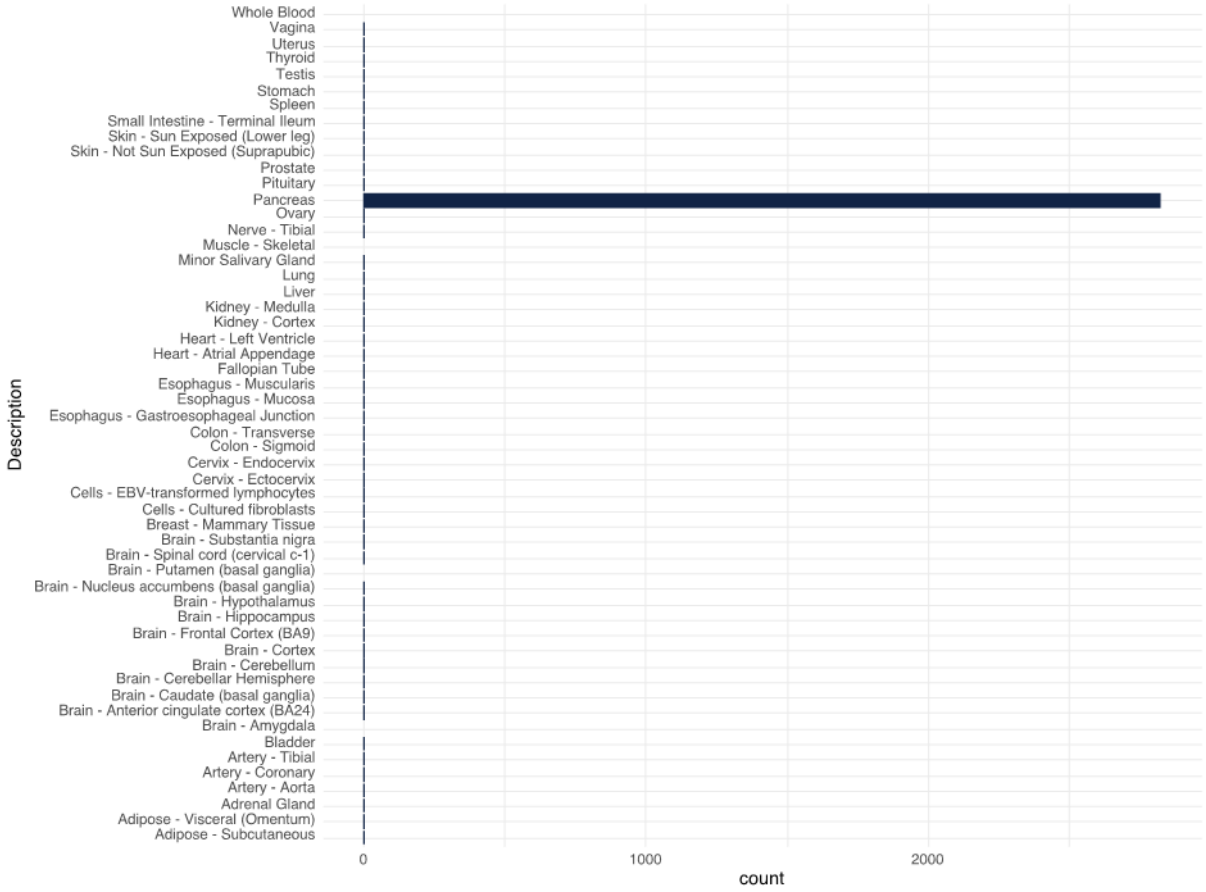
Supplementary figure 9: Biological age estimation using Horvath's DNA methylation age clock. Upper plot shows the DNA methylation age of all samples. Lower plot indicates DNA methylation age split by diabetes status. Purple indicates non diabetics, orange indicates individuals with type 2 diabetes.



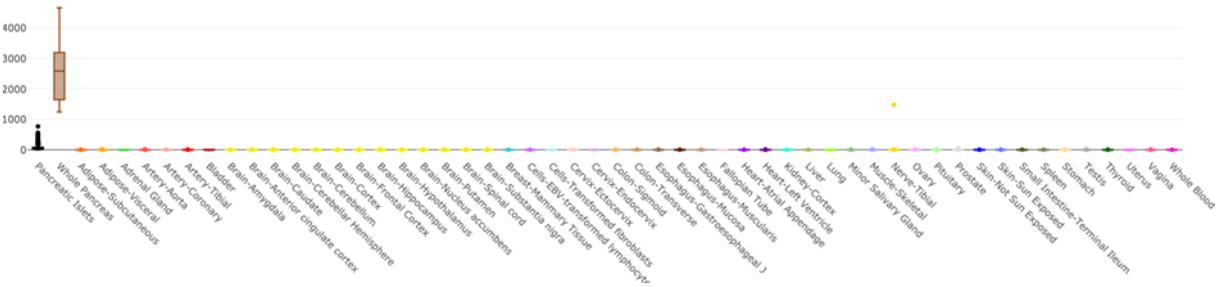
Supplementary figure 10: Relationship between glycaemia and DNA methylation at cg15549216. The X axis indicates glycaemia, and the Y axis indicates β-value of cg15549216. Purple circle indicate control individuals and green triangles indicates individuals with type 2 diabetes.



Supplementary figure 11**: Phenotype associations for common PNLIPRP1 variants from the Type 2 Diabetes Knowledge portal** (https://t2d.hugeamp.org/)**.**

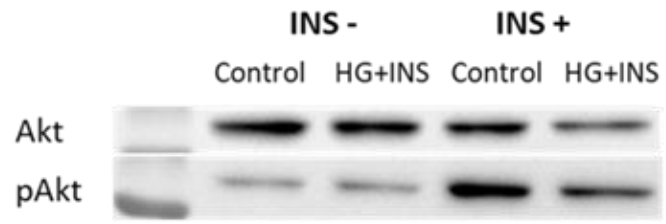


Supplementary figure 12: GTEX gene expression data for PNLIPRP1. Counts indicate transcript per million value for the gene.

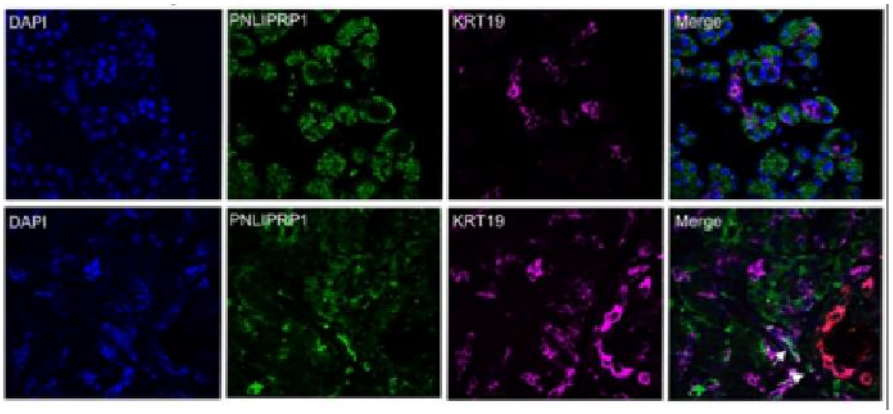


Supplementary figure 13: PNLIPRP1 gene expression from the TIGER database.

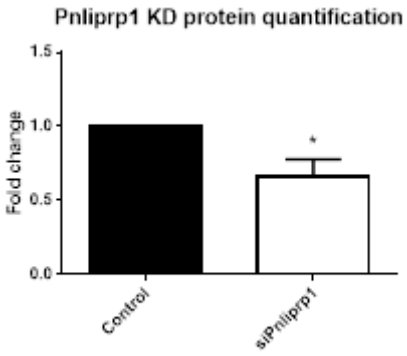
PNLIPRP1 is lowly expressed in pancreatic islets, transcript per million (TPM) = 36, and highly expressed in whole pancreas tissue, TPM = 2581.



Supplementary figure 14: Western blot of Akt and phosphorylated Akt (pAkt) in AR42J. The cells were initially treated with high glucose and insulin for 48 hours. Subsequently, medium was removed and replaced along with 200 nM insulin for 1 hour to stimulate the cells.



Supplementary figure 15: Immunofluorescence of human whole pancreas tissue. The upper row shows healthy whole pancreas tissue whilst the lower row shows whole pancreas tissue from an individual with type 2 diabetes. Blue (DAPI) indicates nuclei. Green indicates PNLIPRP1. Purple indicates KRT19. Red indicates insulin. Arrows indicate sites of PNLIPRP1 and KRT19 co-expression.



Supplementary figure 16: Protein quantification of Pnliprp1 following knockdown with siRNA. AR42J cells were treated with siRNA for 48 hours prior to protein

Supplementary tables

All supplementary tables are available in the excel document provided with the thesis.

14 Published article

PNLIPRP1 Hypermethylation in Exocrine Pancreas Links Type 2 Diabetes and Cholesterol Metabolism

Lucas Maurin,[1,2] Lorella Marselli,[3] Mathilde Boissel,[1,2] Lijiao Ning,[1,2] Raphael Boutry,[1,2] Justine Fernandes,[1,2] Mara Suleiman,[3] Carmela De Luca,[3] Audrey Leloire,[1,2] Vincent Pascat,[1,2] Bénédicte Toussaint,[1,2] Souhila Amanzougarene,[1,2] Mehdi Derhourhi,[1,2] Anne Jörns,[4] Sigurd Lenzen,[4] François Pattou,[1] Julie Kerr-Conte,[5] Mickaël Canouil,[1] Piero Marchetti,[3] Amélie Bonnefond,[1,2,6] Philippe Froguel,[1,2,6] and Amna Khamis[1,2,6]

Diabetes 2024;73:1908–1918 | https://doi.org/10.2337/db24-0215

We postulated that type 2 diabetes (T2D) predisposes patients to exocrine pancreatic diseases through (epi)genetic mechanisms. We explored the methylome (using MethylationEPIC arrays) of the exocrine pancreas in 141 donors, assessing the impact of T2D. An epigenome-wide association study of T2D identified hypermethylation in an enhancer of the pancreatic lipase–related protein 1 (*PNLIPRP1*) gene, associated with decreased *PNLIPRP1* expression. *PNLIPRP1* null variants (found in 191,000 participants in the UK Biobank) were associated with elevated glycemia and LDL cholesterol. Mendelian randomization using 2.5M SNP Omni arrays in 111 donors revealed that T2D was causal of *PNLIPRP1* hypermethylation, which in turn was causal of LDL cholesterol. Additional AR42J rat exocrine cell analyses demonstrated that *Pnliprp1* knockdown induced acinar-to-ductal metaplasia, a known prepancreatic cancer state, and increased cholesterol levels, reversible with statin. This (epi)genetic study suggests a role for *PNLIPRP1* in human metabolism and exocrine pancreatic function, with potential implications for pancreatic diseases.

ARTICLE HIGHLIGHTS

- We performed this study to identify epigenetic changes with type 2 diabetes (T2D) in the pancreas.
- This study addresses whether T2D induces epigenetic changes that could explain why individuals with T2D are more prone to pancreatic diseases.
- We found hypermethylation at *PNLIPRP1* associated with T2D and identified a role of this gene in cholesterol metabolism.
- This study has important implications in the prevention of pancreatic diseases, because their molecular mechanisms remain largely unknown.

The pancreas is an organ featuring distinct endocrine and exocrine compartments. The exocrine pancreas, which comprises 98% of the pancreas (1), is composed primarily of acini that secrete enzymes involved in digestion into the ducts (2). The remaining 2% constitutes the endocrine pancreas, which comprises pancreatic islets, the main function of which is to secrete insulin (3). This distinction is also seen in diseases associated with each compartment; type 2 diabetes (T2D) results from the dysregulation of the endocrine pancreas (3), whereby the dysfunction of the exocrine pancreas induces pancreatitis (i.e., inflammation) or pancreatic cancer, an increasingly leading cause of death (4).

Epidemiologic studies have shown that individuals with T2D have an increased risk of developing pancreatic cancer (5), and the reverse is also true (6). In addition, there

[1]INSERM UMR 1283, CNRS UMR 8199, European Genomic Institute for Diabetes, Institut Pasteur de Lille, Lille, France
[2]Lille University Hospital, University of Lille, Lille, France
[3]Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy
[4]Institute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany
[5]University of Lille, Inserm, Centre Hospitalier Universitaire Lille, Lille Pasteur Institute, U1190, EGID, Lille, France
[6]Section of Genomics of Common Disease, Department of Metabolism, Imperial College London, London, U.K.

Corresponding author: Amna Khamis, amna.khamis@cnrs.fr, or Philippe Froguel, p.froguel@imperial.ac.uk

Received 14 March 2024 and accepted 31 July 2024

P.F. and A.K. contributed equally.

This article contains supplementary material online at https://doi.org/10.2337/figshare.26510815.

is an association between T2D and chronic pancreatitis, leading to acinar atrophy (7–9).

A recent study exploring the effects of T2D and age on the exocrine pancreas found that compared with individuals of controls without diabetes and patients with type 1 diabetes, pancreases of patients with T2D presented with greater acinar-to-ductal metaplasia (ADM), a process whereby exocrine acinar cells transdifferentiate into ductal-like cells (10). Furthermore, ADM correlated positively with T2D duration, and this duration correlated with pancreatic cancer initiation, validating the link between T2D and pancreatic cancer. In this respect, pancreatic ADM is considered a precancerous state of the pancreas. Understanding exocrine–endocrine functional interactions has important implications in the prevention of these deadly pancreatic diseases, because as yet, their molecular mechanisms remain largely unexplored.

We hypothesized that functional alterations in the exocrine pancreas are linked to local epigenetic changes in the context of environmental factors, which may exacerbate the risk of further exocrine disease.

RESEARCH DESIGN AND METHODS

Methylation in Human Samples

We selected 155 whole-pancreas samples, collected based on the Innovative Medicines Initiative for Diabetes (IMIDIA) consortium (11), according to sample availability. Of these samples, 32 had T2D, based on dinical characteristics and tests, according to the American Diabetes Association guideline (2019). Sample collection was followed by next-of-kin informed consent, with the approval of the local ethics committees in Pisa and Hannover.

DNA was extracted from whole-pancreas samples using the NudeoSpin Tissue Kit (cat. no. T740952.50; Macherey-Nagel). Bisulphite conversion was performed in 800 ng DNA using the EZ DNA Methylation Kit (cat. no. 5001; Zymo Research), and DNA was subjected to Illumina 850K MethylationEPIC array. Array data were imported using the *minfi* R package (12). Quality control steps removed CpG probes that were located on sex chromosomes or near single nucleotide polymorphisms (SNPs), cross-hybridizing, or non-CG or had a detection threshold P value <0.01. Samples with <99% probes with a detection P value <0.01 were exduded. Probe-design biases and batch effects were normalized using R packages *Enmix* and *SVA* (13,14). Samples were removed for having a call rate <99%, and two sex-discordant samples were removed. After quality control, 746,912 probes and 141 samples remained (Table 1), comprising 69 females and 72 males. Population structure was evaluated by principal component analysis, with 1000 Genomes as reference (Supplementary Fig. 2).

Comparison With Pancreatic Islet Profiles

To compare the profile of exocrine pancreatic extractions with pancreatic islets, we used pancreatic islet data that were generated using the same methods as those described above and were also collected based on the IMIDIA consortium. The pancreatic islet samples were obtained from 144 organ donors, with an age range of 22 to 96 years (mean age 69; range 22–96 years). There was no statistically significant difference in age, sex, BMI, or T2D status between the exocrine and pancreatic islet donors. Characteristics of the donors are detailed in Supplementary Table 1. Principal component analysis of methylation of exocrine and pancreatic islets was determined from β values from both groups using the *flashPCA* package in R.

Table 1—Clinical characteristics of participants

Characteristic	No diabetes (n = 113)*	T2D (n = 28)*	P†
Sex			0.016
Female	61 (54)	8 (29)	
Male	52 (46)	20 (71)	
Age, years	67 (53, 76)	74 (69, 79)	0.004
BMI, kg/m^2	25.0 (23.1, 27.1)	26.0 (24.2, 28.2)	0.13
T2D duration, years	—	9 (6, 12)	—
T2D treatment	0	25 (89)	—
Statin treatment	10 (17)	4 (29)	0.4

*Data given as n (%) or median (interquartile range). †Pearson χ^2 test or Wilcoxon rank sum test.

Epigenome-Wide Association Study

Cell composition was estimated using the R package *RefFreeEWAS* (15), which estimated eight cell types (Supplementary Fig. 1). We adjusted for J-1 cell types (where J is the number of cell types defined) and, to avoid collinearity, removed the cell type with the lower estimated proportion. A linear regression model was applied for T2D association, corrected for age, sex, BMI, and cellular composition. Bonferroni correction was applied for multiple testing. The study design is summarized in Fig. 1.

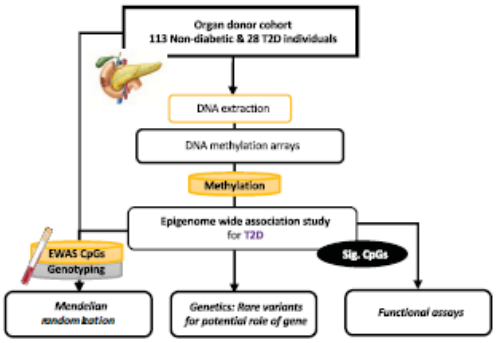


Figure 1—Overview of the study design. Whole-pancreas tissue was extracted from organ donors, and DNA methylation data were generated using the MethylationEPIC 850K array to perform epigenome-wide association studies (EWAS) of T2D in organ donors. Further genetic and functional validation was performed.

UK Biobank Used to Identify Rare Variant Associations in *PNLIPRP1*

Exome sequencing data from up to 191,000 participants in the UK Biobank were used to identify rare null variants (minor allele frequency [MAF] <1%) in *PNLIPRP1*. This research is part of UK Biobank research application 67575. We tested the association of null variants for several traits, including BMI, glucose, and lipid traits, using the MiST method (16), which tests rare variants in a single cluster (at the gene scale). Score π represents the mean effect of the cluster; τ represents the heterogeneous effect of the cluster. The overall *P* value tests the association between the set of variants and the trait of interest. For each trait, we adjusted for relevant covariates. We considered a trait to be significant if the *P* value associated with the direct burden effect of the cluster was significant ($P\pi$ value [i.e., $P < 0.05$]) and the direction of effect for the variants was consistent, revealed by the absence of heterogeneity (τ *P* value [i.e., $P > 0.05$]).

Mendelian Randomization

To assess the direction of causality between traits of interest (i.e., T2D, LDL cholesterol [LDL-C], and CpG methylation), we performed bidirectional two-sample Mendelian randomization (MR). To obtain genetic associations for both T2D and LDL-C, we consulted previously published large-scale European genome-wide association studies of T2D (17) and LDL-C (18). All MR analyses were performed using the TwoSampleMR (version 0.5.7) R software package.

To obtain genetic associations between SNPs and DNA methylation, we genotyped 111 control samples from our organ donors (Illumina HumanOmni2.5 arrays) using the Illumina iScan (method detailed in Supplementary Methods). Methylation quantitative trait loci (mQTL) analysis was performed using QTLtools software (19), adjusting for age, sex, and BMI. Forward and reverse MRs are detailed in Supplementary Methods. The inverse variance weighted (IVW) method was applied to compute the causal effect estimate. The F statistic was used to verify the strength of instrument. The MR-Egger method was applied to investigate the presence of horizontal pleiotropy. Linkage disequilibrium was assessed with the ld_clump() function from the R software package ieugwasr (version 0.1.5).

The leave-one-out analysis was used to verify whether any variant was driving our findings. Heterogeneity was assessed using the Cochran Q test.

RNA Expression of *PNLIPRP1* in Organ Donors

Eleven samples were processed (six controls and five individuals with T2D), matched for age, sex, and BMI. RNA was extracted using Trizol (cat. no. 15596-026; Thermo Fisher Scientific). A detailed description of the results is provided in Supplementary Methods.

Functional Characterization

All in vitro assays were performed using the rat AR42J acinar cell line (cat. no. CRL-1492; American Type Culture Collection). Methods are further detailed in Supplementary Methods.

High Glucose and Insulin Treatment

AR42J cells were treated with 20 mmol/L glucose or 100 nmol/L insulin, or both, for 72 h, and cells were harvested for RT–quantitative (qPCR).

Akt Response and Glucose Uptake

Akt response was tested using Western blot in AR42J cells stimulated with or without 200 nmol/L insulin for 1 h. Details of all antibodies in this study are listed in Supplementary Table 3. Glucose uptake was measured using the colorimetric glucose detection kit (cat. no. EIAGLUC; Invitrogen), following the manufacturer protocol.

siRNA Knockdown

AR42J cells were transfected with *PNLIPRP1* siRNA or nontargeting control siRNA. RNA and protein were harvested after 48 or 72 h (qPCR and Western blotting method detailed in Supplementary Methods).

Immunohistochemistry

Human pancreatic tissue sections, from controls and individuals with T2D, were obtained by the INSERM UMR 1190 unit (University of Lille, Lille, France). We excluded any fibrotic process or pathologic alterations in the pancreatic tissues.

RNA Sequencing

Library preparation from *PNLIPRP1* knockdown (KD) RNA was performed using the KAPA mRNA HyperPrep Preparation Kit (Roche) and sequenced with NovaSeq 6000 (Illumina) (method provided in Supplementary Methods). Differential expression analysis was determined using DeSeq2. Pathway analysis was performed using Metascape or EnrichR. MTS proliferation (cat. no. 197010; Acbam) and Cholesterol Ester-Glo Assay (cat. no. J3190; Promega) methods were performed per protocol instructions in *PNLIPRP1* KD.

Data and Resource Availability

The data sets generated and/or analyzed during the current study are available in the Figshare repository (https://figshare.com/s/16ad75ace2b169a1623c). No applicable resources were generated or analyzed during the current study.

RESULTS

DNA Methylation Changes in the Exocrine Pancreas Associated With T2D

We measured DNA methylation (MethylationEPIC arrays) in whole-pancreas samples obtained from 141 organ donors (age 17–89; median 67 years) of European descent (Table 1 and Supplementary Fig. 2). To validate that the observed whole-pancreas epigenetic profile was indeed

mainly exocrine tissue, we compared the methylation profile with pancreatic islet methylation profiles (Infinium MethylationEPIC arrays) from 125 individuals from the organ donor cohort, matched for age, sex, and BMI (Supplementary Table 1). The methylation profiles of pancreatic islets and whole pancreas were distinct, and this was consistent with four pancreatic islet samples extracted and handled in parallel with the exocrine preparations (Supplementary Fig. 2). Pancreatic islet preparations with low islet purity (i.e., percentage of islets compared with other pancreatic tissues, such as exocrine) did not overlap with exocrine preparations (Supplementary Fig. 3).

We performed an epigenome-wide association study to determine DNA methylation changes associated with T2D. We detected a single CpG significantly associated with T2D: hypermethylation of cg15549216 CpG (Bonferroni-corrected $P = 0.025$) (Fig. 2*A*). This CpG was located in the gene body of *PNLIPRP1*, encoding pancreatic lipase–related protein 1, and was 11.4% more methylated in patients with T2D compared with controls (estimate 0.6; SE 0.1) (Fig. 2*B*). The cg15549216 CpG was also positively correlated with glucose levels ($P = 1.34 \times 10^{-4}$) (Supplementary Fig. 4). We detected one differentially methylated region associated with increased T2D risk, which encompassed this cg15549216 CpG, and two flanking CpGs, cg06606475 and cg08580014 (Fig. 2*C* and *D*), that were consistent in direction of effect: the cg06606475 CpG located 921 bp upstream of the cg15549216 probe (mean difference 9.0%; $P = 5.9 \times 10^{-5}$; estimate 0.38; SE 0.09) and the cg08580014 CpG located 370 bp downstream of the cg15549216 probe (mean difference 6.3%; $P = 4.0 \times 10^{-4}$; estimate 0.27; SE 0.07). The Genehancer and dbSUPER enhancer

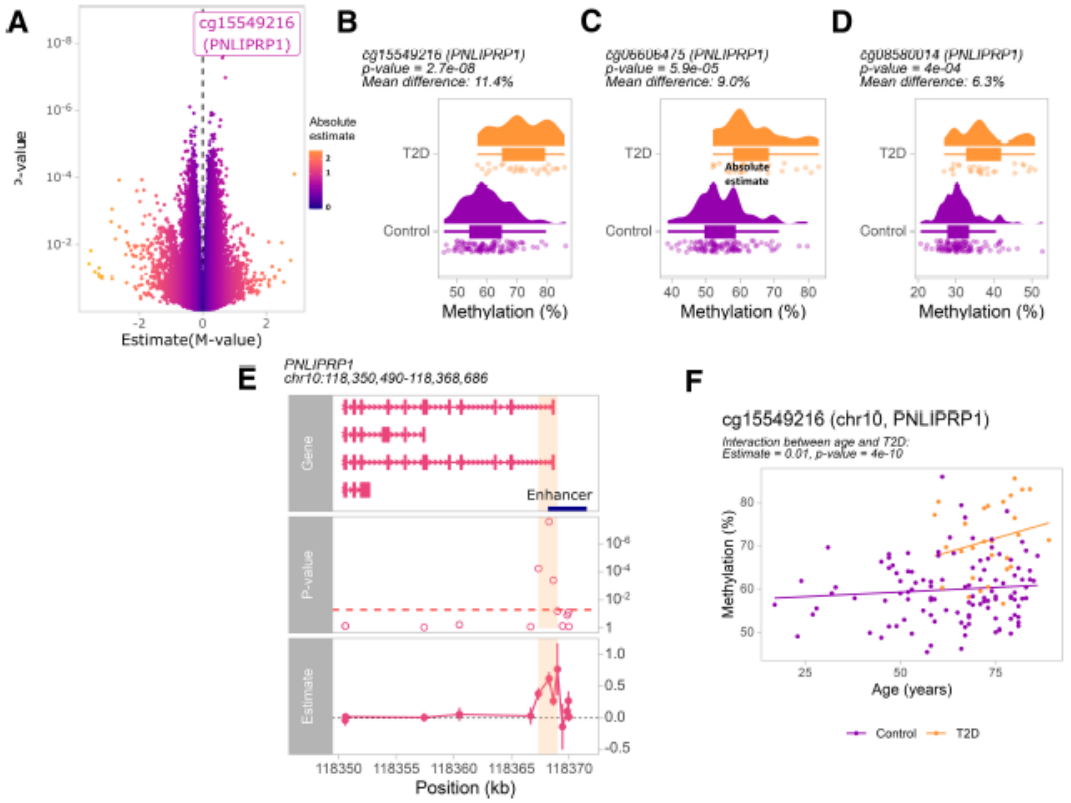


Figure 2— Epigenome-wide association study in whole pancreas in T2D. *A*: Volcano plot depicting the differentially methylated CpGs associated with T2D, highlighting the only significant CpG (Bonferroni-corrected $P < 0.05$). Colors indicate the absolute estimates (*M* values). *B*: Boxplot showing the density of cg15549216 methylation levels in individuals with T2D (orange) and controls without diabetes (purple), for each individual. The *x*-axis represents the percentage methylation β value. *B–D*: Boxplots showing the density of cg06606475 and cg08580014 methylation levels, respectively, in patients with T2D and controls without diabetes, for each individual. *E*: Representation of the *PNLIPRP1* gene, with the differentially methylated region shaded in orange. The enhancer region is depicted in purple, and the estimates and *P* values of all CpGs in the *PNLIPRP1* gene are shown. *F*: Scatterplot depicting the interaction between methylation of cg15549216 and age in organ donor individuals. Controls without diabetes are shown in purple, and patients with T2D are shown in orange.

databases showed that in human whole-pancreas tissue, the region in and surrounding the cg15549216 probe lay in a 12-kb stretch of the only known super-enhancer of the *PNLIPRP1* gene (20,21) (Fig. 2*E*).

Because aging is one of the major risk factors for T2D (22), we verified whether the cg15549216 CpG was associated with age in our cohort. We found that age was nominally associated with the hypermethylation of the cg15549216 probe (unadjusted $P = 0.010$); however, this association was not significant after adjustment for T2D status (unadjusted $P = 0.93$). We considered whether both T2D and age could be contributing factors in cg15549216 methylation by including an interaction term between T2D status and age in our model. At the cg15549216 probe, we found an increase in methylation in individuals with T2D as age increased compared with controls without diabetes (M value estimate 0.01; $P = 4 \times 10^{-10}$) (Fig. 2*F*). The cg15549216 probe was not associated with BMI ($P = 0.54$), sex ($P = 0.32$), T2D duration ($P = 0.39$), or statin treatment ($P = 0.32$).

Rare *PNLIPRP1* Null Variants Are Associated With Metabolic Traits

To investigate the putative role of *PNLIPRP1* in glucose or lipid metabolism, we tested whether rare variants (MAF <1%) in the *PNLIPRP1* gene were associated with T2D and related metabolic traits, including glucose and lipid traits. Using whole-exome sequencing data from up to 191,000 participants in the UK Biobank, we identified 44 null variants (i.e., nonsense, frameshift, canonical ±1 or 2 splice sites) (Supplementary Table 4).

We found that *PNLIPRP1* null variants were associated with increased glycemia ($P_\pi = 1.1 \times 10^{-3}$; effect size 0.13; SE 0.040). We also found that *PNLIPRP1* rare variants were significantly associated with increased LDL-C ($P_\pi = 0.034$; effect size 0.10; SE 0.049), HDL-C ($P_\pi = 0.026$; effect size 0.05; SE 0.022), waist-to-hip ratio ($P_\pi = 2.7 \times 10^{-3}$; effect size 0.23; SE 0.006), waist circumference ($P_\pi = 7.6 \times 10^{3}$; effect size 2.19; SE 0.820), BMI ($P_\pi = 3.1 \times 10^{-3}$; effect size 0.039; SE 0.011), diastolic blood pressure ($P_\pi = 2.02 \times 10^{-5}$; effect size 2.89; SE 0.677), and systolic blood pressure ($P_\pi = 2.9 \times 10^{-3}$; effect size 3.30; SE 1.112). Null variants were not associated with T2D risk ($P_\pi = 0.48$), obesity ($P_\pi = 0.29$), hypertension ($P_\pi = 0.94$), or triglyceride levels ($P_\pi = 0.46$) (Table 2).

PNLIPRP1 Common Variants Are Associated With LDL-C but Not T2D- or Glucose-Related Traits

Using the Type 2 Diabetes Knowledge Portal (23), we found that common SNPs (MAF >1%) located at the *PNLIPRP1* locus were strongly associated with increased LDL-C ($P = 2.0 \times 10^{-14}$) (Supplementary Tables 5 and 6). In addition, common variants in *PNLIPRP1* were associated with non-HDL cholesterol, cholesterol levels, and apolipoprotein B after multiple testing ($P \leq 2.5 \times 10^{-6}$), but not with T2D- or glucose-related traits, in up to 1.61 million participants (Supplementary Fig. 5).

Causal Relationship Between *PNLIPRP1* Methylation and T2D-Related Traits

To identify whether *PNLIPRP1* methylation was causally associated with T2D and LDL-C, and the reverse, we performed MR. To identify SNPs acting as proxies for *PNLIPRP1* methylation, we genotyped 111 control individuals in our cohort and performed mQTL. For T2D, we used 118 proxy SNPs associated with increased T2D risk and six proxy SNPs associated with *PNLIPRP1* rmethylation (Supplementary Tables 7 and 8). Using the IVW method, we found evidence that increased T2D risk was causal of cg15549216 hypermethylation, with an estimate of 0.23 (95% CI 0.029–0.43; $P = 0.025$) (Fig. 3*A*). However, we did not find a significant association using MR methods following different assumptions, such as the

Table 2—*PNLIPRP1* null variant associations

Trait	*n* of individuals	Π	SE	P_π	*n* of variants
Diastolic BP	168,374	2.889	0.677	2.0×10^{-5}	39
WHR	190,739	0.023	0.006	2.7×10^{-3}	41
Log BMI	187,727	0.039	0.011	3.1×10^{-3}	40
Glucose	159,764	0.133	0.041	1.1×10^{-3}	38
Systolic BP	168,367	3.303	1.1119	2.9×10^{-3}	39
Waist circumference	190,751	2.188	0.820	7.6×10^{-3}	41
HDL	156,077	0.050	0.022	2.6×10^{-2}	38
LDL	169,625	0.104	0.049	3.4×10^{-2}	38
Obesity	107,219	−0.146	0.138	0.29	31
Log triglycerides	169,815	−0.009	0.011	0.46	39
T2D	171,651	0.177	0.252	0.48	39
Hypertension	180,290	0.010	0.126	0.94	40

BP, blood pressure; WHR, waist-to-hip ratio.

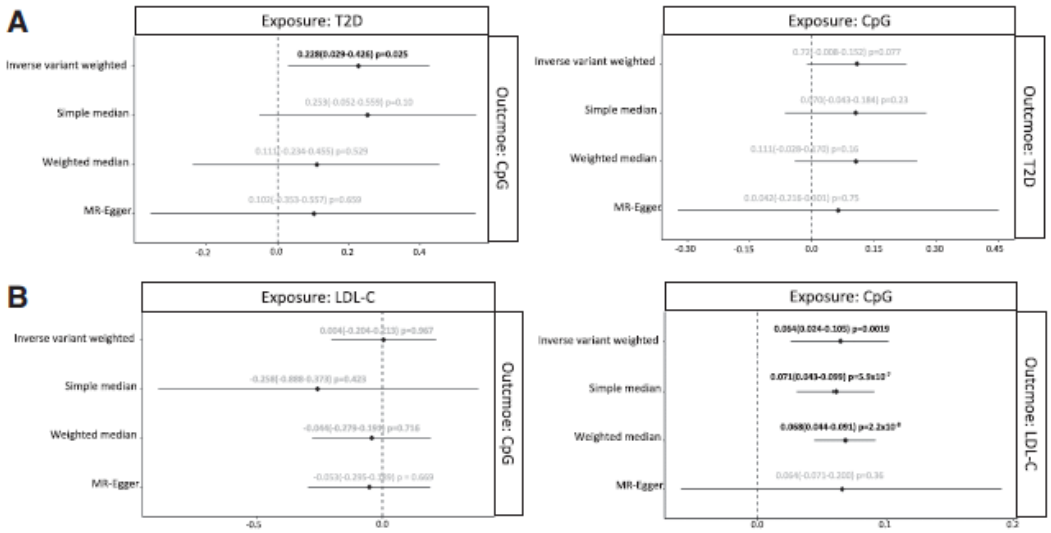


Figure 3— MR to determine causality. Bidirectional causality between T2D and *PNLIPRP1* methylation (*A*) and between LDL-C and *PNLIPRP1* methylation (*B*) using the IVW, simple median, weighted median, and MR-Egger methods. The *x*-axis represents the estimates. Statistically significant associations are shown in black text, and nonsignificant values are shown in gray, with estimates and 95% CIs in parentheses.

simple median and weighted median methods, possibly pointing to the presence of horizontal pleiotropy or heterogeneity in our genetic instruments. To ensure the validity of our results, we used another robust MR method, MR-Egger, which showed no evidence of causal association (estimate 0.102; 95% CI −0.353 to 0.557; P = 0.659), and ruled out the possibility of horizontal pleiotropy (intercept 0.011; 95% CI −0.024 to 0.045; P = 0.55). In the reverse direction, we found no evidence of a causal effect of *PNLIPRP1* methylation on T2D risk (Fig. 3*A*).

For LDL-C, we used nine proxy SNPs to represent *PNLIPRP1* methylation and 241 proxy SNPs for increased LDL-C levels (Supplementary Tables 9 and 10). We found no evidence of a causal effect of LDL-C on *PNLIPRP1* methylation; however, *PNLIPRP1* methylation was associated with increased LDL-C levels, with an estimate of 0.064 (95% CI 0.024–0.105; P = 0.0019) (Fig. 3*B*). This result was consistent with simple median (estimate 0.071; 95% CI 0.043–0.099; $P = 5.9 \times 10^{-7}$) and weighted median methods (estimate 0.068; 95% CI 0.044–0.091; $P = 2.2 \times 10^{-8}$). The MR-Egger intercept showed no evidence of horizontal pleiotropy (estimate 0.000; 95% CI −0.012 to 0.012; P = 0.31) (Fig. 3*B*). In all significant MR analyses, the leave-one-out analysis found that no single SNP altered the results, suggesting that the observed association was robust (Supplementary Tables 11 and 12). Altogether, our data provide some evidence that *PNLIPRP1* hypermethylation increases LDL-C levels and suggest a trend that T2D status increases cg15549216 methylation.

Correlation of cg15549216 CpG With *PNLIPRP1* Gene Expression in Whole Pancreas

To determine whether the expression of *PNLIPRP1* was dysregulated in T2D, we performed gene expression quantification of *PNLIPRP1* in extracted RNA from whole-pancreas tissue in a subset of five donors with T2D and six controls without diabetes, who were matched for age, sex, and BMI (Supplementary Table 13). We found that *PNLIPRP1* expression was downregulated in donors with T2D (53%; P = 0.011) (Fig. 4*A*). *PNLIPRP1* expression was not associated with age (P = 0.17). We performed linear regression to determine whether methylation levels were associated with *PNLIPRP1* gene expression. We found that cg15549216 methylation was significantly correlated with decreased *PNLIPRP1* gene expression (P = 0.042; R^2 = 0.35) (Fig. 4*B*).

PNLIPRP1 Expression in Whole Pancreas Is Restricted to Acinar Cells

To identify target cell types of *PNLIPRP1* dysregulation, we examined the expression of *PNLIPRP1*. In the GTEx database (24), *PNLIPRP1* was exclusively expressed in the whole pancreas, compared with 53 tissues (Supplementary Fig. 6), and the TIGER database (25) showed that *PNLIPPR1* was mainly expressed in the whole pancreas (median transcript per million [TPM] 2,581), compared with pancreatic islets (median TPM 36) (Supplementary Fig. 7). We analyzed several single-cell pancreas data sets (26,27) and found *PNLIPRP1* was mainly expressed in acinar cells (Supplementary Table 14). To confirm at the protein level the acinar expression of *PNLIPRP1*, we performed immunofluorescence in human pancreatic tissue stained

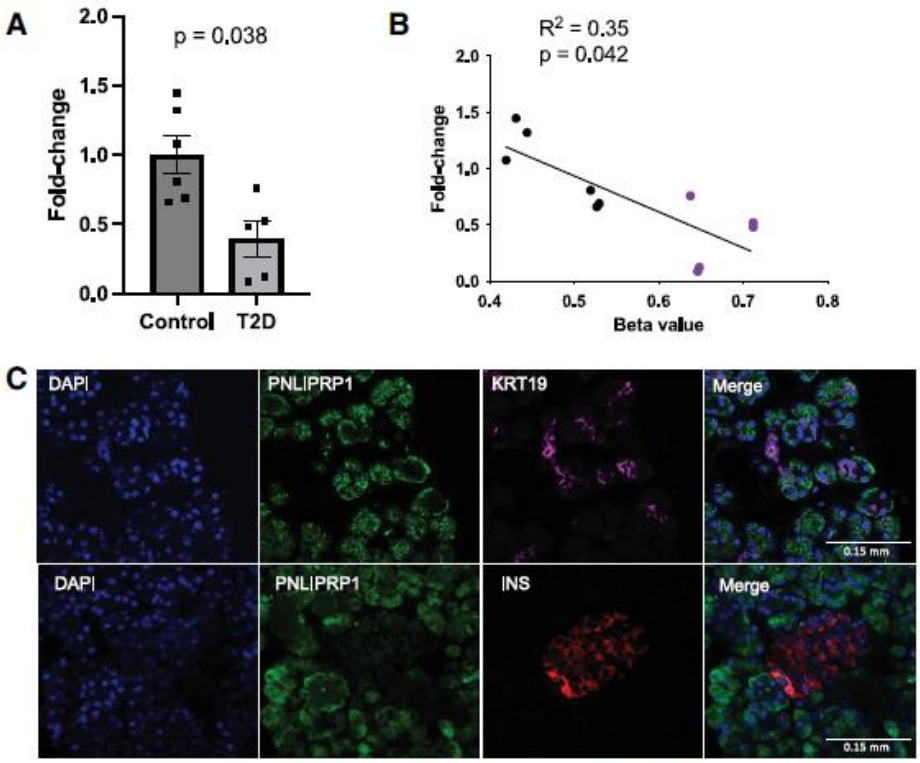


Figure 4—Expression of PNLIPRP1 at the RNA and protein levels. *A*: RNA expression of *PNLIPRP1* in a subset of 11 individuals (five with T2D vs. six controls without diabetes), matched for age, sex, and BMI. *t* tests were performed to test for differences between T2D and nondiabetic samples for the *PNLIPRP1* gene. Error bars represent SEs. *B*: Correlation between the cg15549216 probe and RNA expression of the *PNLIPRP1* gene determined by qPCR and compared with housekeeping gene *RPLP0*. Dots shown in black are donors without diabetes, and those in purple are individuals with T2D. *C*: Immunofluorescence of healthy human whole-pancreas tissue samples stained for PNLIPRP1 and KRT19, a ductal marker (top), and insulin (INS; a marker for pancreatic islets). Nuclei were stained with DAPI.

for PNLIPRP1, KRT19, a marker of ductal cells, and insulin to mark pancreatic islets. This revealed that the PNLIPRP1 protein was not expressed in pancreatic islets or ductal cells but exclusively in acinar cells (Fig. 4*C*).

PNLIPRP1 KD Induces Decreased Cell Cycle and Increases Cholesterol Biosynthesis Genes

Based on the acinar-specific expression pattern of *PNLIPRP1*, we explored the functional role of *PNLIPRP1* in the rat acinar cell line AR42J. To investigate the downstream consequences of *PNLIPRP1* dysregulation, we performed KD of *PNLIPRP1* in the AR42J cell line. We confirmed that our *PNLIPRP1* KD induced a 60% reduction in gene expression of *PNLIPRP1* (Fig. 5*A*) and a 34% decrease at the protein level (Supplementary Fig. 8). RNA sequencing confirmed that *PNLIPRP1* was one of the most significant downregulated genes (false discovery rate 0.025). We explored potential dysregulated pathways by focusing on the 1,024 genes with an unadjusted $P < 0.05$ (Supplementary Table 15). Using enrichR, we found that the top pathway for the downregulated genes (587 genes) was the cell-cycle pathway (adjusted $P = 2.21 \times 10^{-12}$) (Supplementary Table 16), and the most upregulated pathways (437 genes) were cholesterol biosynthesis and SREBP control of lipid biosynthesis (adjusted $P = 0.0044$) (Supplementary Table 17 and Fig. 5*B*). We confirmed using MTS proliferation that *PNLIPRP1* KD led to a 22% reduction in cell proliferation ($P = 0.0027$) (Fig. 5*C*). We also confirmed that total cholesterol content was increased by 29% after *PNLIPRP1* KD ($P = 0.0017$) (Fig. 5*D*). We then found that treatment of *PNLIPRP1* KD with simvastatin (i.e., a statin drug that inhibits the cholesterol synthesis) was able to rescue the increase in cholesterol biosynthesis to levels of the control (Fig. 5*E*). Taken together, these findings demonstrate that the invalidation of *PNLIPRP1* dysregulates in an inverse direction cholesterol metabolism and cell-cycle regulation.

An increase in cholesterol content and a reduction in proliferation are indicative of ADM, a response to acinar

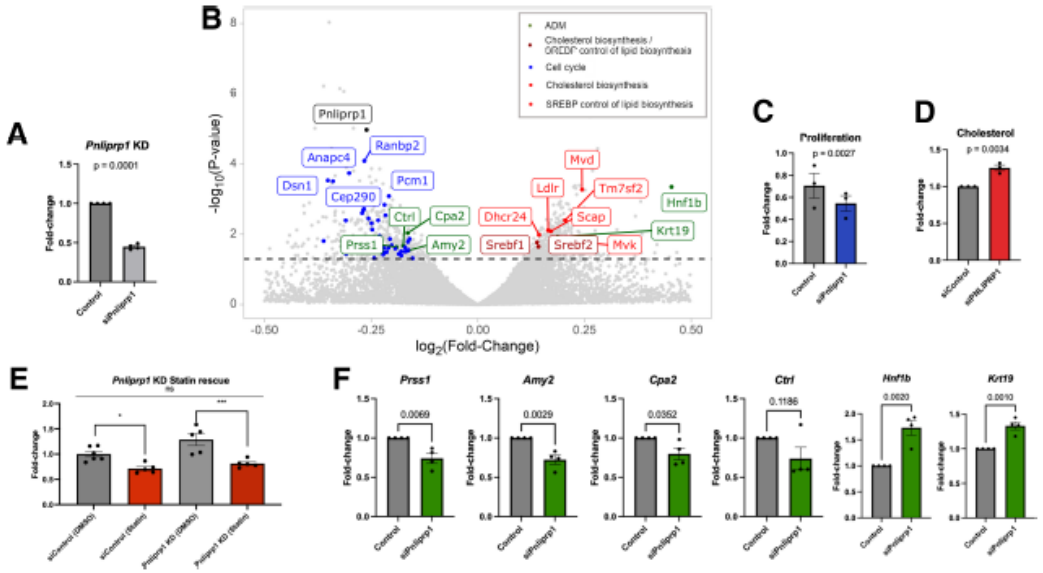


Figure 5— PNLIPRP1 KD in AR42J rat cell line. *A*: Confirmation of *PNLIPRP1* KD in four biologic replicates using qPCR and tested using a *t* test. *B*: Volcano plot of RNA sequencing results, depicting *PNLIPRP1* (black) and downregulated cell-cycle genes (blue). The top five most significant genes are labeled. Upregulated genes from cholesterol synthesis and SREBP control of lipid biosynthesis genes (red and dark red) and ADM genes (green) are labeled. *C*: MTS proliferation assay in AR42J in *PNLIPRP1* KD after 72 h of siRNA treatment. *D*: Total cholesterol measured after 48 h of *PNLIPRP1* KD. All performed in three biologic replicates. *E*: Statin treatment after *PNLIPRP1* KD, compared with untreated DMSO controls. Performed in two biologic replicates. Error bars represent SEs. *F*: Gene expression of acinar and ductal markers using qPCR in AR42J in PNLIPRP1 KD after 72 h of siRNA treatment. *P* values are from *t* tests performed between each gene in the siRNA of PNLIPRP1 (siPnliprp1) and untargeted control (siControl). The *y*-axis shows the fold change, and error bars represent SEs. Experiments were performed in four biologic replicates. $*P < 0.05$, $***P < 0.001$. ns, not significant.

cells stress (28,29). Therefore, we assessed the expression of acinar and ductal markers in our RNA sequencing data, data subsequently confirmed by qPCR. We found that the expression of four acinar markers was downregulated: *Prss1* ($P = 0.0066$), *Amy2* ($P = 0.0029$), *Cpa2* ($P = 0.035$), and *Ctrl* (*Ctrl* was downregulated in three of four replicates [$P = 0.12$]). Two ductal markers were upregulated: *Krt19* ($P = 0.0010$) and *Hnf1b* ($P = 0.020$) (Fig. 5*F*). Taken together, these results suggest that *PNLIPRP1* invalidation induces ADM.

Reduction of *PNLIPRP1* Expression in Response to Diabetogenic Exposure

To address whether T2D induces dysregulation of *PNLIPRP1* and mimics the phenotype observed in *PNLIPRP1* KD, we assessed the impact of a diabetogenic environment (i.e., high glucose and high insulin) on AR42J cells. We confirmed that AR42J cells were responsive to insulin treatment by analyzing phosphorylated Akt (Supplementary Fig. 9). We found that high glucose and insulin induced a decrease in *PNLIPRP1* expression (Fig. 6*A*). We tested whether the treatment induced glucose uptake and found it decreased glucose in the medium of treated cells ($P < 0.0001$) (Fig. 6*B*), indicating increased glucose uptake by the cells. We confirmed using the MTS proliferation assay that the diabetogenic milieu led to a 25% reduction in cell proliferation ($P = 0.0089$) (Fig. 6*C*). Total cholesterol content was unchanged ($P = 0.1002$) (Fig. 6*D*). We also assessed acinar and ductal markers and found that the expression of two acinar markers was downregulated (*Cpa2* [$P = 0.0281$] and *Ctrl* [$P = 0.0022$]), whereas *Prss1* and *Amy2* ($P > 0.05$) were unchanged. Ductal markers were significantly upregulated: *Krt19* ($P = 0.0140$) and *Hnf1b* ($P = 0.0344$) (Fig. 6*E*). Taken together, these results suggest that high glucose and insulin treatment induced downregulation of *PNLIPRP1*, a decrease in proliferation, and ADM in these cells. We performed immunofluorescence in a T2D sample, compared with a healthy control, and did not find any notable differences in PNLIPRP1 expression, but we did observe that some cells expressed both ductal marker KRT19 and PNLIPRP1 expression, suggesting an ADM phenotype (Supplementary Fig. 10).

DISCUSSION

To our knowledge, this is the first epigenetic study of the impact of T2D on the whole pancreas. We found a strong association between T2D and hypermethylation of the cg15549216 probe, located in the only known enhancer of the *PNLIPRP1* gene.

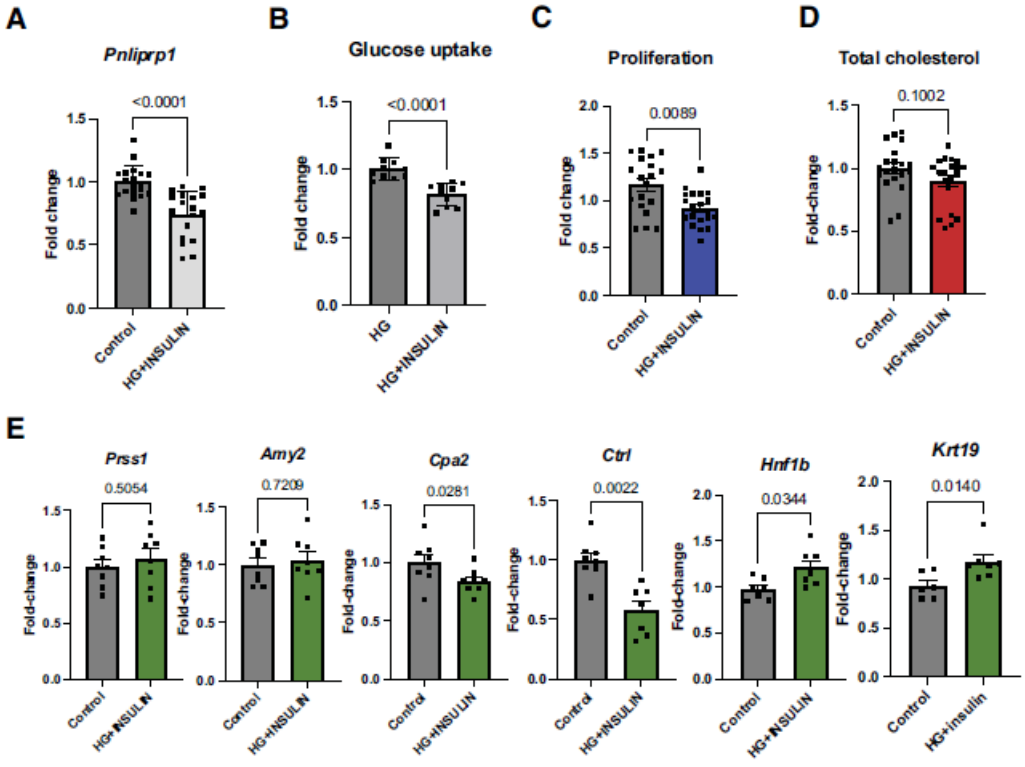


Figure 6—High-glucose (HG) and insulin treatment of AR42J rat cell line. *A*: RNA expression of *PNLIPRP1* after HG and insulin treatment, determined by qPCR; statistical significance was determined using a two-tailed *t* test. Data are representative of four biologic replicates. *B*: Glucose uptake of PNLIPRP1 after HG (20 mmol/L glucose) and insulin (100 nmol/L) treatment, compared with untreated controls, determined by qPCR; statistical significance was determined using a two-tailed *t* test. Data are representative of six biologic replicates. *C*: MTS proliferation assay in 48 h AR42J treated with HG and insulin, compared with controls. *D*: Total cholesterol measured after 48 h of HG and insulin, compared with controls. *E*: Gene expression of acinar and ductal markers in HG- and insulin-treated cells, compared with controls, in AR42J. *P* values are from *t* tests performed for each gene. The *y*-axis shows the fold change, and error bars represent SEs. Experiments were performed in three biologic replicates.

Additional observations revealed that *PNLIPRP1* is a strong candidate for T2D-driven epigenetic changes, leading to deleterious events, particularly in acinar cell cholesterol metabolism. This is supported by several lines of evidence: *1*) cg15549216 hypermethylation was causal of increased LDL-C, *2*) partial extinction of *PNLIPRP1* increased cholesterol metabolism in functional studies, and *3*) this cholesterol impairment was rescued by statin treatment in vitro. Furthermore, our *PNLIPRP1* extinction studies also showed that acinar cells started to express ductal markers and cell-cycle arrest, which both suggest ADM. Cholesterol synthesis and ADM are functionally linked; several studies have shown that cholesterol biosynthesis triggers ADM, and this is reversed in vitro through statin (28).

Although ADM should in theory be reversible, during sustained stress these cells are strongly associated with progression to pancreatic ductal adenocarcinoma (PDAC), the most common malignancy of the exocrine pancreas (30–33). A recent study followed the time course of mouse PDAC using single-cell transcriptomics and revealed that after ADM, metaplastic cells had a signature distinct from that of ductal cells and that *PNLIPRP1* was significantly downregulated in these metaplastic cells compared with acinar cells (34). Additionally, *PNLIPRP1* is one of the most significantly downregulated genes in PDAC tumors (35), and there are somatic deletions in *PNLIPRP1* in PDAC tissue (36), pointing to a role of this gene in pancreatic reprogramming toward malignancy. Additionally, in vitro and in vivo models have shown that statins delayed the formation of PDAC in mutated mice (37,38).

This is reinforced by human studies, where statin use resulted in a significant reduction in pancreatic cancer risk and improved patient survival (39). For instance, a

meta-analysis that included 14 studies showed that statin use was associated with an overall reduction in death in individuals with PDAC, particularly significant in individuals who had undergone surgery, but not in those with advanced disease (40). Indeed, in patients with advanced cancer, the late addition of statins to standard anti-cancer therapy did not show improvement in overall survival. These findings reveal the protective properties of statins, especially when administered early and before disease development, highlighting the importance of studying the early events in pancreatic injury (41).

We found that rare and common genetic variants were associated with increased LDL-C and metabolic traits. Specifically, rare variants were associated with LDL-C, HDL-C, BMI, glucose levels, waist-to-hip ratio, waist circumference, diastolic blood pressure, and systolic blood pressure, indicating a systemic effect on metabolic disease. However, we did not find an association with disease states, including T2D, obesity, or hypertension. This could be due to the involvement of *PNLIPRP1* in intermediate metabolic pathways or disease progression stages. This could also be due to limitations in statistical power. Further research with larger cohorts and more detailed phenotypic data may help clarify these relationships. Our null variant data are supported by *PNLIPRP1* knockout mice, which display mild hyperglycemia and impaired insulin sensitivity, exacerbated by a high-fat diet (42). Additionally, evolutionary studies have shown that *PNLIPRP1* was consistently lost in herbivore or carnivore species consuming a low-fat diet (43), suggesting a role in lipid metabolism. Taken together, our study data suggest that both nature (genetic variation) and nurture (through epigenetic marks) contribute to the regulation of *PNLIPRP1* and various aspects of systemic metabolic health.

This study has several important limitations to consider. First, the sample size was relatively small, underscoring the need for larger exocrine pancreas biobanks to better explore and understand the DNA methylation landscape in this tissue. Second, functional experiments were conducted using the AR42J rat cell line because of challenges associated with working with human exocrine tissue. Because the methylation region is not conserved in rat cells, further research is required to directly link DNA methylation with *PNLIPRP1* gene expression in human cells. Finally, although our findings indicate an association between DNA methylation and gene expression, we lacked sufficient tissue to investigate the impact of DNA methylation on protein expression. Addressing this link is crucial for a more complete understanding.

Understanding the sequence of events leading to pancreatic disease in humans is challenging and necessitates a comprehensive and integrative approach. Our findings not only underscore the importance of understanding the underlying molecular mechanisms relating T2D and pancreatic disease but also suggest preventive strategies (e.g., statin or other lipid metabolism modifiers to protect exocrine tissue). Despite the challenges inherent in studying the exocrine pancreas, unraveling the molecular intricacies linking T2D and pancreatic diseases remains crucial for advancing our understanding of these conditions.

Funding. This study was funded by the Société Francophone du Diabète (SFD-R21100EE/RAK21020EEA) and French National Research Agency (Agence Nationale de la Recherche [ANR]–10-LABX-46 [European Genomics Institute for Diabetes] and ANR-10-EQPX-07-01 [Lille Integrated Genomics Advanced Network for Personalized Medicine]), the EU Horizon Europe Research and Innovation Programme (OBELISK grant agreement 101080465), the France Génomique consortium (ANR-10-INBS-009). This work was supported by the National Center for Precision Diabetic Medicine – PreciDIAB (ANR-18-IBHU-0001; FEDER 2000189I/NP0025517; MEL 2019_ESR_11). A.B. is supported by the European Research Council (Opi0 101043671).

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

Author Contributions. L.Mau., M.B., L.N., M.C., and A.K. performed the statistical analyses and generated the figures. L.Mau., R.B., A.L., and J.F. performed the functional work. L.Mau and V.P. performed the MR analyses. L.Mar., C.D.L., A.J., S.L., M.S., and P.M. provided organ donor samples and clinical data. B.T. and S.A. performed RNA sequencing. M.D. and A.B. performed UK Biobank analyses. F.P. and J.K.-C. provided pancreatic tissue samples. P.F. and A.K. designed the study. All authors critically reviewed and edited the manuscript. A.K. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the 56th Annual Meeting of the European Association for the Study of Diabetes (EASD), virtual conference, 22–25 September 2020; 8th Meeting of the Study Group on the Genetics of Diabetes. Lille, France, 11–13 May 2022; Société Francophone de Diabète (SFD), Montpellier, France, 21–24 March 2023; SFD, Toulouse, France, 19–22 March 2024; and 59th Annual Meeting of the EASD, Hamburg, Germany, 3–6 October 2023.

References

1. Wynne K, Devereaux B, Dornhorst A. Diabetes of the exocrine pancreas. J Gastroenterol Hepatol 2019;34:346–354
2. Gál E, Dolenšek J, Stožer A, et al. Mechanisms of post-pancreatitis diabetes mellitus and cystic fibrosis-related diabetes: a review of preclinical studies. Front Endocrinol (Lausanne)2021;12:715043
3. Wondmkun YT. Obesity, insulin resistance, and type 2 diabetes: associations and therapeutic implications. Diabetes Metab Syndr Obes 2020;13:3611–3616
4. Rawla P, Sunkara T, Gaduputi V. Epidemiology of pancreatic cancer: global trends, etiology and risk factors. World J Oncol 2019;10:10–27
5. Yuan S, Kar S, Carter P, et al. Is type 2 diabetes causally associated with cancer risk? Evidence from a two-sample Mendelian randomization study. Diabetes 2020;69:1588–1596
6. Ewald N, Bretzel RG. Diabetes mellitus secondary to pancreatic diseases (type 3c): are we neglecting an important disease? Eur J Intern Med 2013; 24:203–206
7. Zechner D, Knapp N, Bobrowski A, et al. Diabetes increases pancreatic fibrosis during chronic inflammation. Exp Biol Med (Maywood) 2014;239:670–676
8. Solanki NS, Barreto SG, Saccone GTP. Acute pancreatitis due to diabetes: the role of hyperglycaemia and insulin resistance. Pancreatology 2012;12: 234–239
9. Noel RA, Braun DK, Patterson RE, Bloomgren GL. Increased risk of acute pancreatitis and biliary disease observed in patients with type 2 diabetes: a retrospective cohort study. Diabetes Care 2009;32:834–838

10. Wright JJ, Eskaros A, Windon A, et al. Exocrine pancreas in type 1 and type 2 diabetes: different patterns of fibrosis, metaplasia, angiopathy, and adiposity. Diabetes 2024;73:1140–1152
11. Solimena M, Schulte AM, Marselli L, et al. Systems biology of the IMIDIA biobank from organ donors and pancreatectomised patients defines a novel transcriptomic signature of islets from individuals with type 2 diabetes. Diabetologia 2018;61:641–657
12. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30:1363–1369
13. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic Acids Res 2016;44:e20
14. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28:882–883
15. Houseman EA, Kile ML, Christiani DC, et al. Reference-free deconvolution of DNA methylation data and mediation by cell composition effects. BMC Bioinformatics 2016;17:259
16. Sun J, Zheng Y, Hsu L. A unified mixed-effects model for rare-variant association in sequencing studies. Genet Epidemiol 2013;37:334–344
17. Xue A, Wu Y, Zhu Z, et al.; eQTLGen Consortium. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat Commun 2018;9:2941
18. Graham SE, Clarke SL, Wu K-HH, et al.; Global Lipids Genetics Consortium. The power of genetic diversity in genome-wide association studies of lipids. Nature 2021;600:675–679
19. Delaneau O, Ongen H, Brown AA, et al. A complete tool set for molecular QTL discovery and analysis. Nat Commun 2017;8:15452
20. Fishilevich S, Nudel R, Rappaport N, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford) 2017;2017:bax028
21. Khan A, Zhang X. dbSUPER: a database of super-enhancers in mouse and human genome. Nucleic Acids Res 2016;44:D164–D171
22. Fazeli PK, Lee H, Steinhauser ML. Aging is a powerful risk factor for type 2 diabetes mellitus independent of body mass index. Gerontology 2020;66:209–210
23. Costanzo MC, von Grotthuss M, Massung J, et al.; AMP-T2D Consortium. The Type 2 Diabetes Knowledge Portal: an open access genetic resource dedicated to type 2 diabetes and related traits. Cell Metab 2023;35:695–710.e6
24. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580–585
25. Alonso L, Piron A, Morán I, et al. TIGER: the gene expression regulatory variation landscape of human pancreatic islets. Cell Rep 2021;37:109807
26. Baron M, Veres A, Wolock SL, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Syst 2016;3:346–360.e4
27. Li J, Klughammer J, Farlik M, et al. Single-cell transcriptomes reveal characteristic features of human pancreatic islet cell types. EMBO Rep 2016;17:178–187
28. Grisan F, Spacci M, Paoli C, et al. Cholesterol activates cyclic AMP signaling in metaplastic acinar cells. Metabolites 2021;11:141
29. Carrer A, Trefely S, Zhao S, et al. Acetyl-CoA metabolism supports multistep pancreatic tumorigenesis. Cancer Discov 2019;9:416–435
30. Neuhöfer P, Roake CM, Kim SJ, et al. Acinar cell clonal expansion in pancreas homeostasis and carcinogenesis. Nature 2021;597:715–719
31. Chuvin N, Vincent DF, Pommier RM, et al. Acinar-to-ductal metaplasia induced by transforming growth factor beta facilitates KRASG12D-driven pancreatic tumorigenesis. Cell Mol Gastroenterol Hepatol 2017;4:263–282
32. Liu J, Akanuma N, Liu C, et al. TGF-β1 promotes acinar to ductal metaplasia of human pancreatic acinar cells. Sci Rep 2016;6:30904
33. Shi G, DiRenzo D, Qu C, et al. Maintenance of acinar cell organization is critical to preventing Kras-induced acinar-ductal metaplasia. Oncogene 2013;32:1950–1958
34. Schlesinger Y, Yosefov-Levi O, Kolodkin-Gal D, et al. Single-cell transcriptomes of pancreatic preinvasive lesions and cancer reveal acinar metaplastic cells' heterogeneity. Nat Commun 2020;11:4516
35. Zhang G, He P, Tan H, et al. Integration of metabolomics and transcriptomics revealed a fatty acid network exerting growth inhibitory effects in human pancreatic cancer. Clin Cancer Res 2013;19:4983–4993
36. Tong Y, Sun M, Chen L, et al. Proteogenomic insights into the biology and treatment of pancreatic ductal adenocarcinoma. J Hematol Oncol 2022;15:168
37. Guillaumond F, Bidaut G, Ouaissi M, et al. Cholesterol uptake disruption, in association with chemotherapy, is a promising combined metabolic therapy for pancreatic adenocarcinoma. Proc Natl Acad Sci USA 2015;112:2473–2478
38. Hao F, Xu Q, Wang J, et al. Lipophilic statins inhibit YAP nuclear localization, co-activator activity and colony formation in pancreatic cancer cells and prevent the initial stages of pancreatic ductal adenocarcinoma in KrasG12D mice. PLoS One 2019;14:e0216603
39. Zhang Y, Liang M, Sun C, et al. Statin use and risk of pancreatic cancer: an updated meta-analysis of 26 studies. Pancreas 2019;48:142–150
40. Tamburrino D, Crippa S, Partelli S, et al. Statin use improves survival in patients with pancreatic ductal adenocarcinoma: a meta-analysis. Dig Liver Dis 2020;52:392–399
41. Farooqi MAM, Malhotra N, Mukherjee SD, et al. Statin therapy in the treatment of active cancer: a systematic review and meta-analysis of randomized controlled trials. PLoS One 2018;13:e0209486
42. Ren J, Chen Z, Zhang W, et al. Increased fat mass and insulin resistance in mice lacking pancreatic lipase-related protein 1. J Nutr Biochem 2011;22:691–698
43. Wagner F, Ruf I, Lehmann T, et al. Reconstruction of evolutionary changes in fat and toxin consumption reveals associations with gene losses in mammals: a case study for the lipase inhibitor PNLIPRP1 and the xenobiotic receptor NR1I3. J Evol Biol 2022;35:225–239