Abrrevations

ALP: Alkaline Phosphatase ALT: Alanine Aminotransferase AST: Aspartate Aminotransferase BCAA: Branched-Chain Amino Acids CRP: C-Reactive Protein DAG: Directed Acyclic Graph DEXA: Dual-Energy X-ray Absorptiometry GGT: Gamma-Glutamyl Transferase GIP: Gastric Inhibitory Polypeptide GLP-1: Glucagon-Like Peptide 1 IL-6: Interleukin 6 IRR: Incidence Rate Ratio LC-MS/MS: Liquid Chromatography–Mass Spectrometry/Mass Spectrometry MAFLD: Metabolic Associated Fatty Liver Disease MR: Mendelian Randomization NMR: Nuclear Magnetic Resonance OGTT: Oral Glucose Tolerance Test PCA: Principal Component Analysis PRS: Polygenic Risk Score UMAP: Uniform Manifold Approximation and Projection

Project title

Metabolic Signatures: Decoding incretin and glucagon pathways and their interaction with metabolic traits in cardiometabolic disease progression

Brief project description (max 2000 characters including spaces. Current count: 2088)

People at high risk of type 2 diabetes are not currently being systematically identified, primarily because of variation in who progresses to diabetes. This project investigates the causal role of natural incretin and glucagon responses, aiming to identify the primary drivers of progression toward diabetes and diabetes related complications. The project will use the deep phenotyped data in the Fenland Study.

The Fenland Study is well suited as the selected database because it combines OGTT based incretin and glucagon measurements, genetics, proteomics, metabolomics, DEXA derived adiposity, liver function biomarkers, and long term cardiometabolic outcomes in a general population cohort. The study includes three phases of examination from 2005 to 2025, which allows for the assessment of individual trajectories in cardiometabolic risk. We will validate findings in the Danish ADDITION PRO study and the UK Biobank. The project is divided into three parts. First, using traditional and newer methods of causal inference, including genetically predicted and phenotypically measured traits, we aim to determine the causal relationship between incretin and glucagon responses and long term cardiometabolic outcomes and to identify mediating pathways such as liver health, inflammation and metabolite profiles. Second, we will investigate incretin and glucagon responses in conjunction with metabolic traits that contribute to heterogeneity in metabolic dysfunction risk, such as liver health and inflammation, and aim to decipher the impact of each metabolic trait on long term cardiometabolic risk.

Third, we will identify and predict machine learning based metabolic clusters using clinical and omics data.

By combining new epidemiological methods and machine learning techniques to characterize individual risk of progression toward cardiometabolic disease, and by clarifying the causal role of natural incretin and glucagon release, including important mediating pathways, this proposal aims to help identify individuals who have the greatest potential benefit from early intensive intervention and to pinpoint targetable biomarkers.

Lay project description (max 1000 characters including spaces. Current count: 637)

Type 2 diabetes develops differently from person to person, and today we cannot easily identify who is most at risk. This project uses data from the large Fenland Study to understand how two natural gut hormones, incretin and glucagon, influence the body's ability to control blood sugar. By combining hormone measurements with information on liver health, body fat, inflammation, genes and metabolism, we aim to discover why some people progress to diabetes or other cardiometabolic diseases while others do not. The project will also create simple tools that can help identify people who may benefit from early and targeted prevention.

Project description (max 20,000 characters including spaces. Current count: 36,000)

Specific aim

This Inter-SUSTAIN Part 2 project aims to use the Fenland study to elucidate the role of natural incretin and glucagon responses as drivers of heterogeneity in the long-term risk of diabetes and related complications, and to map clinical and biological traits associated with these responses, with particular emphasis on liver health, adiposity, and inflammation. In doing so, we aim to identify patterns of metabolic dysfunction that may predict progression to type 2 diabetes and inform personalized treatment strategies.

The specific aims are to:

a) Phenotypically and genotypically investigate the causal role of natural incretin and glucagon responses to an oral glucose tolerance test (OGTT) in determining the risk of regression to normoglycemia, progression to type 2 diabetes, and the development of diabetes-related complications. This will be conducted using data from the Fenland study and the Danish ADDITION-PRO cohort, linked to comprehensive outcomes from the Danish National Health Registries.

b) Identify metabolites and biomarkers for liver health, adiposity and inflammation that mediate the association between phenotypically measured (Fenland study, ADDITION-PRO) and genetically predicted (Fenland study, ADDITION-PRO, UK Biobank) incretin and glucagon responses and cardiometabolic outcomes, with particular emphasis on direct and indirect markers of liver function.

c) Map dimensions of metabolic traits in conjunction with incretin and glucagon responses to decipher their role in deteriorating glucose metabolism and their relationship to long-term cardiometabolic health.

d) Predict clustered dimensions of metabolic traits using clinical markers, metabolites, and genes.

Background

Pre-diabetes is a complex state associated with increased risk of progression to type 2 diabetes and related complications[1–3]. However, not all people with pre-diabetes progress to diabetes; many persist in the prediabetic state or regress to normoglycemia over the course of various years[2]. Remission to normoglycemia has been shown to reduce the risk of cardiovascular morbidity and mortality[4], however, it remains unclear to which extend if the remission is driven by lifestyle changes or by a heterogeneous low-risk subgroup.

Recently the idea of metabolic heterogeneity in type 2 diabetes has been extended to pre-diabetes, and traits that distinguish individuals most likely to progress to diabetes, such as insulin-resistant fatty liver and visceral adiposity-related renal dysfunction, have been identified[1]. In recent years, incretins (Glucagon Like Peptide 1 [GLP 1] and Gastric Inhibitory Polypeptide [GIP]) and the hormone glucagon have gained increasing attention as targets in therapeutics for prevention of type 2 diabetes and diabetes-related complications[5–9]. GLP-1 and GIP enhance post-meal insulin secretion, supporting efficient nutrient handling[10]. Glucagon serves as a counter-regulatory hormone to insulin and maintains glucose balance both during fasting and following glucose intake[11]. However, in the context of heterogeneity in pre-diabetes, the role of natural incretin and glucagon responses for the progression to type 2 diabetes and diabetes-related complications remains unexplored. Prior research in the Danish ADDITION-PRO study showed that, individuals with pre-diabetes and type 2 diabetes exhibited up to 25% lower natural GLP-1 responses to oral glucose[12]. Additionally, impaired glucagon suppression and elevated fasting glucagon levels have been observed in individuals with insulin resistance and early glucose dysregulation[13], involving hepatic insulin resistance that impairs the turnover of branched-chain amino acids[14]. However, these findings have not been limited investigated in relation to prospective cardiometabolic outcomes[15]. A limitation of the ADDITION-PRO cohort is the selection of individuals at high risk of diabetes based on diabetes-specific risk scores, which limits the generalizability of the findings to the broader populations.

The Fenland Study represents a pioneering population-based cohort that has substantially advanced the precision of identifying individuals at high risk of type 2 diabetes[16] and elucidating causal pathways underlying cardiometabolic disease [17][18][19][20]. Leveraging extensive phenotypic and genetic data, the study has established causal relationships between liver function markers, alanine aminotransferase (ALT), aspartate aminotransferase (AST), and alkaline phosphatase (ALP), and type 2 diabetes[17], as well as between Dual Energy X-ray Absorptiometry (DEXA) derived adiposity traits, fat-mass ratios, and metabolic dysfunction-associated liver disease[20]. Moreover, Fenland data have uncovered mechanistic links between interleukin-6–mediated inflammation and type 2 diabetes[21], highlighting the interplay between inflammatory and metabolic pathways. The cohort's precise characterization of adiposity, inflammation, and hepatic health, combined with detailed assessments of incretin, glucagon, insulin, and glucose dynamics in a general population, provides an unparalleled resource for investigating the constellation of metabolic traits and their long-term impact on cardiometabolic health.

This project is grounded in the hypothesis that dysregulated incretin secretion and inadequate glucagon suppression, together with insulin resistance, accelerate the progression of metabolic dysfunction by activating hepatic and inflammatory pathways in individuals with at higher risk of diabetes.

Work Package 1: Investigate causal relationship of neutral release of incretin and glucagon and long-term cardiometabolic disease and identify mediating pathways.

Hypothesis for Work Package 1: Natural incretin and glucagon responses play a causal role as drivers of the risk of progression from pre-diabetes to type 2 diabetes and diabetes-related complications as well influence cardiometabolic trajectories. In this pathway, biomarker and metabolites related to liver health play a mediating role.

Natural incretin and glucagon reposes in epidemiological studies have been mostly investigated cross-sectionally, with a few exceptions[15]. The Fenland study includes ~12,000 individuals at baseline who underwent measurement of natural incretin and glucagon responses during a 2-point (0, 120 min) oral glucose tolerance test (OGTT) in 2005-2015, i.e. before the introduction of incretin based therapies. This unique timing provides an opportunity to prospectively analyze the associations between individuals's natural hormonal responses to glucose and progression to diabetes, regression from prediabetes to normoglycemia, and risk of diabetes-related complications. Preliminary findings from ADDITION PRO in a population with high risk of diabetes showed that per SD higher glucagon (incidence rate ratio IRR: 1.38 CI: 1.15 to 1.67) and GLP-1 responses (IRR: 1.34 CI: 1.05 to 1.71) at 120 minutes during the OGTT are associated with incident heart failure over an 11-year follow-up period.

As part of the collaboration between the MRC Epidemiology Unit and Novo Nordisk Foundation Center for Basic Metabolic Research, a research group are currently developing PRS for GLP-1, GIP, and glucagon. Genetic variants, clinical biomarkers and metabolites serve

as valuable indicators of individual predisposition for specific phenotypes and mediating pathways (metabolites and clinical biomarkers) that link phenotypes to disease outcomes. We want to use newly developed PRS to strengthen causal inference by applying them as instrumental variables in Mendelian randomization studies[22]. Identification of metabolic pathways that mediate the effect of genetically determined phenotypes on the incidence of diseases such as type 2 diabetes helps us understand the biological mechanisms that link traits to disease. This understanding in turn supports the development of new targeted interventions based on metabolic pathways.

Research design and methods

To quantify the causal relationship between incretins, glucagon and cardiometabolic disease, we will first apply traditional epidemiological time-to-event analysis methods (Poisson and Cox regression) within the Fenland Study, using adjustments informed by pathways outlined in directed acyclic graphs (DAGs). In addition, we aim to characterize trajectories of cardiometabolic profiles including glucose, lipids, kidney function, adiposity markers from DEXA, liver markers and inflammatory biomarkers using linear mixed effects models. Evaluating the risk associated with incretin and glucagon responses to a glucose load in isolation may be limited without considering concurrent insulin and glucose regulation. To address this in WP1, principal component analysis (PCA) and unifold manifold approximation and projection (UMAP) will be applied to extract the strongest sources of variance in the concurrent OGTT response across all 4 hormones and glucose.

To strengthen causal inference and account for unmeasured confounding, we will conduct Mendelian randomization using pathway specific polygenic risk scores (PRS) as instruments for incretins and glucagon [@daveysmith2014], leveraging data from the Fenland Study, ADDITION-PRO and the UK Biobank [@mrcepidemiologyunit].

To sequence mediating pathways and identify key metabolites linking incretin and glucagon responses to cardiometabolic disease, we will apply a structural causal algorithm (NetCoupler)[23] to identify sequential causal pathways involving metabolites. Based on the identified metabolites, structured causal mediation analyses[24] will be conducted to assess and quantify the direct and indirect effects of individual metabolites seen as most likely to be causal mediators in NetCoupler on cardiometabolic outcomes.

Work Package 2: To map dimensions of metabolic traits in conjunction with incretin and glucagon responses, in order to decipher their role in the context of deteriorated glucose metabolism

Hypothesis for Work Package 2: natural incretin and glucagon responses may exert protective, compensatory effects under conditions of obesity, hyperglycemia, insulin-resistance, and loss of beta-cell function. We further hypothesize that liver health and low-grade inflammation may

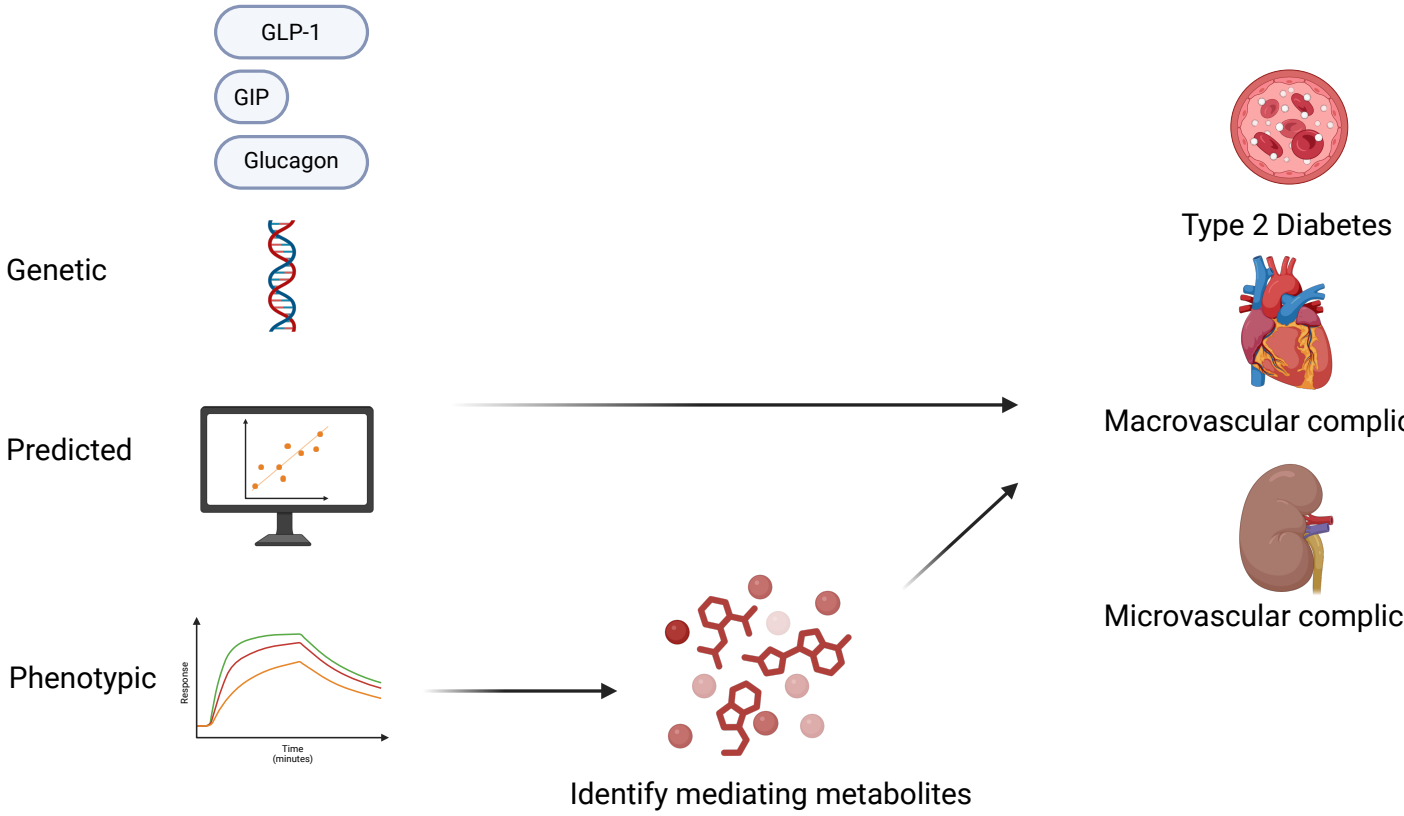

Figure 1: **Figure 1: Structured causal analysis framework of incretin and glucagon responses and their association with the risk of progression to cardiometabolic disease**

act as an effect modifier in the associations between incretin-glucagon responses and metabolic outcomes.

Obesity and insulin resistance have long been regarded as central components in the development of type 2 diabetes. Recent clustering analyses have identified liver fat as a main determinant heterogeneity among people with pre-diabetes with regard to their risk of progression to type 2 diabetes[1]. Reduced liver function interacts with insulin resistance, glucose and glucagon responses, and is accompanied by low grade inflammation and morphological changes to the liver, particularly MAFLD (Metabolically Associated Fatty Liver Disease)[25;26;11][14]. The biological actions of GLP-1 and GIP include reducing hepatic inflammation[27]. Liver function therefore appears to be a central pathway modulating diabetes risk and interacting with or mediating the impact of natural incretin and glucagon responses. The Fenland study includes baseline measurements of -glutamyl transferase and alanine transaminase, and 11,559 participants had DEXA[28] performed allowing to estimate regional fat mass.

[Deep learning proccessing of ultrasound images] In conjunction with liver function biomarkers, this will enable the calculation of a general indication of liver health.

We will further integrate dimensions of metabolically relevant low-grade inflammation by integrating markers of Interleukin 6 (IL-6), adiponectin, and C-reactive protein (CRP)[29] in our multidimensional characterization of liver health. In addition, fasting acylcarnitines, amines, sphingolipids and phospholipids were obtained in Fenland study. These data allow us to characterize aspects of metabolic function, such as liver fat accumulation[30]and tissue-specific insulin resistance[31,32] .

Based on phenotypes, metabolites, and genotypes, we want to identify metabolic traits that, in conjunction with incretin and glucagon responses, contribute to cardiometabolic risk. Machine learning–based dimensionality reduction techniques help identify and characterize distinct constellations of metabolic traits across individuals. From phenotypic clusters, we want to map dimensions of metabolic function that either provide compensatory protection against or contribute to cardiometabolic disease. Then, we will add dimensions of metabolite-based profiles to the existing clusters to enhance their characterization.

To support phenotypic clustering, genetic data will be incorporated through partitioned polygenic risk scores (PRS) for type 2 diabetes[33][34]. These scores will be added with underlying metabolic traits such as glucagon and incretin, inflammation, beta-cell function, obesity, liver fat, and insulin resistance. This approach enables us to decipher phenotypes driven by genetic risk scores that contribute to cardiometabolic disease.

Research design and methods

We use data from previous work in The Fenland study and ADDITION-PRO, including genotypes, metabolomic profiles, and markers of metabolism and inflammation[12,13,29,35–38]. By

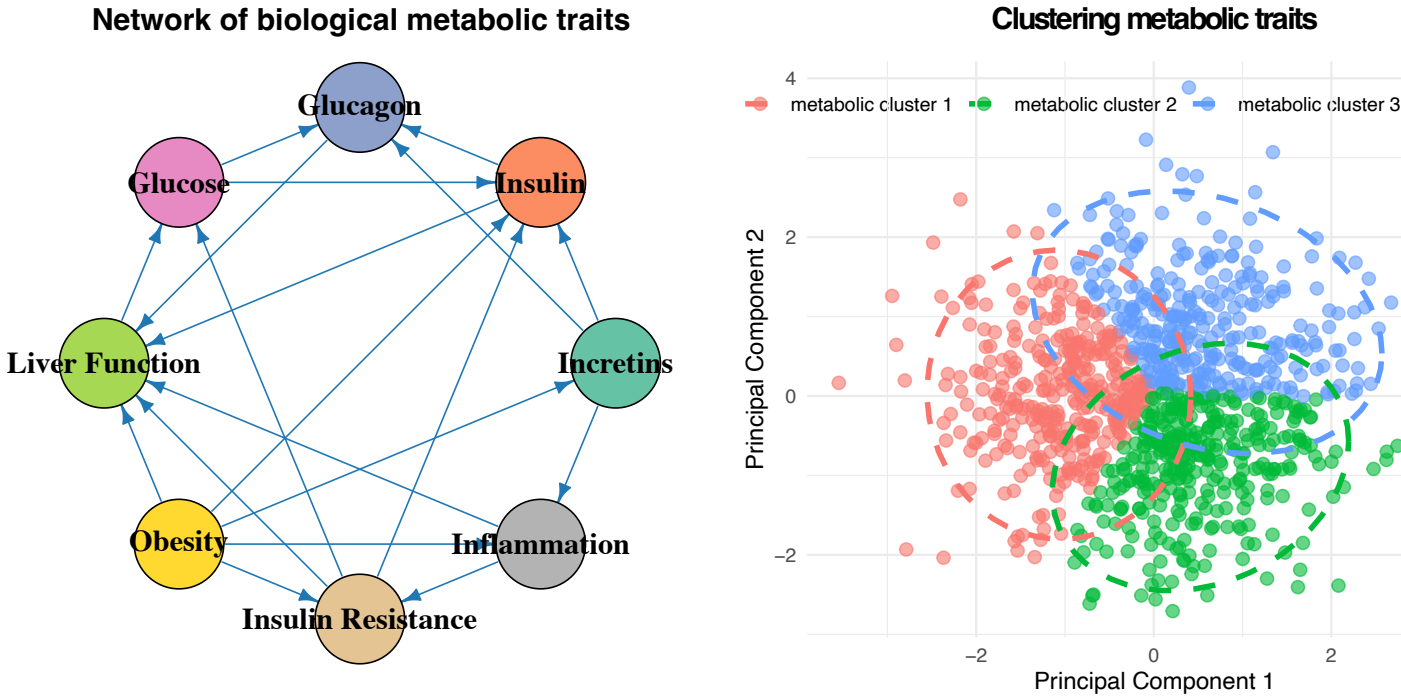

Figure 2: Figure 2: Illustration of dimensionality reduction of metabolic traits into metabolic clusters

combing different liver function biomarkers from images and blood measurements, we will be enable to generate a general, multidimensional indication of liver health.

We will apply PCA and UMAP to integrate variables and identify a restricted number of metabolic dimensions that reflect the interplay between incretin and glucagon responses along with concurrent glucose-insulin dynamics during the OGTT[39]. We will further add dimensions of obesity, insulin resistance, inflammation, liver health markers, subsequently examine these dimensions in relation to diabetes risk[40]. Finally, metabolite-based profiles and partitioned PRS will be analyzed separately in relation to cardiometabolic outcomes and subsequently incorporated to enrich the dimensional metabolic profiling.

Work Package 3: To predict incretin and glucagon responses in conjuction with metabolic traits based clinical and biological traits

Hypothesis for Work Package 3: population heterogeneity of incretin and glucagon responses in conjunction with metabolic traits can be predicted to a clinically useful degree based on clinical, genetic and metabolic biomarkers, and these predictions can stratify individuals by their risk of progression to type 2 diabetes and related complications.

Given the fact that it is impractical to assess a full incretin and glucagon response in clinical practice, WP3 aims to establish a practical set of biomarkers that jointly have the capacity to characterize incretin and glucagon responses. Analyses in the Fenland study will be carried out across three levels: (1) traditional clinical characteristics, (2) metabolomics profiles, and (3) genomic data. The utility of these predicted traits will be investigated in relation to cardiometabolic outcomes in large population-based cohorts that include the same biomarkers but have not measured incretin and glucagon responses, such as the UK Biobank, as well as using available markers in Danish registries.

[Add in conjuction with CM traits]

Research design and methods

To predict incretin and glucagon responses based on clinical and metabolite data, we will apply a two-layered modeling approach in the Fenland study. Clinical and metabolic profiles will be used separately and in combination to predict responses. For variable selection in both layers, we will apply two strategies: restricting variance and using linear LASSO regression to shrink and select the most predictive variables. To support this, we will employ machine learning models to capture undefined interactions and identify important predictors based on variable importance[41]. Our approach follows established principles in machine learning and statistical modeling, adhering to standardized practices for prediction, reporting, and validation[42–45].

These PRSs will be incorporated into the aforementioned predictive layers as composite predictors of incretin and glucagon responses. PRS for other relevant traits such as obesity, insulin

resistance, beta-cell function, liver function and low-grade inflammation will also be considered where relevant.

Study description

Feature	Fenland study	ADDITION-PRO Study
Baseline data collection period	2005-2015 (Phase 1)	2009–2011
Follow-up	2014-2020 (Phase 2)2023-2025 (Phase 3)	Danish National Registries (up to 2024)
Included participants	UK general population aged 30-55 years	Individuals at high risk of diabetes
Number of participants	12,435	2,082
Measure from OGTT	GLP-1, GIP, glucagon, glucose and insulin (t= 0, 120 min)	GLP-1, GIP, glucagon, glucose and insulin (t= 0, 30, 120 min)[12,13,37]
Other relevant metabolic measures	**Adiposity from DEXA scan:**- Visceral and subcutaneous fat- Bone density**Liver function:**- Ultrasound liver images- GGT, ALT- glucagon – alanine index- Alkaline Phosphatase**Inflammation:**- Interleukin 6 (IL-6), high sensitivity CRP, adiponectin	**Adiposity:**- Visceral and subcutaneous fat from ultrasonography- adiponectin**Liver function:**- Ultrasound liver images (still B-mode images with liver protocol)- GGT, ALT- glucagon – alanine index**Inflammation:**- soluble CD163, and high sensitivity CRP
Assessment method of metabolites	LC electrospray ionization and flow-injection analysis tandem MS (ref), targeted metabolomics	Proton nuclear magnetic resonance spectroscopy, targeted metabolomics[35]
Sample tissue	Fasting plasma blood samples	Both fasting and during the OGTT[35,36]
Number of metabolites	175 (acylcarnitines, amines, sphingolipids and phospholipids)	231 lipid-related and 3 BCAA (fasting isoleucine, leucine and valine levels)
Genotyping	Yes	Yes
Follow-up period	10-15 years	13-years