University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Interfacing Food & Medicine

Single-Cell Transcriptomics of Intestinal Epithelial Cells: Insights into the Prediabetic condition

Glenn Ross-Dolan

B.A. (Mod) Microbiology - Trinity College Dublin

Project Thesis in partial fulfilment for the degree of Masters in Bioinformatics and Computational Biology

Supervised by Dr. Silvia Melgar, APC Microbiome Ireland

Table of Contents

List of abbreviations

Acknowledgements

Abstract

Single-Cell Transcriptomics of Intestinal Epithelial Cells: Insights into the Prediabetic condition
Supervisor: Silvia Melgar, APC Microbiome Ireland, S.Melgar@ucc.ie

Background: Obesity is a chronic condition characterised by excess adiposity that may be accompanied by different structural and functional abnormalities and with increased comorbidity and premature mortality risk thus reducing quality of life (Jastreboff et al., 2019; Lustig et al., 2022).

The main cause of obesity is the energetic imbalance due to increased caloric intake and little expenditure. This induces metabolic and hormonal changes e.g. increase in blood sugar levels that induces a prediabetic status with increased risk of developing type 2 diabetes mellitus (T2DM), heart disease, and stroke.

T2DM is a major non-communicable disease and one of the world's fastest growing health problems, with a projected increase in the number of diabetic patients to 700 million by 2045 (Saeedi et al., 2019). T2DM is associated with significant morbidity, including increased risk of cardiovascular diseases (CVD) and stroke, hypertension, etc. These place an enormous burden on individuals, society, and the healthcare system (Brorsson & Pociot, 2015). T2DM is a non-reversible but preventable condition with overweight and obesity being major risk factors. The onset of T2DM is gradual, with most individuals progressing from normoglycemia through a pre-diabetic state.

There is substantial evidence for the role of gut microbiota and impaired barrier in metabolic diseases including T2DM (Brunkwall & Orho-Melander, 2017). Recent clinical trials using Glucagon-like peptide-1 (GLP-1) receptor agonists (an incretin hormone produced by the enteroendocrine L cells in the distal intestine) have shown benefits to patients with these conditions. However, a role

for other intestinal epithelial cell subsets in obesity and diabetes is yet to be determined. Thus, the aim of this project is to:

1) Identify alterations in intestinal epithelial cell subsets
2) Identify alterations in pathways contributing to obesity and/or diabetes pathophysiology.

For this purpose, the student will be performing a bioinformatics analysis on in house and publicly available bulk and single cells RNASeq data sets. Findings from this project, will provide a better understanding on how gut epithelial cells, which are in close contact to the microbiota, might support and/or sustain the development of these chronic conditions.

Aim

1) Identify alterations in intestinal epithelial cell subsets
2) Identify alterations in pathways contributing obesity and/or diabetes pathophysiology.

Chapter 1. The global epidemic of metabolic disease. Prediabetes and Diabetes Type 2

1. Type II Diabetes Mellitus

Type II Diabetes Mellitus (T2DM), is a global health concern growing to pandemic proportions. The number of people with T2DM has increased four fold over the last three decades and is the ninth leading cause of death and approximately x in 100 people above the age of 18 have T2DM in the United States of America. With the rapid developments in technology and urbanisation the lifestyle of the average person across the globe has changed to one with increased obesity rates, sedentary lifestyles and increased ageing, the driving forces behind the disease. Particularly over the last two decades, obesity has become a pressing issue in every country across the globe leaving no exceptions and leading to a vast array of health complications. For these reasons, there is an urgent demand to take actions in slowing down the rate of obesity together with the complications associated with obesity, one of which being T2DM. T2DM is currently defined as a chronic metabolic disorder characterised by insulin resistance and an inadequate insulin secretory response ultimately leading to hyperglycemia. This definition follows the scheme set by the American Diabetes Association which classifies diabetes into three categories: type 1, type 2 and other specific forms such as genetic defects or secondary to other conditions. Characterised by insulin resistance and impaired insulin secretion, T2DM aetiology involves a intricate mix of genetic, environmental and lifestyle factors.

Risk Factors

Overweight and obesity in adults is defined as a body mass index greater than 25 kg/m2 and x kg/m2 respectively. Although high BMI is not a direct measure of health but rather size, a large portion of deaths from non communicable diseases are driven by high BMI. Approximately 80% of which are from diabetes, stroke, heart disease and cancer. Three quarters of this death and

disease in adults is occurring in middle-income countries. In essence, the majority of people living and dying with non-communicable diseases have high BMI and are living in developing/lower resource regions of the globe. Current data from the world obesity atlas reports that data are showing that there is a positive correlation of increased GDP and increased levels of overweight/obesity. There also appears to be positive correlations between high BMI and GHG emissions, increased urbanisation, increased plastic waste, sedentary lifestyle, and consumption of animal proteins, sugars and sweeteners. Gradual weight gain serves as the trigger for later metabolic issues, with Type 2 diabetes mellitus (T2DM) being the most closely linked to obesity. In some cases however, T2DM may develop before obesity in those with innate insulin resistance, where increased glucose production and insulin levels ultimately trigger obesity.

Diet seems to be intimately related to the progression of various diseases, particularly metabolic disorders and with this, the search for high quality diets has been steadily increasing over the last number of decades. One group led by Toi et al. performed an umbrella review to unveil evidence in diet interventions and factors for preventing the progression of T2DM. In this study, sixty systematic reviews with meta analyses of randomised controlled trial/observational studies were analysed. The effect of each dietary intervention was analysed in its effect on the risk of T2DM. Results of the review indicate that dietary patterns such as Mediterranean (diets high in fruit, vegetables, legumes, beans, cereals, grains, fish, foods with high unsaturated fats like olive oil) and Dietary Approaches to Stop Hypertension diets, and diets with a high healthy eating index were beneficial in preventing T2DM. Specific food groups which exhibited beneficial effects in preventing T2DM included whole grain, olive oil, low fat dairy, fibre, magnesium and flavonoids. Contrastingly, diets exhibiting high glycemic index/glycemic load as well as red meats, processed meats, sugar, and artificial sweeteners indicate an accelerating effect of progressing T2DM. Physical activity. Frequent physical activity plays a crucial role in managing type 2 diabetes mellitus. Numerous studies have shown that aerobic exercise, resistance training as well as a combination of the two can

lead to significant improvements of the various systems in the bodies of people with T2DM. Often reported are the improvements in blood pressure, glucose regulation, insulin sensitivity, HbA1c levels as well as reduction in body fat/ visceral adipose tissue. Moreover, the extent of these improvements are proportional to the total energy consumed during the training rather than the intensity or the length of the training exclusively.

Smoking of tobacco has also been identified as a contributing factor in the progression of type 2 diabetes mellitus. Epidemiological research reveals distinct links between tobacco smoking and risk of developing T2DM. Similarly, medical research suggests that tobacco smoking and nicotine use directly affects insulin sensitivity, body composition and pancreatic beta cell mechanisms. Rimm et al. revealed that men smoking more than 25 cigarettes per day had nearly double the risk of developing T2DM compared to non-smokers. Another study investigated the impact of both active and passive smoking on diabetes risk using data from the HIPOP-OHP Study in Japan. Over a median follow-up of 3.4 years, the study found that individuals exposed to passive smoke at work and active smokers had significantly higher risks of developing diabetes compared to non-exposed individuals. Additionally, cross-sectional studies including EPIC-Norfolk and the National Health and Nutrition Examination Survey, found that smoking is linked to higher HbA1c levels indicating poorer glycemic control.

Although these epidemiological studies have shown that the incidence of T2DM is correlated with the incidence of smoking, this alone does not demonstrate that smoking drives the progression of the disease. Numerous studies have also investigated physical mechanisms affirming these links. Although studies show that smoking results in a reduced body weight, it also has a negative effect on body composition. People that smoked more than 20 cigarettes per day exhibited an adjusted odds ratio of 1.93 for abdominal obesity compared to those who have never smoked. Similarly, other cross-sectional studies of men and women in the UK as well as Japanese men exhibit a greater waist-to-hip ratio compared to non-smokers. These changes in body composition appear to

be caused primarily by nicotine signalling suggesting that non-tobacco alternatives may not mitigate these risk factors. Smoking also exhibits negative effects on glucose tolerance and insulin sensitivity. A number of studies demonstrate that the glucose tolerance and insulin sensitivity index was reduced after smoking in a group of persistent smokers as well as total body glucose being reduced in smokers with diabetes versus non-smokers with diabetes. Furthermore other studies demonstrate that clamp analysis shows that the total glucose disposal in the body was lowered in smokers compared to those who do not. It is also apparent that use of nicotine, such as use through nicotine patches also decreases the efficacy of insulin.

Finally, clinical studies note that smoking disrupts B cell function. In studies involving Japanese men as well as studies on Swedish men, higher incidence of impaired insulin secretion as well as lower B cell function as measured by HOMA-B is exhibited compared to non-smokers. These effects are observed even after adjusting for other factors such as age, BMI, alcohol intake and physical activity.

These links may stretch past just the use of tobacco smoking into the realm of smokeless tobacco use, specifically in the use of tobacco pouches called ‘snus’. A number of studies are reporting an increased incidence of T2DM in people that use more than 5 boxes of snus per week at an odds ratio of 3.3. There has been mixed results however with other studies reporting no apparent correlation although, a recent meta analysis reported a hazard ratio of 1.15 in snus users compared to those who have never used snus. This is a concerning figure given the lack of knowledge on tobacco free nicotine products, as well as the immense growth that the snus industry has seen in recent years with sales increasing from 126 million units in 2019 to 808 million units in march 2022 in the USA.

Alcohol intake has a complex relationship with risk of T2DM. Moderate consumption of alcohol is associated with a lower risk of developing T2DM compared to abstainers suggesting a potential protective effect although excessive alcohol consumption is linked to increased risk due to its contribution to obesity and insulin resistance. Several studies point to this ‘U-shaped’

relationship between alcohol consumption and T2DM risk. Reductions in risk in alcohol consumers seems to be somewhat confined to women and non-Asian populations.

As discussed previously, most risks associated with T2DM can be mitigated through lifestyle and dietary changes. There still however remain some risks which are not entirely in the individual's control, genomics. Genomics plays a significant role in T2DM susceptibility with genetic factors influencing a persons predisposition to overweight/obesity, insulin resistance and overall metabolic profile. The genetic susceptibility of Type 2 Diabetes mellitus is well documented and studied through twin and family based studies. Typically this involves mutations in genes regulating glucose levels and glucose-homeostasis hormones. However, most cases of T2DM involving genetic factors are driven by both the genetic and environmental factors together, otherwise known as epigenetics.

Microbiome is also linked to T2DM as a risk factor.

Diagnoses

Management of T2DM

Complications of T2DM

2. Prediabetes and Metabolic Syndrome

Defining Prediabetes

Prediabetes is a condition similar to diabetes where the metrics used to diagnose diabetes such as glucose homeostasis and blood glucose levels are elevated past normal levels but not elevated enough to classify someone as being diabetic. More specifically, prediabetes is characterised by impaired fasting glucose, impaired glucose tolerance, raised HbA1c levels or a combination of the three. The exact specifications of these three metrics are not

agreed upon in the scientific literature leading to some discrepancies in prevalence estimates. The state of being prediabetic is associated with obesity, dyslipidemia, and hypertension and is considered a risk factor for developing T2DM as well as some other pathologies such as cardiovascular disease. Both prediabetes and T2DM share common risk factors including overweight, sedentary lifestyle, smoking, and genetics and both are associated with increased cardiovascular risks. However the cardiovascular risk is significantly higher in those with T2DM compared to those with prediabetes.

Prediabetes Pathologies

The state of being prediabetic is described as being physiologically stressful, having pathological changes linked to several systems in the body most notably the circulatory, nervous, digestive, urinary and endocrine systems.

The most common pathologies associated with prediabetes include microvascular complications, neurological disorders, cardiovascular disorders and metabolic disorders.

Microvascular disorders such as nephropathy, retinopathy and neuropathy are prevalent in individuals with prediabetes. Research is indicating that prediabetes particularly with impaired glucose tolerance is independently linked to peripheral neuropathy and nerve dysfunction likely due to an increased oxidative stress and activation of neurotoxic pathways (Lee et al., 2015). A systematic review and meta-analysis suggests a possible, modest increase in the risk of chronic kidney disease (CKD) in individuals with prediabetes, though the specific risk associated with impaired glucose tolerance (IGT) compared to impaired fasting glucose (IFG) remains unclear, warranting further research to clarify this relationship (Echouffo-Tcheugui et al., 2016). A systematic review suggests that the prevalence of retinopathy in prediabetic populations ranges from 0.3% to 20.9%, with a median of 8.1%, indicating higher retinopathy rates compared to individuals with normal glucose tolerance (Sune et al., n.d.).

T2DM is known to be a driver of cardiovascular issues and issues seem to also present during prediabetes. A review suggests that prediabetes is a toxic cardiometabolic state linked to increased cardiovascular risk, yet it is often underdiagnosed and undertreated, highlighting the need for improved screening, lifestyle interventions, and care coordination to prevent disease progression (Brannick and Dagogo-Jack, 2018). Supporting this is another review article suggesting that prediabetes is associated with an atherogenic lipid profile and increased risk of atherosclerotic cardiovascular disease, and recommends intensive lifestyle modification and statin therapy, particularly in high-risk individuals, to reduce cardiovascular risk and prevent progression to diabetes (Neves et al., 2022).

Metabolic disturbances are characteristic of both prediabetes and T2DM. One systematic review and meta-analysis supports this suggesting alterations in several key metabolite classes, including amino acids, lipids, and carbohydrates are characteristic of the prediabetic condition. Specifically, elevated levels of branched-chain amino acids (BCAAs), aromatic amino acids, and certain lipids (e.g., medium- and long-chain fatty acids, acylcarnitines) are associated with prediabetes and type 2 diabetes. These metabolic alterations may contribute to insulin resistance, dysregulated glucose metabolism, and increased cardiovascular risk, reflecting early signals of deteriorating glycemic control (Guasch-Ferré et al., 2016).

Prediabetes and T2DM represent different stages in the continuum of glucose metabolism disorders and although they are similar in many aspects they still have distinct pathological characteristics. Prediabetes is defined by fasting plasma glucose levels of 100–125 mg/dL (IFG), 2-hour plasma glucose during an oral glucose tolerance test (OGTT) of 140–199 mg/dL (IGT), or an A1C level of 5.7–6.4%. In contrast diabetes can be diagnosed if fasting plasma glucose is ≥126 mg/dL, 2-hour plasma glucose during an oral glucose tolerance test (OGTT) is ≥200 mg/dL, A1C is ≥6.5%, or if the individual with classic hyperglycemia symptoms has a random plasma glucose level ≥200 mg/dL (American Diabetes Association, 2021). Moreover, while prediabetes can

regress to normal glucose homeostasis through lifestyle interventions, T2DM often required pharmacological management to control blood glucose levels (Perreault et al., 2014).

Evidence has been emerging over the last decade exhibiting pathologies within the gut epithelium in the prediabetic subject. This is the focus of this study. Overall, prediabetes

Chapter 2: The intimate link between the intestinal epithelium and prediabetes.

3. The Intestinal Epithelium in Metabolic Health and Disease

Structure and function of the intestinal epithelium

The intestinal epithelium is a single layer of cells, forming the outermost lining of the small and large intestine. Since this layer is the first point of contact between the body and external environment, it plays several important and specialised roles including digestion, absorption as well as immune defense while remaining selectively permeable to nutrients. Furthermore the structure of the intestinal epithelium differs between the small and large intestinge where the small intestine contain villi and the large intstine does not. The villi are long protrusions which increase the surface area and subsequently the ability to absorb nutrients. These are absent in the large intestine where nutrient absorption is no longer as necessary and are unsuitable for solid debris which may damage them. The IE in the SI and LI also share similar structures, one being the presence of intestinal ‘crypts’ which harbour intestinal stem cells. Intestinal stem cells are undifferentiated cells which mature into various other cell types with specific functions. The axis from the crypt to the top of the villi is termed the ‘transition zone’ where ISCs give rise to daughter progenitor cells (AKA transit-amplifying cells), becoming increasingly mature/ differentiated. Mature IE cells eventually undergo apoptosis (anoikis) and shed into the lumen at top of the transition zone. The crypts are in a constant renewal cycle replacng itself every 4-5 days.

Various mature cell types are found in the intestinal epithelium with each possessing a unique role and with varying proportions of each across the gastrointestinal tract.

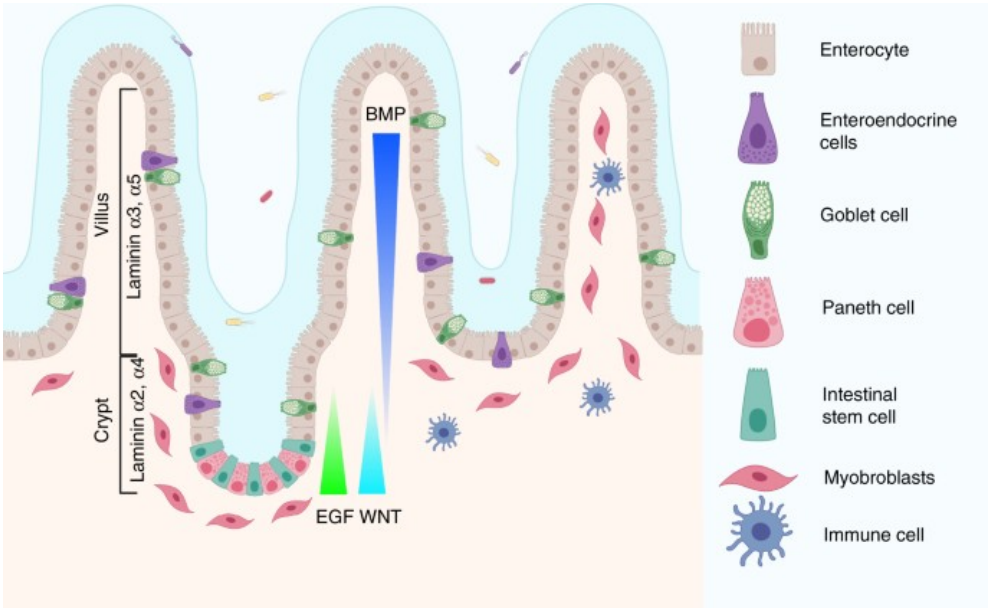

Figure 1: Intestinal Epithelium

The cell types include the **intestinal stem cells** involved in regenerating and producing new cells, **enterocytes** responsible for absorbing nutrients, water and electrolytes from digested foods. Secretory cells are also present such as **goblet cells** that secrete mucin protecting and lubricating the mucosal surface, **enteroendocrine cells** that secret hormones and signalling molecules regulating digestion and metabolism, and **paneth cells** which secrete antimicrobial peptides and enzymes maintaining gut immunity and maintaining ISC function.

Alterations in intestinal epithelium during prediabetes

Gaps in current knowledge

4. Key Signalling Pathways in Intestinal Epithelial Cells and Prediabetes

Intestinal Stem Cells

As mentioned previously, intestinal stems cells have an important role in maintaining the integrity and function of the intestinal epithelium. The Lgr5 gene which encodes a receptor fr the growth factor R-spondin, is a key marker of these stem cells and is regulated by the Wnt/B-catening signalling pathways. The expression of Lgr5 is most prominent in cells at the crypt base, marking them as the primary canonical stem cells responsible for the continuous turnover of the epithelium (Choi et al., 2023). Additionally, together with Wnt/B-catenin; mTORC1 (expressed by Paneth cells) and Notch signalling pathways collectively regulate the balance between stem cell self-renewal and differentiation (McCauley et al., 2023; VanDussen et al., 2012). In the context of prediabetes, recent studies are showing that the transcriptional profile of ISC's are altered contributing to the dysregulation of the intestinal epithelium. Research is indicating that the HFD increases ISC counts and enhances their self-renewal and stemness properties, potentially enriching intestinal regeneration (Xie et al., 2020). Other studies show that HFHSD causes ISCs and progenitors to become hyperproliferative, leading to faster differentiation and cell turnover in the intestinal epithelium (Aliluev et al., 2021). These changes in ISC behavior and intestinal homeostasis under high-fat diets may partly explain the higher incidence of intestinal tumors in obese individuals.

Enterocytes

Enterocytes manage various functions of nutrient absorption as well as protecting the gut from the bacteria-dense environment. The transcriptional profile of enterocytes is complex and regulated by a number of transcription factors. Notably, Hnf4g has been identified as a key transcription factor, specifically expressed in mature enterocytes, while Hnf4a is associated with intestinal stem cells (Lindeboom et al., 2018). Additionally the GATA

transcription factors are essential for the differentiation and proliferation of enterocytes along the transition zone as it regulates gene expression patterns critical for intestinal growth and function (Beuling et al., 2011; Middendorp et al., 2014). Similar to ISCs, the transcriptional landscape is also influenced by signalling pathways such as Wnt and Notch. Increased expression of the Notch pathway in differentiating cells promotes maturation of enterocytes . Moreover, c-MAF is also noted as being a key regulator of enterocyte functions, coordinating transcriptional mechanisms for nutrient uptake in mature enterocytes. The transcription factor is expressed in specific locations of the villi regulating the absorption of different nutrients, highlighting the spatial/functional dynamics of enterocytes within the villi (González-Loyola et al., 2022). TLR4 activation, a key process of the innate immune system particularly in bacterial infections has been reported to significantly inhibit B-catenin signalling resulting in reduced proliferation of enterocytes (Sodhi et al., 2010). This suggests that enterocyte proliferation is not only determined by intrinsic factors but potentially external factors such as the microbiome and diet. Furthermore, the roles of enterocytes differ between the various areas of the gut. Enterocytes in the proximal region, or upstream region of the small intestine, are involved in absorbing carbohydrates, fats, proteins, and iron, where distal, or the downstream region of the small intestine, enterocytes are more involved in absorbing bile acids and vitamin B12 (Haber et al., 2017). As the enterocytes mature along the transition zone, they reconfigure their transcriptional profile from antimicrobial to nutrient absorbing then finally to an immune-regulating profile.

The transcriptional profile of enterocytes under the prediabetic condition is altered. Enterocytes within the prediabetic gut are reported to have increased expression of fatty acid binding proteins, ~~shifts in glucose transporter expression as well as varying reports in inflammatory responses suggesting the gut is In metabolic dysregulation~~. A key finding in this context is the upregulation if FABPs in enterocytes, particularly FABP2. Studies are showing that chronic exposure to lipotoxic conditions, such as high levels of palmitate results in lipotoxicity-induced enterocyte dysfunction at an early stage of enterocytes in progenitors (Filippello et al., 2022).

Goblet Cells

Intestinal goblet cells are resposible for the secretion of mucus playing a unique role in maintaining the intestinal lining from pathogens and mechanical damage. Goblet cells are characterised by their transcriptional profile involving transcription factors regulating their maturation, mucin producing genes, and immune response genes.

Goblet cells differentiate from ISCs through the action of specific transcription factors. A key transcription factor involved in the maturation of the goblet cell is SAM pointed domain ETS factor (SPDEF). In inflammatory conditions, SPDEF is often downregulated leading to a reduction in goblet cells and an impaired mucus barrier compromising the epithelium to the harsh bacteria dense environment . Furthermore, Klf4

Goblet cells secrete intestinal mucus layers primarily composed of MUC2 mucin, forming the backbone of the intestinal mucus layer. This layer acts as the first line of defense against irritants as well as microbes while allowing for nutrient absorption. The layers also contain other products of the goblet cell (TFF3, Fcgbp, and RELMB), antimicrobial peptides (lysozymes, and B-defensin) from paneth cells, as well as IgA from enterocytes (Kim and Ho, 2010). The mucus layer produced by goblet cells is critical for maintaining gut barrier function and regulating the interaction between the host and the gut microbiota. It provides an energy source and attachment site for certain bacteria while also protecting the epithelium from direct contact with potentially harmful microorganisms.

High-fat diet-induced obesity has been associated with alterations in the gut microbiota composition and changes in the intestinal mucus layer [reference]. Some studies have reported that HFD feeding leads to a reduced thickness and increased penetrability of the intestinal mucus layer [reference]. However, the effects of HFD on goblet cell number and Muc2 expression have shown inconsistent results across different studies, possibly due to variations in animal models and experiment durations.

While some research has indicated a decrease in goblet cell number and Muc2 expression in HFD-fed animals, [ref] et. al. did not find significant differences in these parameters between control and HFD groups. The inconsistency in findings suggests that the effects of HFD on goblet cells and mucus production may be influenced by various factors and require further investigation (Paone et al., 2022).

Enteroendocrine Cells

EECs develop from common secretory progenitors expressing mouse atonal homolog 1 [ref], and commit to an enteroendocrine fate upon expressing neurogenin 3 (Ngn3) [ref]. These cells then express markers like neurogenic differentiation factor 1 or forkhead box A1 and A2 [ref], before terminally differentiating into hormone-producing cells. These include K cells (glucose-dependent insulinotropic polypeptide), I cells (cholecystokinin), enterochromaffin cells (serotonin), D cells (somatostatin), and L cells (glucagon-like peptide-1 (GLP-1) and peptide YY (PYY)) [ref]. These hormones have various beneficial metabolic effects, with GLP-1 regulating insulin secretion, appetite, food intake, and gastrointestinal motility [ref]. Enhancing EEC differentiation could potentially treat type 2 diabetes and obesity.

Sirtuin 1 (SIRT1), a mammalian homolog of yeast Sir2, is a NAD+-dependent protein deacetylase [ref] involved in regulating metabolism, development, tumorigenesis, aging, and longevity [ref]. Recent studies suggest SIRT1 plays a critical role in intestinal tissue homeostasis [ref]. Intestinal SIRT1 is necessary for increasing ISC self-renewal and expansion under calorie restriction, leading to a reduction in differentiated villus cells, including EECs [ref]. However, SIRT1's role in EEC lineage specification remains unclear.

This study demonstrates that SIRT1 in EECs is involved in enteroendocrine progenitor cell (EEPC) proliferation and determines EEC numbers by regulating Wnt/β-catenin signaling in high-fat diet (HFD)-fed mice. SIRT1 deficiency in EECs increased EEC and L cell numbers, resulting in elevated plasma GLP-1 levels and improved metabolic status. Conversely, fasting-induced SIRT1 activation decreased EEC and L cell numbers by inhibiting EEPC proliferation.

These findings reveal a novel mechanism controlling EEC numbers through EEPC cell-cycle regulation.

Paneth Cells

Chapter 3: Interrogation methods for revealing alterations in the metabolic disease state.

5. Multi-omics Approaches in Prediabetes & T2DM Research

Bioinformatics, computational biology, multi-omics techniologies, and the advancements of each over the last two decades are redefining molecular biology research. Together they form a powerful driving force behind modern research and emerging advancements across many fields. They simultaneously integrate several omics methods characterising the transcriptome, proteome, metabolome, epigenome, genome and others.

Transcriptomics and scRNA-seq

Transcriptomics involves the employment of RNA-sequencing technologies to capture a detailed view of the gene expression landscape of cells. [Explain more about transcriptomics] Single cell transcriptomics is the most developed of the single cell omics techniques and is often used in conjunction with other omics techniques to understand the link between the expression of genes and phenotypes. Two main single-cell-RNA-seq methods are used currently, microwell based and microfluidics droplet-based based transcriptomics. Both methods are analogous to each other and involve the reverse transcription of RNA into cDNA allowing for amplification through a polymerase chain reaction assay. The amplicon fragments are ideal for sequencing analysis. Droplet based methods are often chosen over plate based as it is not limited by plate size and can analyse at a very high throughput.

Genomics

Proteomics

Proteomics is a field which focusses on the identification and quatification of proteins within a cell, tissue or organism. It complements other omics technologies such as genomisc and transcriptomics by providing insights into protein identity, structure and function. Techniques employed in proteomics are applied in various contexts such as detecting novel diagnostic markers, candidates for new vaccines, understanding mechanisms of pathogenicity, stimulus response patterns as well as protein pathways involved in specific diseases. The complexities of proteomics is derived from the wide range and diversity in protein structure and interaction dynamics, requiring sophisticated techniques to study adequately. Typical proteomics studies pegin with chromatography based purification followed by conventional analysis techniques such as ELISA, western blotting and protein microarrays. Characterisation of proteins is carried out via gel based approaches and mass spectrometry. Protein sequence analysis is done to determine the exact amino-acid sequence of peptides or proteins. Proteomics also employs various quantification techniques to determine the proportion of various proteins in the proteome. X-ray crystallography and NMR spectroscopy are primary high throughput techniques for providing a 3D structure of a protein helping understand the biological importance and function of the protein. With the advent of high-throughput technologies, large volumes of data are being produced requiring sensible methods for organising the data. Bioinformatics databases have been established to organise the data accordingly, as well as tools for predicting 3D structure, protein motifs and domains as well as the analysis of protein-protein interactions. Proteins contain post-translational modifcations which do not get captured by transcriptomics offering another perspective on the syystem landscape of the cell. Proteomics is often implemented at the bulk cell level. To fully capture cellular heterogeneity, it is important to also investigate cells at the single cell resolution. Single-cell proteomics is in a more premature state than sc transcriptomics. The primary reason for this is because proteins cannot be easily amplified in the same way

the transcripts can be, furthermore, the range in copy number of proteins is much wider than transcripts meaning many low copy number proteins do not get captured.

Metabolomics

Integrating Multiple Omics Techniques

Applications

6. Mouse Models in Prediabetes Research

As discussed in sections 1 and 2, there are many risk factors which lead to prediabetes and subsequent T2DM with the primary risk factors being diet, activity, and weight. As a result, when one is divising a method for modelling these systems, one must choose sensible risk factors accordingly to induce these metabolic changes. Currently, the primary means of inducing a diabetic system is through diet interventions, chemical actions, and genetic modifications.

Diet-Induced Models

Diet-induced models are commonly used to study prediabetes and T2DM pathophysiology. These models primarily utilize high-carbohydrate, high-sucrose, high-fructose, or high-fat diets (HFD). HFDs, with fat concentrations ranging from 20% to 60%, are frequently employed and lead to increased triglyceride formation [ref]. Fructose consumption induces increased energy intake, body weight, adiposity, and various metabolic disturbances [ref]. Sucrose-enriched diets, ranging from 20% to 77%, are used to induce obesity and insulin resistance, although they are generally less effective than equivalent amounts of fructose [ref].

C57BL/6J mice fed a 60% HFD for 11-16 weeks typically develop obesity, hyperglycemia, insulin resistance, and hypertension [ref]. Extended HFD feeding (36 weeks) can lead to additional abnormalities in the renin-angiotensin system and cardiac remodeling [ref]. Combinations of diet and drug

interventions, such as HFD with nitric oxide synthase inhibitors, have also been employed to induce more comprehensive cardiometabolic syndrome phenotypes [ref].

It's important to note that diet-induced phenotypes can vary based on genetic background, gut microbiota, diet composition, duration of intervention, age, and sex of the mice [ref]. These factors should be considered when interpreting results from diet-induced models. Additionally, different diets may trigger varying metabolic signatures, simulating different aspects of cardiometabolic complications rather than the complete syndrome [ref].

Not all high-fat diets produce equivalent metabolic effects. A study comparing isocaloric diets with different fat sources and fructose content in C57/BL6 male mice revealed distinct metabolic outcomes. A diet high in soybean oil (rich in polyunsaturated fatty acids) proved to be more obesogenic and diabetogenic than diets high in coconut oil or fructose. It induced more significant weight gain, adiposity, diabetes, glucose intolerance, and insulin resistance compared to a diet primarily composed of coconut oil [ref]. The soybean oil diet also caused more severe hepatic steatosis and alterations in gene expression related to obesity, diabetes, and inflammation. In contrast, a high-fructose diet, while less impactful on overall weight gain and diabetes markers, led to its own set of metabolic disturbances. These findings underscore the importance of considering specific dietary components, rather than just total fat content, when designing and interpreting diet-induced metabolic studies [ref].

Chemical/Drug-Induced Models

Chemical and drug-induced models offer alternative approaches to studying prediabetes and T2DM. Streptozotocin administration is a widely used method to induce diabetes in mice. It acts by destroying pancreatic β-cells, accumulating in these cells via the GLUT2 glucose transporter, and primarily leads to a Type 1 diabetes-like state [ref]. Another approach involves glucocorticoid-induced metabolic syndrome. Glucocorticoids regulate glucose homeostasis by promoting gluconeogenesis in the liver and decreasing glucose uptake and utilization in skeletal muscle and white adipose tissue. Mice treated

with glucocorticoids exhibit glucose intolerance, reduced insulin sensitivity, weight gain, dyslipidemia, central and peripheral fat accumulation, and hypertension [ref]. These chemical-induced models provide valuable tools for studying specific aspects of diabetes and metabolic syndrome, although they may not fully recapitulate the complex etiology of human T2DM.

Genetic Modifications

Genetic engineering strategies for mouse models of prediabetes and T2DM focus on altering lipid metabolism, weight regulation, glucose homeostasis, and blood pressure. Two well-characterized dyslipidemic models are the low-density lipoprotein receptor (Ldlr) and apolipoprotein E (Apoe) deficient mice. Ldlr-/- mice develop moderate hypercholesterolemia on a normal diet and are responsive to atherogenic diets, developing obesity, insulin resistance, and impaired glucose tolerance [ref]. Apoe-/- mice exhibit severe hyperlipidemia and spontaneous atherosclerosis, but typically do not become obese or insulin resistant without specific dietary interventions [ref].

Obesity models include leptin-deficient (ob/ob) and leptin receptor-deficient (db/db) mice. These models display increased food intake, extreme obesity, and reduced energy expenditure, with strain-dependent effects on glucose metabolism [ref]. The agouti yellow obese (Ay/a) mouse and melanocortin 4 receptor (MC4-R) knockout mouse are additional models exhibiting adult-onset obesity, hyperinsulinemia, and glucose intolerance [ref].

Nonobese lipodystrophic models, such as A-ZIP F-1 and aP2-SREBP-1c mice, feature restricted white adipose tissue capacity. These models demonstrate that fat ablation can lead to liver steatosis, diabetes, and elevated blood pressure [ref].

Together, these three model classes form a solid foundation for interrogating the prediabetic / T2DM metabolic system. There still however remain limitations and questions. Do these models truly capture the prediabetic / T2DM system exactly how it pertains to human populations? As discussed, there are several risk factors which are attributed to these metabolic disturbances, some of which are not captured by these models. Smoking, ethnicity, human genetics, varying

human diets, alcohol consumption are not captured within these models and resultingly may not display the intricacies of the true human diabetic system. Some alterations of the human system may not be seen in these models, similarly some alterations may be seen which are not typically seen in the human system. Perhaps for this reason it is worth aggregating studies from multiple different diabetic models as well as human studies when investigating the landscape of disease systems.

Chapter 4: The potential for synthetic and natural pharmacological agents in reversing metabolic dissarray.

7. Current Pharmacological Interventions for Prediabetes and T2DM

Currently, the best options for preventing T2DM is through the screening of prediabetes in populations and subsequent lifestyle changes in reducing the usual risk factors of T2DM. There do however exist preventative/management options regarding pharmacological interventions. These usually involve drugs promoting weight loss and/or the reduction of glucose levels. Typical pharmacological interventions include metformin, a-glucosidase inhibitors, thiazolidinodiones, SCLT2 inhibitors, and GLP-1 agonists.

Metformin

GLP-1 agonists

a-glucosidase inhibitors

thiazolidinodiones

SCLT2 inhibitors

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10033948/

8. Emerging Therapies: Focus on Brown Seaweed Extract

Brown seaweed contains the highest quantity of antioxidants, polyphenols and other bioactive compounds out of the three seaweed classes (brown, red and green). Evidence is emerging attributing a plethora of pharmacological benefits

through supplementation of extracts of these marine organisms. More specifically, numerous studies, in vitro, in vivo and clinical studies have been carried out investigating the potential for brown algae to ameliorate symptoms of the prediabetic and diabetic system although with mixed results. Algae strains which have been shown to demonstrate the largest benefits are *A. nod* and *F. ves*.

One in vitro study involved the extraction of inhibitors from 3 species …

Another study reviewed the potential of species1 and species 2 in manageing MetS

Another study investigated the effect of brown seaweed extract on weight-loss, blood glucose homeostasis in association with the metabolic and inflammatory response.

One other study investigated the ability for borwn seaweed algae to control salmonella infections in pig populations, a common seasonal issue in the agricultural industry during the stressful weaning period of young pigs. This is of particular relevance to diabetics who are more susceptible to Salmonella infections.

Materials and Methods

Mouse models

The dataset used throughout this study consists of scRNA-seq (10X) data from the small intestinal crypts of control diet and HFHSD obese FVF-enriched mice (Foxa2-FVF/FVF) as well as villus samples from the small intestine of control diet and HFHSD C57BL/6N mice.

All mice subjects in the dataset are male. Mice living conditions consisted of 2 to 4 mice per group, 23C 45-65% humidity with a 12 hr light/dark cycle.

Mice diets are reported to have begun at 10-12 weeks old and were randomised into different test groups matched for body weight with similar variance and given ad libitum access to either the control diet or HFHS diet for 11-13 weeks.

Glucose tolerance and insulin secretion tests.

To assess glucose tolerance, an oral glucose tolerance test (oGTT) was performed on FVF mice after 12 weeks on either a control diet (CD) or a high-fat high-sugar diet (HFHSD). Following a 6-hour fast, each mouse received an oral dose of glucose (1.5 mg/g body weight of 20% (wt/v) d-(+)-glucose solution in PBS). Blood glucose levels were measured at 0, 15, 30, 60, and 120 minutes post-glucose administration using a handheld glucometer (Abbott). For insulin secretion analysis, blood samples were collected from the tail vein at 0, 15, and 30 minutes during the oGTT. Plasma was isolated by centrifuging the blood samples (3,500 rpm, 15 minutes, 4°C), and insulin levels were quantified using the Ultra-Sensitive Mouse Insulin ELISA Kit (Crystal Chem, 90080) following the manufacturer's protocol.

Insulin resistance and beta-cell function were evaluated using the homeostasis model assessment of insulin resistance (HOMA-IR) and HOMA-β, respectively. These indices were calculated 12 weeks after diet initiation, using fasting blood

glucose and plasma insulin levels obtained after a 6-hour fast. The conventional formulas applied were HOMA-IR = fasting blood glucose (mg/dL) × fasting insulin (μU/mL) / 405 and HOMA-β = fasting insulin (μU/mL) × 360 / (fasting glucose (mg/dL) - 63).

Crypt and Villus Isolation and Single-Cell Preparation

The isolation of small intestinal crypts was performed following established protocols. Briefly, small intestines (SIs) were excised and rinsed with cold PBS. The villi were carefully scraped away using a glass slide. The remaining tissue was then cut into 2-cm sections, repeatedly washed with cold PBS, and incubated in 2 mM EDTA/PBS for 35 minutes at 4°C with gentle agitation. Crypts were released by vigorous shaking and filtered through a 70-μm mesh to remove any villous debris. For the preparation of single cells, the isolated crypts were treated with TrypLE (Life Technologies, no. 12605) first on ice for 5 minutes, followed by 5 minutes at 37°C, and subsequently incubated with 10 μg/mL DNase in crypt complete medium (DMEM/F-12 with 10% FCS) for 5 minutes at 37°C. The resulting single-cell suspension was achieved by gentle, repeated pipetting. Cells were then washed twice with 2% FCS in PBS and pelleted by centrifugation at 300g for 5 minutes at 4°C. For flow cytometry analysis, cells were resuspended in 1-2 mL of FACS buffer (2% FCS, 2 mM EDTA in PBS; Sigma-Aldrich, no. Y0503) and passed through 40-μm cell strainer caps attached to FACS tubes.

For the isolation of villus cells, the villi were scraped and processed into a single-cell suspension using the same TrypLE treatment protocol as described for crypt cells.

Flow Cytometry

For gene expression analyses, including microarray, single-cell transcriptomics, and western blotting, small intestinal crypt cells were sorted using a FACS-Aria III (BD Bioscience) with FACSDiva software v.6.1.3 and a 100-µm nozzle. In all experiments, cells were gated based on their forward scatter area (FSC-A) and side scatter area (SSC-A). Singlets were identified using forward scatter width (FSC-W) and forward scatter height (FSC-H), and dead cells were excluded using 7-AAD (eBioscience, no. 00-6993-50). For quantitative PCR with reverse transcription (qRT-PCR), cells were directly sorted into Qiazol lysis reagent (QIAGEN, no. 79306). To enrich FVF-positive small intestinal crypt cells for scRNA-seq, 30,000 FVF+ (FVFlow and FVFhigh) cells were sorted along with 30,000 live crypt cells per sample. The sorted cells were collected in modified FACS buffer (2% FCS, 0.02 mM EDTA in PBS).

scRNA-seq

Crypt and villus samples were prepared as described above. Dead cells were excluded via flow cytometry after 7AAD labelling. Dead cell exclusion was controlled by trypan blue staining and sorted cells were counted. Single-cell libraries were generated using the Chromium Single cell 3′ library and gel bead kit v2 (10X Genomics, no. 120237) according to the manufacturer’s instructions. Libraries were sequenced on a HiSeq4000 (Illumina) with 150-bp paired-end sequencing of read 2.

Upstream scRNA-seq pipeline

Quality control was performed using the Scanpy package to ensure the integrity and accuracy of the single-cell RNA sequencing (scRNA-seq) data. The quality control steps focused on identifying and mitigating potential confounding factors, such as the presence of mitochondrial and ribosomal, which can indicate cellular stress or contamination. Identification of Mitochondrial and Ribosomal genes: Mitochondrial genes were identified using the prefix "mt-" for mouse genes. Ribosomal genes were identified using the prefixes "Rps" and

"Rpl". The `calculate_qc_metrics()` function from Scanpy was used to compute common QC metrics, including the percentage of counts attributed to mitochondrial and ribosomal. The metrics were calculated and stored in the AnnData object for subsequent analysis.

Cells were removed with greater than 10% mitochondrial gene content, less than 200 genes by counts as well as cells with less than 500 total counts.

To identify potential doublets, the scrublet tool was used utilising a nearest neighbour classifier of observed transcriptomes and simulated doublets. Predicted doublets were subsequently removed from the dataset.

Counts were normalised per cell by total counts over all genes so that every cell has the same total count after normalisation. Each cell is normalised to a total count equal to the median of total counts for cells before normalisation. The data are then logarithmised using log1p where each count is transformed to the natural log of 1 plus the original count value. This method allows for the data to be transformed as well as zero values.

PCA Nearest neighbour graph and UMAP visualisation.

Filtering lymphocytes and ambient genes.

Mapping published annotations to the data

Downstream scRNA-seq Pipeline

KEGG pathway analysis
GO Term Analysis

Results

Discussion

References

Aliluev A, Tritschler S, Sterr M, Oppenländer L, Hinterdobler J, Greisle T, Irmler M, Beckers J, Sun N, Walch A, Stemmer K, Kindt A, Krumsiek J, Tschöp MH, Luecken MD, Theis FJ, Lickert H, Böttcher A (2021) Diet-induced alteration of intestinal stem cell function underlies obesity and prediabetes in mice. Nat Metab 3:1202–1216.

American Diabetes Association (2021) 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2021. Diabetes Care 44:S15–S33.

Beuling E, Baffour-Awuah NYA, Stapleton KA, Aronson BE, Noah TK, Shroyer NF, Duncan SA, Fleet JC, Krasinski SD (2011) GATA factors regulate proliferation, differentiation, and gene expression in small intestine of mature mice. Gastroenterology 140:1219-1229.e1-2.

Brannick B, Dagogo-Jack S (2018) Prediabetes and Cardiovascular Disease: Pathophysiology and Interventions for Prevention and Risk Reduction. Endocrinol Metab Clin North Am 47:33–50.

Choi J, Houston M, Wang R, Ye K, Li W, Zhang X, Huffman DM, Augenlicht LH (2023) Intestinal stem cell aging at single-cell resolution: Transcriptional perturbations alter cell developmental trajectory reversed by gerotherapeutics. Aging Cell 22:e13802.

Echouffo-Tcheugui JB, Narayan KM, Weisman D, Golden SH, Jaar BG (2016) Association between prediabetes and risk of chronic kidney disease: a systematic review and meta-analysis. Diabet Med J Br Diabet Assoc 33:1615–1624.

Filippello A, Di Mauro S, Scamporrino A, Torrisi SA, Leggio GM, Di Pino A, Scicali R, Di Marco M, Malaguarnera R, Purrello F, Piro S (2022) Molecular Effects of Chronic Exposure to Palmitate in Intestinal Organoids: A New Model to Study Obesity and Diabetes. Int J Mol Sci 23:7751.

González-Loyola A, Bernier-Latmani J, Roci I, Wyss T, Langer J, Durot S, Munoz O, Prat-Luri B, Delorenzi M, Lutolf MP, Zamboni N, Verdeil G, Petrova TV (2022) c-MAF coordinates enterocyte zonation and nutrient uptake transcriptional programs. J Exp Med 219:e20212418.

Guasch-Ferré M, Hruby A, Toledo E, Clish CB, Martínez-González MA, Salas-Salvadó J, Hu FB (2016) Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. Diabetes Care 39:833–846.

Haber AL et al. (2017) A single-cell survey of the small intestinal epithelium. Nature 551:333–339.

Kim YS, Ho SB (2010) Intestinal Goblet Cells and Mucins in Health and Disease: Recent Insights and Progress. Curr Gastroenterol Rep 12:319–330.

Lee CC, Perkins BA, Kayaniyil S, Harris SB, Retnakaran R, Gerstein HC, Zinman B, Hanley AJ (2015) Peripheral Neuropathy and Nerve Dysfunction in Individuals at High Risk for Type 2 Diabetes: The PROMISE Cohort. Diabetes Care 38:793–800.

Lindeboom RG, van Voorthuijsen L, Oost KC, Rodríguez-Colman MJ, Luna-Velez MV, Furlan C, Baraille F, Jansen PW, Ribeiro A, Burgering BM, Snippert HJ, Vermeulen M (2018) Integrative multi-omics analysis of intestinal organoid differentiation. Mol Syst Biol 14:e8227.

McCauley HA, Riedman AM, Enriquez JR, Zhang X, Watanabe-Chailland M, Sanchez JG, Kechele DO, Paul EF, Riley K, Burger C, Lang RA, Wells JM (2023) Enteroendocrine Cells Protect the Stem Cell Niche by Regulating Crypt Metabolism in Response to Nutrients. Cell Mol Gastroenterol Hepatol 15:1293–1310.

Middendorp S, Schneeberger K, Wiegerinck CL, Mokry M, Akkerman RDL, van Wijngaarden S, Clevers H, Nieuwenhuis EES (2014) Adult stem cells in the small intestine are intrinsically programmed with their location-specific function. Stem Cells Dayt Ohio 32:1083–1091.

Neves JS, Newman C, Bostrom JA, Buysschaert M, Newman JD, Medina JL, Goldberg IJ, Bergman M (2022) Management of dyslipidemia and atherosclerotic cardiovascular risk in prediabetes. Diabetes Res Clin Pract 190:109980.

Paone P, Suriano F, Jian C, Korpela K, Delzenne NM, Van Hul M, Salonen A, Cani PD (2022) Prebiotic oligofructose protects against high-fat diet-induced obesity by changing the gut microbiota, intestinal mucus production, glycosylation and secretion. Gut Microbes 14:2152307.

Perreault L, Temprosa M, Mather KJ, Horton E, Kitabchi A, Larkin M, Montez MG, Thayer D, Orchard TJ, Hamman RF, Goldberg RB, Diabetes

Prevention Program Research Group (2014) Regression from prediabetes to normal glucose regulation is associated with reduction in cardiovascular risk: results from the Diabetes Prevention Program outcomes study. Diabetes Care 37:2622–2631.

Sodhi CP, Shi X-H, Richardson WM, Grant ZS, Shapiro RA, Prindle T, Branca M, Russo A, Gribar SC, Ma C, Hackam DJ (2010) Toll-like receptor-4 inhibits enterocyte proliferation via impaired beta-catenin signaling in necrotizing enterocolitis. Gastroenterology 138:185–196.

Sune MP, Sune M, Sune P, Dhok A (n.d.) Prevalence of Retinopathy in Prediabetic Populations: A Systematic Review and Meta-Analysis. Cureus 15:e49602.

VanDussen KL, Carulli AJ, Keeley TM, Patel SR, Puthoff BJ, Magness ST, Tran IT, Maillard I, Siebel C, Kolterud Å, Grosse AS, Gumucio DL, Ernst SA, Tsai Y-H, Dempsey PJ, Samuelson LC (2012) Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells. Dev Camb Engl 139:488–497.

Xie Y, Ding F, Di W, Lv Y, Xia F, Sheng Y, Yu J, Ding G (2020) Impact of a high-fat diet on intestinal stem cells and epithelial barrier function in middle-aged female mice. Mol Med Rep 21:1133–1144.

Appendix