Epigenetics in Type 2 Diabetes

Jitte Jennekens

Summary

Every year more and more people are being diagnosed with type 2 diabetes (T2D). These numbers are getting so high, it might become unsustainable for the health care system. Currently there is no cure and patients need life-long treatment, which is extremely expensive. The main clinical feature associate with T2D is impaired glucose homeostasis. This means that glucose, which is present in the blood and rises after food intake, cannot be taken up efficiently by tissues that need this glucose for energy, mainly the muscles. There are two main causes that contribute to impaired glucose homeostasis. First, not enough insulin is secreted by the beta cells of the pancreas. Insulin is the hormone that stimulates the tissues to take up glucose. Second, the tissues need more and more insulin to stimulate the uptake of the same amount of glucose as healthy people. This is called insulin resistance. Before diagnosis, these two mechanisms are already developing. In the early phase of T2D development, the insulin producing beta cells are able to secrete more insulin. This higher amount of insulin in the blood is enough to stimulate normal glucose uptake by the tissues, even though these tissues have become less sensitive to insulin (i.e. insulin resistant). However, when the insulin resistance becomes worse, the beta cells can no longer keep up their insulin secretion with the insulin resistance and they start to lose their function. At this point, the insulin levels drop and glucose levels of the blood start to rise, since the tissues cannot take up normal amounts of glucose. It is thought that the number of T2D patients is growing so fast because of the changes in life style of the past decades, such as unhealthy eating and not enough physical exercise. A large part of the T2D patients is also obese, which has been defined as a major risk factor for developing the disease. Because the environment is thought to play such an important role in the development of T2D, a lot of attention has been attracted to something called epigenetics. In the cells of the human body, the information on what to do is provided by the DNA. Identical twins have the same DNA in their cells, but they can still be physically different. This is because of epigenetics, which are factors that interact with the DNA. Epigenetic factors can activate or inactivate parts of the DNA. If you think of the DNA as the ‘cookbook of life’ and epigenetics as the cook, different cooks will make the recipes differently because of different ingredients or kitchens, which can be compared with different environment influences. There are different methods available to study these epigenetic changes in T2D. Currently, those methods are used to study tissue samples containing a number of different cell types. For example, a lot of studies use pancreatic islets, which contain the insulin producing beta cells, to study

epigenetics in T2D. However, these pancreatic islets also contain other cells next to the beta cells. Therefore, these techniques cannot provide information on the epigenetic changes in a specific cell type. To completely understand how epigenetic changes can contribute to T2D it is needed to be able to study individual cells. Methods are now being developed that are able to study single cells. However, there are still some technical difficulties that have to be overcome.

Abstract

This literature review focusses on the role of epigenetics in T2D and the methodology used to study different aspects of epigenetic modifications. The main goal is to emphasize the importance of single-cell techniques to elucidate the underlying epigenetic mechanisms of such a complex and multifactorial disease. T2D has become a major health issue over the past decades, and ultimately this could benefit the discovery of new treatment or prevention strategies to relieve the pressure on the health care system. Long before T2D onset, pathologic alterations are already present. The two primary mechanisms leading to T2D are insulin resistance and impaired insulin secretion by the beta cells of the pancreas. Both defects are required in order to progress to T2D. In the early phase of T2D development, beta cells can still compensate for the insulin resistance by increasing their insulin secretion. However, this has a limit and after the maximal capacity is reached, beta cells lose their function and insulin secretion declines. Epigenetics have been implicated in T2D development, since they are known to link environmental factors to the genome. Epigenetics include DNA methylation, histone modifications, and chromatin structure remodeling. Bulk studies have already identified a number of epigenetic alterations in disease-relevant tissues for T2D. With the technological advances also single-cell methods are being developed and used. The main reason that such techniques are important is the heterogeneity of tissue samples currently used for study, which masks important single-cell effects. However, adaptation to single-cell methods still faces a number of challenges, yet with the rapid technological developments the future for single-cell epigenomic studies looks promising.

Introduction

The global prevalence of T2D has been increasing over the last decades and rapidly continues to increase. An increase of more than 50% is expected between 2017 and 2045, equivalent to 693 million people in total with diabetes by 2045. This rapid change in diabetes prevalence, primarily T2D, is associated with a tremendous increase in global healthcare expenditure. In 2017 a total of USD 850 billion was spent on diabetes. However, a major contributor to this

economic burden is the treatment of the related complications. By 2045 this number is expected to increase by 7% resulting in an annual global healthcare expenditure of USD 958 billion on diabetes alone. Hence, this T2D epidemic will become unsustainable for the health care system if this trend continues (1). T2D is a heterogenous metabolic disease characterized by hyperglycemia as a result of insulin resistance and impaired insulin secretion by pancreatic beta cells. It is known that genetic predisposition is involved in T2D development, yet this does not explain the alarming rate at which T2D prevalence is increasing. Therefore, rather environmental factors might be driving the rise in T2D patient numbers. The environment exerts its influence via epigenetic modifications. Gene expression and development of an organism are regulated via such epigenetic modifications without changing the DNA sequence itself. Common epigenetic mechanisms include DNA methylation, histone modifications, and RNA-mediated processes. Over the last decades a lot of attention has been attracted to the study of epigenetics in metabolic diseases (2). With regard to T2D, a strong link with environmental influences was already suggested, since a large population of T2D patients is also obese. Significant evidence now supports this idea of epigenetic alterations as important contributors to disease development. Obesity is considered an important risk factor for T2D and the increasing prevalence of T2D goes hand-in-hand with the increasing number of obese people. This obesity epidemic that has arisen in the last decades is predominantly related to drastic lifestyle changes that have led to sedentary behavior, unhealthy food consumption, and reduced physical activity. Together with these changes in lifestyle, the ageing population is an important driver behind the growing number of T2D patients. As a consequence, the disease now also exhibits a more heterogenous pathophysiology including both lean and obese, and young and old patients (3). Thus, it is of great importance to understand how such environmental factors contribute to the development of T2D, in order to adequately treat or prevent this disease. A lot of effort has been made over the last 10 years to identify epigenetic dysregulation in T2D by studying various disease-relevant tissues. However, current studies on epigenetic mechanisms in obesity and T2D have mainly used bulk analysis of tissue samples containing multiple cell types, which masks cell type specific effects. Pancreatic islets, which are widely being used to study T2D pathology, are a good example to illustrate how analysis of such a heterogenous tissue sample leads to averaged results (figure 1). Pancreatic islets consist of a number of functionally different cell types, such as alpha, beta, and delta cells. Epigenetic alterations associated with T2D may manifest differently between these different cell types. Hence, it is of great importance to obtain information on epigenetic modifications at a single cell level.

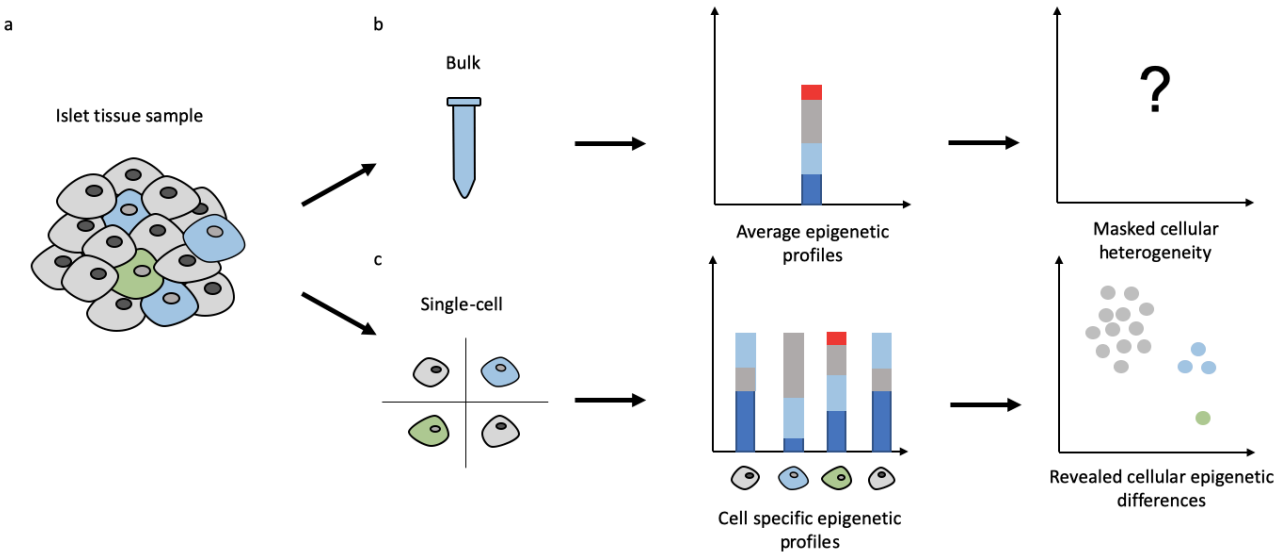

Figure 1. **Pancreatic islet heterogeneity in bulk vs. single-cell analysis. a,** A pancreatic islet tissue sample contains multiple cell types. **b**, Bulk analysis of an islet sample leads to averaged results and cells cannot be clustered based on their epigenetic differences. **c,** Single-cell analysis of an islet sample enables the identification of cell type specific epigenetic modifications.

With the rapid technological advances of the last decade such single-cell methods have become available and are already widely used in the field of transcriptomics. However, epigenome single-cell methods are still in its infancy. This review highlights the current status of the field of epigenomics in T2D and discusses the importance of single-cell methods to further advance this field. What are the main limitations of the current bulk techniques and what are the challenges that still have to be overcome regarding the single-cell methods? The first part of the review is an overview of T2D pathology necessary to illustrate the complexity of this disease and the challenges associated with studying such a multifactorial disease. The second and third part focus on epigenetics in T2D and the bulk versus the single-cell methodology used to study these epigenetic effects in T2D.

Type 2 diabetes

The pathophysiology of T2D is characterized by progressive insulin resistance and impaired insulin secretion as a result of reduced pancreatic beta-cell function over time. This metabolic disease starts to develop long before the onset of T2D and is termed prediabetes before a patient can be diagnosed as diabetic. Prediabetes is associated with both impaired fasting glucose (IFG) and impaired glucose tolerance (IGT) resulting from impaired glucose homeostasis. Although leading to augmented blood glucose levels, the values do not yet reach the diabetic threshold (4). Prevalence of prediabetes and progression to T2D is well-studied in

prevention trials. However, the multifactorial nature of the disease complicates the understanding of the rate and underlying mechanisms of such progression. Several known factors associated with progression to diabetes are elevated free fatty acids (FFA), high body mass index (BMI), and age (5). Once T2D is manifest, patients are often at risk for developing a number of life-threatening complications, with reduced quality of life and a high mortality risk as a consequence. These complications are the result of day-long exposure to high glucose levels, which predominantly affects the eyes, kidneys, nerves, and cardiovascular system. There is no cure for T2D and patients require insulin to control blood glucose levels in order to prevent such complications. With mild T2D, dietary changes or weight reduction can have sufficient beneficial effects and no further treatment is necessary (3).

A set of criteria for the screening and diagnosis of prediabetes and diabetes is based on the recommendations of the American Diabetic Association (ADA) of 1997 and recommendations of the World Health Organization (WHO) and the International Diabetes Federation (IDF) of 2006.

The criteria for the diagnosis of diabetes include:

a) A fasting plasma glucose (FPG) value ≥7.0 mmol/L (126 mg/dL)
b) A 2-h postload glucose concentration ≥11.1 mmol/L (200 mg/dL) during an oral glucose tolerance test (OGTT)
c) Symptoms of diabetes and a casual (i.e., at any moment irrespective of time since last meal) plasma glucose concentration ≥11.1 mmol/L (200 mg/dL)

In 2009, the International Expert Committee recommended measurement of Hemoglobin (Hb) A1c for diabetes diagnosis, since it is a reliable reflection of the average blood glucose over the previous two to three months. The ADA and WHO have endorsed the use of HbA1c for diagnosis of diabetes (6-9).

Insulin resistance and impaired insulin secretion are the two fundamental defects underlying T2D. Insulin resistance is defined as a decreased response of cells to the insulin hormone, also commonly described as insulin insensitivity. Decades before T2D onset, both insulin resistance and alterations in beta-cell function are present. A much-debated topic in research on T2D development is the time point of manifestation and relative contribution of insulin sensitivity and impaired beta-cell function. Fully understanding the sequence and timing of pathologic events will help to define more effective prevention strategies. Multiple longitudinal studies have suggested insulin resistance to be the primary defect, rather than beta-cell dysfunction (10). However, results of other studies have indicated impaired beta-cell function, likely associated with genetic predisposition, as the primary underlying defect (11). These conflicting results illustrate the heterogeneity of this disease evident by the existing variation

between patients regarding insulin resistance and beta-cell function. The variation in disease phenotype mainly results from the large number of factors involved in determining disease outcome, such as age, sex, and BMI. To understand the pathological mechanisms responsible for T2D one has to understand how normal glucose homeostasis is maintained in the human body.

Normal glucose homeostasis

Normal glucose homeostasis depends on the combined effects of the beta-cell insulin response and the sensitivity of target tissues to both glucose and insulin to promote glucose uptake. Under basal conditions, whole-body glucose homeostasis is maintained by precisely regulated endogenous glucose production (EGP). In the basal state, 50% of the total glucose is taken up by the brain in an insulin-independent manner. Both the liver, which is also insulin-independent, and the muscles, which are insulin-dependent, are responsible for the uptake of another 25% of total glucose each. The EGP primarily occurs in the liver (85%) and the remaining amount is derived from the kidneys (15%). There are two different pathways involved in the EGP, glycogenolysis and gluconeogenesis. Glycogenolysis is the production of glucose from glycogen. Gluconeogenesis is the production of glucose from non-carbohydrate precursors (e.g. amino acids). The balance between the EGP and tissue glucose utilization is disrupted following food ingestion. An increase in blood glucose triggers the beta cells to secrete insulin, resulting in temporary blood hyperinsulinemia and hyperglycemia. Insulin is the key hormone for glucose homeostasis. Both high blood insulin and high blood glucose levels are independently responsible for three precisely regulated mechanism:

1. Suppression of the EGP
2. Stimulation of glucose uptake by the liver
3. Stimulation of glucose uptake by the peripheral tissues (primarily muscle)

Hepatic glucose production (HGP) is suppressed in response to insulin to prevent superimposed glucose inputs from both food digestion and liver glucose production. Important for the suppression of the HGP is the inhibition of glucagon by insulin. For the maintenance of adequate basal glucose levels, the HGP depends for approximately 50% on glucagon stimulation.

The muscles are the predominant site of insulin-mediated glucose uptake in the post-prandial state, responsible for 80% of the whole-body glucose disposal. Adipocyte tissue metabolizes approximately only 4%-5% of the total blood glucose. After glucose transportation into the cell, two metabolic pathways are predominantly activated: non-oxidative glucose metabolism (i.e. glycogen synthesis), which is responsible for 75-80% of glucose metabolism, and glucose

oxidation (i.e. glycolysis) into carbon dioxide and water, which is responsible for the remaining 20-25% (12).
Although adipocyte tissue plays a minor role in the uptake of glucose, it has a powerful effect on glucose homeostasis through the regulation of FFA. FFAs are released in a process called lipolysis, in which triglycerides are hydrolyzed into a glycerol and three FFAs. When glucose is limited, most organs use the breakdown of triglycerides as their primary energy source. FFA release is strongly inhibited by insulin, which favors muscle glucose uptake and contributes to the inhibition of HGP (13).

In summary, glucose homeostasis depends on the precisely adjusted glucose-stimulated insulin response to the insulin sensitivity of target tissues. T2D occurs when this balance can no longer be maintained because of the inability of beta cells to secrete sufficient insulin matching the increasing insulin requirements to store glucose.

Insulin resistance

Longitudinal studies have shown that already decades before the onset, people with T2D develop both gradually increasing post-absorptive and postprandial hyperglycemia (14). Under post-absorptive conditions, fasting hyperglycemia results from an excessive rate of EGP, primarily derived from the liver. The excessive rate of HGP, despite elevated insulin and glucose levels, reflects the resistance of the liver to both these hormones, which have been demonstrated to be potent inhibitors of HGP in healthy individuals (15,16). Together with the developed resistance to hyperinsulinemia and hyperglycemia, it is believed that other indirect mechanisms are involved in the accelerated HGP, mainly the stimulation of excessive gluconeogenesis, as shown by studies employing ^{13}C-magnetic resonance imaging (17,18). An increase in gluconeogenic factors (e.g. lactate), hyperglucagonemia, an increase in FFA oxidation, and enhanced glucagon sensitivity all contribute to a disturbed balance in hepatic glucose output (19,20).

In the postprandial state, skeletal muscle represents the major site of insulin resistance reflected by impaired insulin-stimulated storage of ingested glucose. This defect primarily results from reduced non-oxidative glucose metabolism into muscle glycogen. In T2D patients the insulin action is markedly delayed and more insulin is required to dispose of the same amount of glucose compared to healthy individuals (21). The contribution of both inadequate suppression of HGP and reduced insulin-mediated glucose uptake by the muscle to the impaired glucose homeostasis has been shown to be equal following glucose ingestion (22).

Although the kidneys contribute to the total EGP, studies have shown that renal glucose production is not involved in the excessive EGP in T2D patients, which could be completely

attributed to an increase in HGP (23). Besides, the kidneys contribute to hyperglycemia in T2D via a different and more relevant process. In T2D the elevated glucose load faced by the proximal tubular lumen of the kidneys leads to the upregulation of SGLT2, the main cotransporter involved in renal glucose reabsorption. This results in an increase in tubular glucose transport capacity of 20%. Consequently, a significant amount of glucose is being transported back into the blood. This renal contribution to T2D pathogenesis should not be overlooked and presents an important treatment target (24).

Insulin secretion

Along with the decline in insulin sensitivity early in the course of T2D development, the insulin secretion by beta cells is impaired. Both defects have to be present for the progression from normal glucose tolerance to prediabetes and ultimately T2D to occur. The natural history of T2D is characterized by worsening IFG and IGT in the presence of hyperinsulinemia. DeFronzo has used an inverted-U-shaped curve to describe the insulin response in relation to the FPG concentration in progression towards T2D, with two distinguishable phases. During the first phase, progressively worsening insulin resistance occurs, as cells become gradually less sensitive to insulin. Initially, the beta cells are still capable to augment their insulin secretion in proportion to the severity of insulin insensitivity of target tissues, thereby maintaining normoglycemia. As the FPG concentration starts to increase from 80 mg/dl in normal individuals to 120 mg/dl characteristic for early diabetic patients, twice as much insulin is secreted by the beta cells, resulting in hyperinsulinemia. It should be emphasized, that even though the beta cells are still able to compensate for the insulin resistance in the first phase, their function at this point cannot be classified as healthy and is declining over time already before the onset of hyperglycemia. The beta cells are unable to compensate for any further increment in the FPG concentration (>120 mg/dl) as this value reflects the limit of their secretory capacity. When beta cells can no longer compensate for the insulin resistance, the beginning of a second phase, characterized by a decline in insulin secretion, is marked. At this point the body fails to maintain normal blood glucose levels and the body reaches a state of hyperglycemia, which is moderate to begin with, but worsens over time. Because an FPG concentration exceeding 120 mg/dl is associated with a progressive decline in the insulin levels in response to glucose, a person with an FPG concentration of 150-160 mg/dl will resemble a normal nondiabetic individual with regard to insulin secretion values. However, it is important to note that normal insulin secretion levels in absolute values in response to fasting hyperglycemia is far from normal. Ultimately, beta-cell dysfunction progresses and is well advanced by the time the FPG concentration exceeds 200-220 mg/dl. At this point a person falls within the diabetic range and their beta-cell function is already reduced with 50% (12,25).

In summary, T2D onset is preceded by hyperinsulinemia to counterbalance the insulin resistance, which is the major driver behind the progression towards prediabetes. However, beta-cell dysfunction is the main defect leading to overt T2D, which is associated with no further progression in insulin resistance. Thus, an individual will only develop overt diabetes when the beta cells can no longer adjust their insulin secretion to overcome the insulin resistance (26). In the next paragraph, the nature of beta-cell defects will be discussed.

Beta cell failure

Several studies have been investigating the cause of beta-cell dysfunction. However, the exact underlying mechanisms remain to be established. A number of hypotheses have been proposed to explain the deterioration of beta cells, manifest in both loss of function and a decline in beta-cell mass. During the compensatory phase, beta cells increase their capacity to counterbalance the decline in tissue sensitivity. Failure of this compensation process, however, sets in when beta cells can no longer maintain the high rate of insulin secretion due to genetic or acquired factors making them susceptible for beta-cell failure. Besides the inadequate insulin response to glucose, additional defects in the beta cells which affects their function and viability, have been shown. Factors that have been implicated in beta-cell failure, including hyperglycemia, dyslipidemia, mitochondrial dysfunction, and islet amyloid polypeptide (IAPP), can trigger apoptosis in beta cells via activation of the oxidative or ER stress response. Apoptosis has been demonstrated as the primary cause for a decline in beta-cell mass. Post-mortem studies found that subjects with T2D had a 40-60% reduction in beta-cell mass compared to weight-matched controls (27).

One of the known pathogenic mechanisms leading to beta-cell death is chronic exposure to high concentrations of glucose, which was shown to alter gene expression and increase ROS production. Similarly, overexposure of beta cells to elevated levels of FFAs also impairs beta-cell function and promotes apoptosis. While these pathogenic factors alone were shown to have adverse effects, the combination of those defects has an even larger impact on the deterioration of beta cells. This synergistic process is defined as glucolipotoxicity. In addition, high levels of IAPP result in excessive depositing of amyloid in the islets, which can also activate the oxidative and ER stress response. The ER of beta cells is relatively susceptible to ER stress because of its highly developed nature. The combinatorial action of these adverse processes in combination with islet inflammation creates a diabetic environment for beta cells, ultimately leading to loss of beta cell mass (28).

While beta-cell apoptosis has been suggested by many studies as the primary cause for loss in beta-cell mass in T2D, more recent studies are also reporting that there might be other

mechanisms causing a decline in beta-cell function as a result of exposure to stress. Growing evidence supporting this more complex concept suggests that loss of beta-cell identity and beta cell de-differentiation in addition to apoptosis could contribute to beta-cell failure. Loss of cell identity has been demonstrated by a reduction in proteins related to mature beta-cell function, such as insulin and several transcription factors (e.g. MAFA, PDX1, and NKX6.1). Moreover, de-differentiation of beta cells to progenitor-like cells and trans-differentiation into other islet cell types, predominantly alpha cells, have been implicated in beta cell failure. Understanding the underlying mechanisms could provide new potential treatment options by preventing dedifferentiation or trans-differentiation or even re-differentiation into beta cells (29,30).

Adipocytes

Obesity, defined as having a BMI of greater than 30 kg/m^2, is frequently linked to T2D, which is not surprising since the majority of T2D patients is also obese. Being obese increases the risk for developing T2D, mainly via the pathogenic effects of the increased fat mass on glucose homeostasis. Work of the past several years has made a convincing case for the active role of adipose tissue in obesity and T2D and the adverse impact of alterations in adiposity on T2D is now commonly recognized. People with obesity or T2D have day-long elevated FFA levels. In a normal individual, insulin secretion inhibits lipolysis in the adipocytes. However, in patients with T2D, this inhibitory action is impaired because of developed resistance to insulin in the adipocytes. The resulting chronically elevated plasma FFA concentration contributes to further worsening of the main pathogenic mechanisms underlying T2D by affecting both muscle and liver insulin resistance and insulin secretion (31).

Disturbances in lipid metabolites as a result of accelerated FFA oxidation inhibits muscle glucose transport, glucose oxidation, and glycogen synthesis. This inverse correlation between increased FFA oxidation and decreased glucose oxidation contributes to insulin resistance (32). Next to insulin and glucagon, plasma FFA is a potent regulator of HGP. In obesity and T2D the chronically elevated FFA levels present a continuous stimulus for gluconeogenesis in the liver, even in the presence of insulin. Acetyl CoA, a product of FFA oxidation, stimulates pyruvate carboxylase. Normally, depletion of this enzyme limits gluconeogenesis. However, the increased stimulation of pyruvate carboxylase enables an increase in gluconeogenesis and ultimately HGP. (13,33,34).

Another profound implication of adipocytes in T2D pathogenesis is their role in the secretion of biologically active adipokines. In the diabetic state adipokines can negatively influence triglyceride and fatty acid metabolism in the adipocytes by triggering an inflammatory

response. The secretion of cytokines, which act to attract immune cells, affects the adipocyte function. With the adipocytes being in an inflammatory state, lipolysis is increased leading to elevated FFA and triglyceride concentrations. Thus, a high percentage of fat tissue can aggravate the primary pathogenic mechanisms of T2D, insulin resistance and impaired insulin secretion, because of the lipotoxic effects of accelerated FFA oxidation (31).

Epigenetics

The field of genetics, which has been a popular field of study for decades, has known remarkable advancements, mainly associated with the technological improvements that have enabled for example complete human genome sequencing. However, the genetic sequence alone cannot explain the diversity in phenotypes. It is now well-established that additional layers to gene regulation exist, known as epigenetics. Although being a relatively young field of research, epigenetics has attracted a lot of attention. The increasing interest in the field is fueled by the growing knowledge on the importance of epigenetics and the technological advancements that have enabled its study. Epigenetics has been defined as the study of heritable modifications in gene expression that does not involve changes in the genetic code itself. Multiple epigenetic mechanisms are involved in gene regulation, including DNA methylation, post-translational histone modifications, chromatin structure remodeling, and non-coding RNAs. The underlying processes of gene regulation encompasses more epigenetic mechanisms than can be described in this review. Therefore, the main focus will be on DNA methylation, histone modifications, and chromatin structure, since these are the most extensively studied mechanisms.

<u>DNA methylation</u>

DNA methylation is the most studied epigenetic modification for the mere reason that it is the easiest mechanism to study because of its chemical stability. DNA methylation is the addition of a methyl group to a base in the genetic code. This primarily occurs on cytosine nucleotides that are followed by a guanine nucleotide in the 5' to 3' direction resulting in 5-methylcytosine (5-mC). This CpG site, which is located on a single DNA strand as a cytosine linked to a guanine by a single phosphate group, should not be confused with the C-G base pair (bp). Three highly specialized enzymes have been identified that are capable of the methylation of cytosine nucleotides. These enzymes belong to the DNA methyl transferase family and include DNMT1, DNMT3a, and DNMT3b. DNMT1 can only maintain existing methylation profiles during replication. DNMT3a and DNMT3b can both maintain existing methylation and perform de novo methylation. Demethylation of cytosine residues occurs via either an active or passive process. How exactly active demethylation is carried out, is still a topic of debate. Several possible mechanisms have been proposed, including the action of enzymes such as TET (ten-

eleven-translocation), and TGD (thymine DNA glycosylase). Depending on the position of methylation in the genome its function is primarily repressive but has been shown to be activating in certain cases.

CpG dinucleotides are rare in the genome based on what would be expected by random chance. When methylated, a cytosine residue is susceptible to spontaneous deamination to thymine, which causes a mutation as a result of mismatched bases. The resulting CpG-poor genome is alternated by genomic regions with a relatively high occurrence of these CpG sites, which are called CpG islands (CGI). CGIs are regions of 200 bp and contain CpGs with a density of 50%. It is an important regulatory element for gene expression, evident by the high percentage (60-70%) of CGIs in human promotors. Methylation of CGIs in the promotor region near the transcription start site (TSS) has been shown to block transcription and is associated with long-term silencing of genes. However, CGIs do not only regulate transcription from the promotor near the transcription start site (TSS) but also from a more remote position, such as distant enhancer sites (48).

CGIs are often unmethylated which favors transcription initiation. However, during embryonic development DNA methylation of CGIs silences genes in processes such as genomic imprinting, X-chromosome inactivation, and establishing cell identity. Hypermethylation of CGIs prevents the association of DNA binding proteins that are required for the activation of the transcription or the recruitment of other transcriptional activators, thereby silencing genes. In addition, DNA methylation contributes to chromosome and genome stability and chromatin compaction via the recruitment of histone modifying and chromatin remodeling complexes that can alter the chromatin structure, which ultimately affects gene activity (35).

DNA hydroxy methylation

Recently, studies have found another type of DNA modification where oxidation of the 5-mC at CpG sites can form 5-hydroxymethylcytosine (5-hmC). Since this reaction is catalyzed by the family of TET proteins, 5-hmC has been suggested as an intermediate step in the process of demethylation. However, DNMT enzymes have also been implicated in the generation of 5-hmC, which suggests a possible regulatory function in gene expression (36).

Chromatin structure remodeling

A 146 bp segment of DNA is wrapped around histone proteins into nucleosomes, which are further condensed to form chromatin structures. Each nucleosome contains four core histone proteins (H3, H4, H2A, H2B). Chromatin compaction depends on a combination of factors including DNA methylation, histone modifications, histone variants, the distance between

nucleosomes, and chromatin remodeling complexes. When chromatin is in a condensed state, the DNA is inaccessible to regulatory processes such as transcription, and DNA replication.

Histone modifications are covalent post-translational modifications to the amino-terminal tail of the histone proteins, which can be acetylation, methylation, phosphorylation, and ubiquitination as well as other less studied modifications. These modifications can affect the condensation of the chromatin depending on the type of modification and the locus and gene. Organization into lightly packed chromatin is associated with transcriptionally active regions, also referred to as euchromatin. Tightly packed chromatin or heterochromatin, on the other hand, is associated with transcriptionally inactive regions.

While only two de novo DNA methyltransferases exist, cells contain many enzymes that can modify histones. Specific histone modifications have been identified either associated with active of silent genes. Acetylation is a histone modification typically found associated with euchromatin. In addition, H3K4me3, and H3K79me3 are well-known markers for active regions. Heterochromatin, on the other hand, is associated with low levels of acetylation and characteristic histone marks for inactive regions such as H3K9me3 and H3K27me3. Similar to DNA methylation, post-translational modifications can recruit histone modifying and chromatin remodeling enzymes to regulate gene expression. Together with DNA methylation, the histone marks provide a complex landscape with multiple regulatory inputs affecting gene transcription.

The unmethylated CGIs in human cells, typically associated with an active gene, are characterized by nucleosome-depleted regions (NDRs) at the transcription start site (TSS) of that gene. These NDRs at the TSS are frequently flanked by nucleosomes marked by H3K4me3, a modification characteristic for active genes. Figure 2 gives an example of how DNA methylation and histone modifications regulate gene expression. Alterations in the known histone marks and their effect on transcription can be studied to identify regulatory patterns in histone marks but also differently modified histones in disease (37,38).

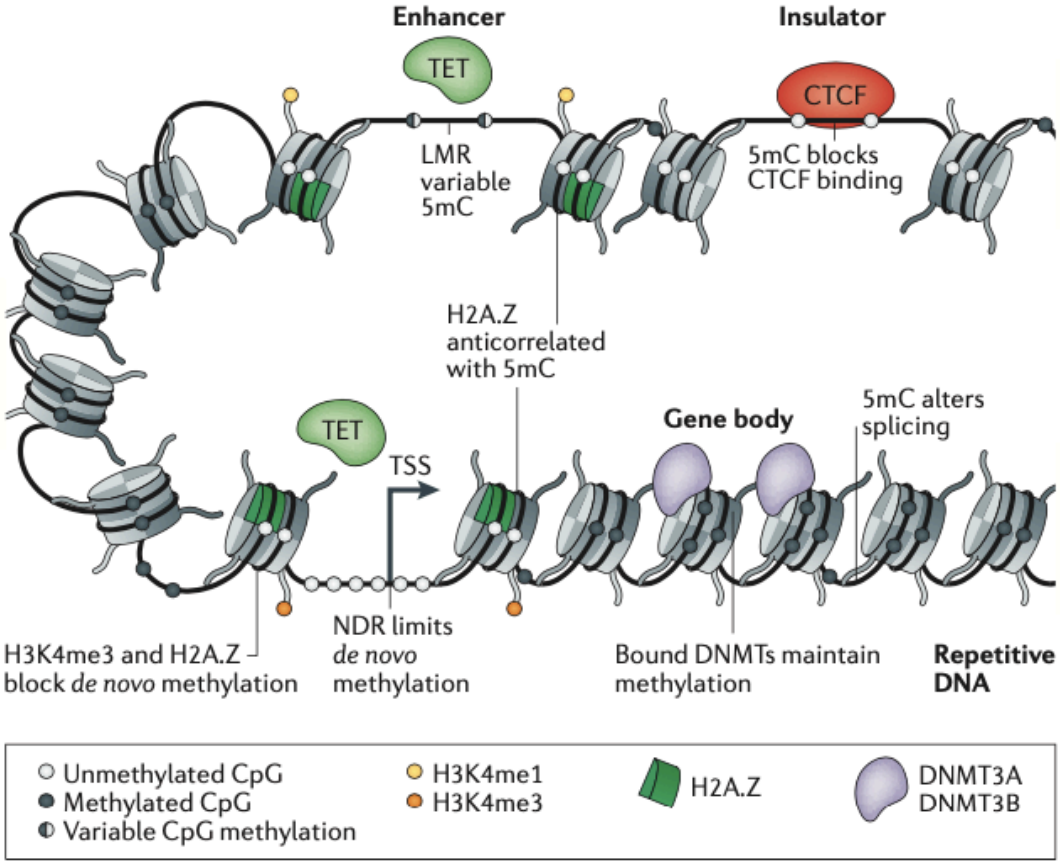

Figure 2. Epigenetic regulation of transcription by DNA methylation and histone modifications. The CpG sites at the TSS are often characterized by NDR. In addition, these CGIs are unmethylated which favors transcription. Specific histone modifications associated with transcriptional activation, such as H3K4me1 and H3K4me3 are located near this promotor region. Another signature mark for active genes is the histone variant H2A.Z. Note: reprinted from Functions of DNA methylation: islands, start sites, gene bodies and beyond, by Jones,P, 2012, Nature Review Genetics, 13, 484-492 *(39)*.

Epigenetics in T2D

The importance of epigenetics in normal functioning but also disease development is well-established. With regard to T2D, epigenetic abnormalities might explain the rapid increase in prevalence. This concept of epigenetic alterations contributing to the pathogenesis of T2D emanated from studies such as in offspring of mothers during the Dutch hunger winter, who were at high risk for developing T2D. Besides, studies on diabetes and pregnancy showed that intrauterine exposure to hyperglycemia increases the risk for T2D (40).

Identifying different epigenetic marks in T2D patients compared to healthy individuals might provide the key to understanding the contribution of epigenetics to the development of the disease and the disease progression. It can provide information on the effect of environment related risk factors (e.g. obesity) on the genome and explain the differences in T2D susceptibility and pathogenesis. This in turn could result in new treatment therapies or even

prevention methods. This section will highlight the current knowledge on epigenetics in T2D for the disease-relevant tissues (pancreatic islets, skeletal muscle, liver, and adipocytes), based on studies on DNA methylation, histone modifications and chromatin structure.

DNA methylation in T2D pancreatic islets

The first studies investigated the methylome of candidate genes, as identified by genome-wide association studies (GWAS), in human pancreatic islets samples from T2D and non-diabetic donors. These studies identified hypermethylated promotor regions in genes shown to be crucial in beta-cell function in the islets of donors with T2D. The genes that were found to have differently methylated regions (DMRs) included INS (encoding insulin), PDX1, PPARGC1A, and GLP1R. Functional studies have demonstrated the effect on gene expression, evident by the reduced mRNA levels of these genes, showing that the epigenetic alterations indeed affected gene expression.

While the first studies on T2D-related epigenetic variations in human islets analyzed DNA methylation of candidate genes only, the technological advancements made it possible to identify methylation of CpG sites of genes that were not directly implicated in T2D (41,42). Development of Illumina's Infinium arrays, a state-of the art technique, has provided new insights in altered DNA methylation sites in T2D patients compared to healthy individuals for all the different target tissues. In diabetic islets, Dayeh et al. identified 853 genes with differential DNA methylation, of which 102 genes were differentially expressed. A decrease in methylation of the genes CDKN1A, PDE7B, and SEPT9 was associated with an increase in gene expression. Overexpression of these newly identified candidate genes in beta cells provided functional proof for the impaired insulin response to glucose. Furthermore, Dayeh et al. confirmed candidate genes for T2D and obesity, as identified by earlier GWAS, to be differentially methylated (43,44).

The arrays used in the studies described above, however, have one major limitation. These methods cover only up to 1.5% of the total number of CpG sites. Only with genome-wide analysis of DNA methylation one could assemble the complete picture of the epigenome and further understand its role in T2D. Recently, Volkov et al. carried out such a comprehensive study on genome-wide DNA methylation in pancreatic islets. They made use of a method called whole-genome bisulfate sequencing (WGBS) and were able to analyze 83% of all the CpG sites in the genome. With a single-nucleotide resolution, this is the most extensive method to study DNA methylation and is considered to be the golden standard. This next-generation sequencing method can identify DNA methylation of cytosine residues by treating the DNA with sodium bisulfate prior to sequencing. Unmethylated cytosines are converted to uracil upon

sodium bisulfate treatment and all the remaining cytosines are thus methylated. The study by Volkov et al, which is one of the most, if not the most, exhaustive study on DNA methylation in pancreatic islets, identified 25,820 DMRs in diabetic islets. These DMRs were located on genes with known importance in islet function, such as PDX1, TCF7L2, and ADCY5. More importantly, they identified 457 genes, that not only had DMRs, but also significant changes in gene expression. They showed that the altered DNA methylation of these genes resulted in impaired insulin secretion by overexpression or silencing of the candidate genes. Taken together, there is now strong evidence underlying the concept that abnormal DNA methylation contributes to islet dysfunction in T2D (45).

Besides investigating DNA methylation in pancreatic islets, studies have also analyzed DNA methylation in liver, skeletal muscle, and adipose tissue in the context of T2D. The Illumina arrays have identified differently methylated CpG sites in these disease-relevant tissues. The integrated data on alterations in DNA methylation for the different tissues supports the role of epigenetics in this multifactorial disease (46,47).

Chromatin structure in T2D pancreatic islets

In addition to DNA methylation, identifying alterations in chromatin structure and histone modifications in T2D is necessary to further understand the contribution of epigenetics to the disease. Four epigenetic methods that are often used to study different aspects of chromatin organization categorized into histone modifications, chromatin accessibility, and chromatin conformation, are chromatin immunoprecipitation sequencing (ChIP-seq), chromatin immunocleavage sequencing (ChIC-seq), assay for transposase-accessible chromatin sequencing (ATAC-seq), and high-throughput chromosome conformation capture (Hi-C) respectively. However, with regard to T2D, there is a large gap in knowledge on genome wide histone modifications and chromatin structure, since most studies employing these techniques have only studied healthy pancreatic islets. These studies have identified epigenetic landscapes in healthy pancreatic islets, but data remains limited on the alterations that occur in T2D islets or other target tissues.

ChIP-seq is a powerful genome-wide tool to map histone modifications, transcription factor binding sites, and nucleosome positioning. The epigenetic state of chromatin regulates the accessibility of transcription factors and other proteins complexes to the DNA. The ChIP-seq technique can detect specific protein binding regions with high resolution. This is relevant for epigenetics since positions of histones with specific modifications but also methylated cytosines can be determined (48). The numerous data sets that can be generated with this technique can be combined into genome-wide profiles that provide insight into the organization

and regulation of the genome. This technique uses antibodies that can bind to specific proteins of interest. These can be histones or other protein complexes attached to the DNA. First, DNA is fragmented resulting in segments with or without the protein of interest. Next, specific antibodies bind to the fragments containing the protein of interest. The DNA segments with the protein-antibody complexes can be extracted and analyzed to identify the specific protein binding sites.

While the ChIP-seq methods is a widely used technique that has been applied in many studies to determine the interactions between protein complexes and the chromatin, experiments have reported limitations of this technique associated with the fragmentation-solubilization step. Therefore, researchers have developed an alternative method called ChIC-seq that omits this step. This method involves the addition of a fusion protein, consisting of protein A and micrococcal nuclease (MNase), to the protein of interest via recruitment by a specific antibody. Addition of Ca2+ ions activates the specifically marked nucleases, which locally causes double-stranded DNA breaks. These breaks can be mapped with high resolution and specificity (49).

ATAC-seq is a relatively new high-throughput method that enables profiling of chromatin accessibility. ATAC-seq directly identifies open chromatin regions, which can provide powerful insights on key factors of gene regulation. Hyperactive Tn5 transposases simultaneously cut accessible chromatin and attach DNA sequencing adapters to mark open regions of the genome, a process called tagmentation. The tagged DNA fragments are purified, PCR-amplified, and sequenced. The read values resulting from the sequencing data represent how open specific regions of the genome were. With a single-nucleotide resolution, this technique reveals how accessibility is associated with gene regulation via DNA binding proteins, nucleosomes, and higher order chromatin compaction (50).

The Hi-C technology reveals the interaction between the spatial orientation of chromatin and gene regulation. This technique captures these sequence interactions across the entire genome by crosslinking with formaldehyde. Crosslinked DNA is digested by endonucleases and the ends are biotinylated and ligated to form chimeric DNA molecules. The ends that are not ligated, are depleted for biotin. Next, the molecules with biotinylated junctions are purified and sequenced (51).

Methodology

Bulk sequencing vs. single cell sequencing

The final part of this review will focus on the methodology used to study the epigenome, particularly in the context of T2D. Important for this discussion are the technological advances that enable the analysis of the mechanisms underlying gene regulation at a single-cell level. The previously mentioned techniques, WGBS, ChIP-seq, ChIC-seq, ATAC-seq, and Hi-C have been applied to characterize a plethora of regulatory elements, including chromatin structure, histone modifications, and DNA methylation. However, epigenetic regulation is cell specific and the effect size of each individual epigenetic alteration is small. This makes it difficult to determine the effect on gene regulation of cell-specific epigenetic alterations in profiles from heterogenous tissue samples. As a result, the data only represents mixture averages in these types of bulk studies, thereby masking possible regulatory differences between cells. The importance of this cellular heterogeneity within tissues is now being recognized, which will help to define the regulatory mechanisms underlying transcriptional variation between individual cells in development and disease (52).

This major limitation with bulk studies also affects the epigenome studies in pancreatic islets for T2D. Pancreatic islets are a mixture of cell types, including alpha, beta, delta, gamma, and epsilon cells. These different endocrine cell types each secrete a unique hormone, that together contribute to the precise regulation of glucose homeostasis. Because of this tissue heterogeneity it is difficult to understand how cell type specific epigenetic alterations contribute to overall disease etiology, as primarily beta-cell defects are important in T2D. In addition, with T2D disease progression, the cellular composition of the islets changes and there is a large reduction in beta-cell mass. This has a major effect on bulk analysis of islet tissue samples of T2D donors, since these tissue samples might contain only a small number of beta cells or even almost no beta cells at all with severe T2D. This illustrates the importance to isolate specific cell types.

Analysis of DNA methylation, chromatin accessibility, or histone modifications with a single-cell resolution can lead to new discoveries of connections between regulatory epilayers and gene activity. Moreover, the contribution of the 3D organization of chromatin to gene regulation can be further explored with single-cell resolutions. Capturing sequences that are in close proximity in a 3D configuration, allows for the identification of promotor and enhancer regions. In addition, non-coding genetic variants, while initially not thought to be of importance for gene expression, have been shown to influence genes via the spatial organization of the chromatin, which enables physical accessibility to the gene. This knowledge has led to the identification

of new disease-associated variants outside of the coding regions, that can still affect the activity of a target gene, also in T2D. This emphasizes the importance of methods that capture the 3D chromatin regulatory landscape, such as Hi-C, especially at a single-cell resolution (53).

The same is true for DNA methylation, histone modifications, and chromatin accessibility. Studying these regulatory features with a single-cell resolution can provide new insights on their effect on gene transcription. This is particularly relevant for such a heterogenous disease with a large interindividual variation as T2D. The study of single cells in T2D tissue samples could help decipher the individual contributions of epigenetic alterations that act together to result in T2D phenotype. Currently, the type and the extent of epigenetic alterations needed to change gene expression remains to be defined (52).

Current limitations of single-cell sequencing methods

Recent advances in single-cell epigenomic methods now enable the analysis of gene regulatory elements at a single cell level. However, such single-cell studies face several challenges because data interpretation is complicated and reconfiguration of the existing methods to a single cell resolution is associated with experimental difficulties.

To develop valuable single-cell assays, a number of parameters have to be optimized such as the throughput, the robustness, the complexity, and the accuracy. However, due to the distinct biochemical nature of the different epigenetic marks, each requires a specific capture approach. As a consequence, the quality of the output and the ability to integrate different data sets from the same cell can vary for the different epigenetic modifications. Each approach requires its own consideration of the depth and the throughput of the assay necessary to generate the most informative results. The complexity, which indicates the number of distinct molecules that can be captured from each cell, is particularly challenging for single-cell epigenomics because of the extremely sparse signal resulting from only two genomic copies. Per locus per cell only 0, 1, or 2 reads can be observed which markedly complicates data analysis. Thus, not only technical and experimental innovations are needed to allow high-throughput assays of good quality and large coverage of single cells, but also equally advanced computational tools are necessary (54).

Bisulfate sequencing (BS-seq) is a well-established technique that is commonly employed to generate whole-genome DNA methylation profiles. Two other commonly used DNA methylation profiling methods are affinity purification and methylation specific restriction enzymes. However, BS-seq has a higher resolution, covers more CpGs, and provides more reads. Although being considered the golden-standard, BS-seq has some major limitations.

Firstly, BS-seq cannot distinguish between 5-methylcytosine and the recently discovered 5-hydroxymethylcytosine DNA modification, since both modifications are read as a cytosine nucleotide when sequenced. Modified BS-seq protocols, such as TET-assisted BS-seq and 5hmC-specific restriction enzymes, have been developed to allow the analysis of 5-hmC. Secondly, the single nucleotide resolution of this method relies on the complete conversion of all the unmethylated cytosine residues to uracil. As a consequence, incomplete conversion will produce false positives. Finally, a major limitation, which also largely complicates the adaptation to a protocol for single cells, is the DNA degradation that occurs with the bisulfate induced conversion. Therefore, complete coverage of the genome is currently impossible. With the recent technological advances in single cell BS-seq (scBS-seq) protocols a coverage of up to 50% of the CpG sites has been achieved (55). However, these assays are low-throughput because each cell has to be treated as a separate sample. Combinatorial cell barcoding (figure 3b) with subsequent BS-seq has markedly increased the throughput, enabling the parallel processing of thousands of individual cells for whole-genome DNA methylation profiles in a single experiment. However, the depth of information per cell is limited. The readouts can be traced back to individual cells, because each cell is labelled with a unique combination of barcodes. Moreover, both scBS-seq and scRNA-seq can be performed on the same single cell, which provides information on the relationship between transcription and epigenetics.

In addition to single cell DNA methylation studies, a number of techniques have been adapted to study the chromatin state at a single cell level. For the single cell analysis of chromatin accessibility, the most high-throughput measure (>10.000 samples) available has been adapted from the previously mentioned ATAC-seq assay. Similar to scBS-seq, barcoding technologies, as shown in figure 3, that introduce cell-identifying labels, are integrated in single cell ATAC-seq (scATAC-seq) protocols for the identification of individual cells within the large number of readouts. In addition to combinatorial barcoding, droplet microfluidics (figure 3a) is among the most promising techniques to improve scalability in single-cell techniques. This method isolates the cells in droplets and introduces a cell specific barcode. As a result, the DNA fragments that are collected in the pooling process can be linked to the cell of origin. Compared with the combinatorial barcoding, the microfluidics approach for scATAC-seq has a markedly higher resolution, with an average of 70,000 reads per cell compared with 3000, yet this is at the expense of the throughput. The read depth (coverage), describing the number of times a unique nucleotide is sequenced, and the throughput level of single cell experiments, describing the number of cells that is included, are important determinants for the strength of the data set. Low-throughput or low coverage limits the identification of cellular variation in chromatin accessibility, because of sparse or missing data.

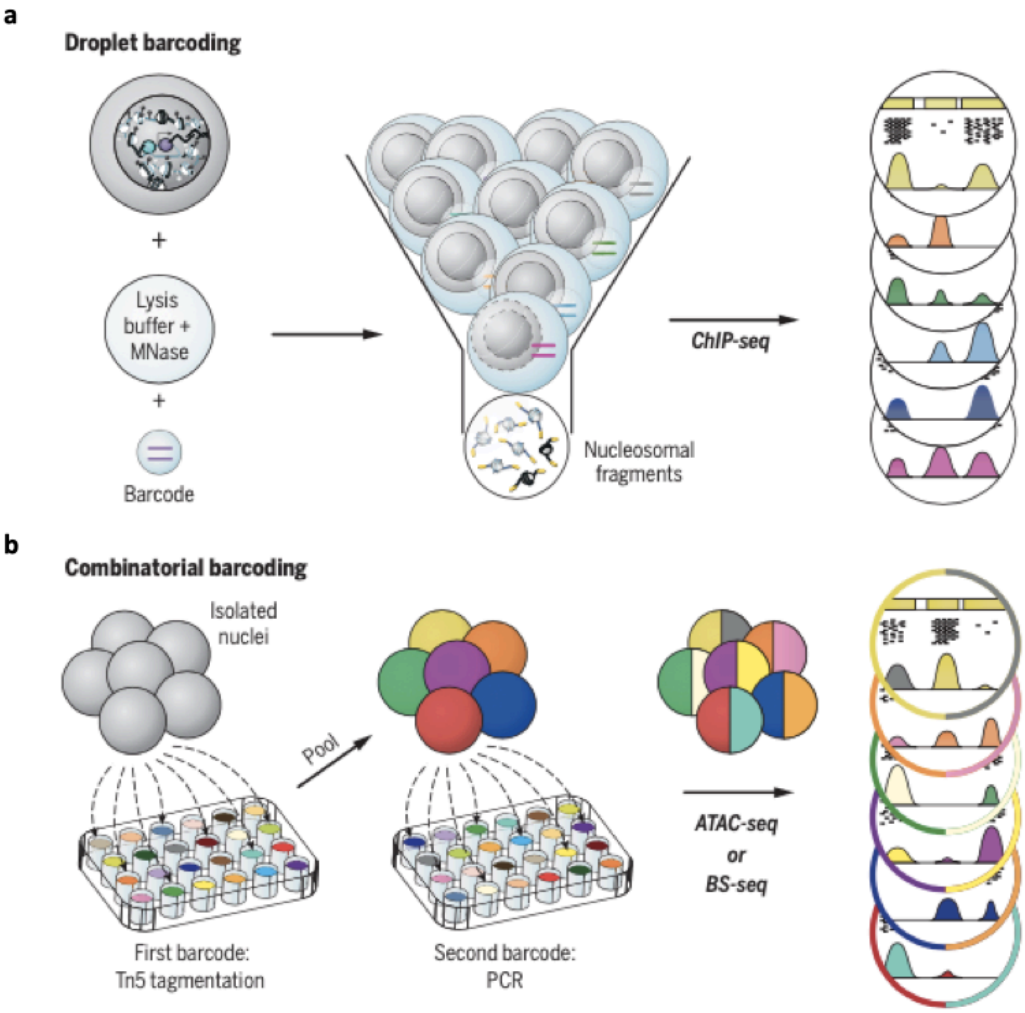

Figure 3. Different barcoding methods to enable high throughput analysis. a, Droplet barcoding. Single cells are isolated in droplets and the chromatin of individual cell is labelled. This enables the analysis of thousands of cells, by pooling and subsequent bulk analysis. **b,** Combinatorial barcoding. Readouts can be traced back to individual cells by the unique combinations of barcodes that has been introduced to each cell. This allows for the analysis of large numbers of cells. However, the depth of information for each cell is limited. Note: adapted from Single-cell epigenomics: Recording the past and predicting the future, by Kelsey,G. et al., 2017, Science, 358, 69-75 *(54)*.

Studying histone modifications at a single-cell level remains the most challenging, because adaptation of the ChIP-seq method, which is most commonly used to study histone modifications, faces problems with specificity. This technique depends on the binding of antibodies to pull down specific proteins of interest and associated DNA fragments. The nonspecific pull-down of off-target antibodies in the immunoprecipitation step results in experimental noise. Especially with small-input experiments, such as single cell assays, this problem is exacerbated, which markedly limits its compatibility for single-cell studies. In a recent study this major limitation was partly overcome by using the droplet approach. Isolated cells were lysed and the chromatin was fragmented by MNase within droplets. Barcode droplets were introduced to these chromatin containing droplets, thus uniquely labelling cells before immunoprecipitation. This enables sequencing of large numbers of pool of cells, which

avoids nonspecific noise associated with low-input samples. Although this method allows for the detection of cell-cell variation based on known peaks, the single cell profile is not yet sensitive enough to identify de novo peaks (56).

A more promising single-cell method to evaluate chromatin states can be adapted from the ChIC-seq method. Single cell ChIC-seq (scChIC-seq) consists of specific MNase cleavage, which is recruited by a specific antibody, at sites of histone modifications or other bound protein targets. The target DNA fragments can then be separated from the non-target DNA fragments by selective PCR amplification. Ku et. al. (2019) developed this method and showed that reliable identification of H3K4me3 and H3K27me3 histone targets in single human white blood cells was possible (57). The single-cell methods discussed in this review and their application are shown in figure 4.

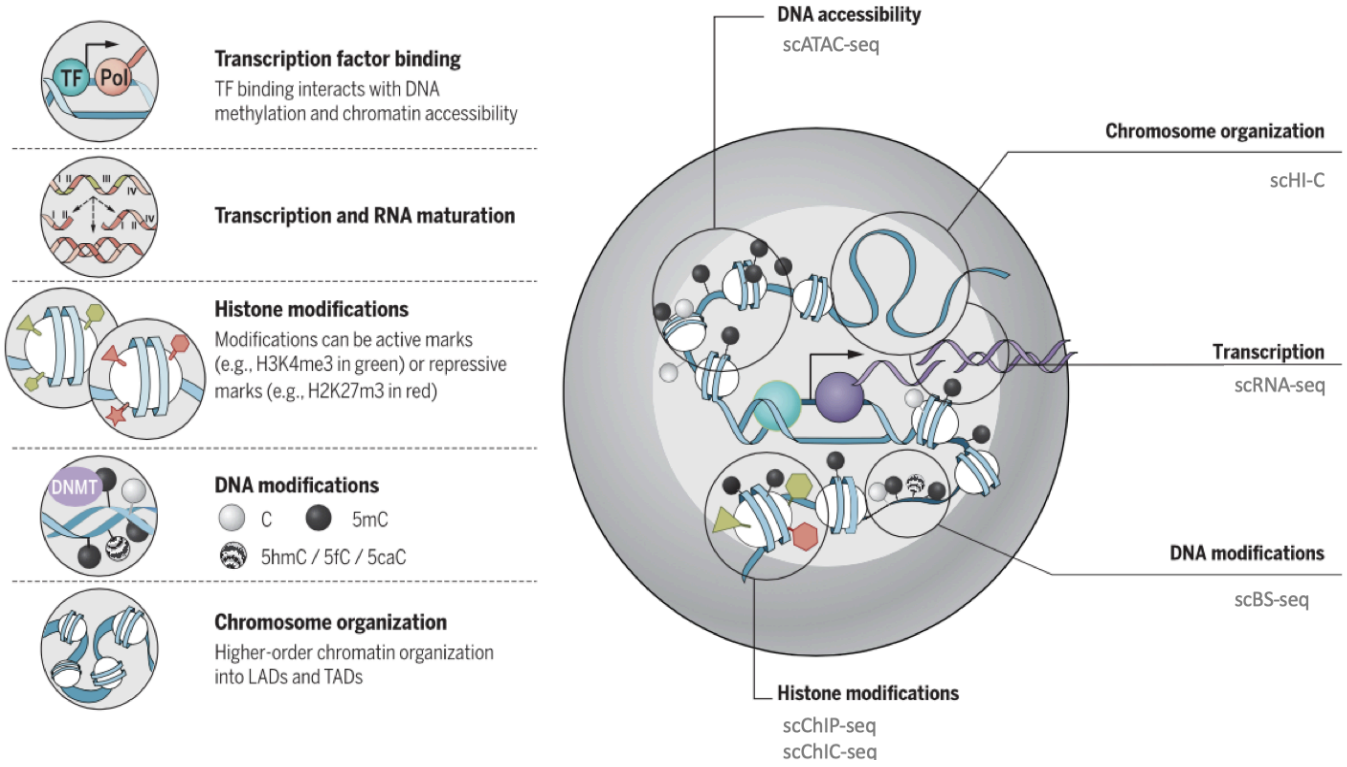

Figure 4. Single-cell methods for the different gene regulatory elements. Left, overview of the different epigenetic mechanisms involved in gene regulation. **Right,** the single-cell methods that have been developed to study the different epigenetic layers of gene regulation within a cell. Note: adapted from Single-cell epigenomics: Recording the past and predicting the future, by Kelsey,G. et al., 2017, Science, 358, 69-75 *(54)*.

The adaptation of existing epigenetic techniques to obtain single-cell resolution data, although being in its infancy, has already shown that these approaches are cutting-edge tools, capable of elucidating unknown gene regulatory processes. However, the currently available single-cell techniques are not yet advanced enough to provide information on cell to cell variability in

the context of epigenetic gene regulation. Further development of these single-cell techniques, which will inevitably be seen in the future, is largely dependent on the rate of developments in the technological field. Ultimately, the most useful datasets will be obtained if transcriptomics and epigenetics can be analyzed in parallel. However, extracting RNA from cells while maintaining both the RNA integrity and the chromatin structure, remains an enormous challenge. Moreover, since data from scATAC-seq or single cell ChIP-seq (scChIP-seq) methods is sparse, acquiring data on both gene expression and chromatin structure at a specific locus is very difficult. To obtain sufficient parallel data, analysis of large numbers of cells is required.

In addition to the methods mentioned above, it is also possible to define the 3D orientation of the chromatin at a single cell level with the single-cell Hi-C (scHi-C) method. Despite recent advances and optimizations, this technique currently provides very limited information. However, it still enables the identification of chromosome conformation and compartmentalization. Also, this method can describe interaction between different chromosomes within a single cell. Further technical improvements could help to more specifically elucidate the regulatory elements that were obscured in bulk-cell data, such as promoter-enhancer interactions, especially in combination with functional studies.

These single-cell methods face a set of similar limitations that require both technical and computational improvements to overcome these challenges. Currently, a major limitation is the low coverage, which means that data analysis relies on computational tools to fill in the missing information per cell. Another challenge is the mappability of the sequence reads. Low mappability, which means that sequence reads are unmappable to a reference genome, limits de novo assembly.

In summary, epigenome-based single cell methods are not as advanced as transcriptomic methods such as scRNA-seq in terms of throughput and coverage. However, with the recent advances that have been made, such as barcoding of cells, and with the rapid evolving technological field, the epigenetic techniques described in this review will become as powerful as single cell transcriptomic methods.

Single-cell studies in T2D

Even though single-cell epigenome-based methods are in its infancy, they can potentially provide new insights in development and disease processes, also in T2D which is thought to be partly caused by epigenetic dysregulation. To this date, no epigenetic studies have been performed with these single-cell methods to investigate cells from T2D tissue samples. There

are, however, a few studies using these techniques to study the chromatin of non-diabetic pancreatic islet cells, which is a disease-relevant tissue for T2D. The next paragraph will highlight these findings to illustrate the importance of single cell epigenetic studies for the field of T2D. A recent study by Rai et al. used scATAC-seq in combination with combinatorial barcoding to investigate the chromatin accessibility in pancreatic alpha, beta and delta cells. Recent GWAS have identified over 400 risk-associated variants for T2D, most of which are in non-coding regions of the genome. These non-coding genetic variants are known to contribute to disease risk by altering gene regulation or chromatin structure. However, for most of the identified risk variants the exact regulatory molecular mechanisms remain elusive. Rai et al. were able to link target genes of T2D GWAS variants to cell-specific open-chromatin regions. Interestingly, T2D genetic variants were significantly enriched in beta-cell specific open-chromatin regions, indicating this to be the islet cell type of action based on genetic variants associated with T2D predisposition. A similar study performed by Greenwald et al. further demonstrates the power of single-cell epigenome methods to elucidate complex disease genetics such as T2D. They generated accessible chromatin profiles for each cell, and similar to the study by Rai et al., they were able to identify regulatory elements and genes affected by non-coding risk variants derived from GWAS. Identifying these target genes and their regulatory elements for T2D is important for understanding the genetic basis and molecular dysregulation of such a complex disease (58,59).

To summarize, gene expression is the outcome of precisely regulated transcription by several regulatory elements, including DNA methylation, histone modifications, and chromatin organization. Regulatory regions together with other factors, such as transcription factors, determine the gene activity. High-throughput epigenomic profiling methods have already largely contributed to the elucidation of epigenetic gene regulation and there is now enough evidence that supports the role of epigenetics in T2D. However, gene expression is cell specific as is the closely related epigenomic profile. Understanding cell specific epigenetic gene regulation is paramount for understanding the development of complex diseases such as T2D. However, this cell-to-cell variation is masked in bulk measurements of heterogenous tissue samples, which limits the understanding of how cell specific epigenetic dysregulation could contribute to disease. With recent technological advances, single-cell epigenomic measurements have become available and their potential to link regulatory elements to target genes has already been demonstrated. The study of single cells could determine the extent of epigenetic alterations needed to result in disease phenotype, which could provide new treatment options. Yet, single-cell studies in the context of T2D are extremely limited. However, the advancements regarding epigenetic dysregulation in T2D should not be undermined, since it has only been 10 years that studies have been performed in this research field. With that in

mind, tremendous progress has been made, and hand-in-hand with the rapidly evolving technology, the future looks promising and might provide answers to several remaining questions that can only be addressed with single-cell epigenome studies. For example, what is the relationship between transcriptional and epigenetic variation between cells, do epigenetic modifications affect transcription or are epigenetic modifications the result of transcriptional changes, are both transcriptomics and epigenetics equally capable to identify cell populations? Most important, defining epigenetic dysregulation in T2D can provide interesting opportunities for new therapeutic or prevention methods, because of its reversible nature.

References

(1) Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes, J. D., Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018 Apr;138:271-281.

(2) Groop L, Pociot F. Genetics of diabetes--are we missing the genes or the disease? Mol Cell Endocrinol 2014 Jan 25;382(1):726-739.

(3) Gregg EW, Sattar N, Ali MK. The changing face of diabetes complications. Lancet Diabetes Endocrinol 2016 Jun;4(6):537-547.

(4) Fonseca VA. Identification and treatment of prediabetes to prevent progression to type 2 diabetes. Clin Cornerstone 2007;8(2):10-20.

(5) Fonseca VA. Defining and characterizing the progression of type 2 diabetes. Diabetes Care 2009;32 Suppl 2:S151-S156.

(6) ,. Standards of Medical Care in Diabetes—2010. Diabetes Care 2010;33:S11.

(7) ,. International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes. Diabetes Care 2009;32(7):1327.

(8) ,. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 1997;20(7):1183.

(9) World HO, International DF. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia : report of a WHO/IDF consultation. 2006.

(10) Weyer C, Bogardus C, Mott DM, Pratley RE. The natural history of insulin secretory dysfunction and insulin resistance in the pathogenesis of type 2 diabetes mellitus. J Clin Invest 1999 Sep;104(6):787-794.

(11) Lyssenko V, Almgren P, Anevski D, Perfekt R, Lahti K, Nissén M, et al. Predictors of and longitudinal changes in insulin sensitivity and secretion preceding onset of type 2 diabetes. Diabetes 2005 Jan;54(1):166-174.

(12) DeFronzo RA. Lilly lecture 1987. The triumvirate: beta-cell, muscle, liver. A collusion responsible for NIDDM. Diabetes 1988 Jun;37(6):667-687.

(13) Boden G. Role of fatty acids in the pathogenesis of insulin resistance and NIDDM. Diabetes 1997 Jan;46(1):3-10.

(14) Ohn JH, Kwak SH, Cho YM, Lim S, Jang HC, Park KS, et al. 10-Year Trajectory of β-Cell Function and Insulin Sensitivity in the Development of Type 2 Diabetes: a Community-Based Prospective Cohort Study. Lancet Diabetes Endocrinol 2016 Jan;4(1):27-34.

(15) DeFronzo RA, Ferrannini E, Hendler R, Felig P, Wahren J. Regulation of splanchnic and peripheral glucose uptake by insulin and hyperglycemia in man. Diabetes 1983 Jan;32(1):35-45.

(16) DeFronzo RA, Ferrannini E. Regulation of hepatic glucose metabolism in humans. Diabetes Metab Rev 1987 Apr;3(2):415-459.

(17) Gastaldelli A, Toschi E, Pettiti M, Frascerra S, Quiñones-Galvan A, Sironi AM, et al. Effect of physiological hyperinsulinemia on gluconeogenesis in nondiabetic subjects and in type 2 diabetic patients. Diabetes 2001 Aug;50(8):1807-1812.

(18) Gastaldelli A, Baldi S, Pettiti M, Toschi E, Camastra S, Natali A, et al. Influence of obesity and type 2 diabetes on gluconeogenesis and glucose output in humans: a quantitative study. Diabetes 2000 Aug;49(8):1367-1373.

(19) Campbell PJ, Mandarino LJ, Gerich JE. Quantification of the relative impairment in actions of insulin on hepatic glucose production and peripheral glucose uptake in non-insulin-dependent diabetes mellitus. Metabolism 1988 Jan;37(1):15-21.

(20) Baron AD, Schaeffer L, Shragg P, Kolterman OG. Role of hyperglucagonemia in maintenance of increased rates of hepatic glucose output in type II diabetics. Diabetes 1987 Mar;36(3):274-283.

(21) DeFronzo RA, Tripathy D. Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. Diabetes Care 2009 Nov;32 Suppl 2(Suppl 2):157.

(22) Ferrannini E, Simonson DC, Katz LD, Reichard G,Jr, Bevilacqua S, Barrett EJ, et al. The disposal of an oral glucose load in patients with non-insulin-dependent diabetes. Metabolism 1988 Jan;37(1):79-85.

(23) Gerich JE, Meyer C, Woerle HJ, Stumvoll M. Renal gluconeogenesis: its importance in human glucose homeostasis. Diabetes Care 2001 Feb;24(2):382-391.

(24) Vallon V, Thomson SC. Targeting renal glucose reabsorption to treat hyperglycaemia: the pleiotropic effects of SGLT2 inhibition. Diabetologia 2017 Feb;60(2):215-225.

(25) DeFronzo RA. Pathogenesis of type 2 diabetes mellitus. Med Clin North Am 2004 Jul;88(4):787-835, ix.

(26) Reaven GM, Hollenbeck CB, Chen YD. Relationship between glucose tolerance, insulin secretion, and insulin action in non-obese individuals with varying degrees of glucose tolerance. Diabetologia 1989 Jan;32(1):52-55.

(27) Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, Butler PC. β-Cell Deficit and Increased β-Cell Apoptosis in Humans With Type 2 Diabetes. Diabetes 2003;52(1):102.

(28) Prentki M, Nolan CJ. Islet beta cell failure in type 2 diabetes. J Clin Invest 2006;116(7):1802-1812.

(29) Efrat S. Beta-Cell Dedifferentiation in Type 2 Diabetes: Concise Review. Stem Cells 2019;37(10):1267-1272.

(30) Hunter CS, Stein RW. Evidence for Loss in Identity, De-Differentiation, and Trans-Differentiation of Islet β-Cells in Type 2 Diabetes. Frontiers in Genetics 2017;8:35.

(31) Guilherme A, Virbasius JV, Puri V, Czech MP. Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. Nat Rev Mol Cell Biol 2008 May;9(5):367-377.

(32) Itani SI, Ruderman NB, Schmieder F, Boden G. Lipid-induced insulin resistance in human muscle is associated with changes in diacylglycerol, protein kinase C, and IkappaB-alpha. Diabetes 2002 Jul;51(7):2005-2011.

(33) RANDLE PJ, GARLAND PB, HALES CN, NEWSHOLME EA. The glucose fatty-acid cycle. Its role in insulin sensitivity and the metabolic disturbances of diabetes mellitus. Lancet 1963 Apr 13;1(7285):785-789.

(34) Reaven GM. The fourth musketeer--from Alexandre Dumas to Claude Bernard. Diabetologia 1995 Jan;38(1):3-13.

(35) Li E, Zhang Y. DNA methylation in mammals. Cold Spring Harb Perspect Biol 2014 May 1;6(5):a019133.

(36) Richa R, Sinha RP. Hydroxymethylation of DNA: an epigenetic marker. EXCLI J 2014 May 27;13:592-610.

(37) Margueron R, Reinberg D. Chromatin structure and the inheritance of epigenetic information. Nat Rev Genet 2010 Apr;11(4):285-296.

(38) Miller JL, Grant PA. The role of DNA methylation and histone modifications in transcriptional regulation in humans. Subcell Biochem 2013;61:289-317.

(39) Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature Reviews Genetics 2012;13(7):484-492.

(40) Hjort L, Novakovic B, Grunnet LG, Maple-Brown L, Damm P, Desoye G, et al. Diabetes in pregnancy and epigenetic mechanisms-how the first 9 months from conception might affect the child's epigenome and later risk of disease. Lancet Diabetes Endocrinol 2019 Oct;7(10):796-806.

(41) Yang BT, Dayeh TA, Kirkpatrick CL, Taneera J, Kumar R, Groop L, et al. Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA(1c) levels in human pancreatic islets. Diabetologia 2011 Feb;54(2):360-367.

(42) Hall E, Dayeh T, Kirkpatrick CL, Wollheim CB, Dekker Nitert M, Ling C. DNA methylation of the glucagon-like peptide 1 receptor (GLP1R) in human pancreatic islets. BMC Med Genet 2013 Jul 23;14:76-76.

(43) Volkmar M, Dedeurwaerder S, Cunha DA, Ndlovu MN, Defrance M, Deplus R, et al. DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients. EMBO J 2012;31(6):1405-1426.

(44) Volkov P, Olsson AH, Gillberg L, Jørgensen SW, Brøns C, Eriksson KF, et al. A Genome-Wide mQTL Analysis in Human Adipose Tissue Identifies Genetic Variants Associated with DNA Methylation, Gene Expression and Metabolic Traits. PLoS One 2016 Jun 20;11(6):e0157776.

(45) Volkov P, Bacos K, Ofori JK, Esguerra JL, Eliasson L, Rönn T, et al. Whole-Genome Bisulfite Sequencing of Human Pancreatic Islets Reveals Novel Differentially Methylated Regions in Type 2 Diabetes Pathogenesis. Diabetes 2017 Apr;66(4):1074-1085.

(46) Nilsson E, Jansson PA, Perfilyev A, Volkov P, Pedersen M, Svensson MK, et al. Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. Diabetes 2014 Sep;63(9):2962-2976.

(47) Nilsson E, Matte A, Perfilyev A, de Mello VD, Käkelä P, Pihlajamäki J, et al. Epigenetic Alterations in Human Liver From Subjects With Type 2 Diabetes in Parallel With Reduced Folate Levels. J Clin Endocrinol Metab 2015 Nov;100(11):1491.

(48) Yan H, Tian S, Slager SL, Sun Z. ChIP-seq in studying epigenetic mechanisms of disease and promoting precision medicine: progresses and future directions. Epigenomics 2016;8(9):1239-1258.

(49) Schmid M, Durussel T, Laemmli UK. ChIC and ChEC: Genomic Mapping of Chromatin Proteins. Mol Cell 2004;16(1):147-157.

(50) Sun Y, Miao N, Sun T. Detect accessible chromatin using ATAC-sequencing, from principle to applications. Hereditas 2019 Aug 15;156:29-9. eCollection 2019.

(51) Pal K, Forcato M, Ferrari F. Hi-C analysis: from data generation to integration. Biophysical Reviews 2019;11(1):67-78.

(52) Shema E, Bernstein BE, Buenrostro JD. Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. Nat Genet 2019;51(1):19-25.

(53) Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 2015 Jul 23;523(7561):486-490.

(54) Kelsey G, Stegle O, Reik W. Single-cell epigenomics: Recording the past and predicting the future. Science 2017 Oct 6;358(6359):69-75.

(55) Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods 2014 Aug;11(8):817-820.

(56) Ku WL, Nakamura K, Gao W, Cui K, Hu G, Tang Q, et al. Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. Nature Methods 2019;16(4):323-325.

(57) Rai V, Quang DX, Erdos MR, Cusanovich DA, Daza RM, Narisu N, et al. Single cell ATAC-seq in human pancreatic islets and deep learning upscaling of rare cells reveals cell-specific type 2 diabetes regulatory signatures. bioRxiv 2019:749283.

(58) Greenwald WW, Chiou J, Yan J, Qiu Y, Dai N, Wang A, et al. Pancreatic islet chromatin accessibility and conformation reveals distal enhancer networks of type 2 diabetes risk. Nature Communications 2019;10(1):2078.