Cape Peninsula University of Technology

GENOME-WIDE DNA METHYLATION PROFILING IN MIXED ANCESTRY SOUTH AFRICANS WITH DIABETES OR PREDIABETES

By

WENDY LYNN SOLOMON

Thesis submitted in fulfilment of the requirements for the degree

Doctor of Philosophy (PhD): Biomedical Sciences

In the Faculty of Health & Wellness Sciences

At the Cape Peninsula University of Technology

Supervisor: Prof TE Matsha

Co-supervisor: Dr SBE Hector

Bellville

March 2021

DECLARATION

I, **WENDY LYNN SOLOMON**, declare that the contents of this thesis represent my own unaided work and that the thesis has not previously been submitted for academic examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of the Cape Peninsula University of Technology.

Signed  **Date:** 24 March 2021

ABSTRACT

Type 2 diabetes mellitus (T2DM) is a global health concern and has shown to be a rising public health crisis in Africa. South Africa, in particular, has shown the highest prevalence of the disease with a large percentage of undiagnosed individuals. There is a need for intervention strategies that include early detection of at-risk individuals and the prevention of disease progression. Although studies have shown a relationship between the occurrence of T2DM and genetic and lifestyle factors, it has been suggested that epigenetic mechanisms such as DNA methylation contribute to the pathogenesis of T2DM through its association with the transcriptional activity of genes. There are however limited studies focussing on the genome-wide DNA methylation profiling of diabetic and prediabetic subjects from Africa and in particular, South Africa. The aim of the current study therefore was to conduct genome-wide DNA methylation in South African subjects with varying glucose tolerance and investigate the relationship between the observed DMRs and cardiometabolic risk factors.

A cross-sectional case-control study design was used to conduct genome-wide DNA methylation profiling and lncRNA analysis on the peripheral blood of 48 South African individuals from the Bellville South community in the Western Cape. The participants were classified according to their glucose tolerance status and comprised 12 participants with known diabetes on metformin treatment, 12 screen-detected (newly diagnosed) diabetics, 12 prediabetics and 12 participants with normal glucose tolerance (NGT). After DNA extraction, DNA immunoprecipitation sequencing was used to identify statistically significant differentially methylated regions (DMRs) and lncRNA-associated DMRs, followed by gene ontology and KEGG pathway analysis. Lastly, the significant DMRs identified were validated by performing pyrosequencing of bisulphite converted DNA.

The study identified several DMRs and functional pathways affected in subjects with diabetes and prediabetes. A total of 366 DMRs and 641 lncRNA-associated DMRs were observed, of which 63% were hypermethylated and 37% hypomethylated. Gene ontology and KEGG pathway analyses identified hypermethylation in cardiovascular processes in screen-detected diabetic subjects indicating that a decrease in the expression of the associated genes may be associated with these functions in diabetes. Furthermore, hypomethylation was evident in purine metabolism in the screen-detected diabetic subjects highlighting the risk of increased reactive oxygen species production, inflammation and cell damage associated with excess uric acid in diabetic individuals. The study also identified DMRs and their functional pathways which showed a possible progression from prediabetes to diabetes. Hypomethylation of LBP (Lipopolysaccharide Binding Protein) in prediabetic subjects suggested that increased expression of this DMR could potentially be used as a biomarker for

the progression to diabetes due to its association with increased inflammatory cytokines, insulin resistance and beta-cell dysfunction. Furthermore, hypomethylation of the regulation of wound healing and blood coagulation and their associated DMRs, SERPINF2 and DMTN, was observed in prediabetic subjects. Elevated levels of SERPINF2 (alpha 2-antiplasmin) and DMNT (dematin) expression in hyperglycaemic subjects could be used as an indicator of the increased risk for cardiovascular disease and red blood cell stability.

Metformin, which is used as the first line of treatment for T2DM, has been shown to be affected by DNA methylation patterns. Hypermethylation of the cytokine-cytokine receptor interaction and oxidative phosphorylation pathways in known diabetics on metformin treatment suggested that metformin may have an inhibitory effect on complement-mediated inflammation and mitochondrial oxidative phosphorylation in diabetic individuals. The cAMP signalling pathway was also found to be hypermethylated in metformin-treated subjects which suggest the use of DNA methylation as a potential marker to monitor the effects of metformin on glucose homeostasis. Furthermore, hypermethylation of the functional pathways and DMRs identified in metformin-treated subjects when compared to the untreated screen-detected subjects were associated with Diabetic Peripheral Neuropathy (DPN).

Several novel lncRNAs were observed when comparing the lncRNA-associated DMRs identified amongst individuals with varying degrees of glucose tolerance. Hypomethylation of the complement component C4 was observed in hyperglycaemic subjects suggesting a possible association with the cardiometabolic risk factors and complement-mediated inflammation associated with T2DM. The lncRNA-associated DMRs observed in metformin-treated subjects included the mitochondrial ATP synthase-coupling factor 6 (ATP5J) enzyme thought to be involved in the oxidative phosphorylation pathway with associations to T2DM and hypertension.

In conclusion, these findings show that DNA methylation patterns differ amongst individuals with varying degrees of glucose tolerance within a South African population. Furthermore, the study showed that DNA methylation patterns are associated with certain cardiometabolic traits and diabetic complications, and could be used as potential biomarkers for the occurrence and progression of T2DM. The expression of several DMRs and lncRNA-associated DNA methylation regions observed in metformin-treated T2DM may also be potential targets for therapeutic monitoring in patients with diabetes.

ACKNOWLEDGEMENTS

I wish to thank:

- Professor Tandi Matsha, my supervisor for her support, encouragement and mentorship during the PhD. Thank you for believing in me and always providing opportunities for upliftment and future growth, even before this PhD journey.
- Dr Stanton Hector, my co-supervisor for the support and words of encouragement during the PhD project. Thank you for providing a sounding board especially during times of doubt and uncertainty.
- My colleagues in the Department of Biomedical Sciences for all the support during the PhD process. A special thanks to the head of department, Associate Professor Glenda Davison for understanding the challenges involved in the PhD process and providing the necessary support and encouragement.
- The SAMRC-CPUT Cardiometabolic Health Research Unit staff and fellow post-graduate students. Thank you for all the support and help. The weekly meetings during the Covid-19 lockdown period provided the ideal platform for reflection, discussions and a place to vent.
- Yvonne Prince, my colleague and friend, who walked this PhD journey with me. This journey would have been much more challenging without your support, motivation and friendship.
- My parents, Mathew and Irene, for their support, love and unfailing confidence in my ability to reach my goals. To my sisters, Tracy and Leslee, for their support during the good and bad times.
- My husband, Noel, for his unfailing support during this PhD process. Thank you for providing advice and being my sounding board. Without your love and support, I would not have had the strength to walk this path. To Emma and Kyle for all their love, hugs and kisses during this time. I appreciate the understanding and patience through all the sacrifices made during my studies.

Acknowledgements of funding:

- This research project was supported by a grant from the South African Medical Research Council (SAMRC), with funds from National Treasury under its Economic Competitiveness and Support Package (MRC-RFA-UFSP-01-2013/VMH Study), South African National Research Foundation (SANRF) (Grant no. 115450). Any opinions, findings, conclusions, or recommendations expressed in this thesis are

those of the author and are not necessarily to be attributed to the SAMRC and/or SANRF.

- The financial assistance of the University Research Funding (URF) towards this research is acknowledged. Opinions expressed in this thesis and the conclusions arrived at, are those of the author, and are not necessarily to be attributed to the URF.
- The financial assistance of the University Capacity Development Grant (UCDG) - Improvement of Qualifications Programme (IQP) towards staff replacement in 2020 is acknowledged.

DEDICATION

For Noel, Emma and Kyle

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS, SYMBOLS AND UNITS

%	Percentage
°C	Degrees Celsius
cm	Centimetres
kg/m^2	Kilograms per meter squared
IU/L	International units per litre
mg/mL	Milligrams per millilitres
mmHg	Millimetres of mercury
mmol/L	Millimoles per litre
µg	Microgram
µL	Microlitre
µM	Micromolar
mM	Millimolar
M	Molar concentration
pM	Picomolar
5-hmC	5-hydroxymethylcytosine
5-mC	5-methylcytosine
ADA	American Diabetes Association
ADK	Adenosine kinase
AGEs	Advanced glycation end-products
AMP	Adenosine monophosphate
AMPK	AMP-activated protein kinase
APS	Adenosine 5′ phosphosulfate
ATP	Adenosine triphosphate
BMI	Body mass index
bp	Base pairs
BP	Blood pressure
Ca^{2+}	Calcium
cAMP	Cyclic adenosine monophosphate
CpG	Cytosine-Guanine dinucleotide
CpGI	CpG island
CPUT	Cape Peninsula University of Technology
CRP	C-reactive protein
CVD	Cardiovascular disease
DAG	Directed Acyclic Graphs
DBP	Diastolic blood pressure
DE	Differentially enriched
DKD	Diabetic kidney disease
DM	Diabetes mellitus
DMRs	Differentially methylated regions
DNA	Deoxyribonucleic acid
DNMTs	DNA methyltransferases
DPN	Diabetic Peripheral Neuropathy
ECM	Extracellular matrix

EDTA	Ethylenediaminetetraacetic acid
EWAS	Epigenome-wide association studies
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDM	Gestational diabetes mellitus
GO	Gene ontology
GWAS	Genome-wide association studies
HbA1c	Glycosylated haemoglobin A1C
HDL	High-density lipoproteins
HPLC	High-performance liquid chromatography
IDF	International Diabetes Federation
IFG	Impaired fasting glucose
IGT	Impaired glucose tolerance
IRS1	Insulin receptor substrate 1
JNK	Jun N-terminal kinase
K^+	Potassium
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LADA	Latent autoimmune diabetes of adults
LDL	Low-density lipoproteins
LINE-1	Long interspersed nucleotide element-1
LMICs	Low- and middle-income countries
LncRNA	Long non-coding RNA
LUMA	Luminometric methylation assay
MAPK	Mitogen-activated protein kinase
MBDs	Methyl binding domains
MeCP2	Methyl-CpG-binding protein 2
MeDIP	Methylated DNA immunoprecipitation
MeDIP-seq	Methylated DNA immunoprecipitation sequencing
miRNA	microRNA
MODY	Maturity Onset Diabetes of the Young
mRNA	Messenger RNA
MS	Mass spectrophotometry
Na^+	Sodium
NaOH	Sodium hydroxide
NCDs	Non-communicable diseases
ncRNAs	Non-coding RNAs
NGT	Normal glucose tolerance
OGTT	Oral glucose tolerance test
PCR	Polymerase chain reaction
PI3K-AKT	Phosphatidylinositol 3-kinase - protein kinase B (AKT)
PKA	protein kinase A
PPi	Pyrophosphate
qPCR	Quantitative Polymerase Chain Reaction
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rpm	Revolutions per minute
SAH	*S*-adenosyl-L-homocysteine
SAHH	S-adenosyl-L-homocysteine hydrolase

SAM	S-adenosyl-methionine
SBP	Systolic blood pressure
SNPs	Single nucleotide polymorphisms
SSA	Sub-Saharan Africa
ssDNA	Single-stranded DNA
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
TET	Ten-eleven translocation
TSS	Transcription start site
TTS	Transcription termination site
UDP-GlcNAc	UDP-N-acetylglucosamine
UDP-Xyl	UDP-xylose
VMH	Vascular and Metabolic Health
WHO	World Health Organisation

CHAPTER 1

LITERATURE REVIEW

1.1 Background

The global prevalence of diabetes is growing rapidly and it has been predicted that Africa will be one of the regions with the greatest increase in diabetes prevalence (International Diabetes Federation, 2019). Type 2 diabetes mellitus (T2DM) in Africa is further characterised by the high number of undiagnosed cases as well as a high number of individuals who are at risk of developing the condition in the future. Although T2DM has been strongly linked to environmental factors, recent genome-wide association studies have shown that genetic variation plays a role in an individual's susceptibility to diabetes. In addition, it has also been suggested that epigenetics plays a role in the link between environment and genetic factors for many diseases including T2DM (Davegårdh *et al.*, 2018). Despite the large volume of reports on the topic of epigenetics, there are limited studies on the clinical utility of these genetic factors for the prediction and prevention of T2DM.

The cohort of South Africans from a Bellville South community in this study came from a larger study group where the distribution of traditional risk factors did not differ fundamentally between diabetic and non-diabetic individuals (Erasmus *et al.*, 2012). Furthermore, additional studies showed that the increase in glucose tolerance deterioration over time in the same community could not be explained by the known determinants of diabetes such as obesity (Matsha *et al.*, 2013). This study, therefore, provided some insight into the epigenetic mechanism associated with T2DM within the Bellville South cohort.

The literature review aims to provide an overview of T2DM, its epidemiology and the environmental and genetic factors affecting this disorder. Lastly, the review will highlight the relationship between T2DM and the epigenetic mechanism of DNA methylation.

1.2 Diabetes mellitus

Diabetes mellitus is a complex multi-organ disease often difficult to predict or cure and can be characterised by features such as beta-cell dysfunction, hyperglycaemia, insufficient insulin secretion and insulin resistance (Arpón *et al.*, 2019). Diabetes is associated with numerous

complications including microvascular and macrovascular complications, increased susceptibility to infections and slow wound healing (Berbudi *et al.*, 2019). Furthermore, the global incidence, prevalence, mortality and disability of diabetes has increased over the years affecting both public health and social-economic development (Lin *et al.*, 2020). Diabetes is ranked as one of the top ten causes of death in the world and together with cardiovascular disease, cancer and respiratory disease, it accounts for 80% of death by non-communicable diseases (NCDs) (Lin *et al.*, 2020). Diabetes mellitus is broadly divided into various subtypes including type 1 diabetes mellitus (T1DM), type 2 diabetes mellitus (T2DM), gestational diabetes mellitus (GDM) and some monogenic diabetes such as maturity-onset diabetes of the young (MODY). There are also less common forms of diabetes which include genetic defects of insulin action, diseases of the pancreas and diabetes caused by drugs or infection (American Diabetes Association, 2019).

1.2.1 Type 1 diabetes mellitus (T1DM)

Globally, T1DM accounts for approximately 5-10% of diabetes cases and is often diagnosed in younger individuals (American Diabetes Association, 2020). It is also known as insulin-dependent diabetes mellitus and occurs due to T-cell mediated autoimmune destruction of the pancreatic beta-cells. This leads to a loss of functional beta-cells resulting in insulin insufficiency and hyperglycaemia (DiMeglio *et al*., 2018). Individuals with T1DM therefore require insulin therapy. Although genetic predisposition was thought to be the main cause of T1DM, it has more recently been suggested that an interplay between environmental factors and the microbiome, genome, metabolism and immune systems may play a role in T1DM (Rewers and Ludvigsson, 2016). Even though the condition mainly occurs in childhood or adolescence, onset in adulthood has been experienced by individuals and is referred to as latent autoimmune diabetes in adults (LADA) (Pozzilli and Pieralice, 2018). The autoimmune process in these older individuals appears to be milder with a slower progression of beta-cell failure. Since these individuals often do not require insulin therapy at diagnosis, they appear to be clinically similar to patients with T2DM (Carlsson, 2019).

1.2.2 Gestational diabetes mellitus

According to the International Diabetes Federation in 2019, 16% of pregnancies, which equates to approximately 20 million births worldwide have been affected by gestational diabetes mellitus (GDM) (International Diabetes Federation, 2019). Gestational diabetes mellitus is defined as impaired glucose tolerance which occurs during pregnancy in women without previously diagnosed diabetes and is a result of beta-cell dysfunction and insulin resistance during pregnancy (Plows *et al.*, 2018). The risk factors for GDM include obesity, advanced maternal age and a family history of diabetes. Furthermore, GDM may result in an increased risk for pregnancy outcomes such as macrosomia (high birth weight), preeclampsia and caesarean delivery (Szmuilowicz *et al*., 2019). Intrauterine exposure to hyperglycaemia is associated with consequences such as a predisposition to obesity, metabolic syndrome and diabetes later in life for both the mother and the foetus (Kang *et al.*, 2017). With increasing rates of obesity among women of child-bearing age, the global prevalence of GDM is expected to continue to increase which may further perpetuate the cycle of diabetes between mother and child (Zheng *et al*., 2018).

1.2.3 Monogenic diabetes

Monogenic diabetes such as neonatal diabetes mellitus and MODY which result from a single gene is less common and accounts for 1.5-2% of all diabetes cases (International Diabetes Federation, 2019; Standl *et al.*, 2019). Neonatal diabetes occurs under 6 months of age and may be transient or permanent with some cases requiring insulin therapy (Hattersley and Patel, 2017). Maturity-onset diabetes of the young (MODY) is characterised by hyperglycaemia onset before 25 years and is mostly due to autosomal dominant mutations which influence beta-cell dysfunction (Hattersley and Patel, 2017; American Diabetes Association, 2020).

1.2.4 Type 2 diabetes mellitus

Type 2 diabetes, the more common type of diabetes mellitus which accounts for 90% of diabetes cases globally (American Diabetes Association, 2020), has contributed to the burden of mortality and disability worldwide (Zheng *et al*., 2018). It is a chronic metabolic disease whereby individuals with T2DM are characterised by varying degrees of insulin deficiency in

conjunction with insulin resistance (Standl *et al.*, 2019). There is an inability of cells to respond to insulin resulting in decreased glucose uptake by skeletal and adipose tissue and increased glucose production by the liver resulting in hyperglycaemia (Ling and Rönn, 2019). Initially, pancreatic beta-cells compensate by the overproduction and secretion of insulin, however, this overproduction and high glucose concentrations may eventually lead to a decrease in beta-cell mass and function. When the beta-cells can no longer compensate with increased insulin secretion, it results in the manifestation of diabetes (Skyler *et al.*, 2017).

Chronic and inadequate glycaemic control in T2DM can lead to long term damage and dysfunction of organs such as the eyes, kidneys, heart, blood vessels and liver. These in turn can result in renal failure, cardiovascular disease, diabetic foot disorders as well as microvascular complications such as retinopathy or neuropathy (Pantalone *et al.*, 2015; Cole and Florez, 2020). Furthermore, the pathophysiology and complications of T2DM may be evident up to ten years before clinical diagnosis. The progression of T2DM is depicted in Figure 1.1, where factors such as increased insulin resistance and declining beta-cell function are evident and progressing before diagnosis. Also, the progression shows that microvascular and macrovascular complications of T2DM can be evident five to ten years before clinical diagnosis (Ali *et al.*, 2017a). Taking the progression of T2DM into consideration as well as the complications associated with T2DM, early detection and treatment interventions are of great importance in combating the disease.

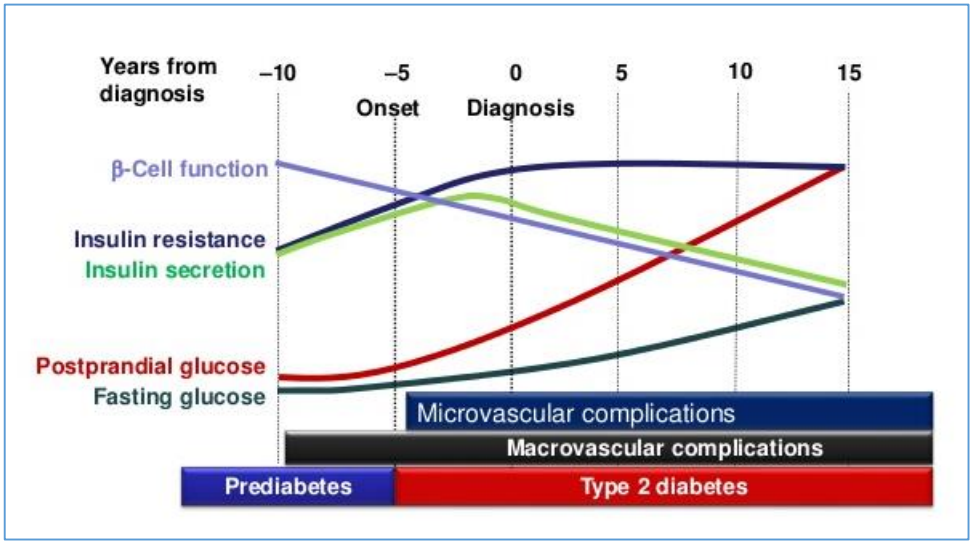


Figure 1.1: The progression of T2DM. The timelines indicate that the presence of complications and the effects of T2DM may be evident up to ten years before diagnosis (adapted from Ramlo-Halsted and Edelman, 1999)

Treatments for T2DM include modification of lifestyle such as dietary intervention, weight loss and physical activity as well as glucose-lowering drugs (Kolb and Martin, 2017; Zheng *et al*., 2018). These interventions aim to stabilise glucose levels and thereby reduce or prevent diabetes-induced complications (Galicia-Garcia *et al.*, 2020). Oral anti-diabetic medication includes among others, Biguanides, Thiazolidinediones, Sulfonylureas and Alpha-glucosidase inhibitors which aim to reduce blood glucose levels. This is achieved by either stimulating insulin secretion from the pancreas, increasing tissue sensitivity to insulin or by delaying glucose absorption from the gastrointestinal tract (Maruthur *et al.*, 2014; Marín-Peñalver *et al.*, 2016). If blood glucose levels cannot be controlled by oral medication alone, insulin therapy may be administered to provide sufficient control (Marín-Peñalver *et al.*, 2016). Lastly, education in the form of adequate information about T2DM and training on self-monitoring of blood glucose levels could also play a role in the management of the disease (Marín-Peñalver *et al.*, 2016).

1.2.5 Impaired glucose tolerance and impaired fasting glucose

Impaired glucose tolerance (IGT) or impaired fasting glucose (IFG), also referred to as prediabetes, characterises glucose levels that do not meet the criteria for diabetes but are higher than normal. Although not a clinical entity, prediabetes presents an increased risk for the development of T2DM as well as cardiovascular disease (Skyler *et al.*, 2017). Individuals who have prediabetes are at a 5-12 times higher risk for developing diabetes than the general population (Ali *et al.*, 2017a). Prediabetes is a reversible condition and if proper measures are taken during this phase, individuals can be spared the long-term complications of diabetes (Khan *et al.*, 2019).

When evaluating the different types of diabetes mellitus described, it is evident that T2DM is more prevalent. It affects more people globally in comparison to the other types which only account for 10% of cases globally. With its effects and complications manifesting years before diagnosis and the need for treatment interventions after diagnosis, T2DM is a prime target for research to expand on the approaches in addressing this global health concern. These reasons have provided the rationale for the current study's focus on T2DM.

1.3 The epidemiology of Type 2 diabetes mellitus

1.3.1 Global

Type 2 diabetes mellitus (T2DM) is a global health problem with the number of diabetic individuals increasing rapidly due to population growth, ageing, urbanization, increasing physical inactivity, and obesity (Guariguata *et al.*, 2014; Cho *et al.*, 2018). The International Diabetes Federation (IDF) estimated that worldwide in 2019, there were 463 million people with diabetes and that this number is expected to rise to 578 million in 2030 and 700 million in 2045 (Saeedi *et al.*, 2019). Furthermore, the IDF showed that the global diabetes prevalence for the age group 20-79 years was 9.3% in 2019 with an expected increase to 10.9% in 2045. This global prevalence showed an increasing trend with age where adults aged 20-24 years had the lowest prevalence (1.4%) in 2019 in comparison to adults aged 70-70 years at 19.9% who were predicted to be 20.5% in 2045. Also, the estimated prevalence of women (9.0%) aged 20-79 years was lower than men (9.6%) in 2019 with expected increases in both genders by 2045 (Saeedi *et al.*, 2019). Moreover, there are more people with diabetes living in urban areas with a global prevalence of 10.8% when compared to rural areas at 7.2% in 2019 (Saeedi *et al.*, 2019).

The prevalence of diabetes matches the socio-economic status of a country with developed and high-income regions showing higher prevalence rates (Khan *et al.*, 2020). The global prevalence of diabetes in adults aged 20-79 years according to World Bank income in 2019 was found to be 10.4% in high-income counties, 9.5% in middle-income counties and 4.0% in lower-income counties with an expected increase to 11.9%, 11.8% and 4.7% respectively in 2045 (Saeedi *et al.*, 2019). This increase could not only be attributed to socio-economic changes such as urbanisation but may also be affected by advancements in healthcare which improve the life expectancy of individuals with diabetes (Cho *et al.*, 2018). Cho *et al* (2018) further showed that people with diabetes in low- and middle-income counties were predominately found to be below the age of 65 years while those in higher-income countries were above the working age (Cho *et al.*, 2018). Although the prevalence of diabetes in low- and middle-income countries (LMICs) is less than that of high-income countries, these countries are projected to suffer the greatest increase in diabetes prevalence (Guariguata *et al.*, 2014). This increase can be attributed to the rapid development and changes in lifestyle in developing countries coupled with the slow development of health systems in terms of early detection and treatment of diabetes. Furthermore, more than half (50.1%) of the individuals worldwide living with diabetes are unaware of their condition. A high proportion of undiagnosed diabetes is found in LMICs with the African region showing the highest percentage worldwide (Forouhi and Wareham, 2019; Saeedi *et al.*, 2019).

1.3.2 Africa

According to the IDF, Africa has the highest percentage of people with diabetes (56.7%) who are undiagnosed and living unaware of their condition (International Diabetes Federation, 2019). Furthermore, diabetes prevalence is higher in urban areas (5.9%) than in rural areas (2.4%) with the highest prevalence (8.8%) of diabetes being amongst adults aged 65-69 years (Mapa-Tassou *et al.*, 2019). The increase in urbanisation, rapid population growth, as well as lifestyle changes, are suggested to be responsible for the expected increase of diabetes prevalence in Africa from 3.9% in 2019 to an expected 4.4% in 2045 (Goedecke *et al.*, 2017; Mapa-Tassou *et al.*, 2019). It has been further projected that while all regions of the world would experience an increase in diabetes prevalence, Sub-Saharan Africa (SSA) would experience the greatest increase between 2019 and 2045 (Ekoru *et al.*, 2019; Kibirige *et al.*, 2019). The five African countries with the highest number of people with diabetes include South Africa, Nigeria, the Democratic Republic of Congo, Ethiopia and the United Republic of Tanzania (see Table 1.1) with South Africa showing the highest number of people with diabetes (4.6 million) as well as the highest age-adjusted prevalence at 12.7% (International Diabetes Federation, 2019).

Table 1.1: The top five African counties for the number of people with diabetes in 2019

Country	Number of people with diabetes (20-79 years)	Diabetes prevalence age-adjusted (20-79 years)
South Africa	4.6 million	12.7
Nigeria	2.7 million	3.1
Democratic Republic of Congo	1.8 million	6
Ethiopia	1.7 million	4.3
United Republic of Tanzania	1.0 million	5.7

The availability of data on the prevalence of diabetes in SSA is however limited as reported by the IDF in 2017. The report showed that more than three-quarters of the African region's countries lacked primary data on diabetes prevalence in adults. Data obtained from the countries like South Africa, Ethiopia and the Democratic Republic of Congo with the highest

numbers were limited in the number of data sources and could not be used to accurately predict numbers for other African countries (Dessie *et al.*, 2020). Furthermore, Atun *et al* (2017) reported on the absence of reliable data on the prevalence, age of onset and progression of diabetes in SSA which has been attributed to limited sample sizes and differences in methodologies used to ascertain the diagnosis and biomarkers for diabetes (Atun *et al.*, 2017). Although there is evidence that diabetes numbers are increasing within the African region, data on the extent of this rise is still limited for a large number of SSA countries. Furthermore, it has been shown that the prevalence of undiagnosed diabetes is not consistent in the different African countries due to social, economic and genetic differences (Asmelash and Asmelash, 2019).

1.3.3 South Africa

South Africa has been deemed one of the LMICs where there is an alarming rise in the prevalence of diabetes and prediabetes with only a low percentage of individuals being aware of their diagnosis (Shen *et al.*, 2016; Mutyambizi *et al.*, 2019). In addition, South Africa has a quadruple disease burden which includes high rates of Human immunodeficiency virus (HIV)/Acquired Immunodeficiency Syndrome (AIDS) and tuberculosis (TB), poverty-related illnesses, non-communicable diseases as well as injury-related disorders. This has resulted in some challenges experienced by the health system with health outcomes worse than that of many lower-income countries (Coovadia *et al.*, 2009; Pillay-van Wyk *et al.*, 2016).

Although South Africa has been ranked as an upper-middle income country by the World Bank (2017), economic and health inequalities exist amongst populations due to the country's history of racial segregation (Pillay-van Wyk *et al.*, 2016). In South Africa, diabetes is increasingly common among disadvantaged or previously disadvantaged populations who were thought to be at lower risk of such conditions. With the improvement of the political and democratic condition in South Africa, the prevalence of diabetes has risen due to factors such as an ageing population, urbanisation, increased obesity and physical inactivity amongst these populations (Mutyambizi *et al.*, 2019). For instance, data from a South African study in 2012 showed that the prevalence of diabetes (28.2%) in a mixed-ancestry population had a twofold increase within a decade. Moreover, about two-thirds of those with the disease (18.1%) were not aware of their condition, and therefore were not receiving interventions to improve the adverse health consequences of diabetes (Erasmus *et al.*, 2012). Peer *et al* (2012) also reported similar trends in individuals of African descent where a substantial increase in

diabetes prevalence compared to that of two decades ago was observed. Among the study participants with diabetes, only 57.9% were aware of their condition and 38.6% were on treatment. Furthermore, the study also revealed an age-adjusted diabetes prevalence of 13.1% which was amongst the highest in SSA (Peer *et al.*, 2012). Similarly, a study by Hird *et al* (2016) reported a high age-adjusted prevalence (12.9%) of diabetes in an urban black South African population which was more than double the prevalence (5.3%) found in a study conducted 30 years earlier in the same region (Omar *et al.*, 1993; Hird *et al.*, 2016). Furthermore, the highest prevalence was observed in females over the age of 65 years (39.3%) and in males aged 55-64 years (29.0%) with 31,1% of the total participants being undiagnosed with diabetes(Hird *et al.*, 2016). The above-mentioned studies confirm the rise of the diabetes burden in South Africa, especially amongst the previously disadvantaged population groups.

1.4 The aetiology of Type 2 diabetes mellitus

The global rise in the number of individuals with T2DM has been linked to both genetic and environmental risk factors (Ronn and Ling, 2015; Zheng *et al*., 2018). These traditional risk factors however do not fully explain the rapid increase in T2DM in African populations. It has become evident that epigenetic mechanisms play a role in the gene-environmental interaction in T2DM and could provide insight into the pathogenesis of the disease (Ling and Groop, 2009; Matsha *et al.*, 2016b; Ling and Rönn, 2019).

1.4.1 Environmental factors

The environmental risk factors for T2DM include both modifiable and non-modifiable risk factors. Non-modifiable risk factors include factors such as family history or genetic predisposition, age and ethnicity (Boles *et al*., 2017; Misra and Misra, 2020). For years the genetic origin of diabetes has been linked to family history with the risk increasing depending on the number of relatives affected or whether it stemmed from the paternal or maternal side (Meigs *et al*., 2000; Lauenborg *et al.*, 2011; Prasad and Groop, 2015). Studies have also shown that ethnicity plays a role in the development of T2DM and this has been linked to various factors including the differences in socioeconomic status and access to healthcare in less developed countries (Shen *et al.*, 2016; Bavuma *et al.*, 2019). Certain ethnicities are more

prone to T2DM than others and migration to developed countries is also associated with higher risks for T2DM. Globally Asian-Indian immigrants have shown a higher prevalence of T2DM compared to Caucasian populations in the United States and Europe (Sattar and Gill, 2015). Similarly, studies in Tanzania and South Africa have shown that the frequency of T2DM is higher in migrant Asian-Indians when compared to the indigenous African population (Mbanya *et al.*, 2010). In addition, the increased risk amongst different ethnicities has also been linked to a familial component whereby increased genetic susceptibility to the disease is passed from one generation to the next (Unnikrishnan *et al.*, 2017; Galicia-Garcia *et al.*, 2020). The risk of diabetes has also shown to increase with age and studies have shown that higher rates of T2DM occur over the age of 45 (Werfalli *et al.*, 2016; Boles *et al*., 2017). Often this increased risk is augmented by slowing metabolism, physical inactivity, obesity and the employing of unhealthy lifestyle choices over the years (Shrivastava and Ghorpade, 2014). Since T2DM cannot be acted upon by non-modifiable risk factors, most studies have focussed on improving the modifiable risk factors (Deepa *et al*., 2017; Issaka *et al*., 2018).

Modifiable risk factors which play a role in the development of diabetes include factors such as sedentary lifestyle, physical inactivity, smoking, alcohol consumption, obesity, poor diet, stress and urbanisation, (Kengne *et al.*, 2013; Wu *et al.*, 2014; Zheng *et al*., 2018). Within an African context, these modifiable risk factors could be extended to broader determinants such as cultural, social, environmental and economic influences (Issaka *et al*., 2018). The extensive use of tobacco products and alcohol consumption has been associated with an increased risk of developing T2DM (Kolb and Martin, 2017). In African countries, these products have contributed to the increased risk of T2DM as alcohol plays a central role in certain cultural, traditional and social aspects and smokeless tobacco products are particularly popular due to being cheaper than cigarettes (Issaka *et al*., 2018). Urbanisation and sedentary lifestyles coupled with more westernised diets have been implicated as a contributor to T2DM and other cardiovascular diseases (Kengne *et al.*, 2013; Kolb and Martin, 2017). Furthermore, these factors increase weight, body mass index and visceral fat, further increasing the chances of obesity. Obesity influences metabolic abnormalities resulting in insulin resistance and it has been shown that approximately 90% of T2DM cases are related to excess body weight, therefore making obesity the strongest risk factor for T2DM (Wu *et al.*, 2014; Galicia-Garcia *et al.*, 2020). Although obesity, poor diet and lack of exercise increase the susceptibility to T2DM, many individuals displaying these risk factors, do not develop T2DM. It is therefore thought that T2DM results from an interaction between environmental and genetic factors (Matsha *et al.*, 2016a).

1.4.2 Genetic Factors

There have been various approaches in understanding the genetics of T2DM and these were based on family-based linkage analysis, candidate gene approach and genome-wide association studies (GWAS) (Park, 2011; Muka *et al.*, 2016). Although the heritability for T2DM has been estimated to be between 15 and 85%, genetic loci identified to date has only explained 5-10% of this heritability (Kwak and Park, 2016; Arpón *et al.*, 2019). Earlier family-based linkage studies used to identify T2DM susceptibility genes, identified loci associated with T2DM such as Calpain 1 (CAPN1) and Transcription Factor 7 Like 2 (TCF7L2) genes. These findings were however not consistent across the study population and the high-risk regions could not be reliably associated with T2DM (Park, 2011; Grant, 2019). The candidate gene approach investigated several candidate genes based on their known biological function. Although hundreds of candidate genes were investigated, only a few such as Peroxisome proliferator-activated receptor gamma (PPARG), Potassium Inwardly Rectifying Channel Subfamily J Member 11 (KCNJ11) and TCF7L2 showed to be associated with T2DM (Park, 2011; Grant, 2019; Khan *et al.*, 2019). More recently, genome-wide association studies have provided a breakthrough in identifying a genetic link to T2DM. It allowed for hundreds of thousands of single-nucleotide polymorphisms (SNPs) to be tested for association with T2DM in large scale analyses (Morris *et al.*, 2012; Flannick and Florez, 2016). Besides the identification of several novel gene loci, known T2DM genes such as PPARG, KCNJ11 and TCF7L2 were also confirmed by GWAS (Park, 2011; Prasad and Groop, 2015; Khan *et al.*, 2019). In 2015, Prasad and Groop indicated that GWAS provided approximately 153 variants for T2DM mapping to more than 120 loci as well as several loci for glucose and insulin-related traits (see Figure 1.2) (Prasad and Groop, 2015). The majority of known T2DM related genetic risk variants were found to be associated with insulin secretion and not insulin resistance (Carlsson, 2019). Chen *et al* (2019) conducted GWAS in 4347 Africans from South Africa, Nigeria, Ghana and Kenya to provide insight into the genetic architecture of T2DM in Africa as the 400 risk loci identified to date were from populations of European and Asian ancestry. The study identified a novel association signal at AGMO as well as replicated the widely known association at TCF7L2 (Chen *et al.*, 2019a).

Although GWAS has aided in identifying a large number of genes related to T2DM, their expression cannot specifically predict the risk of T2DM, indicating that there may be other factors involved in the development of T2DM. It has also been shown that genetically identical twins exposed to different environmental factors while growing up display different susceptibilities to T2DM (Tan *et al.*, 2013; Wu *et al.*, 2014). These findings indicate that genetic predisposition may not be the only factor contributing to T2DM. Furthermore, the increased

prevalence of T2DM over the last decades cannot be explained by genetics alone as it is unlikely that the human genome has changed during this short period (Jin *et al.*, 2019). Unlike single-gene disorders whose expression is affected by a mutation at one gene locus, the disease expression of T2DM is dependent on many gene loci with varying effects (Lyssenko and Laakso, 2013; Flannick and Florez, 2016). As T2DM is a multifactorial disease, the predisposition to the disease could be as a result of a combination of genetic variants and environmental factors which limits the use of these genes for the prediction and prevention of T2DM. This leads to the thought that epigenetic mechanisms may be a vital interface between genetic predisposition and environmental factors in the development of T2DM (Ronn and Ling, 2015; Jin *et al.*, 2019).

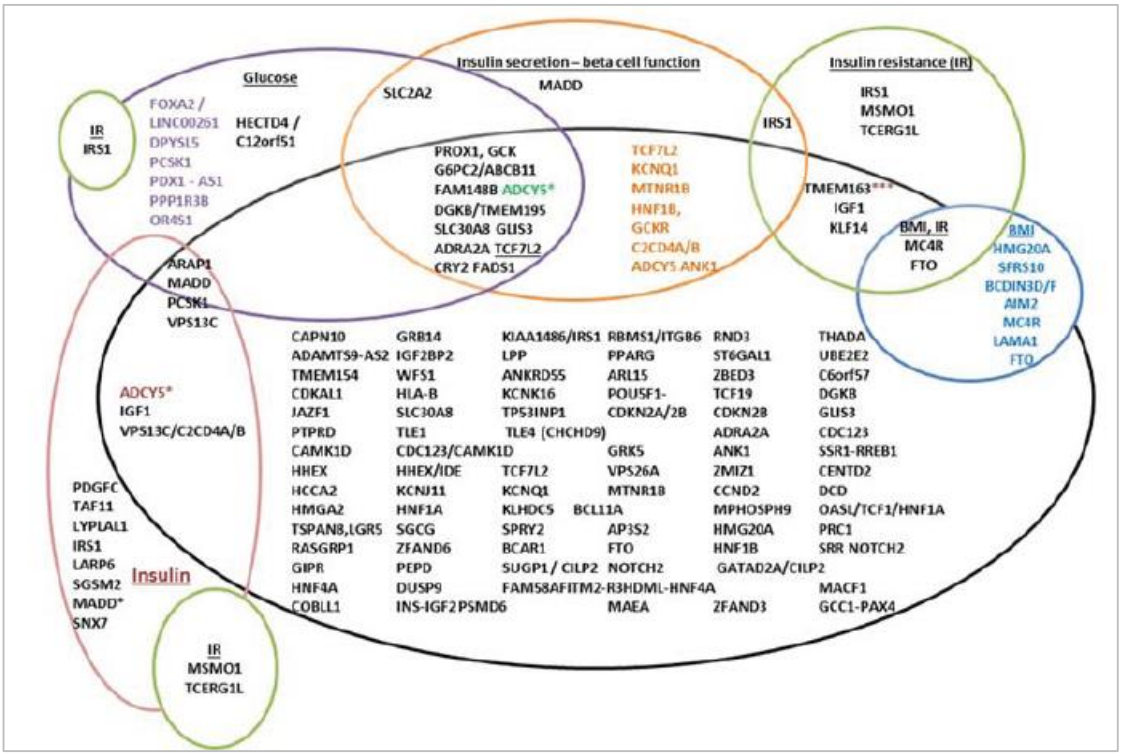


Figure 1.2: The T2DM and glycaemic trait-associated variants. Gene names are used to represent the variants that may either be located in the gene or the vicinity of the gene. The black circle represents the variants only associated with T2DM. The overlapping circles show additional associations for that variant. For example, TCF7L2, KCNQ1, MTNR1B, HNF18, GCKR, C2CD4A/B and ADCY5 ANK1 (in the brown circle) are associated with T2DM and with beta-cell dysfunction. In addition to T2DM, a variant of ADCY5 is associated with 2-hour insulin adjusted for 2-hour glucose and 2-hour glucose. Also, variants of TMEM163 are associated with T2DM and fasting insulin and TCFL2 is associated with fasting and 2-hour glucose (Prasad and Groop, 2015)

1.5 Epigenetics in diabetes mellitus

1.5.1 Epigenetics

Epigenetics is the study of heritable changes in gene expression or cellular phenotype without changing the primary nucleotide sequence. These changes can survive cell division and may be passed from one cell generation to the next (Liyanage *et al.*, 2014; Martínez *et al.*, 2014; Ling and Rönn, 2019). Essentially, gene expressions start with the transcription of DNA into RNA followed by the translation into proteins. Epigenetic changes can either increase or inhibit transcription and thereby affect the expression of genes (Arpón *et al.*, 2019). All cells in the body have the same genetic content but with individual or distinct epigenomes. Epigenomes are influenced by factors such as diet, chemical exposure and medication; and the accumulation of these environmental changes may affect individuals differently throughout the ageing process (Arpón *et al.*, 2019). This explains why identical twins who start with the same genome may experience differences in their susceptibility to disease (Nilsson *et al.*, 2014b).

Characteristic features of epigenetic change are that it is reversible and adaptable, meaning that it can be altered by environmental factors. Although epigenetic changes are part of normal development, they can lead to disease. This makes epigenetic markers a possible therapeutic target and diagnostic indicator for the risk and prognosis of disease (Kelly *et al*., 2012; Kwak and Park, 2016).

The major epigenetic modifications have been classified to include cytosine methylation of DNA (DNA methylation), histone post-translational modifications (PTMs), and non-coding RNAs (ncRNAs) (Reddy and Natarajan, 2015).

1.5.2 Histone modifications

Chromatin is made up of a complex of chromosomal DNA wrapped around core histones (proteins) and is considered to be the major site affected by epigenetic changes (Wegner *et al.*, 2014; Martire and Banaszynski, 2020). Although the core histones are densely packed (Figure 1.3), they can be modified by histone-modifying enzymes resulting in acetylation, methylation, phosphorylation, sumoylation or ubiquitination. The families of enzymes involved include histone deacetylases (HDACs), K-acetyltransferases (KATs), K-methyltransferases

(KMTs), and K-demethylases (KDMs) (Pons *et al.*, 2009; Dayeh *et al.*, 2014; Costantino *et al.*, 2019). These modifications determine the accessibility of the DNA for transcription and could result in the activation or suppression of special genes (Ling and Groop, 2009; Gilbert and Liu, 2012; Al-Haddad *et al.*, 2016). Post-translational modification of histones therefore plays an important epigenetic role in maintaining cellular transcriptional patterns (Sun *et al.*, 2017). New techniques have made it easier to analyse histone modifications on a genome-wide scale and these may be useful when studying the impact of epigenetics on the pathogenesis of T2DM (Reddy and Natarajan, 2015; Sun *et al.*, 2017; Zhang and Pollin, 2018). Recent studies have shown that histone modifications, especially acetylation and methylation modifications, seem to play a role in the pathobiology of Diabetic Kidney Disease (DKD). Histone modifications were shown to influence the progression of renal fibrosis of DKD through regulating the expression of extracellular matrix proteins (Sun *et al.*, 2017; Ling and Rönn, 2019).

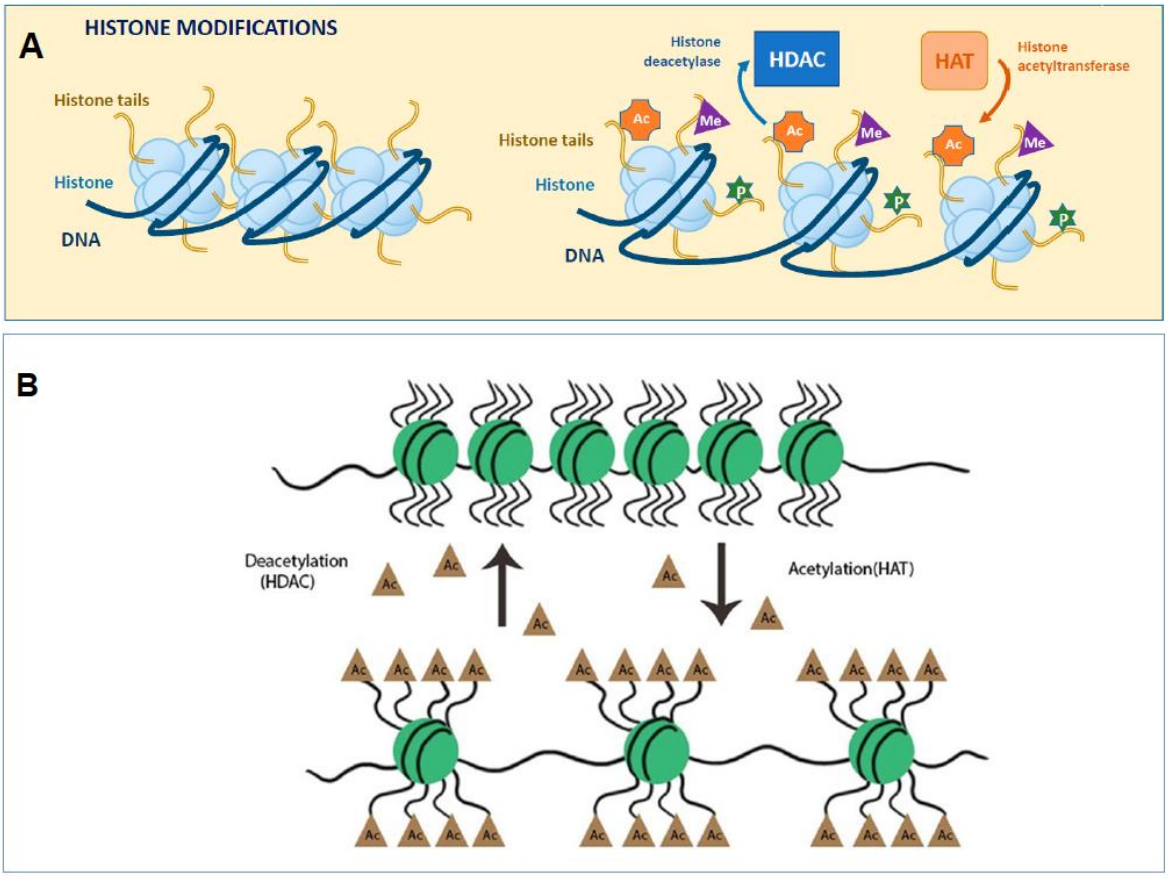


Figure 1.3: An overview of histone modifications. (A) Chromosomal DNA is packaged around histone proteins to form nucleosomes. Post-translational modifications of histone tails which include phosphorylation, ubiquitination, acetylation (Ac) and methylation are responsible for regulating the opening of nucleosomes and their accessibility to nuclear factors. (B) Acetylation reactions on lysine residues by histone acetyl-transferase (HAT) result in chromatin having a less condensed or more open chromatin formation which favours transcription. In contrast, deacetylation reactions by histone deacetylase (HDAC) result in a more compact or closed chromatin structure which may prevent DNA transcription (Adapted from Lakshmaiah et al., 2014; Coco et al., 2019)

1.5.3 Non-coding RNAs

Non-coding RNAs (ncRNAs) play a role in the post-translational regulation of transcription and are broadly classified according to their size as small ncRNAs (<200 nucleotides) or long ncRNAs (200bp) (Coco *et al.*, 2019). In terms of their function, ncRNAs can be grouped in constitutive housekeeping molecules, for example, ribosomal (rRNA) and transfer (tRNA) RNAs as well as regulatory molecules such as micro (miRNA) and long non-coding (lncRNA) RNAs. Of these molecules, miRNA are the most extensively studied (Mukherjee *et al*., 2015; Coco *et al.*, 2019).

1.5.3.1 MicroRNA

MicroRNA (miRNA) are small non-coding RNAs of approximately 20-25 nucleotides in length and are involved in several diverse biological functions including proliferation, differentiation, apoptosis and metabolic functioning (Gilbert and Liu, 2012; O 'Connell and Markunas, 2016). They have shown to play a significant role in gene regulation by acting as repressors as well as activators, mainly at the post-transcriptional level (Olivieri *et al.*, 2013; Witkowski *et al.*, 2018). MicroRNAs circulate in the bloodstream in a remarkably stable form and have been reported to be minimally invasive, reproducible and consistent amongst individuals as well as inexpensive biomarkers of complex processes like age-related diseases including T2DM and its complications (Olivieri *et al.*, 2013; O 'Connell and Markunas, 2016; Vasu *et al.*, 2019).

1.5.3.2 Long non-coding RNAs

Long non-coding RNAs (lncRNAs) are transcription products greater than 200 nucleotides with limited protein function (Huang *et al.*, 2020). They exhibit tissue-specific expression and have the ability to fold into complex secondary and tertiary structures which interact with various molecules such as DNA, RNA and proteins and participate in numerous regulatory networks (Rinn and Chang, 2012; Coco *et al.*, 2019). It is through these interactions that lncRNAs play a role in the regulation of gene expression at the epigenetic, transcriptional and post-transcriptional level (Zhang *et al.*, 2019). The functions of lncRNAs include acting as a signal, decoy, scaffold, guide, enhancer RNAs and short peptides (see Figure 1.4) (Fang and Fullwood, 2016). Signal lncRNAs function as a molecular signal to regulate transcription in

response to stimuli and their presence can serve as an indicator of transcriptional activity (Li *et al*., 2016a). Decoy lncRNAs limit the availability of regulatory factors like transcription factors, catalytic proteins and miRNA by binding and sequestering these factors and thereby limiting their function (Lin *et al.*, 2018). Scaffold lncRNAs provide platforms for the assembly of multiple-component complexes such as ribonucleoprotein (RNP) complexes which could either result in transcriptional repression or activation whereas guide lncRNAs play a role in the localisation of RPNs by directing them to specific target genes (Fang and Fullwood, 2016). Enhancer RNAs (eRNAs) influence the chromatin interactions and may tether the interacting proteins to enhancer regions (Malik and Feng, 2016). Lastly, lncRNAs can encode functional small peptides involved in cellular functions (Lin *et al.*, 2018).

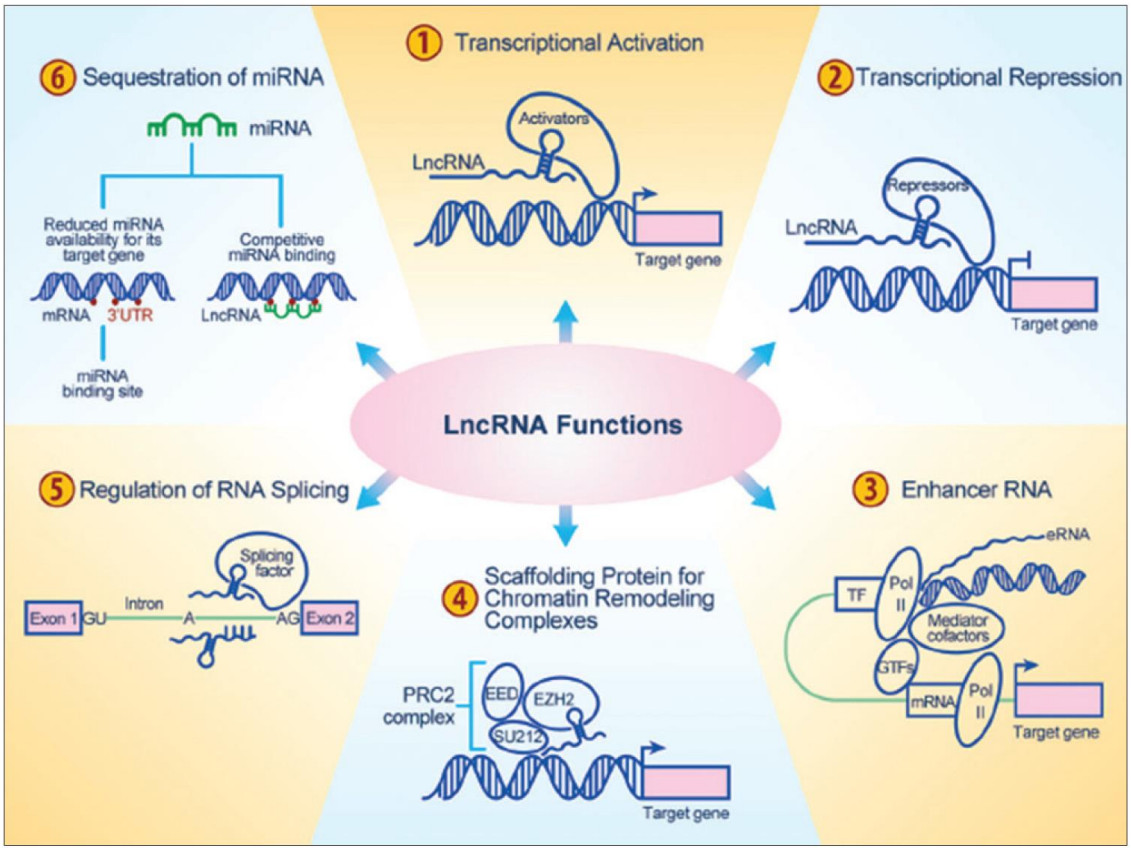


Figure 1.4: The function of long non-coding RNAs (lncRNAs). Long non-coding RNAs regulate gene expression in several ways including the following: (1) LncRNAs interact with transcriptional activators leading to target gene activation. (2) LncRNAs can mediate transcriptional repression by acting as a decoy to withhold transcriptional activators from chromatin. (3) Enhancer lncRNAs recruits lineage-specific complexes and regulates signalling. (4) LncRNAs may function as scaffolding proteins by recruiting chromatin remodelling complexes such as Polycomb Repressive Complexes 1 and 2 (PCR1 and PCR2). (5) LncRNAs may either interact with splicing factors or bind the splicing junctions of pre-mRNA and thereby regulate RNA splicing. (6) LncRNAs can act as molecular sponges by possessing the binding sites for miRNA and occupy their mRNA targets (Malik and Feng, 2016)

LncRNAs were initially viewed as ‘junk genes’ lacking coding capacity, however further studies revealed their importance in human diseases (Fang and Fullwood, 2016; Pengyu *et al.*, 2020). Due to their impact on biological and pathologic processes, lncRNAs have been linked to numerous diseases including cancers, neurodegeneration, autoimmune disease and metabolic disorders (Xu *et al.*, 2020). Although fairly limited, studies show that lncRNAs may play a role in the therapeutic and diagnostic management of diabetes due to the involvement of lncRNAs in the regulatory processes and complications of T2DM (Leti and DiStefano, 2017; Sathishkumar *et al.*, 2018). One study showed increased levels of lncRNAs in T2DM patients, including HOTAIR, MEG3, LET, MALAT1, MIAT, CDKN2BAS1/ANRIL, XIST, PANDA, GAS5, Linc-p21, ENST00000550337.1, PLUTO, NBR2THRIL, and SALRNA1. The majority of these lncRNAs were involved in cell cycle regulation and senescence with their expression levels correlating to poor glycaemic control, insulin resistance, and inflammation (Sathishkumar *et al.*, 2018). Another study demonstrated that H19, a lncRNA expressed abundantly in skeletal muscle but with decreased expression in insulin-resistant rodents and T2DM individuals, functions to control or modulate DNA methylation by inhibiting S-adenosylhomocysteine hydrolase (SAHH) (Geng *et al.*, 2018). Pengyu *et al* (2020) investigated the differentially expressed lncRNAs in T2DM patients when compared to non-diseased subjects and found 68 763 significantly up-regulated lncRNAs. In addition, gene ontology (GO) and KEGG pathway analyses revealed lncRNAs mainly involved in the phagocytic signalling pathway (Pengyu *et al.*, 2020). Chang *et al* (2020) showed that higher levels of the lncRNA MEG3 were evident in peripheral blood mononuclear cells of diabetic patients with vascular complications when compared to diabetic patients with no vascular complications and the healthy control group. Also, MEG3 was positively correlated with HbA1c indicating that abnormal expression of MEG3 may be related to high glucose levels. The study highlighted the use of MEG3 upregulation as a prognostic tool in T2DM patients with vascular complications (Chang *et al.*, 2020). Using a mouse cell model, Takahashi *et al* (2019) showed the role of lncRNAs by looking at the glucoregulatory effects of metformin on skeletal muscle cells. Metformin treatment reduced the expression of the lncRNA, Dreh, which in turn enabled glucose uptake via the increased expression of glucose 4 transporter (GLUT4) on the cell surface of skeletal muscle cells. The study identified Dreh as a new lncRNA with a role in glucose metabolism as well as a potential therapeutic target for diabetes (Takahashi *et al.*, 2019).

It has been suggested that the expression of lncRNAs vary in healthy and diabetic subjects and could be used as diagnostic biomarkers for T2DM (He *et al.*, 2017). This was evident in a study where the lncRNA, GAS5 was highlighted as a potential prognostic biomarker as its expression could be used to distinguish between diabetic and non-diabetic patient samples (Carter *et al.*, 2015). In addition to being easy to extract and detect with greater specificity in

comparison to proteins, lncRNAs expression is tissue and cell-specific in humans (He *et al.*, 2017). This makes the use of lncRNAs as potential T2DM biomarkers and targets for drug research promising. Furthermore, a large number of lncRNAs have not been well-characterised in terms of their role in the pathogenesis of T2DM as well as a therapeutic intervention and therefore, lncRNAs and their relationship to diabetes warrants further investigation.

1.5.4 DNA methylation

1.5.4.1 Mechanism of DNA methylation

DNA methylation is the biochemical reaction whereby a methyl group ($-CH_3$) is transferred from S-adenyl methionine (SAM) through DNA methyltransferases (DNMTs) to the 5' position of cytosine, forming 5-methylcytosine (5-mC) (Gilbert and Liu, 2012; Moore *et al*., 2012; Samblas *et al*., 2019). In mammals, the targets of DNA methylation are CpG dinucleotides (cytosine and guanine separated by phosphate) which can cluster together as repetitive sequences called CpG islands (CpGI), found in the gene promoter regions (Jones, 2012; Mahna *et al*., 2018). Sequences in the genome can be divided into CpG poor regions and CpG islands. The intergenic and the intronic regions of the genome is considered to be CpG poor. In humans, most of the CpG dinucleotides are methylated whereas the CpG islands are on average 1000bp long, usually unmethylated and associated with active gene expression (Deaton and Bird, 2011; Rao *et al.*, 2018). Once methylated, CpG island promoters are associated with gene repression (Liyanage *et al.*, 2014; Du *et al.*, 2015). Where hypermethylation of promoter CpG islands can result in suppression of gene expression, hypomethylation is associated with the transcriptional activation of affected genes (Pasquier *et al.*, 2015).

DNA methylation can be catalysed by three DNA methyltransferase enzymes (DNMTs) namely DNMT1, DNMT3a and DNMT3b. The DNMTs are essential for establishing DNA methylation patterns in early development as well as methylation maintenance. The *de novo* methyltransferases, DNMT3a and DNMT3b establish methylation in previously unmethylated cytosines, while DNMT1 maintains a state of methylation, by copying methylation patterns from parent to daughter strands (Bansal and Pinney, 2017; Bridgeman *et al.*, 2018). It has also been observed that 5-mC can be oxidised to 5-hydroxymethylcytosine (5-hmC) and further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by a family of dioxygenases named Ten-eleven translocation (TETs) (Tan and Shi, 2012; López *et al*., 2017). Initially, the conversion

from 5-mC to 5-hmC was considered the first step in the DNA methylation pathway, however more recent studies have highlighted 5-hmC as an epigenetic marker with its own distinct regulatory functions (Li *et al.*, 2017; López *et al*., 2017).

DNA methylation plays a role in several processes such as gene imprinting, embryonic development, genomic stability, X-chromosome gene silencing and regulation of gene expression (Miranda and Jones, 2007; Moore *et al*., 2012; De Mello *et al.*, 2014; Docherty *et al.*, 2014; Jin and Liu, 2018). In addition, DNA methylation has also been associated with transcriptional silencing (Jones, 2012; Ponnaluri *et al.*, 2017). For transcription to take place, the promoter regions of genes should be accessible to transcription factors. DNA methylation can however block the binding of transcription factors and thereby affect gene expression (see Figure 1.5) (Bansal and Pinney, 2017). Besides the direct inhibition of transcription factor binding, DNA methylation can also recruit methyl DNA binding domains (MBDs) proteins that specifically bind to methylated CpGs. This MBD family of proteins involved in transcriptional repression include MBD1, MBD2, MBD3, MBD4 and methyl-CpG-binding protein 2 (MeCP2) (Du *et al.*, 2015).

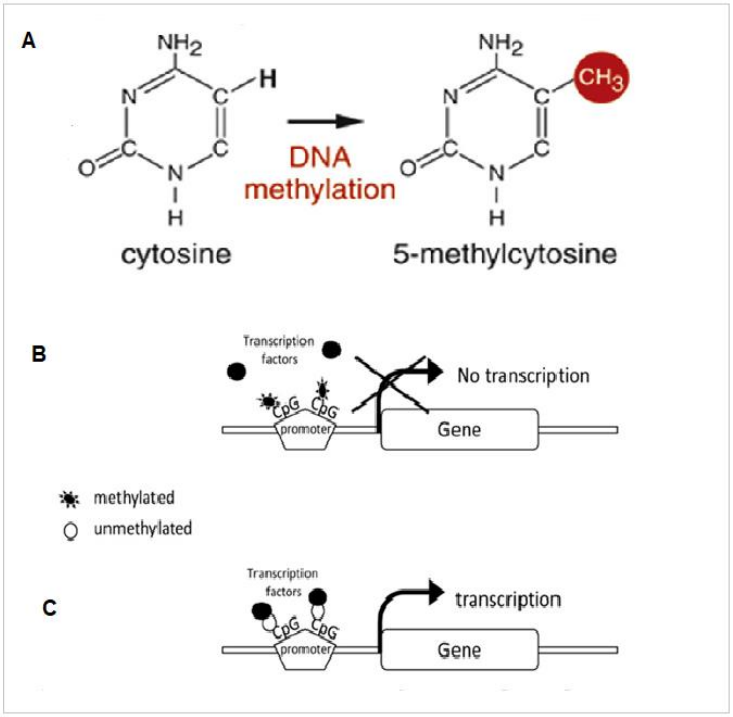


Figure 1.5: Methylated and unmethylated CpGs at gene promoter sites. (A) DNA methylation is facilitated by the adding of a methyl group to the carbon-5-position of cytosine with the aid of DNA methyltransferase (DMNT) enzymes in CpG islands. (B) Methylation of CpG islands may silence genes by preventing the binding of transcription factors to the promoter resulting in no transcription taking place. (C) Transcription may be enhanced by no or reduced methylation at the promoter regions of CpG islands, thereby allowing gene expression. (Mukherjee *et al*., 2015; Bansal and Pinney, 2017)

1.5.4.2 DNA methylation and disease

Regulated DNA methylation is needed for normal development and plays an important role in tissue-specific gene regulation and transcription throughout life. DNA methylation can, however, be affected by genetic, environmental and lifestyle factors and thereby impact the development of diseases (Breitling *et al.*, 2011; Rakyan *et al.*, 2012; Dhingra *et al.*, 2018; Liu *et al.*, 2018). Changes in DNA methylation have been associated with certain pathologic conditions such as cancer, cardiovascular disease, infections and autoimmune diseases (Estellar, 2008; Bierne *et al*., 2012; Dang *et al*., 2013; Dayeh *et al.*, 2014; Pasquier *et al.*, 2015). Aberrant DNA methylation patterns are apparent in most human cancers and it has been shown that DNA hypermethylation leads to transcriptional suppression whereas global hypomethylation in gene bodies, intergenic regions as well as repetitive sequences can lead to genomic instability and altered gene expression (Kelly *et al*., 2010; Pan *et al.*, 2018). Moreover, these changes in DNA methylation play a role in tumour development, the proliferation of cancer and metastasis as well as a potential biomarker for cancer diagnosis (Jin and Liu, 2018). Similarly, abnormal changes in DNA methylation affect the normal functioning of the immune system and play a role in autoimmune diseases. An example of such abnormal changes is the hypermethylation at gene regions, CCR6 and SPTBN1, in blood cells or synovial fibroblasts which are associated with rheumatoid arthritis (Ehrlich, 2019). Evidence has also suggested that DNA methylation is involved in the development and progression of diabetes (Dayeh *et al.*, 2014; Van Otterdijk *et al.*, 2017; Davegårdh *et al.*, 2018; Ahmed *et al.*, 2020). It has been proposed that the transcriptional activity of genes affected by epigenetic modifications may contribute to the T2DM phenotype, including the response to anti-diabetic agents and the occurrence of diabetes complications (Raciti *et al.*, 2015; Bansal and Pinney, 2017). Epigenetics may, therefore, explain how environmental factors contribute to T2DM at the genomic level as well as provide information on the variance in T2DM susceptibility amongst individuals (Prasad and Groop, 2015).

1.5.4.3 DNA methylation in T2DM

Several human studies have shown evidence for associations between DNA methylation and T2DM. For instance, Kuroda *et al* (2009) showed in mouse and human studies using pancreatic beta cells that CpG methylation affected beta-cell maturation and insulin gene expression. Demethylation of the mouse insulin 2 gene and the human insulin gene promoters were found in insulin-producing cells and DNA methylation resulted in the suppression of insulin reporter gene activity (Kuroda *et al.*, 2009). In a study by Ling *et al* (2008), a two-fold

increase in DNA methylation of PPARGC1A gene promoter was detected when comparing pancreatic islets of ten T2DM patients and nine control subjects showing that epigenetic factors affected PPARGC1A expression and insulin secretion (Ling *et al.*, 2008). Another study by Ribel-Madsen *et al* (2012) showed methylation changes in promoter regions of genes related to T2DM and included PPARGC1A in muscle and HNF4A in adipose tissue (Ribel-Madsen *et al.*, 2012). Yang *et al* (2012) showed that DNA methylation was increased in the PDX-1(pancreatic and duodenal homeobox 1) promoter and enhancer regions in islets from nine T2DM patients when compared with 55 non-diabetic individuals. The increased DNA methylation was associated with decreased PDX-1 expression. Additionally, the study showed that hyperglycaemia decreased gene expression and increased DNA methylation of PDX-1 as the glycated haemoglobin (HbA1C) correlated negatively with mRNA expression and positively with DNA methylation (Yang *et al.*, 2012). Dayeh and co-workers showed that regions closer to the transcription start site (TSS) displayed low levels of methylation whereas the regions further away from the TSS showed higher levels of methylation. They identified 853 differentially methylated regions including TCF7L2, FTO, KCNQ1 of which 102 including CDKN1A, PDE7B, SEPT9 and EXOC3L2, were differentially expressed in T2DM islets. Many of these differentially expressed genes are involved in the production and release of insulin (Dayeh *et al.*, 2014). Furthermore in a study reported by Volkmar and co-workers in 2012 who used the pancreatic islets of five individuals with T2DM, 276 differentially methylated regions (DMRs) were identified of which 96% located in the promoter region were hypomethylated. The DMRs were associated with beta-cell function, cell death and adaption to metabolic stress (Volkmar *et al.*, 2012). In 2013, Dayeh *et al* found that 19 of 40 T2DM associated SNPs introduced or removed CpG sites and all the CpG-SNPs were associated with differential DNA methylation of the CpG-SNP site. Moreover, CpG-SNPs of TCF7L2, KCNQ1, CDKN2A, ADCY5, WFS1 and HMGA2 were also associated with DNA methylation of surrounding CpG sites (Dayeh *et al.*, 2013). Hall *et al* studied DNA methylation in human pancreatic islets from 13 donors treated with palmitate in comparison to non-treated cells. The study showed that palmitate treatment affected the DNA methylation pattern in human islets and 290 genes, including TCF7L2 and GLIS3, showed altered gene expression and changes in DNA methylation levels which may contribute to impaired insulin secretion and T2DM (Hall *et al.*, 2014). Van Otterdijk *et al* (2017) studied the role of DNA methylation as a biomarker for T2DM and metabolic syndrome in peripheral blood leukocytes from 25 T2DM individuals when compared to 11 healthy control subjects. Amongst others, hypermethylation of the PPARG gene was observed which is known to affect insulin sensitivity as well as hypomethylation of the PDK4 gene which may affect the regulation of glucose metabolism and mitochondrial function (Van Otterdijk *et al.*, 2017). Ortiz *et al* (2018) observed that higher methylation of FKBP5 in 43 T2DM individuals was associated with risk factors for metabolic syndrome and cardiovascular disease. DNA methylation of FKBP5 was associated with higher levels of

HbA1c, LDL-cholesterol, BMI (body mass index) and waist circumference indicating that DNA methylation could be used as a marker of cardiovascular risk in T2DM (Ortiz *et al.*, 2018). Similar findings were observed by Willmer *et al* (2020) where hypermethylation of FKBP5 was observed in obese South African women when compared to those with normal weight. Hypermethylation of FKBP5 correlated with adiposity (BMI and waist circumference), insulin resistance and systemic inflammation (Willmer *et al.*, 2020). In another South African study, genome-wide DNA methylation was performed on three prediabetic, three diabetic and three normoglycaemic subjects and showed altered DNA methylation patterns in genes related to the immune system, signal transduction, glucose transport and pancreas development in both prediabetic and diabetic subjects (Matsha *et al.*, 2016a). Pheiffer *et al* (2016) conducted a study using peripheral blood DNA from the same South African cohort as Matsha *et al* (2016) and showed increased DNA methylation in intergenic regions when compared to gene bodies and promoter regions. Furthermore, 3 081 of the differentially methylated regions identified were associated with miRNAs, including miR-9, miR-34, miR-124 and miR1297 which have been linked to T2DM (Pheiffer *et al.*, 2016). Matsha *et al* (2016) also studied global DNA methylation in a South African mixed ancestry cohort and investigated the relationship with diabetes and polymorphisms in genes involved in DNA methylation and folate metabolism. They found global methylation to be higher in prediabetes and newly diagnosed diabetes when compared to the participants with normal glucose tolerance and in newly diagnosed diabetes when compared to diabetics on treatment. The increased global DNA methylation was associated with the NOS3 gene polymorphism G894T which indicated that vascular complications could develop in T2DM despite glycaemic control (Matsha *et al.*, 2016b). Although DNA methylation is the most extensively investigated epigenetic modification, genome-wide DNA methylation profiling in individuals with T2DM warrants further investigations.

1.5.4.4 DNA methylation and antidiabetic therapy

1.5.4.4.1 Metformin

Metformin, an oral anti-diabetic biguanide drug with minimal side effects, is considered the first line of treatment in the management of T2DM (Zhou *et al.*, 2018; Yendapally *et al.*, 2020). It can increase insulin sensitivity, reduce hepatic gluconeogenesis, maintain beta-cell function and enhance peripheral glucose uptake resulting in lowered blood glucose (Priya and Kalra, 2018). Metformin's action includes promoting the phosphorylation and activation of AMP-activated protein kinase (AMPK) which results in the inhibition of gluconeogenic genes (He and Wondisford, 2015). In addition to glucose metabolism, AMPK activation impacts other

pathways such as lipid metabolism, mitochondrial biogenesis, autophagy, cell growth and circadian rhythm (Bridgeman *et al.*, 2018).

Figure 1.6 depicts the various mechanisms identified for the action of metformin in hepatic gluconeogenesis and glucose production. Metformin is transported into hepatocytes through organic cation transporter 1 (OCT1) encoded by the SLC22A1 gene (García-Calzón *et al.*, 2017). Once inside the cells, it can indirectly activate AMPK through the inhibition of the mitochondrial respiratory-chain complex 1 resulting in the depletion of ATP and an increase in AMP levels (Foretz *et al*., 2019). The change in the AMP:ATP ratio causes the phosphorylation of AMPK by liver kinase B1 (LKB1) which activates AMPK. This activation in turn can inhibit gluconeogenic gene expression (He and Wondisford, 2015). Gluconeogenesis, which requires six ATP equivalents per molecule of glucose synthesised, is affected by the deficit of ATP. Furthermore, the elevated levels of AMP inhibit glucose production through its inhibition of the key gluconeogenic enzyme, fructose-1, 6-bisphosphatase (FBPase) (Foretz *et al.*, 2014). The elevated AMP levels also inhibit adenylate cyclase which decreases cyclic AMP (cAMP) synthesis and protein kinase A (PKA) activity. Gluconeogenesis is therefore suppressed by both the decrease in gluconeogenic enzyme activity and the inhibition of gluconeogenic gene expression (Foretz *et al*., 2019). The metformin-induced activation of AMPK also inhibits lipogenesis through inhibitory phosphorylation of acetyl-CoA carboxylase (ACC) which decreases the synthesis of malonyl-CoA and stimulates fatty acid oxidation. Over time, this AMPK activation decreases hepatic steatosis and leads to improved insulin resistance and hyperglycaemia (Li *et al.*, 2018a; Foretz *et al*., 2019). Metformin also suppresses gluconeogenesis through the direct inhibition of the mitochondrial glycerolphosphate dehydrogenase (mGPD) enzyme involved in the glycerol-phosphate shuttle. This inhibition disrupts the glucose production from glycerol and affects the cytosolic redox potential (NADH:NAD^+ ratio) which impairs the incorporation of lactate into glucose (Foretz *et al.*, 2014; Foretz *et al*., 2019).

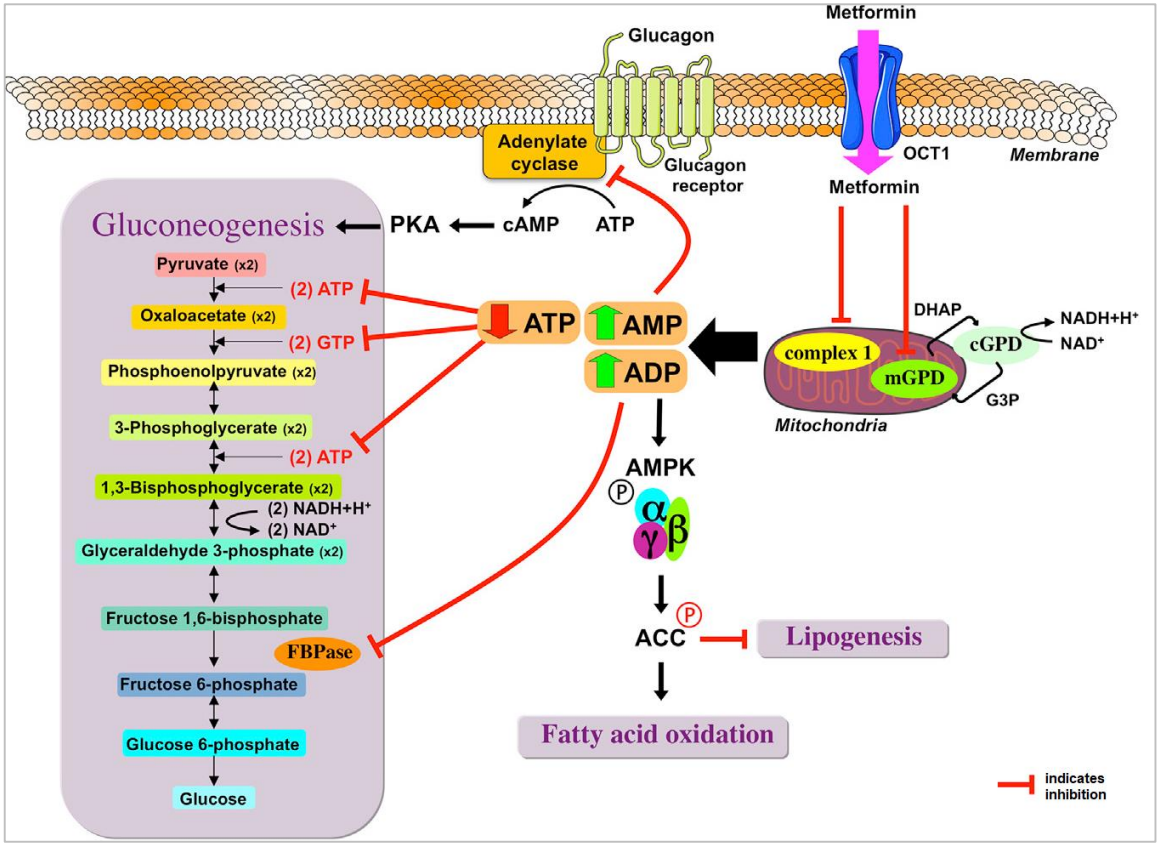


Figure 1.6: The actions of metformin. Metformin is transported to hepatic cells through organic transporter 1 (OCT1) and inhibits the mitochondrial respiratory chain complex 1. This results in decreased levels of ATP and an accumulation of AMP. The decrease in ATP results in a reduction of gluconeogenesis, thereby limiting glucose synthesis, while the increased levels of AMP inhibit the activity of the gluconeogenic enzyme, fructose-1, 6-bisphosphatase (FBPase) and affect the cyclic AMP-protein kinase A (cAMP-PKA) pathway by inhibiting adenylate cyclase activity. Elevated levels of AMP also activate AMP-activated protein kinase (AMPK) which inhibit lipogenesis and increase fatty acid oxidation through the phosphorylation of acetyl-CoA carboxylase (ACC). Metformin also inhibits the mitochondrial glycerophosphate dehydrogenase (mGPD) thereby contributing to the reduced conversion of glycerol to glucose (Adapted from Foretz et al., 2014; Li et al., 2018a)

1.5.4.4.2 DNA methylation and metformin

Metformin is considered as the gold standard for the treatment of T2DM and has shown to affect epigenetic processes through its activation of AMP-activated protein kinase (AMPK) (García-Calzón *et al.*, 2017; Bridgeman *et al.*, 2018; Elbere *et al.*, 2018). Upon activation, AMPK phosphorylates epigenetic enzymes such as DNA methyltransferases (DNMTs) resulting in their inhibition and thereby affecting epigenetic processes. Alterations in DNA methylation due to metformin include both hypomethylation and hypermethylation at the promoters of different genes (Ishikawa *et al.*, 2015; García-Calzón *et al.*, 2017; Zhong *et al.*, 2017). A decrease in DNA methylation at the insulin gene (Ins1) promoter in a beta-cell line, cultured with high glucose concentrations was observed following metformin treatment in a

study by Ishikawa *et al* (2015). Their study showed an up-regulation of the insulin gene with metformin (Ishikawa *et al.*, 2015). Garcia-Calzón *et al* (2017) observed decreased DNA methylation of the metformin transporter genes, SLC22A1, SLC22A3, and SLC47A1 in the human liver when comparing T2DM patients on metformin to untreated subjects (García-Calzón *et al.*, 2017).

In addition to its anti-diabetic properties, it has been suggested that metformin may induce alterations in DNA methylation of cancer cells thereby displaying anti-cancer properties (Zhong *et al.*, 2017; Cuyàs *et al.*, 2018). Banerjee *et al* (2016) observed reduced DNA methylation and increase expression of the tumour suppressor gene E-cadherin in both the cancer cell lines and white blood cells from diabetic patients on metformin (Banerjee *et al.*, 2016). Furthermore, studies have reported that cancer cells exposed to metformin have led to hypermethylation of tumour promoting pathway genes and cell proliferation inhibition due to the restraining effect of metformin on S-adenosylhomocysteine hydrolase (SAHH) activity (Zhong *et al.*, 2017). Metformin has also been shown to decrease the regeneration ability and tumorigenicity of osteosarcoma stem cells by directly acting on mitochondria and inducing reactive oxygen species-mediated apoptosis and autophagy (Zhao *et al.*, 2019). These studies highlight the role of DNA methylation in the anti-diabetic and anti-cancer actions of metformin.

1.5.4.5 Factors affecting DNA methylation

Changes in DNA methylation can occur in response to numerous factors including biological, lifestyle and environmental factors. Ageing has been associated with changes in DNA methylation and a decrease in DNA methylation with age as well as a loss of the epigenetic state in CpG sites have been observed (Horvath, 2013; Dugué *et al.*, 2018). Alterations in DNA methylation have been linked to environmental exposures that occur in the prenatal stage (Martin and Fry, 2016). A study by Heijmans *et al* was the first study to show that prenatal exposure to famine during the Dutch Hunger Winter had an impact on insulin growth factor 2 (IGF2) methylation which was associated with lowered birth weight, predisposition to obesity and adverse metabolic health outcomes later in life (Heijmans *et al.*, 2008). Chemicals have also been observed to affect methylation and the alterations to DNA methylation have been associated with lung conditions, cardiovascular disease, diabetes, cancers and immune deficiencies amongst others (Argos *et al.*, 2015; Jung *et al.*, 2017; Jin and Liu, 2018; Martin and Fry, 2018). In addition, urban development and traffic-related air pollution have been linked to alterations in DNA methylation levels of the mitogen-activated

protein kinase (MAPK) and adult exposures to the constituents of air pollution have been associated with global hypomethylation (Carmona *et al.*, 2014; Plusquin *et al.*, 2017). Tobacco smoke, a known carcinogen, has been associated with global hypomethylation with both in utero and adult exposures. The methylation targets of tobacco smoke include cancer, cell growth and metabolism-related genes (Rotroff *et al.*, 2016; Tsai *et al.*, 2018; Bakulski *et al.*, 2019).

Alterations in DNA methylation can also occur due to nutritional factors by either the modification of DNA methyltransferase (DNMT) enzymes or by the availability or lack of methyl donors needed for S-adenosyl-methionine (SAM) synthesis (Elgendy *et al.*, 2018). Supplementation with methyl donors such as methionine, folate, betaine and choline appears to increase global methylation levels and their deficiency is associated with global hypomethylation (Martin and Fry, 2018). Maternal dietary deficiencies have also been shown to affect DNA methylation patterns in the foetus resulting in increased risk of cardiovascular diseases, obesity and insulin resistance in the offspring (McKay and Mathers, 2016). Furthermore, paternal diet, lifestyle, obesity and diabetes may also result in DNA methylation alterations in the offspring which could lead to metabolic consequences (Samblas *et al*., 2019). In contrast, over-nutrition associated with urbanisation and the adoption of higher fat diets may alter DNA methylation patterns. An example is the supplementation of saturated fatty acids like palmitic acid induces hypermethylation in human pancreatic islets thereby impairing insulin secretion and increasing the risk of type 2 diabetes (González-Becerra *et al.*, 2019). Excess dietary trans fatty acids induce TNF hypomethylation and PPARG1 hypermethylation and the increase and decrease in expression of these genes lead to an inflammatory environment in adipose tissue (Flores-Sierra *et al.*, 2016).

Alcohol consumption influences DNA methylation which in turn can affect gene expression and influence the risk of chronic diseases (Wilson *et al.*, 2019). SAM levels are decreased by chronic alcohol consumption and result in DNA hypomethylation. In addition, the metabolism of ethanol promotes the formation of reactive oxygen species and acetate which impact regulatory mechanisms of epigenetics (Zakhari, 2012; Mahna *et al*., 2018).

1.5.4.6 Methods for detecting DNA methylation

Several methods have been described for DNA methylation determination. There are, however, a variety of factors that influence the choice of method for DNA methylation analysis and these include the potential for bias, cost, reproducibility, sensitivity and specificity, amount and quality of DNA available, availability of equipment and bioinformatics software (Kurdyukov and Bullock, 2016). Previous methods used for DNA methylation determination were based on mass spectrophotometry (MS), high-performance liquid (HPLC), luminometric methylation assay (LUMA) and LINE-1/pyrosequencing which has the highest specificity and sensitivity with minimal variability for measuring global DNA methylation (Kurdyukov and Bullock, 2016). Recent advancements have however made DNA methylation more reliable and reproducible and the experimental techniques used to differentiate between methylated and unmethylated DNA can be divided into three groups which include restriction enzyme-based, affinity enrichment-based and bisulphite conversion-based methods (Rauluseviciute *et al*., 2019). For this thesis, the DNA methylation techniques addressed include bisulphite conversion, pyrosequencing and methylated DNA immunoprecipitation and array-based technologies.

1.5.4.6.1 Bisulphite conversion

Although several methods have been developed over the years to study DNA methylation, bisulphite conversion has been considered the 'gold standard' in the identification of DNA methylation at single-nucleotide resolution (Olova *et al.*, 2018). In this process, DNA is treated with sodium bisulphite leading to the conversion of unmethylated cytosine to uracil which is then displayed as thymine in the PCR amplification and subsequent sequencing. The methylated cytosine remains unchanged during bisulphite conversion and can therefore be distinguished from unmethylated cytosine (Gupta *et al*., 2010; Raine *et al*., 2018). Numerous methodologies using bisulphite converted DNA has been developed due to the advantages of bisulphite conversion which include the allowance of quantitative DNA methylation analyses anywhere in the genome, single CpG resolution and detection of strand-specific methylation (Olkhov-Mitsel and Bapat, 2012; Wreczycka *et al.*, 2017). Furthermore, bisulphite conversion has more recently been used in combination with next-generation sequencing (NGS) for the study of DNA methylation (Raine *et al*., 2018).

Challenges are also presented when using bisulphite conversion such as the reduction of genome complexity to three nucleotides which affects the processing of bisulphite sequencing

data (Rauluseviciute *et al*., 2019). Also, high concentrations of bisulphite and long incubation periods can lead to DNA fragmentation and degradation of up to 90% of the incubated DNA. This may result in biased sequence representation and affect the estimation of methylated cytosine levels (Kurdyukov and Bullock, 2016; Olova *et al.*, 2018). Furthermore, false-positive results may also occur as a result of incomplete conversions during bisulphite treatment when unconverted unmethylated cytosines are interpreted as being methylated (Wreczycka *et al.*, 2017). Bisulphite conversion is also unable to discriminate between 5-methylcytosine and 5-hydroxymethylcytosine (Yong *et al*., 2016). Moreover, extensive bioinformatics for base calling, sequence alignment and statistical analysis is needed by methods that rely on bisulphite sequencing (Olkhov-Mitsel and Bapat, 2012).

1.5.4.6.2 Pyrosequencing

Pyrosequencing, a next-generation DNA sequencing technology, has been considered a gold standard for the identification of allele-specific methylation patterns and is suitable for both CpG rich and CpG poor regions (Frommer *et al.*, 1992; Reed *et al.*, 2010; De Chiara *et al.*, 2020). It uses a platform that can interrogate many CpG sites within an amplicon in real-time and can be conducted to identify methylated and unmethylated cytosines following bisulphite conversion in a sequencing-by-synthesis process (Kong, 2014; Delaney *et al*., 2015). The availability of commercial bisulphite conversion kits and PCR amplification makes the use of pyrosequencing cost-effective and more accessible in comparison to next-generation sequencing and the more labour intensive Sanger sequencing. Moreover, sequences from different samples can be identified in the same run which increases efficiency and throughput while decreasing costs (Siqueira *et al*., 2012). Furthermore, studies comparing various methylation techniques found pyrosequencing together with bisulphite amplicon sequencing to be the best method for methylation marker validation and development (Bock *et al.*, 2016; Šestáková *et al*., 2019). A disadvantage, however, is that the method is more time-consuming as it involves a three-step process of PCR amplification and tagging using a biotinylated primer; isolation of the PCR product with streptavidin beads and hybridization with a sequencing primer, and sequencing (Delaney *et al*., 2015; Šestáková *et al*., 2019). Furthermore, pyrosequencing only allows for the analysis of shorter regions (maximum of 350bp) as longer amplicons could result in the formation of secondary structures and loops which could impede the sequencing process (Fakruddin and Chowdhury, 2012). Another challenge in pyrosequencing is the detection of long homopolymers, which is a string of more than three to four repeated nucleotide in the run (Ivády *et al.*, 2018). These homopolymer regions may cause sequencing errors as they influence the synchronised synthesis of the DNA

strand resulting in irregular sequence peak heights which affect the read length (Balzer *et al*., 2011; Heather and Chain, 2016). The design of primers for PCR and sequencing is important for pyrosequencing as sequencing and base-calling complications can arise if self-looping, primer-primer hybridisation (primer dimers) or cross-hybridisation where more than one sequencing primer is used, occurs (Fakruddin *et al.*, 2013). Pyrosequencing assays often require extensive PCR optimisation to gain sufficient PCR product. This amplified PCR product together with a negative PCR control should always be checked by an agarose gel electrophoresis to prevent complications in pyrosequencing (Šestáková *et al*., 2019).

1.5.4.6.3 Methylated DNA immunoprecipitation sequencing (MeDIP-seq)

Methylated DNA immunoprecipitation sequencing (MeDIP-seq) is an enrichment-based method that can be used to measure DNA methylation at a genome-wide level (Xing *et al.*, 2018). MeDIP-seq uses a monoclonal antibody against 5-methylcytosine to enrich methylated DNA, after which the immunoprecipitated DNA can be sequenced (Chen *et al*., 2016). This technique is beneficial as it can distinguish between 5-methylcytosine and 5-hydroxymethylcytosine, an oxidation product of 5-methylcytosine (Willmer *et al.*, 2018). In addition, MeDIP is impartial to a specific nucleotide sequence other than CpGs (Xing *et al.*, 2018). Moreover, MeDIP-seq can detect up to 70% of all CpG dinucleotides in the human genome which makes it suitable for exploring the 60-68% methylated CpGs in the human genome. In addition, the assay is more sensitive in high CpG density regions as DNA sequences with more adjacent methylated CpG sites are more effectively captured (Soozangar *et al.*, 2018). Another advantage of MeDIP-seq is that it can be used to profile DNA methylation in small DNA samples, rare cell types and micro-dissected tissues as the assay is possible with low amounts of starting DNA (Yong *et al*., 2016). Although it is a cost-effective and useful technique that can bypass the need for bisulphite conversion, it displays a bias toward highly methylated regions. MeDIP CpG-rich fragments are more likely to be enriched than CpG-poor fragments leading to PCR bias and inaccurate estimation of methylation which then warrants the use of additional computational correction to normalise CpG content across the various densities (Soto *et al.*, 2016; Yong *et al*., 2016).

Although MeDIP-seq has its limitations, these are outweighed by the advantages of using this technique which include the relatively straightforward methodology with the generation of data that is relatively easy to analyse and interpret. With its specificity for 5-methylcytosine, the assay is useful for genome-wide analysis of DNA methylation (Soozangar *et al.*, 2018).

1.5.4.6.4 Array-based technology

Methylation arrays were initially developed to observe methylation at a single-base resolution and currently the most widely used assay is the Infinium HumanMethylation450 BeadChip developed by Illumina (Soto *et al.*, 2016). This array involves the bisulphite conversion of genomic DNA and amplification, followed by the hybridization of the converted DNA to arrays containing predesigned probes to distinguish between methylated and unmethylated cytosines (Yong *et al*., 2016; Willmer *et al.*, 2018). With the probes, this array can assess more than 485 000 CpG sites across the whole genome and detect the fluorescent intensity of the methylated or unmethylated status for each CpG site (Chen *et al*., 2016). The main advantage of using array-based technologies is the ease with which experiments can be performed, even for those with limited experience. Also, data generated from these array-based technologies are easy to interpret with the use of software programs that do not require extensive computational skills (Leti *et al.*, 2018). Array-based technologies also offer high sample throughput and provision of sensitivity and specificity that cannot be achieved by sequencing methods at a similar cost (Rauluseviciute *et al*., 2019). Limitations of array-based technologies however include the requirement for high-quality input DNA and possible DNA degradation by bisulphite treatment. In addition, DNA methylation is often found in repetitive sequences and the designing of probes that can distinguish between repetitive elements is not easy (Gupta *et al*., 2010; Soozangar *et al.*, 2018).

In terms of genome-wide methylation, array techniques are methylation-state dependant whereas sequencing methods provide the possibility of exploring methylation patterns beyond the single-site methylation. For this reason, sequencing supersedes arrays as the method of choice for methylation profiling even though data are more complicated to analyse (Soto *et al.*, 2016; Leti *et al.*, 2018).

1.6 Significance and local relevance of the study

An overview of T2DM has been presented by describing the various types of diabetes mellitus as well as highlighting the global, African and South African burden of the disease. The chapter

also addressed the epigenetic mechanisms associated with T2DM with a focus on DNA methylation, which is the more commonly studied mechanism.

Type 2 diabetes mellitus is a complex and multifactorial metabolic disease for which the clinical significance in an African setting is still poorly understood. This could be attributed to the fact that most of the studies looking at T2DM risk factors have been conducted within western countries and therefore it is unclear whether current knowledge applies to Africa. As T2DM has shown to be a rising public health crisis in Africa and particularly in South Africa, there is a need for intervention strategies that include early detection of at-risk individuals and the prevention of disease progression.

Many studies have recognised a positive correlation between the occurrence of T2DM and the interaction between genetic and environmental factors such as dietary habits and lifestyle. Recent evidence has suggested the involvement of epigenetic mechanisms in the link between the effects of genetic predisposition and environmental factors. Due to the increase in the prevalence of T2DM within a South African population and a large number of undiagnosed cases, epigenetic studies present opportunities to improve the diagnosis and treatment management of the disease. Furthermore, the South African population has also been reported to have high rates of obesity and metabolic syndrome, and a high risk of developing cardiovascular diseases, particularly amongst individuals of mixed ancestral descent (Erasmus *et al.*, 2012; Matsha *et al.*, 2012). Follow up studies within South Africa communities have shown that the deterioration of glucose tolerance status overtime is not explained by the known determinants of T2DM occurrence in these individuals (Matsha *et al.*, 2013). Therefore the identification of individuals who are at risk of developing T2DM could facilitate risk stratification and prevention of T2DM as well as facilitate improved classification of the subtypes of diabetes and reduce the burden of this disease.

1.7 Hypothesis

We hypothesize that epigenetics plays a major role in the genesis of diabetes in South Africans. We further hypothesize that altered epigenetic patterns are established early in the development of diabetes and may play a role as predictive markers.

1.8 Aims and objectives

1. Conduct the genome-wide DNA methylation profile in 48 mixed ancestry individuals within a South African population with varying degrees of glucose tolerance.
2. Validate significant differential methylated regions (DMRs) using pyrosequencing in the same 48 individuals.
3. Investigate the relationship between the DMRs observed and cardiometabolic risk factors.
4. Investigate the relationship between the lncRNAs observed and diabetes.

CHAPTER 2

METHODOLOGY

2.1 Ethical considerations and confidentiality

This study was part of the Vascular and Metabolic Health (VMH) study that has been registered with the research ethics committees of the Cape Peninsula University of Technology (CPUT) Human Research Ethics Committee (Ref #: NHREC: REC - 230 408 – 014, renewed for 2017: CPUT/HW-REC 2015/H01 (renewal) and Stellenbosch University (Ref #: N14/01/003; approved on 21 May 2018). Specific ethical approval for this sub-project was sought from CPUT (CPUT/HW-REC 2017/H29) to extend the use of the samples obtained in the main study. The Code of Ethics of the World Medical Association (Declaration of Helsinki) was applied to the study. There was no additional recruitment or contact or participant identifier information beyond what has been ethically mandated in the main project. All participants gave informed written consent to participate in the study after the procedures were explained in the language of their choice. They were also informed that they had the right to withdraw from the study at any stage. Participants were allocated coded identifiers to keep personal details confidential. All data gathered was kept confidential and storage was on a password protected computer locked in the office of the project manager.

2.2 Study design and settings

This was a cross-sectional case-control study.

The research setting is defined as Ward 009 by the City of Cape Town and it includes the areas of Bellville South, Bellville South Industrial, Cape Peninsula University of Technology (CPUT), Glenhaven, Greenlands, Sack's Circle Industrial and Vogelvlei. The township was formed in the late 1950s and comprised largely of individuals of mixed ancestry. According to the 2011 population census, its population stood at approximately 24 642 with an average household of 4.84 individuals. The population is mainly individuals of mixed ancestry followed by those of African descent and with individuals of Caucasian and Asian ancestry making up a small percentage (City of Cape Town Census, 2013). The socio-economic condition of the

individuals in the community is average with 37% of households having a monthly income of R3 200 or less. Data from the Bellville South community has indicated an increased number of inhabitants with diabetes. The target population for the VMH study included subjects aged 20 or older, both male and female, which comprises 16 168 individuals of which 14 352 were mixed ancestry individuals.

2.3 Inclusion criteria

All consenting adults, male and female aged 20 years or older who resided in the Bellville South community were included in the VMH study. For this particular study, the participants were all female and matched for both age and body mass index.

2.4 Exclusion criteria

All non-consenting participants as well as pregnant or breastfeeding women and bedridden patients were excluded from the study. In addition, any individuals with conditions prohibiting any of the study investigations such as blood sample collection or ambulatory blood pressure monitoring were also excluded from the study.

2.5 Study sample and justification

There is still uncertainty as to the estimation of statistical power and sample size for epigenome-wide association studies (EWAS) (Wang, 2011; Rakyan *et al.*, 2012; Tsai and Bell, 2015). Simulation studies by Tsai and Bell (2015) showed that a case-control EWAS design provided an 80% power to detect a mean DNA methylation difference of 1%, assuming a p-value threshold of 0.05 for single-locus analysis, and $1x10^6$ for genome-wide significance (Tsai and Bell, 2015). It was further suggested by Chambers *et al* (2015) in a nested case-control study among Asian-Indians and Europeans that the mean difference in DNA methylation between people with type 2 diabetes and those without ranged between 0.05 and 0.11

(Chambers *et al.*, 2015). Based on the above specifications, we initially aimed to select a minimum of 200 subjects each of normal glucose tolerance, prediabetic, screen-detected or newly diagnosed and treated diabetic participants from the existing VMH sample of about 2000 individuals. Due to the budget available for this study, these targets were not possible. For genome-wide DNA methylation a convenient sample of 48 subjects was selected, 12 for each of the following glucose tolerance categories; 12 known diabetics on metformin treatment, 12 screen-detected or newly diagnosed diabetics, 12 prediabetics and 12 subjects with normal glucose tolerance (NGT). Based on previous findings by Matsha *et al*, where increased global DNA methylation was observed in screen-detected diabetes compared to those on treatment the inclusion of diabetics on treatment was important in explaining those epigenetic changes that may be due to treatment (Matsha *et al*., 2016b).

2.6 Study procedure

2.6.1 Questionnaire and data collection

Trained personnel presented and explained a questionnaire to each participant to obtain information relevant to the study. Information requested included gender, age, previous medical history of diabetes, any family history of CVD and diabetes mellitus as well as lifestyle factors such as smoking, diet, alcohol consumption and physical activity. A detailed drug history was also obtained either by interrogation or by examining the clinic cards and record of drugs brought by participants to the study site.

2.6.2 Blood pressure measurements

Blood pressure (BP) measurements were performed by using an automatic blood pressure monitor (Omron M6 Comfort-performed Cuff Blood Pressure Monitor) on the arm of the participant while at rest in a sitting position. The correct adult cuff size for each participant was selected according to the circumference of the individual's arm and placed 2cm above the elbow joint to ensure accurate readings. The systolic (SBP) and diastolic (DBP) readings (mmHg) were measured in triplicate at one-minute intervals according to the WHO guidelines (World Health Organisation, 1999). All readings were recorded, with the lowest reading chosen as the participant's blood pressure.

2.6.3 Anthropometric measurements

All anthropometric measurements were performed by a trained research assistant and performed in triplicate. The average of the readings was calculated and used for the final analysis. The participants were asked to wear light clothing with no shoes and socks for the anthropometric measurements. The height of each participant was measured using a portable stadiometer and recorded in centimetres (cm) to the nearest 0.5cm. The body weight (to the nearest 0.1kg) of each participant was then measured using an Omron body fat meter HBF-511 digital bathroom scale which was calibrated and standardised before use by using a weight of a known mass. The waist circumference was measured in centimetres (cm) by placing a non-elastic tape around the narrowest part of the abdomen (between the lowest rib and the top of the iliac crest) of each participant while in a standing position. This was followed by the measurement of the hip circumference (cm) which was performed by placing the tape around the widest circumference of the hips and buttocks of the participant while in a standing position.

2.6.4 Body Mass Index (BMI)

The body mass index (BMI) of each participant was calculated by dividing the weight (kg) of the individual by the square of the height (m), BMI = $\frac{Body\ mass\ (Kg)}{height\ (m)^2}$. The international classification of adult obesity using BMI according to the WHO criteria of 2004 (updated in 2016) is as follows (Philip *et al.*, 2004):

- Underweight: BMI less than 18.50 kg/m^2
- Normal range: BMI 18.50 to 24.99 kg/m^2
- Overweight: BMI greater or equal to 25.00 - 29.99 kg/m^2
- Obese: BMI greater or equal to 30.00 kg/m^2

2.6.5 Biochemical analysis

Fasting and postprandial blood samples were collected from each participant by a qualified registered nursing sister. The known diabetic subjects were self-reported and confirmed by their medical records or medication and only fasting blood samples were drawn from these

subjects. The samples were sent to Pathcare Reference Laboratory, Cape Town, South Africa where biochemical analyses of the parameters listed in Table 2.1 were performed in ISO15189 accredited pathology laboratories.

Table 2.1: The biochemical parameters measured at Pathcare Reference Laboratory

Analyte	Method used	Equipment/Analyser used
Plasma glucose (mmol/L)	Enzymatic Hexokinase	Beckman AU; Beckman Coulter, South Africa
Glycated haemoglobin (HbA1c) (%)	High-Performance Liquid Chromatography	Bio-Rad Variant Turbo, Bio-Rad, South Africa
Serum insulin (mmol/L)	Paramagnetic Particle Chemiluminescence Assay	Beckman AU; Beckman Coulter, South Africa
Total cholesterol (mmol/L)	Enzymatic Immuno-inhibition	Beckman AU; Beckman Coulter, South Africa
High-density lipoprotein cholesterol (HDL-c) (mmol/L)	Enzymatic Immuno-inhibition	Beckman AU; Beckman Coulter, South Africa
Low density lipoprotein cholesterol (LDL) (mmol/L)	Enzymatic Selective Protection	Beckman AU; Beckman Coulter, South Africa
Triglycerides (mmol/L)	Glycerol Phosphate Oxidase-peroxidase	Beckman AU; Beckman Coulter, South Africa
Ultra-sensitive C-reactive protein (U-CRP) ((mg/L)	Latex Particle Immunoturbidimetric	Beckman AU; Beckman Coulter, South Africa
Serum Cotinine (ng/mL)	Competitive Chemiluminescent	Immulite 2000, Siemens, South Africa
Gamma-glutamyl transferase (IU/L)	IFCC standardised method	Beckman AU; Beckman Coulter, South Africa

The oral glucose tolerance test (OGTT) was performed on participants with no previous diagnosis of diabetes mellitus. The participants were required to fast overnight and the OGTT was performed the next day according to the WHO guidelines (World Health Organisation, 1999). Fasting blood samples were collected from each participant after which they were given 75 grams of anhydrous glucose dissolved in 250-300ml of water. Once ingested, the time was recorded and a second blood sample (postprandial) was collected after two hours.

In addition to the biochemical parameters analysed by Pathcare Reference Laboratory, a full blood count was also done for all participants. Furthermore, ethylenediaminetetraacetic acid (EDTA) treated blood samples were stored at −20 degrees Celsius for DNA extraction and analysis.

2.6.6 Type 2 diabetes mellitus classification

The classification of participants was done according to the revised WHO criteria of 1999 (Alberti and Zimmet, 1998; World Health Organisation, 1999; World Health Organisation, 2006). The participants were categorised as being normoglycaemic or normal glucose tolerance (NGT), prediabetes, new diabetes or screen-detected diabetes and known diabetes on treatment. The diagnosis of diabetes was dependant on a history of diabetes and the fasting and postprandial glucose concentrations as categorised below:

- T2DM: if fasting plasma glucose is ≥7.0 mmol/L or post-2-hour plasma glucose is ≥11.1 mmol/L
- Impaired glucose tolerance: if fasting plasma glucose is <7.0 mmol/L and post 2-hour plasma glucose is between ≥7.8 mmol/L and <11.1 mmol/L
- Impaired fasting glucose: if fasting plasma glucose is between 6.1 and 6.9 mmol/L and if measured, post-2-hour plasma glucose is <7.8 mmol/L

2.6.7 Statistical analysis

Descriptive statistics for the general characteristics of participants were performed using the IBM SPSS Statistics for Windows (version 26) software (IBM Corp, Armonk, New York, USA). The distribution of data was evaluated. Data for the general characteristics that displayed normal distribution was reported as the mean. Standard deviation and one way ANOVA (analysis of variation) was used to determine the p-value. General characteristics which displayed a skewed distribution of data was reported with the median and interquartile range (25^{th} and 75^{th} percentile) and p-values were determined by using the Kruskal-Wallis test. A p-value of <0.05 was considered significant.

2.6.8 DNA extraction

Genomic DNA was extracted from peripheral blood samples collected in the EDTA tubes using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. White blood cells were lysed with the Nuclei Lysis Solution followed by the removal of cellular proteins by salt precipitation. The high molecular weight genomic DNA left in solution was concentrated and desalted by isopropanol precipitation. DNA was then quantified using the NanoDrop One Spectrophotometer (ThermoFisher Scientific, Wilmington, USA) which measures the reflection or transmission of material and aides in estimating protein contamination in DNA samples. In terms of the A_{260}/A_{280} and A_{260}/A_{230} ratios, protein contamination is excluded to the greatest possible extent when the values are between 1.8 and 2.

2.6.9 Genome-wide DNA methylation sequencing

For this particular study, DNA methylation was investigated using whole-genome methylated DNA immunoprecipitation sequencing (MeDIP-seq) (Illumina, San Diego, CA, USA). Although bisulphite sequencing is considered the gold standard, it does not distinguish between 5-methylcytosine (5-mC) and 5-hydroxymethylcytosine (5-hmC) which is possible with MeDIP-seq (Olova *et al.*, 2018). Also, MeDIP-seq can detect up to 70% of all CpG nucleotides in the human genome, making it more suitable for large-scale applications generating a huge amount of data which is not the case for whole-genome bisulphite sequencing. In addition, MeDIP bypasses the need for a bisulphite conversion step and can be utilised at relatively low costs (Jeong *et al.*, 2016). Although MeDIP-seq has some limitations which include a low resolution and inability to precisely locate the methylated CpG sites in the genome, the benefits of using this method was suited for the study.

Genome-wide DNA methylation techniques involve three steps, namely the library preparation, sequencing and lastly data analysis. The workflow process depicted in Figure 2.1 was utilised in this study and commenced once DNA extraction had been completed.

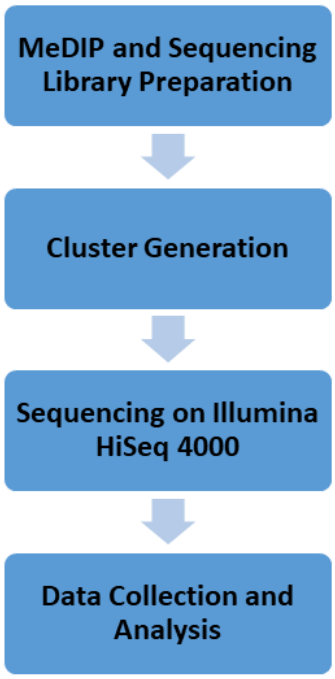


Figure 2.1: Overview of the experimental workflow for genome-wide DNA methylation using Methylated DNA immunoprecipitation

2.6.9.1 Methylated DNA Immunoprecipitation and sequencing library preparation

Methylated DNA immunoprecipitation (MeDIP) was performed by Arraystar Inc. (Rockville, MD, USA) according to Down et al.(2008), with minor modifications (Down *et al.*, 2008). A minimum of 2 µg of DNA (concentrations ranging between 70 and 130 ng/µL) with A260/A280 and A260/A230 ratios ≥ 1.8 was shipped frozen on dry ice, as instructed by Arraystar Inc. (Rockville, MD, USA). Genomic DNA was sonicated to approximately 200–500bp fragments with a Bioruptor sonicator (Diagenode, Denville, NJ, USA) and then 1µg of fragmented DNA was prepared for Illumina HiSeq 4000 sequencing. Firstly, the sonicated DNA was end-repaired with T4 DNA polymerase, Klenow DNA polymerase and T4 polynucleotide kinase. A single adenine base was added to the 3' ends with Klenow (exo minus) polymerase. Illumina's single-end genomic adapters were ligated to the DNA fragments and agarose gel size-selection was used to remove unligated adapters. The adaptor-ligated DNA fragments were immunoprecipitated by using a mouse monoclonal anti-5-methylcytosine antibody (Diagenode). DNA was heat-denatured at 94°C for 10 minutes, rapidly cooled on ice, and immunoprecipitated with 1 µL primary antibody overnight at 4°C with rocking agitation in 400 µL immunoprecipitation buffer (0.5% BSA in Phosphate buffered saline). The immunoprecipitated DNA fragments were recovered by adding 100 µL of protein

G magnetic beads (Life Technologies, Carlsbad, CA, USA) followed by incubation for an additional 2 hours at 4°C with agitation. After immunoprecipitation, a total of five immunoprecipitation washes were performed with ice-cold immunoprecipitation buffer. A nonspecific Mouse IgG immunoprecipitation was performed in parallel to methyl DNA immunoprecipitation as a negative control. Washed beads were resuspended in TE buffer with 0.25% SDS and 0.25 mg/mL proteinase K for 2 hours at 65°C and then allowed to cool down to room temperature. MeDIP and supernatant DNA were purified using Qiagen MinElute columns and eluted in 16 µL elution buffer (Qiagen, Germantown, MD, USA). Fourteen cycles of PCR were performed using 5 µL of the immunoprecipitated DNA and single-end Illumina PCR primers. The resulting PCR reactions were purified with Qiagen MinElute columns, after which a final size selection of approximately 300-600 bp was performed using agarose gel electrophoresis 2% agarose (w/v). Library size was quality controlled by the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Quality assessment of the MeDIP procedure involved normalization of each library to 5 ng/µL of which 1 µL was used for real-time PCR confirmation of successful methylation region enrichment. For the qPCR DNA immunoprecipitation enrichment assessment specific methylated sites and non-methylated sites at the H19 and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene loci were used, respectively.

2.6.9.2 Sequencing

The library was denatured with 0.1 M NaOH to generate single-stranded DNA molecules and loaded onto channels of the flow cell at 8 pM concentration, amplified *in situ* using the TruSeq Rapid SR Cluster Kit (Illumina, San Diego, CA, USA). Sequencing was carried out by running 100 cycles on the Illumina HiSeq 4000 according to the manufacturer's instructions.

2.6.9.3 MeDIP-Seq data analysis

2.6.9.3.1 Analysis of sequencing data

After generation of the sequencing images, image analysis and base calling were performed using Off-Line Basecaller software (OLB V1.8). First, the reads had to pass through the Solexa CHASTITY quality filter. Individual bases generated from original image files had quality scores which reflected the probability of whether a base-call is correct or not. The

quality score was calculated by the CHASTITY Formula shown in Figure 2.2. In short, the CHASTITY (C) of each base in the short reads is determined by the intensity of four colours (I_A, I_C, I_G, and I_T), using the following formula: the ratio of the highest (I_C) of the four (base type) intensities to the sum of the highest two (I_C and I_G). The CHASTITY (C) should be no less than 0.6 in the first 25 bases. After passing a Solexa CHASTITY quality filter, the clean reads were aligned to the human genome (UCSC HG19) using HISAT2 software (V2.1.0).

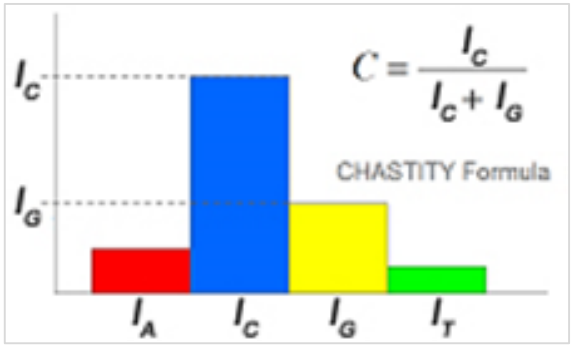


Figure 2.2: Diagram showing the CHASTITY Formula

2.6.9.3.2 Detection of peaks

Statistically significant MeDIP-enriched regions (peaks) were identified by MACS v2 (Model-based Analysis of ChIP-Seq) software for each sample by comparison to input background, using a q-value threshold of 10^{-5}. The peaks in samples were annotated by the nearest gene using the newest UCSC RefSeq database (https://genome.ucsc.edu/). The peaks were divided into three classes based on their distances to the UCSC RefSeq genes:

1. Promoter peaks had their centres within promoter regions which were identified as 2000 bp upstream and downstream from the transcription start site (TSS).
2. Gene body peaks were found in the gene body region defined as + 2000 bp downstream of the transcription start site (TSS) to the transcription termination site (TTS).
3. Intergenic peaks had their centres within intergenic regions which were defined as other genomic regions not included in the promoter or gene body regions.

In addition, the MACS v2 software was used to identify the statistically significant lncRNA-associated MeDIP enriched peaks for each sample using a q-value threshold of 10^{-5} and annotated by the nearest gene using the newest UCSC RefSeq database.

2.6.9.3.3 Differentially methylated regions (DMRs)

Differentially methylated regions (DMRs) located within gene promoters (TSS − 2000 bp, TSS + 2000 bp) with statistical significance between two groups were identified by diffReps (Cut-off: log2FC = 1.0, p-value = 10−4). Both mRNA and lncRNA-associated DMRs within promoters were annotated by the nearest gene using the UCSC RefSeq and database of multiple databases integration. There were six comparison pairs for both the mRNA and lncRNA-associated DMRs provided in the study. They were as follows:

1. Known diabetes versus NGT
2. Screen-detected diabetes (newly diagnosed) versus NGT
3. Prediabetes versus NGT
4. Known diabetes versus screen-detected diabetes
5. Known diabetes versus prediabetes
6. Prediabetes versus screen-detected diabetes (newly diagnosed)

2.6.9.4 Gene Ontology (GO) analysis

Following the identification of the statistically significant DMRs, gene ontology analysis was performed for each of the six comparison groups. The gene ontology project (http://www.geneontology.org) is a structured vocabulary of terms used to describe the gene and gene product attributes in any organism (Gene Ontology Consortium, 2006). The ontology covers three domains, namely biological process, cellular component, and molecular function. The biological process is the largest of the three GO domains and represents the larger processes or specific objectives an organism is programmed to achieve. Biological processes are usually achieved by several molecular processes carried out by specific gene products in a regulated manner and sequence (Thomas PD, 2017). The molecular function refers to molecular activities carried out by a single gene product or it may be part of a larger process

utilising multiple gene products (Thomas PD, 2017). The cellular component refers to the location related to the cellular structures or anatomy where the gene product performs its function (Thomas PD, 2017). Gene ontology structure is that of a Directed Acyclic Graph (DAG) where genes are linked to the most specific set of terms that describes its functionality. Each term can have multiple relationships to broader parent terms and more specific child terms which makes this a more flexible structure in comparison to a hierarchy (Du Plessis *et al*., 2011).

In the current study, Fisher's exact test was used to determine whether there was more overlap between the differentially enriched (DE) list and the GO annotation list than would be expected by chance. The p-value denoted the significance of GO terms enrichment in the DE genes. The lower the p-value, the more significant the GO term. A p-value of ≤ 0.05 was considered significant. Annotation was performed using standard workflow according to http://www.geneontology.org/.

2.6.9.5 KEGG pathway analysis

Pathway analysis for each of the six comparison groups was done using the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database. KEGG is a collection of databases linking genomic information with higher-level systemic functional information (Kanehisa and Goto, 2000). The integrated database initially consisted of three generic categories of systems information, genomic information and chemical information, but has grown and now includes 18 databases in four categories (Kanehisa *et al.*, 2019).

In this study, the p-value (EASE score, Fisher's p-value, or hypergeometric p-value) denoted the significance of the pathway correlated to the conditions. The lower the p-value is, the more significant the pathway. A p-value of ≤ 0.05 was considered significant.

2.6.10 Validation of significant DMRs using pyrosequencing

The significant DMRs identified in the 48 participant samples by the MeDIP sequencing was validated using pyrosequencing. Pyrosequencing has been considered a golden standard for the quantitative methylation analysis of bisulphite converted DNA (De Chiara *et al.*, 2020). The advantages and disadvantages of this sequencing technique have been addressed in

chapter one of the thesis and have aided in the decision to use pyrosequencing as a validation method. Also, in a study by Šestáková and co-workers (2019) where different DNA methylation validation methods were compared, pyrosequencing was deemed one of the most suitable methods in terms of overall feasibility, the DNA methylation information obtained and the consistency across various methylation levels (Šestáková *et al*., 2019). Despite it being a more time-consuming technique as well as costly due to the instrument required, pyrosequencing has more advantages in terms of assessing DNA methylation of a specific locus.

Pyrosequencing assays require three main steps namely, bisulphite conversion of unmethylated cytosine to uracil, PCR to generate an amplicon and a pyrosequencing reaction to analyse the nucleotide content of the amplified fragment.

2.6.10.1 Bisulphite conversion

Bisulphite conversion was first performed on the 48 participant samples using the EpiTect Fast DNA Bisulphite Kit (Qiagen, Germantown, MD, USA).

2.6.10.1.1 Principle of bisulphite conversion

Bisulphite conversion has long been described as the gold standard for detecting DNA methylation (Frommer *et al.*, 1992; Leti *et al.*, 2018). DNA is treated with bisulphite which converts cytosine residues to uracil while methylated cytosine residues such as 5-methylcytosines, remain unaffected. Following the conversion, sequencing can then distinguish between unmethylated cytosines which are displayed as thymine and 5-methylcytosines which are displayed as cytosines in the resultant amplified sequence of the sense strand. The bisulphite conversion chemistry comprises three steps namely, sulphonation, hydrolytic deamination, and desulphonation. Sulphonation involves the addition of bisulphite to the 5-6 double bond of cytosine. This is followed by the hydrolytic deamination of the cytosine-bisulphite derivative resulting in a uracil-bisulphite derivative. Lastly, desulphonation is the removal of the sulphonate group by an alkali treatment to give uracil (Patterson *et al.*, 2011).

2.6.10.1.2 Laboratory procedure

The EpiTect Fast DNA Bisulphite Kit (Qiagen, Germantown, MD, USA) was applied to 500ng of DNA according to the manufacturer's instructions. The bisulphite reactions were prepared in 200 µL PCR tubes by adding the following components: 500ng DNA in RNAse-free water up to a maximum volume of 20 µL; 85 µL bisulphite mix and 35 µL DNA Protect Buffer. After thoroughly mixing the PCR tube contents at room temperature and confirming the DNA Protect Buffer colour change from green to blue indicating that sufficient mixing and the correct pH for bisulphite conversion has been reached, the tubes were placed in a thermal cycler. The thermal cycler with a heated lid was programmed according to Table 2.2.

Table 2.2: Bisulphite conversion thermal cycler conditions

STEP	TIME	TEMPERATURE
Denaturation	5 min	95°C
Incubation	25 min	60°C
Denaturation	5 min	95°C
Incubation	85 min (1 h 25 min)	60°C
Denaturation	5 min	95°C
Incubation	175 min (2 h 55 min)	60°C
Hold	Indefinite	*20°C

*Converted DNA can be left in the thermal cycler overnight without any loss of performance

Once the bisulphite conversion was completed, the PCR tubes containing the bisulphite reactions were briefly centrifuged and then transferred into clean 1.5 ml microcentrifuge tubes. This was followed by the addition of 560 µL of the freshly prepared Buffer BL containing 10 µg/ml carrier RNA to each sample which was mixed and briefly centrifuged at room temperature. The mixture was then transferred into EpiTect spin columns where the Buffer BL promotes binding of the converted single-stranded DNA to the EpiTect spin column membrane. The spin columns were then centrifuged at maximum speed for 1 minute after which the flow-through was discarded and the spin columns placed back in collection tubes.

Next, 500 µL Buffer BW was added to each spin column, centrifuged at maximum speed for 1 minute and the flow-through discarded. The function of the Buffer wash was to efficiently remove residual sodium bisulphite as part of the washing process. Then 500 µl Buffer BD, the desulfonation agent, was added to each spin column, incubated for 15 minutes at room temperature, centrifuged at maximum speed for 1 minute and the flow-through discarded. This was followed by two washing steps using the Buffer BW to further desalt the DNA. Additional centrifugation at maximum speed for 1 minute, as well as incubation on a heating block for 5 minutes at 56°C, was done to remove any residual liquid before the spin columns were placed in new and clean 1.5 ml microcentrifuge tubes. Next, 20 µL Elution buffer EB was dispensed onto the centre of each spin column membrane to elute the purified DNA by centrifugation for 1 minute at approximately 15,000 x g (12,000 rpm). The purified DNA could then be stored for up to 24 hours at 2-8°C or longer periods at -15 to -30°C.

2.6.10.2 Amplification of the bisulphite converted DNA

The amplification of bisulphite converted DNA was conducted using the PyroMark PCR Kit (Qiagen, Germantown, MD, USA). This step aimed to ensure specific and unbiased amplification of the template DNA for pyrosequencing analysis. The reaction mixture was set up by adding 12.5 µL PyroMark PCR Master Mix, 2.5 µL CoralLoad Concentrate, 2.5 µL primer solutions (0.2 µM concentration), 2.5 µL $MgCl_2$ (4 mM concentration) and 3 µL RNAse-free water in a PCR tube. Then 2 µL of the DNA template (500 ng/25 µL concentration) was added to the mix and the tubes were placed in the thermal cycler which was programmed according to Table 2.3. The EpiTect PCR Control DNA Set (Qiagen, Hilden, Germany) for methylated and unmethylated DNA was used as part of the quality control process.

Table 2.3: Thermal cycler program for the amplification of the bisulphite converted DNA

STEP	TIME	TEMPERATURE	ADDITIONAL COMMENTS
Initial PCR activation step	15 minutes	95°C	DNA Polymerase is activated by this heating step
3-step cycling:			
Denaturation	30 seconds	94°C	
Annealing	30 seconds	56°C	Cycle repeated 45 times
Extension	30 seconds	72°C	
Final extension	10 minutes	72°C	

The Pyromark CpG Assay primers (Qiagen, Hilden, Germany) used were designed specifically to bisulphite-modified DNA and were HPLC-purified biotinylated primers. One primer must be biotinylated at its 5' end to prepare a single-stranded PCR product for use in the subsequent Pyrosequencing procedure.

After amplification, the samples were ready for pyrosequencing but could also be stored overnight at 2–8°C or -20°C for longer periods. The PCR product integrity was verified before pyrosequencing by agarose gel electrophoresis (1% agarose w/v). Verification of single, strong bands with no unincorporated primers was an indication that PCR was successful.

2.6.10.3 Pyrosequencing

Pyrosequencing was conducted using the PyroMark Q48 Autoprep instrument, the PyroMark Q48 Autoprep reagents (Qiagen, Hilden, Germany) and PyroMark PCR Reagents (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions. Primers for selectively identified DMRs was used in pyrosequencing analysis to validate methylation levels detected during MeDIP analysis.

2.6.10.3.1 Principle of pyrosequencing

Pyrosequencing is a sequencing method used for the quantitative methylation analysis of bisulphite converted DNA (Šestáková *et al*., 2019). It uses a sequencing-by-synthesis method whereby nucleotides are dispensed one at a time and incorporated into an extending strand and then degraded before the dispensation of the next nucleotide. Pyrosequencing requires a single-stranded PCR amplicon that serves as a DNA template, four different enzymes including DNA polymerase, ATP sulfurylase, luciferase, and apyrase, as well as two different substrates including adenosine 5′ phosphosulfate (APS) and luciferin (Delaney *et a*l., 2015). During the first step, a sequencing primer is annealed to a single-stranded DNA (ssDNA) template. A single nucleotide is then subsequently incorporated by DNA polymerase, releasing pyrophosphate (PPi). The liberated PPi is used by ATP sulfurylase in the presence of APS to generate ATP, which activates the luciferase-mediated conversion of luciferin to the light-emitting oxyluciferin. The light that is emitted is proportionate to the number of nucleotides that are added to the elongating strand. The excess nucleotide is degraded by apyrase followed by the dispensation of the next nucleotide. The light intensity produced is detected and translated as a peak on a pyrogram. The amount of methylation can be determined by comparing the peak light emission of a cytosine (methylated signal) or thymine (unmethylated signal) incorporation at a CpG site within the amplicon. This will provide a precise measure of the amount of methylation at that position within the sample.

2.6.10.3.2 Pyrosequencing analysis

The PyroMark Q48 Autoprep instrument (Qiagen, Hilden, Germany) and related PyroMark Q48 Autoprep reagents (Qiagen, Hilden, Germany) were used to perform the pyrosequencing. The PyroMark Control Oligo (Qiagen, Hilden, Germany) was used for quality control and performed by using the Pyromark CpG Assay Primers (Qiagen, Hilden, Germany) (Table 2.4) selected for pyrosequencing. Thereafter, the prepared amplified DNA segments were analysed.

Table 2.4: The Pyromark CpG Assay Primers used in the pyrosequencing analysis

Pyromark CpG Assay Name	Gene symbol	Accession number	Amplicon length	Number of CpGs	Primer sequence
Hs_CYP1A1_01_PM	CYP1A1	ENSG00000140465	102	4	GTCCGCATTTTCGGTCCACGCC TGTGGCASGACACGA
Hs_MIIP_01_PM	MIIP	ENSG00000116691	97	4	CGTGCGAAAACGGCGA
Hs_SEMA4F_01_PM	SEMA4F	ENSG00000135622	118	5	CGGAACTTCGAAATGAACGACCT CGGGCAGCCCCCACCCGA
Hs_GUK1_01_PM	GUK1	ENSG00000143774	210	5	CGTCGGGAAGGCCTGCTGTTTC CCCAAGTCCCCGTCTCTGCGCC GT
Hs_AL118506.2_03_PM	ABHD16B	ENSG00000183260	84	5	GTGTTGGACGCCACCTTCGACG ACCTTGTGCCGCTGGCGC
Hs_KIAA1467_02_PM	KIAA1467	ENSG00000084444	241	3	TCCCCTCCACCTAACGGCAGCA ACGACGC
Hs_SLC10A3_01_PM	SLC10A3	ENSG00000126903	107	3	GGCTGGGTGCCCGGGGTTAGG GTTTTGGCCACGTTGCCGC
Hs_ZNF574_01_PM	ZNF574	ENSG00000105732	191	6	CGTCGAGATTCCCCGCAGCGCC CAACCAATCCTTTAGCGTCG
Hs_TFE3_01_PM	TFE3	ENSG00000068323	195	6	GCGCGGTCTAGGGCTCAGATAT TRCAAATATTGGCCGAAAGAGG GCGCGGTTCCG
Hs_FXYD5_01_PM	FXYD5	ENSG00000089327	132	3	CGCCTCCCGCGC
Hs_ADK_01_PM	ADK	ENSG00000156110	233	5	TGGAGGCGCCGCAAGCGCTGA GGTGAGCGCTGCCGGA
Hs_TNFSF9_01_PM	TNFSF9	ENSG00000125657	143	4	CGAGTGGAGAAAATTCCGCAGA GTCACGGGGACGA

Setting up the assay involved preparing the PyroMark Q48 Autoprep instrument which was guided by the PyroMark Q48 Autoprep software on the instrument itself. This preparation included the insertion of the absorber strip, injector loading, disc insertion and bead and template loading. Insertion of the PyroMark Q48 absorber strip involved the sliding of the strip into place ensuring that it was level. The main function of the absorber strip was to collect liquids spun out from the disc wells during a wash cycle as well as liquid waste from the priming and cleaning steps. Injector priming was performed to ensure that the injectors were operating correctly. Reagents were pipetted into the designated injectors according to the volumes shown on the instrument touchscreen. This ensured that sufficient reagents were available for the initial injector priming as well as the run. Following the injector priming, all injectors underwent test shots which had to be passed in order to proceed to the next step. Before

insertion of the PyroMark Q48 Disc, 10 µL of the biotinylated PCR product and 3 µL PyroMark Q48 Magnetic beads were loaded into the wells of the disc. The instrument software indicated the well position for each sample. When setting up the Control Oligo assay, 20 µL of the PyroMark Control Oligo and 3 µL PyroMark Q48 Magnetic beads were dispensed into the correct wells of the PyroMark Q48 Disc. The PyroMark Q48 disc was then inserted into the instrument by matching the disc grips with those on the rotor and locking the disc into position by screwing down the lock nut. Once this was completed, the instrument was ready to proceed with the run. The run was initiated by selecting 'Start' on the instrument. Once the run was completed, the lid automatically opened and 'Run Complete' was displayed on the instrument touchscreen. The data was now ready for analysis. The sample run and analysis of data was completed in the CpG mode of the instrument and the resultant report at the end included the methylation percentages of all the passed CpG sites. Similarly, the analysis of the Control Oligo was completed in the AQ mode of the instrument and the report generated showed the percentages of the T (thymine) and C (cytosine) calculated.

CHAPTER 3

RESULTS

3.1 General characteristics of participants

General characteristics of the 48 female participants are presented in Table 3.1. The study group comprised 12 normoglycaemic (normal glucose tolerance-NGT), 12 prediabetics, 12 screen-detected (newly diagnosed) and 12 known (on treatment) diabetic individuals who were all matched for age and body mass index. As expected there was a significant increase in plasma glucose ($p = <0.0001$) and HbA1c levels ($p = <0.0001$) in the participants with hyperglycaemia (prediabetes, screen-detected diabetes and known diabetes) when comparing to the normoglycaemic group. The lipid profile only showed a significant increase in the triglycerides for the screen-detected and known diabetic groups ($p= 0.004$). The ultra-sensitivity C - reactive protein (us-CRP) showed a significant difference ($p = 0.024$) amongst the groups.

Table 3.1: Table of characteristics of the 48 participants

Characteristics	Total	Normoglycaemic	IGT	New DM	Known DM	p-value
n	48	12	12	12	12	
Age (years)	53.38(8.20)	52.08 (7.83)	53.50 (8.53)	54.75 (7.52)	53.17 (9.61)	0.893
Weight (kg)	75.49 (18.83)	70.18 (18.89)	79.31 (21.7)	80.80 (18.30)	71.66 (14.53)	0.403
Height (cm)	155.86 (6.22)	157.15 (7.86)	154.38 (5.98)	156.05 (5.93)	155.96 (5.54)	0.771
Body Mass Index (kg/m^2)	30.95 (7.68)	27.35 (5.80)	33.27 (9.07)	33.52 (8.93)	29.37 (5.00)	0.145
Ave Waist (cm)	93.32 (16.97)	83.21 (18.53)	97.10 (13.78)	101.29 (19.70)	91.68 (10.50)	0.049
Ave Hip (cm)	106.14 (16.44)	102.00 (18.42)	109.92 (18.85)	109.38 (16.57)	103.28 (11.90)	0.540
Systolic blood pressure (mmHg)	137.85 (26.85)	135.25 (30.78)	137.00 (18.52)	142.92 (32.91)	136.25 (25.86)	0.904
Diastolic blood pressure (mmHg)	86.23 (16.56)	78.00 (16.06)	88.92 (11.68)	94.17 (22.07)	83.83 (11.52)	0.094
Pulse rate (bpm)	73.19 (14.14)	64.83 (12.19)	71.25 (5.82)	80.08 (16.40)	76.58 (16.09)	0.041
Fasting plasma glucose (mmol/L)	5.70 (4.90-8.10)	4.75 (4.08-4.90)	5.30 (4.93-5.78)	7.90 (6.08-11.48)	9.25 (5.85-16.95)	<0.0001
Glucose 2 HRs Post Prandial (mmol/L	9.00 (7.20-13.30)	5.20 (4.00-7.20)	9.00 (8.48-10.05)	15.10 (12.70-19.75)		<0.0001
HbA1c(%)*	6.30 (5.60-7.45)	5.55 (5.10-5.83)	6.00 (5.28-6.55)	7.05 (6.43-9.48)	9.20 (6.53-10.85)	<0.0001
Fasting serum insulin (mIU/L)*	8.00 (5.85-14.18)	5.65 (2.83-7.98)	9.45 (7.38-12.50)	14.10 (5.90-21.20)	7.90 (5.90-16.20)	0.012
2HR Insulin (mIU/L)*	53.20 (27.50-104.55)	28.50 (15.80-43.60)	111.20 (60.20-187.20)	48.80 (30.33-90.60)		0.003
Triglycerides-S (mmol/L)*	1.44 (1.04-2.10)	0.94 (0.70-1.27)	1.50 (1.07-2.00)	1.72 (1.37-2.95)	2.06 (1.20-2.36)	0.004
LDL Cholesterol (Measured) (mmol/L	3.75 (1.04)	3.30 (0.96)	3.46 (0.96)	4.18 (0.84)	4.08 (1.22)	0.093
Cholesterol HDL-S (mmol/L)*	1.30 (1.10-1.60)	1.40 (1.20-1.85)	1.25 (1.10-1.40)	1.15 (1.10-1.48)	1.30 (0.90-1.60)	0.253
Cholesterol-S (mmol/L)	5.83 (1.15)	5.51 (1.10)	5.39 (0.98)	6.36 (0.85)	6.10 (1.47)	0.119
High-sensitivity CRP (mg/L)*	5.19 (2.24-14.66)	2.47 (1.47-3.41)	7.90 (4.89-14.03)	6.78 (1.75-18.67)	15.14 (3.04-24.16)	0.024
Cotinine (ng/mL)*	15.55 (10.00-216.75)	22.50 (10.00-268.00)	91.00 (10.00-222.00)	85.00 (10.00-208.00)	15.55 (10.00-189.50)	0.975
Gamma-glutamyl transferase (IU/L)*	30.00 (19.00-61.00)	19.50 (12.50-29.25)	37.50 (21.00-89.50)	45.00 (18.75-85.25)	33.00 (20.00-61.00)	0.118

p-values were determined using one-way ANOVA;
*, median [25th-75th percentiles], and p-values from Kruskal Wallis test.

3.2 Differentially Methylated Regions (DMRs)

There was a total of 954820 statistically significant MeDIP-enriched regions (peaks) identified for the 48 samples (Figure 3.1). When compared to the gene body and intergenic regions, the promoter region (TSS-2000bp, TSS+2000bp) as expected, showed the least amount of peaks for all the groups. The statistically significant differentially methylated regions (DMRs) between promoters of two groups were identified by diffReps using a limit of log2FC=1.0 and p-value=0.0001.

The DMRs identified for both hypermethylation and hypomethylation have been presented in Figure 3.1. There are six comparison pairs provided in the data, namely known diabetes versus NGT; screen-detected diabetes versus NGT; prediabetes versus NGT; known diabetes versus

screen-detected diabetes; known diabetes versus prediabetes and prediabetes versus screen-detected diabetes. A total of 366 DMRs have been observed, of which 63% were hypermethylated and 37% hypomethylated. The screen-detected and known diabetes subjects showed the highest number of DMRs and more than 70% of these DMRs were hypermethylated when compared to the NGT group.

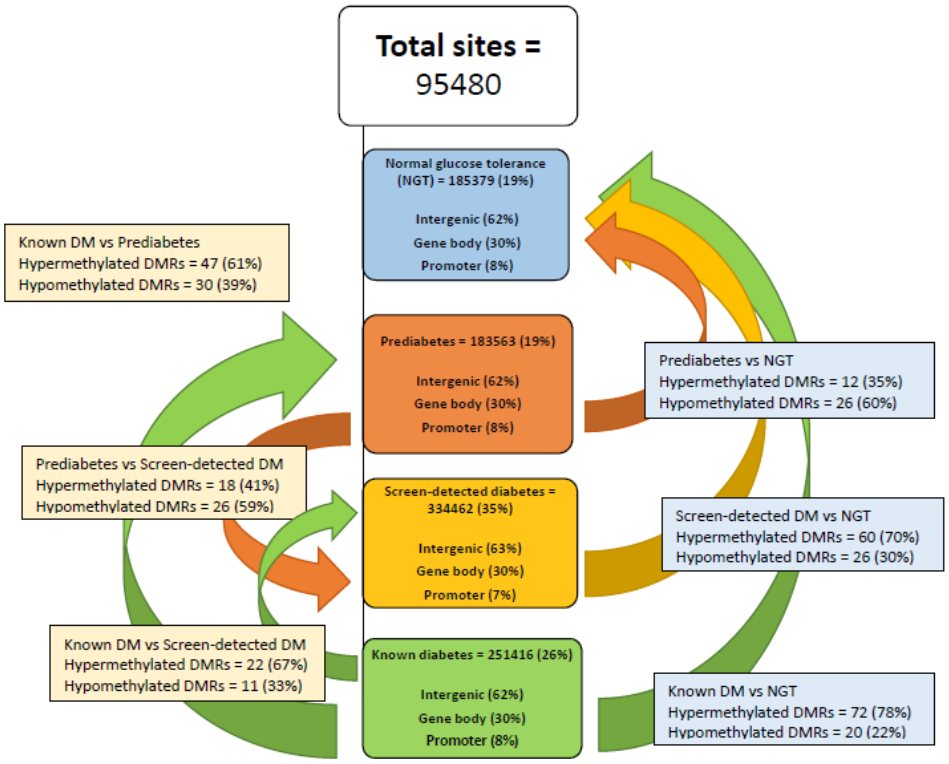


Figure 3.1: MeDIP enriched regions (peaks) identified in all the 48 samples. The total sites are the sum of the peak number for the subjects with known diabetes (known DM), screen-detected diabetes (screen-detected DM); prediabetes and normal glucose tolerance (NGT). The distribution of the peaks in the intergenic, gene body and promoter regions are shown. The hypermethylated and hypomethylated DMRs within the gene promoter for the various glucose tolerance groups are shown

The DMRs were also grouped according to their chromosomal location (Figure 3.2). When looking at the six comparison groups, the most hypermethylated DMRs were found on chromosomes 19, particularly when comparing the known diabetes and screen-detected diabetes to the NGT group. In addition, hypermethylated DMRs were more common in chromosome 1 when observing the known diabetes to prediabetes subjects. On the other hand, most hypomethylated DMRs were commonly found in chromosomes X and 1.

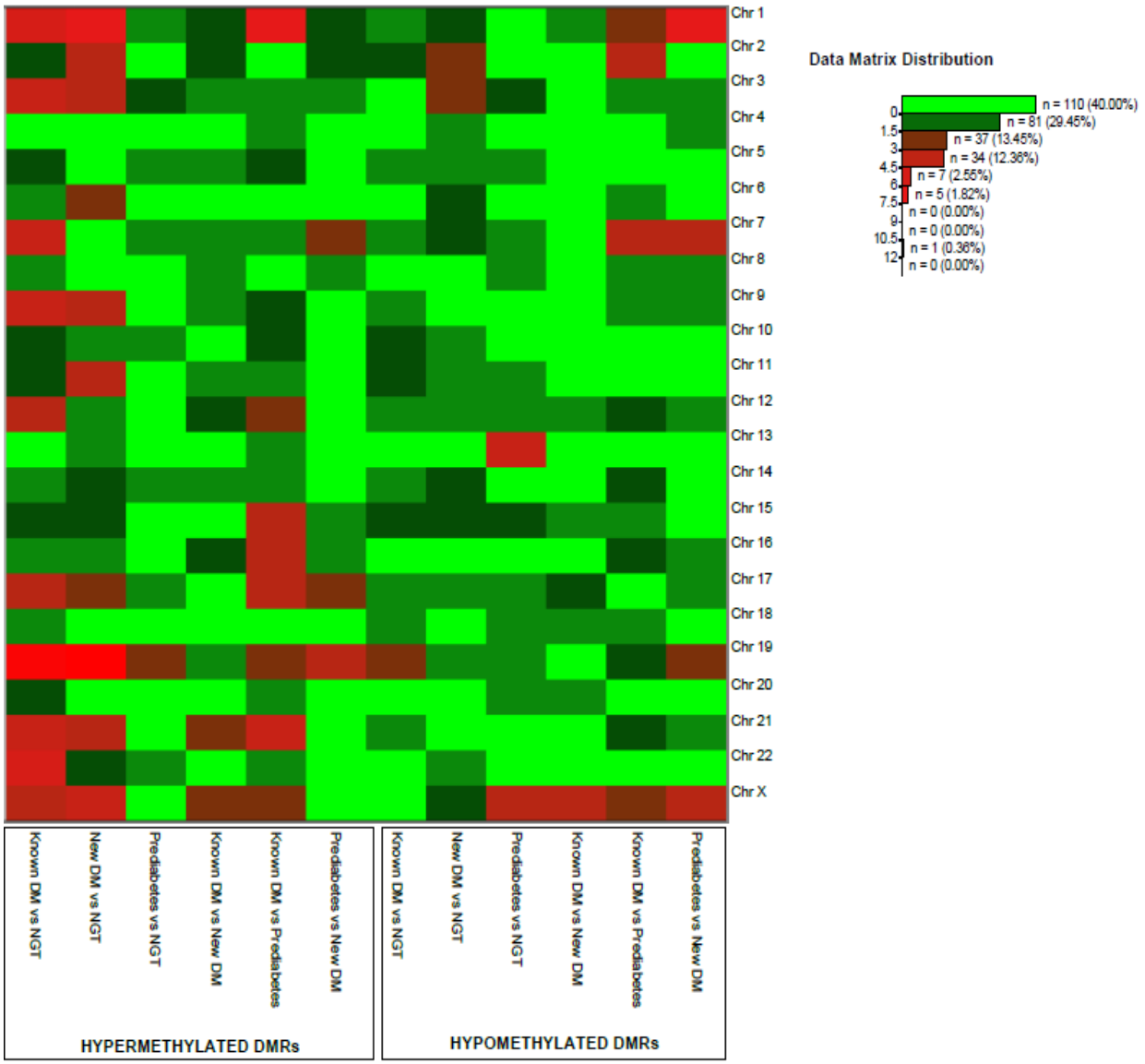


Figure 3.2: Heatmap presenting the distributions of differentially methylated regions (DMRs) among the chromosomes for all six comparison groups, namely known diabetes versus NGT; screen-detected diabetes versus NGT; prediabetes versus NGT; known diabetes versus screen-detected diabetes; known diabetes versus prediabetes and prediabetes versus screen-detected diabetes. Both hypermethylated and hypomethylated DMRs are presented

The DMRs were further summarised according to the top ten DMRs per fold change for each of the six comparison groups (Table 3.2 and 3.3). As expected all the hypomethylated DMRs had a negative fold change (log2FC values less than -1).

Furthermore, when observing all the DMRs for the diabetic subjects (known diabetes and screen-detected diabetes), there were 11 hypermethylated and 7 hypomethylated DMRs commonly found when comparing them to the NGT group (Table 3.4). The hypermethylated DMRs included BAGE2, BAGE3, BAGE4, BAGE5, CD248, COL8A2, IGSF9, PACSIN1,

SPACA3, SYT3 and TMEM89 whereas the hypomethylated DMRs included ADK, CCDC53, POTEB, POTEB2, RHOBTB3, RNF169 and SLC35B4. When the groups were further categorised to include non-diabetic hyperglycaemia and compared to NGT, 1 hypermethylated DMR (TMEM89) and 4 hypomethylated DMRs (CCDC53, POTEB, POTEB2 and RHOBTB3) were observed (Table. 3.5).

When comparing the known diabetic subjects on metformin treatment to the screen-detected (newly diagnosed) diabetic subjects, there were 22 hypermethylated DMRs and 11 hypomethylated DMRs identified (Table 3.6). The hypomethylated DMRs of interest are SLC25A35 and SLC28A1.

Table 3.2: The top ten hypermethylated DMRs for all comparison groups, known diabetes versus NGT; known diabetes versus screen-detected diabetes; known diabetes versus prediabetes; screen-detected diabetes versus NGT; prediabetes versus NGT and prediabetes versus screen-detected diabetes based on the fold change (log2FC)

HYPERMETHYLATED DMRs

Known DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
RACGAP1	chr12:50417381-50417620	2.43	1.88×10^{-7}
ZNF346	chr5:176448161-176448360	2.28	4.35×10^{-7}
GPR112	chrX:135382161-135382420	2.04	2.03×10^{-8}
CTAGE15	chr7:143268761-143269080	1.99	1.42×10^{-7}
GAGE2A	chrX:49355621-49355840	1.98	9.02×10^{-6}
TNFSF9	chr19:6529841-6530080	1.97	3.7×10^{-7}
SBK3	chr19:56056701-56056920	1.97	1.9×10^{-6}
HIRIP3	chr16:30005501-30005720	1.95	2.01×10^{-7}
PARVB	chr22:44394581-44394780	1.92	6.05×10^{-5}
TJP3	chr19:3720641-3720860	1.91	3.43×10^{-7}

Known DM vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
XAGE1E	chrX:52260741-52261020	1.66	5.36×10^{-7}
XAGE1B	chrX:52260741-52261020	1.66	5.36×10^{-7}
KIAA1467	chr12:13198981-13199200	1.62	2.21×10^{-7}
ASB2	chr14:94442921-94443160	1.6	3.33×10^{-7}
GABPA	chr21:27105221-27105420	1.56	6.07×10^{-6}
ZNF346	chr5:176448161-176448360	1.47	6.99×10^{-6}
FKBP8	chr19:18655321-18655520	1.4	3.83×10^{-7}
CTAGE15	chr7:143268761-143269080	1.38	2.62×10^{-7}
VIPR1	chr3:42531141-42531360	1.37	9.82×10^{-5}
TMEM204	chr16:1584901-1585100	1.31	2.77×10^{-6}

Known DM vs Prediabetes

Gene Name	Genomic Coordinates	log2FC	p-value
POMP	chr13:29231801-29232000	2.84	9×10^{-8}
KIAA1467	chr12:13198981-13199200	2.17	1.12×10^{-7}
ZNF346	chr5:176448161-176448360	2.14	8.49×10^{-5}
TMEM204	chr16:1584901-1585100	2.06	1.61×10^{-7}
RMDN3	chr15:41046161-41046380	2.05	1.54×10^{-7}
VIPR1	chr3:42531141-42531360	2.04	1.92×10^{-7}
KLK13	chr19:51567281-51567500	1.95	8.58×10^{-7}
POU3F1	chr1:38513121-38513340	1.87	2.43×10^{-7}
KCNAB2	chr1:6095841-6096040	1.84	1.97×10^{-7}
C9orf152	chr9:112969721-112969920	1.77	2.75×10^{-8}

Screen-detected DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
FXYD5	chr19:35647241-35647460	2.84	2.16×10^{-9}
DBH	chr9:136500621-136500820	2.77	7.52×10^{-10}
VBP1	chrX:154443841-154444040	2.69	2.02×10^{-7}
SHKBP1	chr19:41084061-41084280	2.16	5.75×10^{-10}
SLC28A1	chr15:85429461-85429660	1.98	1.01×10^{-6}
CSF2RA	chrX:1389401-1389620	1.95	8.97×10^{-5}
AWAT1	chrX:69454821-69455080	1.94	2.03×10^{-8}
CRX	chr19:48323441-48323640	1.77	5.07×10^{-7}
PITPNM2	chr12:123595921-123596120	1.75	7.62×10^{-5}
SPACA3	chr17:31317561-31317800	1.73	1.31×10^{-5}

Prediabetes vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
ZIM2	chr19:57352241-57352440	2.16	9.46×10^{-8}
PLSCR4	chr3:145969281-145969500	1.85	6.59×10^{-7}
TMEM89	chr3:48658161-48658360	1.76	2.6×10^{-7}
ZNF574	chr19:42578781-42578980	1.68	4.45×10^{-5}
CFHR1	chr1:196789721-196789920	1.66	2.98×10^{-5}
TCL1A	chr14:96179621-96179840	1.54	4.62×10^{-5}
TDRD1	chr10:115938921-115939120	1.46	9.82×10^{-5}
SLC4A2	chr7:150759481-150759680	1.46	1.69×10^{-5}
CCDC144NL	chr17:20799641-20799840	1.4	1.97×10^{-5}
C22orf15	chr22:24103821-24104060	1.26	9.51×10^{-5}

Prediabetes vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
SEMA4F	chr2:74879901-74880100	1.88	2.2×10^{-7}
CFHR1	chr1:196789721-196789920	1.45	8.78×10^{-5}
MBD3L4	chr19:7039841-7040040	1.4	1.03×10^{-6}
RGPD5	chr2:110553321-110553580	1.39	9.85×10^{-8}
METTL23	chr17:74724801-74725000	1.38	1.5×10^{-5}
RSPH10B	chr7:6793401-6793600	1.37	2.47×10^{-5}
RSPH10B2	chr7:6793401-6793600	1.37	2.47×10^{-5}
ZIM2	chr19:57352241-57352440	1.28	3.28×10^{-6}
MUC20	chr3:195447581-195447800	1.24	7.96×10^{-5}
C15orf41	chr15:36873481-36873700	1.23	5.12×10^{-7}

Table 3.3: The top ten hypomethylated DMRs for all comparison groups, known diabetes versus NGT; known diabetes versus screen-detected diabetes; known diabetes versus prediabetes; screen-detected diabetes versus NGT; prediabetes versus NGT and prediabetes versus screen-detected diabetes

HYPOMETHYLATED DMRs

Known DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
DHTKD1	chr10:12109861-12110060	-1.01	9.54×10^{-7}
METTL5	chr2:170682181-170682420	-1.03	1.33×10^{-6}
IGSF5	chr21:41117521-41117780	-1.08	9.04×10^{-7}
ADK	chr10:75934581-75934840	-1.1	1.88×10^{-6}
TMEM214	chr2:27253761-27253980	-1.13	3.31×10^{-7}
EIF2B3	chr1:45452441-45452640	-1.14	2.67×10^{-7}
VRK3	chr19:50527041-50527300	-1.16	2.26×10^{-6}
ZNF560	chr19:9609581-9609800	-1.21	9.9×10^{-7}
TRAF1	chr9:123691681-123691960	-1.21	3.51×10^{-5}
POTEB	chr15:21070661-21071020	-1.22	7.61×10^{-5}

Known DM vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
TPD52L2	chr20:62497561-62497920	-1	5.33×10^{-6}
GAGE7	chrX:49217161-49217580	-1.15	2.95×10^{-5}
NUDT10	chrX:51075781-51076040	-1.37	3.93×10^{-7}
OPN1MW2	chrX:153446941-153447140	-1.39	1.64×10^{-5}
OPN1MW	chrX:153446941-153447140	-1.39	1.64×10^{-5}
BRDT	chr1:92415321-92415520	-1.39	7.79×10^{-5}
ELAC2	chr17:12919641-12919840	-1.45	3.08×10^{-5}
SLC25A35	chr17:8196461-8196680	-1.52	7.49×10^{-6}
C18orf8	chr18:21081741-21081940	-1.55	2.33×10^{-6}
SLC28A1	chr15:85429461-85429660	-1.67	1.78×10^{-5}

Known DM vs Prediabetes

Gene Name	Genomic Coordinates	log2FC	p-value
COL6A6	chr3:130280381-130280600	-1.02	1.43×10^{-6}
CCPG1	chr15:55698761-55699200	-1.03	6.89×10^{-7}
ARHGEF35	chr7:143890801-143891200	-1.03	2.16×10^{-5}
GUK1	chr1:228326721-228326980	-1.07	5.97×10^{-6}
RNF187	chr1:228673261-228673560	-1.15	1.05×10^{-5}
WRB	chr21:40750961-40751200	-1.16	5.57×10^{-7}
MTM1	chrX:149737281-149737620	-1.17	2.2×10^{-5}
GNA14	chr9:80261681-80261900	-1.17	1.6×10^{-6}
SNAPC1	chr14:62227581-62227800	-1.21	9.45×10^{-5}
RPS7	chr2:3621621-3621820	-1.23	1.95×10^{-5}

Screen-detected DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
TSPAN10	chr17:79605201-79605400	-1	4.62×10^{-5}
TFE3	chrX:48900201-48900480	-1	6.62×10^{-6}
POTEB	chr15:21070661-21070880	-1.01	2.68×10^{-5}
POTEB2	chr15:21070661-21070880	-1.01	2.68×10^{-5}
AQP12B	chr2:241624101-241624300	-1.02	2.03×10^{-5}
RNF169	chr11:74461501-74461760	-1.05	1.33×10^{-6}
CCDC53	chr12:102454161-102454440	-1.05	1.04×10^{-6}
SLC35B4	chr7:134002221-134002500	-1.06	3.1×10^{-6}
SLC10A3	chrX:153719261-153719600	-1.06	3.38×10^{-6}
CIPC	chr14:77562681-77562960	-1.06	2.06×10^{-7}

Prediabetes vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
ZNF556	chr19:2867381-2867620	-1.02	5.88×10^{-7}
EFCC1	chr3:128721801-128722040	-1.09	4.52×10^{-5}
SLC10A3	chrX:153719261-153719600	-1.1	1.1×10^{-5}
RPS13	chr11:17100501-17100760	-1.15	9.31×10^{-5}
TMUB1	chr7:150781381-150781700	-1.2	9.31×10^{-7}
HAUS1	chr18:43685801-43686080	-1.22	5.32×10^{-5}
WFIKKN2	chr17:48913101-48913380	-1.33	9.4×10^{-5}
SPACA5B	chrX:47867301-47867560	-1.35	5.58×10^{-5}
ERICH6	chr3:150419901-150420200	-1.36	6.96×10^{-7}
POTEB	chr15:21070661-21070880	-1.44	2.72×10^{-5}

Prediabetes vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
H2AFB2	chrX:154609901-154610540	-1	1.58×10^{-5}
MIIP	chr1:12080981-12081260	-1.09	5.18×10^{-5}
DMTN	chr8:21910161-21910540	-1.12	5.72×10^{-5}
SLC41A3	chr3:125777161-125777520	-1.14	8.29×10^{-8}
C1orf167	chr1:11823621-11823820	-1.23	5.25×10^{-6}
SERPINF2	chr17:1647601-1647900	-1.25	9.84×10^{-6}
CXorf40B	chrX:149107901-149108120	-1.25	3.29×10^{-5}
TRMT44	chr4:8443121-8443380	-1.26	4.56×10^{-7}
SYNC	chr1:33169061-33169380	-1.31	3.89×10^{-5}
C9orf152	chr9:112970201-112970480	-1.31	4.22×10^{-6}

Table 3.4: Hypermethylated and hypomethylated DMRs in all diabetic individuals (known diabetes and screen-detected diabetes) versus NGT

Gene Name	Accession number	Known DM vs NGT fold change	p-value	Screen-detected DM vs NGT fold change	p-value
Hypermethylated DMRs					
BAGE2	NM_182482	1.44	1.64×10^{-9}	1.18	1.48×10^{-5}
BAGE3	NM_182481	1.44	1.64×10^{-9}	1.18	1.48×10^{-5}
BAGE4	NM_181704	1.44	1.64×10^{-9}	1.18	1.48×10^{-5}
BAGE5	NM_182484	1.44	1.64×10^{-9}	1.18	1.48×10^{-5}
CD248	NM_020404	1.56	6.86×10^{-6}	1.47	3.22×10^{-5}
COL8A2	NM_005202	1.04	3.65×10^{-5}	1.02	2.75×10^{-5}
IGSF9	NM_020789	1.53	9.07×10^{-7}	1.54	5.45×10^{-6}
PACSIN1	NM_020804	1.66	1.75×10^{-5}	1.39	1.27×10^{-8}
SPACA3	NM_173847	1.36	8.65 x 10-^{7}	1.73	1.31×10^{-5}
SYT3	NM_032298	1.49	1.49×10^{-9}	1.39	3.52×10^{-8}
TMEM89	NM_001008269	1.01	2.02×10^{-5}	1.04	3.14×10^{-6}
Hypomethylated DMRs					
ADK	NM_001202449	-1.1	1.87×10^{-6}	-1.32	2.55×10^{-6}
CCDC53	NM_016053	-1.28	1.52×10^{-6}	-1.05	1.03×10^{-6}
POTEB	NM_001277304	-1.22	7.61×10^{-5}	-1.01	2.68×10^{-5}
POTEB2	NM_001277303	-1.22	7.61×10^{-5}	-1.01	2.68×10^{-5}
RHOBTB3	NM_014899	-2.02	1.93×10^{-7}	-1.82	2.69×10^{-7}
RNF169	NM_001098638	-1.23	6.75×10^{-7}	-1.05	1.33×10^{-6}
SLC35B4	NM_032826	-1.45	2.01×10^{-5}	-1.06	3.10×10^{-6}

Table 3.5: Hypermethylated and hypomethylated DMRs in known diabetes, screen-detected diabetes and prediabetes versus NGT

Gene Name	Accession number	Known DM vs NGT fold change	p-value	Screen-detected DM vs NGT fold change	p-value	Prediabetes vs NGT fold change	p-value
Hypermethylated DMRs							
TMEM89	NM_001008269	1.01	2.02×10^{-5}	1.04	3.14×10^{-6}	1.76	2.60×10^{-7}
Hypomethylated DMRs							
CCDC53	NM_016053	-1.28	1.52×10^{-6}	-1.05	1.03×10^{-6}	-1.61	4.91×10^{-7}
POTEB	NM_001277304	-1.22	7.61×10^{-5}	-1.01	2.68×10^{-5}	-1.44	2.72×10^{-5}
POTEB2	NM_001277303	-1.22	7.61×10^{-5}	-1.01	2.68×10^{-5}	-1.44	2.72×10^{-5}
RHOBTB3	NM_014899	-2.02	1.93×10^{-7}	-1.82	2.69×10^{-7}	-1.88	2.24×10^{-7}

Table 3.6: Hypermethylated and hypomethylated DMRs in known diabetes versus screen-detected diabetes

Gene Name	Accession number	Known DM vs Screen-detected DM fold change	p-value
Hypermethylated DMRs			
XAGE1E	NM_001097604	1.66	5.36×10^{-7}
XAGE1B	NM_001097594	1.66	5.36×10^{-7}
KIAA1467	NM_020853	1.62	2.21×10^{-7}
ASB2	NM_001202429	1.6	3.33×10^{-7}
GABPA	NM_001197297	1.56	6.07×10^{-6}
ZNF346	NM_012279	1.47	6.99×10^{-6}
FKBP8	NM_012181	1.4	3.83×10^{-7}
CTAGE15	NM_001008747	1.38	2.62×10^{-7}
VIPR1	NM_001251882	1.37	9.82×10^{-5}
TMEM204	NM_024600	1.31	2.77×10^{-6}
RNF103-CHMP3	NM_001198954	1.31	2.43×10^{-5}
PARVB	NM_001003828	1.26	6.86×10^{-6}
POTED	NM_174981	1.23	5.31×10^{-7}
STAG2	NM_006603	1.21	1.9×10^{-5}
KCNQ3	NM_001204824	1.19	9.67×10^{-5}
TBCE	NM_003193	1.16	7.23×10^{-7}
GAREML	NM_001168241	1.14	3.93×10^{-6}
Sep-12	NM_001154458	1.13	4.44×10^{-6}
OR6C3	NM_054104	1.11	2.94×10^{-5}
PPP1R32	NM_001170753	1.1	4.38×10^{-6}
ZNF169	NM_194320	1.07	8.86×10^{-5}
TAS1R1	NM_177540	1.07	2.86×10^{-7}
Hypomethylated DMRs			
TPD52L2	NM_001243895	-1	5.33×10^{-6}
GAGE7	NM_021123	-1.15	2.95×10^{-5}
NUDT10	NM_153183	-1.37	3.93×10^{-7}
OPN1MW2	NM_001048181	-1.39	1.64×10^{-5}
OPN1MW	NM_000513	-1.39	1.64×10^{-5}
BRDT	NM_001242808	-1.39	7.79×10^{-5}
ELAC2	NM_001165962	-1.45	3.08×10^{-5}
SLC25A35	NM_201520	-1.52	7.49×10^{-6}
C18orf8	NM_013326	-1.55	2.33×10^{-6}
SLC28A1	NM_001287761	-1.67	1.78×10^{-5}
FBXW8	NM_153348	-1.79	1.7×10^{-5}

3.3 Gene Ontology (GO) analysis

Gene Ontology analysis was performed and enabled the classification of the DMRs according to three domains, namely biological process, cellular component and molecular function. The GO analyses were completed for each of the six comparison pairs and highlighted all three domains for both the hypermethylated and hypomethylated DMRs. Fisher's exact test was used to identify whether there was more overlap between the differentially expressed (DE) list and the GO annotation list than would be expected by chance. The p-value represents the significance of GO terms enrichment in the DE genes. A p-value of <= 0.05 was considered significant and the lower the p-value, the more significant the GO Term.

Furthermore, GO Directed Acyclic Graphs (DAG) were formulated and displayed as p-value trees. The gene annotation moves from more general to more specific as one moves from the parent (precursor) nodes to child nodes. The top ten terms with the lowest p-value and their parents are shown in the p-value trees. The terms with pane marks are significant enrichment and the more red means more significant. Any p-values lower than 1 $x10^{-20}$ will be displayed as $1x10^{-20}$ in the p-value tree.

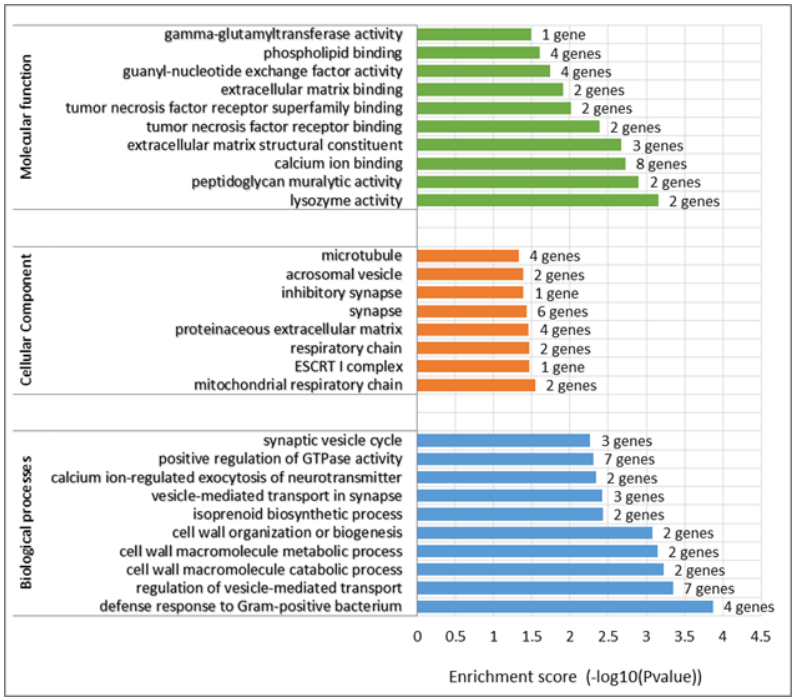
3.3.1 Known diabetes versus NGT

The top 10 GO enriched terms for the known diabetic group versus the NGT group categorised into biological processes, molecular functions and cellular components are illustrated in Figure 3.3 and 3.4. Defense response to Gram-positive bacterium (biological process), mitochondrial respiratory chain (cellular component) and lysozyme activity (molecular function) have the highest-ranked enrichment score amongst the hypermethylated DMRs (Figure 3.3a). On the other hand, regulation of translational initiation (biological process), dihydrolipoyl dehydrogenase complex (cellular component) and translation initiation factor activity (molecular function) have the highest-ranked enrichment score among the hypomethylated DMRs (Figure 3.4a).

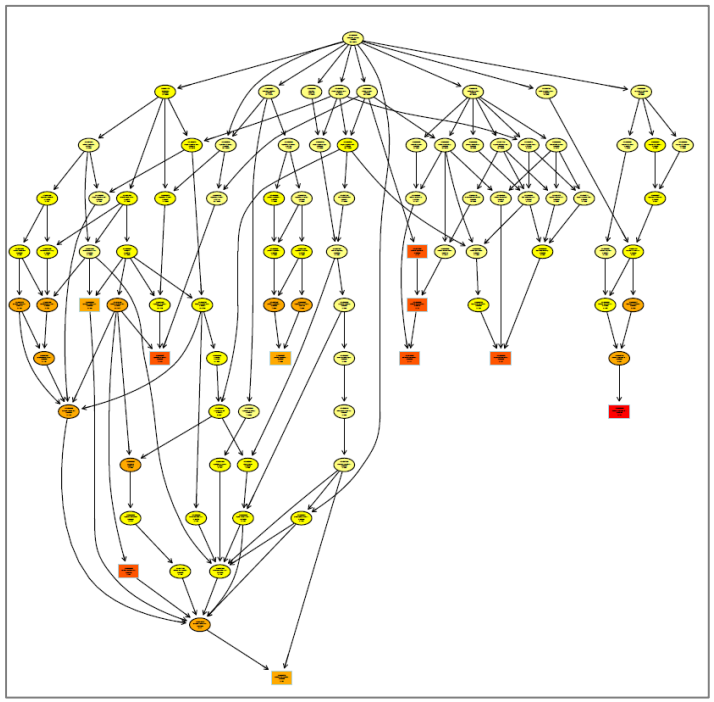
P-value trees showing the hierarchy of gene ontology biological processes, cellular components and molecular functions are displayed for both hypermethylated (Figure 3.3 b-d) and hypomethylated (Figure 3.4 b-d) enriched terms. The most significant enrichment (lowest p-value) is displayed by the red blocks which show the biological process of defense response

to Gram-positive bacterium (hypermethylated) (Figure 3.3b) as well as regulation of translational initiation and regulation to cell shape (hypomethylated) (Figure 3.4b). The significant enrichment for cellular components includes the mitochondrial respiratory chain, ESCRT1 complex, acrosomal vesicle and respiratory chain (hypermethylated) (Figure 3.3c) as well as dihydrolipoyl dehydrogenase complex, cell−cell adherens junction, site of the double−strand break, cytosol, cell−cell junction, sperm principal piece, tricarboxylic acid cycle enzyme complex and WASH complex (hypomethylated) (Figure 3.4c). Additionally, the significant enrichment for molecular functions includes lysozyme activity, peptidoglycan muralytic activity, calcium ion binding and extracellular matrix structural constituent (hypermethylated) (Figure 3.3d) as well as translation initiation factor activity and translation factor activity, RNA binding (hypomethylation) (Figure 3.4d).

(a)


Molecular function
gamma-glutamyltransferase activity 1 gene
phospholipid binding 4 genes
guanyl-nucleotide exchange factor activity 4 genes
extracellular matrix binding 2 genes
tumor necrosis factor receptor superfamily binding 2 genes
tumor necrosis factor receptor binding 2 genes
extracellular matrix structural constituent 3 genes
calcium ion binding 8 genes
peptidoglycan muralytic activity 2 genes
lysozyme activity 2 genes
Cellular Component
microtubule 4 genes
acrosomal vesicle 2 genes
inhibitory synapse 1 gene
synapse 6 genes
proteinaceous extracellular matrix 4 genes
respiratory chain 2 genes
ESCRT I complex 1 gene
mitochondrial respiratory chain 2 genes
Biological processes
synaptic vesicle cycle 3 genes
positive regulation of GTPase activity 7 genes
calcium ion-regulated exocytosis of neurotransmitter 2 genes
vesicle-mediated transport in synapse 3 genes
isoprenoid biosynthetic process 2 genes
cell wall organization or biogenesis 2 genes
cell wall macromolecule metabolic process 2 genes
cell wall macromolecule catabolic process 2 genes
regulation of vesicle-mediated transport 7 genes
defense response to Gram-positive bacterium 4 genes
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Enrichment score (-log10(Pvalue))

(b)



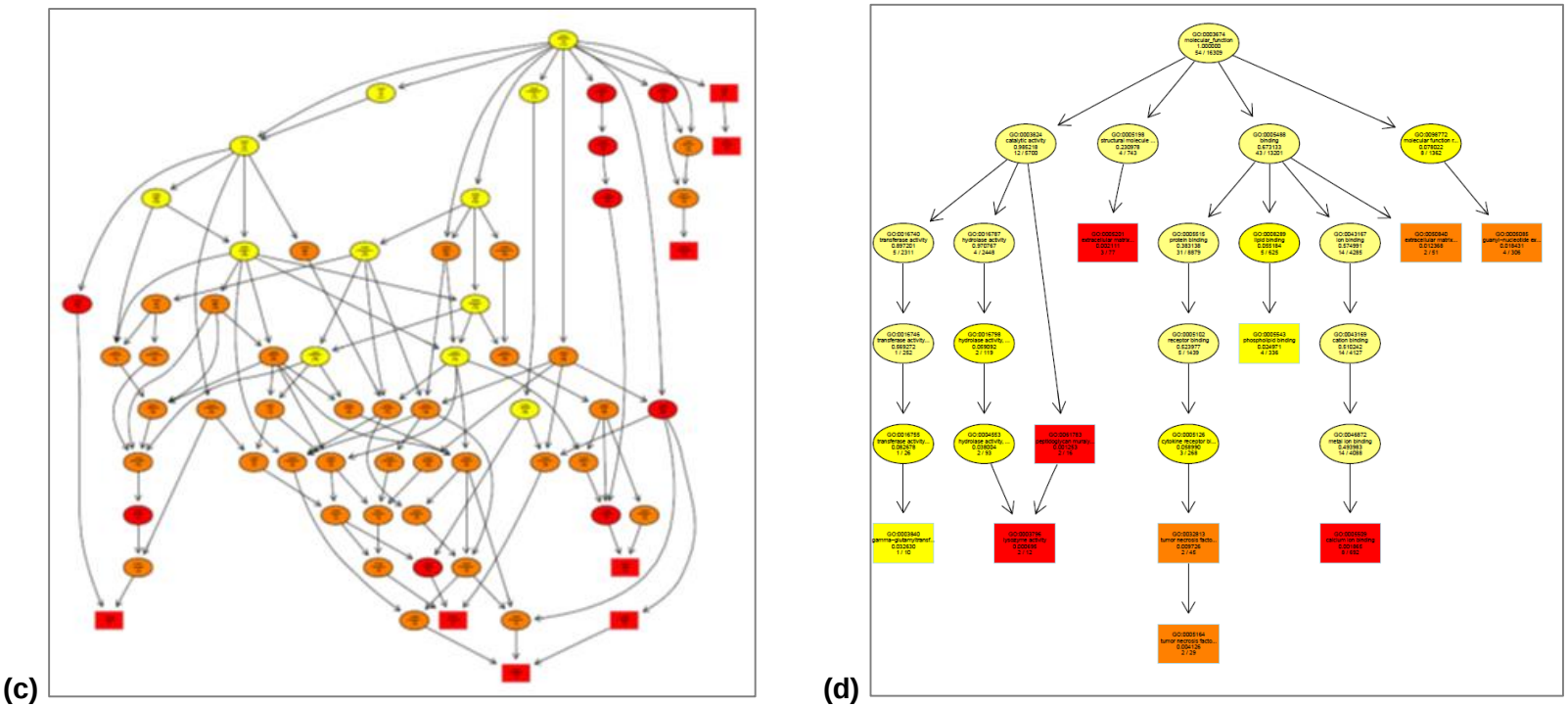
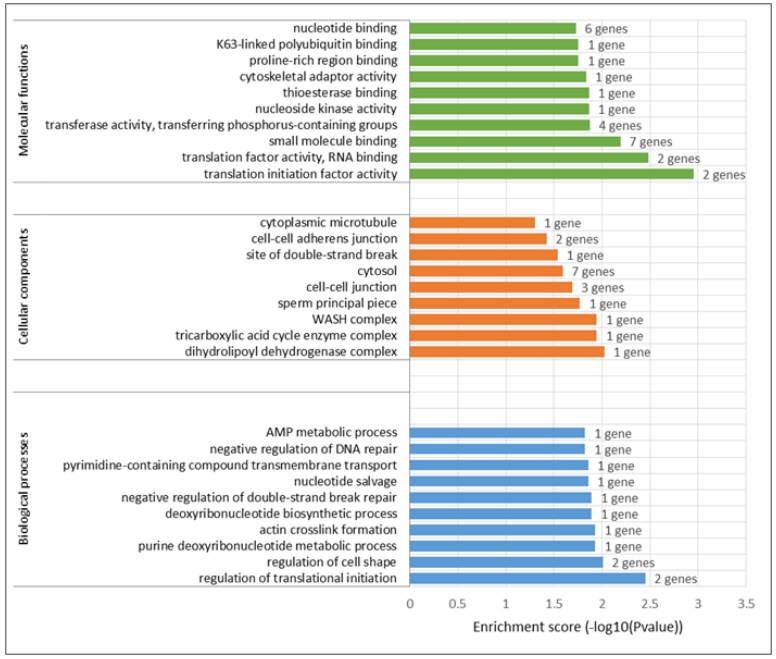
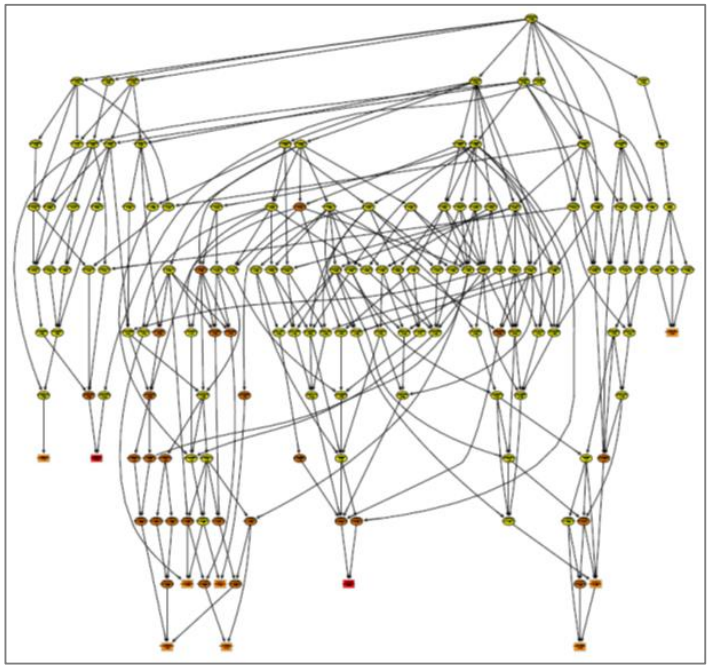


Figure 3.3: Gene Ontology (GO) enrichment analysis of the hypermethylated differentially methylated genes in known diabetics compared to the normoglycaemic control group (NGT). (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

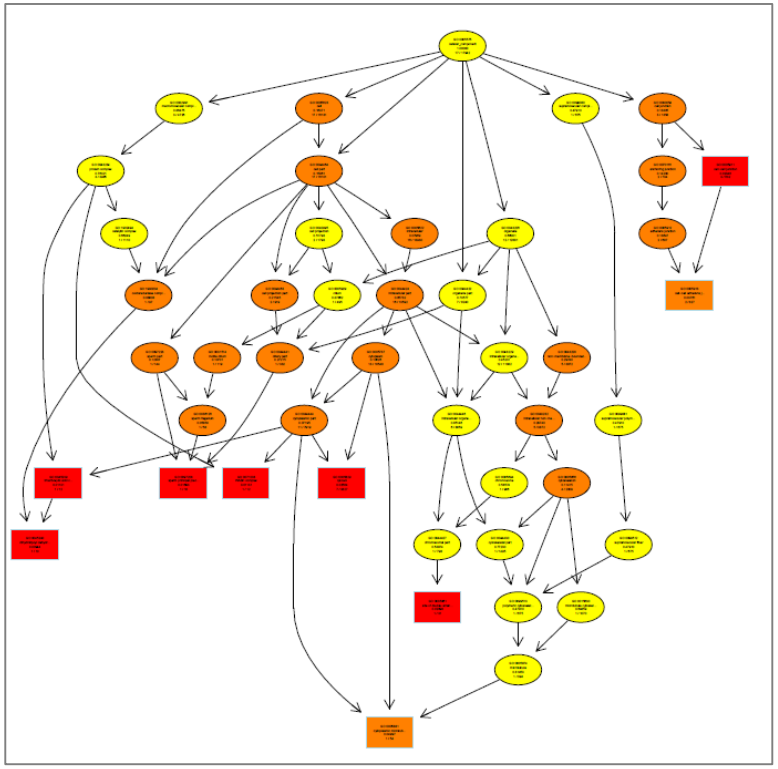

Molecular functions
nucleotide binding 6 genes
K63-linked polyubiquitin binding 1 gene
proline-rich region binding 1 gene
cytoskeletal adaptor activity 1 gene
thioesterase binding 1 gene
nucleoside kinase activity 1 gene
transferase activity, transferring phosphorus-containing groups 4 genes
small molecule binding 7 genes
translation factor activity, RNA binding 2 genes
translation initiation factor activity 2 genes
Cellular components
cytoplasmic microtubule 1 gene
cell-cell adherens junction 2 genes
site of double-strand break 1 gene
cytosol 7 genes
cell-cell junction 3 genes
sperm principal piece 1 gene
WASH complex 1 gene
tricarboxylic acid cycle enzyme complex 1 gene
dihydrolipoyl dehydrogenase complex 1 gene
Biological processes
AMP metabolic process 1 gene
negative regulation of DNA repair 1 gene
pyrimidine-containing compound transmembrane transport 1 gene
nucleotide salvage 1 gene
negative regulation of double-strand break repair 1 gene
deoxyribonucleotide biosynthetic process 1 gene
actin crosslink formation 1 gene
purine deoxyribonucleotide metabolic process 1 gene
regulation of cell shape 2 genes
regulation of translational initiation 2 genes
0 0.5 1 1.5 2 2.5 3 3.5
Enrichment score (-log10(Pvalue))

(a)



(b)

(c)



(d)

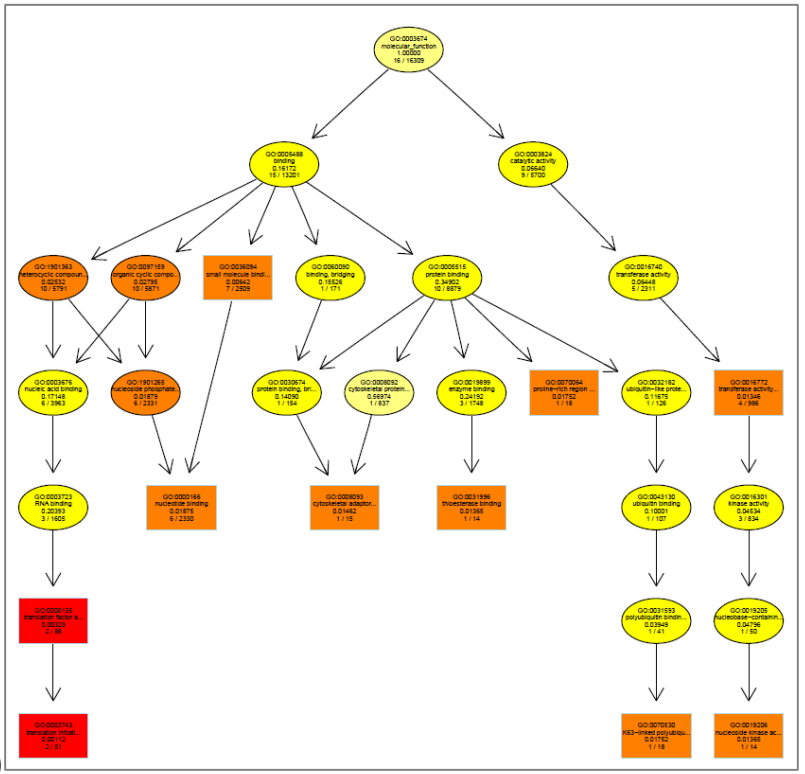


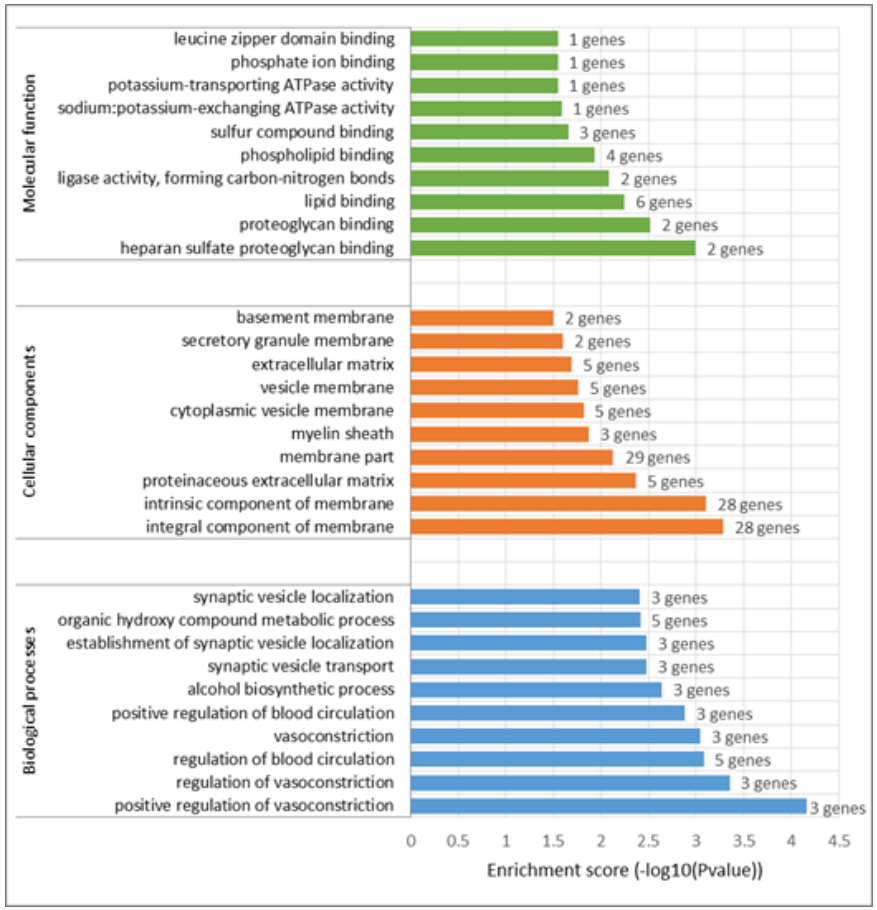
Figure 3.4: Gene Ontology (GO) enrichment analysis of the hypomethylated differentially methylated genes in known diabetics compared to the normoglycaemic control group (NGT). (a)The bars plot shows the top ten enrichment score values of the significant enrichment terms categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

3.3.2 Screen-detected diabetes versus NGT

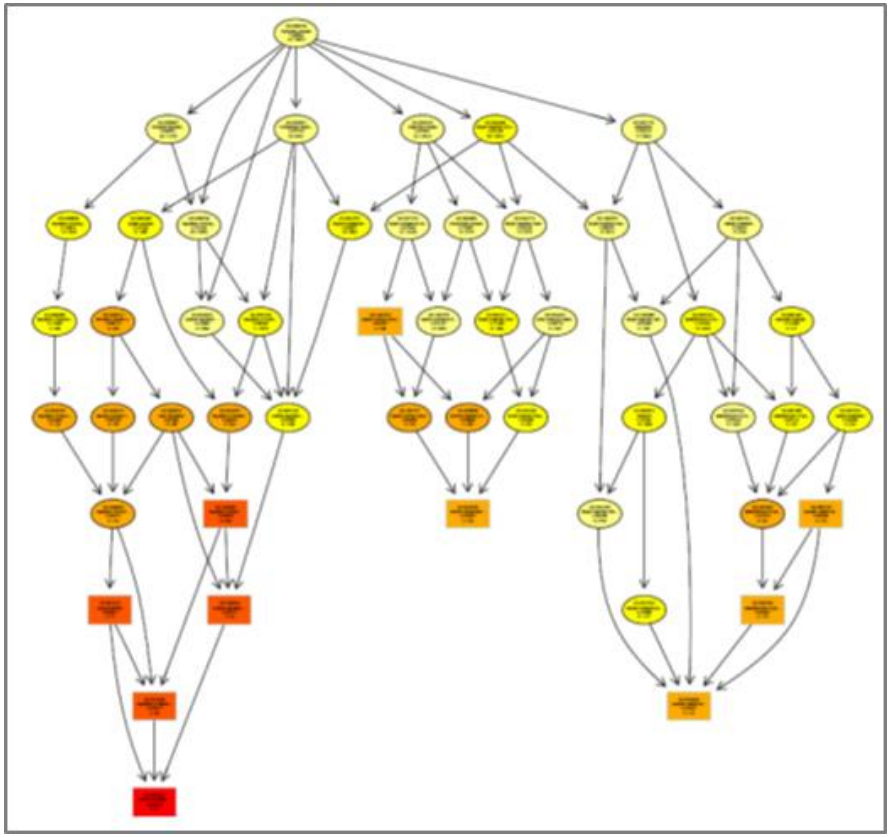
The top 10 GO enriched terms for the screen-detected diabetic group versus the NGT group categorised into biological processes, molecular functions and cellular components are illustrated in Figure 3.5 and 3.6. Positive regulation of vasoconstriction (biological process), integral component of membrane (cellular component) and heparin sulphate proteoglycan binding (molecular function) have the highest-ranked enrichment score amongst the hypermethylated DMRs (Figure 3.5a). On the other hand, pyrimidine-containing compound transmembrane transport (biological process), integral component of endoplasmic reticulum membrane (cellular component) and endoribonuclease activity, producing 3'-phosphomonoesters (molecular function) has the highest-ranked enrichment score among the hypomethylated DMRs (Figure 3.6a).

The hierarchy of gene ontology biological processes, cellular components and molecular functions are displayed using p-value trees for both the hypermethylated (Figure 3.5 b-d) and hypomethylated (Figure 3.6 b-d) enriched terms. The most significant enrichment (lowest p-value) is displayed by the red blocks which show the biological process of positive regulation of vasoconstriction (hypermethylated) (Figure3.5b) and pyrimidine-containing compound transmembrane transport (hypomethylated) (Figure 3.6b). The significant enrichment for cellular components includes the integral component of membrane and intrinsic component of membrane (hypermethylated) (Figure 3.5c) as well as integral component of endoplasmic reticulum membrane, intrinsic component of endoplasmic reticulum membrane, WASH complex and DNA-directed RNA polymerase III complex (hypomethylated) (Figure 3.6c). The significant enrichment for molecular functions include heparin sulphate proteoglycan binding and proteoglycan binding (hypermethylated) (Figure 3.5d) as well as endoribonuclease activity, producing 3'-phosphomonoesters, nucleoside kinase activity, bile acid transmembrane transporter activity, water channel activity, water transmembrane transporter activity, RNA polymerase III activity, macrolide binding, FK506 binding and K63-linked polyubiquitin binding (hypomethylation) (Figure 3.6d).

(a)


Molecular function
leucine zipper domain binding 1 genes
phosphate ion binding 1 genes
potassium-transporting ATPase activity 1 genes
sodium:potassium-exchanging ATPase activity 1 genes
sulfur compound binding 3 genes
phospholipid binding 4 genes
ligase activity, forming carbon-nitrogen bonds 2 genes
lipid binding 6 genes
proteoglycan binding 2 genes
heparan sulfate proteoglycan binding 2 genes
Cellular components
basement membrane 2 genes
secretory granule membrane 2 genes
extracellular matrix 5 genes
vesicle membrane 5 genes
cytoplasmic vesicle membrane 5 genes
myelin sheath 3 genes
membrane part 29 genes
proteinaceous extracellular matrix 5 genes
intrinsic component of membrane 28 genes
integral component of membrane 28 genes
Biological processes
synaptic vesicle localization 3 genes
organic hydroxy compound metabolic process 5 genes
establishment of synaptic vesicle localization 3 genes
synaptic vesicle transport 3 genes
alcohol biosynthetic process 3 genes
positive regulation of blood circulation 3 genes
vasoconstriction 3 genes
regulation of blood circulation 5 genes
regulation of vasoconstriction 3 genes
positive regulation of vasoconstriction 3 genes
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Enrichment score (-log10(Pvalue))

(b)



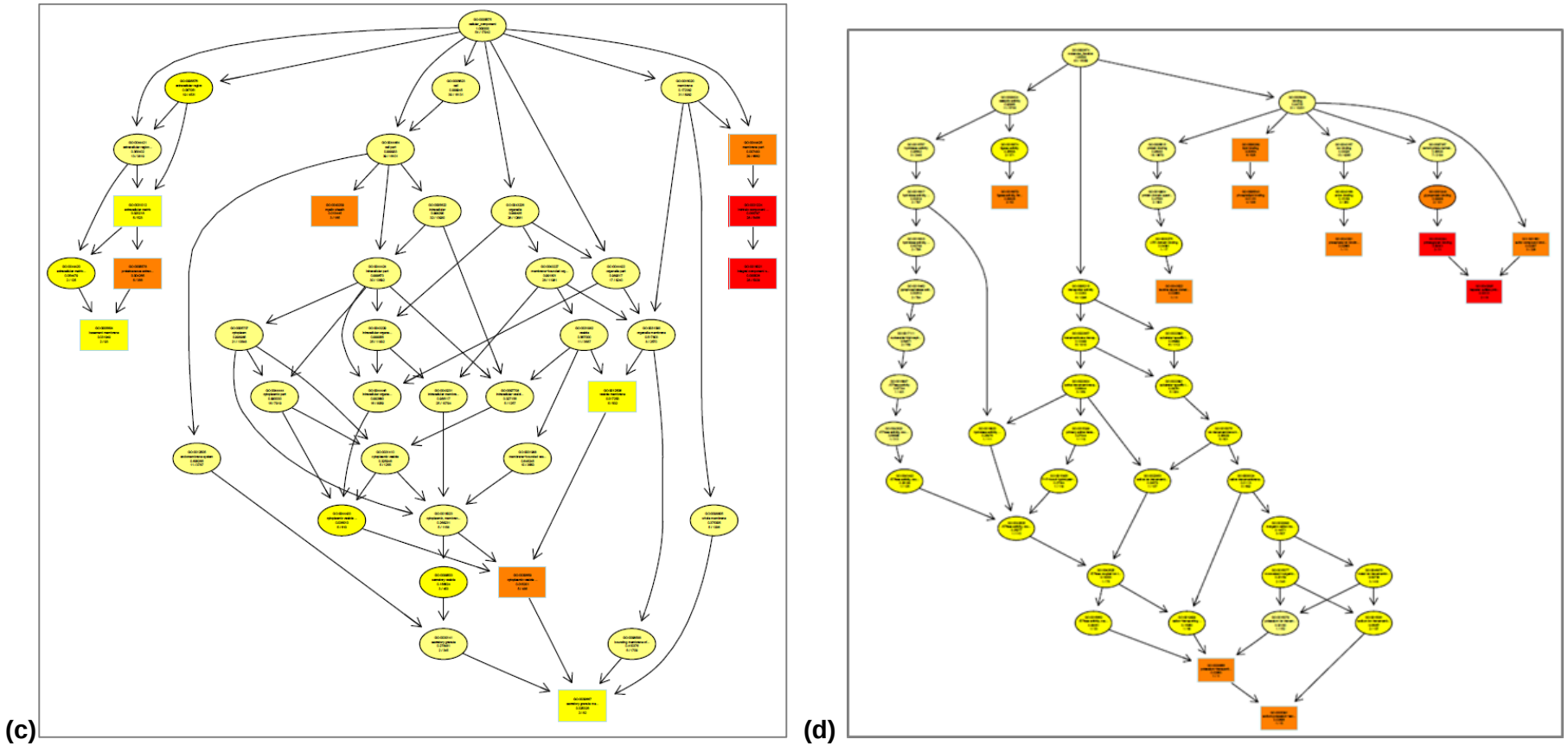
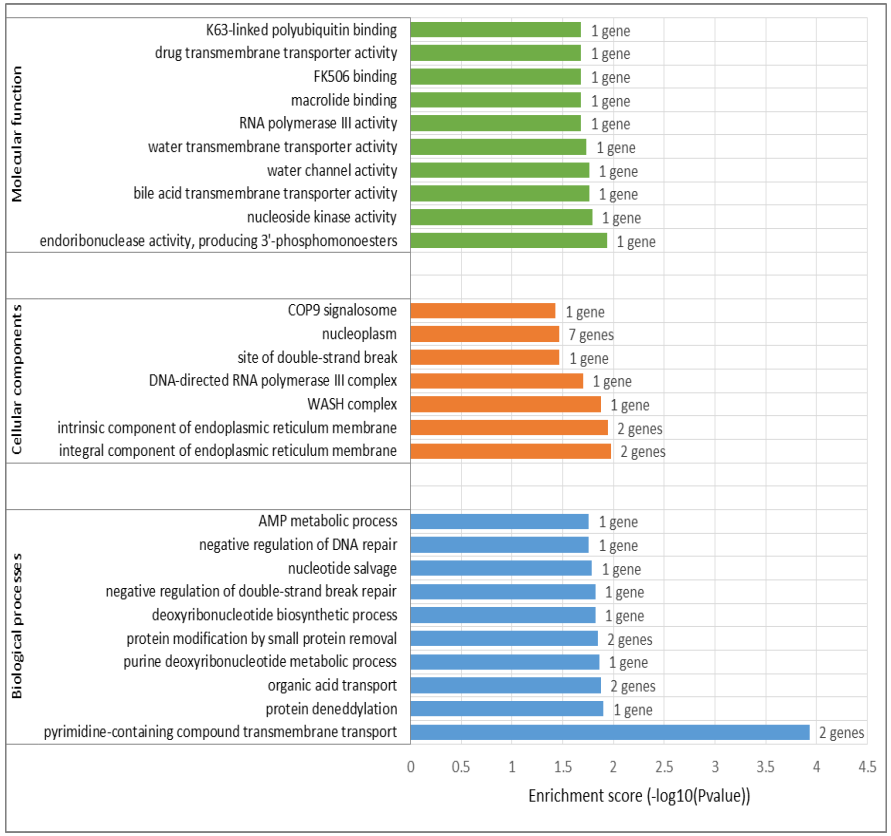
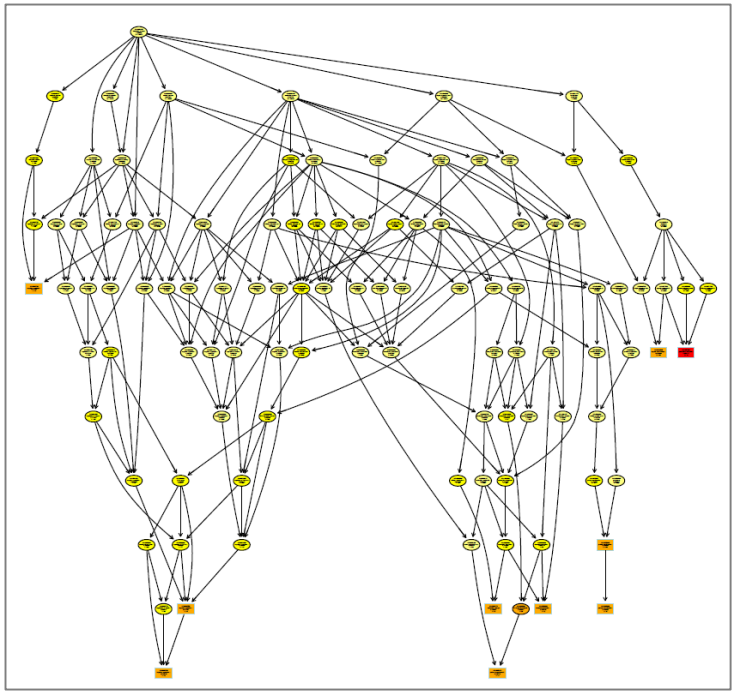


Figure 3.5: Gene Ontology (GO) enrichment analysis of the hypermethylated differentially methylated genes in screen-detected diabetics when compared to the normoglycaemic control group (NGT). (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

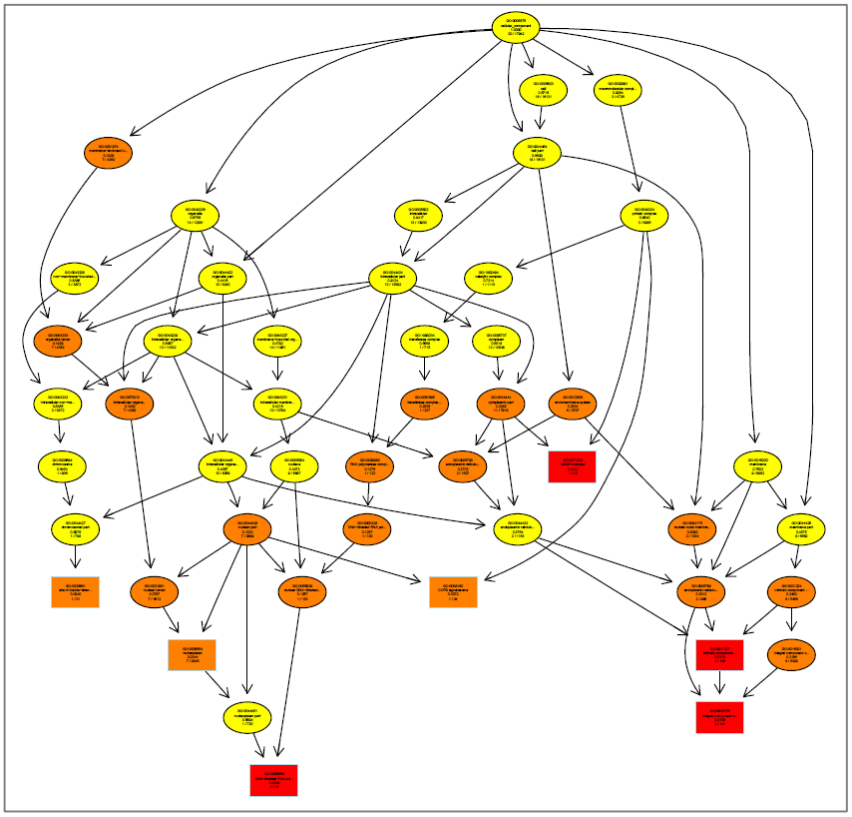
(a)


Molecular function
K63-linked polyubiquitin binding
drug transmembrane transporter activity
FK506 binding
macrolide binding
RNA polymerase III activity
water transmembrane transporter activity
water channel activity
bile acid transmembrane transporter activity
nucleoside kinase activity
endoribonuclease activity, producing 3'-phosphomonoesters
1 gene
Cellular components
COP9 signalosome
nucleoplasm
site of double-strand break
DNA-directed RNA polymerase III complex
WASH complex
intrinsic component of endoplasmic reticulum membrane
integral component of endoplasmic reticulum membrane
7 genes
2 genes
Biological processes
AMP metabolic process
negative regulation of DNA repair
nucleotide salvage
negative regulation of double-strand break repair
deoxyribonucleotide biosynthetic process
protein modification by small protein removal
purine deoxyribonucleotide metabolic process
organic acid transport
protein deneddylation
pyrimidine-containing compound transmembrane transport
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
Enrichment score (-log10(Pvalue))

(b)



(c)



(d)

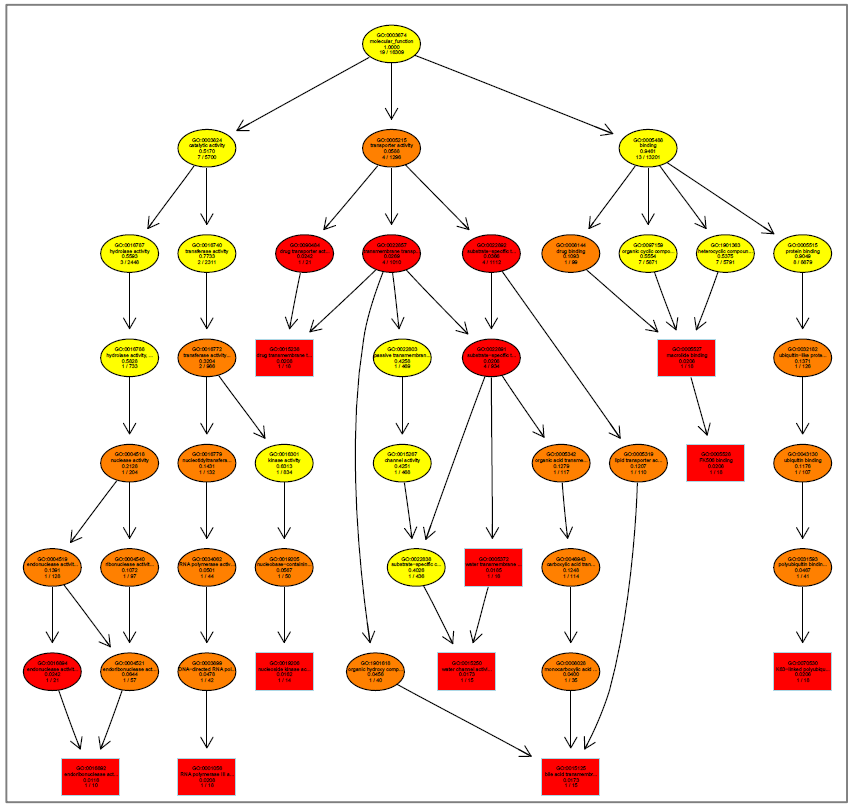


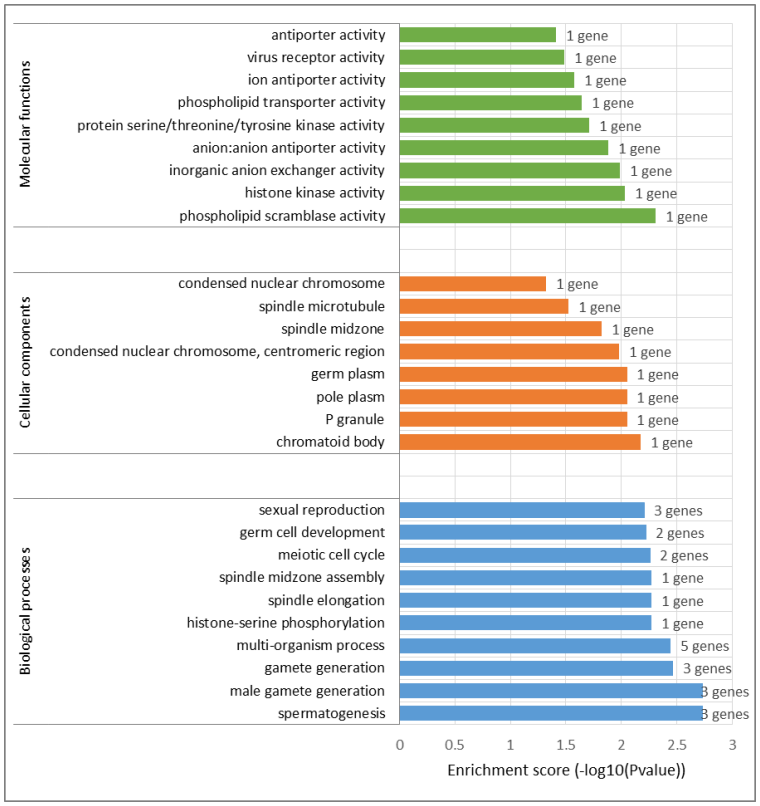
Figure 3.6: Gene Ontology (GO) enrichment analysis of the hypomethylated differentially methylated genes in screen-detected diabetics when compared to the normoglycaemic control group (NGT). (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

3.3.3 Prediabetes versus NGT

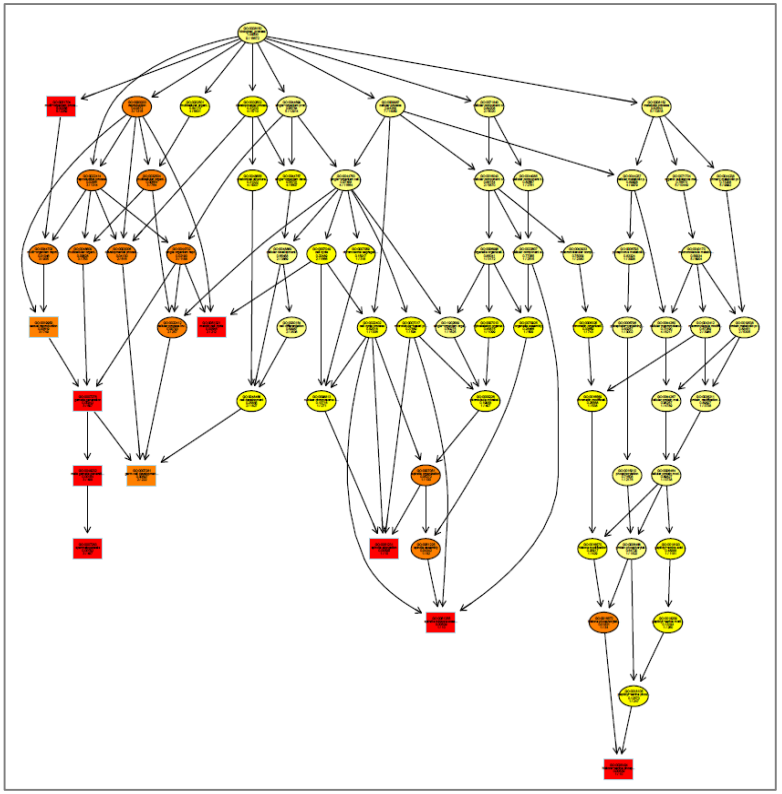
The top 10 GO enriched terms for the prediabetic group versus the NGT group categorised into biological processes, molecular functions and cellular components are illustrated in Figure 3.7 and 3.8. Spermatogenesis (biological process), chromatoid body (cellular component) and phospholipid scramblase activity (molecular function) have the highest-ranked enrichment score amongst the hypermethylated DMRs (Figure 3.7a). On the other hand, defense response to Gram-negative bacterium (biological process), WASH complex (cellular component) and histone demethylase activity (H3-K9 specific) (molecular function) has the highest-ranked enrichment score among the hypomethylated DMRs (Figure 3.8a).

P-value trees showing the hierarchy of gene ontology biological processes, cellular components and molecular functions are presented for both hypermethylated (Figure 3.7 b-d) and hypomethylated (Figure 3.8 b-d) enriched terms. The most significant enrichment (lowest p-value) is displayed by the red blocks which show the biological process of spermatogenesis, male gamete generation, gamete generation, multi-organism process, histone-serine phosphorylation, spindle elongation, spindle midzone assembly and meiotic cell cycle (hypermethylated) (Figure 3.7b) as well as defense response to Gram-negative bacterium and defense response to Gram-positive bacterium (hypomethylated) (Figure 3.8b). The significant enrichment for cellular components includes chromatoid body, P granule, pole plasm, germ plasm, condensed nuclear chromosome, centromeric region and spindle midzone (hypermethylated) (Figure 3.7c) as well as WASH complex (hypomethylated) (Figure 3.8c). The significant enrichment for molecular functions includes phospholipid scramblase activity, histone kinase activity, inorganic anion exchanger activity and anion:anion antiporters activity (hypermethylated) (Figure 3.7d) as well as histone demethylase activity (H3-K9 specific), lysozyme activity, metalloendopeptidase inhibitor activity, bile acid transmembrane transporter activity, peptidoglycan muralytic activity, lipopolysaccharide binding, organic acid:sodium symporter activity and histone demethylase activity (hypomethylation) (Figure 3.8d).

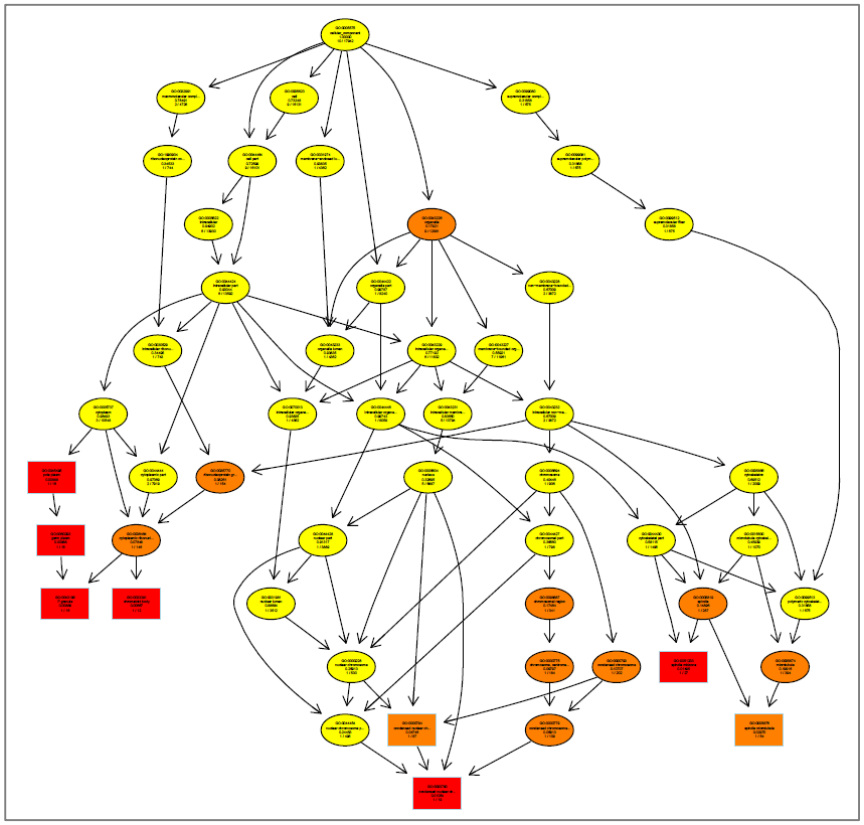
(a)


Molecular functions
antiporter activity
virus receptor activity
ion antiporter activity
phospholipid transporter activity
protein serine/threonine/tyrosine kinase activity
anion:anion antiporter activity
inorganic anion exchanger activity
histone kinase activity
phospholipid scramblase activity
1 gene
Cellular components
condensed nuclear chromosome
spindle microtubule
spindle midzone
condensed nuclear chromosome, centromeric region
germ plasm
pole plasm
P granule
chromatoid body
1 gene
Biological processes
sexual reproduction
germ cell development
meiotic cell cycle
spindle midzone assembly
spindle elongation
histone-serine phosphorylation
multi-organism process
gamete generation
male gamete generation
spermatogenesis
3 genes
2 genes
5 genes
0
0.5
1
1.5
2
2.5
3
Enrichment score (-log10(Pvalue))

(b)



(c)



(d)

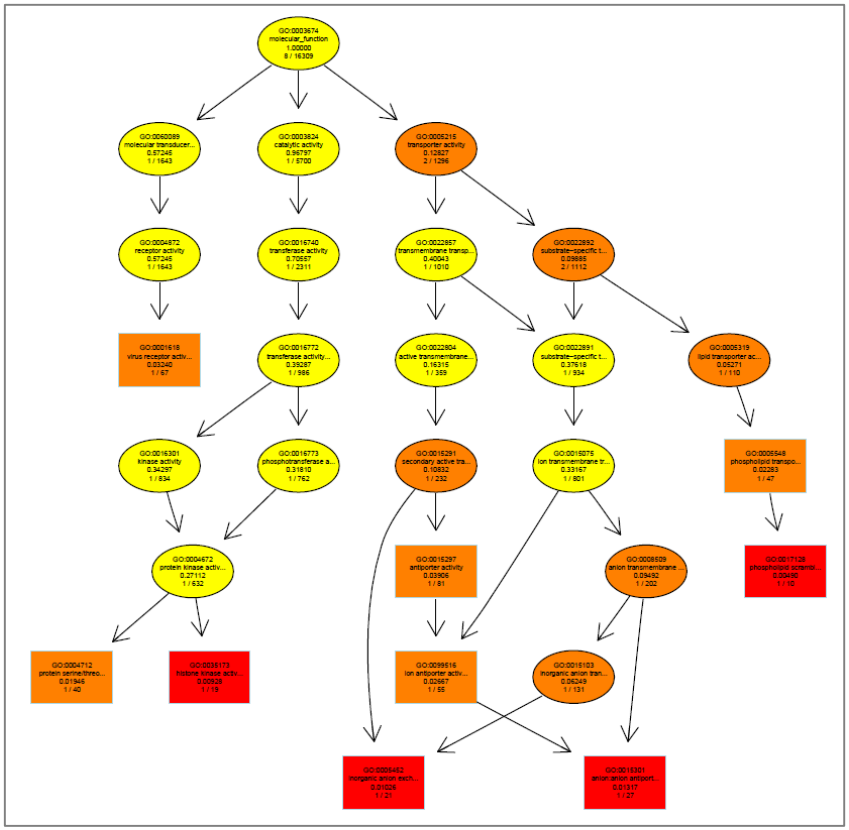
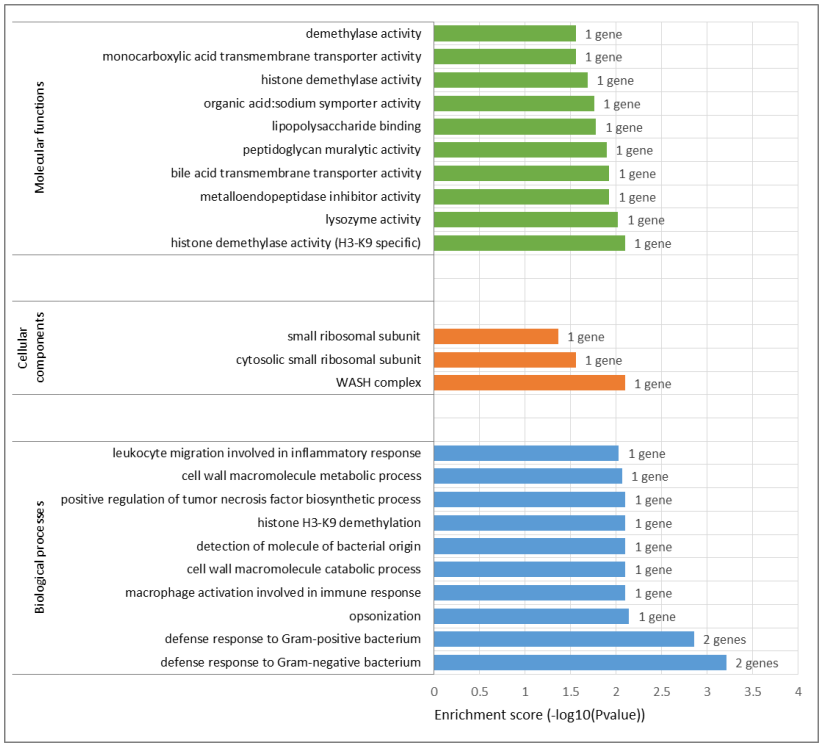
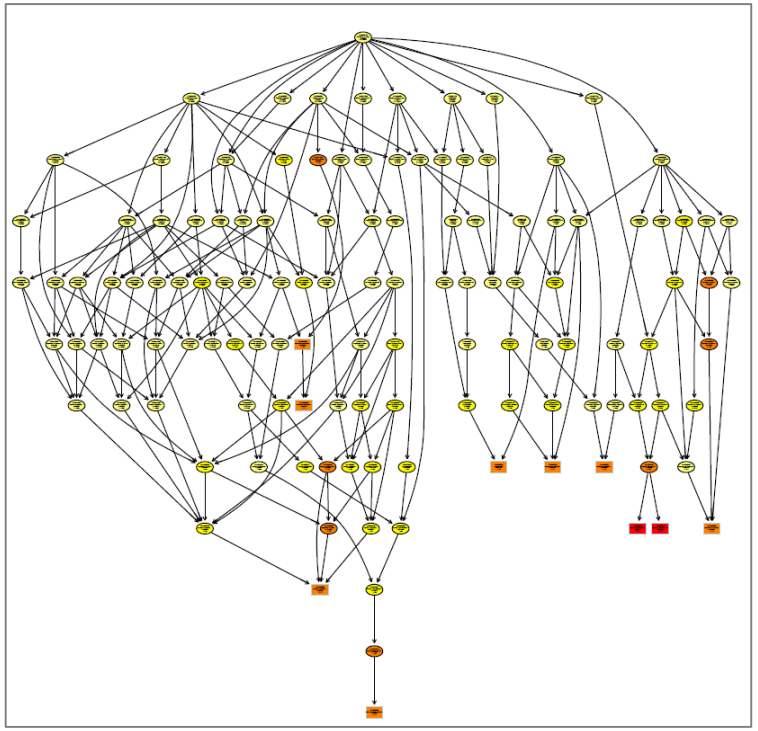


Figure 3.7: Gene Ontology (GO) enrichment analysis of the hypermethylated differentially methylated genes in prediabetes when compared to the normoglycaemic control group (NGT). (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

(a)


Molecular functions
demethylase activity
monocarboxylic acid transmembrane transporter activity
histone demethylase activity
organic acid:sodium symporter activity
lipopolysaccharide binding
peptidoglycan muralytic activity
bile acid transmembrane transporter activity
metalloendopeptidase inhibitor activity
lysozyme activity
histone demethylase activity (H3-K9 specific)
Cellular components
small ribosomal subunit
cytosolic small ribosomal subunit
WASH complex
Biological processes
leukocyte migration involved in inflammatory response
cell wall macromolecule metabolic process
positive regulation of tumor necrosis factor biosynthetic process
histone H3-K9 demethylation
detection of molecule of bacterial origin
cell wall macromolecule catabolic process
macrophage activation involved in immune response
opsonization
defense response to Gram-positive bacterium
defense response to Gram-negative bacterium
1 gene
2 genes
0
0.5
1
1.5
2
2.5
3
3.5
4
Enrichment score (-log10(Pvalue))

(b)



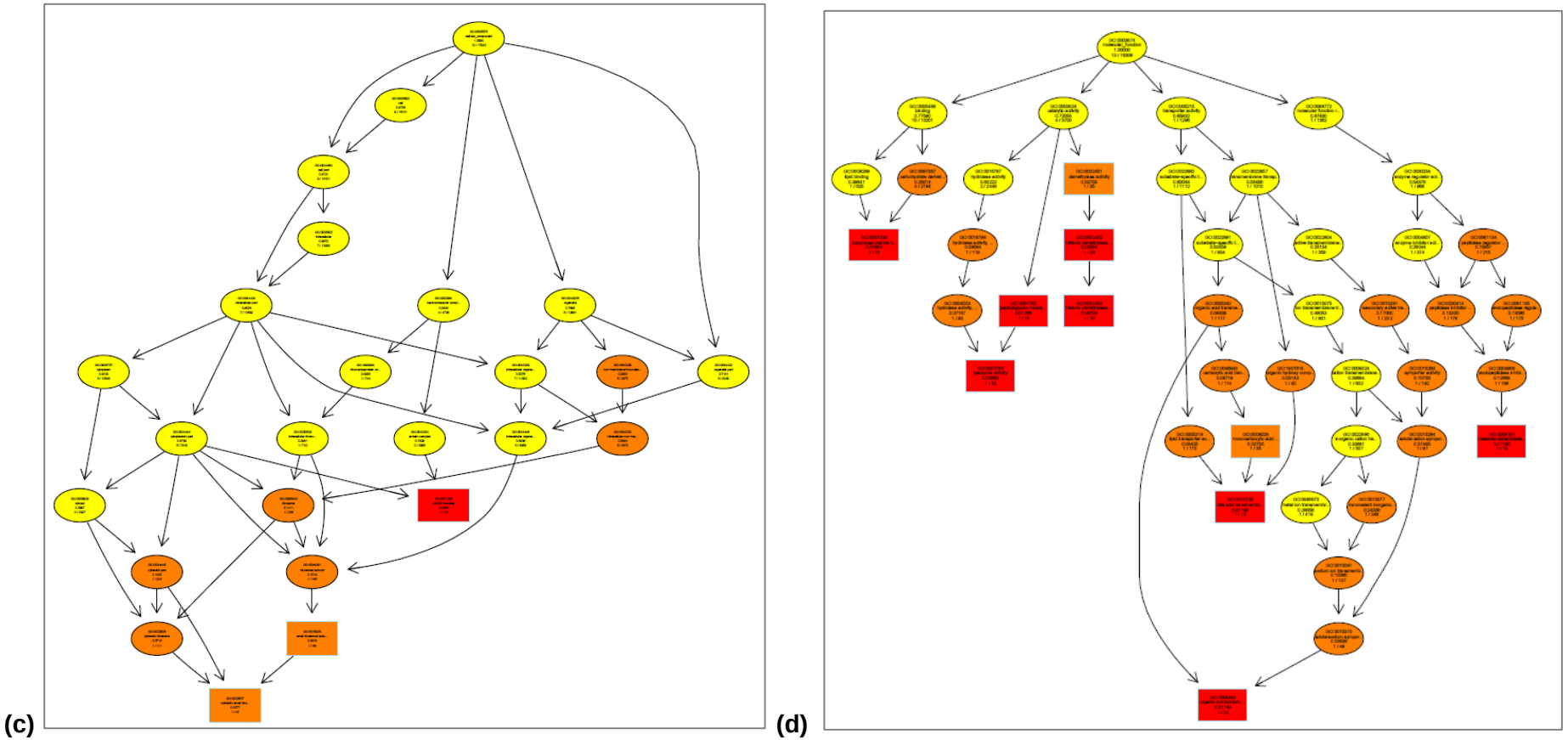


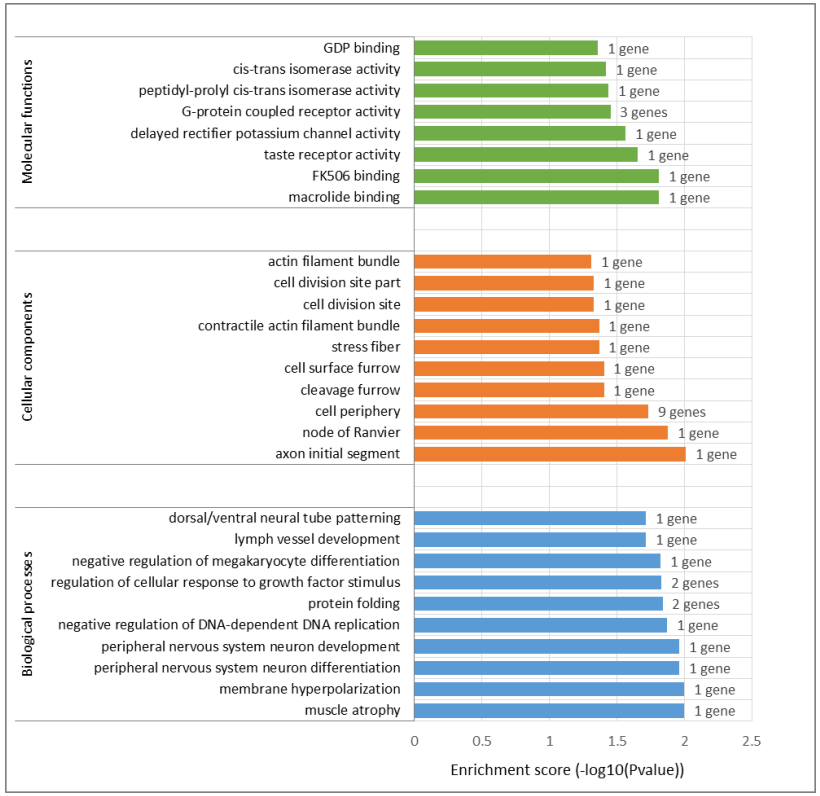
Figure 3.8: Gene Ontology (GO) enrichment analysis of the hypomethylated differentially methylated genes in prediabetes when compared to the normoglycaemic control group (NGT). (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

3.3.4 Known diabetes versus screen-detected diabetes

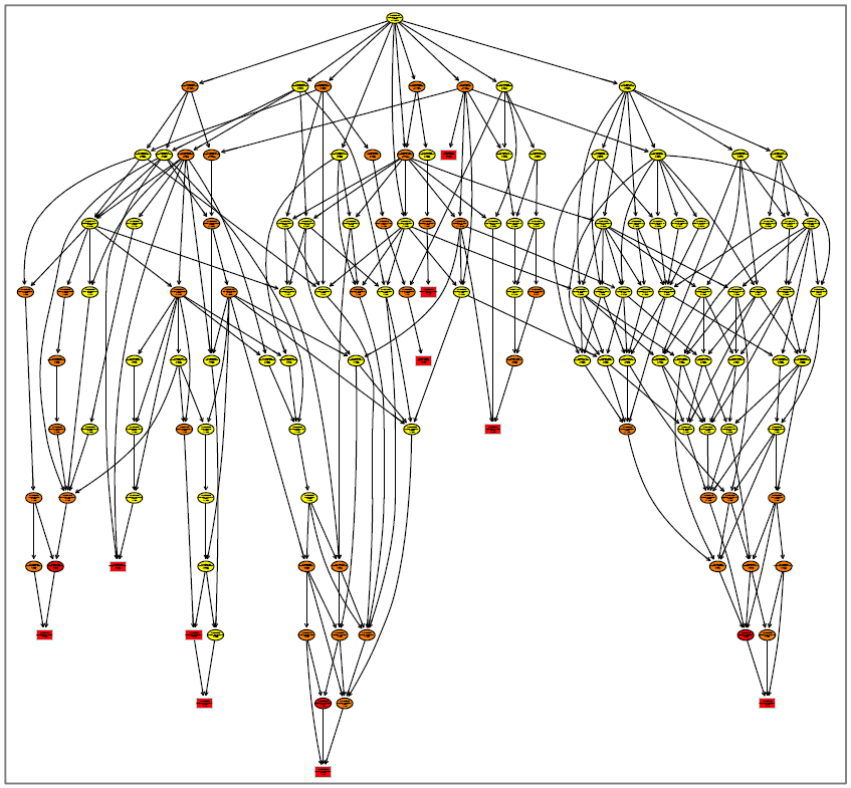
The top 10 GO enriched terms for the known diabetic group versus the screen-detected diabetic group categorised into biological processes, molecular functions and cellular components are illustrated in Figure 3.9 and 3.10. Muscle atrophy (biological process), axon initial segment (cellular component) and macrolide binding (molecular function) have the highest-ranked enrichment score amongst the hypermethylated DMRs (Figure 3.9a). On the other hand, protein-chromophore linkage (biological process), photoreceptor outer segment membrane (cellular component) and G-protein coupled photoreceptor activity (molecular function) have the highest-ranked enrichment score among the hypomethylated DMRs (Figure 3.10a).

The hierarchy of gene ontology biological processes, cellular components and molecular functions are displayed using p-value trees for both the hypermethylated (Figure 3.9 b-d) and hypomethylated (Figure 3.10 b-d) enriched terms. The most significant enrichment (lowest p-value) is displayed by the red blocks which show the biological process of muscle atrophy, membrane hyperpolarization, peripheral nervous system neuron differentiation, peripheral nervous system neuron development, negative regulation of DNA-dependent DNA replication, protein folding, regulation of cellular response to growth factor stimulus, negative regulation of megakaryocyte differentiation, lymph vessel development and dorsal/ventral neural tube patterning (hypermethylated) (Figure3.9b) as well as protein-chromophore linkage (hypomethylated) (Figure 3.10b). The significant enrichment for cellular components includes axon initial segment, node of Ranvier and cell periphery (hypermethylated) (Figure 3.9c) as well as photoreceptor outer segment membrane (hypomethylated) (Figure 3.10c). The significant enrichment for molecular functions includes macrolide binding, FK506 binding, taste receptor activity, delayed rectifier potassium, channel activity, G-protein coupled receptor activity, peptidyl-prolyl cis-trans isomerase activity, cis-trans isomerase activity and GDP binding (hypermethylated) (Figure 3.9d), as well as G-protein coupled photoreceptor activity and photoreceptor activity (hypomethylation) (Figure 3.10d).

(a)


Molecular functions
GDP binding 1 gene
cis-trans isomerase activity 1 gene
peptidyl-prolyl cis-trans isomerase activity 1 gene
G-protein coupled receptor activity 3 genes
delayed rectifier potassium channel activity 1 gene
taste receptor activity 1 gene
FK506 binding 1 gene
macrolide binding 1 gene
Cellular components
actin filament bundle 1 gene
cell division site part 1 gene
cell division site 1 gene
contractile actin filament bundle 1 gene
stress fiber 1 gene
cell surface furrow 1 gene
cleavage furrow 1 gene
cell periphery 9 genes
node of Ranvier 1 gene
axon initial segment 1 gene
Biological processes
dorsal/ventral neural tube patterning 1 gene
lymph vessel development 1 gene
negative regulation of megakaryocyte differentiation 1 gene
regulation of cellular response to growth factor stimulus 2 genes
protein folding 2 genes
negative regulation of DNA-dependent DNA replication 1 gene
peripheral nervous system neuron development 1 gene
peripheral nervous system neuron differentiation 1 gene
membrane hyperpolarization 1 gene
muscle atrophy 1 gene
0 0.5 1 1.5 2 2.5
Enrichment score (-log10(Pvalue))

(b)



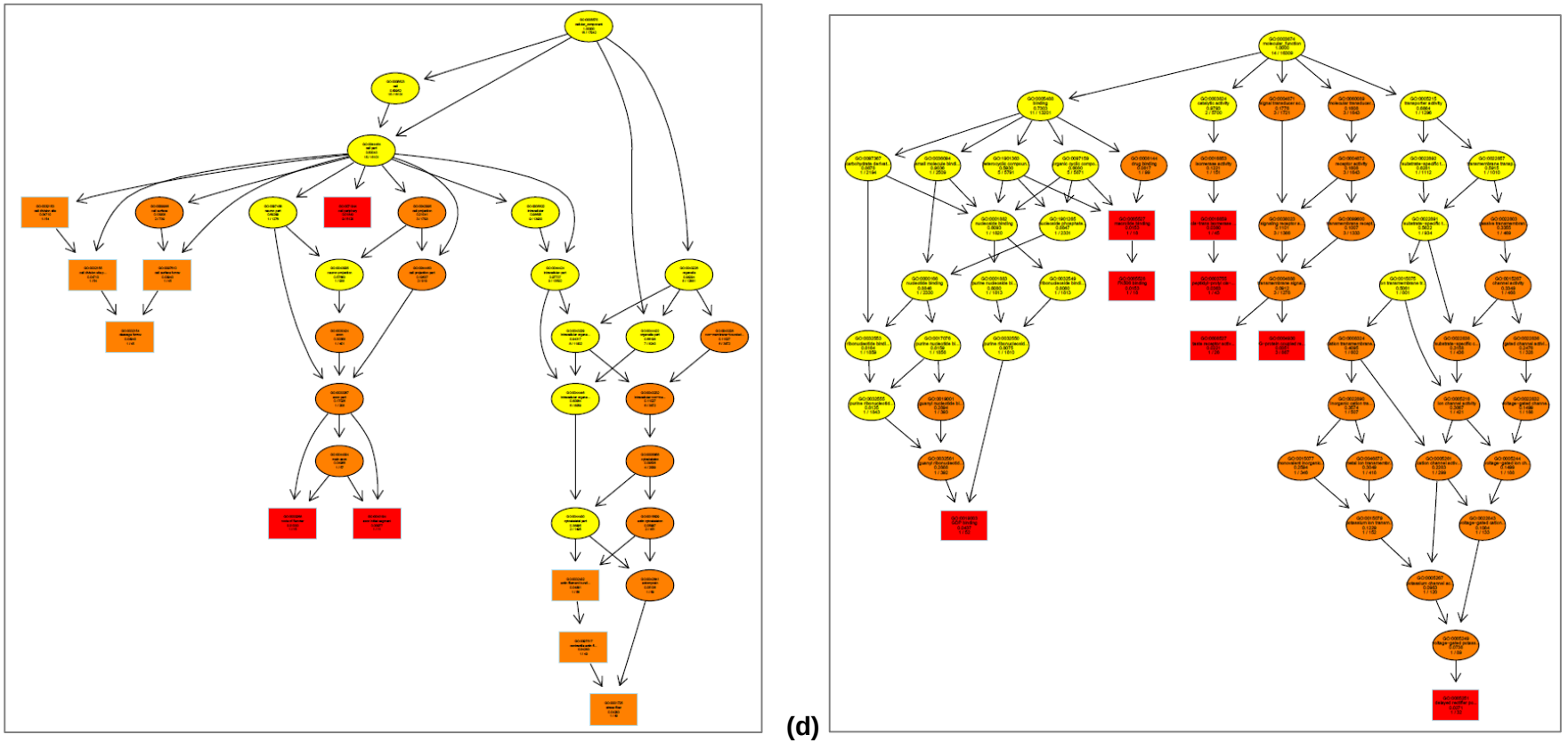
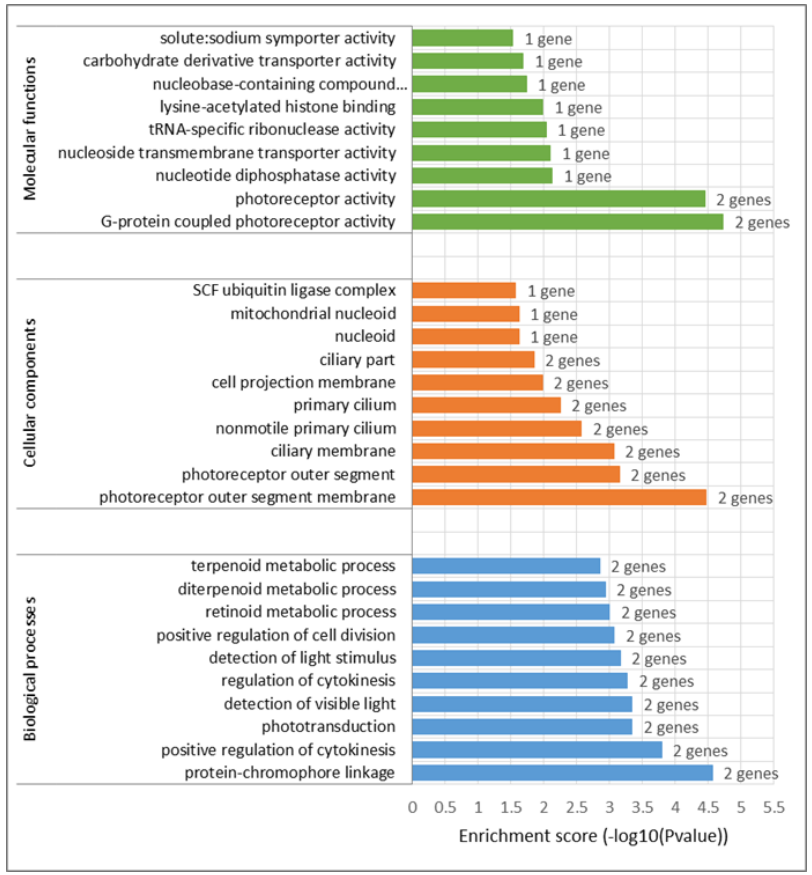
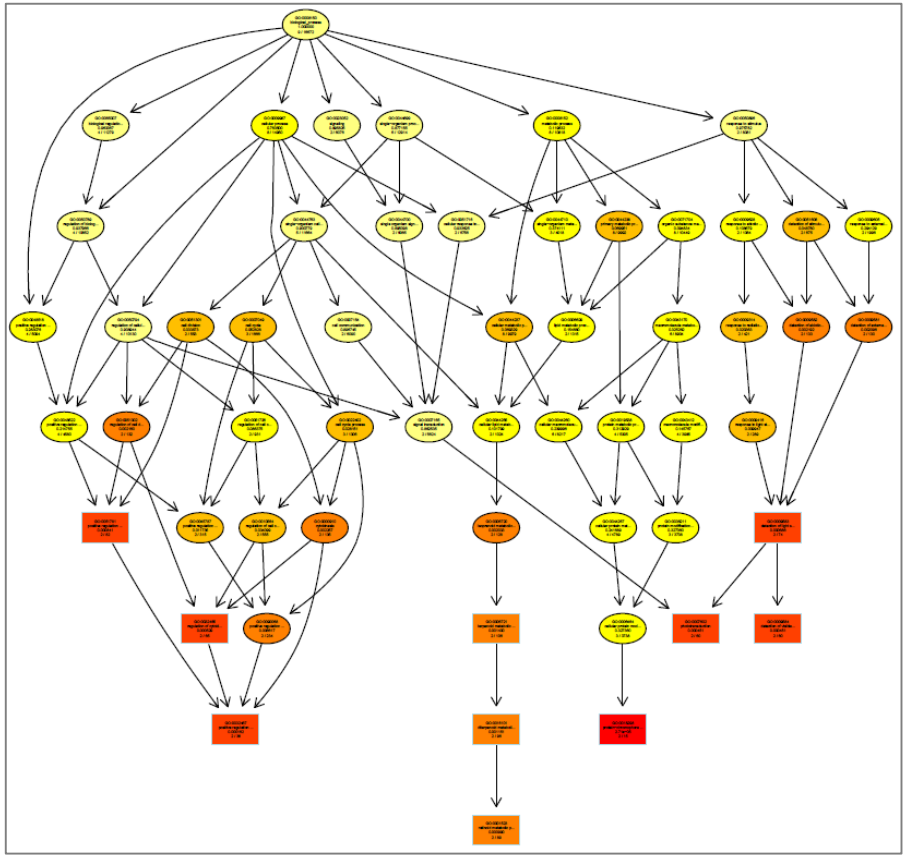


Figure 3.9: Gene Ontology (GO) enrichment analysis of the hypermethylated differentially methylated genes in known diabetes when compared to the screen-detected diabetes group. (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions


Molecular functions
solute:sodium symporter activity
carbohydrate derivative transporter activity
nucleobase-containing compound...
lysine-acetylated histone binding
tRNA-specific ribonuclease activity
nucleoside transmembrane transporter activity
nucleotide diphosphatase activity
photoreceptor activity
G-protein coupled photoreceptor activity
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
2 genes
2 genes
Cellular components
SCF ubiquitin ligase complex
mitochondrial nucleoid
nucleoid
ciliary part
cell projection membrane
primary cilium
nonmotile primary cilium
ciliary membrane
photoreceptor outer segment
photoreceptor outer segment membrane
1 gene
1 gene
1 gene
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
Biological processes
terpenoid metabolic process
diterpenoid metabolic process
retinoid metabolic process
positive regulation of cell division
detection of light stimulus
regulation of cytokinesis
detection of visible light
phototransduction
positive regulation of cytokinesis
protein-chromophore linkage
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5
Enrichment score (-log10(Pvalue))

(a)



(b)

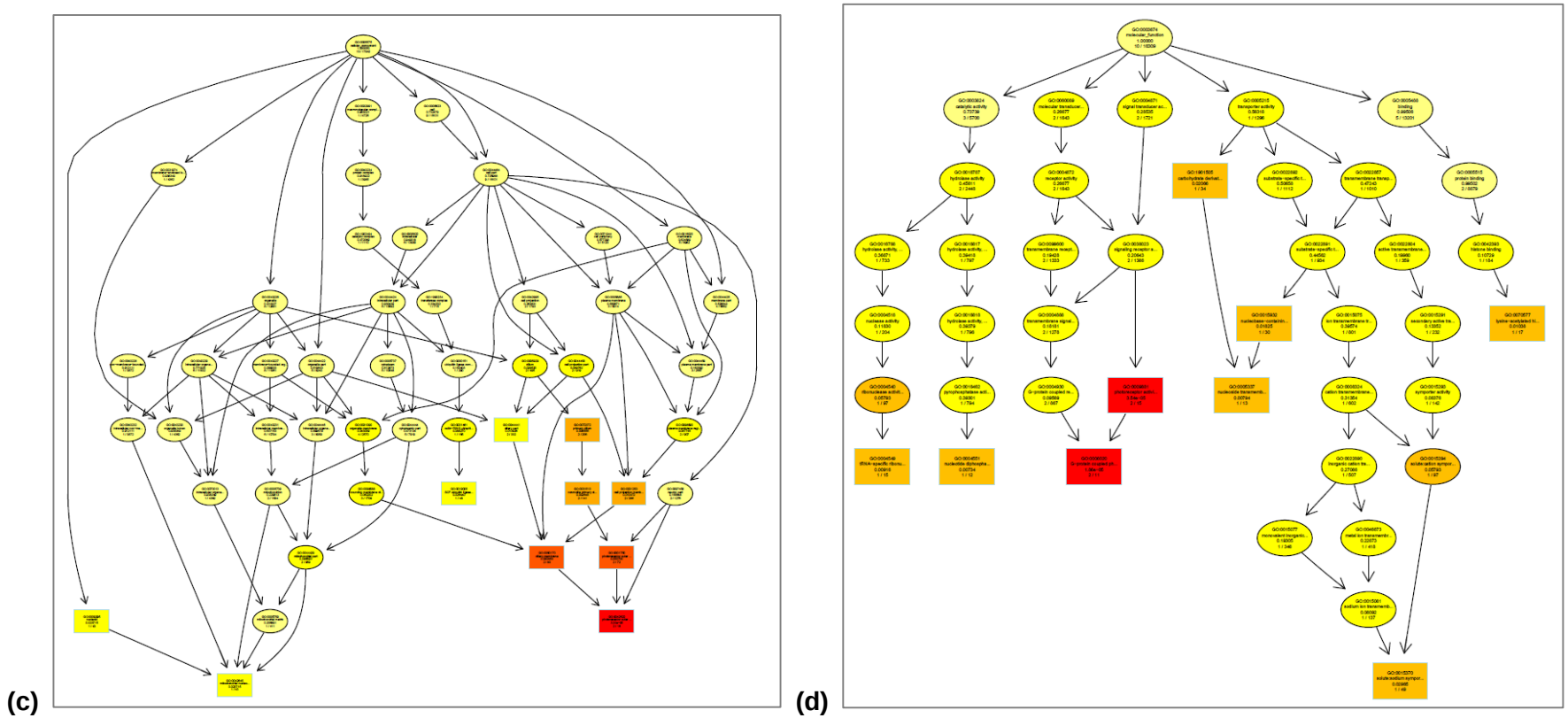


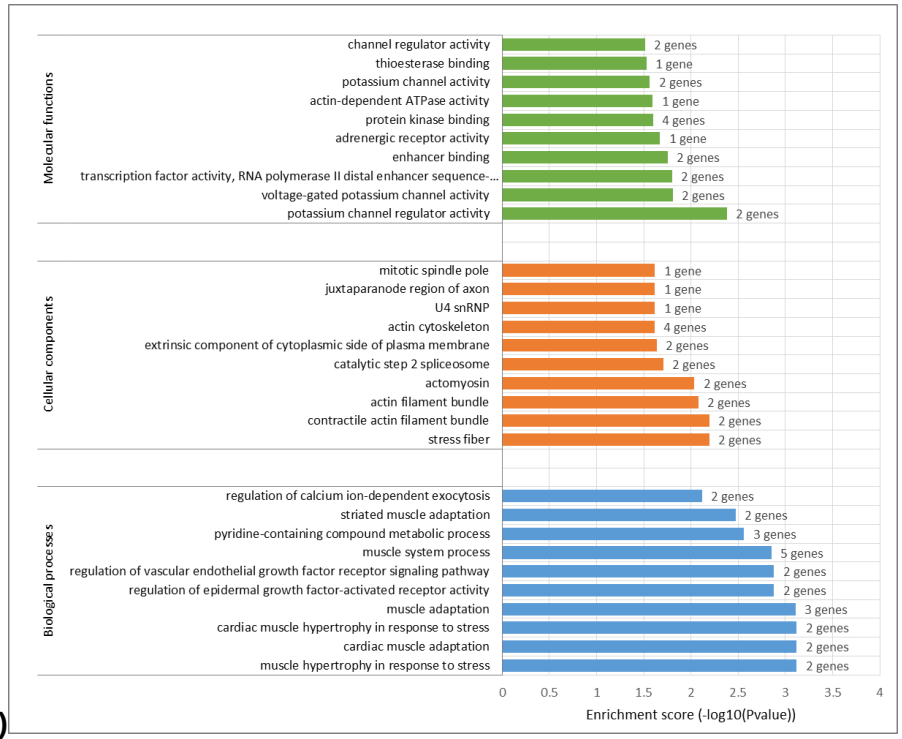
Figure 3.10: Gene Ontology (GO) enrichment analysis of the hypomethylated differentially methylated genes in known diabetes when compared to the screen-detected diabetes group. (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

3.3.5 Known diabetes versus prediabetes

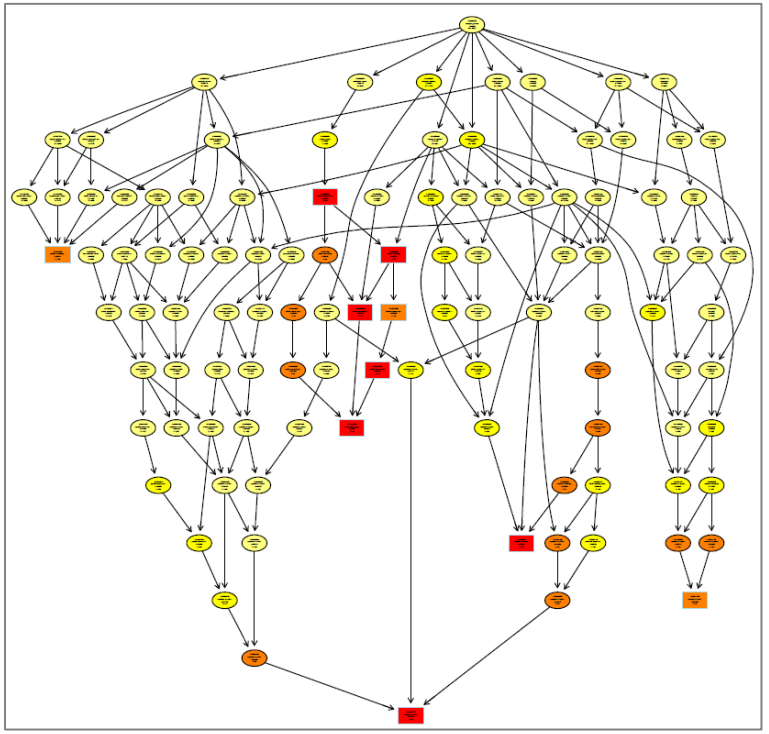
The top 10 GO enriched terms for the known diabetic group versus the prediabetes group categorised into biological processes, molecular functions and cellular components are illustrated in Figure 3.11 and 3.12. Muscle hypertrophy in response to stress (biological process), stress fiber (cellular component) and potassium channel regulator activity (molecular function) have the highest-ranked enrichment score amongst the hypermethylated DMRs (Figure 3.11a). On the other hand, carbohydrate derivative transport (biological process), catalytic step 2 spliceosome (cellular component) and nucleobase-containing compound transmembrane transporter activity (molecular function) have the highest-ranked enrichment score among the hypomethylated DMRs (Figure 3.12a).

P-value trees showing the hierarchy of gene ontology biological processes, cellular components and molecular functions are displayed for both hypermethylated (Figure 3.11 b-d) and hypomethylated (Figure 3.12 b-d) terms. The most significant enrichment (lowest p-value) is displayed by the red blocks which show the biological process of muscle hypertrophy in response to stress, cardiac muscle adaptation, cardiac muscle hypertrophy in response to stress, muscle adaptation, regulation of epidermal growth factor-activated receptor activity, regulation of vascular endothelial growth factor receptor signalling pathway and muscle system process (hypermethylated) (Figure3.11b) as well as carbohydrate derivative transport and regulation of carbohydrate biosynthetic process (hypomethylated) (Figure 3.12b). The significant enrichment for cellular components includes stress fiber, contractile actin filament bundle, actin filament bundle, actomyosin and catalytic step 2 spliceosome (hypermethylated) (Figure 3.11c) as well as catalytic step 2 spliceosome, endoplasmic reticulum, oligosaccharyltransferase complex and nuclear transcription factor complex (hypomethylated) (Figure 3.12c). The significant enrichment for molecular functions includes potassium channel regulator activity (hypermethylated) (Figure 3.11d) as well as nucleobase-containing compound transmembrane transporter activity (hypomethylation) (Figure 3.12d).

(a)


Molecular functions
channel regulator activity
thioesterase binding
potassium channel activity
actin-dependent ATPase activity
protein kinase binding
adrenergic receptor activity
enhancer binding
transcription factor activity, RNA polymerase II distal enhancer sequence-...
voltage-gated potassium channel activity
potassium channel regulator activity
2 genes
1 gene
2 genes
1 gene
4 genes
1 gene
2 genes
2 genes
2 genes
2 genes
Cellular components
mitotic spindle pole
juxtaparanode region of axon
U4 snRNP
actin cytoskeleton
extrinsic component of cytoplasmic side of plasma membrane
catalytic step 2 spliceosome
actomyosin
actin filament bundle
contractile actin filament bundle
stress fiber
1 gene
1 gene
1 gene
4 genes
2 genes
2 genes
2 genes
2 genes
2 genes
2 genes
Biological processes
regulation of calcium ion-dependent exocytosis
striated muscle adaptation
pyridine-containing compound metabolic process
muscle system process
regulation of vascular endothelial growth factor receptor signaling pathway
regulation of epidermal growth factor-activated receptor activity
muscle adaptation
cardiac muscle hypertrophy in response to stress
cardiac muscle adaptation
muscle hypertrophy in response to stress
2 genes
2 genes
3 genes
5 genes
2 genes
2 genes
3 genes
2 genes
2 genes
2 genes
0
0.5
1
1.5
2
2.5
3
3.5
4
Enrichment score (-log10(Pvalue))

(b)



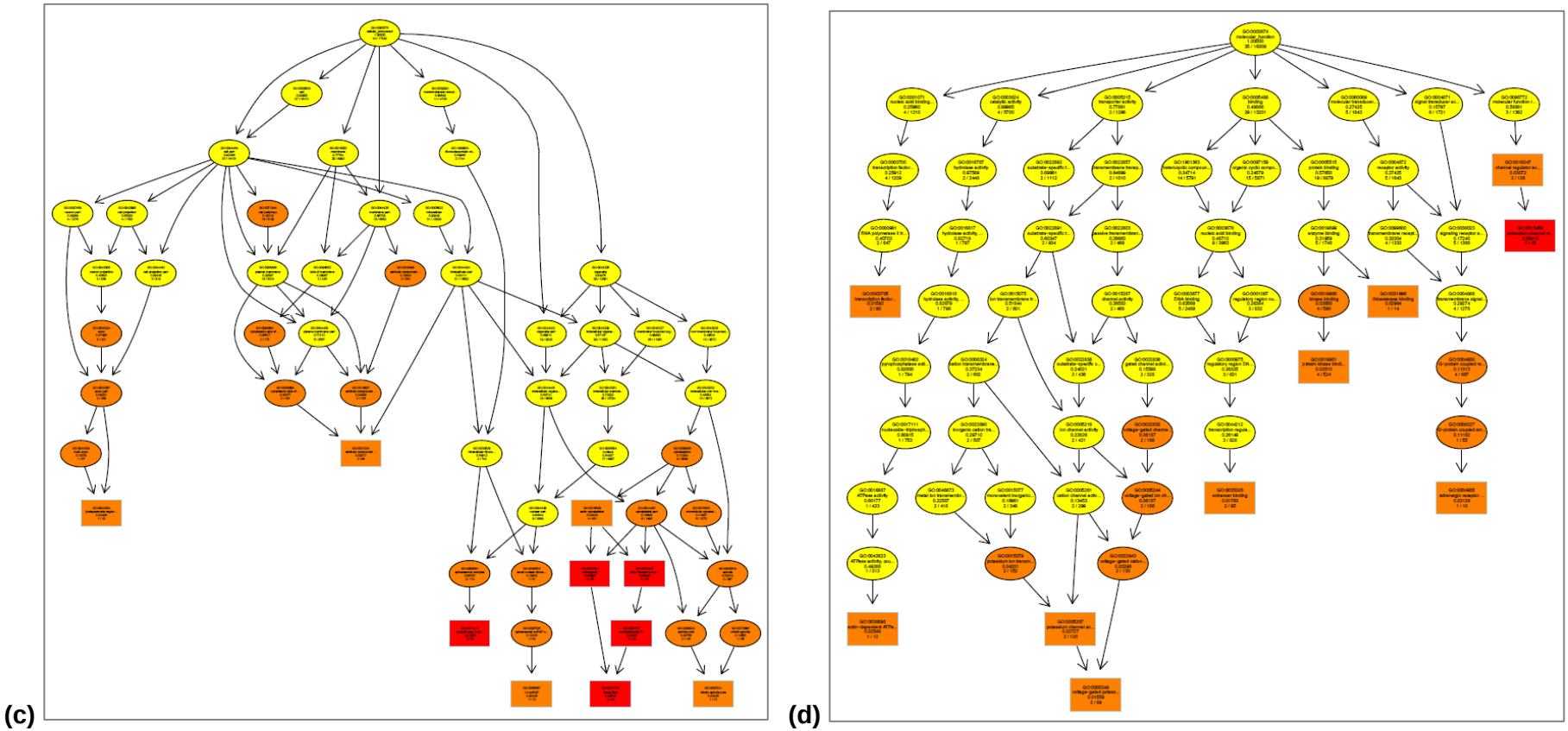
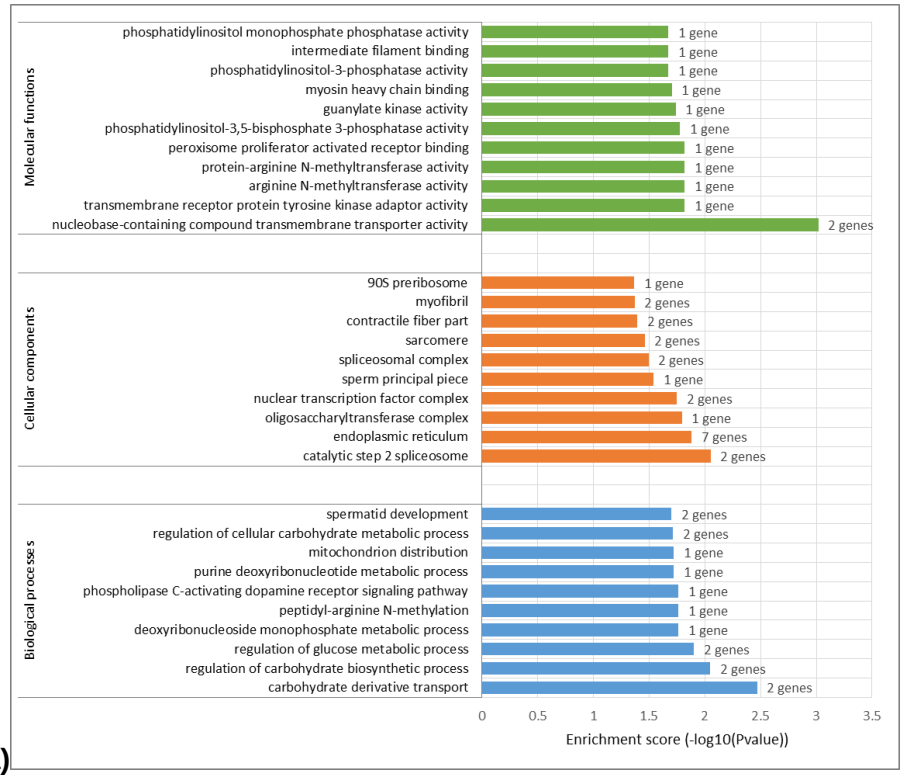
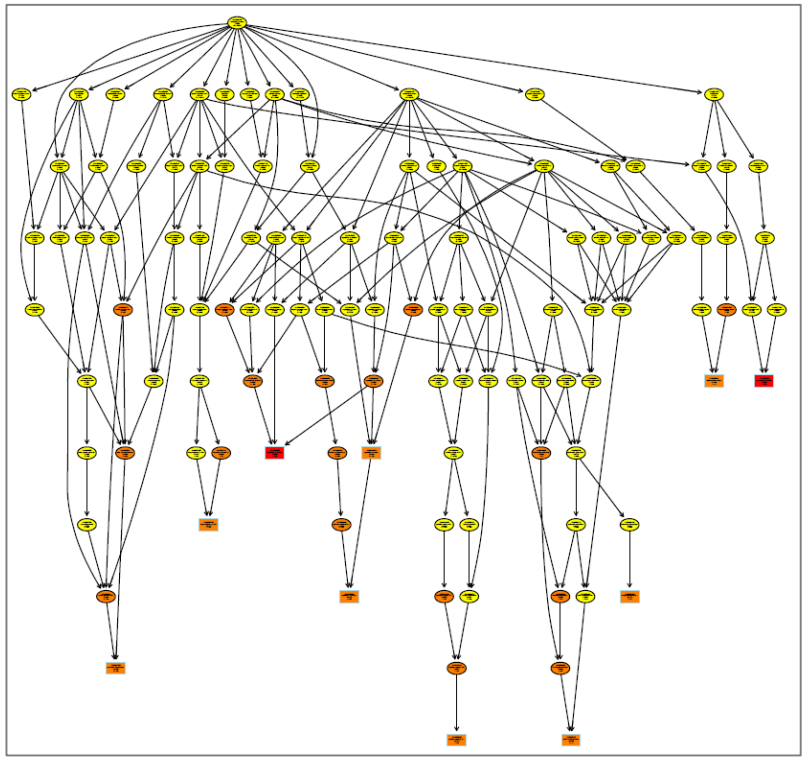


Figure 3.11: Gene Ontology (GO) enrichment analysis of the hypermethylated differentially methylated genes in known diabetes when compared to the prediabetes group. (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions


Molecular functions
phosphatidylinositol monophosphate phosphatase activity
1 gene
intermediate filament binding
1 gene
phosphatidylinositol-3-phosphatase activity
1 gene
myosin heavy chain binding
1 gene
guanylate kinase activity
1 gene
phosphatidylinositol-3,5-bisphosphate 3-phosphatase activity
1 gene
peroxisome proliferator activated receptor binding
1 gene
protein-arginine N-methyltransferase activity
1 gene
arginine N-methyltransferase activity
1 gene
transmembrane receptor protein tyrosine kinase adaptor activity
1 gene
nucleobase-containing compound transmembrane transporter activity
2 genes
Cellular components
90S preribosome
1 gene
myofibril
2 genes
contractile fiber part
2 genes
sarcomere
2 genes
spliceosomal complex
2 genes
sperm principal piece
1 gene
nuclear transcription factor complex
2 genes
oligosaccharyltransferase complex
1 gene
endoplasmic reticulum
7 genes
catalytic step 2 spliceosome
2 genes
Biological processes
spermatid development
2 genes
regulation of cellular carbohydrate metabolic process
2 genes
mitochondrion distribution
1 gene
purine deoxyribonucleotide metabolic process
1 gene
phospholipase C-activating dopamine receptor signaling pathway
1 gene
peptidyl-arginine N-methylation
1 gene
deoxyribonucleoside monophosphate metabolic process
1 gene
regulation of glucose metabolic process
2 genes
regulation of carbohydrate biosynthetic process
2 genes
carbohydrate derivative transport
2 genes
0
0.5
1
1.5
2
2.5
3
3.5
Enrichment score (-log10(Pvalue))

(a)



(b)

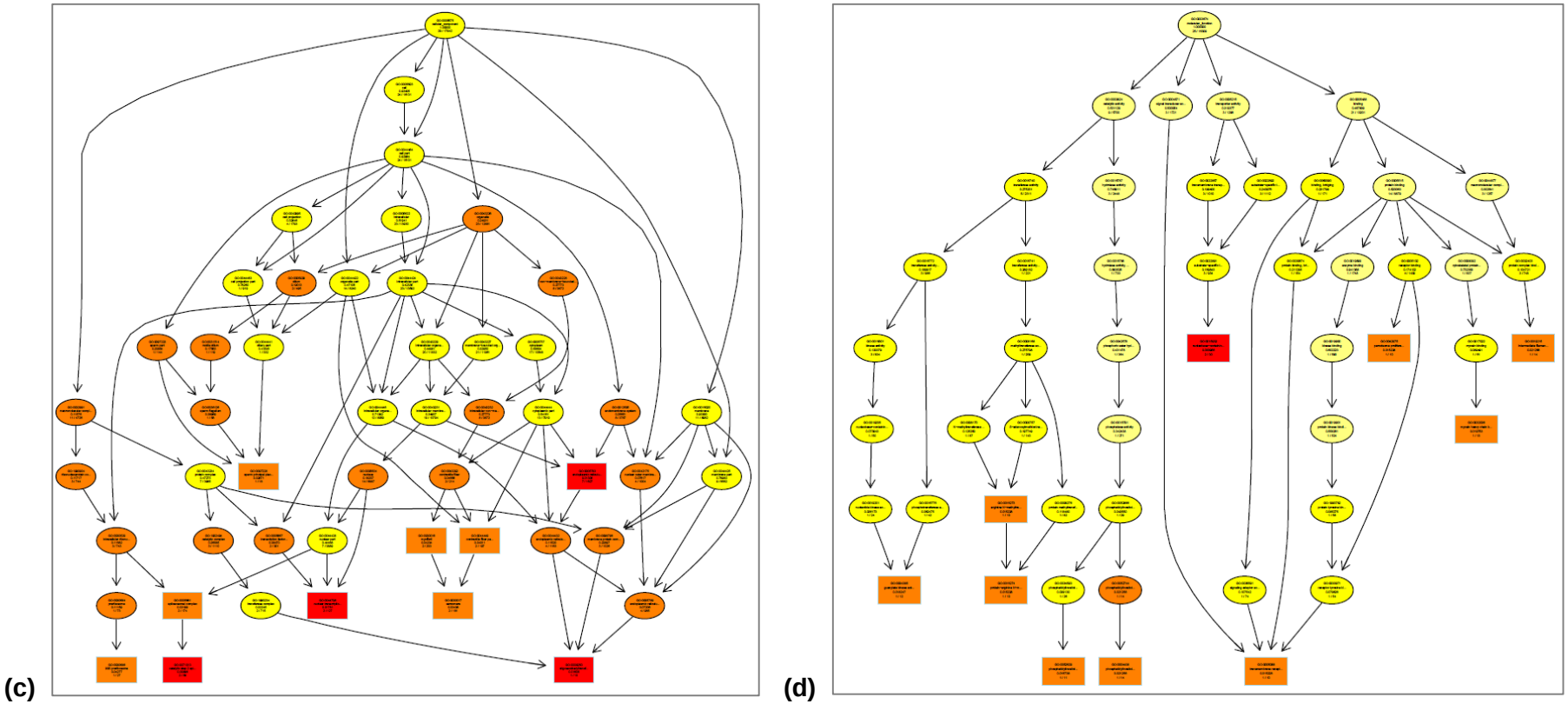


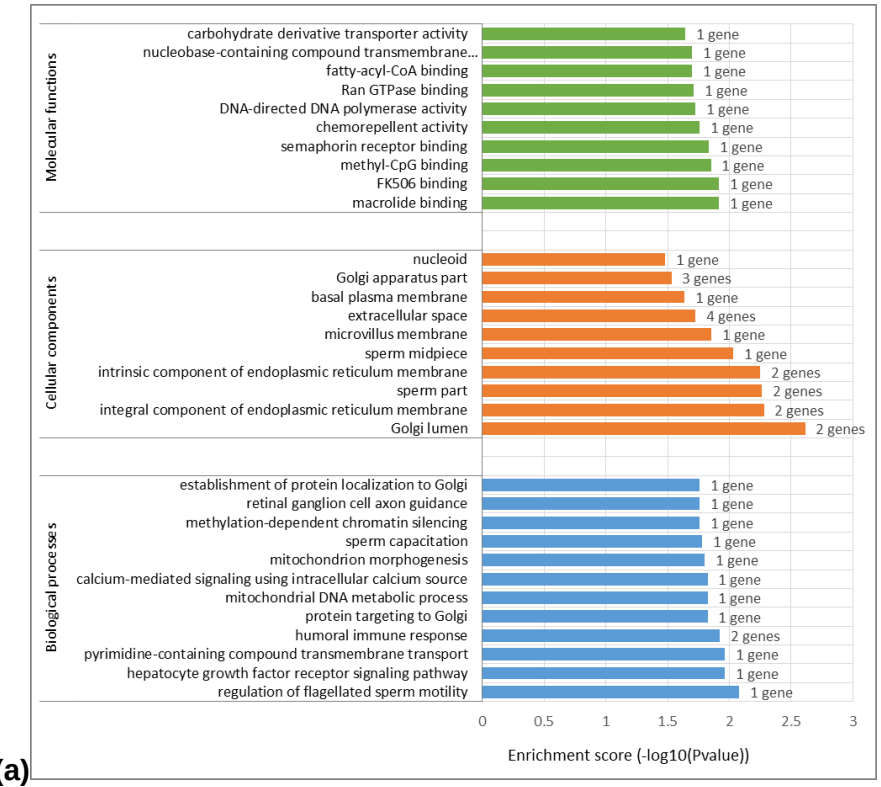
Figure 3.12: Gene Ontology (GO) enrichment analysis of the hypomethylated differentially methylated genes in known diabetes when compared to the prediabetes group. (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

3.3.6 Prediabetes versus screen-detected diabetes

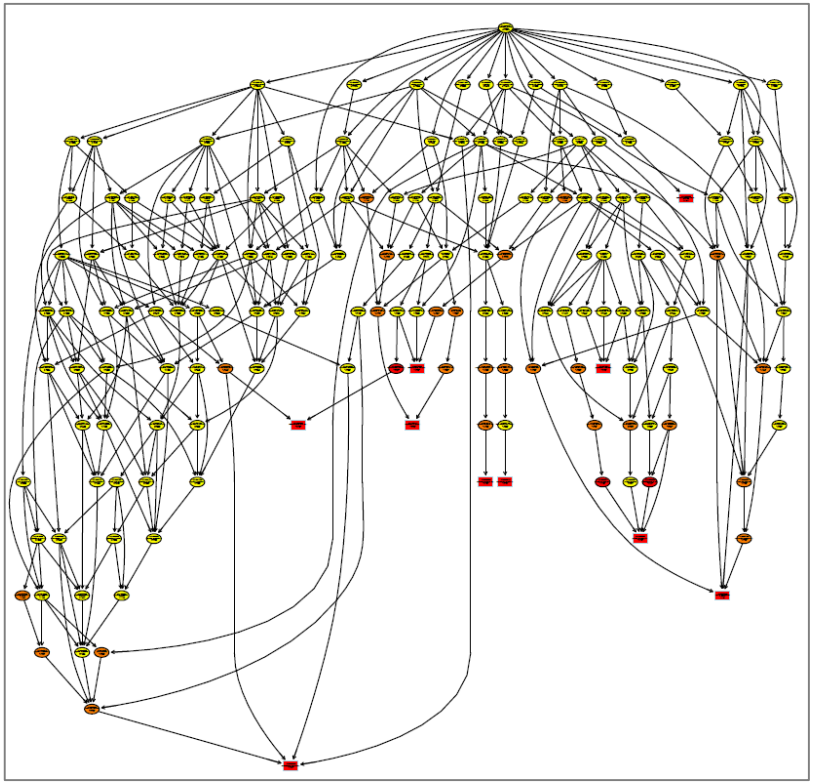
The top 10 GO enriched terms for the prediabetic group versus the screen-detected diabetic group categorised into biological processes, molecular functions and cellular components are illustrated in figure 3.13 and 3.14. Regulation of flagellated sperm motility (biological process), Golgi lumen (cellular component) and macrolide binding (molecular function) have the highest-ranked enrichment score amongst the hypermethylated DMRs (Figure 3.13a). On the other hand, positive regulation of wound healing (biological process), sarcolemma (cellular component) and hormone activity (molecular function) have the highest-ranked enrichment score among the hypomethylated DMRs (Figure 3.14a).

The hierarchy of gene ontology biological processes, cellular components and molecular functions are displayed using p-value trees for both the hypermethylated (Figure 3.13 b-d) and hypomethylated (Figure 3.14 b-d) enriched terms. The most significant enrichment (lowest p-value) is displayed by the red blocks which show the biological process of regulation of flagellated sperm motility, hepatocyte growth factor receptor signalling pathway, pyrimidine-containing compound transmembrane transport, humoral immune response, protein targeting to Golgi, mitochondrial DNA metabolic process, calcium-mediated signalling using intracellular calcium source, mitochondrion morphogenesis, sperm capacitation and methylation-dependent chromatin silencing (hypermethylated) (Figure3.13b) as well as positive regulation of wound healing and positive regulation of response to wounding (hypomethylated) (Figure 3.14b). The significant enrichment for cellular components includes Golgi lumen, integral component of endoplasmic reticulum membrane, sperm part and intrinsic component of endoplasmic reticulum membrane (hypermethylated) (Figure 3.13c) as well as sarcolemma, host cell part, platelet dense tubular network, inhibitory synapse, host and host cell (hypomethylated) (Figure 3.14c). The significant enrichment for molecular functions includes macrolide binding, FK506 binding, methyl-CpG binding, semaphorin receptor binding, chemorepellent activity, DNA-directed DNA polymerase activity, Ran GTPase binding, fatty-acyl-CoA binding, nucleobase-containing compound transmembrane transporter activity and carbohydrate derivative transporter activity (hypermethylated) (Figure 3.13d) as well as hormone activity (hypomethylation) (Figure 3.14d).

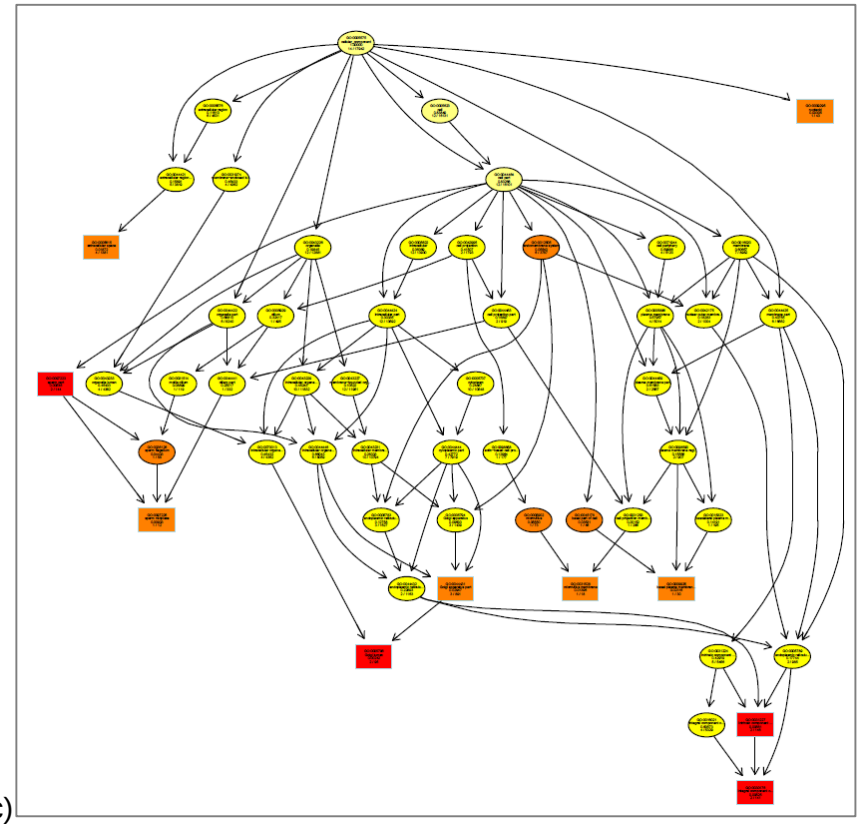
(a)


Molecular functions
carbohydrate derivative transporter activity
nucleobase-containing compound transmembrane...
fatty-acyl-CoA binding
Ran GTPase binding
DNA-directed DNA polymerase activity
chemorepellent activity
semaphorin receptor binding
methyl-CpG binding
FK506 binding
macrolide binding
1 gene
Cellular components
nucleoid
Golgi apparatus part
basal plasma membrane
extracellular space
microvillus membrane
sperm midpiece
intrinsic component of endoplasmic reticulum membrane
sperm part
integral component of endoplasmic reticulum membrane
Golgi lumen
1 gene
3 genes
4 genes
2 genes
Biological processes
establishment of protein localization to Golgi
retinal ganglion cell axon guidance
methylation-dependent chromatin silencing
sperm capacitation
mitochondrion morphogenesis
calcium-mediated signaling using intracellular calcium source
mitochondrial DNA metabolic process
protein targeting to Golgi
humoral immune response
pyrimidine-containing compound transmembrane transport
hepatocyte growth factor receptor signaling pathway
regulation of flagellated sperm motility
0
0.5
1
1.5
2
2.5
3
Enrichment score (-log10(Pvalue))

(b)



(c)



(d)

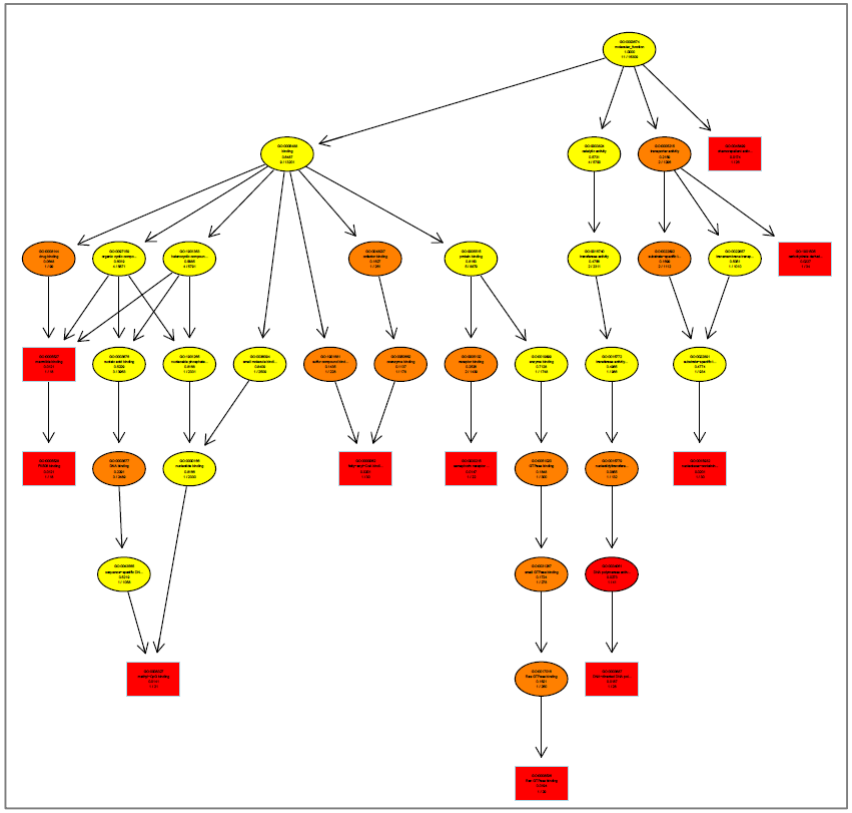
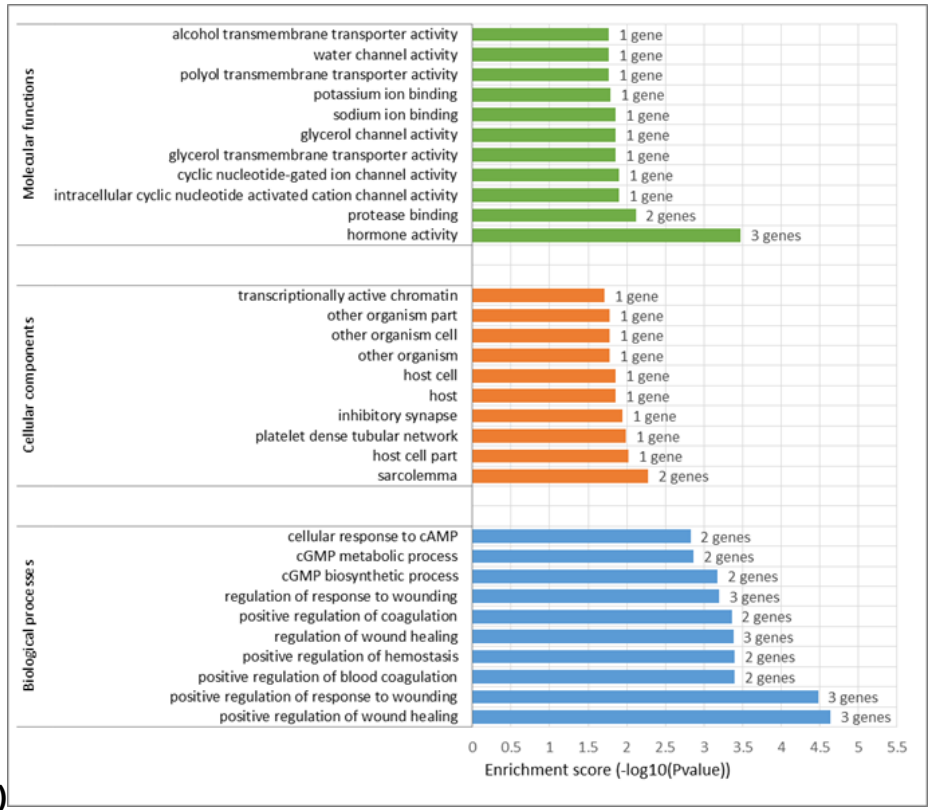
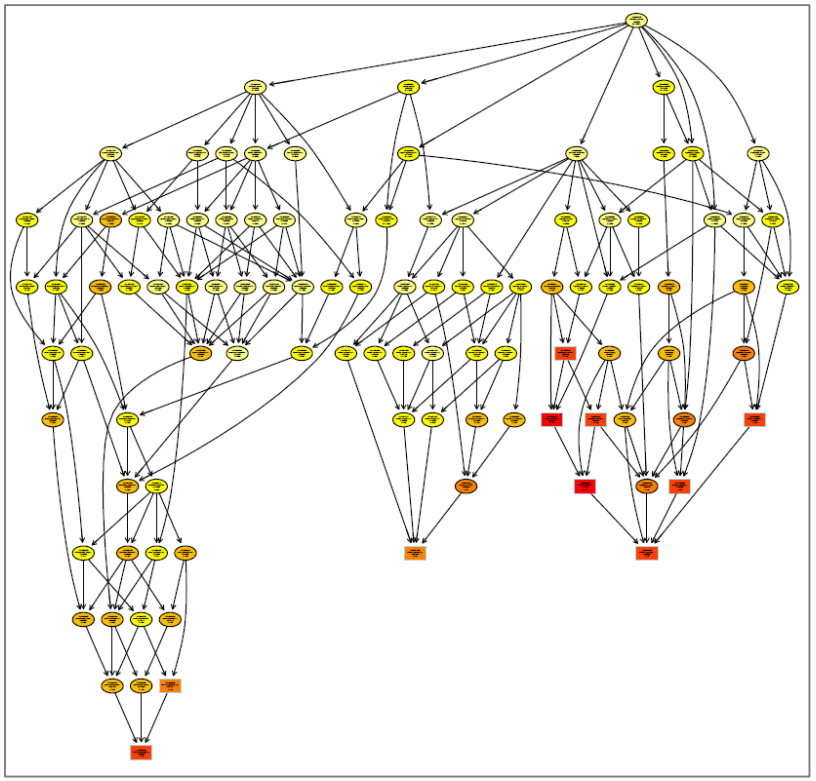


Figure 3.13: Gene Ontology (GO) enrichment analysis of the hypermethylated differentially methylated genes in the prediabetes group when compared to the screen-detected group. (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

(a)


Molecular functions
alcohol transmembrane transporter activity
water channel activity
polyol transmembrane transporter activity
potassium ion binding
sodium ion binding
glycerol channel activity
glycerol transmembrane transporter activity
cyclic nucleotide-gated ion channel activity
intracellular cyclic nucleotide activated cation channel activity
protease binding
hormone activity
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
2 genes
3 genes
Cellular components
transcriptionally active chromatin
other organism part
other organism cell
other organism
host cell
host
inhibitory synapse
platelet dense tubular network
host cell part
sarcolemma
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
1 gene
2 genes
Biological processes
cellular response to cAMP
cGMP metabolic process
cGMP biosynthetic process
regulation of response to wounding
positive regulation of coagulation
regulation of wound healing
positive regulation of hemostasis
positive regulation of blood coagulation
positive regulation of response to wounding
positive regulation of wound healing
2 genes
2 genes
2 genes
3 genes
2 genes
3 genes
2 genes
2 genes
3 genes
3 genes
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5
Enrichment score (-log10(Pvalue))

(b)



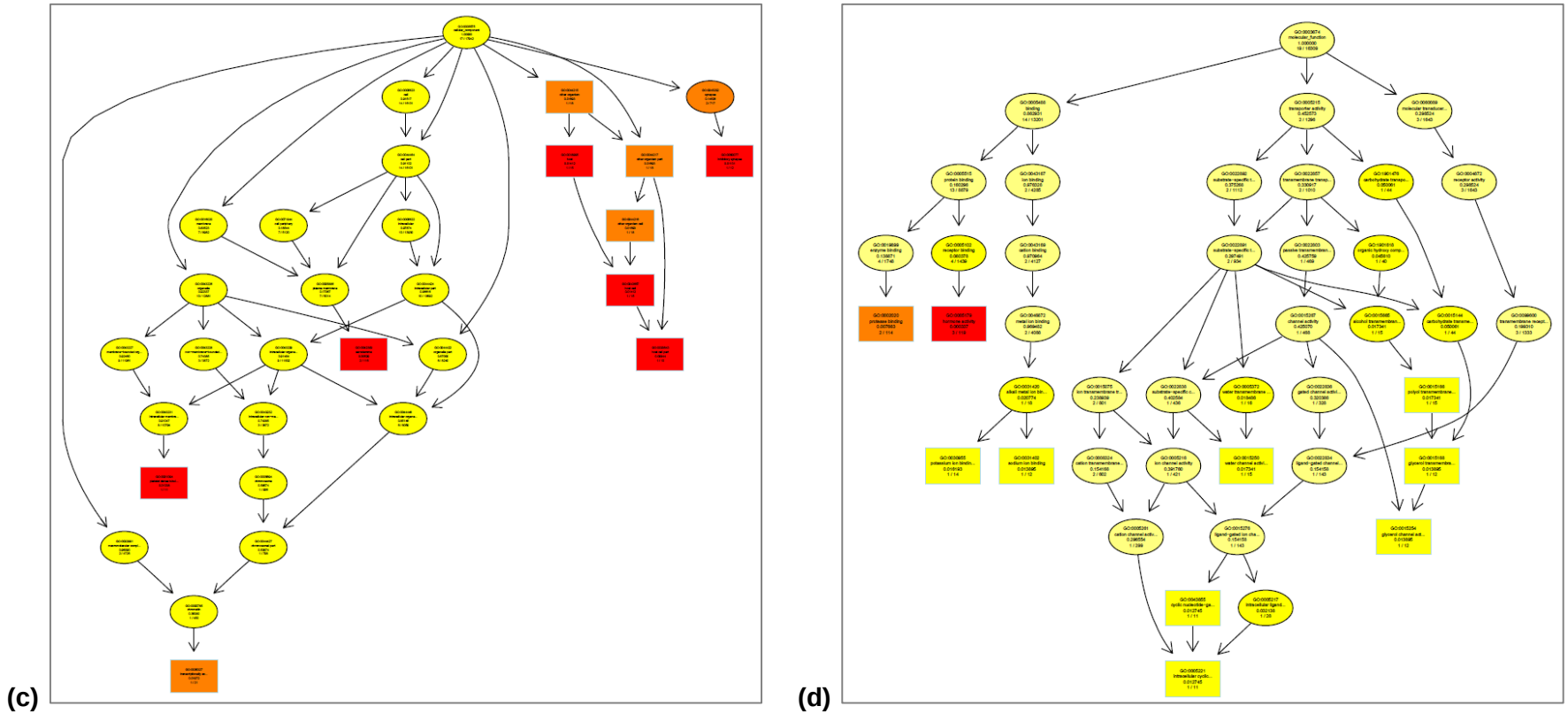


Figure 3.14: Gene Ontology (GO) enrichment analysis of the hypomethylated differentially methylated genes in the prediabetes group when compared to the screen-detected group. (a) The bars plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components and molecular function. Data are presented as enriched scores expressed as -log10 (p-value). P-value trees are presented for the (b) biological processes; (c) cellular components and (d) molecular functions

3.4 KEGG Pathway analysis

The functional analysis mapping of genes to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were performed. The functional analysis was completed for each of the six comparison pairs and highlighted the significant pathways for both the hypermethylated and hypomethylated DMRs (Table 3.7). The p-value (Fisher P-value) denotes the significance of the pathway correlated to the conditions. A p-value of <= 0.05 was considered significant and the lower the p-value, the more significant is the pathway.

The cytokine-cytokine receptor interaction pathway and cAMP signalling pathway were significantly hypermethylated in the known diabetic group when compared to the normoglycaemic and prediabetic groups, whereas RNA transport and PI3K-Akt signalling pathways were hypomethylated. In addition, the notch signalling pathway was hypermethylated in the newly diagnosed group and purine metabolism was hypomethylated when comparing to the control subjects. No enriched pathways were observed for the prediabetic group when comparing to the control group and similarly when comparing the known diabetic group to newly diagnosed diabetic participants.

Table 3.7: KEGG analysis of the top ten enrichment score values (−log 10(P-value)) of the significantly enriched pathway

Pathway ID	Definition	Fisher-Pvalue	Enrichment Score	Genes
Hypermethylated Known diabetes versus control				
hsa04060	Cytokine-cytokine receptor interaction	0.002346103	2.629653	IL12A//IL1R2//TNFSF8//TNFSF9
hsa05146	Amoebiasis - Homo sapiens	0.01963637	1.706939	IL12A//IL1R2
hsa00190	Oxidative phosphorylation	0.03465798	1.460197	COX7A1//NDUFB2
hsa05012	Parkinson's disease	0.03906889	1.408169	COX7A1//NDUFB2
hsa04932	Non-alcoholic fatty liver disease (NAFLD)	0.04368588	1.359659	COX7A1//NDUFB2
Hypomethylated Known diabetes versus controls				
hsa03013	RNA transport	0.0110018	1.958536	EIF2B3//EIF5
Hypermethylated New diabetes versus Controls				
hsa04330	Notch signaling pathway	0.003803428	2.419825	MFNG//NUMBL
hsa04918	Thyroid hormone synthesis	0.00884694	2.053207	ATP1A3//IYD
hsa04260	Cardiac muscle contraction	0.00979235	2.009113	ATP1A3//CACNG8
hsa04261	Adrenergic signaling in cardiomyocytes	0.03322194	1.478575	ATP1A3//CACNG8
Hypomethylated New diabetes versus Controls				
hsa00230	Purine metabolism	0.005656938	2.247419	ADK//POLR3H
Hypermethylated Prediabetes versus Controls				
There is no enriched pathway				
Hypomethylated Prediabetes versus Controls				
There is no enriched pathway				
Hypermethylated Known diabetes versus New diabetes				
There is no enriched pathway				
Hypomethylated Known diabetes versus New diabetes				
There is no enriched pathway				
Hypermethylated Known diabetes versus prediabetes				
hsa04024	cAMP signaling pathway	0.04261372	1.370451	AMH//GPR119
Hypomethylated Known diabetes versus prediabetes				
hsa04151	PI3K-Akt signaling pathway	0.01327285	1.877036	COL6A6//IRS1//TCL1A
hsa04510	Focal adhesion	0.03649743	1.437738	COL6A6//MYL2
Hypermethylated prediabetes versus new diabetes				
There is no enriched pathway				
Hypomethylated prediabetes versus new diabetes				
hsa04914	Progesterone-mediated oocyte maturation	0.00013244	3.87798	SPDYE2//SPDYE2B//SPDYE6
hsa04114	Oocyte meiosis	0.000266361	3.574529	SPDYE2//SPDYE2B//SPDYE6

3.5 Validation of significant DMRs using pyrosequencing

The significant DMRs identified in the 48 participant samples by the MeDIP sequencing were validated using pyrosequencing. Selected Pyromark CpG Assay Primers (Qiagen, Hilden, Germany) (refer to Table 2.4 in Chapter 2: Methodology) were used in the pyrosequencing analysis to validate whether methylation existed and could be detected in the selected regions. Samples from participants in the various group were analysed and areas of hypermethylation

and hypomethylation were detected. In addition, the PyroMark Control Oligo (Qiagen, Hilden, Germany) was used for quality control for the experimental procedure. The results for the pyrosequencing are presented in pyrograms whereby the y-axis displays the light intensity and the x-axis presents the sequence of nucleotides added during the pyrosequencing run.

The pyrogram for the Control Oligo run is presented in Figure 3.15. The results from the control run show that the percentage of C (cytosine) was 49% and T (thymine) was 51%. These results confirmed that the Control Oligo run passed the quality control check.

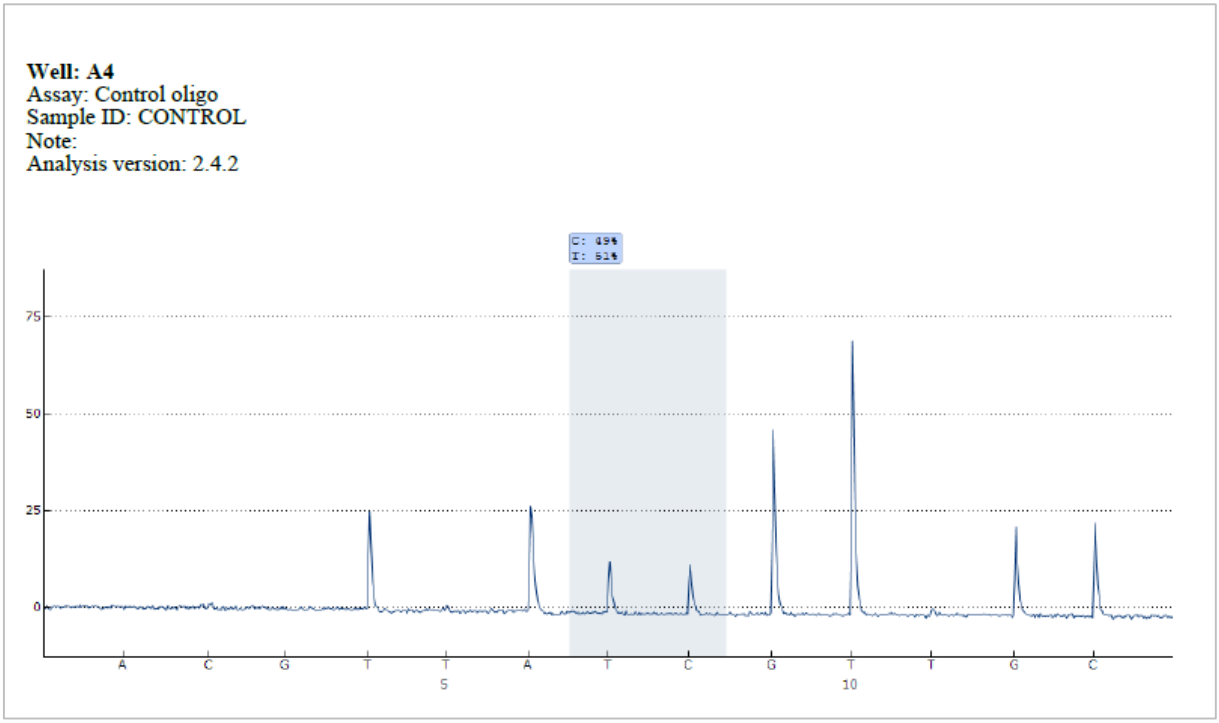


Figure 3.15: Pyrogram showing the results from the Pyromark Control Oligo test run. The percentages shown for C (cytosine) and T (thymine) indicate that the control run has passed the quality control check

Pyrogram results from participant samples where the selected Pyromark CpG Assay Primers were applied to the pyrosequencing can be seen in the examples presented in Figure 3.16 and 3.17. In Figure 3.16, the low percentage methylation detected in the sample run at the five CpG regions range from 5-12% and is indicative of hypomethylation. Readings at the first three CpG regions passed the quality check whereas readings at the last two CpG regions were flagged due to high peak deviation at dispensation 30. After evaluating the percentage methylation and the peak height detected at the fourth and fifth CpG region, the readings were manually passed. In Figure 3.17, the higher percentage of methylation detected in the sample

run at the four CpG regions ranged from 36-71% and indicated hypermethylation at these regions. The readings at all four CpG regions passed the quality check.

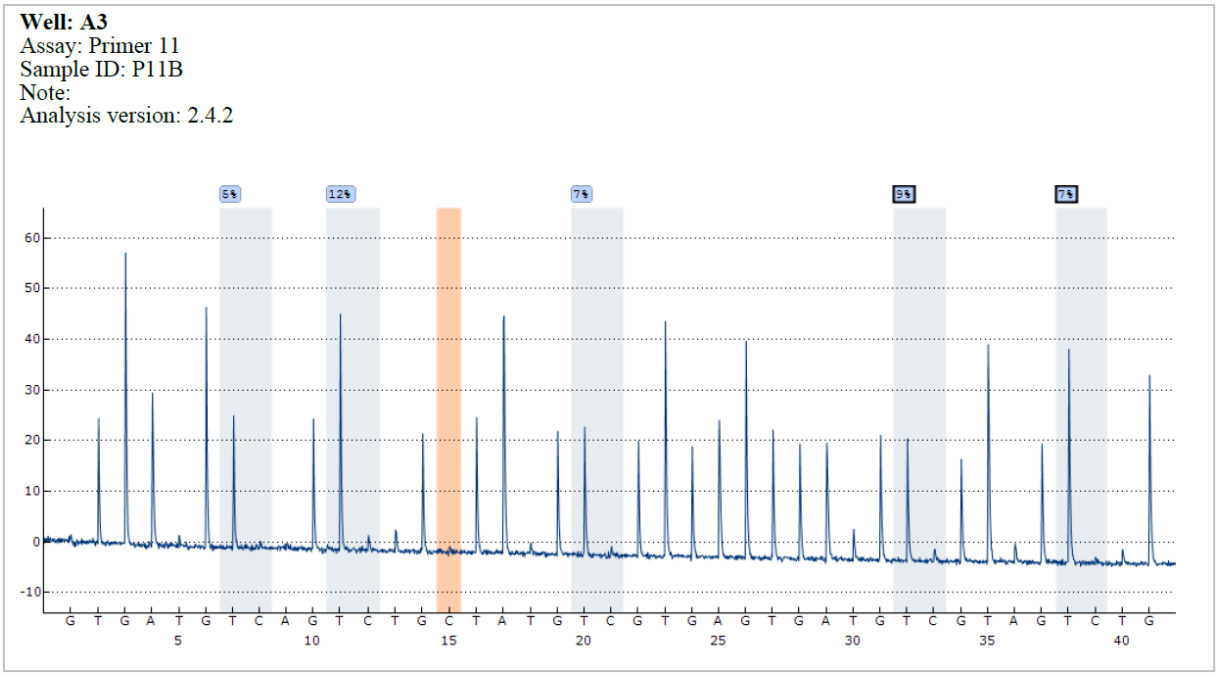


Figure 3.16: Pyrogram showing the percentage methylation at five CpG regions. The percentage of methylation ranged from 5-12% across the five CpG regions indicating low levels of DNA methylation

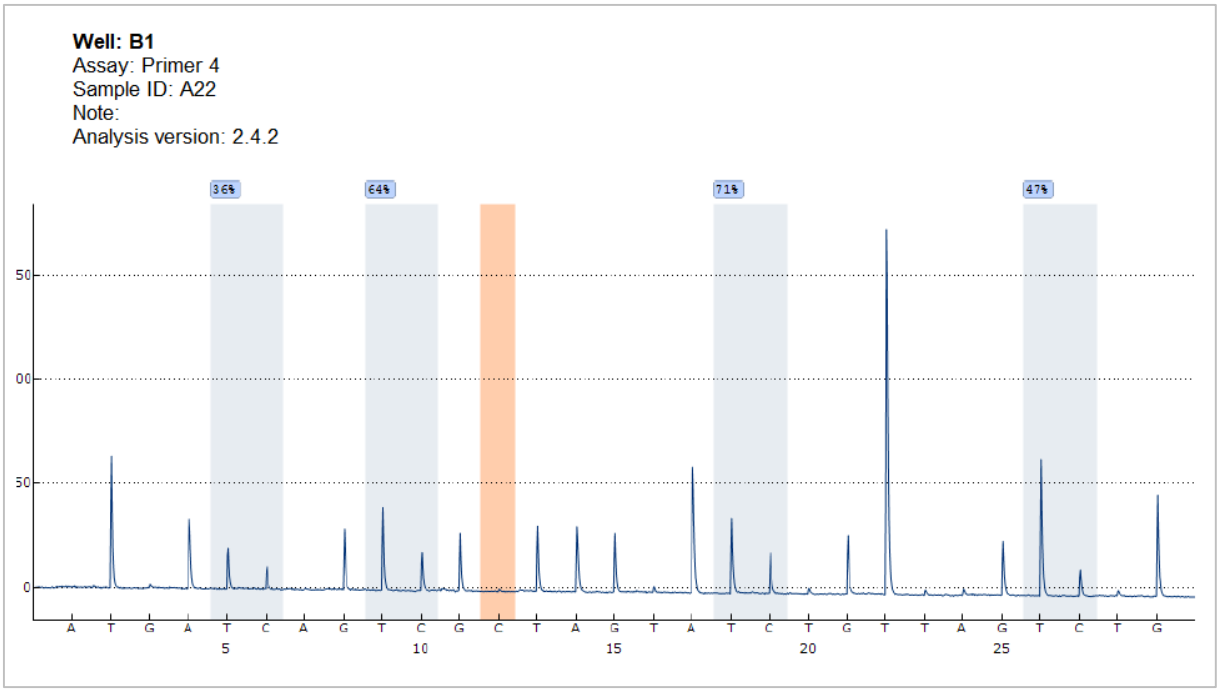
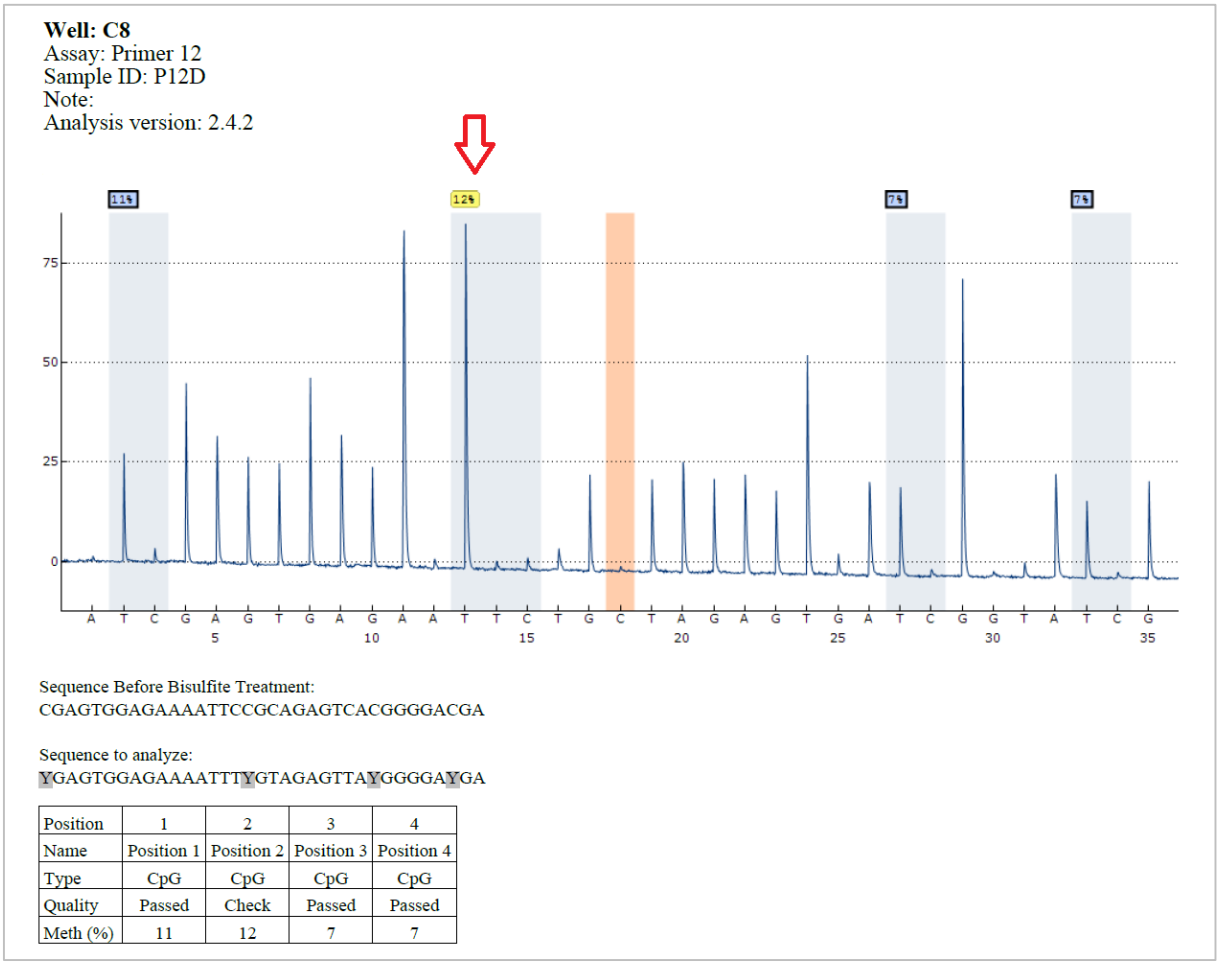


Figure 3.17: Pyrogram showing the percentage methylation at four CpG regions. The percentage of methylation ranged from 36-71% across the four CpG regions indicating higher levels of DNA methylation

Flagged messages due to high peak height deviation could at times be as a result of homopolymer regions whereby a string of repeated nucleotides are detected in the run. Homopolymer regions in a run affect the synchronised synthesis of the DNA strand and may result in irregular sequence peak heights which affect the read length (Heather and Chain, 2016). In such cases, the instrument software is undecided and cannot discern between two bases next to each other. This is a common base-calling error that occurs in pyrosequencing (Miller *et al.*, 2008; Balzer *et al*., 2011; Heather and Chain, 2016). An example of a homopolymer stretch can be seen in Figure 3.18 where the repeated occurrence of the A (adenosine) and T (thymine) nucleotide in the sample run resulted in exaggerated peak heights and uncertainty in the quality check at the second CpG region. This resulted in a reading of 12% which was flagged, as indicated by the yellow highlighting of the reading in Figure 3.18. Upon further investigation of the peak height and percentage at this CPG region, the reading was disregarded as it was unclear whether the peak reading was a reflection of a T (thymine) or C (cytosine) due to the homopolymer stretch present. Repeating the run resulted in the same outcome indicating that this was a possible base calling error.



Position	1	2	3	4
Name	Position 1	Position 2	Position 3	Position 4
Type	CpG	CpG	CpG	CpG
Quality	Passed	Check	Passed	Passed
Meth (%)	11	12	7	7

Figure 3.18: Pyrogram showing the percentage methylation at four CpG regions. The reading at the second CpG region (indicated by the red arrow) was flagged and disregarded

3.6 Long non-coding RNAs (lncRNAs)

In addition to the previous DNA methylation regions identified, lncRNA-associated DNA methylation was also measured. There was a total of 943947 statistically significant lncRNA-associated MeDIP-enriched regions (peaks) identified for the 48 samples. As expected, the promoter region (TSS-2000bp, TSS+2000bp) showed the least amount of peaks for all the groups when compared to the gene body and intergenic regions. Statistically significant lncRNA-associated DMRs between promoters of the various groups were identified by diffReps using a limit of log2FC=1.0 and p-value=0.0001. Hypermethylated and hypomethylated DMRS are presented in Figure 3.19 with comparisons amongst the same six groups, namely known diabetes versus NGT; screen-detected diabetes versus NGT; prediabetes versus NGT; known diabetes versus screen-detected diabetes; known diabetes versus prediabetes and prediabetes versus screen-detected diabetes. Overall 641 lncRNA -associated DMRs were identified of which 63% were hypermethylated and 37% hypomethylated. The highest number of hypermethylated lncRNA-associated DMRs were observed by the known diabetes group (77%) when compared to the NGT group. When grouping the lncRNA-associated DMRs according to their chromosomal location (Figure 3.20), the most hypermethylated and hypomethylated DMRs were found on chromosome 1.

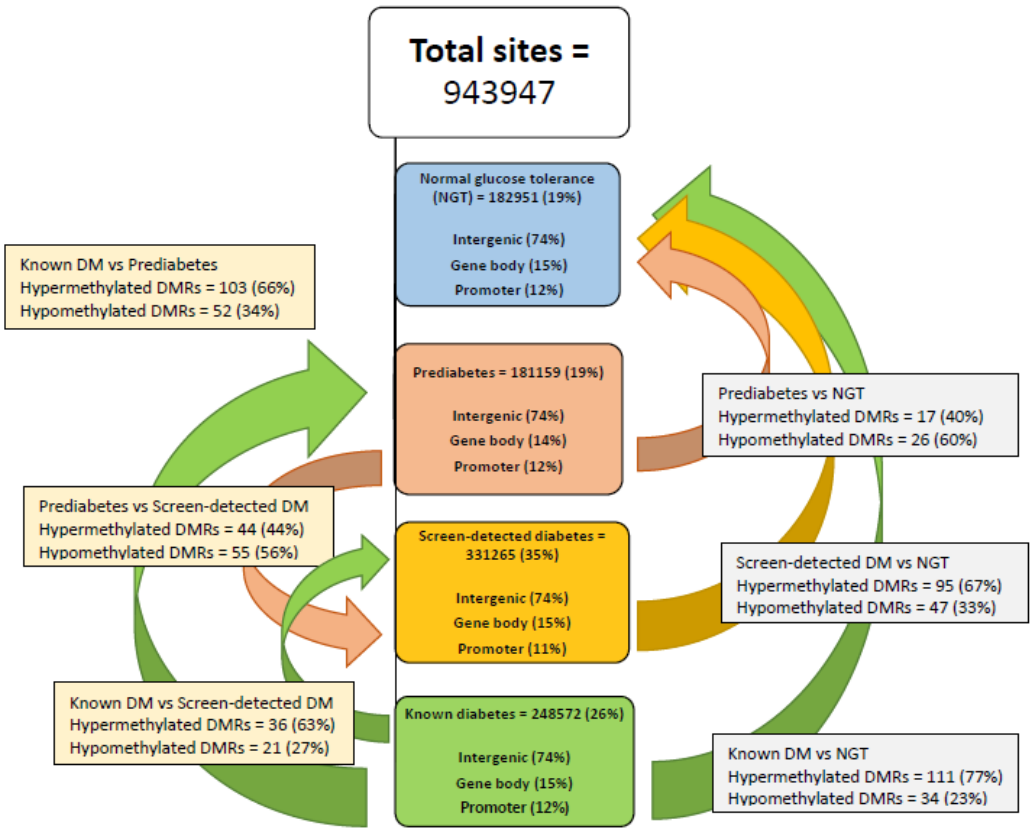


Figure 3.19: LncRNA-associated MeDIP enriched regions (peaks) identified in all 48 samples. The total sites are the sum of the peak number for the subjects with known diabetes (known DM), screen-detected diabetes (screen-detected DM); prediabetes and normal glucose tolerance (NGT). The distribution of the peaks in the intergenic, gene body and promoter regions are shown. The hypermethylated and hypomethylated lncRNA-associated DMRs within the gene promoter for the various glucose tolerance groups are shown

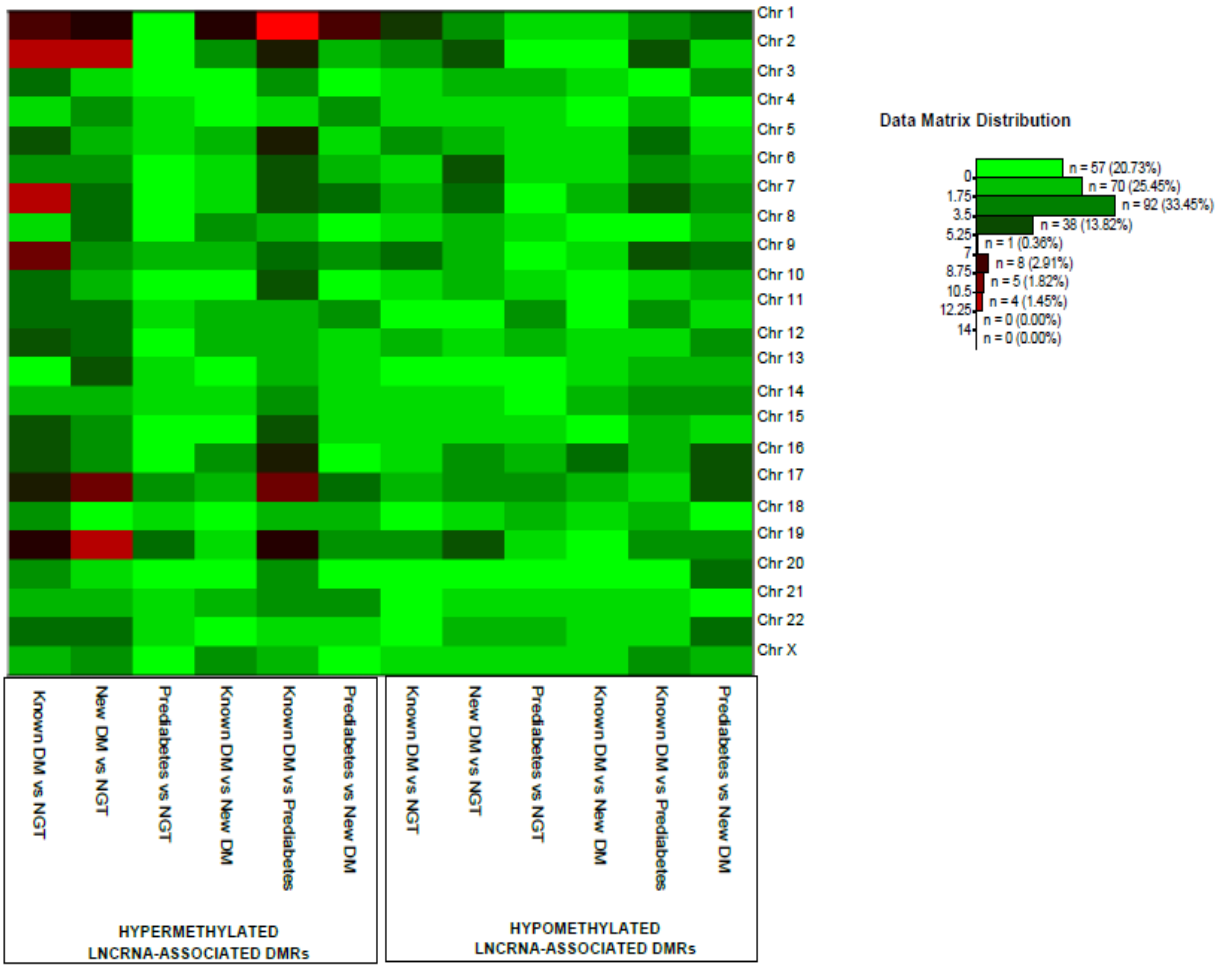


Figure 3.20: Heatmap presenting the distributions of lncRNA-associated differentially methylated regions (DMRs) among the chromosomes for all six comparison groups, namely known diabetes versus NGT; screen-detected diabetes versus NGT; prediabetes versus NGT; known diabetes versus screen-detected diabetes; known diabetes versus prediabetes and prediabetes versus screen-detected diabetes. Both hypermethylated and hypomethylated DMRs are presented

Tables 3.8 and 3.9 summarises the lncRNA-associated DMRs according to the top ten DMRs per fold change for each of the six comparison groups. When observing all the common lncRNA-associated DMRS for the diabetic subjects (known diabetes and screen-detected diabetes), there were 11 hypermethylated and 14 hypomethylated DMRS when compared to the NGT subjects (Table 3.10). The hypermethylated lncRNA-associated DMRs included AC096669.2, AK126635, ARL17B, BAGE3, MLXIP, NGEF, PCMTD2, RPTOR, TMEM9, XLOC_002527 and XLOC_007696 whereas the hypomethylated lncRNA-associated DMRs included BC029473, C4A, CCDC53, CROCCP2, CTD-2154I11.2, FAM225B, MBD3, POMT1, POTEB2, RP11-119F7.5, RP11-219A15.4, RP11-429J17.4, RP11-77K12.5 and SLC35B4.

Table 3.8: The top 10 hypermethylated lncRNA-associated DMRs for all comparison groups, known diabetes versus NGT; known diabetes versus screen-detected diabetes; known diabetes versus prediabetes; screen-detected diabetes versus NGT; prediabetes versus NGT and prediabetes versus screen-detected diabetes

HYPERMETHYLATED lncRNA-associated DMRs

Known DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
SLC26A9	chr1:205895421-205895620	2.69	1.32×10^{-7}
ZNF346	chr5:176448161-176448360	2.28	4.34×10^{-7}
LOC100130075	chr12:69198841-69199100	2.28	1.48×10^{-7}
AF258560	chr16:24930681-24930880	2.2	2.33×10^{-7}
ARL17B	chr17:44429021-44429220	2.19	2.74×10^{-7}
PRRC2B	chr9:134364781-134365100	2.17	5.34×10^{-9}
XLOC_002527	chr2:238039241-238039560	2.05	1.20×10^{-7}
LINC00521	chr14:94461821-94462080	2.04	1.11×10^{-7}
CTD-2354A18.1	chr18:70993201-70993440	2.02	3.63×10^{-8}
IGH	chr16:33753381-33753580	1.99	8.29×10^{-7}

Known DM vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
SLC26A9	chr1:205895421-205895620	1.91	2.22×10^{-7}
FAM223A	chrX:153859601-153859800	1.82	3.06×10^{-7}
SDK2	chr17:71432461-71432680	1.69	2.75×10^{-7}
XAGE1B	chrX:52260741-52261020	1.66	5.35×10^{-7}
SCRIB	chr8:144877441-144877660	1.62	4.72×10^{-7}
KIAA1467	chr12:13198981-13199200	1.62	2.21×10^{-7}
AK092098	chr11:63591421-63591720	1.58	1.11×10^{-7}
ATP5J	chr21:27105221-27105420	1.56	6.07×10^{-6}
AF420437	chr1:146216561-146217120	1.49	2.52×10^{-6}
ZNF346	chr5:176448161-176448360	1.47	6.99×10^{-6}

Known DM vs Prediabetes

Gene Name	Genomic Coordinates	log2FC	p-value
POMP	chr13:29231801-29232000	2.84	9.00×10^{-8}
XLOC_008014	chrX:73047161-73047360	2.66	8.29×10^{-8}
AK128525	chr2:89160101-89160340	2.29	1.49×10^{-7}
KIAA1467	chr12:13198981-13199200	2.17	1.12×10^{-7}
ZNF346	chr5:176448161-176448360	2.14	8.49×10^{-5}
RP11-502H18.2	chr1:180093741-180094000	2.13	3.34×10^{-7}
KLK13	chr19:51567281-51567500	1.95	8.58×10^{-7}
XLOC_001908	chr2:238120261-238120720	1.92	2.28×10^{-7}
AX748236	chr10:6568081-6568280	1.92	1.14×10^{-7}
RP11-944C7.1	chr14:90101381-90101600	1.9	1.25×10^{-7}

Screen-detected DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
FXYD5	chr19:35647241-35647460	2.84	2.16×10^{-9}
CTD-2534I21.9	chr17:43061181-43061480	2.4	1.48×10^{-7}
ARL17B	chr17:44430841-44431040	2.38	2.38×10^{-7}
VANGL2	chr1:160389941-160390140	2.28	1.04×10^{-6}
RP11-332J15.2	chr5:6867621-6867880	2.24	7.23×10^{-9}
XLOC_013291	chr19:30186641-30186900	2.18	6.88×10^{-7}
LINC00442	chr13:19582301-19582500	2.17	5.42×10^{-8}
SHKBP1	chr19:41084061-41084280	2.16	5.75×10^{-10}
GTF2I	chr7:74168921-74169120	2.06	3.08×10^{-6}
LTBP4	chr19:41116601-41116800	2.05	4.81×10^{-5}

Prediabetes vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
MIMT1	chr19:57352241-57352440	2.16	9.45×10^{-8}
XLOC_l2_014834	chr9:69728801-69729040	1.99	3.91×10^{-7}
DQ572107	chr17:62120181-62120380	1.99	4.31×10^{-7}
RP11-357H3.1	chr18:73727201-73727400	1.88	3.87×10^{-10}
LOC100288123	chr19:1820041-1820240	1.85	2.33×10^{-7}
LINC01508	chr9:93195001-93195220	1.8	6.04×10^{-7}
RP11-676J12.8	chr17:694041-694240	1.79	1.13×10^{-6}
GRTP1-AS1	chr13:114005801-114006000	1.72	1.15×10^{-6}
AP001046.6	chr21:44783281-44783480	1.72	7.68×10^{-7}
RP11-75C10.7	chr17:75466201-75466440	1.67	3.92×10^{-7}

Prediabetes vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
SEMA4F	chr2:74879901-74880100	1.88	2.20×10^{-7}
MAN1C1	chr1:26012381-26012620	1.57	1.95×10^{-7}
XLOC_000777	chr1:33173461-33173660	1.53	2.92×10^{-7}
EHMT1-IT1	chr9:140656061-140656260	1.52	1.35×10^{-7}
RP11-782C8.5	chr1:143226161-143226400	1.48	2.74×10^{-6}
BC071797	chr21:9709901-9710100	1.45	4.48×10^{-6}
ADAMTS1	chr21:28212361-28212640	1.45	7.40×10^{-8}
RP5-905H7.8	chr7:62815041-62815240	1.44	2.64×10^{-7}
XLOC_l2_011193	chr4:93104781-93105020	1.42	1.51×10^{-6}
METTL23	chr17:74724801-74725000	1.38	1.50×10^{-5}

Table 3.9: The top 10 hypomethylated lncRNA-associated DMRs for all comparison groups, known diabetes versus NGT; known diabetes versus screen-detected diabetes; known diabetes versus prediabetes; screen-detected diabetes versus NGT; prediabetes versus NGT and prediabetes versus screen-detected diabetes

HYPOMETHYLATED lncRNA-associated DMRs

Known DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
XLOC_004212	chr4:186941421-186941700	-1.02	3.71×10^{-6}
BRF1	chr14:105689281-105689540	-1.03	9.73×10^{-5}
THAP4	chr2:242540201-242540400	-1.06	3.16×10^{-5}
CROCCP2	chr1:16959961-16960280	-1.06	1.27×10^{-6}
EIF4A3	chr17:78113501-78113740	-1.11	7.76×10^{-6}
TMEM214	chr2:27253761-27253980	-1.13	3.31×10^{-7}
EIF2B3	chr1:45452441-45452640	-1.14	2.67×10^{-7}
HOXA3	chr7:27179421-27179740	-1.15	1.34×10^{-6}
CUX2	chr12:111651881-111652080	-1.15	4.87×10^{-5}
VRK3	chr19:50527041-50527300	-1.16	2.26×10^{-6}

Known DM vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
SSH1	chr12:109199901-109200160	-1.12	7.43×10^{-6}
XLOC_005639	chr6:21980601-21980860	-1.13	9.27×10^{-5}
RP11-458D21.1	chr1:145380441-145380780	-1.19	2.76×10^{-5}
LOC100506603	chr14:77252181-77252640	-1.19	1.55×10^{-6}
AK125727	chr14:77252181-77252640	-1.19	1.55×10^{-6}
AP001476.3	chr21:47477561-47477760	-1.24	9.48×10^{-5}
BC034416	chr3:180586661-180586880	-1.31	1.44×10^{-5}
RN7SL367P	chr16:1946361-1946700	-1.35	6.51×10^{-6}
RP11-586K12.4	chr16:32752701-32752900	-1.37	1.00×10^{-5}
EIF3B	chr7:2412041-2412260	-1.37	5.80×10^{-6}

Known DM vs Prediabetes

Gene Name	Genomic Coordinates	log2FC	p-value
RP11-79P5.3	chr5:72705681-72705960	-1.01	1.78×10^{-6}
RP11-67H24.2	chr16:32821721-32822020	-1.01	5.89×10^{-5}
XLOC_005521	chr6:159549481-159549940	-1.02	7.46×10^{-6}
RP4-545C24.1	chr7:143890801-143891200	-1.03	2.16×10^{-5}
RCAN3	chr1:24861401-24861600	-1.06	7.29×10^{-6}
CHRNB4	chr15:78953061-78953460	-1.06	4.32×10^{-6}
ADAMTS1	chr21:28212361-28212640	-1.11	4.90×10^{-6}
RN7SKP204	chr6:13547661-13548000	-1.12	1.67×10^{-7}
BC015435	chr1:228673261-228673560	-1.15	1.05×10^{-5}
PMS2L14	chr7:74928941-74929300	-1.16	3.25×10^{-5}

Screen-detected DM vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
TRMU	chr22:46748281-46748680	-1	6.26×10^{-7}
TFE3	chrX:48900201-48900480	-1	6.62×10^{-6}
SNAR-A2	chr19:48438741-48438980	-1	2.21×10^{-5}
RP11-503C24.2	chr6:168628021-168628220	-1	9.35×10^{-6}
RP11-27N21.3	chr8:80697521-80697720	-1.01	1.61×10^{-5}
POTEB2	chr15:21070661-21070880	-1.01	2.68×10^{-5}
AC011298.2	chr2:241624101-241624300	-1.02	2.03×10^{-5}
RP11-219A15.4	chr17:16689761-16690120	-1.03	1.32×10^{-5}
PASK	chr2:242072801-242073060	-1.04	2.59×10^{-6}
RP11-119F7.5	chr10:70456261-70456580	-1.05	2.10×10^{-6}

Prediabetes vs NGT

Gene Name	Genomic Coordinates	log2FC	p-value
ZNF556	chr19:2867381-2867620	-1.02	5.88×10^{-7}
KIAA1257	chr3:128721801-128722040	-1.09	4.52×10^{-5}
MIR210HG	chr11:570301-570520	-1.12	5.30×10^{-5}
WBP2	chr17:73847441-73847820	-1.13	5.23×10^{-5}
MFSD7	chr4:678821-679060	-1.14	5.47×10^{-5}
RPS13	chr11:17100501-17100760	-1.15	9.31×10^{-5}
RP11-524F11.1	chr17:17409041-17409240	-1.2	8.63×10^{-6}
HAUS1	chr18:43685801-43686080	-1.22	5.32×10^{-5}
ADSL	chr22:40759661-40759880	-1.26	7.70×10^{-7}
TFF3	chr21:43732601-43732860	-1.29	1.51×10^{-5}

Prediabetes vs Screen-detected DM

Gene Name	Genomic Coordinates	log2FC	p-value
XLOC_009483	chr11:69000461-69000660	-1.02	8.53×10^{-6}
AL928742.12	chr14:106065161-106065560	-1.04	4.80×10^{-5}
XLOC_l2_013648	chr7:149730661-149731040	-1.05	8.38×10^{-5}
MIIP	chr1:12080981-12081260	-1.09	5.18×10^{-5}
RP11-1260E13.1	chr17:171381-171760	-1.11	1.39×10^{-5}
KIAA0586	chr14:58908521-58908840	-1.11	6.72×10^{-5}
RDH8	chr19:10126581-10126940	-1.12	6.10×10^{-5}
RP11-158I23.1	chr3:125777161-125777520	-1.14	8.29×10^{-8}
Z99756.1	chr22:43671621-43672000	-1.17	1.68×10^{-6}
RP4-751H13.7	chr7:149558981-149559180	-1.2	8.05×10^{-5}

Table 3.10: Hypermethylated and hypomethylated lncRNA-associated DMRs in known diabetes and screen-detected diabetes versus NGT

Gene Name	Accession number	Known DM vs NGT fold change	p-value	Screen-detected DM vs NGT fold change	p-value
Hypermethylated lncRNA-associated DMRs					
AC096669.2	ENST00000414300	1.47	4.05×10^{-8}	1.19	2.70×10^{-5}
AK126635	uc001lsk.1	1.04	7.04×10^{-5}	1.22	4.58×10^{-6}
ARL17B	ENST00000572991	2.19	2.74×10^{-7}	2.38	2.38×10^{-7}
BAGE3	uc002yix.2	1.44	1.64×10^{-9}	1.18	1.48×10^{-5}
MLXIP	ENST00000541750	1.27	9.71×10^{-7}	1.2	4.71×10^{-7}
NGEF	ENST00000489127	1.21	3.82×10^{-5}	1.19	2.99×10^{-5}
PCMTD2	ENST00000266078	1.74	1.23×10^{-8}	1.15	9.99×10^{-6}
RPTOR	ENST00000575542	1.53	1.78×10^{-6}	1.46	2.93×10^{-7}
TMEM9	ENST00000472411	1.56	4.64×10^{-6}	1.77	5.46×10^{-8}
XLOC_002527	TCONS_00004618	2.05	1.20×10^{-7}	1.65	2.90×10^{-7}
XLOC_007696	TCONS_00016322	1.38	2.27×10^{-8}	1.38	1.16×10^{-7}
Hypomethylated lncRNA-associated DMRs					
BC029473	uc001ejd.1	-2.27	3.11×10^{-7}	-1.93	5.96×10^{-7}
C4A	ENST00000463034	-1.81	7.70×10^{-7}	-1.13	1.27×10^{-5}
CCDC53	ENST00000544341	-1.28	1.53×10^{-6}	-1.05	1.04×10^{-6}
CROCCP2	ENST00000412962	-1.06	1.27×10^{-6}	-1.51	8.05×10^{-7}
CTD-2154I11.2	ENST00000512486	-2.02	1.94×10^{-7}	-1.82	2.69×10^{-7}
FAM225B	ENST00000439875	-1.51	1.07×10^{-6}	-1.47	7.54×10^{-7}
MBD3	ENST00000590830	-1.33	2.98×10^{-6}	-1.26	6.72×10^{-7}
POMT1	ENST00000485278	-1.71	5.42×10^{-7}	-1.25	7.40×10^{-7}
POTEB2	NR_102390_1	-1.22	7.61×10^{-5}	-1.01	2.68×10^{-5}
RP11-119F7.5	ENST00000562082	-1.6	4.45×10^{-7}	-1.05	2.10×10^{-6}
RP11-219A15.4	ENST00000602730	-2.56	1.65×10^{-7}	-1.03	1.32×10^{-5}
RP11-429J17.4	ENST00000527579	-1.29	7.95×10^{-5}	-1.6	3.65×10^{-7}
RP11-77K12.5	ENST00000575421	-1.57	7.12×10^{-7}	-1.38	4.76×10^{-7}
SLC35B4	ENST00000416907	-1.45	2.02×10^{-5}	-1.06	3.10×10^{-6}

Furthermore, when including the non-diabetic hyperglycaemia to the diabetic subjects, the common hypomethylated lncRNA-associated DMRs found were C4A, CCDC53, CTD-2154I11.2, POTEB2 and RP11-429J17.4 when compared to the NGT subjects (Table 3.11). LncRNA-associated DMRs were also summarised for the known diabetic subjects on metformin treatment compared to the screen-detected (newly diagnosed) diabetic subjects (Table 3.12). There were 36 hypermethylated and 21 hypomethylated lncRNA-associated DMRs identified. Amongst the hypermethylated lncRNA-associated DMRs, ATP5J and LOXL2 were highlighted as having a possible association with diabetes.

Table 3.11: Hypermethylated and hypomethylated lncRNA-associated DMRs in known diabetes, screen-detected diabetes and prediabetes versus NGT

Gene Name	Accession number	Known DM vs NGT fold change	p-value	Screen-detected DM vs NGT fold change	p-value	Prediabetes vs NGT fold change	p-value
Hypermethylated lncRNA-associated DMRs							
none							
Hypomethylated lncRNA-associated DMRs							
C4A	ENST00000463034	-1.81	7.70×10^{-7}	-1.13	1.27×10^{-5}	-1.43	1.13×10^{-6}
CCDC53	ENST00000544341	-1.28	1.53×10^{-6}	-1.05	1.04×10^{-6}	-1.61	4.91×10^{-7}
CTD-2154I11.2	ENST00000512486	-2.02	1.94×10^{-7}	-1.82	2.69×10^{-7}	-1.88	2.24×10^{-7}
POTEB2	NR_102390_1	-1.22	7.61×10^{-5}	-1.01	2.68×10^{-5}	-1.44	2.72×10^{-5}
RP11-429J17.4	ENST00000527579	-1.29	7.95×10^{-5}	-1.6	3.65×10^{-7}	-1.53	3.44×10^{-7}

Table 3.12: Hypermethylated and hypomethylated lncRNA-associated DMRs in known diabetes versus screen-detected diabetes

Gene Name	Accession number	Known DM vs Screen-detected DM fold change	p-value
Hypermethylated lncRNA-associated DMRs			
SLC26A9	ENST00000491127	1.91	2.22×10^{-7}
FAM223A	NR_027401_2	1.82	3.06×10^{-7}
SDK2	ENST00000479356	1.69	2.75×10^{-7}
XAGE1B	NR_033254_2	1.66	5.36×10^{-7}
SCRIB	ENST00000525051	1.62	4.72×10^{-7}
KIAA1467	ENST00000416494	1.62	2.21×10^{-7}
AK092098	uc001nxu.1	1.58	1.11×10^{-7}
ATP5J	ENST00000486002	1.56	6.07×10^{-6}
AF420437	uc021ove.1	1.49	2.52×10^{-6}
ZNF346	NR_131773	1.47	6.99×10^{-6}
AX747590	uc003wwb.1	1.46	9.42×10^{-7}
AK128525	uc002sti.2	1.45	3.82×10^{-7}
XLOC_007349	TCONS_00015975	1.4	6.19×10^{-7}
FKBP8	ENST00000601844	1.4	3.83×10^{-7}
LOC101927468	NR_120331	1.38	5.35×10^{-7}
CTAGE15	ENST00000447022	1.38	2.62×10^{-7}
AF258560	uc002dnd.1	1.38	8.09×10^{-6}
LOXL2	ENST00000520925	1.35	4.57×10^{-6}
AC016644.1	ENST00000438553	1.35	6×10^{-5}
RP11-14N7.2	ENST00000457390	1.34	1.16×10^{-5}
AP001476.4	ENST00000429512	1.34	6.69×10^{-5}
RP3-399L15.2	ENST00000435802	1.31	2.95×10^{-5}
AK310441	uc009wkv.1	1.28	8.61×10^{-7}
RP11-423O2.7	ENST00000424640	1.25	5.17×10^{-5}
LOC101928402	NR_130770	1.21	1.9×10^{-5}
LINC00521	NR_024182	1.18	1.07×10^{-5}
TBCE	uc010pxq.1	1.16	7.23×10^{-7}
SMIM22	ENST00000591004	1.13	4.44×10^{-6}
XLOC_l2_000395	TCONS_l2_00000554	1.11	9.64×10^{-5}
SEMA4C	ENST00000482925	1.11	3.28×10^{-5}
LOC101929378	NR_110250	1.08	2.24×10^{-5}
ZNF169	ENST00000492115	1.07	8.86×10^{-5}
GNPTG	ENST00000527076	1.07	6.09×10^{-6}
TIMELESS	ENST00000557589	1.05	2.72×10^{-6}
RP13-638C3.3	ENST00000575085	1.02	2.17×10^{-5}
XLOC_009584	TCONS_00019787	1	1.07×10^{-5}

Hypomethylated lncRNA-associated DMRs			
SSH1	ENST00000546433	-1.12	7.44×10^{-6}
XLOC_005639	TCONS_00011466	-1.13	9.27×10^{-5}
RP11-458D21.1	ENST00000433081	-1.19	2.76×10^{-5}
LOC100506603	NR_104183	-1.19	1.55×10^{-6}
AK125727	uc001xsu.1	-1.19	1.55×10^{-6}
AP001476.3	ENST00000435738	-1.24	9.48×10^{-5}
BC034416	uc011bqi.2	-1.31	1.44×10^{-5}
RN7SL367P	ENST00000584097	-1.35	6.51×10^{-6}
RP11-586K12.4	ENST00000561479	-1.37	1×10^{-5}
EIF3B	ENST00000475415	-1.37	5.8×10^{-6}
ANKIB1	ENST00000422095	-1.37	3.43×10^{-5}
XLOC_010373	TCONS_00021545	-1.38	3.32×10^{-5}
RP11-510M2.5	ENST00000568523	-1.38	6.23×10^{-6}
OPN1MW	ENST00000468495	-1.39	1.64×10^{-5}
ELAC2	ENST00000484122	-1.45	3.08×10^{-5}
SLC25A35	ENST00000585311	-1.52	7.49×10^{-6}
AK095057	uc021yjh.1	-1.53	8.5×10^{-5}
C18orf8	ENST00000590870	-1.55	2.33×10^{-6}
RP11-168K11.3	ENST00000437712	-1.62	8.8×10^{-6}
LL22NC03-N27C7.1	ENST00000602816	-1.63	2.06×10^{-5}
CRAMP1L	ENST00000467286	-1.76	3.72×10^{-5}

CHAPTER 4

DISCUSSION

Epigenetic changes have been shown to predispose to disease or occur once the disease has developed (Davegårdh *et al.*, 2018). DNA methylation, the most widely studied epigenetic mechanism, targets unmethylated CpG nucleotides with the aid of DNA methyltransferases (DNMTs) enzymes (Ahmed *et al.*, 2020). These CpG nucleotides which occur at high-frequency in the promoter regions of genes can undergo hyper- or hypomethylation which can result in transcriptional suppression or activation of the affected genes (Du *et al.*, 2015; Pasquier *et al.*, 2015). Several studies suggest that these epigenetic modifications may alter the transcriptional activity of genes and contribute to pathogenic conditions, such as the type 2 diabetes mellitus (T2DM) phenotype (Muka *et al.*, 2016; Bansal and Pinney, 2017). The exact mechanism of DNA methylation's involvement in the pathogenesis and development of T2DM and its association with cardiometabolic traits is still unclear, especially within an Africa context. The aim of the current study was therefore to conduct genome-wide DNA methylation in 48 subjects within a South African context, with varying glucose tolerance and investigate the relationship between the observed DMRs and cardiometabolic risk factors. The subjects were divided into four groups according to whether they were known diabetics on metformin treatment; newly diagnosed or screen-detected diabetics; had prediabetes and finally those with normal glucose tolerance (NGT). Comparisons were made between all four categories for both hypermethylated and hypomethylated DMRs as well as lncRNA-associated DMRs observed. Methylation observed in the DMRs was validated using pyrosequencing. Furthermore, gene ontology and KEGG pathway analysis were performed for all four categories.

4.1 Differentially methylated regions (DMRs)

Following the completion of the genome-wide DNA methylation analyses, several DMRs were observed. Overall 366 DMRs were identified, of which 63% were hypermethylated and 37% hypomethylated. The hypermethylated DMRs were more commonly observed on chromosome 19 whereas a large number of hypomethylated DMRs were found on chromosomes X and 1. More than 70% of the DMRs identified in the known diabetes and screen-detected diabetes were hypermethylated which could contribute to the hyperglycaemic

profile of these individuals. When observing all the DMRs for the diabetic subjects (known diabetes and screen-detected diabetes), there were 11 hypermethylated and 7 hypomethylated DMRs commonly found when comparing them to the NGT group. When the groups were further categorised to include non-diabetic subjects with hyperglycaemia (prediabetes) and compared to NGT, 1 hypermethylated DMR and 4 hypomethylated DMRs were observed.

4.1.1 DMRs identified in hyperglycaemic subjects

The common hypermethylated DMRs identified in the diabetic and non-diabetic hyperglycaemic subjects (known, screen-detected and prediabetes) when compared to the NGT subjects included one hypermethylated DMR, TMEM89 and four hypomethylated DMRs namely CCDC53, POTEB, POTEB2 and RHOBTB3. Although the fold change for the hypermethylated DMR TMEM89 was highest in the prediabetic subjects and suggests possible repression of gene expression, no direct link to diabetes is evident. TMEM (transmembrane) proteins are thought to be components of various cell membranes including mitochondrial, endoplasmic reticulum, lysosomes and Golgi membranes serving as channels for the transport of substances across these membranes (Schmit and Michiels, 2018). TMEM genes other than that of TMEM89 have been implicated in diabetes. Genome-wide association (GWAS) studies have loosely associated the involvement of TMEM135, TMEM195 and TMEM163 in T2DM in individuals from diverse population groups (Ramos *et al.*, 2011; Tabassum *et al.*, 2013; Chidambaram *et al.*, 2016). Ramos *et al* (2011) showed that several single nucleotide polymorphisms (SNPs) including DKGB-TMEM195, which mainly affects beta-cell function, was associated with fasting plasma glucose and T2DM in a population of African-Americans. The study identified shared loci associated with fasting plasma glucose across populations by replicating SNPs associated with T2DM in populations with European ancestry (Ramos *et al.*, 2011). The study by Tabassum *et al* (2013) observed variants within TMEM163 that showed association with decreased fasting plasma insulin and impaired insulin secretion, thereby modulating T2DM susceptibility in the Indian cohort studied. Since TMEM163 is expressed in certain brain regions and within subpopulations of nerve terminals, the loci identified in the study suggested a role in the neurologic aetiology of T2DM (Tabassum *et al.*, 2013). Similarly, Chidambaram and co-workers (2016) evaluated the association of genetic variants previously associated with T2DM in populations of European ancestry in an Asian-Indian population with early-onset T2DM. Variants in TMEM135 was suggested to have

a loose association with T2DM as it had a less robust replication in the GWAS in comparison to other genes. The reason could be that certain variants may differ in their frequency between different ethnic groups as well as various forms of diabetes such as early- and late-onset T2DM (Chidambaram *et al.*, 2016).

The association of TMEM89 observed in the current study warrants further investigation particularly in prediabetic individuals as a potential marker for the progression to T2DM. Hypermethylation of this DMR with subsequent suppression of the transmembrane protein function could suggest dysfunction of the transport of ions and glucose across cellular membranes. Additionally, studies such as the ones mentioned above have highlighted the benefits of investigating genes in diverse populations other than those from European ancestry.

4.1.2 DMRs identified in subjects with diabetes (known and newly diagnosed)

Amongst the hypomethylated DMRs identified in diabetic individuals (includes known and screen-detected subjects) was Adenosine Kinase (ADK) which encodes the metabolic enzyme adenosine kinase. ADK plays a role in cellular metabolism by regulating the intracellular and extracellular concentrations of adenosine through the phosphorylation of adenosine to AMP (Boison, 2013; El-Kharrag *et al*., 2019). Adenosine, a purine nucleoside, plays a role in glucose metabolism through interactions with insulin, glucagon and lipolysis (Boison, 2013). It has been shown to accumulate under stress, tissue injury and inflammation, especially in cells and tissue of the liver, pancreas, muscle and fat. In addition, any decrease in ATP has shown to cause an increase in adenosine levels (Antonioli *et al.*, 2015). Any changes in ADK activity impacts adenosine concentration and ADK dysfunction has been linked to pathological disorders such as epilepsy, cancer and diabetes (Kiese *et al.*, 2016). In the current study, ADK was hypomethylated which could indicate an increase in expression of the enzyme activity in both known diabetics on treatment and screen-detected diabetics. This would be expected if one assumed that adenosine levels would be increased due to the stress caused by diabetes on cells and tissue. ADK has also shown to increase when greater clearance of adenosine is required (Kiese *et al.*, 2016). Adenosine accumulation has also been shown to affect DNA methylation. Adenosine is an end product of DNA methylation where the transfer of a methyl group from S-adenosylmethionine (SAM) to DNA which results

in the formation of *S*-adenosyl-L-homocysteine (SAH). Increased levels of adenosine reverse the direction of the *S*-adenosyl-L-homocysteine (SAH)-hydrolase reaction resulting in an increase in SAH which is known to inhibit DNA methyltransferases (DMNTs) (Boison, 2013).

Another hypomethylated DMR observed in all diabetic subjects was SLC35B4 which is part of the SLC35 solute transporter family. SLC35B4 is involved in the transport of nucleotide sugars, particularly UDP-xylose (UDP-Xyl) and UDP-N-acetylglucosamine (UDP-GlcNAc) from the cytoplasm to the Golgi apparatus where they are utilised (Ashikov *et al.*, 2005; Hadley *et al.*, 2019). UDP-GlcNAc is the end-product of the hexosamine biosynthesis pathway whereas UDP-Xyl is involved in glycosaminoglycan biosynthesis (Wex *et al.*, 2018). UDP-GlcNAc is considered to be a sensor for the nutritional state of the cell as it integrates glucose, glutamine, fatty acids, uridine and ATP metabolism (Wex *et al.*, 2018). The intracellular concentration levels of UDP-GlcNAc have been linked to diabetes and insulin resistance. Hyperglycaemia leads to elevated UDP-GlcNAc synthesis through the hexosamine biosynthesis pathway and in turn increased protein *O*-GlcNAcylation. O-GlcNAcylation regulates many cellular functions including signal transduction, gene expression, protein degradation and stress response, and increased levels of *O*-GlcNAcylation lead to insulin resistance (Vigetti *et al.*, 2012; Ma and Hart, 2013). It has been suggested that O-GlcNAcylation may be a useful screening tool to assess varying glucose metabolism amongst individuals (Park *et al*., 2010; Springhorn *et al.*, 2012; Myslicki *et al.*, 2014). In a study by Myslicki *et al* (2014), it was demonstrated that O-linked β-N-acetylglucosamine (O-GlcNAc) had greater sensitivity to metabolic status when compared to HbA1C which is considered the gold standard for monitoring diabetes diagnosis and progression (Myslicki *et al.*, 2014). Similarly, Park *et al* (2010) observed increased expression of O-linked β-N-acetylglucosaminidase (O-GlcNAcase), an enzyme involved in O-GlcNAcylation, in erythrocytes of individuals with prediabetes and diabetes when compared to control subjects (Park *et al*., 2010). Springhorn *et al* (2012) also demonstrated that leukocytes from prediabetic and diabetic individuals had elevated global O-GlcNAc levels compared to that of the healthy controls (Springhorn *et al.*, 2012). It has been found that HbA1C is not as sensitive in detecting prediabetes as it does not significantly reflect the glycaemic variance but only the mean blood glucose (Derr *et al.*, 2003; Wang *et al.*, 2009; Myslicki *et al.*, 2014)). The above studies, therefore, show that besides HbA1C, the hexosamine biosynthesis pathway and its components such as O-GlcNAc have the potential to be used in the earlier detection of prediabetes and full-blown diabetes.

In a study by Chen and co-workers, increased SLC35B4 expression was found in subcutaneous adipose tissue in obese subjects, linking SLC35B4 to obesity-induced T2DM (Chen *et al.*, 2013). Furthermore, it was also shown that SLC35B4 controlled hepatic glucose production *in vitro* in liver-derived tissue culture cells (Yazbek *et al.*, 2011; Wex *et al.*, 2018). In addition, Yazbek and co-workers also observed both *in vivo* and *in vitro* studies in mice showed that decreased SLC35B4 expression was associated with decreased gluconeogenesis (Yazbek *et al.*, 2011). In the current study due to hypomethylation, the SLC35B4 expression in diabetic subjects is thought to be increased. This is in line with the profile of T2DM as increased gluconeogenesis is considered to be the main cause of fasting hyperglycaemia (Hatting *et al.*, 2018). Moreover increased expression of SLC35B4 transporters would be expected with hyperglycaemia due to its association with the nutrient sensor UDP-GlcNAc. Hyperglycaemia leads to increased UDP-GlcNAc which in turns needs to be transported by SLC35B4 transporters to the Golgi apparatus for the glycosylation process.

The common hypermethylated DMRs identified in diabetic individuals included BAGE2, BAGE3, BAGE4, BAGE5, CD248, COL8A2, IGSF9, PACSIN1, SPACA3, SYT3 and TMEM89. These DMRs do not appear to have a direct relationship with T2DM, however, a number of them, namely CD248 and TMEM89 are related to cancer in terms of fibroblast activity and metastasis, respectively (Schmit and Michiels, 2018; Teicher, 2019).

4.1.3 DMRs in known versus screen-detected diabetic subjects

Genes in the SLC family, specifically SLC25A35 and SLC28A1, were observed as hypomethylated DMRs when comparing known diabetic subjects on metformin treatment to the screen-detected diabetic subjects. The SLC family is known for its importance in drug development and its association with metabolic diseases (Rives *et al*., 2017; Zhang *et al.*, 2018). Proteins encoded by the SLC family include passive transporters, symporters and antiporters and facilitate the movement of a specific substrate across the cell or organelle membranes (Zhang *et al.*, 2018).

SLC28A1 encodes the sodium dependant concentrative nucleoside transporter (CNTs) proteins. The CNTs play a role in nucleoside homeostasis and are involved in mediating the

transport of nucleosides necessary for nucleic acid synthesis (Pastor-Anglada *et al.*, 2008; Pastor-Anglada and Pérez-Torras, 2015). SLC28A1, therefore, indirectly plays a role in the control of cell and tissue growth. In addition, members of the SLC28 transporter family also have a pharmacological role as they mediate the transport of many clinical drugs associated with anticancer and antiviral therapies (Gray *et al*., 2004; Wang and Buolamwini, 2019). CNT-1, a high-affinity pyrimidine nucleoside transporter is localised mainly in epithelial tissue and predominantly expressed in the kidney and jejunum. The involvement of SLC28A1 and its transporter CNT-1 in renal absorption, has shown to be impaired by diabetes, with this impairment being an early marker in the development of diabetic nephropathy (Rodríguez-Mulero *et al.*, 2005).

On the other hand, sequence analysis of SLC25A35 indicates that it likely functions as an oxaloacetate carrier which implies mitochondrial association (Haitina *et al.*, 2006). The SLC25 family of genes encode mitochondrial carriers that are localized in the inner mitochondrial membrane and these mitochondrial transporters have been functionally implicated in some metabolic pathways including oxidative phosphorylation, TCA cycle, fatty acid oxidation, gluconeogenesis, lipogenesis, urea synthesis, amino acid degradation and regulation of nucleotide and deoxynucleotide pools in the mitochondrial matrix amongst others (Palmieri, 2013; Palmieri and Monné, 2016). Other members of the SLC25 family of mitochondrial carriers have been associated with diabetes. These include the uncoupling proteins (UCPs) which play a role in energy homeostasis. UCP1, UCP2 and UCP3 encoded by SLC25A7, SLC25A8 and SLC25A9 respectively have been linked to obesity and T2DM in human and animal models suggesting that the expression of SLC25A8 and SLC25A9 could predict the onset of diabetes in humans (Palmieri *et al*., 2020; Schumann *et al.*, 2020).

Besides ATP generation, mitochondria also play a role in reactive oxygen species (ROS) – mediated signalling, apoptosis, calcium signalling and haem synthesis amongst others (Montgomery, 2019). Any defects in these processes will affect cellular energy metabolism as well as affect several tissues and systems. Mitochondria dysfunction has been associated with changes in gene expression of mitochondrial markers, decrease in mitochondrial biogenesis, decreased mitochondrial content, reduced enzymatic activities of mitochondrial proteins, ROS generation, and decrease in mitochondrial activity such as oxidative phosphorylation (Montgomery and Turner, 2015; Montgomery, 2019). Mitochondrial dysfunction with a decrease in oxidative phosphorylation and ATP production and a high production of ROS has been linked to insulin resistance and T2DM (Rovira-Llopis *et al.*, 2017; Sergi *et al.*, 2019).

As SLC proteins transport various solutes across the mitochondrial membrane to partake in several metabolic pathways (Palmieri, 2013; Palmieri and Monné, 2016), the decrease in methylation and subsequent increase in gene expression of SLC transporters in this study could be indicative of the antidiabetic effect of metformin treatment. It is therefore likely that metformin in its demethylation action of SLC mitochondrial carriers could aid cell repair in these patients, however, this requires further investigation. Metformin treatment has been associated with lower methylation levels in SLC transporter genes as was shown in a study conducted on metformin transporter genes in liver tissue (García-Calzón *et al.*, 2017). Lower methylation levels in SLC22A1, SLC22A3 and SLC47A1 which encode for the three metformin transporters OCT1, OCT3, and MATE1 in the human liver had a similar DNA methylation pattern when compared to that of non-diabetic subjects and diabetics not on any treatment (García-Calzón *et al.*, 2017). Although the specific transporter genes SLC22A1, SLC22A3 and SLC47A1 were not detected in the current study, DNA methylation has shown to be tissue-specific which has previously been noted in other studies (Zhang *et al.*, 2013; Dujic *et al.*, 2017). The absence of these other SLC DMRs in the current study's dataset could be attributed to this phenomenon.

GABPA was identified as a hypermethylated DMR when comparing known diabetic subjects on metformin treatment to the screen-detected diabetic subjects. This gene encodes GA-binding protein transcription factor subunit alpha, also known as the nuclear respiratory factor 2 (NRF2) and thought to be involved in the respiratory chain and oxidative phosphorylation (Yang *et al.*, 2014). Furthermore, GABPA (NRF2) may be involved in the activation of cytochrome oxidase expression and thereby impact mitochondrial function (Ongwijitwat *et al.*, 2006). Cytochrome oxidase, also known as complex IV, is an enzyme in the final stages of the electron transport chain responsible for energy production in the form of ATP (Helling *et al.*, 2008; Van Der Schueren *et al.*, 2015). The absence of cytochrome oxidase would result in no energy production and could ultimately lead to cell death. Due to the hypermethylation of GABPA in the current study, it is assumed that the expression of this DMR has been suppressed. A decrease in GABPA expression would affect cytochrome oxidase activity and result in an increase in reactive oxygen species (ROS) and a decrease in ATP production. Reactive oxygen species, although a by-product of ATP generation, can result in damage to cellular components and ultimately lead to apoptosis and autophagy (Li *et al.*, 2020). Hyperglycaemia in diabetics has been shown to cause an increase in ROS which leads to activation of the stress signalling Jun N-terminal kinase (JNK) pathway (Solinas and Becattini,

2017). This JNK pathway can also be activated by the increase of free fatty acids, inflammatory cytokines like tumour necrosis factor-alpha (TNF-α) and endoplasmic reticulum stress induced by hyperglycaemia (Kaneto *et al.*, 2010). It is suggested that this stress signalling pathway is involved in the development of insulin resistance and beta-cell dysfunction. Hyperglycaemia-induced generation of ROS and oxidative stress have been associated with cell death (apoptosis) of beta-cells as well as cells from the heart, retina, kidneys and nervous system thereby implicating ROS in the complications of diabetes (Volpe *et al.*, 2018).

4.2 Gene Ontology and KEGG pathway analyses

Gene Ontology (GO) analysis was performed and enabled the classification of the observed DMRs according to the biological process, cellular component and molecular function domains. The GO analyses were completed for each of the six comparison pairs and highlighted all three domains for both the hypermethylated and hypomethylated DMRs. In addition, GO Directed Acyclic Graphs (DAG) was formulated and displayed as p-value trees showing the hierarchy of gene ontology for each of the GO domains. The top ten terms with the lowest p-value and were displayed in the p-value trees with the more significant terms highlighted in red. The functional analysis mapping of genes to the Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways were also performed. The functional analysis was completed for each of the six comparison pairs and highlighted the significant pathways for both the hypermethylated and hypomethylated DMRs.

4.2.1 Known diabetes versus NGT

The hypermethylated DMRS identified in known diabetics on metformin treatment when compared to the NGT subjects were associated with the cytokine-cytokine receptor interaction and the oxidative phosphorylation pathways. The cytokine-cytokine receptor interaction pathway included DMRs, IL12A1 and TNFSF8. IL12A1 encodes a subunit of Interleukin-12 (IL-12), an immune-regulatory cytokine that plays a role in the link between innate and acquired immune system as well as stimulates interferon production, particularly interferon-gamma (IFN-γ). In addition, IL-12 also promotes the maturation of cytotoxic T lymphocytes

and activates natural killer cells (Shen *et al.*, 2017). Elevated levels of IL-12 have been associated with T2DM and insulin resistance as exposure of beta-cells to these inflammatory cytokines have been shown to mediate beta-cell dysfunction and apoptosis (Wegner *et al.*, 2008; Weaver *et al*., 2015). Weaver *et al* (2015) have also shown that the disruption of IL-12 could act as a protection mechanism for beta-cells against the apoptosis associated with inflammatory cytokines (Weaver *et al*., 2015). TNFSF8, on the other hand, encodes for a cytokine in the tumour necrosis factor ligand super family that induces the proliferation of T-cells and acts as a pro-inflammatory cytokine (Chu, 2013). Studies have shown pro-inflammatory cytokines such as IL-12 and tumour necrosis factor (TNF) are increased in the hyperglycaemic environments encountered in T2DM subjects (Wen *et al.*, 2006; Wu *et al.*, 2010; Akash *et al*., 2018). These pro-inflammatory cytokines enhance insulin resistance in adipocytes, muscle and hepatic cells which leads to impaired insulin sensitivity and glucose homeostasis (Al-Shukaili *et al.*, 2013). Metformin has shown to decrease the production of pro-inflammatory cytokines such as tumour necrosis factor and interleukins (Hyun *et al.*, 2013) and thereby reduce inflammation in T2DM subjects (Cameron *et al.*, 2016). In addition, metformin has an antimicrobial effect facilitated by its inhibition of the mitochondrial complex 1 activity which affects ATP production and gluconeogenesis (Malik *et al.*, 2018; Xiao *et al.*, 2020). In this study, the hypermethylated DMRs and their related GO terms support the action of metformin. IL12A1 and TNFSF8 were associated with the GO biological process of defence to Gram-positive bacteria, the molecular function of tumour necrosis factor receptor binding and the cellular component of the mitochondrial respiratory chain. These GO terms were all hypermethylated suggesting that the expression of the related genes was repressed by the increased methylation. In a study by Park and co-workers (2019), it was postulated that methylation levels could play a role in controlling the expression of TNF as increased levels were shown to induce tissue damage and death resulting in the development of Diabetic kidney disease (Park *et al.*, 2019). Furthermore, both IL-12 and TNFSF8 play a role in the progression of atherosclerosis where chronic inflammation is associated with a continuous influx of immune cells. Increased glucose levels stimulate inflammatory reactions related to these cytokines and they in turn stimulate T-lymphocyte proliferation and migration to atherosclerotic plagues (Wegner *et al.*, 2008; Foks *et al.*, 2012). A reduction of these cytokines have been associated with a reduction of atherosclerosis (Wen *et al.*, 2006) and metformin itself has also been shown to have a cardio-protective effect reducing atherosclerosis (Luo *et al.*, 2019). Therefore, in this current study, the reduction of TNFSF8 and IL-12 expression and inflammatory effects could be a result of metformin treatment.

KEGG analysis also highlighted the hypermethylation of the oxidative phosphorylation pathway. The process of oxidative phosphorylation takes place in the inner mitochondrial

membrane and comprises four electron transfer chain complexes (I-IV), ATP synthase (complex V) and the electron carriers ubiquinone and cytochrome c (Srinivasan and Avadhani, 2012). The oxidative phosphorylation pathway in this study included the hypermethylated DMRs, COX7A1 and NDUFB2. NDUFB2 encodes the NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 2 enzyme whereas COX7A1 encodes the Cytochrome c oxidase polypeptide 7A1 enzyme. Both these enzymes play a role in the electron transport process within the mitochondrial respiratory chain with NDUFB2 and COX7A1 being involved in complex 1 and complex IV, respectively (Wirth *et al.*, 2016; Guerrero-Castillo *et al.*, 2017). Studies have shown a decrease in the expression of oxidative phosphorylation genes in T2DM (Dahlman *et al.*, 2006; Ling *et al.*, 2007; Rönn *et al.*, 2008). Studies looking at tissue-specific DNA methylation has shown that COX7A1 is only expressed in skeletal and heart muscle, with COX7A1 being down-regulated in diabetic muscle (Mootha *et al.*, 2003; Chalaya *et al.*, 2006). Rönn *et al* (2008) observed an increase in DNA methylation of COX7A1 with a decrease in the gene expression and increased insulin resistance in the skeletal muscle of elderly individuals (Rönn *et al.*, 2008). Similar findings were observed for NDUFB6, a gene also involved in complex 1 of the respiratory chain by Ling *et al* (2007). NDUFB6 showed increased methylation and decreased gene expression in muscle from patients with T2DM (Ling *et al.*, 2007). Additionally, metformin is known to inhibit mitochondrial oxidative phosphorylation by decreasing the mitochondrial respiratory chain (Vial *et al*., 2019). In the current study, the results showed hypermethylation in these two genes as well as the gene ontology cellular components, mitochondria respiratory chain, associated with them. Collectively, these have shown increased methylation and a possible decreased expression of the oxidative phosphorylation pathway in known diabetic subjects on metformin treatment. The hypermethylation in this study could therefore be a result of the metformin treatment.

The KEGG pathway analysis for the hypomethylated DMRs observed in known diabetics on metformin treatment when comparing to the NGT subjects was associated with RNA transport. The DMRS identified were EIF2B3 and EIF5 which were also associated with the biological process of regulation of translational initiation. RNA transport and translation play an important role in gene expression with translation being a key component in the regulation of protein synthesis (Roux and Topisirovic, 2018). In addition, regulation of translation allows cellular adaptation to occur during stress or physiological conditions such as hypoxia, endoplasmic reticulum stress, apoptosis and heat shock where an immediate change in protein levels are required (Sharma *et al.*, 2016). The initiation of eukaryotic translation is a multistep process involving the assembly of a pre-initiation complex consisting of a small (40S) ribosomal subunit, methionyl initiator tRNA and numerous eukaryotic initiation factors (EIFs)

(Llácer *et al.*, 2018). Eukaryotic translation initiation factors EIF2B3 and EIF5 play a role in translational control. EIF5 is a GTPase-activating protein specific for EIF2 and ensures that the correct AUG (start codon) selection occurs during the translation initiation. EIF2B3 on the other hand binds GTP and delivers the methionyl initiator tRNA to the small ribosomal subunit for the assembly of the pre-initiation complex (Nanda *et al.*, 2013; Sharma *et al.*, 2016). In a study by Stenvers *et al* (2019) where gene expression was measured, individuals with T2DM showed decreased activity of the EIF2 signalling in translational initiation when compared to a healthy control group (Stenvers *et al.*, 2019). This is in contrast to the current study where hypomethylation of the eukaryotic initiation factors EIF5 and EIF2B3 could indicate a possible increase in their expression and activity. Overexpression of EIF5 has been associated with the induction of translation for several genes, thereby enhancing the survival of normal and cancer cells under stress conditions (Ali *et al.*, 2017b).

4.2.2 New diabetes versus NGT

The KEGG pathway analysis in screen-detected diabetics, when compared to the NGT subjects, was associated with hypermethylation of the Notch signalling pathway. Notch signalling plays a role in intercellular communication and is essential for tissue homeostasis and the regulation of metabolic processes in certain organs (Guruharsha *et al*., 2012; Bi and Kuang, 2015). This signalling pathway links the fate of one cell to that of a neighbouring cell and can promote or suppress cell proliferation, cell death, differentiation or the attainment of specific cell fates (Hori *et al*., 2013; Hasan *et al.*, 2020). The Notch signalling pathway involves four Notch receptors (Notch 1-4) and five ligands (Delta 1, 3, 4 and Jagged 1-2) and is initiated when the Notch receptor interacts with its ligand on an adjacent cell (Mirtschink and Chavakis, 2018). Since Notch signalling plays a major role in many processes across a wide range of tissue, the abnormal gain or loss of Notch signalling has been linked to several disorders including genetic disorders and cancers (Bi and Kuang, 2015). A study by Hasan *et al* (2020) showed that Notch signalling in endothelial cells regulated insulin transport to muscle cells. Continued Notch signalling in endothelial cells lowered insulin sensitivity and increased blood glucose levels whereas inhibition of Notch signalling increased insulin sensitivity and improved glucose tolerance and uptake in muscle cells in a high-fat induced insulin resistance model (Hasan *et al.*, 2020). Similarly, it was also shown that Notch signalling regulates glucose metabolism in liver and adipose tissue and over-activation in these tissues impaired insulin sensitivity (Bi and Kuang, 2015). These studies implied that Notch signalling could be used as

a therapeutic target for diabetes due to the control of insulin sensitivity and glucose homeostasis. Notch signalling also affects pancreatic tissue. The function and mass of pancreatic beta-cells adapt to the insulin requirements of an organism to maintain glucose homeostasis. The development of diabetes is associated with a decrease in beta-cell mass and function. Bartolome *et al* (2019) found that the inhibition of notch signalling enhanced beta-cell proliferation and maturity, and improved glucose tolerance whereas activation of the pathway was associated with insulin resistance in the beta-cells of mice and human islets (Bartolome *et al.*, 2019). Similarly, Eom *et al* (2020) also showed that notch signalling played a role in maintaining beta-cell mass by regulating beta-cell function and proliferation in the mice model and thereby play a role in diabetes (Eom *et al.*, 2020). A study by Zheng and co-workers showed that hyperglycaemia caused Notch 1 signalling to be over-activated in diabetic skin and mediate the inhibitory effect of diabetes on wound healing. By blocking the Notch 1 signalling using genetic or chemical modes, wound healing was improved in the diabetic subjects showing that notch signalling could be a potential therapeutic target for the treatment of diabetic foot ulcerations (Zheng *et al.*, 2019). In the current study, Notch signalling was hypermethylated which could possibly indicate that the expression of this pathway was suppressed. Since Notch signalling plays role in many processes across a wide range of tissue, it is unclear whether the suppressed signalling pathway is due to an attempt to regulate insulin and maintain glucose homeostasis in the hyperglycaemic subjects. Furthermore, it is also known that individuals are often diagnosed many years after the onset of diabetes. Therefore, in terms of the current study, this statement would need to be further investigated.

When looking at the GO analyses of screen-detected diabetic subjects compared to the NGT subjects, the hypermethylated biological processes included the regulation of vasoconstriction and blood circulation. The hypermethylated DMRs associated with these biological processes were DBH, SMTNL1, PTAFR, CACGN8 and ATP1A3. In addition, ATP1A3 and CACNG8 were identified as hypermethylated DMRs in the KEGG pathway of cardiac muscle contraction. Hypermethylation of these DMRs and pathway indicate that these processes may be repressed in diabetic individuals when comparing them to the NGT subjects.

The DMR DBH encodes an enzyme, dopamine beta-hydroxylase, which catalyses the conversion of dopamine to norepinephrine and plays a role in blood pressure regulation through its association with the sympathetic nervous system (Abe *et al.*, 2005; Bozek *et al.*, 2017). Studies have shown an association between the inhibition of DBH with hypotension

due to the imbalance between dopamine and norepinephrine caused by the DBH inhibition (Abe *et al.*, 2005; Barrie *et al.*, 2014). Furthermore, a study by Arnold *et al* (2017) observed that a DBH deficiency in mice resulted in hyperinsulinaemia, lower plasma glucose levels and insulin resistance. They went on to further show impairment in cardiovascular autonomic regulation, hyperinsulinaemia and insulin resistance in a genetically deficient human patient (Arnold *et al.*, 2017). The DMR SMTNL1 encodes for proteins involved in smooth muscle contraction and has been linked to the adaptation of vascular contractility in conditions such as hypertension, pregnancy and exercise (Turner and Macdonald, 2014). PTAFR, on the other hand, encodes for the platelet-activating factor receptor involved in smooth muscle contraction, inflammation, immune response and hypotensive activity (Cao *et al.*, 2018). Calcium-voltage gated channel subunit gamma 8 encoded by CACGN8, regulates the gating properties of AMPA-selective glutamate receptors which mediates the synaptic transmission in the central nervous system (Deng *et al.*, 2006; Watson *et al*., 2017). ATP1A3 encodes a transmembrane protein, sodium-potassium (Na^+/K^+) ATPase, responsible for the transport of sodium and potassium across the membrane and plays a role in the electrical excitability of nerves and muscles (Ju *et al.*, 2016). Together these DMRs play a role in the cardiac muscle contraction pathway which was also identified as being hypermethylated in the KEGG pathway analysis. Disruptions of this pathway have been associated with diabetic cardiomyopathy as a result of cellular damage caused by hyperglycaemia (Ward, 2014). The process of cardiac muscle contraction is regulated by the control of calcium (Ca^{2+}) into and out of the myocytic cell and sarcoplasmic reticulum. Decrease in mitochondrial function in diabetes with resultant defects in the calcium transport and uptake of calcium by the mitochondria together with reduced sarcoplasmic/endoplasmic reticulum calcium ATPase (SERCA) pump activity results in calcium sequestration during cardiomyocytes diastolic relaxation. This leads to slowed cytosolic Ca^{2+} release and defective excitation-contraction coupling which contributes to the myocardial dysfunction in diabetic cardiomyopathy (Lebeche *et al*., 2008; Singh *et al*., 2018). In addition, it has been shown that cardiac dysfunction in diabetes is related to disruption of the Ca^{2+} homeostasis in the diabetic heart (Singh *et al.*, 2018). Cardiac ion channels are essential in the contraction of heart muscles and it has been shown that down-regulation of CACNG8 which encodes a subunit of calcium channels was associated with left ventricular dysfunction in cardiomyopathy patients (Ortega *et al.*, 2015). In a study by Dewey *et al* (2013) examining diabetic cardiomyopathy in mice, ATP1A3 was found to be suppressed. A decrease in ATP1A3, a subunit of the Na^+/K^+-ATPase membrane protein, resulted in a decrease in the capacity to maintain the sodium and potassium balance in the cell needed for maintaining the cellular electrochemical gradient. This in turn affected the Ca^{2+} regulation and cardiac muscle contraction (Dewey *et al.*, 2013). These observations support the findings of the current study

where the hypermethylation of the cardiac muscle contraction and related DMRs suggest that they may be suppressed in diabetic individuals when comparing them to the NGT subjects.

Pathway analysis in the screen-detected subjects, when compared to the NGT subjects, revealed hypomethylation of purine metabolism. The main function of purine metabolism is to maintain the level of nucleotides in tissues needed for biochemical processes, energy metabolism and the regulation of metabolic pathways (Papandreou *et al.*, 2019). The end product of purine metabolism is uric acid and its overproduction has been associated with diseases such as gout, renal dysfunction, hypertension, hyperlipidaemia, diabetes and obesity (Maiuolo *et al.*, 2016; Xiong *et al*., 2019). In terms of T2DM, hyperuricaemia has been associated with insulin resistance, impaired glucose intolerance and early onset of diabetic nephropathy (Ekpenyong and Akpan, 2014). Furthermore, hyperuricaemia has been shown to influence inflammation by its positive association with inflammatory cytokines such as interleukin-6, interleukin- 1β and tumour necrosis factor alpha as well as C-reactive protein (Kirilmaz *et al.*, 2010; Xiong *et al*., 2019). Excessive uric acid also leads to an increase in ROS production which leads to inflammation and cellular damage. In the current study hypomethylation of the purine metabolism pathway suggests that the end product of its metabolism would be overexpressed. Uric acid levels would be higher and therefore enhance the pro-inflammatory effect in the study subjects as well as put them at risk of developing gout and renal dysfunction. This enhanced pro-inflammatory effect is in contrast to previously mentioned findings in the current study where known diabetic subjects on metformin treatment possibly showed decreased levels of the pro-inflammatory cytokines from the tumour necrosis factor and interleukin family. This contrast could be explained by the role metformin plays in reducing effect of the pro-inflammatory cytokines which is not the case for the untreated newly diagnosed subjects.

Additionally, the hypomethylated DMR, ADK, identified in diabetic individuals (known and newly diagnosed) is associated with purine metabolism. As mentioned before in section 4.1.2, ADK levels tend to increase when more adenosine clearance is needed in the time of stress, tissue injury and inflammation.

4.2.3 Prediabetes versus NGT

The hypermethylated DMRs observed in prediabetic subjects when compared to the NGT group did not appear to have a relevant association with diabetes as they were associated with biological processes such as spermatogenesis and male sexual reproduction. On the other hand, the hypomethylated biological processes showed defense to Gram-negative bacterium and included the DMR LBP ((Lipopolysaccharide Binding Protein). The gene LBP encodes a protein involved in the innate immune response to gram-negative bacteria which promotes the release of cytokines once it binds to lipopolysaccharides (LPS) on the outer cell wall (Sakura *et al.*, 2017). The binding of LPS to LBP initiates the immune response by presenting LPS to CD14 which then associates with Toll-like receptor 4 (TLR4) on macrophages and initiates the inflammatory response, mostly by cytokines such as interleukin-1 (IL1), interleukin-6 (IL6) and tumour necrosis factor alpha (TNFα) (Tilves *et al.*, 2016). LBP is considered a biomarker for LPS and endotoxaemia (increased LPS in the bloodstream) as measuring LBP concentrations is an indirect way of assessing LPS (Gomes *et al*., 2017). Tilves *et al* (2016) showed that LBP levels were higher in African ancestry males with obesity, diabetes and insulin resistance. Furthermore higher LBP levels in individuals with normal serum glucose were associated with an increase in insulin resistance and increased fasting glucose at the six years follow up assessment suggesting that LBP could be a marker for prediabetes (Tilves *et al.*, 2016). The hypomethylation of LBP in the current study suggests that LBP levels may be increased in the prediabetic subjects and these findings are comparable to those of Tilves *et al* (2016). The increased LBP would indirectly indicate an increase in LPS and this may be a marker of future cardiovascular disease as LPS is a source of vascular inflammation in atherosclerosis (Sakura *et al.*, 2017). The release of pro-inflammatory cytokines by the LBP-LPS binding could also impair beta-cell function as previously indicated by the effects of the JNK pathway. The JNK pathway can be activated by the increase of inflammatory cytokines like tumour necrosis factor alpha (TNF-α) which is suggested to be involved in the development of insulin resistance and beta-cell dysfunction.

No enriched KEGG pathways were observed for both the hyper- and hypomethylated DMRs when comparing the prediabetic subjects to those with normal glucose tolerance.

4.2.4 Known diabetes versus new diabetes

Functional pathway analysis observed in this study when comparing known diabetics on metformin treatment with the newly diagnosed diabetic subjects was consistent with the basic pathological abnormalities in Diabetic Peripheral Neuropathy (DPN). Diabetic Peripheral Neuropathy is a chronic progressive disorder characterised by complications of axonal degeneration and demyelination, lack of sensation, numbness, paraesthesia (pins and needles), pain and allodynia in diabetic individuals (Guo *et al.*, 2019). Cell death of nerves in DPN results from multifactorial metabolic imbalances associated with diabetes. Hyperglycaemia, dyslipidaemia and insulin resistance may affect various pathways such as the protein kinase C (PKC), polyol, AGE, poly(ADP-ribose) polymerase (PARP), and hexosamine pathways which in turn may result in mitochondrial dysfunction (Miranda-Massari *et al.*, 2011). The resulting mitochondrial dysfunction through a series of cascade effects involving AMP-activated protein kinase (AMPK), sirtuin (SIRT), and peroxisome proliferator-activated receptor-γ coactivator α (PGCα) suppresses mitochondrial oxidative phosphorylation, resulting in neuronal and axonal degeneration through increased oxidative injury (Fernyhough, 2015; Fujimaki and Kuwabara, 2017). Treatment with metformin was shown to decrease the incidence of DPN as was observed by the Bypass Angioplasty Revascularization Investigation 2 Diabetes trial (Pop-Busui *et al.*, 2013). Although metformin cannot reverse the nerve damage caused by diabetes, it could assist in managing blood glucose levels and improving the symptoms for patients.

The DMR, KCQN3 was associated with the hypermethylated GO cellular components of the axon initial segment. The KCNQ family of voltage-activated potassium (K^+) channels encoded by KCNQ genes play a role in the hyperpolarisation and stabilisation of cell membrane potential. Due to this action, the KCNQ channels have a role in regulating the excitability of neurons, smooth muscle cells and cardiomyocytes (Fosmo and Skraastad, 2017). KCNQ3 has shown an association with the development of encephalopathy and epilepsy in humans as well as an association with diabetic neuropathic pain (Fosmo and Skraastad, 2017; Yu *et al.*, 2018). Diabetic neuropathic pain is one of the main symptoms of diabetic neuropathy and has shown to be associated with suppression of KCNQ3 channel expression. This down-regulation of expression causes enhanced neuronal excitability which results in the allodynia and neuropathic pain experienced by diabetics (Yu *et al.*, 2018). It is suggested that the activation of the KCNQ3 channels could be used as a therapeutic target for the treatment of neuropathic pain (Abd-Elsayed *et al.*, 2019). In the current study, KCNQ3 was hypermethylated which suggests that the expression of the KCNQ3 channels may be

suppressed. These findings are in line with those of the previous studies mentioned by Yu et (2018) and Abd-Elsayed *et al* (2019). Furthermore, it is unclear whether the metformin treatment has any effect on the improvement of the KCNQ3 channel activity in these study subjects. Although metformin exerts a neuroprotective effect, it could enhance peripheral neuropathy by inducing vitamin B12 deficiency (Chung *et al.*, 2015; Ahmed *et al*., 2017). Previous studies have shown that prolonged treatment with metformin led to more severe neuropathy by inducing vitamin B12 deficiency in diabetic patients (Wile and Toth, 2010; Russo *et al.*, 2016). To date, there are conflicting findings on the link between vitamin B12 deficiency and metformin treatment based on the different tools used to assess neuropathy (Ahmed *et al*., 2017). Cumulatively, it has been reported that there was no difference found in the risk of peripheral neuropathy between metformin-treated individuals with vitamin B12 deficiency and those with normal vitamin levels (Ahmed *et al*., 2016; Russo *et al.*, 2016).

No enriched KEGG pathways were observed for both the hyper- and hypomethylated DMRs when comparing the known diabetics on metformin treatment to the newly diagnosed diabetic subjects.

4.2.5 Known versus prediabetes

The KEGG pathway analysis revealed hypermethylation of the cAMP signalling pathway when comparing known diabetics on metformin treatment to the prediabetic subjects. One of the DMRs associated with this pathway was GPR119 (G Protein-Coupled Receptor 119).

Cyclic adenosine monophosphate (cAMP) is an intracellular second messenger playing a role in the regulation of insulin from beta-cells and glucagon from alpha-cells of the pancreas. It is synthesised from adenosine triphosphate (ATP) by adenylate cyclase (AC) which is activated by the binding of hormones and neurotransmitters to G-protein coupled receptors (GPCRs) (Yang and Yang, 2016). The physiological effects of cAMP are mediated by the activation of cAMP-dependent protein kinase (PKA), which in turn phosphorylates and regulates the functions of downstream protein targets including ion channels, enzymes, and transcription factors. The cAMP/PKA signalling pathway regulates glucose homeostasis including insulin and glucagon levels, glucose uptake, glycogen synthesis and gluconeogenesis breakdown (Deb *et al*., 2017). GPR119 is a class A member of the

rhodopsin family of G-protein coupled receptors whose expression is mainly found in cells of the pancreatic islets and gastrointestinal tract (Moran *et al*., 2016). Activation of GPR119 is mainly by fatty acid ethanolamides which results in the stimulation of the adenylate cyclase pathway leading to cyclic adenosine monophosphate (cAMP) production in cells expressing GPR119 (Moran *et al*., 2016). Upon activation of GPR119, endogenous lysophospholipids stimulate insulin secretion by acting directly on beta-cells and stimulating the secretion of glucagon-like peptide 1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP). This leads to the regulation of glucose homeostasis and improved glucose control. In addition, the stimulation of intestinal secretion of incretin is also part of the glucoregulatory effects of GPR119 (Zhou *et al.*, 2019). These factors have led to the suggestion of GPR119 as a therapeutic target for T2DM (N. X. Li *et al.*, 2018b).

Glucagon levels, which are elevated in diabetic individuals, is responsible for activating adenylate cyclase which results in cAMP synthesis and activation of PKA activity (Tengholm and Gylfe, 2017). Metformin has been shown to inhibit the stimulation of cAMP levels by glucagon and disrupting the PKA activity. This is through the action of increasing AMP levels and the inhibition of mitochondrial complex 1 (Pernicova and Korbonits, 2014). Inhibition of complex 1 reduces NADH oxidation and the proton-driven synthesis of ATP causing the equilibrium between ATP, ADP and AMP to shift towards AMP synthesis by adenylate kinase. Increased AMP levels inhibit adenylate cyclase which catalyses the action of converting ATP to cAMP. In this way, metformin reduces cAMP levels and glucagon signalling (Pernicova and Korbonits, 2014). In the current study, the hypermethylation of the cAMP signalling pathway and related GPR119 in known diabetics on metformin treatment suggests possible suppression of the pathway. This is expected for the known diabetics as metformin has been shown to decrease cAMP synthesis and PKA activity.

The KEGG pathway analysis also showed hypomethylation of the PI3K-AKT signalling pathway when comparing known diabetics on metformin to prediabetic subjects. In addition, the hypomethylated DMR, Insulin Receptor Substrate 1 (IRS1) was associated with this pathway. Phosphatidylinositol 3-kinase (PI3K)-AKT signalling is an intracellular signal transduction pathway that plays a role in several processes such as glucose homeostasis, lipid metabolism, protein synthesis, angiogenesis, cell growth and survival, in response to extracellular signals (Huang *et al.*, 2018). Phosphatidylinositol 3-kinases (PI3Ks) are part of a lipid kinase family that phosphorylates a cell membrane component, phosphatidylinositol. The activation of PI3Ks is mediated by several receptors such as G-protein coupled receptors

(GPCR), receptor tyrosine kinases (RTKs), cytokines, hormones and growth factors amongst others. Once activated, PI3K recruits signalling proteins including AKT (protein kinase B) (Pompura and Dominguez-Villar, 2018). AKT regulates glucose and lipid metabolism and when activated, it is expressed in insulin-responsive tissues and promotes the translation of glucose transporter 4 (GLUT4) (Huang *et al.*, 2018). The IRS1 is a cytoplasmic substrate for insulin and has found to be dysregulated in obesity and T2DM (Besse-Patin *et al.*, 2019). Insulin is needed for many metabolic processes and when it binds to the insulin receptor, it leads to recruiting and phosphorylation of insulin receptor substrate proteins such as IRS1. Phosphorylated IRS1 activates PI3K producing phosphatidylinositol-3,4,5- triphosphate (PIP3) from phosphatidylinositol-4,5-bisphosphate which in turn activates the protein kinase B, AKT (Copps and White, 2012; Keshavarzi and Golsheh, 2019). Since activated AKT regulates glucose metabolism, the role of IRS1 in the insulin signalling pathway of its target tissue has a role in the development of T2DM as any disruption in this signalling process may result in insulin resistance.

Metformin has shown to inactivate IRS1 and subsequently the PI3K-AKT signalling pathway thereby inhibiting hyperinsulinaemia-associated neoplastic activity in diabetic cancer patients. By inhibiting protein translation via the PI3K-AKT-mTOR pathway, cancer cell proliferation is restricted (Saini and Yang, 2018). Activation of AMPK by metformin has also been associated with a decrease in insulin sensitivity, so one would expect metformin to have an effect on the PI3K-AKT pathway. However, the PI3K-AKT pathway is also involved in many cellular processes which include cell proliferation and survival which means that its activation is context dependant (Schultze *et al.*, 2012). Therefore the effect of metformin on the PI3K-AKT pathway could vary according to the context of its activation and whether the activation is through insulin, growth factors, cytokines or environmental stresses. In the current study known diabetic subjects on metformin displayed hypomethylation of the PI3K-AKT pathway and IRS1 when compared to prediabetic subjects suggesting that their expression might be increased. Since disturbances in the PI3K-AKT signalling results in insulin resistance, it is suggested that metformin treatment in these known diabetics may aid in the management of insulin resistance by enhancing the PI3K-AKT signalling pathway. This would need further investigation as it has also been shown that inappropriate activation of this pathway is associated with tumours and that the inhibition of the PI3K-AKT signalling pathway could be utilised as a potential treatment option for cancer and other diseases (Schultze *et al.*, 2012).

4.2.6 Prediabetes versus new diabetes

The hypermethylated DMRs observed in prediabetic subjects when compared to the NGT group did not appear to have a relevant association with diabetes as they were associated with biological processes such as sperm motility. Furthermore, the KEGG pathway analysis revealed no enriched hypermethylated pathways whereas the hypomethylated KEGG pathway analysis was associated with oocyte meiosis and maturation. On the other hand, the hypomethylated biological processes showed positive regulation of wound healing and blood coagulation associated with the DMRs DMTN and SERPINF2. SERPINF2 also known as alpha 2-antiplasmin encodes a serine protease inhibitor that inhibits plasmin and plays a role in regulating fibrinolysis (Viganò *et al.*, 2018). Diabetes has been linked to haemostasis disorders as hyperglycaemia has been shown to enhance thrombosis and decreased fibrinolysis (Alzahrani and Ajjan, 2010; Kearney *et al.*, 2017). Elevated levels of antiplasmin in diabetes have been associated with an increased risk for myocardial infarction (Meltzer *et al.*, 2010; Agren *et al.*, 2014). As antiplasmin regulates fibrinolysis, increased incorporation of antiplasmin into fibrin results in a more stable fibrin network. Furthermore, the binding of antiplasmin to fibrin will reduce fibrin degradation through its binding and inhibition of plasmin and this increased action could therefore contribute to diabetic complications (Agren *et al.*, 2014). In the current study, prediabetic subjects with hyperglycaemia have shown hypomethylation of the antiplasmin DMR which suggest possible elevation of its action. This could be used as an indicator of progression to full-blown diabetes and a risk for cardiovascular disease. Furthermore, targeting antiplasmin could have benefits as a potential therapeutic target in the treatment of diabetes.

Dematin actin binding protein (DMTN) encodes a cytoskeletal protein which plays a structural role in maintaining erythrocytes shape and membrane stability as well as regulating cell movement (Lu *et al.*, 2016). Deficient dematin has shown to result in defects in erythrocyte shape, membrane instability and haemolysis (Khan *et al.*, 2008; Lu *et al.*, 2016). Proteins in the erythrocyte membrane are organised in structural complexes anchoring the cytoskeleton to the lipid bilayer. In addition, glucose transporter 1 (GLUT1), which aids in regulating glucose transport in erythroid, endothelial, epithelial and neuronal cells, interacts with dematin at the actin-spectrin junctional complex in erythrocyte membranes (Khan *et al.*, 2008; Guizouarn and Allegrini, 2020). It is suggested that hyperglycaemia causes modifications in the cytoskeleton and protein complexes around GLUT1 in red blood cells which could interfere with glucose uptake through GLUT1 in diabetic individuals (Guizouarn and Allegrini, 2020). In addition to its role in haematopoietic cells, DMNT also plays a role in regulating the actin cytoskeleton in

non-erythroid cells (Ye *et al.*, 2018). In a study by Mohseni *et al* (2008), it was also shown that DMTN regulates cell shape, motility and wound healing by controlling RhoA activation which participates in cytoskeletal remodelling. The study revealed that the reepithelialisation and granulation in wound healing were delayed by the absence of dematin in fibroblasts (Mohseni and Chishti, 2008). In the current study, DMTN was found to be hypomethylated which could suggest an increase in expression. The increased expression of dematin could be linked to an increase in stability and structure of the red blood cell membranes in prediabetic subjects when compared to the newly diagnosed diabetic subjects. In addition, the increased expression of dematin and its link with the GLUT1 transporter could be an indication of the glucose transport mechanism in prediabetic individuals. The expression of dematin could also be monitored as an indicator of red blood cell stability with worsening glucose tolerance. Since diabetes is associated with delayed wound healing (Patel *et al.*, 2019), increased expression of dematin could aid in the epithelialisation and granulation processes in the wound healing of the prediabetic subjects.

4.3 Long non-coding RNAs (lncRNAs)

The impact of long non-coding RNAs (lncRNAs) on biological and pathological processes have been associated with several conditions including metabolic diseases. For this reason, lncRNA-associated DNA methylation was investigated in the current study. A total of 943947 statistically significant lncRNA-associated MeDIP-enriched regions (peaks) were identified for the 48 samples. Of the 641 lncRNA-associated DMRs identified, 63% were hypermethylated and 37% hypomethylated of which the highest number of hypermethylated lncRNA-associated DMRs were observed by the known diabetes group (77%) when compared to the NGT group. In addition, most of the hypermethylated and hypomethylated DMRs were found on chromosome 1.

4.3.1 LncRNA-associated DMRs identified in hyperglycaemic subjects

The common hypomethylated DMRs identified in the diabetic and non-diabetic hyperglycaemic subjects (known, screen-detected and prediabetes) when compared to the NGT subjects included C4A. Complement component 4A (C4A) is a polypeptide derived from

the cleavage of C4 upon activation of the classical and lectin pathways of the complement system involved in the innate immune system (Barnum, 2015; Wang *et al*., 2017a). Although the complement system is an essential part of the immune system, prolonged activation has shown to play a role in inflammation and cardiometabolic disease (Copenhaver *et al.*, 2020). In the study by Copenhaver *et al* (2020), complement affected endothelial cells, especially in the earlier stages of endothelial dysfunction where higher levels of C4 were associated with impaired endothelial function due to poor reactive hyperaemic responses. Furthermore, inflammation was shown to increase as the markers of inflammation which included white blood cell and neutrophil counts as well as CRP and interleukin-6 (IL6) levels corresponded to the increased levels of C4 (Copenhaver *et al.*, 2020). In addition, it has been shown that C4A increased stress fibre formation and the permeability of endothelial cells through its activation of the G-coupled protein receptor protease-activated receptors 1 and 4 (PAR1 and PAR4) which have a pro-inflammatory effect (Wang *et al*., 2017a). Although components of the complement system are involved in both diabetes mellitus type 1 and type 2, more studies have linked C4 to type 1 diabetes (Szilagyi *et al.*, 2006; Wang *et al.*, 2010; Mason *et al.*, 2014). In the case of T2DM, increased levels of C4 was found when comparing T2DM patients to individuals with normal glucose tolerance. It was suggested that complement-mediated inflammation contributed to the acceleration of diabetic microangiopathy after seeing that the earlier phase of the complement pathway was excessively activated in T2DM patients (Shim *et al.*, 2020). Furthermore, in a study by Nilsson *et al* (2014) the relationship between C4 and metabolic cardiovascular risk factors were assessed and found to be positively correlated. It was shown that raised C4 levels were associated with cardiometabolic risk factors such as obesity, blood pressure, blood lipids and metabolic syndrome (MetS) (B. Nilsson *et al.*, 2014a). As these risk factors including MetS are associated with T2DM (Liu *et al.*, 2016), one could suggest that C4 as part of the complement system plays a role in the development of the metabolic profile of individuals with T2DM. Liu *et al* (2016) found C4 levels to be increased in study participants with MetS and that the raised levels of C4 could be used as an indicator of increased risk for developing MetS (Liu *et al.*, 2016). In the current study, the lncRNA-associated DMR C4A was hypomethylated in all subjects with hyperglycaemia when compared to the control subjects which may suggest increased expression of this complement component. This possible increase in C4A expression correlates with the findings of the other studies mentioned where increased C4 was associated with T2DM and cardiometabolic risk factors.

4.3.2 LncRNA-associated DMRs identified in subjects with diabetes (known and newly diagnosed)

The hypermethylated DMRs identified in diabetic individuals (known and screen-detected subjects) included MLXIP and RPTOR. MondoA-interacting protein (MLXIP) plays a role in the transcriptional response to intracellular glucose concentration (Richards *et al.*, 2017). It maintains glucose homeostasis under normal nutrient conditions, however, chronic nutrient overload causes dysregulation of MondoA and could result in insulin resistance and T2D (Song *et al.*, 2019). MondoA is highly expressed in skeletal muscle where it regulates the glycolytic pathway (Ran *et al.*, 2020). It has been shown that once activated, MondoA inhibits the glucose uptake in skeletal muscle whereas repression of MondoA in skeletal myotubes has shown to increase glucose uptake (Ahn *et al.*, 2019). Furthermore, the loss of MondoA in mice muscles have shown to improve high-fat-diet-induced glucose tolerance and insulin resistance by decreasing muscle lipid accumulation and increasing muscle insulin signalling and glucose uptake (Ahn *et al.*, 2016). Richards *et al* (2018) also observed that MondoA is a glucose-responsive transcription factor in human pancreatic beta-cells. During high glucose conditions, MondoA up-regulated the expression of the genes, thioredoxin interacting protein (TXNIP) and arrestin domain-containing protein 4 (ARRDC4), which play a role in the inhibition of glucose uptake (Richards *et al.*, 2018). Studies have shown that metformin reduces the binding of the MondoA transcription factor complex to TXNIP and thereby reducing the expression of TXNIP. Since TNXIP acts by inhibiting cellular glucose uptake, metformin thereby aids in lowering blood glucose levels (Chai *et al.*, 2012; Li *et al.*, 2015). With its glucose-sensing transcriptional activity playing a role in glucose metabolism and insulin sensitivity, MondoA may therefore provide a potential target for anti-diabetic therapy through its inhibition in various tissues. In the current study MLXIP was hypermethylated which may imply repression of the MondoA activity in both the known diabetics on metformin treatment and the newly diagnosed diabetic subjects. As noted in the previous studies, inhibition of MondoA improved glucose uptake by tissue. Treatment by metformin could therefore explain the possible decrease in expression of MLXIP in the known diabetic subjects. Since the newly diagnosed subjects were not on metformin treatment, it is unclear as to why the MLXIP expression may be repressed. A possible explanation could be that MLXIP expression varies amongst the different metabolic tissues as explained by Song *et al* (2019) who observed opposing effects of MondoA in the muscle and pancreas in comparison to the liver and white adipose tissue (Song *et al.*, 2019).

Another hypermethylated lncRNA-associated DMR observed was the Regulatory-Associated Protein of mTOR (RPTOR), also known as RAPTOR. RAPTOR forms part of the mTOR complex 1 (mTORC1), which is one of two large protein complexes of the nutrient-sensing serine-threonine protein kinase (Yin *et al.*, 2020). mTORC1 is an important driver of anabolic metabolism in response to growth factors and nutrients and plays a role in the promotion of protein synthesis, lipid biogenesis and metabolism as well as the reduction of autophagy (Howell *et al.*, 2017). Although mTORC1 is essential for beta-cell development, growth and function, it's over activation by nutrient overload in the form of glucose has been shown to lead to beta-cell exhaustion, functional loss and eventually cell death in T2DM (Ni *et al.*, 2017; Yuan *et al.*, 2017). Yuan *et al* (2017) showed that mTORC1 activity was increased in human pancreatic islets from T2DM patients and that this elevated mTORC1 activation led to impaired beta-cell function and survival in response to metabolic stress (Yuan *et al.*, 2017). In a study by Ni *et al* (2017), it was found that the deletion of RAPTOR and inactivation of the mTORC1 in murine beta cells affected their maturation and also resulted in hypoinsulinaemia and decreased glucose tolerance. Furthermore, the deficiency of RAPTOR showed reduced expression of the DNA-methyltransferase 3a resulting in hypomethylation and transcriptional activation of beta cell-specific disallowed genes (*Hk1*, *Dlk1*, *Pdgfra*, *Oat* and *Mylk*) involved in glucose metabolism, insulin secretion and beta-cell maturation (Ni *et al.*, 2017). Metformin has shown to inhibit mTORC1 in the liver by its activation of AMPK (Soliman *et al*., 2016). An effect of metformin is to decrease cellular energy levels by inhibiting the mitochondrial complex 1 and thereby decreasing cellular respiration and ATP levels. This stimulates the release of AMPK which in turn promotes ATP-producing catabolic processes as well as inhibiting anabolic processes which consume ATP. Processes stimulated by mTORC1 are reliant on energy provided by ATP and therefore are inhibited by the metformin-induced activation of AMPK during cellular energy depletion (Howell *et al.*, 2017; Ardestani *et al.*, 2018). In the current study, the DMR RAPTOR was hypermethylated which suggests that its expression and the activity of mTORC1 was suppressed in both known diabetics on metformin treatment and newly diagnosed diabetic subjects. In the metformin-treated subjects, the possible decrease in RAPTOR and mTORC1 activity can be explained by the inhibitory effect of AMPK. It is however unclear as to why there is a possible decrease in RAPTOR and subsequently, mTORC1 activity in the newly diagnosed diabetic subjects as no known inhibitors such as pharmaceutical measures was in place. A possible explanation could be that mTORC1 activity affects metabolism and energy homeostasis in a tissue-specific manner. This was suggested by Blandino-Rosano *et al* (2017) as tissue-specific deletion of RAPTOR resulted in a loss of mTORC1 signalling in liver, muscle and adipose tissue with various outcomes (Blandino-Rosano *et al.*, 2017).

The hypomethylated lncRNA-associated DMR observed included C4A (previously discussed for all hyperglycaemic subjects) and Methyl-CpG Binding Domain Protein 3 (MBD3). The methyl-CpG binding domain (MBD) family of proteins, including MBD3, are considered to be protein readers of methylation involved in interpreting DNA methylation into functional outcomes (Song *et al.*, 2020). The members of the MBD family have no similarity in sequence except for MBD2 and MBD3 who are paralog genes and have a 70% amino acid sequence similarity (Leighton and Williams Jr, 2020). As opposed to the others, MBD3 is the only family member which does not bind to methylated DNA but instead can bind to non-methylated CpG-rich promoters and enhancers (Gunther *et al.*, 2013; Menafra and Stunnenberg, 2014). Furthermore, it has been suggested that there is potential coordination between MBD2 and MBD3 as the presence of MBD2 at genomic regions seems to be influenced by MBD3 at the same sites (Gunther *et al.*, 2013; Du *et al.*, 2015). A study by Brown *et al* (2008) showed that overexpression of MBD3 induced demethylation in specific targets of the genome and resulted in the activation of these genes. The study utilised HEK 293 cells and found that overexpression of MBD3 resulted in a loss of methylation in the promoters of the MMP24 and VEGF-C genes associated with certain cancers (Brown *et al.*, 2008). Cui *et al* (2015) showed that MBD3 played a role in DNA methylation homeostasis as together with MBD2 and DNA methyltransferase 1 (DNMT1), it provided favourable conditions for continuing DNA maintenance methylation. The demethylating property of MBD3 provided a protective mechanism in the S-G2 phases of the cell cycle and insufficient MBD3 resulted in a disruption of DNA methylation homeostasis (Cui and Irudayaraj, 2015). Although there is no direct association between MDB3 and T2DM in literature, it was observed that MBD2 expression levels were increased in patients with T2DM when compared to control subjects. This study indicated that dysregulation of the methylation process occurred as high glucose levels induced DNA methylation by the up-regulation of DNMTs and MBD2 (Karachanak-Yankova *et al.*, 2015). Since a close relationship has been shown between MBD2 and MBD3, one could speculate whether MBD3 expression levels are also increased in T2DM. This suggestion correlates with the findings of the current study, where MBD3 expression is suggested to be increased as the lncRNA-associated DMR was found to be hypomethylated in diabetic subjects. Furthermore additional studies would need to be conducted to determine which genes and their functional outcomes are affected by the suggested increased expression of MBD3 in the current study.

4.3.3 LncRNA-associated DMRs identified in known diabetes versus new diabetes

When comparing known diabetics on metformin treatment to newly diagnosed diabetic subjects, hypermethylation of the mitochondrial ATP synthase-coupling factor 6 (ATP5J) enzyme was observed. ATP5J is a protein connecting two components, F0 and F1, of ATP synthase which is a crucial enzyme promoting oxidative phosphorylation in the mitochondria. Therefore ATP5J expression is related to both ATP synthase synthesis and the synthesis of mitochondrial ATP and any change in their expression could mirror the functioning of the mitochondria and ATP synthase (Wang *et al*., 2017b; Bai *et al*., 2018). Previous studies have investigated the role of ATP5J in cardiovascular diseases. A study by Morena-Viedman *et al* (2016) identified dysregulation of the oxidative phosphorylation pathway and associated genes including ATP5J in an insulin resistant-atherosclerosis mouse model linking ATP5J to cardiovascular and metabolic diseases (Moreno-Viedma *et al.*, 2016). Furthermore, a study by Osanai *et al* (2012) showed increased plasma levels of ATP5J, also known as coupling factor 6, was associated with diabetes and hypertension due to tissue acidosis. Overexpression of the coupling factor 6 (ATP5J) causes a decrease in intracellular pH in tissue expressing the coupling factor 6 receptor and this metabolic acidosis induces hypertension and insulin resistance (Osanai *et al.*, 2012). Findings in the current study show hypermethylation of ATP5J in metformin-treated subjects possibly indicating suppression of this lncRNA. The suppression of the ATP5J expression could be a result of the metformin as previous studies have associated diabetes with increased levels of ATP5J. This, however, warrants further investigation.

Another hypermethylated lncRNA-associated DMR identified in known diabetics on metformin treatment when compared to newly diagnosed subjects was lysyl oxidase-like 2 (LOXL2). Lysyl oxidase (LOX) proteins consist of a family of copper-dependent enzymes which play a role in extracellular matrix (ECM) homeostasis and remodelling (Rodriguez and Martinez-Gonzalez, 2019). The ECM plays an important role in maintaining tissue integrity and housing mediators of cell signalling and growth and disruptions may lead to increased matrix synthesis and disease progression (Stangenberg *et al.*, 2018). One of the four LOX-like enzymes, LOXL2, has been associated with cardiovascular diseases which are characterised by fibrosis and decreased cardiac function (Yang *et al.*, 2016; Rodriguez and Martinez-Gonzalez, 2019; Erasmus *et al.*, 2020). Fibrosis can be caused by several factors which include increased oxidative stress, inflammation, hyperglycaemia, hypertension and these factors may mediate epigenetic modifications and affect gene expression (Erasmus *et al.*, 2020). Large amounts of ECM protein deposits, of which collagen is a major component,

is a risk factor for heart failure and T2DM has been associated with alterations in ECM patterns (Frangogiannis, 2019). Diabetes is a known risk factor for cardiovascular disease and it has been shown that cardiac damage induced by diabetes results in abnormally increased collagen deposits between the myocytes, resulting in thickening of the ventricle wall and stiffening of the heart muscles. Furthermore, the increased oxidative stress in cells and hyperglycaemia in diabetes lead to the production of advanced glycation end-products (AGEs), a complex and heterogeneous group of compounds associated with diabetes related complications (Chen *et al*., 2019b). Through its interaction with a number of pathways hyperglycaemia-induced AGEs as well as the activation of the receptors for AGEs have been linked to increased activation of inflammation and oxidative stress (Chen *et al*., 2019). The accumulation of AGEs in various organs such as the kidneys and the heart have been implicated in vascular and endothelial disorders (Perera and Handuwalage, 2015). The aggregation of AGEs on proteins involved in fibrotic processes, such as collagen, can affect the normal degradation of proteins (Erasmus *et al*., 2020). The combination of AGEs and collagen affects the elasticity and stiffness of heart tissue leading to myocardial fibrosis associated with LOXL2 gene expression (Erasmus *et al.*, 2020). Johnson *et al* (2020) showed that increased expression of LOXL2 correlated with increased tissue fibrosis and suggested that LOXL2 could be used as a predictive marker to detect the early onset of diabetic cardiomyopathy (Johnson *et al.*, 2020). Also, their study found that suppression of LOXL2 reduced the expression of genes such as COL1A associated with increased tissue fibrosis and enhanced collagen formation. This indicated that collagen formation was possibly controlled by LOXL2 (Johnson *et al.*, 2020). Similarly, Stangenberg *et al* (2018) found that increased levels of LOXL2 was associated with renal fibrosis and the development and progression of diabetic nephropathy in an animal model. By inhibiting LOXL2 expression, an improvement in glomerular structure and function was observed, once again indicating that LOXL2 could be used as an therapeutic target for renal fibrosis and more particularly diabetic kidney fibrosis (Stangenberg *et al.*, 2018).

In the current study, LOXL2 was hypermethylated which may suggest that expression of this lncRNA-associated DMR was suppressed. This could be as a result of the metformin treatment as studies listed above have shown that diabetes is associated with increased fibrosis and ECM deposition. Studies by Li *et al* (2016) showed that metformin down-regulated the expression of LOXL2 and reduced the collagen deposition in adipose tissue in mice, thereby reducing fibrosis (Li *et al.*, 2016b). It is therefore suggested that metformin may have decreased the expression of LOXL2 in the known diabetic subjects of the current study.

CHAPTER 5

CONCLUSION

5.1 Strengths of the study

The inclusion of the diabetic and pre-diabetic groups add strength to the study, especially within a South African context where limited studies on DNA methylation have been conducted. The classification of the study subjects was also well characterised with the use of the oral glucose tolerance test (OGTT) as well as the measurement of HbA1C for determining the glucose status.

In this study, peripheral blood was used to assess DNA methylation. Although it has been shown that DNA methylation is tissue-specific, studies using peripheral blood DNA has shown consistent methylation patterns with other organs (Farré *et al.*, 2015; Crujeiras *et al.*, 2017). Therefore the use of blood which is an accessible and convenient biological material in a clinical setting may be a strength in DNA methylation studies. In addition, suitable amounts of quality DNA can be extracted from peripheral blood and banked for future use or extensions of the current study (Li *et al.*, 2012). This highlights the use of blood-based DNA methylation as a potential non-invasive screening test for T2DM which can aid in the prevention, diagnosis and treatment of T2DM.

Another strength of the study was the genome-wide approach. The benefit of using genome-wide DNA methylation profiling instead of candidate gene methylation was that it provided a comprehensive view of the DNA methylation landscape within a South African population. It also provided the foundation for future studies whereby the identified DMRs and lncRNA-associated DMRs could be further explored. Furthermore, the combination of utilising DNA methylation and pathway analysis methods in this study provided insight into the potential mechanisms influencing glucose regulation and the pathogenesis of T2DM.

5.2 Limitations of the study

Limitations of the study include the use of only female participants from a South African community and therefore the generalizability of these findings to males is unknown. Although the small sample size can be considered a limitation, it allowed for comparison and limited error which may occur during statistical manipulation of small sample size by sex. Larger sample sizes may however have increased the statistical power of the study. The study participants were all from mixed genetic origin and therefore the possible effect of genetic heterogeneity on DNA methylation cannot be ignored. The mixed ancestry community in South Africa has genetic contributions from Europeans, South Asians, Indonesians and a population genetically close to the isiXhosa sub-Saharan Bantu (Patterson *et al.*, 2009). Since it has been shown that genetic ancestry plays a role in DNA methylation patterns (Chu and Yang, 2017; Galanter *et al.*, 2017), the identification of ancestry markers particularly within this heterogeneous South African population warrants further investigation.

As the study participants were all matched for age, gender, ethnicity and body mass index (BMI), it was assumed that the environmental factors were similar for the participants living within this particular South African community. Since DNA methylation is affected by environmental and lifestyle factors such as physical activity, smoking, diet and alcohol consumption (Martin and Fry, 2018; González-Becerra *et al.*, 2019; Wilson *et al.*, 2019), a limitation of the study was the assumption of participant honesty during the questionnaire data collection phase. The current study did however measure serum cotinine levels which matched the participants' responses on smoking habits. Further studies are however needed to determine whether these environmental and lifestyle factors had any effect on the DNA methylation patterns observed in this particular cohort.

The study also used a cross-sectional study design which may not effectively demonstrate causality between the changes in DNA methylation and metabolic traits and therefore warrants further investigation with the use of a longitudinal study design. Although longitudinal studies are more suited to follow the long-term effects on DNA methylation patterns, the high cost and time challenges need to be considered. Also, the reversible nature of epigenetic changes over time as well as the long-time course for T2DM development would need to be considered. Since the development of T2DM may occur years before clinical diagnosis (Porta *et al.*, 2014; Ali *et al.*, 2017a), the methylation patterns observed in the diabetic (treated and untreated)

participants should be interpreted with caution as it is unclear whether the findings are as a result of the duration and severity of diabetes.

While the current study has highlighted the strengths of using blood-based DNA methylation, it should be noted that blood is a heterogeneous tissue consisting of many different cell types (Jaffe and Irizarry, 2014). Since DNA methylation varies with cell type, the cellular heterogeneity of blood may affect DNA methylation patterns as each cell type possesses its own epigenetic signature (Houseman *et al.*, 2015; Husby, 2020). The use of peripheral blood could therefore be noted as a limitation of the current study.

The MeDIP sequencing technology used to measure genome-wide DNA methylation in the study has its limitations but these are outweighed by the advantages which included the lack of a bisulphite conversion step and the generation of large datasets which are easy to analyse and interpret. This resulted in MeDIP-seq being used as the primary method for DNA methylation over pyrosequencing, which while highly sensitive and reliable, only allows for the investigation of small genomic regions. Although pyrosequencing was used to validate the methylation status of the participant samples, a previous study comparing four sequencing-based DNA methylation methods including MeDIP-seq showed comparable methylation in all four methods (Harris *et al.*, 2010). It is however recommended that DNA methylation studies should use more than one sequencing-based DNA methylation profiling method to verify DNA methylation regions.

5.3 Conclusion of the study

Type 2 diabetes mellitus is a global health concern and it has become increasingly necessary to identify ways of recognising populations at risk of developing T2DM as well as improve the clinical diagnosis and treatment regimes for patients. Moreover, the number of people with diabetes in Africa is growing faster in comparison to other countries, with South Africa showing the highest age-adjusted prevalence of the disease and a large percentage of undiagnosed individuals (Asmelash and Asmelash, 2019; International Diabetes Federation, 2019; Mutyambizi *et al.*, 2019). Although studies have shown a relationship between the occurrence of T2DM and genetic and lifestyle factors, it has been suggested that epigenetic mechanisms such as DNA methylation contribute to the pathogenesis of T2DM through its association with

the transcriptional activity of genes. There are however limited studies focussing on the genome-wide DNA methylation profiling of diabetic and prediabetic subjects from Africa and in particular, South Africa. The aim of the current study therefore was to conduct genome-wide DNA methylation in South African subjects with varying glucose tolerance and investigate the relationship between the observed DMRs and cardiometabolic risk factors. Genome-wide DNA methylation profiling was conducted in 48 South African individuals from the Bellville South community in the Western Cape, using DNA immunoprecipitation sequencing (MeDIP-seq). Following the identification of statistically significant differentially methylated regions (DMRs) and lncRNA-associated DMRs, gene ontology and KEGG pathway analysis was performed for the participants, comparing the findings between those with known diabetes on treatment, screen-detected (newly diagnosed) diabetes, prediabetes and normal glucose tolerance. Also, the significant DMRs identified were validated by performing pyrosequencing of bisulphite converted DNA.

The study identified several DMRs and functional pathways affected in subjects with diabetes and prediabetes. These findings show that DNA methylation patterns differ amongst individuals with varying degrees of glucose tolerance within a South African population. Furthermore, the study showed that DNA methylation patterns are associated with certain cardiometabolic traits and diabetic complications, and could be used as potential biomarkers for the occurrence and progression of T2DM. For example, biological processes associated with cardiac muscle contraction, blood circulation and vasoconstriction were found to be hypermethylated in the screen-detected diabetic subjects when compared to the subjects with normal glucose tolerance. These findings suggest that cardiovascular processes may be affected in diabetic individuals due to a suggested decrease in the expression of the genes associated with these functions. Also, Notch signalling, which plays a role in many processes across a wide range of tissue, was found to be hypermethylated in the screen-detected diabetic subjects. The implied decrease in Notch signalling could be an attempt to regulate insulin and glucose homeostasis in diabetic individuals, however, this statement warrants further investigation. Pathway analysis in the screen-detected subjects revealed hypomethylation of purine metabolism and its associated DMR, ADK when compared to the subjects with normal glucose tolerance. The overexpression of purine metabolism and its end-product uric acid has been linked to diseases such as gout, obesity, hypertension, hyperlipidaemia, renal dysfunction and diabetes (Maiuolo *et al.*, 2016; Xiong *et al*., 2019). The hypomethylation and implied increase in expression of purine metabolism in this study suggests a role for DNA methylation in highlighting the risk of increased ROS production, inflammation and cell damage associated with excess uric acid in diabetic individuals.

The study also identified DMRs and their functional pathways which showed a possible progression from prediabetes to diabetes. The defense to Gram-negative bacterium and its associated DMR, LBP (Lipopolysaccharide Binding Protein) was hypomethylated in prediabetic subjects suggesting an increased expression of LBP levels in these subjects. Increased LBP levels have been associated with the release of increased inflammatory cytokines, insulin resistance and beta-cell dysfunction in diabetes (Tilves *et al.*, 2016). Therefore the increased expression of LBP in this study indicates the potential use of this DMR as a biomarker for prediabetes with the potential for progression to diabetes. When comparing prediabetic subjects to the screen-detected diabetic subjects, hypomethylation of the regulation of wound healing and blood coagulation and their associated DMRs, SERPINF2 and DMTN, was observed. As SERPINF2 plays a role in regulating fibrinolysis (Viganò *et al.*, 2018), the hypomethylation and elevated levels of SERPINF2 (alpha 2-antiplasmin) expression in subjects with hyperglycaemia could be used as an indicator of progression to full-blown diabetes and a risk for cardiovascular disease. Furthermore, DMTN (Dematin), which plays a role in erythrocyte shape and membrane stability is affected by hyperglycaemia resulting in changes in the glucose uptake through GLUT1 in red blood cells (Lu *et al.*, 2016; Guizouarn and Allegrini, 2020). Therefore the increased expression of DMTN in this study indicates the potential use of this DMR as an indicator of red blood cell stability with worsening glucose tolerance.

The expression of several DMRs and pathways were affected by the metformin treatment in the known diabetic subjects, showing that DNA methylation could be utilized as potential biomarkers for monitoring treatment regimes or developing new strategies. For example, the cytokine-cytokine receptor interaction and oxidative phosphorylation pathways were hypermethylated in known diabetics on metformin treatment when compared to the subjects with normal glucose tolerance. This suggested that metformin may have an inhibitory effect on complement-mediated inflammation and mitochondrial oxidative phosphorylation in diabetic individuals. In addition, hypermethylation of the functional pathways and DMRs identified in metformin-treated subjects when compared to the untreated screen-detected (newly diagnosed) subjects were associated with Diabetic Peripheral Neuropathy (DPN), suggesting that metformin may have a role in dampening the effects of DPN. When compared to prediabetic individuals, hypermethylation of the cAMP signalling pathway in known diabetics on treatment suggested a decrease in cAMP synthesis and cAMP-dependent protein kinase (PKA) activity. These findings correlate with a previous study on metformin's inhibitory effect on cAMP and PKA activity (Pernicova and Korbonits, 2014). This, therefore, showed that the DNA methylation patterns observed in this study could be used to monitor the effects of

metformin on glucose homeostasis. The PI3K-AKT signalling pathway and its associated DMR, Insulin Receptor Substrate 1 (IRS1), was hypomethylated when comparing known diabetics on metformin treatment to prediabetic subjects. Metformin has shown to inactivate IRS1 and the PI3K-AKT signalling pathway in diabetic cancer patients through its activation of AMPK (Saini and Yang, 2018). Therefore, the hypomethylation and implied increase in the PI3K-AKT signalling pathway in metformin-treated diabetics in this study suggests that metformin treatment may aid in the management of insulin resistance by enhancing the PI3K-AKT signalling pathway. Since activation of the PI3K-AKT signalling pathway is known to be context-dependent, these findings warrant further investigation.

Several novel lncRNAs were observed when comparing the lncRNA-associated DMRs identified amongst individuals with varying degrees of glucose tolerance. When comparing hyperglycaemic individuals to those with normal glucose tolerance, the complement component C4 was hypomethylated suggesting a possible association with the cardiometabolic risk factors and complement-mediated inflammation associated with T2DM. This association was examined in other studies which linked increased levels of C4 to cardiometabolic risk factors such as obesity, blood pressure, blood lipid profile and metabolic syndrome (Nilsson *et al.*, 2014a; Liu *et al.*, 2016). The findings in the current study therefore suggest that increased expression of C4 could be used as a marker for the higher risk of developing cardiometabolic traits associated with T2DM. Furthermore, lncRNA-associated DMRs in diabetic individuals when compared to those with normal glucose tolerance showed hypermethylation of MLXIP and RPTOR. Both of these lncRNA-associated DMRs were found to be tissue-specific in their actions. The MondoA-interacting protein MLXIP plays a role in glucose-sensing transcriptional activity (Richards *et al.*, 2017) and could be a potential target for anti-diabetic therapy through its inhibition in various tissues. In the current study, the hypermethylation and possible repression of MLXIP in diabetic individuals could be used as a potential marker for monitoring improved glucose uptake in tissues while on anti-diabetic treatment. Metformin has also shown to inhibit RPTOR, a part of the mTOR complex 1, which when elevated in certain tissues has been associated with increased beta-cell dysfunction and metabolic stress in T2DM (Soliman *et al*., 2016; Yuan *et al.*, 2017). In the current study, the hypermethylation and implied suppression of RPTOR in diabetic subjects on metformin treatment suggest that RPTOR expression may have an influence on the metabolism and energy homeostasis in these individuals.

When comparing known diabetics on metformin treatment to the newly diagnosed diabetic subjects, the lncRNA-associated DMRs observed included the mitochondrial ATP synthase-coupling factor 6 (ATP5J) enzyme. Increased levels of the ATP5J enzyme thought to be involved in the oxidative phosphorylation pathway, has been associated with T2DM and metabolic acidosis in hypertension (Osanai *et al.*, 2012). Further investigation is needed in the current study to determine whether the hypermethylation of this lncRNA suggests possible repression of ATP5J in metformin-treated diabetics. Lastly, the hypomethylated lncRNA-associated DMRs identified in metformin-treated diabetics when compared to untreated newly diagnosed subjects included lysyl oxidase-like 2 (LOXL2). Increased levels of LOXL2 are associated with increased tissue fibrosis and have recently been implicated in diabetic nephropathy and diabetic cardiomyopathy (Erasmus *et al.*, 2020). Furthermore, the effect of metformin on LOXL2 has been observed in a previous animal model where metformin down-regulated the expression of LOXL2 (Li *et al.*, 2016b). In the current study, the hypermethylation of LOXL2 in the known diabetic subjects therefore suggests that metformin treatment may aid in the decreased expression of LOXL2. Moreover, LOXL2 may serve as a potential marker for the early detection of diabetic-related complications. In conclusion, the lncRNA-associated DMRs observed in this study may serve as potential targets for the detection of diabetic-related complications as well as the therapeutic monitoring in diabetic patients.

5.4 Future recommendations

The study has identified several DMRs and lncRNA-associated DNA methylation regions within a South African population which could be used as potential biomarkers for the monitoring of T2DM. These findings will however benefit from longitudinal studies which can ascertain the relationship between DNA methylation patterns and cardiometabolic risk factors. Moreover, the longitudinal studies can incorporate participants from varying age groups, gender and ethnicity and monitor the DNA methylation changes in the progression of the disease. Longitudinal studies will also aid in determining whether the DMRs and lncRNA-associated DMRS identified in metformin-treated subjects were not influenced by the duration and severity of diabetes.

In terms of the genetic heterogeneity and its effect on DNA methylation within South Africa populations, future studies investigating methylation quantitative trail loci (mQTLs) together

with ancestry markers for population stratification may aid in expanding the genome-wide association studies of diabetes within Africa. Furthermore, meta-analyses using larger sample sizes will permit a better understanding of the DNA methylation profile of T2DM within a South African context and perhaps identify additional DMRs not detected in this study.

Future DNA methylation studies may also benefit from the generation of additional transcription or expression data to ascertain whether the candidate genes identified are actually suppressed or overly expressed in cases of hypermethylation and hypomethylation. Further investigation is also needed on the DMRs and pathways identified in the current study where the results were not conclusive. Moreover, the novel lncRNAs identified in this study warrants further investigation to establish their possible roles in T2DM.

Although blood is still a valuable tissue used in genome-wide DNA methylation assays, additional methods are required to explore the effect of its cellular heterogeneity on DNA methylation. These methods may include direct measurement of counts of the various cell types which could be expensive and time-consuming in large study populations. Alternatively, there is a need for advanced digital sequencing platforms with established reference data on cell type-specific epigenomic profiles or the development of statistical methods that account for cell-type composition.

In conclusion, T2DM is often diagnosed years after its onset and manifestation of micro- and macrovascular complications. It is, therefore, necessary to understand the relationship between DNA methylation and the early pathogenic mechanisms of the disease to develop preventative mechanisms. The DNA methylation patterns identified in this study may therefore aid in the detection of early and established risk factors for T2DM as well as contribute to the therapeutic monitoring of the disease.

REFERENCES

Abd-Elsayed, A., Jackson, M., Gu, S.L., Fiala, K. & Gu, J. 2019. Neuropathic pain and Kv7 voltage-gated potassium channels: The potential role of Kv7 activators in the treatment of neuropathic pain. *Molecular Pain*, 15: 1–8.

Abe, M., Wu, Z., Yamamoto, M., Jin, J.J., Tabara, Y., Mogi, M., Kohara, K., Miki, T. & Nakura, J. 2005. Association of dopamine β-hydroxylase polymorphism with hypertension through interaction with fasting plasma glucose in Japanese. *Hypertension Research*, 28(3): 215–221.

Agren, A., Jorneskog, G., Elgue, G., Henriksson, P., Wallen, H. & Wiman, B. 2014. Increased incorporation of antiplasmin into the fibrin network in patients with Type 1 diabetes. *Diabetes Care*, 37(7): 2007–2014.

Ahmed, M.A., Muntingh, G. & Rheeder, P. 2016. Vitamin B12 deficiency in metformin-treated type-2 diabetes patients, prevalence and association with peripheral neuropathy. *BMC Pharmacology and Toxicology*, 17(1): 1–10.

Ahmed, M.A., Muntingh, G.L. & Rheeder, P. 2017. Perspectives on peripheral neuropathy as a consequence of metformin-induced vitamin B12 deficiency in T2DM. *International Journal of Endocrinology*, 2452853.

Ahmed, S., Ansari, S., Mensah-Brown, E. & Emerald, B. 2020. The role of DNA methylation in the pathogenesis of type 2 diabetes mellitus. C*linical Epigenetics*, 12(104): 1–23.

Ahn, B., Vega, R.B., Kelly, D.P., Ahn, B., Soundarapandian, M.M., Sessions, H., Peddibhotla, S., Roth, G.P., Li, J., Sugarman, E., Koo, A., Malany, S., Wang, M., Yea, K., Brooks, J., Leone, T.C., Han, X., Vega, R.B. & Kelly, D.P. 2016. MondoA coordinately regulates skeletal myocyte lipid homeostasis and insulin signaling. *The Journal of Clinical Investigation*, 126(9): 3567–3579.

Ahn, B., Won, K.J., Kelly, D.P., Ahn, B., Wan, S., Jaiswal, N., Vega, R.B. & Ayer, D.E. 2019. MondoA drives muscle lipid accumulation and insulin resistance. *JCI Insight*, 4(15): e129119.

Akash, M.S.H., Rehman, K. & Liaqat, A. 2018. Tumor Necrosis Factor-Alpha: Role in development of insulin resistance and pathogenesis of Type 2 Diabetes mellitus. *Journal of Cellular Biochemistry*, 119: 105–110.

Alberti, K.G.M.M. & Zimmet, P.Z. 1998. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. *Diabetic Medicine*, 15(7): 539–553.

Al-Haddad, R., Karnib, N., Assaad, R.A., Bilen, Y., Emmanuel, N., Ghanem, A., Younes, J., Zibara, V., Stephan, J.S. & Sleiman, S.F. 2016. Epigenetic changes in diabetes. *Neuroscience Letters*, 625: 64–69.

Ali, M.K., Siegel, K.R., Chandrasekar, E., Tandon, N., Montoya, P.A., Mbanya, J.C., Chan, J., Zhang, P. & Narayan, K.M.V. 2017a. Diabetes: An update on the pandemic and potential solutions. In Prabhakaran D, S. Anand, S. Gaziano, T. A., Mbanya, J. C., Wu, Y & Nugent, R (eds). *Disease Control Priorities, Third Edition (Volume 5): Cardiovascular, Respiratory, and Related Disorders*. Washington (DC): 79–99.

Ali, M.U., Ur Rahman, M.S., Jia, Z. & Jiang, C. 2017b. Eukaryotic translation initiation factors and cancer. *Tumor Biology*, 39(6): 1010428317709805.

Al-Shukaili, A., Al-Ghafri, S., Al-Marhoobi, S., Al-Abri, S., Al-Lawati, J. & Al-Maskari, M. 2013. Analysis of inflammatory mediators in type 2 diabetes patients. *International Journal of Endocrinology*, 2013: 8–10.

Alzahrani, S.H. & Ajjan, R.A. 2010. Coagulation and fibrinolysis in diabetes. *Diabetes and Vascular Disease Research*, 7(4): 260–273.

American Diabetes Association. 2019. Classification and diagnosis of diabetes: Standards of medical care in diabetes 2019. *Diabetes Care*, 42(Supplement 1): S13–S28.

American Diabetes Association. 2020. Classification and diagnosis of diabetes: Standards of medical care in diabetes 2020. *Diabetes Care*, 43(Suppl. 1): S14–S31.

Antonioli, L., Blandizzi, C., Csoka, B., Pacher, P. & Hasko, G. 2015. Adenosine signalling in diabetes mellitus - pathophysiology and therapeutic considerations. *Nature Reviews Endocrinology*, 11(4): 228–241.

Ardestani, A., Lupse, B., Kido, Y., Leibowitz, G. & Maedler, K. 2017. mTORC1 Signaling: A double-edged sword in diabetic β cells. *Cell Metabolism*, 27(2): 314–331.

Argos, M., Chen, L., Jasmine, F., Tong, L., Pierce, B.L., Roy, S., Paul-Brutus, R., Gamble, M. V., Harper, K.N., Parvez, F., Rahman, M., Rakibuz-Zaman, M., Slavkovich, V., Baron, J.A., Graziano, J.H., Kibriya, M.G. & Ahsan, H. 2015. Gene-specific differential DNA methylation and chronic arsenic exposure in an epigenome-wide association study of adults in Bangladesh. *Environmental Health Perspectives*, 123(1): 64–71.

Arnold, A.C., Garland, E.M., Celedonio, J.E., Raj, S.R., Abumrad, N.N., Biaggioni, I., Robertson, D., Luther, J.M. & Shibao, C.A. 2017. Hyperinsulinemia and insulin resistance in dopamine β-hydroxylase deficiency. *Journal of Clinical Endocrinology and Metabolism*, 102(1): 10–14.

Arpón, A., Milagro, F.I., Ramos-Lopez, O., Mansego, M.L., Santos, J.L., Riezu-Boj, J.I. & Martínez, J.A. 2019. Epigenome-wide association study in peripheral white blood cells involving insulin resistance. *Scientific Reports*, 9(1): 1–11.

Ashikov, A., Routier, F., Fuhlrott, J., Helmus, Y., Wild, M., Gerardy-Schahn, R. & Bakker, H. 2005. The human solute carrier gene SLC35B4 encodes a bifunctional nucleotide sugar transporter with specificity for UDP-xylose and UDP-N-acetylglucosamine. *Journal of Biological Chemistry*, 280(29): 27230–27235.

Asmelash, D. & Asmelash, Y. 2019. The burden of undiagnosed Diabetes mellitus in adult African population : A systematic review and meta-analysis. *Journal of Diabetes Research*: 4134937.

Atun, R., Davies, J.I., Gale, E.A.M., Bärnighausen, T., Beran, D., Kengne, A.P., Levitt, N.S., Mangugu, F.W., Nyirenda, M.J., Ogle, G.D., Ramaiya, K., Sewankambo, N.K., Sobngwi, E., Tesfaye, S., Yudkin, J.S., Basu, S., Bommer, C., Heesemann, E., Manne-Goehler, J., Postolovska, I., Sagalova, V., Vollmer, S., Abbas, Z.G., Ammon, B., Angamo, M.T., Annamreddi, A., Awasthi, A., Besançon, S., Bhadriraju, S., Binagwaho, A., Burgess, P.I., Burton, M.J., Chai, J., Chilunga, F.P., Chipendo, P., Conn, A., Joel, D.R., Eagan, A.W.,

Gishoma, C., Ho, J., Jong, S., Kakarmath, S.S., Khan, Y., Kharel, R., Kyle, M.A., Lee, S.C., Lichtman, A., Malm, C.P., Mbaye, M.N., Muhimpundu, M.A., Mwagomba, B.M., Mwangi, K.J., Nair, M., Niyonsenga, S.P., Njuguna, B., Okafor, O.L.O., Okunade, O., Park, P.H., Pastakia, S.D., Pekny, C., Reja, A., Rotimi, C.N., Rwunganira, S., Sando, D., Sarriera, G., Sharma, A., Sidibe, A., Siraj, E.S., Syed, A.S., Van Acker, K. & Werfalli, M. 2017. Diabetes in sub-Saharan Africa: from clinical care to health policy. *The Lancet Diabetes and Endocrinology*, 5(8): 622–667.

Bai, B., Xie, B., Pan, Z., Shan, L., Zhao, J. & Zhu, H. 2018. Identification of candidate genes and long non-coding RNAs associated with the effect of ATP5J in colorectal cancer. *International Journal of Oncology*, 52(4): 1129–1138.

Bakulski, K.M., Dou, J., Lin, N., London, S.J. & Colacino, J.A. 2019. DNA methylation signature of smoking in lung cancer is enriched for exposure signatures in newborn and adult blood. *Scientific Reports*, 9(4576).

Balzer, S., Malde, K. & Jonassen, I. 2011. Systematic exploration of error sources in pyrosequencing flowgram data. *Bioinformatics*, 27(13): i304–i309.

Banerjee, P., Surendran, H., Chowdhury, D.R., Prabhakar, K. & Pal, R. 2016. Metformin mediated reversal of epithelial to mesenchymal transition is triggered by epigenetic changes in E-cadherin promoter. *Journal of Molecular Medicine*, 94(12): 1397–1409.

Bansal, A. & Pinney, S.E. 2017. DNA methylation and its role in the pathogenesis of diabetes. *Pediatric Diabetes*, 18(3): 167–177.

Barnum, S.R. 2015. C4a: An anaphylatoxin in name only. *Journal of Innate Immunity*, 7(4): 333–339.

Barrie, E.S., Weinshenker, D., Verma, A., Pendergrass, S.A., Lange, L.A., Ritchie, M.D., Wilson, J.G., Kuivaniemi, H., Tromp, G., Carey, D.J., Gerhard, G.S., Brilliant, M.H., Hebbring, S.J., Cubells, J.F., Pinsonneault, J.K., Norman, G.J. & Sadee, W. 2014. Regulatory polymorphisms in human DBH affect peripheral gene expression and sympathetic activity. *Circulation Research*, 115(12): 1017–1025.

Bartolome, A., Zhu, C., Sussel, L. & Pajvani, U.B. 2019. Notch signaling dynamically regulates adult β cell proliferation and maturity. *Journal of Clinical Investigation*, 129(1): 268–280.

Bavuma, C., Sahabandu, D., Musafiri, S., Danquah, I., McQuillan, R. & Wild, S. 2019. Atypical forms of diabetes mellitus in Africans and other non-European ethnic populations in low- and middle-income countries: a systematic literature review. *Journal of Global Health*, 9(2): 020401.

Berbudi, A., Rahmadika, N., Tjahjadi, A.I. & Ruslami, R. 2019. Type 2 diabetes and its impact on the immune system. *Current Diabetes Reviews*, 16(5): 442–449.

Besse-Patin, A., Jeromson, S., Levesque-Damphousse, P., Secco, B., Laplante, M. & Estall, J.L. 2019. PGC1A regulates the IRS1:IRS2 ratio during fasting to influence hepatic metabolism downstream of insulin. *Proceedings of the National Academy of Sciences of the United States of America*, 116(10): 4285–4290.

Bi, P. & Kuang, S. 2015. Notch signaling as a novel regulator of metabolism. *Trends in Endocrinology and Metabolism*, 26(5): 248–255.

Bierne, H., Hamon, M. & Cossart, P. 2012. Epigenetics and bacterial infections. *Cold Spring Harbor Perspectives in Medicine*, 2(12): a010272.

Blandino-Rosano, M., Barbaresso, R., Jimenez-Palomares, M., Bozadjieva, N., Werneck-de-Castro, J.P., Hatanaka, M., Mirmira, R.G., Sonenberg, N., Hall, M.N., Bernal-Mizrachi, E., Liu, M. & Ruegg, M.A. 2017. Loss of mTORC1 signalling impairs β-cell homeostasis and insulin processing. *Nature Communications*, 8: 16014.

Bock, C., Halbritter, F., Carmona, F.J., Tierling, S., Datlinger, P., Assenov, Y., Berdasco, M., Bergmann, A.K., Booher, K., Busato, F., Campan, M., Dahl, C., Dahmcke, C.M., Diep, D., Fernández, A.F., Gerhauser, C., Haake, A., Heilmann, K., Holcomb, T., Hussmann, D., Ito, M., Kläver, R., Kreutz, M., Kulis, M., Lopez, V., Nair, S.S., Paul, D.S., Plongthongkum, N., Qu, W., Queirós, A.C., Reinicke, F., Sauter, G., Schlomm, T., Statham, A., Stirzaker, C., Strogantsev, R., Urdinguio, R.G., Walter, K., Weichenhan, D., Weisenberger, D.J., Beck, S., Clark, S.J., Esteller, M., Ferguson-Smith, A.C., Fraga, M.F., Guldberg, P., Hansen, L.L., Laird, P.W., Martín-Subero, J.I., Nygren, A.O.H., Peist, R., Plass, C., Shames, D.S., Siebert, R., Sun, X., Tost, J., Walter, J. & Zhang, K. 2016. Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. *Nature Biotechnology*, 34(7): 726–737.

Boison, D. 2013. Adenosine kinase: Exploitation for therapeutic gain. *Pharmacological Reviews*, 65(3): 906–943.

Boles, A., Kandimalla, R. & Reddy, P.H. 2017. Dynamics of diabetes and obesity: Epidemiological perspective. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1863(5): 1026–1036.

Bozek, T., Blazekovic, A., Perkovic, M.N., Jercic, K.G., Sustar, A., Smircic-Duvnjak, L., Outeiro, T.F., Pivac, N. & Borovecki, F. 2017. The influence of dopamine-beta-hydroxylase and catechol O-methyltransferase gene polymorphism on the efficacy of insulin detemir therapy in patients with type 2 diabetes mellitus. *Diabetology and Metabolic Syndrome*, 9: 97.

Breitling, L.P., Yang, R., Korn, B., Burwinkel, B. & Brenner, H. 2011. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *American Journal of Human Genetics*, 88(4): 450–457.

Bridgeman, S.C., Ellison, G.C., Melton, P.E., Newsholme, P. & Mamotte, C.D.S. 2018. Epigenetic effects of metformin: From molecular mechanisms to clinical implications. *Diabetes, Obesity and Metabolism*, 20(7): 1553-1562.

Brown, S.E., Suderman, M.J., Hallett, M. & Szyf, M. 2008. DNA demethylation induced by the methyl-CpG-binding domain protein MBD3. *Gene*, 420(2): 99–106.

Cameron, A.R., Morrison, V.L., Levin, D., Mohan, M., Forteath, C., Beall, C., McNeilly, A.D., Balfour, D.J.K., Savinko, T., Wong, A.K.F., Viollet, B., Sakamoto, K., Fagerholm, S.C., Foretz, M., Lang, C.C. & Rena, G. 2016. Anti-inflammatory effects of metformin irrespective of diabetes status. *Circulation Research*, 119(5): 652–665.

Cao, C., Tan, Q., Xu, C., He, L., Yang, L., Zhou, Ye, Zhou, Yiwei, Qiao, A., Lu, M., Yi, C., Han, G.W., Wang, X., Li, X., Yang, H., Rao, Z., Jiang, H., Zhao, Y., Liu, J. & Stevens, R.C. 2018. Structural basis for signal recognition and transduction by platelet-activating-factor receptor. *Nature Structural & Molecular Biology*, 25: 488–495.

Carlsson, S. 2019. Etiology and pathogenesis of latent autoimmune diabetes in adults (LADA) compared to type 2 diabetes. *Frontiers in Physiology*, 10(MAR).

Carmona, J.J., Sofer, T., Hutchinson, J., Cantone, L., Coull, B., Maity, A., Vokonas, P., Lin, X., Schwartz, J. & Baccarelli, A.A. 2014. Short-term airborne particulate matter exposure alters the epigenetic landscape of human genes associated with the mitogen-activated protein kinase network: A cross-sectional study. *Environmental Health: A Global Access Science Source*, 13(94).

Carter, G., Miladinovic, B., Patel, A.A., Deland, L., Mastorides, S. & Patel, N.A. 2015. Circulating long noncoding RNA GAS5 levels are correlated to prevalence of type 2 diabetes mellitus. *BBA Clinical*, 4: 102–107.

Chai, T.F., Hong, S.Y., He, H., Zheng, L., Hagen, T., Luo, Y. & Yu, F.-X. 2012. A potential mechanism of metformin-mediated regulation of glucose homeostasis : Inhibition of Thioredoxin-interacting protein (Txnip) gene expression. *Cellular Signalling*, 24(8): 1700–1705.

Chalaya, T. V., Akopov, S.B., Nikolaev, L.G. & Sverdlov, E.D. 2006. Tissue specificity of methylation of cytosines in regulatory regions of four genes located in the locus FXYD5-COX7A1 of human chromosome 19: Correlation with their expression level. *Biochemistry (Moscow)*, 71(3): 294–299.

Chambers, J.C., Loh, M., Lehne, B., Drong, A., Kriebel, J., Motta, V., Wahl, S., Elliott, H.R., Rota, F., Scott, W.R., Zhang, W., Tan, S.T., Campanella, G., Chadeau-Hyam, M., Yengo, L., Richmond, R.C., Adamowicz-Brice, M., Afzal, U., Bozaoglu, K., Mok, Z.Y., Ng, H.K., Pattou, F., Prokisch, H., Rozario, M.A., Tarantini, L., Abbott, J., Ala-Korpela, M., Albetti, B., Ammerpohl, O., Bertazzi, P.A., Blancher, C., Caiazzo, R., Danesh, J., Gaunt, T.R., de Lusignan, S., Gieger, C., Illig, T., Jha, S., Jones, S., Jowett, J., Kangas, A.J., Kasturiratne, A., Kato, N., Kotea, N., Kowlessur, S., Pitkäniemi, J., Punjabi, P., Saleheen, D., Schafmayer, C., Soininen, P., Tai, E.S., Thorand, B., Tuomilehto, J., Wickremasinghe, A.R., Kyrtopoulos, S.A., Aitman, T.J., Herder, C., Hampe, J., Cauchi, S., Relton, C.L., Froguel, P., Soong, R., Vineis, P., Jarvelin, M.R., Scott, J., Grallert, H., Bollati, V., Elliott, P., McCarthy, M.I. & Kooner, J.S. 2015. Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: A nested case-control study. *The Lancet Diabetes and Endocrinology*, 3(7): 526–534.

Chang, W., Zhang, L., Yao, X., Chen, Y., Zhu, L., Fang, Z., Zhao, Y., Yao, Y. & Jin, Y. 2020. Upregulation of long non-coding RNA MEG3 in type 2 diabetes mellitus complicated with vascular disease: a case-control study. *Molecular and Cellular Biochemistry*, 473(1–2): 93–99.

Chen, D.P., Lin, Y.C. & Fann, C.S.J. 2016. Methods for identifying differentially methylated regions for sequence- and array-based data. *Briefings in Functional Genomics*, 15(6): 485–490.

Chen, J., Meng, Y., Zhou, J., Zhuo, M., Ling, F., Zhang, Y., Du, H. & Wang, X. 2013. Identifying candidate genes for type 2 diabetes mellitus and obesity through gene expression profiling in multiple tissues or cells. *Journal of Diabetes Research*, (970435): 1–9.

Chen, J., Sun, M., Adeyemo, A., Pirie, F., Carstensen, T., Pomilla, C., Doumatey, A.P., Chen, G., Young, E.H., Sandhu, M., Morris, A.P., Barroso, I., McCarthy, M.I., Mahajan, A., Wheeler,

E., Rotimi, C.N. & Motala, A.A. 2019. Genome-wide association study of type 2 diabetes in Africa. *Diabetologia*, 62(7): 1204–1211.

Chen, Y., Zhao, X. and Wu, H. 2019b. Metabolic stress and cardiovascular disease in Diabetes mellitus. *Arteriosclerosis, Thrombosis and Vascular Biology*, 39(10): 1911–1924.

Chidambaram, M., Liju, S., Saboo, B., Sathyavani, K., Viswanathan, V., Pankratz, N., Gross, M., Mohan, V. & Radha, V. 2016. Replication of genome-wide association signals in Asian Indians with early-onset type 2 diabetes. *Acta Diabetologica*, 53(6): 915–923.

Cho, N.H., Shaw, J.E., Karuranga, S., Huang, Y., da Rocha Fernandes, J.D., Ohlrogge, A.W. & Malanda, B. 2018. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Research and Clinical Practice*, 138: 271–281.

Chu, S.K. & Yang, H.C. 2017. Interethnic DNA methylation difference and its implications in pharmacoepigenetics. *Epigenomics*, 9(11): 1437–1454.

Chu, W. 2013. Tumor necrosis factor. *Cancer Letters*, 328(2): 222–225.

Chung, M.M., Chen, Y.L., Pei, D., Cheng, Y.C., Sun, B., Nicol, C.J., Yen, C.H., Chen, H.M., Liang, Y.J. & Chiang, M.C. 2015. The neuroprotective role of metformin in advanced glycation end product treated human neural stem cells is AMPK-dependent. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1852(5): 720–731.

Coco, C., Sgarra, L., Potenza, M.A., Nacci, C., Pasculli, B., Barbano, R., Parrella, P. & Montagnani, M. 2019. Can epigenetics of endothelial dysfunction represent the key to precision medicine in type 2 diabetes mellitus? *International Journal of Molecular Sciences*, 20(12): 1–26.

Cole, J.B. & Florez, J.C. 2020. Genetics of diabetes mellitus and diabetes complications. *Nature Reviews Nephrology*, 16(7): 377–390.

Coovadia, H., Jewkes, R., Barron, P., Sanders, D. & McIntyre, D. 2009. The health and health system of South Africa: historical roots of current public health challenges. *The Lancet*, 374(9692): 817–834.

Copenhaver, M.M., Yu, C.-Y., Zhou, D. & Hoffman, R.P. 2020. Relationships of complement components C3 and C4 and their genetics to cardiometabolic risk in healthy, non-Hispanic white adolescents. *Pediatric Research*, 87(1): 88–94.

Copps, K.D. & White, M.F. 2012. Regulation of insulin sensitivity by serine/threonine phosphorylation of insulin receptor substrate proteins IRS1 and IRS2. *Diabetologia*, 55(10): 2565–2582.

Costantino, S., Mohammed, S.A., Ambrosini, S. & Paneni, F. 2019. Epigenetic processing in cardiometabolic disease. *Atherosclerosis*, 281: 150–158.

Crujeiras, A.B., Diaz-Lagares, A., Sandoval, J., Milagro, F.I., Navas-Carretero, S., Carreira, M.C., Gomez, A., Hervas, D., Monteiro, M.P., Casanueva, F.F., Esteller, M. & Martinez, J.A. 2017. DNA methylation map in circulating leukocytes mirrors subcutaneous adipose tissue methylation pattern: A genome-wide analysis from non-obese and obese patients. *Scientific Reports*, 7(41903): 1–13.

Cui, Y. & Irudayaraj, J. 2015. Dissecting the behavior and function of MBD3 in DNA methylation homeostasis by single-molecule spectroscopy and microscopy. *Nucleic Acids Research*, 43(6): 3046–3055.

Cuyàs, E., Fernández-Arroyo, S., Verdura, S., García, R.Á.F., Stursa, J., Werner, L., Blanco-González, E., Montes-Bayón, M., Joven, J., Viollet, B., Neuzil, J. & Menendez, J.A. 2018. Metformin regulates global DNA methylation via mitochondrial one-carbon metabolism. *Oncogene*, 37(7): 963–970.

Dahlman, I., Forsgren, M., Sjögren, A., Nordström, E.A., Kaaman, M., Näslund, E., Attersand, A. & Arner, P. 2006. Downregulation of electron transport chain genes in visceral adipose tissue in type 2 diabetes independent of obesity and possibly involving tumor necrosis factor-α. *Diabetes*, 55(6): 1792–1799.

Dang, M.N., Buzzetti, R. & Pozzilli, P. 2013. Epigenetics in autoimmune diseases with focus on type 1 diabetes. *Diabetes/Metabolism Research and Reviews*, 29(1): 8–18.

Davegårdh, C., García-Calzón, S., Bacos, K. & Ling, C. 2018. DNA methylation in the pathogenesis of type 2 diabetes in humans. *Molecular Metabolism*: 1–14.

Dayeh, T., Volkov, P., Salo, S., Hall, E., Nilsson, E., Olsson, A.H., Kirkpatrick, C.L., Wollheim, C.B., Eliasson, L., Ronn, T., Bacos, K. & Ling, C. 2014. Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS Genetics*, 10(3): e1004160.

Dayeh, T.A., Olsson, A.H., Volkov, P., Almgren, P., Rönn, T. & Ling, C. 2013. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia*, 56(5): 1036–1046.

De Chiara, L., Leiro-Fernandez, V., Rodríguez-Girondo, M., Valverde, D., Botana-Rial, M.I. & Fernández-Villar, A. 2020. Comparison of bisulfite pyrosequencing and methylation-specific qPCR for methylation assessment. *International Journal of Molecular Sciences*, 21: 9242.

De Mello, V.D.F., Pulkkinen, L., Lalli, M., Kolehmainen, M., Pihlajamäki, J. & Uusitupa, M. 2014. DNA methylation in obesity and type 2 diabetes. *Annals of Medicine*, 46(3): 103–113.

Deaton, A. & Bird, A. 2011. CpG islands and the regulation of transcription. *Genes & Development*, 25(10): 1010–1022.

Deb, D.K., Bao, R. & Li, Y.C. 2017. Critical role of the cAMP-PKA pathway in hyperglycemia-induced epigenetic activation of fibrogenic program in the kidney. *FASEB Journal*, 31(5): 2065–2075.

Deepa, M., Anjana, R.M. & Mohan, V. 2017. Role of lifestyle factors in the epidemic of diabetes: Lessons learnt from India. *European Journal of Clinical Nutrition*, 71(7): 825–831.

Delaney, C., Garg, S. & Yung, R. 2015. Analysis of DNA methylation by pyrosequencing. *Methods in Molecular Biology*, 1343: 249–264.

Deng, F., Price, M.G., Davis, C.F., Mori, M. & Burgess, D.L. 2006. Stargazin and other transmembrane AMPA receptor regulating proteins interact with synaptic scaffolding protein MAGI-2 in brain. *Journal of Neuroscience*, 26(30): 7875–7884.

Derr, R., Garrett, E., Stacy, G.A. & Saudek, C.D. 2003. Is HbA1c affected by glycemic instability? *Diabetes Care*, 26(10): 2728–2733.

Dessie, G., Mulugeta, H., Amare, D., Negesse, A., Wagnew, F., Getaneh, T., Endalamew, A., Adamu, Y.W., Tadesse, G., Workineh, A. & Lebu, S. 2020. A systematic analysis on prevalence and sub-regional distribution of undiagnosed diabetes mellitus among adults in African countries. *Journal of Diabetes and Metabolic Disorders*.

Dewey, S., Lai, X., Witzmann, F.A., Sohal, M. & Gomes, A. V. 2013. Proteomic analysis of hearts from Akita mice suggests that increases in soluble epoxide hydrolase and antioxidative programming are key changes in early stages of diabetic cardiomyopathy. *Journal of Proteome*, 12(9): 3920–3933.

Dhingra, R., Nwanaji-Enwerem, J.C., Samet, M. & Ward-Caviness, C.K. 2018. DNA methylation age - environmental influences, health impacts, and its role in environmental epidemiology. *Current Environmental Health Reports*, 5(3): 317–327.

DiMeglio, L.A., Evans-Molina, C. & Oram, R.A. 2018. Type 1 diabetes. *Lancet*, 391(10138): 2449–2462.

Docherty, L.E., Rezwan, F.I., Poole, R.L., Jagoe, H., Lake, H., Lockett, G.A., Arshad, H., Wilson, D.I., Holloway, J.W., Temple, I.K. & Mackay, D.J.G. 2014. Genome-wide DNA methylation analysis of patients with imprinting disorders identifies differentially methylated regions associated with novel candidate imprinted genes. *Journal of Medical Genetics*, 51(4): 229–238.

Down, T.A., Rakyan, V.K., Turner, D.J., Flicek, P., Li, H., Kulesha, E., Gräf, S., Johnson, N., Herrero, J., Tomazou, E.M., Thorne, N.P., Bäckdahl, L., Herberth, M., Howe, K.L., David, K., Miretti, M.M., Marioni, J.C., Birney, E., Hubbard, T.J.P., Tavaré, S. & Beck, S. 2008. A Bayesian deconvolution strategy for immunoprecipitation- based DNA methylome analysis. *Nature Biotechnology.*, 26(7): 779–785.

Du Plessis, L., Skunca, N. & Dessimoz, C. 2011. The what, where, how and why of gene ontology - a primer for bioinformaticians. *Briefings in Bioinformatics*, 12(6): 723–735.

Du, Q., Luu, P.L., Stirzaker, C. & Clark, S.J. 2015. Methyl-CpG-binding domain proteins: Readers of the epigenome. *Epigenomics*, 7(6): 1051–1073.

Dugué, P.A., Bassett, J.K., Joo, J.E., Baglietto, L., Jung, C.H., Wong, E.M., Fiorito, G., Schmidt, D., Makalic, E., Li, S., Moreno-Betancur, M., Buchanan, D.D., Vineis, P., English, D.R., Hopper, J.L., Severi, G., Southey, M.C., Giles, G.G. & Milne, R.L. 2018. Association of DNA methylation-based biological age with health risk factors and overall and cause-specific mortality. *American Journal of Epidemiology*, 187(3): 529–538.

Dujic, T., Zhou, K., Yee, S.W., van Leeuwen, N., de Keyser, C.E., Javorský, M., Goswami, S., Zaharenko, L., Hougaard Christensen, M.M., Out, M., Tavendale, R., Kubo, M., Hedderson, M.M., van der Heijden, A.A., Klimčáková, L., Pirags, V., Kooy, A., Brøsen, K., Klovins, J., Semiz, S., Tkáč, I., Stricker, B.H., Palmer, C.N.A., Hart, L.M., Giacomini, K.M. & Pearson, E.R. 2017. Variants in pharmacokinetic transporters and glycemic response to metformin: A Metgen meta-analysis. *Clinical Pharmacology and Therapeutics*, 101(6): 763–772.

Ehrlich, M. 2019. DNA hypermethylation in disease: mechanisms and clinical relevance. *Epigenetics*, 14(12): 1141–1163.

Ekoru, K., Doumatey, A., Bentley, A.R., Chen, G., Zhou, J., Shriner, D., Fasanmade, O., Okafor, G., Eghan, B., Agyenim-Boateng, K., Adeleye, J., Balogun, W., Amoah, A., Acheampong, J., Johnson, T., Oli, J., Adebamowo, C., Collins, F., Dunston, G., Adeyemo, A. & Rotimi, C. 2019. Type 2 diabetes complications and comorbidity in Sub-Saharan Africans. *EClinicalMedicine*, 16: 30–41.

Ekpenyong, C. & Akpan, E. 2014. Abnormal serum uric acid levels in health and disease: A double-edged sword. *American Journal of Internal Medicine*, 2(6): 113–130.

Elbere, I., Silamikelis, I., Ustinova, M., Kalnina, I., Zaharenko, L., Peculis, R., Konrade, I., Ciuculete, D.M., Zhukovsky, C., Gudra, D., Radovica-Spalvina, I., Fridmanis, D., Pirags, V., Schiöth, H.B. & Klovins, J. 2018. Significantly altered peripheral blood cell DNA methylation profile as a result of immediate effect of metformin use in healthy individuals. *Clinical Epigenetics*, 10(156): 1–13.

Elgendy, K., Malcomson, F.C., Lara, J.G., Bradburn, D.M. & Mathers, J.C. 2018. Effects of dietary interventions on DNA methylation in adult humans: Systematic review and meta-analysis. *British Journal of Nutrition*, 120(9): 961–976.

El-Kharrag, R., Owen, R. & Boison, D. 2019. Adenosine kinase deficiency increases susceptibility to a carcinogen. *Journal of Caffeine and Adenosine Research*, 9(1): 4–11.

Eom, Y.S., Gwon, A.R., Kwak, K.M., Youn, J.Y., Park, H., Kim, K.W. & Kim, B.J. 2020. Notch1 has an important role in β-cell mass determination and development of diabetes. *Diabetes and Metabolism Journal*, 44: 1–12.

Erasmus, M., Samodien, E., Lecour, S., Cour, M., Lorenzo, O., Dludla, P., Pheiffer, C. & Johnson, R. 2020. Linking LOXL2 to Cardiac Interstitial Fibrosis. *International Journal of Molecular Sciences*, 21(16): 5913.

Erasmus, R.T., Soita, D.J., Hassan, M.S., Blanco-Blanco, E., Vergotine, Z., Kengne, A.P. & Matsha, T.E. 2012. High prevalence of diabetes mellitus and metabolic syndrome in a South African coloured population: Baseline data of a study in Bellville, Cape Town. *South African Medical Journal*, 102(11): 841–844.

Estellar, M. 2008. Epigenetics in cancer. *The New England Journal of Medicine*, 358(1): 1148–1159.

Fakruddin, M. & Chowdhury, A. 2012. Pyrosequencing - An alternative to traditional Sanger sequencing. *American Journal of Biochemistry and Biotechnology*, 8(1): 14–20.

Fakruddin, Mazumdar, R.M., Chowdhury, A., Hossain, N., Mahajan, S. & Islam, S. 2013. Pyrosequencing - A next generation sequencing technology. *World Applied Sciences Journal*, 24(12): 1558–1571.

Fang, Y. & Fullwood, M.J. 2016. Roles, functions, and mechanisms of long non-coding RNAs in cancer. *Genomics, Proteomics and Bioinformatics*, 14(1): 42–54.

Farré, P., Jones, M.J., Meaney, M.J., Emberly, E., Turecki, G. & Kobor, M.S. 2015. Concordant and discordant DNA methylation signatures of aging in human blood and brain.

Epigenetics & Chromatin, 8(19): 1–17.

Fernyhough, P. 2015. Mitochondrial dysfunction in diabetic neuropathy: a series of unfortunate metabolic events. *Current Diabetes Reports*, 15(11): 24–27.

Flannick, J. & Florez, J. 2016. Type 2 diabetes: genetic data sharing to advance complex disease research. *Nature Reviews Genetics*, 17: 535–549.

Flores-Sierra, J., Arredondo-Guerrero, M., Cervantes-Paz, B., Rodríguez-Ríos, D., Alvarado-Caudillo, Y., Nielsen, F.C., Wrobel, Katarzyna, Wrobel, Kazimierz, Zaina, S. & Lund, G. 2016. The trans fatty acid elaidate affects the global DNA methylation profile of cultured cells and in vivo. *Lipids in Health and Disease*, 15(75).

Foks, A.C., Bot, I., Frodermann, V., De Jager, S.C.A., Ter Borg, M., Van Santbrink, P.J., Yagita, H., Kuiper, J. & Van Puijvelde, G.H.M. 2012. Interference of the CD30-CD30L pathway reduces atherosclerosis development. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 32(12): 2862–2868.

Foretz, M., Guigas, B. & Viollet, B. 2019. Understanding the glucoregulatory mechanisms of metformin in type 2 diabetes mellitus. *Nature Reviews Endocrinology*, 15: 569–589.

Foretz, M., Guigas, B., Bertrand, L., Pollak, M. & Viollet, B. 2014. Metformin: From mechanisms of action to therapies. *Cell Metabolism*, 20(6): 953–966.

Forouhi, N.G. & Wareham, N.J. 2019. Epidemiology of diabetes. *Medicine (United Kingdom)*, 47(1): 22–27.

Fosmo, A.L. & Skraastad, Ø.B. 2017. The Kv7 Channel and Cardiovascular Risk Factors. *Frontiers in Cardiovascular Medicine*, 4: 75.

Frangogiannis, N.G. 2019. The extracellular matrix in ischemic and nonischemic heart failure. *Circulation Research*, 125(1): 117–146.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. & Paul, C.L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, 89(5): 1827–1831.

Fujimaki, S. & Kuwabara, T. 2017. Diabetes-induced dysfunction of mitochondria and stem cells in skeletal muscle and the nervous system. *International Journal of Molecular Sciences*, 18(2147): 1–24.

Galanter, J.M., Gignoux, C.R., Oh, S.S., Torgerson, D., Pino-Yanes, M., Thakur, N., Eng, C., Hu, D., Huntsman, S., Farber, H.J., Avila, P.C., Brigino-Buenaventura, E., Lenoir, M.A., Meade, K., Serebrisky, D., Rodríguez-Cintrón, W., Kumar, R., Rodríguez-Santana, J.R., Seibold, M.A., Borrell, L.N., Burchard, E.G. & Zaitlen, N. 2017. Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife*, 6: e20532.

Galicia-Garcia, U., Benito-Vicente, A., Jebari, S., Larrea-Sebal, A., Siddiqi, H., Uribe, K.B., Ostolaza, H. & Martín, C. 2020. Pathophysiology of type 2 diabetes mellitus. *International Journal of Molecular Sciences*, 21(17): 1–34.

García-Calzón, S., Perfilyev, A., Männistö, V., de Mello, V.D., Nilsson, E., Pihlajamäki, J. & Ling, C. 2017. Diabetes medication associates with DNA methylation of metformin transporter genes in the human liver. *Clinical Epigenetics*, 9(102): 1–9.

Gene Ontology Consortium. 2006. The Gene Ontology (GO) project in 2006. *Nucleic Acids Research*, 34: 322–326.

Geng, T., Liu, Y., Xu, Y., Jiang, Y., Zhang, N., Wang, Z., Carmichael, G.G., Taylor, H.S., Li, D. & Huang, Y. 2018. H19 lncRNA promotes skeletal muscle insulin sensitivity in part by targeting AMPK. *Diabetes*, 67(11): 2183–2198.

Gilbert, E.R. & Liu, D. 2012. Epigenetics: the missing link to understanding β-cell dysfunction in the pathogenesis of type 2 diabetes. *Epigenetics*, 7(8): 841–52.

Goedecke, J.H., Mtintsilana, A., Dlamini, S.N. & Pascal Kengne, A. 2017. Type 2 diabetes mellitus in African women. *Diabetes Research and Clinical Practice*, 123: 87–96.

Gomes, J.M.G., Costa, J. de A. & Alfenas, R. de C.G. 2017. Metabolic endotoxemia and diabetes mellitus: A systematic review. *Metabolism: Clinical and Experimental*, 68: 133–144.

González-Becerra, K., Ramos-Lopez, O., Barrón-Cabrera, E., Riezu-Boj, J.I., Milagro, F.I., Martínez-López, E. & Martínez, J.A. 2019. Fatty acids, epigenetic mechanisms and chronic diseases: A systematic review. *Lipids in Health and Disease*, 18: 178.

Grant, S.F.A. 2019. The TCF7L2 locus: A genetic window into the pathogenesis of type 1 and type 2 diabetes. *Diabetes Care*, 42(9): 1624–1629.

Gray, J.H., Owen, R.P. & Giacomini, K.M. 2004. The concentrative nucleoside transporter family, SLC28. *Pflugers Archiv European Journal of Physiology*, 447(5): 728–734.

Guariguata, L., Whiting, D.R., Hambleton, I., Beagley, J., Linnenkamp, U. & Shaw, J.E. 2014. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Research and Clinical Practice*, 103(2): 137–149.

Guerrero-Castillo, S., Baertling, F., Kownatzki, D., Wessels, H.J., Arnold, S., Brandt, U. & Nijtmans, L. 2017. The assembly pathway of mitochondrial respiratory chain complex I. *Cell Metabolism*, 25(1): 128–139.

Guizouarn, H. & Allegrini, B. 2020. Erythroid glucose transport in health and disease. *European Journal of Physiology*, 472: 1371–1383.

Gunther, K., Rust, M., Leers, J., Boettger, T., Scharfe, M., Jarek, M., Bartkuhn, M. & Renkawitz, R. 2013. Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences. *Nucleic Acids Research*, 41(5): 3010–3021.

Guo, K., Elzinga, S., Eid, S., Figueroa-Romero, C., Hinder, L.M., Pacut, C., Feldman, E.L. & Hur, J. 2019. Genome-wide DNA methylation profiling of human diabetic peripheral neuropathy in subjects with type 2 diabetes mellitus. *Epigenetics*, 14(8): 766–779.

Gupta, R., Nagarajan, A. & Wajapeyee, N. 2010. Advances in genome-wide DNA methylation analysis. *Biotechniques*, 49(4).

Guruharsha, K.G., Kankel, M.W. & Artavanis-Tsakonas, S. 2012. The Notch signalling system: recent insights into the complexity of a conserved pathway. *Nature Reviews Genetics*, 13(9): 654–666.

Hadley, B., Litfin, T., Day, C.J., Haselhorst, T., Zhou, Y. & Tiralongo, J. 2019. Nucleotide sugar transporter SLC35 family structure and function. *Computational and Structural Biotechnology Journal*, 17: 1123–1134.

Haitina, T., Lindblom, J., Renström, T. & Fredriksson, R. 2006. Fourteen novel human members of mitochondrial solute carrier family 25 (SLC25) widely expressed in the central nervous system. *Genomics*, 88: 779–790.

Hall, E., Volkov, P., Dayeh, T., Bacos, K., Rönn, T., Nitert, M.D. & Ling, C. 2014. Effects of palmitate on genome-wide mRNA expression and DNA methylation patterns in human pancreatic islets. *BMC Medicine*, 12: 103.

Harris, R.A., Wang, T., Coarfa, C., Nagarajan, R.P., Hong, C., Sara, L., Johnson, B.E., Fouse, S.D., Delaney, A., Zhao, Y., Ballinger, T., Zhou, X., Forsberg, K.J., Gu, J., Echipare, L., O'Geen, H., Lister, R., Pelizzola, M., Xi, Y., Epstein, C.B., Bernstein, B.E., Hawkins, R.D., Ren, B., Chung, W.-Y., Gu, H., Bock, C., Gnirke, A., Zhang, M.Q., Haussler, D., Ecker, J., Li, W., Farnham, P.J., Waterland, R.A., Meissner, A., Marra, M.A. & Costello, J.F. 2010. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnology*, 28(10): 1097–1105.

Hasan, S.S., Jabs, M., Taylor, J., Wiedmann, L., Leibing, T., Nordström, V., Federico, G., Roma, L.P., Carlein, C., Wolff, G., Ekim-Üstünel, B., Brune, M., Moll, I., Tetzlaff, F., Gröne, H., Fleming, T., Géraud, C., Herzig, S., Nawroth, P.P. & Fischer, A. 2020. Endothelial Notch signaling controls insulin transport in muscle. *EMBO Molecular Medicine*, 12(4): e09271.

Hattersley, A.T. & Patel, K.A. 2017. Precision diabetes: learning from monogenic diabetes. *Diabetologia*, 60(5): 769–777.

Hatting, M., Tavares, C.D.J., Sharabi, K., Rines, A.K. & Puigserver, P. 2018. Insulin regulation of gluconeogenesis. *Annals of the New York Academy of Sciences*, 1411(1): 21–35.

He, L. & Wondisford, F.E. 2015. Metformin action: Concentrations matter. *Cell Metabolism*, 21(2): 159–162.

He, X., Ou, C., Xiao, Y., Han, Q., Li, H. & Zhou, S. 2017. LncRNAs: Key players and novel insights into diabetes mellitus. *Oncotarget*, 8(41): 71325–71341.

Heather, J.M. & Chain, B. 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1): 1–8.

Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E. & Lumey, L.H. 2008. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(44): 17046–17049.

Helling, S., Vogt, S., Rhiel, A., Ramzan, R., Wen, L., Marcus, K. & Kadenbach, B. 2008. Phosphorylation and kinetics of mammalian cytochrome c oxidase. *Molecular and Cellular Proteomics*, 7(9): 1714–1724.

Hird, T.R., Pirie, F.J., Esterhuizen, T.M., O'Leary, B., McCarthy, M.I., Young, E.H., Sandhu, M.S. & Motala, A.A. 2016. Burden of diabetes and first evidence for the utility of HbA1c for diagnosis and detection of diabetes in urban black South Africans: The Durban diabetes study. *PLoS ONE*, 11(8): e0161966.

Hori, K., Sen, A. & Artavanis-Tsakonas, S. 2013. Notch signaling at a glance. *Journal of Cell Science*, 126(10): 2135–2140.

Horvath, S. 2013. DNA methylation age of human tissues and cell types. *Genome biology*, 14(10): R115.

Houseman, E.A., Kim, S., Kelsey, K.T. & Wiencke, J.K. 2015. DNA methylation in whole blood: Uses and challenges. *Current Environmental Health Reports*, 2: 145–154.

Howell, J.J., Hellberg, K., Turner, M., Saghatelian, A., Shaw, R.J. & Manning, B.D. 2017. Metformin inhibits hepatic mTORC1 signaling via dose-dependent mechanisms involving AMPK and the TSC complex. *Cell Metabolism*, 25(2): 463–471.

Huang, X., Liu, G., Guo, J. & Su, Z.Q. 2018. The PI3K/AKT pathway in obesity and type 2 diabetes. *International Journal of Biological Sciences*, 14(11): 1483–1496.

Huang, Y., Li, J., Chen, S., Zhao, S., Huang, J., Zhou, J. & Xu, Y. 2020. Identification of Potential Therapeutic Targets and Pathways of Liraglutide Against Type 2 Diabetes Mellitus (T2DM) Based on long non-coding RNA (lncRNA) sequencing. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 26: e922210.

Husby, A. 2020. On the use of blood samples for measuring DNA methylation in ecological epigenetic studies. *Integrative and Comparative Biology*, 60(6): 1558–1566.

Hyun, B., Shin, S., Lee, A., Lee, S., Song, Y., Ha, N.J., Cho, K.H. & Kim, K. 2013. Metformin down-regulates TNF-α secretion via suppression of scavenger receptors in macrophages. *Immune Network*, 13(4): 123.

International Diabetes Federation. 2019. *IDF Diabetes ATLAS 9th edition 2019*.

Ishikawa, K., Tsunekawa, S., Ikeniwa, M., Izumoto, T., Iida, A., Ogata, H., Uenishi, E., Seino, Y., Ozaki, N., Sugimura, Y., Hamada, Y., Kuroda, A., Shinjo, K., Kondo, Y. & Oiso, Y. 2015. Long-term pancreatic beta cell exposure to high levels of glucose but not palmitate induces DNA methylation within the insulin gene promoter and represses transcriptional activity. *PLoS ONE*, 10(2): 1–19.

Issaka, A., Paradies, Y. & Stevenson, C. 2018. Modifiable and emerging risk factors for type 2 diabetes in Africa: A systematic review and meta-analysis protocol. *Systematic Reviews*, 7: 139.

Ivády, G., Madar, L., Dzsudzsák, E., Koczok, K., Kappelmayer, J. & Krulisova, V. 2018. Analytical parameters and validation of homopolymer detection in a pyrosequencing-based next generation sequencing system. *BMC Genomics*, 19: 158.

Jaffe, A.E. & Irizarry, R.A. 2014. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biology*, 15(2): R31.

Jeong, H.M., Lee, S., Chae, H., Kim, R.N., Kwon, M.J., Oh, E., Choi, Y. La, Kim, S. & Shin, Y.K. 2016. Efficiency of methylated DNA immunoprecipitation bisulphite sequencing for whole-genome DNA methylation analysis. *Epigenomics*, 8(8): 1061–1077.

Jin, J., Wang, X., Zhi, X. & Meng, D. 2019. Epigenetic regulation in diabetic vascular complications. *Journal of Molecular Endocrinology*, 63(4): R103–R115.

Jin, Z. & Liu, Y. 2018. DNA methylation in human diseases. *Genes and Diseases*, 5: 1–8.

Johnson, R., Nxele, X., Cour, M., Sangweni, N., Jooste, T., Hadebe, N., Samodien, E., Benjeddou, M., Mazino, M., Louw, J. & Lecour, S. 2020. Identification of potential biomarkers for predicting the early onset of diabetic cardiomyopathy in a mouse model. *Scientific Reports*, 10(1): 1–12.

Jones, P.A. 2012. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, 13(7): 484–492.

Ju, J., Hirose, S., Shi, X.Y., Ishii, A., Hu, L.Y. & Zou, L.P. 2016. Treatment with Oral ATP decreases alternating hemiplegia of childhood with de novo ATP1A3 Mutation. *Orphanet Journal of Rare Diseases*, 11(55).

Jung, K.H., Torrone, D., Lovinsky-Desir, S., Perzanowski, M., Bautista, J., Jezioro, J.R., Hoepner, L., Ross, J., Perera, F.P., Chillrud, S.N. & Miller, R.L. 2017. Short-term exposure to $PM_{2.5}$ and vanadium and changes in asthma gene DNA methylation and lung function decrements among urban children. *Respiratory Research*, 18(63).

Kanehisa, M. & Goto, S. 2000. KEGG : Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1): 27–30.

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. 2019. New approach for understanding genome variations in KEGG. *Nucleic Acids Research*, 47: 590–595.

Kaneto, H., Katakami, N., Matsuhisa, M. & Matsuoka, T. 2010. Role of reactive oxygen species in the progression of type 2 diabetes and atherosclerosis. *Mediators*: 453892.

Kang, J., Lee, C.N., Li, H.Y., Hsu, K.H. & Lin, S.Y. 2017. Genome-wide DNA methylation variation in maternal and cord blood of gestational diabetes population. *Diabetes Research and Clinical Practice*, 132: 127–136.

Karachanak-Yankova, S., Dimova, R., Nikolova, D., Nesheva, D., Koprinarova, M., Maslyankov, S., Tafradjiska, R., Gateva, P., Velizarova, M., Hammoudeh, Z., Stoynev, N., Toncheva, D., Tankova, T. & Dimova, I. 2015. Epigenetic Alterations in patients with type 2 Diabetes mellitus. *Balkan Journal of Medical Genetics*, 18(2): 15–24.

Kearney, K., Tomlinson, D., Smith, K. & Ajjan, R. 2017. Hypofibrinolysis in diabetes : a therapeutic target for the reduction of cardiovascular risk. *Cardiovascular Diabetology*, 16: 34.

Kelly, T.K., De Carvalho, D.D. & Jones, P.A. 2010. Epigenetic modifications as therapeutic targets. *Nature Biotechnology.*, 28(10): 1069–1078.

Kengne, A.P., Echouffo-Tcheugui, J. B., Sobngwi, E. & Mbanya, J.C. 2013. New insights on diabetes mellitus and obesity in Africa - Part 1: prevalence, pathogenesis and comorbidities. *Heart*, 99(14): 1–5.

Keshavarzi, F. & Golsheh, S. 2019. IRS1- rs10498210 G/A and CCR5-59029 A/G polymorphisms in patients with type 2 diabetes in Kurdistan. *Molecular Genetics and Genomic Medicine*, 7(5): e631.

Khan, A.A., Hanada, T., Mohseni, M., Jeong, J., Zeng, L., Gaetani, M., Li, D., Reed, B.C., Speicher, D.W. & Chishti, A.H. 2008. Dematin and Adducin provide a novel link between the spectrin cytoskeleton and human erythrocyte membrane by directly interacting with Glucose Transporter-1. *The Journal of Biological Chemistry*, 283(21): 14600–14609.

Khan, M.A.B., Hashim, M.J., King, J.K., Govender, R.D., Mustafa, H. & Kaabi, J. Al. 2020. Epidemiology of Type 2 diabetes - Global burden of disease and forecasted trends. *Journal of Epidemiology and Global Health*, 10(1): 107–111.

Khan, R.M.M., Chua, Z.J.Y., Tan, J.C., Yang, Y., Liao, Z. & Zhao, Y. 2019. From pre-diabetes to diabetes: Diagnosis, treatments and translational research. *Medicina (Lithuania)*, 55(9): 546.

Kibirige, D., Lumu, W., Jones, A.G., Smeeth, L., Hattersley, A.T. & Nyirenda, M.J. 2019. Understanding the manifestation of diabetes in sub Saharan Africa to inform therapeutic approaches and preventive strategies : a narrative review. *Clinical Diabetes and Endocrinology*, 5: 2.

Kiese, K., Jablonski, J., Boison, D. & Kobow, K. 2016. Dynamic regulation of the adenosine kinase gene during early postnatal brain development and maturation. *Frontiers in Molecular Neuroscience*, 9(99): 1–11.

Kirilmaz, B., Asgun, F., Alioglu, E., Ercan, E., Tengiz, I., Turk, U., Saygi, S. & Özerkan, F. 2010. High inflammatory activity related to the number of metabolic syndrome components. *Journal of Clinical Hypertension*, 12(2): 136–144.

Kolb, H. & Martin, S. 2017. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. *BMC Medicine*, 15(1): 131.

Kong, Y. 2014. Distributions of positive signals in pyrosequencing. *Journal of Mathematical Biology*, 69(1): 39–54.

Kurdyukov, S. & Bullock, M. 2016. DNA methylation analysis: Choosing the right method. *Biology*, 5(1): 3.

Kuroda, A., Rauch, T.A., Todorov, I., Ku, H.T., Al-Abdullah, I.H., Kandeel, F., Mullen, Y., Pfeifer, G.P. & Ferreri, K. 2009. Insulin gene expression is regulated by DNA methylation. *PLoS ONE*, 4(9).

Kwak, S.H. & Park, K.S. 2016. Recent progress in genetic and epigenetic research on type 2 diabetes. *Experimental & Molecular Medicine*, 48(3): e220.

Lakshmaiah, K.C., Jacob, L.A., Aparna, S., Lokanatha, D. & Saldanha, S.C. 2014. Epigenetic therapy of cancer with histone deacetylase inhibitors. *Journal of Cancer Research and Therapeutics*, 10(3): 469–478.

Lauenborg, J., Jørgensen, M.K., Damm, P., Major-Pedersen, A., Eiberg, H., Urhammer, S., Pedersen, O. & Hansen, T. 2011. The influence of parental history of diabetes and offspring birthweight on offspring glucose metabolism in adulthood. *Acta Obstetricia et Gynecologica Scandinavica*, 90(12): 1357–1363.

Lebeche, D., Davidoff, A. & Hajjar, R. 2008. Interplay between impaired calcium regulation and insulin signaling abnormalities in diabetic cardiomyopathy. *Nature Clinical Practice Cardiovascular Medicine*, 5(11): 715–724.

Leighton, G. & Williams Jr, D.C. 2020. The methyl-CpG-binding domain 2 and 3 proteins and formation of the Nucleosome Remodeling and Deacetylase complex. *Journal of Molecular Biology*, 432(6): 1624–1639.

Leti, F. & DiStefano, J.K. 2017. Long noncoding RNAs as diagnostic and therapeutic targets in type 2 diabetes and related complications. *Genes*, 8(207).

Leti, F., Llaci, L., Malenica, I. & DiStefano, J. 2018. Methods for CpG methylation array profiling via bisulfite conversion. *Methods in Molecular Biology*, 1706: 233–254.

Li, B., Zhou, P., Xu, K., Chen, T., Jiao, J., Wei, H., Yang, X. & Xu, W. 2020. Metformin induces cell cycle arrest, apoptosis and autophagy through ROS / JNK signaling pathway in human osteosarcoma. *International Journal of Biological Sciences*, 16(1): 74–84.

Li, L., Choi, J.Y., Lee, K.M., Sung, H., Park, S.K., Oze, I., Pan, K.F., You, W.C., Chen, Y.X., Fang, J.Y., Matsuo, K., Kim, W.H., Yuasa, Y. & Kang, D. 2012. DNA methylation in peripheral blood: A potential biomarker for cancer molecular epidemiology. *Journal of Epidemiology*, 22(5): 384–394.

Li, M., Li, X., Zhang, H. & Lu, Y. 2018a. Molecular mechanisms of metformin for diabetes and cancer treatment. *Frontiers in Physiology*, 9: 1039.

Li, N.X., Brown, S., Kowalski, T., Wu, M., Yang, L., Dai, G., Petrov, A., Ding, Y., Dlugos, T., Wood, H.B., Wang, L., Erion, M., Sherwin, R. & Kelley, D.E. 2018b. GPR119 agonism increases glucagon secretion during insulin-induced hypoglycemia. *Diabetes*, 67(7): 1401–1413.

Li, R., Zhu, H. & Luo, Y. 2016a. Understanding the functions of long non-coding RNAs through their higher-order structures. *International Journal of Molecular Sciences*, 17: 702.

Li, W., Zhang, Xu, Lu, X., You, L., Song, Y., Luo, Z., Zhang, J., Nie, J., Zheng, W., Xu, D., Wang, Y., Dong, Y., Yu, S., Hong, J., Shi, J., Hao, H., Luo, F., Hua, L., Wang, P., Qian, X., Yuan, F., Wei, L., Cui, M., Zhang, T., Liao, Q., Dai, M., Liu, Z., Chen, G., Meckel, K., Adhikari, S., Jia, G., Bissonnette, M.B., Zhang, Xinxiang, Zhao, Y., Zhang, W., He, C. & Liu, J. 2017. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Research*, 27: 1243–1257.

Li, X., Kover, K.L., Heruth, D.P., Watkins, D.J., Moore, W. V, Jackson, K., Zang, M., Clements, M.A. & Yan, Y. 2015. New Insight Into Metformin Action : Regulation of ChREBP and FOXO1 Activities in Endothelial Cells. *Molecular Endocrinology*, 29(8): 1184–1194.

Li, X., Li, J., Wang, L., Li, A., Qiu, Z., Qi, L.W., Kou, J., Liu, K., Liu, B. & Huang, F. 2016b. The role of metformin and resveratrol in the prevention of hypoxia-inducible factor 1α accumulation and fibrosis in hypoxic adipose tissue. *British Journal of Pharmacology*, 173(12): 2001–2015.

Lin, X., Xu, Y., Pan, X., Xu, J., Ding, Y., Sun, X., Song, X., Ren, Y. & Shan, P.F. 2020. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Scientific Reports*, 10: 14790.

Lin, Y.H., Wu, M.H., Yeh, C.T. & Lin, K.H. 2018. Long non-coding RNAs as mediators of tumor microenvironment and liver cancer cell communication. *International Journal of Molecular Sciences*, 19: 3742.

Ling, C. & Groop, L. 2009. Epigenetics: A molecular link between environmental factors and type 2 diabetes. *Diabetes*, 58(12): 2718–2725.

Ling, C. & Rönn, T. 2019. Epigenetics in human obesity and type 2 diabetes. *Cell Metabolism*, 29(5): 1028–1044.

Ling, C., Del Guerra, S., Lupi, R., Rönn, T., Granhall, C., Luthman, H., Masiello, P., Marchetti, P., Groop, L. & Del Prato, S. 2008. Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. *Diabetologia*, 51(4): 615–622.

Ling, C., Poulsen, P., Simonsson, S., Rönn, T., Holmkvist, J., Almgren, P., Hagert, P., Nilsson, E., Mabey, A.G., Nilsson, P., Vaag, A. & Groop, L. 2007. Genetic and epigenetic factors are associated with expression of respiratory chain component NDUFB6 in human skeletal muscle. *Journal of Clinical Investigation*, 117(11): 3427–3435.

Liu, C., Marioni, R.E., Hedman, A.K., Pfeiffer, L., Tsai, P.C., Reynolds, L.M., Just, A.C., Duan, Q., Boer, C.G., Tanaka, T., Elks, C.E., Aslibekyan, S., Brody, J.A., Kühnel, B., Herder, C., Almli, L.M., Zhi, D., Wang, Y., Huan, T., Yao, C., Mendelson, M.M., Joehanes, R., Liang, L., Love, S.A., Guan, W., Shah, S., McRae, A.F., Kretschmer, A., Prokisch, H., Strauch, K., Peters, A., Visscher, P.M., Wray, N.R., Guo, X., Wiggins, K.L., Smith, A.K., Binder, E.B., Ressler, K.J., Irvin, M.R., Absher, D.M., Hernandez, D., Ferrucci, L., Bandinelli, S., Lohman, K., Ding, J., Trevisi, L., Gustafsson, S., Sandling, J.H., Stolk, L., Uitterlinden, A.G., Yet, I., Castillo-Fernandez, J.E., Spector, T.D., Schwartz, J.D., Vokonas, P., Lind, L., Li, Y., Fornage, M., Arnett, D.K., Wareham, N.J., Sotoodehnia, N., Ong, K.K., Van Meurs, J.B.J., Conneely, K.N., Baccarelli, A.A., Deary, I.J., Bell, J.T., North, K.E., Liu, Y., Waldenberger, M., London, S.J., Ingelsson, E. & Levy, D. 2018. A DNA methylation biomarker of alcohol consumption. *Molecular Psychiatry*, 23: 422–433.

Liu, Z., Tan, Q., Wen, J., Tang, Y., Huang, D., Huang, Y., Xie, J., Luo, Y., Liang, M., Wu, C., Lu, Z., Tan, A., Gao, YongWang, Q., Jiang, Y., Yao, Z., Lin, X., Zhang, H., Mo, Z. & Yang, X. 2016. Elevated serum complement factors 3 and 4 are strong inflammatory markers of the metabolic syndrome development : a longitudinal cohort study. *Nature Publishing Group*, 6: 18713.

Liyanage, V.R.B., Jarmasz, J.S., Murugeshan, N., Del Bigio, M.R., Rastegar, M. & Davie, J.R. 2014. DNA modifications: function and applications in normal and disease States. *Biology*, 3(4): 670–723.

Llácer, J.L., Hussain, T., Saini, A.K., Nanda, J.S., Kaur, S., Gordiyenko, Y., Kumar, R., Hinnebusch, A.G., Lorsch, J.R. & Ramakrishnan, V. 2018. Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition. *eLife*, 7: 1–33.

López, V., Fernández, A.F. & Fraga, M.F. 2017. The role of 5-hydroxymethylcytosine in development, aging and age-related diseases. *Ageing Research Reviews*, 37: 28–38.

Lu, Y., Hanada, T., Fujiwara, Y., Nwankwo, J.O., Wieschhaus, A.J., Hartwig, J., Huang, S., Han, J. & Chishti, A.H. 2016. Gene disruption of dematin causes precipitous loss of erythrocyte membrane stability and severe hemolytic anemia. *Blood*, 128(1): 93–103.

Luo, F., Das, A., Chen, J., Wu, P., Li, X. & Fang, Z. 2019. Metformin in patients with and without diabetes: A paradigm shift in cardiovascular disease management. *Cardiovascular Diabetology*, 18(1): 1–9.

Lyssenko, V. & Laakso, M. 2013. Genetic screening for the risk of type 2 diabetes: Worthless or valuable? *Diabetes Care*, 36(SUPPL.2): S120–S126.

Ma, J. & Hart, G.W. 2013. Protein O-GlcNAcylation in diabetes and diabetic complications. *Expert Review of Proteomics*, 10(4): 365–380.

Mahna, D., Puri, S. & Sharma, S. 2018. DNA methylation signatures: Biomarkers of drug and alcohol abuse. *Mutation Research - Reviews in Mutation Research*, 777: 19–28.

Maiuolo, J., Oppedisano, F., Gratteri, S., Muscoli, C. & Mollace, V. 2016. Regulation of uric acid metabolism and excretion. *International Journal of Cardiology*, 213: 8–14.

Malik, B. & Feng, F. 2016. Long noncoding RNAs in prostate cancer: Overview and clinical implications. *Asian Journal of Andrology*, 18(4): 568–574.

Malik, F., Mehdi, S.F., Ali, H., Patel, P., Basharat, A., Kumar, A., Ashok, F., Stein, J., Brima, W., Malhotra, P. & Roth, J. 2018. Is metformin poised for a second career as an antimicrobial? *Diabetes Metabolism Research and Reviews*, 34(4): e2975.

Mapa-Tassou, C., Katte, J.C., Mba Maadjhou, C. & Mbanya, J.C. 2019. Economic impact of diabetes in Africa. *Current Diabetes Reports*, 19(5).

Marín-Peñalver, J.J., Martín-Timón, I., Sevillano-Collantes, C. & Cañizo-Gómez, F.J. del. 2016. Update on the treatment of type 2 diabetes mellitus. *World Journal of Diabetes*, 7(17): 354.

Martin, E.M. & Fry, R.C. 2016. A cross-study analysis of prenatal exposures to environmental contaminants and the epigenome: support for stress-responsive transcription factor occupancy as a mediator of gene-specific CpG methylation patterning. *Environmental Epigenetics*, 2(1): 1-9.

Martin, E.M. & Fry, R.C. 2018. Environmental influences on the epigenome: Exposure-associated DNA methylation in human populations. *Annual Review of Public Health*, 39: 309–333.

Martínez, J.A., Milagro, F.I., Claycombe, K.J. & Schalinske, K.L. 2014. Epigenetics in adipose tissue, obesity, weight loss, and diabetes. *Advances in Nutrition (Bethesda, Md.), 5 (1)*: 71–81.

Martire, S. & Banaszynski, L.A. 2020. The roles of histone variants in fine-tuning chromatin organization and function. *Nature Reviews Molecular Cell Biology*, 21: 522–541.

Maruthur, N.M., Gribble, M.O., Bennett, W.L., Bolen, S., Wilson, L.M., Balakrishnan, P., Sahu, A., Bass, E., Kao, W.H.L. & Clark, J.M. 2014. The pharmacogenetics of Type 2 Diabetes: A systematic review. *Diabetes Care*, 37(3): 876–886.

Mason, M.J., Speake, C., Gersuk, V.H., Nguyen, Q., Odegard, J.M., Buckner, J.H., Greenbaum, C.J., Chaussabel, D. & Nepom, G.T. 2014. Low HERV-K (C4) copy number is associated with type 1 diabetes. *Diabetes*, 63: 1789–1795.

Matsha, T.E., Hassan, M.S., Kidd, M. & Erasmus, R.T. 2012. The 30-year cardiovascular risk profile of South Africans with diagnosed diabetes, undiagnosed diabetes, pre-diabetes or normoglycaemia: The Bellville, South Africa pilot study. *Cardiovascular Journal of Africa*, 23(1): 5–11.

Matsha, T.E., Pheiffer, C., Humphries, S.E., Gamieldien, J., Erasmus, R.T. & Kengne, A.P. 2016a. Genome-wide DNA methylation in mixed ancestry individuals with diabetes and prediabetes from South Africa. *International Journal of Endocrinology*: 3172093.

Matsha, T.E., Pheiffer, C., Mutize, T., Erasmus, R.T. & Kengne, A.P. 2016b. Glucose tolerance, MTHFR C677T and NOS3 G894T polymorphisms, and global DNA methylation in mixed ancestry African individuals. *Journal of Diabetes Research*, 8738072.

Matsha, T.E., Soita, D.J., Hassan, M.S., Hon, G.M., Yako, Y.Y., Kengne, A.P. & Erasmus, R.T. 2013. Three-year's changes in glucose tolerance status in the Bellville South cohort: rates and phenotypes associated with progression. *Diabetes Research and Clinical Practice*, 99(2): 223–30.

Mbanya, J.C.N., Motala, A.A., Sobngwi, E., Assah, F.K. & Enoru, S.T. 2010. Diabetes in sub-Saharan Africa. *The Lancet*, 375(9733): 2254–2266.

McKay, J.A. & Mathers, J.C. 2016. Maternal folate deficiency and metabolic dysfunction in offspring. *Proceedings of the Nutrition Society*, 75(1): 90–95.

Meigs, J.B., Cupples, L.A. & Wilson, P.W.F. 2000. Parental transmission of type 2 diabetes: The Framingham offspring study. *Diabetes*, 49(12): 2201–2207.

Meltzer, M.E., Doggen, C.J.M., Groot, P.G. De, Rosendaal, F.R. & Lisman, T. 2010. Plasma levels of fibrinolytic proteins and the risk of myocardial infarction in men. *Blood*, 116(4): 529–536.

Menafra, R. & Stunnenberg, H.G. 2014. MBD2 and MBD3: elusive functions and mechanisms. *Frontiers in Genetics*, 5: 428.

Miller, J.R., Delcher, A.L., Koren, S., Venter, E., Walenz, B.P., Brownley, A., Johnson, J., Li, K., Mobarry, C. & Sutton, G. 2008. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics*, 24(24): 2818–2824.

Miranda, T.B. & Jones, P.A. 2007. DNA Methylation: The nuts and bolts of repression. *Journal Cellular Physiology*, 213: 384–390.

Miranda-Massari, J.R., Gonzalez, M.J., Jimenez, F.J., Allende-Vigo, M.J. & Duconge, J. 2011. Metabolic correction in the management of Diabetic peripheral neuropathy: Improving clinical results beyond symptom control. *Current Clinical Pharmacology*, 6(4): 260–273.

Mirtschink, P. & Chavakis, T. 2018. The missed notch to bring down diabetes. *Trends in Endocrinology and Metabolism*, 29(7): 448–450.

Misra, B.B. & Misra, A. 2020. The chemical exposome of type 2 diabetes mellitus: Opportunities and challenges in the omics era. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 14(1): 23–38.

Mohseni, M. & Chishti, A.H. 2008. The headpiece domain of dematin regulates cell shape, motility, and wound healing by modulating RhoA activation. *Molecular and Cellular Biology*, 28(15): 4712–4718.

Montgomery, M.K. & Turner, N. 2015. Mitochondrial dysfunction and insulin resistance: An update. *Endocrine Connections*, 4(1): R1-R15.

Montgomery, M.K. 2019. Mitochondrial dysfunction and diabetes: Is mitochondrial transfer a friend or foe? *Biology*, 8(2): 33.

Moore, L.D., Le, T. & Fan, G. 2012. DNA methylation and its basic function. *Neuropsychopharmacology*, 38(1): 23–38.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M.J., Patterson, N., Mesirov, J.P., Golub, T.R., Tamayo, P., Spiegelman, B., Lander, E.S., Hirschhorn, J.N., Altshuler, D. & Groop, L.C. 2003. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics*, 34(3): 267–273.

Moran, B.M., Flatt, P.R. & McKillop, A.M. 2016. G protein-coupled receptors: signalling and regulation by lipid agonists for improved glucose homoeostasis. *Acta Diabetologica*, 53(2): 177–188.

Moreno-Viedma, V., Amor, M., Sarabi, A., Bilban, M., Staffler, G., Zeyda, M. & Stulnig, T.M. 2016. Common dysregulated pathways in obese adipose tissue and atherosclerosis. *Cardiovascular Diabetology*, 15(1): 1–12.

Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segrè, A. V, Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., Prokopenko, I., Kang, H.M., Dina, C., Esko, T., Fraser, R.M., Kanoni, S., Kumar, A., Lagou, V., Langenberg, C., Luan, J., Lindgren, C.M., Müller-Nurasyid, M., Pechlivanis, S., Rayner, W., Scott, L.J., Wiltshire, S., Yengo, L., Kinnunen, L., Rossin, E.J., Raychaudhuri, S., Johnson, A.D., Dimas, A.S., Loos, R.J.F., Vedantam, S., Chen, H., Florez, J.C., Caroline Fox, Liu, C.-T., Rybin, D., Couper, D.J., Kao, W.H.L., Li, M., Cornelis, M.C., Kraft, P., Sun, Q., Dam, R.M. van, Stringham, H.M., Chines, P.S., Fischer, K., Fontanillas, P., Holmen, O.L., Hunt, S.E., Jackson, A.U., Kong, A., Lawrence, R., Meyer, J., Perry, J.R., Platou, C.G., Potter, S., Rehnberg, E., Robertson, N., Sivapalaratnam, S., Stančáková, A., Stirrups, K., Thorleifsson, G., Tikkanen, E., Wood, A.R., Almgren, P., Atalay, M., Benediktsson, R., Bonnycastle, L.L., Burtt, N., Carey, J., Charpentier, G., Crenshaw, A.T., Doney, A.S.F., Dorkhan, M., Edkins, S., Emilsson, V., Eury, E., Forsen, T., Gertow, K., Gigante, B., Grant, G.B., Groves, C.J., Guiducci, C., Herder, C., Hreidarsson, A.B., Hui, J., James, A., Groop, L.C., Stefansson, K., Hu, F., Pankow, J.S., Dupuis, J., Meigs, J.B., Altshuler, D., Boehnke, M. & McCarthy, M.I. 2012. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature Genetics*, 44(9): 981–990.

Muka, T., Nano, J., Voortman, T., Braun, K.V.E., Ligthart, S., Stranges, S., Bramer, W.M., Troup, J., Chowdhury, R., Dehghan, A. & Franco, O.H. 2016. The role of global and regional DNA methylation and histone modifications in glycemic traits and type 2 diabetes: A systematic review. *Nutrition, Metabolism and Cardiovascular Diseases*, 26(7): 553–566.

Mukherjee, K., Twyman, R.M. & Vilcinskas, A. 2015. Insects as models to study the epigenetic basis of disease. *Progress in Biophysics and Molecular Biology*, 118(1–2): 69–78.

Mutyambizi, C., Booysen, F., Stokes, A., Pavlova, M. & Groot, W. 2019. Lifestyle and socio-economic inequalities in diabetes prevalence in South Africa: A decomposition analysis. *PLoS ONE*, 14(1): 1–21.

Myslicki, J.P., Shearer, J., Hittel, D.S., Hughey, C.C. & Belke, D.D. 2014. O-GlcNAc modification is associated with insulin sensitivity in the whole blood of healthy young adult males. *Diabetology and Metabolic Syndrome*, 6(1): 96.

Nanda, J.S., Saini, A.K., Muñoz, A.M., Hinnebusch, A.G. & Lorsch, J.R. 2013. Coordinated movements of eukaryotic translation initiation factors eIF1, eIF1A, and eIF5 trigger phosphate release from eIF2 in response to start codon recognition by the ribosomal preinitiation complex. *Journal of Biological Chemistry*, 288(8): 5316–5329.

Ni, Q., Gu, Y., Xie, Y., Yin, Q., Zhang, H., Nie, A., Li, W., Wang, Y., Ning, G., Wang, W. & Wang, Q. 2017. Raptor regulates functional maturation of murine beta cells. *Nature Communications*, 8: 15755.

Nilsson, B., Hamad, O.A., Ahlstrom, H., Kullberg, J., Johansson, L., Lindhagen, L., Haenni, A., Ekdahl, K.N. & Lind, L. 2014a. C3 and C4 are strongly related to adipose tissue variables and cardiovascular risk factors. *European Journal of Clinical Investigation*, 44(6): 587–597.

Nilsson, E., Jansson, P.A., Perfilyev, A., Volkov, P., Pedersen, M., Svensson, M.K., Poulsen, P., Ribel-Madsen, R., Pedersen, N.L., Almgren, P., Fadista, J., Rönn, T., Pedersen, B.K., Scheele, C., Vaag, A. & Ling, C. 2014b. Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes*, 63(9): 2962–2976.

O ’Connell, T.M. & Markunas, C.A. 2016. DNA methylation and microRNA-based biomarkers for risk of type 2 diabetes. *Current Diabetes Reviews, 12(1)*: 20–29.

Olivieri, F., Rippo, M.R., Procopio, A.D. & Fazioli, F. 2013. Circulating inflamma-miRs in aging and age-related diseases. *Frontiers in Genetics*, 4: 121.

Olkhov-Mitsel, E. & Bapat, B. 2012. Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers. *Cancer Medicine*, 1(2): 237–260.

Olova, N., Krueger, F., Andrews, S., Oxley, D., Berrens, R. V., Branco, M.R. & Reik, W. 2018. Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data. *Genome Biology*, 19(33).

Omar, M. A., Seedat, M. A., Motala, A. A., Dyer, R. B. and Becker, P. 1993. The prevalence of diabetes mellitus and impaired glucose tolerance in a group of urban South African blacks. *South African Medical Journal*, 83 (9): 641-643.

Ongwijitwat, S., Liang, H.L., Graboyes, E.M. & Wong-Riley, M.T.T. 2006. Nuclear respiratory factor 2 senses changing cellular energy demands and its silencing down-regulates cytochrome oxidase and other target gene mRNAs. *Gene*, 374: 39–49.

Ortega, A., Tarazón, E., Roselló-Lletí, E., Gil-Cayuela, C., Lago, F., González-Juanatey, J., Cinca, J., Jorge, E., Martínez-Dolz, L., Portolés, M. & Rivera, M. 2015. Patients with dilated cardiomyopathy and sustained monomorphic ventricular tachycardia show up-regulation of KCNN3 and KCNJ2 genes and CACNG8-linked left ventricular dysfunction. *PLoS ONE*, 10(12): e0145518.

Ortiz, R., Joseph, J.J., Lee, R., Wand, G.S. & Golden, S.H. 2018. Type 2 diabetes and cardiometabolic risk may be associated with increase in DNA methylation of FKBP5. *Clinical Epigenetics*, 10: 82.

Osanai, T., Tanaka, M., Magota, K., Tomita, H. & Okumura, K. 2012. Coupling factor 6-induced activation of ecto-F_1F_0 complex induces insulin resistance, mild glucose intolerance and elevated blood pressure in mice. *Diabetologia*, 55: 520–529.

Palmieri, F. & Monné, M. 2016. Discoveries, metabolic roles and diseases of mitochondrial carriers: A review. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1863(10): 2362–2378.

Palmieri, F. 2013. The mitochondrial transporter family SLC25: Identification, properties and physiopathology. *Molecular Aspects of Medicine*, 34(2–3): 465–484.

Palmieri, F., Scarcia, P. & Monné, M. 2020. Diseases caused by mutations in mitochondrial carrier genes SLC25: A review. *Biomolecules*, 10(4): 1–32.

Pan, Y., Liu, G., Zhou, F., Su, B. & Li, Y. 2018. DNA methylation profiles in cancer diagnosis and therapeutics. *Clinical and Experimental Medicine*, 18(1): 1–14.

Pantalone, K.M., Hobbs, T.M., Wells, B.J., Kong, S.X., Kattan, M.W., Bouchard, J., Yu, C., Sakurada, B., Milinovich, A., Weng, W., Bauman, J.M. & Zimmerman, R.S. 2015. Clinical characteristics, complications, comorbidities and treatment patterns among patients with type 2 diabetes mellitus in a large integrated health system. *BMJ Open Diabetes Research and Care*, 3: e000093.

Papandreou, C., Li, J., Liang, L., Bulló, M., Zheng, Y., Ruiz-Canela, M., Yu, E., Guasch-Ferré, M., Razquin, C., Clish, C., Corella, D., Estruch, R., Ros, E., Fitó, M., Arós, F., Serra-Majem, L., Rosique, N., Martínez-González, M.A., Hu, F.B. & Salas-Salvadó, J. 2019. Metabolites related to purine catabolism and risk of type 2 diabetes incidence; modifying effects of the TCF7L2-rs7903146 polymorphism. *Scientific Reports*, 9(2892).

Park, J., Guan, Y., Sheng, X., Gluck, C., Seasock, M.J., Ari Hakimi, A., Qiu, C., Pullman, J., Verma, A., Li, H., Palmer, M. & Susztak, K. 2019. Functional methylome analysis of human diabetic kidney disease. *JCI Insight*, 4(11): e128886.

Park, K., Saudek, C.D. & Hart, G.W. 2010. Increased expression of beta- N-acetylglucosaminidase in erythrocytes from individuals with pre-diabetes and diabetes. , 59(7): 1845–1850.

Park, K.S. 2011. The search for genetic risk factors of type 2 diabetes mellitus. *Diabetes and Metabolism Journal*, 35(1): 12–22.

Pasquier, J., Hoarau-Vechot, J., Fakhro, K., Rafii, A. & Abi Khalil, C. 2015. Epigenetics and cardiovascular disease in diabetes. *Current Diabetes Reports*, 15(108): 1–12.

Pastor-Anglada, M. & Pérez-Torras, S. 2015. Nucleoside transporter proteins as biomarkers of drug responsiveness and drug targets. *Frontiers in Pharmacology*, 6: 13.

Pastor-Anglada, M., Cano-Soldado, P., Errasti-Murugarren, E. & Casado, F.J. 2008. SLC28 genes and concentrative nucleoside transporter (CNT) proteins. *Xenobiotica*, 38(7–8): 972–994.

Patel, S., Srivastava, S., Singh, M.R. & Singh, D. 2019. Mechanistic insight into diabetic wounds: Pathogenesis, molecular targets and treatment strategies to pace wound healing. *Biomedicine and Pharmacotherapy*, 112(October 2018): 108615.

Patterson, K., Molloy, L., Qu, W. & Clark, S. 2011. DNA methylation: Bisulphite modification and analysis. *Journal of Visualized Experiments*, 56(e3170): 1–9.

Patterson, N., Petersen, D.C., van der Ross, R.E., Sudoyo, H., Glashoff, R.H., Marzuki, S., Reich, D. & Hayes, V.M. 2009. Genetic structure of a unique admixed population: Implications for medical research. *Human Molecular Genetics*, 19(3): 411–419.

Peer, N., Steyn, K., Lombard, C., Lambert, E. V., Vythilingum, B. & Levitt, N.S. 2012. Rising diabetes prevalence among urban-dwelling black South Africans. *PLoS ONE*, 7(9): e43336.

Pengyu, Z., Yan, Y., Xiying, F., Maoguang, Y., Mo, L., Yan, C., Hong, S., Lijuan, W., Xiujuan, Z. & Hanqing, C. 2020. The differential expression of long non-coding RNAs in type 2 Diabetes mellitus and Latent autoimmune diabetes in adults. *International Journal of Endocrinology*, 2020: 1–12.

Perera, H.K.I. & Handuwalage, C.S. 2015. Analysis of glycation induced protein cross-linking inhibitory effects of some antidiabetic plants and spices. *BMC Complementary and Alternative Medicine*, 15(1): 1–9.

Pernicova, I. & Korbonits, M. 2014. Metformin-mode of action and clinical implications for diabetes and cancer. *Nature Reviews Endocrinology*, 10(3): 143–156.

Pheiffer, C., Erasmus, R.T., Kengne, A.P. & Matsha, T.E. 2016. Differential DNA methylation of microRNAs within promoters, intergenic and intragenic regions of type 2 diabetic, pre-diabetic and non-diabetic individuals. *Clinical Biochemistry*, 49(6): 433–438.

Philip, W., James, T., Jackson-Leach, R., Mhurchu, C.N., Kalamara, E., Shayeghi, M., Rigby, N.J., Nishida, C. & Rodgers, A. 2004. Overweight and obesity (high body mass index). In *Comparative Quantification of Health Risks. Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Geneva: World Health Organisation.

Pillay-van Wyk, V., Msemburi, W., Laubscher, R., Dorrington, R.E., Groenewald, P., Glass, T., Nojilana, B., Joubert, J.D., Matzopoulos, R., Prinsloo, M., Nannan, N., Gwebushe, N., Vos, T., Somdyala, N., Sithole, N., Neethling, I., Nicol, E., Rossouw, A. & Bradshaw, D. 2016. Mortality trends and differentials in South Africa from 1997 to 2012: second National Burden of Disease Study. *The Lancet Global Health*, 4(9): e642–e653.

Plows, J.F., Stanley, J.L., Baker, P.N., Reynolds, C.M. & Vickers, M.H. 2018. The pathophysiology of gestational diabetes mellitus. *International Journal of Molecular Sciences*, 19(11): 1–21.

Plusquin, M., Guida, F., Polidoro, S., Vermeulen, R., Raaschou-Nielsen, O., Campanella, G., Hoek, G., Kyrtopoulos, S.A., Georgiadis, P., Naccarati, A., Sacerdote, C., Krogh, V., Bas Bueno-de-Mesquita, H., Monique Verschuren, W.M., Sayols-Baixeras, S., Panni, T., Peters, A., Hebels, D.G.A.J., Kleinjans, J., Vineis, P. & Chadeau-Hyam, M. 2017. DNA methylation and exposure to ambient air pollution in two prospective cohorts. *Environment International*, 108: 127–136.

Pompura, S.L. & Dominguez-Villar, M. 2018. The PI3K/AKT signaling pathway in regulatory T-cell development, stability, and function. *Journal of Leukocyte Biology*, 103(6): 1065–1076.

Ponnaluri, V.K.C., Ehrlich, K.C., Zhang, G., Lacey, M., Johnston, D., Pradhan, S. & Ehrlich, M. 2017. Association of 5-hydroxymethylation and 5-methylation of DNA cytosine with tissue-specific gene expression. *Epigenetics*, 12(2): 123–138.

Pons, D., De Vries, F.R., Van Den Elsen, P.J., Heijmans, B.T., Quax, P.H.A. & Jukema, J.W. 2009. Epigenetic histone acetylation modifiers in vascular remodelling: New targets for therapy in cardiovascular disease. *European Heart Journal*, 30(3): 266–277.

Pop-Busui, R., Lu, J., Brooks, M.M., Albert, S., Althouse, A.D., Escobedo, J., Green, J., Palumbo, P., Perkins, B.A., Whitehouse, F. & Jones, T.L.Z. 2013. Impact of glycemic control strategies ontheprogressionofdiabeticperipheral neuropathy in the bypass angioplasty revascularization investigation 2 diabetes (BARI 2D) Cohort. *Diabetes Care*, 36(10): 3208–3215.

Porta, M., Curletto, G., Cipullo, D., De la Longrais, R.R., Trento, M., Passera, P., Taulaigo, A.V., Di Miceli, S., Cenci, A., Dalmasso, P. & Cavallo, F. 2014. Estimating the delay between onset and oiagnosis of type 2 diabetes from the time course of retinopathy prevalence. *Diabetes Care*, 37: 1668–1674.

Pozzilli, P. & Pieralice, S. 2018. Latent autoimmune diabetes in adults: Current status and new horizons. *Endocrinology and Metabolism*, 33(2): 147–159.

Prasad, R.B. & Groop, L. 2015. Genetics of type 2 diabetes - pitfalls and possibilities. *Genes*, 6: 87–123.

Priya, G. & Kalra, S. 2018. Metformin in the management of diabetes during pregnancy and lactation. *Drugs in Context*, 7: 1–21.

Raciti, G.A., Longo, M., Parrillo, L., Ciccarelli, M., Mirra, P., Ungaro, P., Formisano, P., Miele, C. & Béguinot, F. 2015. Understanding type 2 diabetes: from genetics to epigenetics. *Acta Diabetologica*, 52(5): 821–827.

Raine, A., Liljedahl, U. & Nordlund, J. 2018. Data quality of whole genome bisulfite sequencing on Illumina platforms. *PLoS ONE*, 13(4): e0195972.

Rakyan, V.K., Down, T.A., Balding, D.J. & Beck, S. 2012. Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics.*, 12(8): 529–541.

Ramlo-Halsted, B. A. & Edelman, S. V. 1999. The natural history of type 2 diabetes. Implications for clinical practice. *Primary Care Diabetes*, 26(4): 771–789.

Ramos, E., Chen, G., Shriner, D., Doumatey, A., Gerry, N.P., Herbert, A., Huang, H., Zhou, J., Christman, M.F., Adeyemo, A. & Rotimi, C. 2011. Replication of genome-wide association studies (GWAS) loci for fasting plasma glucose in African-Americans. *Diabetologia*, 54(4): 783–788.

Ran, H., Lu, Y., Zhang, Q., Hu, Q., Zhao, J., Wang, K., Tong, X. & Su, Q. 2020. MondoA is required for normal myogenesis and regulation of the skeletal muscle glycogen content in mice. *Diabetes and Metabolism Journal*, 10.4093/dmj.2019.0212. Advance online publication.

Rao, S., Chiu, T.P., Kribelbauer, J.F., Mann, R.S., Bussemaker, H.J. & Rohs, R. 2018. Systematic prediction of DNA shape changes due to CpG methylation explains epigenetic effects on protein-DNA binding. *Epigenetics and Chromatin*, 11(1): 1–11.

Rauluseviciute, I., Drabløs, F. & Rye, M.B. 2019. DNA methylation data by sequencing: Experimental approaches and recommendations for tools and pipelines for data analysis. *Clinical Epigenetics*, 11(1): 1–13.

Reddy, M.A. & Natarajan, R. 2015. Recent developments in epigenetics of acute and chronic kidney diseases. *Kidney International*, 88(2): 250–261.

Reed, K., Poulin, M.L., Yan, L. & Parissenti, A.M. 2010. Comparison of bisulfite sequencing PCR with pyrosequencing for measuring differences in DNA methylation. *Analytical Biochemistry*, 397(1): 96–106.

Rewers, M. & Ludvigsson, J. 2016. Environmental risk factors for type 1 diabetes. *Lancet.*, 387(10035): 2340–2348.

Ribel-Madsen, R., Fraga, M.F., Jacobsen, S., Bork-Jensen, J., Lara, E., Calvanese, V., Fernandez, A.F., Friedrichsen, M., Vind, B.F., Højlund, K., Beck-Nielsen, H., Esteller, M., Vaag, A. & Poulsen, P. 2012. Genome-wide analysis of DNA methylation differences in muscle and fat from monozygotic twins discordant for type 2 diabetes. *PLoS ONE*, 7(12): e51302.

Richards, P., Ourabaha, S., Montagne, J., Burnol, A., Postic, C. & Guilmeau, S. 2017. MondoA/ChREBP: The usual suspects of transcriptional glucose sensing; Implication in pathophysiology. *Metabolism*, 70: 133–151.

Richards, P., Rachdi, L., Oshima, M., Marchetti, P., Bugliani, M., Armanet, M., Postic, C., Guilmeau, S. & Scharfmann, R. 2018. MondoA is an essential glucose-responsive transcription factor in human pancreatic β-cells. *Diabetes*, 67(3): 461–472.

Rinn, J.L. & Chang, H.Y. 2012. Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry*, 81: 145–166.

Rives, M.L., Javitch, J.A. & Wickenden, A.D. 2017. Potentiating SLC transporter activity: Emerging drug discovery opportunities. *Biochemical Pharmacology*, 135: 1–11.

Rodriguez, C. & Martinez-Gonzalez, J. 2019. The role of lysyl oxidase enzymes in cardiac function and remodeling. *Cells*, 8(12): 1483.

Rodríguez-Mulero, S., Errasti-Murugarren, E., Ballarín, J., Felipe, A., Doucet, A., Casado, F.J. & Pastor-Anglada, M. 2005. Expression of concentrative nucleoside transporters SLC28

(CNT1, CNT2, and CNT3) along the rat nephron: Effect of diabetes. *Kidney International*, 68(2): 665–672.

Ronn, T. & Ling, C. 2015. DNA methylation as a diagnostic and therapeutic target in the battle against Type 2 diabetes. *Epigenomics*, 7(3): 451–460.

Rönn, T., Poulsen, P., Hansson, O., Holmkvist, J., Almgren, P., Nilsson, P., Tuomi, T., Isomaa, B., Groop, L., Vaag, A. & Ling, C. 2008. Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle. *Diabetologia*, 51: 1159–1168.

Rotroff, D.M., Joubert, B.R., Marvel, S.W., Håberg, S.E., Wu, M.C., Nilsen, R.M., Ueland, P.M., Nystad, W., London, S.J. & Motsinger-Reif, A. 2016. Maternal smoking impacts key biological pathways in newborns through epigenetic modification *in Utero*. *BMC Genomics*, 17(976).

Roux, P.P. & Topisirovic, I. 2018. Signaling pathways involved in the regulation of mRNA translation. *Molecular and Cellular Biology*, 38(12): 1–26.

Rovira-Llopis, S., Bañuls, C., Diaz-Morales, N., Hernandez-Mijares, A., Rocha, M. & Victor, V.M. 2017. Mitochondrial dynamics in type 2 diabetes: Pathophysiological implications. *Redox Biology*, 11: 637–645.

Russo, G.T., Giandalia, A., Romeo, E.L., Scarcella, C., Gambadoro, N., Zingale, R., Forte, F., Perdichizzi, G., Alibrandi, A. & Cucinotta, D. 2016. Diabetic neuropathy is not associated with homocysteine, folate, vitamin B12 levels, and MTHFR C677T mutation in type 2 diabetic outpatients taking metformin. *Journal of Endocrinological Investigation*, 39(3): 305–314.

Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., Colagiuri, S., Guariguata, L., Motala, A.A., Ogurtsova, K., Shaw, J.E., Bright, D. & Williams, R. 2019. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice*, 157: 107843.

Saini, N. & Yang, X. 2018. Metformin as an anti-cancer agent: Actions and mechanisms targeting cancer stem cells. *Acta Biochimica et Biophysica Sinica*, 50(2): 133–143.

Sakura, T., Morioka, T., Shioi, A., Kakutani, Y., Miki, Y., Yamazaki, Y., Motoyama, K., Mori, K., Fukumoto, S., Shoji, T., Emoto, M. & Inaba, M. 2017. Lipopolysaccharide-binding protein is associated with arterial stiffness in patients with type 2 diabetes: A cross-sectional study. *Cardiovascular Diabetology*, 16(1): 62.

Samblas, M., Milagro, F.I. & Martínez, A. 2019. DNA methylation markers in obesity, metabolic syndrome, and weight loss. *Epigenetics*, 14(5): 421–444.

Sathishkumar, C., Prabu, P., Mohan, V. & Balasubramanyam, M. 2018. Linking a role of lncRNAs (long non-coding RNAs) with insulin resistance, accelerated senescence, and inflammation in patients with type 2 diabetes. *Human Genomics*, 12(1): 1–9.

Sattar, N. & Gill, J.M.R. 2015. Type 2 diabetes in migrant south Asians: Mechanisms, mitigation, and management. *The Lancet Diabetes and Endocrinology*, 3(12): 1004–1016.

Schmit, K. & Michiels, C. 2018. TMEM proteins in cancer: A review. *Frontiers in Pharmacology*, 9: 1345.

Schultze, S.M., Hemmings, B.A., Niessen, M. & Tschopp, O. 2012. PI3K/AKT, MAPK and AMPK signalling: Protein kinases in glucose homeostasis. *Expert Reviews in Molecular Medicine*, 14: e1.

Schumann, T., König, J., Henke, C., Willmes, D.M., Bornstein, S.R., Jordan, J., Fromm, M.F. & Birkenfeld, A.L. 2020. Solute carrier transporters as potential targets for the treatment of metabolic disease. *Pharmacological Reviews*, 72(1): 343–379.

Sergi, D., Naumovski, N., Heilbronn, L.K., Abeywardena, M., O'Callaghan, N., Lionetti, L. & Luscombe-Marsh, N. 2019. Mitochondrial (dys)function and insulin resistance: From pathophysiological molecular mechanisms to the impact of diet. *Frontiers in Physiology*, 10: 532.

Šestáková, Š., Šálek, C. & Remešová, H. 2019. DNA methylation validation methods: A coherent review with practical comparison. *Biological Procedures Online*, 21(19).

Sharma, D.K., Bressler, K., Patel, H., Balasingam, N. & Thakor, N. 2016. Role of eukaryotic initiation factors during cellular stress and cancer progression. *Journal of Nucleic Acids*: 8235121.

Shen, J., Kondal, D., Rubinstein, A., Irazola, V., Gutierrez, L., Miranda, J.J., Bernabé-Ortiz, A., Lazo-Porras, M., Levitt, N., Steyn, K., Bobrow, K., Ali, M.K., Prabhakaran, D. & Tandon, N. 2016. A multiethnic study of pre-diabetes and diabetes in LMIC. *Global Heart*, 11(1): 61–70.

Shen, T.C., Tsai, C.W., Chang, W.S., Wang, S., Chao, C.Y., Hsiao, C.L., Chen, W.C., Hsia, T.C. & Bau, D.T. 2017. Association of interleukin-12A rs568408 with susceptibility to asthma in Taiwan. *Scientific Reports*, 7(1): 1–7.

Shim, K., Begum, R., Yang, C. & Wang, H. 2020. Complement activation in obesity, insulin resistance, and type 2 diabetes mellitus. *World Journal of Diabetes*, 11(1): 1–12.

Shrivastava, S.R. & Ghorpade, A.G. 2014. High prevalence of type 2 diabetes melitus and its risk factors among the rural population of Pondicherry, South India. *Journal of Research in Health Sciences*, 14(4): 258–263.

Singh, R.M., Waqar, T., Howarth, F.C., Adeghate, E., Bidasee, K. & Singh, J. 2018. Hyperglycemia-induced cardiac contractile dysfunction in the diabetic heart. *Heart Failure Reviews*, 23(1): 37–54.

Siqueira, J.F., Fouad, A.F. & Rôças, I.N. 2012. Pyrosequencing as a tool for better understanding of human microbiomes. *Journal of Oral Microbiology*, 4: 10.3402.

Skyler, J.S., Bakris, G.L., Bonifacio, E., Darsow, T., Eckel, R.H., Groop, L., Groop, P.H., Handelsman, Y., Insel, R.A., Mathieu, C., McElvaine, A.T., Palmer, J.P., Pugliese, A., Schatz, D.A., Sosenko, J.M., Wilding, J.P.H. & Ratner, R.E. 2017. Differentiation of diabetes by pathophysiology, natural history, and prognosis. *Diabetes*, 66(2): 241–255.

Soliman, G.A., Steenson, S.M. & Etekpo, A.H. 2016. Effects of metformin and a mammalian target of rapamycin (mTOR) ATP-competitive inhibitor on targeted metabolomics in pancreatic cancer cell line. *Metabolomics (Los Angel)*, 6(3): 183.

Solinas, G. & Becattini, B. 2017. JNK at the crossroad of obesity, insulin resistance, and cell stress response. *Molecular Metabolism*, 6(2): 174–184.

Song, X., Zhang, Y., Zhong, Q., Zhan, K., Bi, J., Tang, J., Xie, J. & Li, B. 2020. Genomics identification and functional characterization of methyl-CpG binding domain protein from Tribolium castaneum. *Genomics*, 112(3): 2223–2232.

Song, Z., Yang, H., Zhou, L. & Yang, F. 2019. Glucose-sensing transcription factor MondoA / ChREBP as targets for type 2 diabetes : Opportunities and challenges. *International Journal of Molecular Sciences*, 20(20): 5132.

Soozangar, N., Sadeghi, M.R., Jeddi, F., Somi, M.H., Shirmohamadi, M. & Samadi, N. 2018. Comparison of genome-wide analysis techniques to DNA methylation analysis in human cancer. *Journal of Cellular Physiology*, 233(5): 3968–3981.

Soto, J., Rodriguez-Antolin, C., Vallespin, E., De Castro Carpeno, J. & De Caceres, I.I. 2016. The impact of next-generation sequencing on DNA methylation-based translational research. *Translational Research*, 169: 1–18.

Springhorn, C., Matsha, T.E., Erasmus, R.T. & Essop, M.F. 2012. Exploring leukocyte O-GlcNAcylation as a novel diagnostic tool for the earlier detection of type 2 diabetes mellitus. *Journal of Clinical Endocrinology and Metabolism*, 97(12): 4640–4649.

Srinivasan, S. & Avadhani, N.G. 2012. Cytochrome c oxidase dysfunction in oxidative stress. *Free Radical Biology and Medicine*, 15(6): 1252–1263.

Standl, E., Khunti, K., Hansen, T.B. & Schnell, O. 2019. The global epidemics of diabetes in the 21st century: Current situation and perspectives. *European Journal of Preventive Cardiology*, 26(2S): 7–14.

Stangenberg, S., Saad, S., Schilter, H.C., Zaky, A., Gill, A., Pollock, C.A. & Wong, M.G. 2018. Lysyl oxidase-like 2 inhibition ameliorates glomerulosclerosis and albuminuria in diabetic nephropathy. *Scientific Reports*, 8(1): 1–10.

Stenvers, D.J., Jongejan, A., Atiqi, S., Vreijling, J.P., Limonard, E.J., Endert, E., Baas, F., Moerland, P.D., Fliers, E., Kalsbeek, A. & Bisschop, P.H. 2019. Diurnal rhythms in the white adipose tissue transcriptome are disturbed in obese individuals with type 2 diabetes compared with lean control individuals. *Diabetologia*, 62(4): 704–716.

Strategic Development Information and GIS Department. 2013. *City of Cape Town – 2011 Census Suburb Bellville South*. City of Cape Town.

Sun, J., Wang, Y., Cui, W., Lou, Y., Sun, G. & Zhang, D. 2017. Role of epigenetic histone modifications in Diabetic kidney disease involving renal fibrosis. *Journal of Diabetes Research*, 7242384.

Szilagyi, A., Blasko, B., Szilassy, D., Fust, G., Sasvari-Szekely, M. & Ronai, Z. 2006. Real-time PCR quantification of human complement C4A and C4B genes. *BMC Genetics*, 7: 1–9.

Szmuilowicz, E.D., Josefson, J.L. & Metzger, B.E. 2019. Gestational Diabetes Mellitus. *Endocrinology and Metabolism Clinics of North America*, 48(3): 479–493.

Tabassum, R., Chauhan, G., Dwivedi, O.P., Mahajan, A., Jaiswal, A., Kaur, I., Bandesh, K., Singh, T., Mathai, B.J., Pandey, Y., Chidambaram, M., Sharma, A., Chavali, S., Sengupta, S., Ramakrishnan, L., Venkatesh, P., Aggarwal, S.K., Ghosh, S., Prabhakaran, D., Srinath, R.K., Saxena, M., Banerjee, M., Mathur, S., Bhansali, A., Shah, V.N., Madhu, S.V., Marwaha, R.K.,

Basu, A., Scaria, V., McCarthy, M.I., Venkatesan, R., Mohan, V., Tandon, N. & Bharadwaj, D. 2013. Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. *Diabetes*, 62(3): 977–986.

Takahashi, N., Kimura, A.P., Otsuka, K., Ohmura, K., Naito, S., Yoshida, M. & Ieko, M. 2019. Dreh, a long noncoding RNA repressed by metformin, regulates glucose transport in C2C12 skeletal muscle cells. *Life Sciences*, 236: 116906.

Tan, L. & Shi, Y.G. 2012. Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development*, 139: 1895–1902.

Tan, Q., Christiansen, L., Thomassen, M., Kruse, T.A. & Christensen, K. 2013. Twins for epigenetic studies of human aging and development. *Ageing Research Reviews*, 12(1): 182–187.

Teicher, B.A. 2019. CD248: A therapeutic target in cancer and fibrotic diseases. *Oncotarget*, 10(9): 993–1009.

Tengholm, A. & Gylfe, E. 2017. cAMP signalling in insulin and glucagon secretion. *Diabetes, Obesity and Metabolism*, 19: 42–53.

Thomas PD. 2017. The Gene Ontology and the meaning of biological function. In Dessimoz, C. & Škunca, N. (eds). *The Gene Ontology Handbook. Methods in Molecular Biology*. New York: Humana Press: 15–24.

Tilves, C.M., Zmuda, J.M., Kuipers, A.L., Nestlerode, C.S., Evans, R.W., Bunker, C.H., Patrick, A.L. & Miljkovic, I. 2016. Association of lipopolysaccharide-binding protein with aging-related adiposity change and prediabetes among African ancestry men. *Diabetes Care*, 39(3): 385–391.

Tsai, P., Glastonbury, C., Eliot, M., Bollepalli, S., Yet, I., Castillo-Fernandez, J., Carnero-Montoro, E., Hardiman, T., Martin, T., Vickers, A., Mangino, M., Ward, K., Pietiläinen, K., Deloukas, P., Spector, T., ViñuelaX, A., Loucks, E., Ollikainen, M., Kelsey, K., Small, K. & Bell, J. 2018. Smoking induces coordinated DNA methylation and gene expression changes in adipose tissue with consequences for metabolic health. *Clinical Epigenetics*, 10: 126.

Tsai, P.C. & Bell, J.T. 2015. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *International Journal of Epidemiology*, 44(4): 1429–1441.

Turner, S.R. & Macdonald, J.A. 2014. Novel contributions of the smoothelin-like 1 protein in vascular smooth muscle contraction and its potential involvement in myogenic tone. *Microcirculation*, 21(3): 249–258.

Unnikrishnan, R., Pradeepa, R., Joshi, S.R. & Mohan, V. 2017. Type 2 diabetes: Demystifying the global epidemic. *Diabetes*, 66(6): 1432–1442.

Van der Schueren, B., Vangoitsenhoven, R., Geeraert, B., De Keyzer, D., Hulsmans, M., Lannoo, M., Huber, H.J., Mathieu, C. & Holvoet, P. 2015. Low cytochrome oxidase 4I1 links mitochondrial dysfunction to obesity and type 2 diabetes in humans and mice. *International Journal of Obesity*, 39: 1254–1263.

Van Otterdijk, S.D., Binder, A.M., Szarc Vel Szic, K., Schwald, J. & Michels, K.B. 2017. DNA methylation of candidate genes in peripheral blood from patients with type 2 diabetes or the metabolic syndrome. *PLoS ONE*, 12: e0180955.

Vasu, S., Kumano, K., Darden, C.M., Rahman, I., Lawrence, M.C. & Naziruddin, B. 2019. MicroRNA signatures as future biomarkers for diagnosis of diabetes states. *Cells*, 8(12): 1533.

Vial, G., Detaille, D. & Guigas, B. 2019. Role of mitochondria in the mechanism(s) of action of metformin. *Frontiers in Endocrinology*, 10(294).

Viganò, S., D'Andrea, G., Valle, P. Della, Santacroce, R., Margaglione, M. & D'Angelo, A. 2018. A novel allele variant of the SERPINF2 gene responsible for severe plasmin inhibitor (α2-antiplasmin) deficiency in an Italian patient. *Thrombosis Research*, 166: 60–62.

Vigetti, D., Deleonibus, S., Moretto, P., Karousou, E., Viola, M., Bartolini, B., Hascall, V.C., Tammi, M., De Luca, G. & Passi, A. 2012. Role of UDP-N-acetylglucosamine (GlcNAc) and O-GlcNacylation of hyaluronan synthase 2 in the control of chondroitin sulfate and hyaluronan synthesis. *Journal of Biological Chemistry*, 287(42): 35544–35555.

Volkmar, M., Dedeurwaerder, S., Cunha, D.A., Ndlovu, M.N., Defrance, M., Deplus, R., Calonne, E., Volkmar, U., Igoillo-Esteve, M., Naamane, N., Del Guerra, S., Masini, M., Bugliani, M., Marchetti, P., Cnop, M., Eizirik, D.L. & Fuks, F. 2012. DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients. *The EMBO Journal*, 31(6): 1405–1426.

Volpe, C., Villar-Delfin, P., Dos Anjos, P. & Nogueira-Machado, J. 2018. Cellular death, reactive oxygen species (ROS) and diabetic complications. *Cell Death and Disease*, 9(2): 119.

Wang, C. & Buolamwini, J.K. 2019. A novel RNA variant of human concentrative nucleoside transporter 1 (hCNT1) that is a potential cancer biomarker. *Experimental Hematology and Oncology*, 8: 18.

Wang, H., Ricklin, D. & Lambris, J.D. 2017a. Complement-activation fragment C4a mediates effector functions by binding as untethered agonist. *Proceedings of the National Academy of Sciences of the United States of America*, 114(41): 10948–10953.

Wang, H., Tian, Y., Chien, C., Kan, W., Liao, P., Wu, H., Su, S. & Lin, C. 2010. Differential proteomic characterization between normal peritoneal fluid and diabetic peritoneal dialysate. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*, 25(6): 1955–1963.

Wang, H., Zhao, W., Liu, J., Tan, P., Tian, W. & Zhou, B. 2017b. ATP5J and ATP5H proactive expression correlates with cardiomyocyte mitochondrial dysfunction induced by fluoride. *Biological Trace Element Research*, 180(1): 63–69.

Wang, S. 2011. Method to detect differentially methylated loci with case-control designs using Illumina arrays. *Genetic Epidemiology*, 35(7): 686–694.

Wang, Z., Park, K., Comer, F., Hsieh-Wilson, L.C., Saudek, C.D. & Hart, G.W. 2009. Site-specific GlcNAcylation of human erythrocyte proteins: Potential biomarker(s) for diabetes. *Diabetes*, 58(2): 309–317.

Ward, M. 2014. Mechanisms underlying the impaired contractility of diabetic cardiomyopathy. *World Journal of Cardiology*, 6(7): 577–584.

Watson, J.F., Ho, H. & Greger, I.H. 2017. Synaptic transmission and plasticity require AMPA receptor anchoring via its N-terminal domain. *eLife*, 6: e2302.

Weaver, J.R., Nadler, J.L. & Taylor-Fishwick, D.A. 2015. Interleukin-12 (IL-12)/ STAT4 axis is an important element for β -cell dysfunction induced by inflammatory cytokines. *PLoS ONE*, 10(11): e0142735.

Wegner, M., Neddermann, D., Piorunska-Stolzmann, M. & Jagodzinski, P.P. 2014. Role of epigenetic mechanisms in the development of chronic complications of diabetes. *Diabetes Research and Clinical Practice*, 105(2): 164–175.

Wegner, M., Winiarska, H., Bobkiewicz-Kozłowska, T. & Dworacka, M. 2008. IL-12 serum levels in patients with type 2 diabetes treated with sulphonylureas. *Cytokine*, 42(3): 312–316.

Wen, Y., Gu, J., Li, S.L., Reddy, M.A., Natarajan, R. & Nadler, J.L. 2006. Elevated glucose and diabetes promote interleukin-12 cytokine gene expression in mouse macrophages. *Endocrinology*, 147(5): 2518–2525.

Werfalli, M., Engel, M.E., Musekiwa, A., Kengne, A.P. & Levitt, N.S. 2016. The prevalence of type 2 diabetes among older people in Africa: A systematic review. *The Lancet Diabetes and Endocrinology*, 4(1): 72–84.

Wex, B., Safi, R.M., Antonios, G., Zgheib, P.Z., Awad, D.B., Kobeissy, F.H., Mahfouz, R.A., El-Sabban, M.M. & Yazbek, S.N. 2018. SLC35B4, an inhibitor of gluconeogenesis, responds to glucose stimulation and downregulates Hsp60 among other proteins in HepG2 liver cell lines. *Molecules*, 23: 1350.

Wile, D.J. & Toth, C. 2010. Association of metformin, elevated homocysteine, and methylmalonic acid levels and clinically worsened diabetic peripheral neuropathy. *Diabetes Care*, 33(1): 156–161.

Willmer, T., Goedecke, J.H., Dias, S., Louw, J. & Pheiffer, C. 2020. DNA methylation of FKBP5 in South African women: Associations with obesity and insulin resistance. *Clinical Epigenetics*, 12: 141.

Willmer, T., Johnson, R., Louw, J. & Pheiffer, C. 2018. Blood-based DNA methylation biomarkers for type 2 diabetes : Potential for clinical applications. *Frontiers in Endocrinology*, 9: 1–16.

Wilson, L.E., Xu, Z., Harlid, S., White, A.J., Troester, M.A., Sandler, D.P. & Taylor, J.A. 2019. Alcohol and DNA methylation: An epigenome-wide association study in blood and normal breast tissue. *American Journal of Epidemiology*, 188(6): 1055–1065.

Wirth, C., Brandt, U., Hunte, C. & Zickermann, V. 2016. Structure and function of mitochondrial complex I. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(7): 902–914.

Witkowski, Marco, Tabaraie, T., Steffens, D., Friebel, J., Dörner, A., Skurk, C., Witkowski, Mario, Stratmann, B., Tschoepe, D., Landmesser, U. & Rauch, U. 2018. MicroRNA-19a contributes to the epigenetic regulation of tissue factor in diabetes. *Cardiovascular Diabetology*, 17(1): 34.

World Health Organisation. 1999. Definition, diagnosis and classification of Diabetes mellitus and its complications. Report of a WHO Consultation.

World Health Organisation. 2006. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia. Report of a WHO/IDF Consultation.

Wreczycka, K., Gosdschan, A., Yusuf, D., Grüning, B., Assenov, Y. & Akalin, A. 2017. Strategies for analyzing bisulfite sequencing data. *Journal of Biotechnology*, 261: 105–115.

Wu, H.P., Kuo, S.F., Wu, S.Y. & Chuang, D.Y. 2010. High interleukin-12 production from stimulated peripheral blood mononuclear cells of type 2 diabetes patients. *Cytokine*, 51(3): 298–304.

Wu, Y., Ding, Y., Tanaka, Y. & Zhang, W. 2014. Risk factors contributing to type 2 diabetes and recent advances in the treatment and prevention. *International Journal of Medical Sciences*, 11(11): 1185–1200.

Xiao, Y., Liu, F., Li, S., Jiang, N., Yu, C., Zhu, X., Qin, Y., Hui, J., Meng, L., Song, C., Li, X.F. & Liu, Y. 2020. Metformin promotes innate immunity through a conserved PMK-1/p38 MAPK pathway. *Virulence*, 11(1): 39–48.

Xing, X., Zhang, B., Li, D. & Wang, T. 2018. Comprehensive whole DNA methylome analysis by integrating MeDIP-seq and MRE-seq. *Methods in Molecular Biology*, 1708: 209–246.

Xiong, Q., Liu, J. & Xu, Y. 2019. Effects of uric acid on diabetes mellitus and its chronic complications. *International Journal of Endocrinology*, 9691345.

Xu, E., Hu, X., Li, X., Jin, G., Zhuang, L., Wang, Q. & Pei, X. 2020. Analysis of long non-coding RNA expression profiles in high-glucose treated vascular endothelial cells. *BMC Endocrine Disorders*, 20(107): 1–10.

Yang, B.T., Dayeh, T.A., Volkov, P.A., Kirkpatrick, C.L., Malmgren, S., Jing, X., Renström, E., Wollheim, C.B., Nitert, M.D. & Ling, C. 2012. Increased DNA methylation and decreased expression of PDX-1 in pancreatic islets from patients with type 2 diabetes. *Molecular Endocrinology*, 26(7): 1203–1212.

Yang, H. & Yang, L. 2016. Targeting cAMP/PKA pathway for glycemic control and type 2 diabetes therapy. *Journal of Molecular Endocrinology*, 57(2): R93–R108.

Yang, J., Savvatis, K., Kang, J.S., Fan, P., Zhong, H., Schwartz, K., Barry, V., Mikels-Vigdal, A., Karpinski, S., Kornyeyev, D., Adamkewicz, J., Feng, X., Zhou, Q., Shang, C., Kumar, P., Phan, D., Kasner, M., López, B., Diez, J., Wright, K.C., Kovacs, R.L., Chen, P.S., Quertermous, T., Smith, V., Yao, L., Tschöpe, C. & Chang, C.P. 2016. Targeting LOXL2 for cardiac interstitial fibrosis and heart failure treatment. *Nature Communications*, 7: 13710.

Yang, Z., Drumea, K., Mott, S., Wang, J. & Rosmarin, A.G. 2014. GABP transcription factor (Nuclear Respiratory Factor 2) is required for mitochondrial biogenesis. *Molecular and Cellular Biology*, 34(17): 3194–3201.

Yazbek, S.N., Buchner, D.A., Geisinger, J.M., Burrage, L.C., Spiezio, S.H., Zentner, G.E., Hsieh, C.W., Scacheri, P.C., Croniger, C.M. & Nadeau, J.H. 2011. Deep congenic analysis identifies many strong, context-dependent QTLs, one of which, SLC35B4, regulates obesity and glucose homeostasis. *Genome Research*, 21(7): 1065–1073.Ye, Y., Jiao, H., Wang, S.,

Xiao, Z., Zhang, D., Qiu, J., Zhang, L., Zhao, Y., Li, T., Li-Liang, W. L. & Ding, Y. 2018. Hypermethylation of DMTN promotes the metastasis of colorectal cancer cells by regulating the actin cytoskeleton through Rac1 signaling activation. *Journal of Experimental & Clinical Cancer Research*, 37: 299.

Yendapally, R., Sikazwe, D., Kim, S.S., Ramsinghani, S., Fraser-Spears, R., Witte, A.P. & La-Viola, B. 2020. A review of phenformin, metformin, and imeglimin. *Drug Development Research*, 2020: 1–12.

Yin, Q., Ni, Q., Wang, Y., Zhang, H., Li, W., Nie, A., Wang, S., Gu, Y., Wang, Q. & Ning, G. 2020. Raptor determines β-cell identity and plasticity independent of hyperglycemia in mice. *Nature Communications*, 11: 2538.

Yong, W.S., Hsu, F.M. & Chen, P.Y. 2016. Profiling genome-wide DNA methylation. *Epigenetics and Chromatin*, 9(26).

Yu, T., Li, L., Liu, H., Li, H. & Liu, Z. 2018. KCNQ2 / 3 / 5 channels in dorsal root ganglion neurons can be therapeutic targets of neuropathic pain in diabetic rats. *Molecular Pain*, 14: 1–15.

Yuan, T., Rafizadeh, S., Durga, K., Gorrepati, D. & Lupse, B. 2017. Reciprocal regulation of mTOR complexes in pancreatic islets from humans with type 2 diabetes. *Diabetologia*, 60: 668–678.

Zakhari, S. 2012. Alcohol metabolism and epigenetics changes. *Alcohol Research: Current Reviews*, 35(1): 6–16.

Zhang, Bo, Zhou, Y., Lin, N., Lowdon, R.F., Hong, C., Nagarajan, R.P., Cheng, J.B., Li, D., Stevens, M., Lee, H.J., Xing, X., Zhou, J., Sundaram, V., Elliott, G., Gu, J., Shi, T., Gascard, P., Sigaroudinia, M., Tlsty, T.D., Kadlecek, T., Weiss, A., O'Geen, H., Farnham, P.J., Maire, C.L., Ligon, K.L., Madden, P.A.F., Tam, A., Moore, R., Hirst, M., Marra, M.A., Zhang, Baoxue, Costello, J.F. & Wang, T. 2013. Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm. *Genome Research*, 23(9): 1522–1540.

Zhang, H. & Pollin, T.I. 2018. Epigenetics variation and pathogenesis in diabetes. *Current Diabetes Reports*, 18: 121.

Zhang, W., Zheng, J., Hu, X. & Chen, L. 2019. Dysregulated expression of long noncoding RNAs serves as diagnostic biomarkers of type 2 diabetes mellitus. *Endocrine*, 65(3): 494–503.

Zhang, Yong, Zhang, Yuping, Sun, K., Meng, Z. & Chen, L. 2018. The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. *Journal of Molecular Cell Biology*, 11(1): 1–13.

Zhao, B., Luo, J., Wang, Y., Zhou, L., Che, J., Wang, F., Peng, S., Zhang, G. & Shang, P. 2019. Metformin suppresses self-renewal ability and tumorigenicity of osteosarcoma stem cells via reactive oxygen species-mediated apoptosis and autophagy. *Oxidative Medicine and Cellular Longevity*, 9290728.

Zheng, X., Narayanan, S., Sunkari, V.G., Eliasson, S., Botusan, I.R., Grünler, J., Catrina, A.I., Radtke, F., Xu, C., Zhao, A., Ekberg, N.R., Lendahl, U. & Catrina, S.B. 2019. Triggering of a

Dll4–Notch1 loop impairs wound healing in diabetes. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14): 6985–6994.

Zheng, Y., Ley, S.H. & Hu, F.B. 2018. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nature Reviews Endocrinology*, 14(2): 88–98.

Zhong, T., Men, Y., Lu, L., Geng, T., Zhou, J., Mitsuhashi, A., Shozu, M., Maihle, N.J., Carmichael, G.G., Taylor, H.S. & Huang, Y. 2017. Metformin alters DNA methylation genome-wide via the H19/SAHH axis. *Oncogene*, 36(17): 2345–2354.

Zhou, J., Massey, S., Story, D. & Li, L. 2018. Metformin: An old drug with new applications. *International Journal of Molecular Sciences*, 19(10): 2863.

Zhou, Y., Zhu, X., Zhang, L., Tang, C. & Feng, B. 2019. Design, synthesis, and biological evaluation of 2-(4-(methylsulfonyl)phenyl)pyridine derivatives as GPR119 agonists. *Chemical Biology and Drug Design*, 93(1): 67–74.

Appendix A: Participant questionnaire

Title: The Cape Town Diabetes and Cardiovascular Disease Study (VMH)

Principal Investigator: **Prof Tandi Matsha**

Name of Interviewer: **…….**

Date of Interview: **……. /……./…....** **Ref No ……**

To the respondent:

Thank you very much for your willingness to participate in the completion of this questionnaire. The information obtained on this questionnaire will provide us with information on all the possible health, family, life style and dietary risk factors within your house hold that might influence the development of diabetes. This is because many health conditions develop slowly over time yet could be prevented if diagnosed early or if pre-determined. This questionnaire therefore aims at getting information which may be used to determine the extent of diabetes and those likely to develop diabetes in the future. The questionnaire should not take long and we hope you find it interesting and enjoyable. All answers provided will be treated as confidential and anonymous.

Note

No special knowledge is needed to fill this questionnaire. Please feel free to ask for clarification if needed.

Section A: General Questions

1. Language

1. What is your home language?
   - □ English
   - □ Afrikaans
   - □ Xhosa
   - □ Other: ……..

2. Which language would you prefer to be communicated in?
   - □ English
   - □ Afrikaans
   - □ Xhosa
   - □ Other: ……..

2. Personal Questions

1. Gender ___Male ___Female
2. Date of birth? ____/_____/_____
3. What is your relationship status?
   - ☐ Married/registered partnership
   - ☐ Cohabiting (living together)
   - ☐ Unmarried (never married)
   - ☐ Divorced or separated
   - ☐ Widow/widower
4. Including yourself, how many people are there in your household? |__|__| person(s) .
   This includes children who live with you only some of the time.

Section B: General Health

3. General Health

3.1 In general, would you say your health is:

- ☐ Excellent
- ☐ Very good
- ☐ Good
- ☐ Fair
- ☐ Poor

3.2 In general, would you say you are physically active (that is, gardening, jogging etc)?

- ☐ Yes
- ☐ No

3.3 In general, would you say you have emotional problems (such as feeling depressed or anxious)

- ☐ Yes
- ☐ No

3.4 During the past 4 weeks, how much did pain keep you from doing your normal activities?

- ☐ Not at all
- ☐ A little bit
- ☐ Quite a lot
  A lot
- ☐ A very great deal

3.5 The following questions are about how you felt and how you were doing during the past 4 weeks. For each question, please choose the answer that best describes how often you felt this way.

a. During the past 4 weeks, how often did you feel calm and contented?

- ☐ Always
- ☐ Usually
- ☐ Often
- ☐ Sometimes
- ☐ Hardly ever
- ☐ Never

b. During the past 4 weeks, how often did you feel very energetic?

- ☐ Always
- ☐ Usually
- ☐ Often
- ☐ Sometimes
- ☐ Hardly ever
- ☐ Never

c. During the past 4 weeks, how often did you feel down and depressed?

- ☐ Always
- ☐ Usually
- ☐ Often
- ☐ Sometimes
- ☐ Hardly ever
- ☐ Never

3.6 During the past 4 weeks, how often did your physical health or emotional problems limit your social activities (such as visiting friends or family)?

- ☐ Always
- ☐ Usually
- ☐ Often
- ☐ Sometimes
- ☐ Hardly ever
- ☐ Never

3.7 Please indicate the status of your eyesight at the present time, using both eyes (with glasses or contact lenses, if you wear them)

- ☐ Excellent
- ☐ Good
- ☐ Fair
- ☐ Poor
- ☐ Very Poor
- ☐ Completely blind

4. Question for women

If you are a woman, please answer the questions below. If you are a man, please go directly to Question 5.

4.1 How old were you when you had your first period (menstruation)? |___|___|
***If you aren't sure, please try to estimate this*.**

4.2 How old were you the first time your period stopped for a whole year? |___|___|
Do not include times when your period stopped because of pregnancy, breastfeeding, or using birth control.

4.3 Do you currently use contraceptive medication? (Birth control pill)

☐ Yes

☐ No

4.4 Have you given birth to one or more children?

☐ Yes

☐ No- **Go to question 5**

4.5 How many children have you had? |____|____| children
(This includes still born babies)

Section C: Specific Illness and disorders

5. Illness and disorders

Please indicate which of the following illnesses and disorders you have now or that you have had in the past 12 months, and whether or not this was diagnosed by a doctor.

If you don't know or if you have had a certain illness or disorder, please fill in 'No'. Please give an answer for every illness/disorder. [Interviewer: please explain medical terms if necessary]

	No	**Yes, not diagnosed by a doctor**	**Yes, diagnosed by a doctor**
a. stroke, brain hemorrhage, cerebral infarction, or TIA ('transient ischaemic attack': temporary loss of bodily function)	☐	☐	☐
b. heart attack (myocardial infarction)	☐	☐	☐
c. another serious heart condition (for example, heart failure or angina pectoris (severe chest pain))	☐	☐	☐
d. a form of cancer (malignant disorder)	☐	☐	☐
e. migraine or frequent severe headaches	☐	☐	☐
f. severe or chronic fatigue	☐	☐	☐
g. narrowing of the arteries in the belly or legs (artery stenosis)	☐	☐	☐
h. asthma, chronic bronchitis, lung emphysema, or CNSLD (chronic non-specific lung disease) or COPD (chronic obstructive pulmonary disease)	☐	☐	☐
i. serious or persistent intestinal disorders lasting more than 3months	☐	☐	☐
j. chronic inflammation of the joints (inflammatory rheumatism, chronic rheumatism, rheumatoid arthritis)	☐	☐	☐

5.2 Have you had cancer ☐ Yes

☐ No

5.3 What type of cancer _____

_____

_____

6. Medicine and prescription

6.1 Are you currently on medication? *(Do not include birth control and vitamins)*

☐ Yes

☐ No

6.2 How many days in the past have you been able to take your prescribed medicine(s)?

☐ Not a single day

☐ 1 to 2 days

☐ 3 to 4 days

☐ 5 to 6 days

☐ All 7 days

6.3 Now I would like tell me the **<u>names of all medication(s)</u>** you currently taking on a regular basis. I need you to include vitamins and over-the-counter medicine, as well as herbal remedies that you have taken at least once a day over the past two weeks.

1._____	7._____	13._____
2._____	8._____	14._____
3._____	9._____	15._____
4._____	10._____	16._____
5._____	11._____	17._____
6._____	12._____	18._____

6.4 Are you using or have used any prohibited drugs?

☐ Tik

☐ Marijuana (Dagga)

☐ Cocaine

☐ Other:_____

☐ Prefer not to answer

7. Blood pressure

7.1 Have you ever been diagnosed with high blood pressure at a hospital or clinic (by a doctor or nurse etc)?

☐ No- **got to Question 8**

☐ Yes

7.2 How old where you when you were diagnosed with high blood pressure?

___ Years

7.3 Do you use medication for your high blood pressure?

☐ No

☐ Yes

7.4 Are you on a special diet (for example low salt) for your blood pressure

☐ No

☐ Yes

8. Cholesterol

8.1 Have you ever been diagnosed with a high blood cholesterol at a hospital or clinic (by a doctor or nurse etc)?

☐ No- **Go to Question 9**

☐ Yes

☐ I don't know- **go to Question 9**

8.2 When was the last time a doctor checked your cholesterol level?

☐ Never

☐ I don't know

☐ More than 2 years ago

☐ Between 1 and 2 years ago

☐ Between 6 months and 1 year ago

☐ Less than 6 months ago

8.3 How old where you when you were diagnosed with high blood cholesterol?

___ Years

8.4 Do you use cholesterol-lowering medication?

☐ No

8.5 Are you on a cholesterol-lowering diet right now?

- ☐ Yes
- ☐ No
- ☐ Yes

9. Blood sugar and diabetes mellitus

9.1 When was the last time a doctor checked your blood sugar (glucose) level?

- ☐ Never**- go to Question 9.12**
- ☐ I don't know
- ☐ More than 2 years ago
- ☐ Between 1 and 2 years ago
- ☐ Between 6 months and 1 year ago
- ☐ Less than 6 months ago

9.2 Have you ever been diagnosed with diabetes by a doctor or health care worker?

- ☐ No- **go to Question 9.12**
- ☐ Yes
- ☐ I don't know

9.3 Were you diagnosed with diabetes only when you were pregnant?

- ☐ No, I'm a man
- ☐ Yes- **go to Question 9.12**
- ☐ I don't know

9.4 How old were you when you were first diagnosed with diabetes? *If you aren't sure, please try to estimate this.*

|__|__| years old

9.5 Has a doctor or specialist treated you for diabetes in the past 12 months?

- ☐ No
- ☐ Yes

9.6 Do you use tablets for your diabetes?

- ☐ No
- ☐ Yes

9.7 Are you on a diabetic diet right now?

☐ No

☐ Yes

9.8 Do you use insulin injections for your diabetes?

☐ No

☐ Yes

9.9 Did you start insulin injections immediately after being diagnosed with diabetes?

☐ No

☐ Yes

9.10 Did you start insulin injections within 6 months of being diagnosed with diabetes?

☐ No

☐ Yes

9.11 Have you ever been told by a doctor or health care worker that you have eye disease or eye damage as a result of your diabetes (diabetic retinopathy)?

☐ No

☐ Yes

9.12 Has someone in your immediate family (your parents, brothers, sisters, or children) been diagnosed with diabetes?

☐ No

☐ Yes

10. Chest pain

10.1 Have you **ever** had any pain or discomfort in your chest? *This does not include problems caused by a cold, asthma, or a stomach ulcer.*

☐ No- **Go to Question 11**

☐ Yes

10.2 Do you get this pain when you're exerting yourself (for example, when you're climbing stairs, walking fast, or cycling)?

☐ No- **Go to Question 10.8**

☐ Yes

10.3 Do you get this pain when you're just walking along the street at a normal pace?

☐ No

☐ Yes

10.4 If you get pain or discomfort in your chest when walking or cycling, what do you usually do?

☐ Continue at the same pace – **go to Question 10.7**

☐ Slow down or stop

☐ Use tablet or spray under the tongue and continue at the same pace

☐ Use a tablet or spray under the tongue and slow down

10.5 If you stop or slow down, or use a tablet or spray under the tongue, does the pain disappear?

☐ No- go to **Question 10.7**

☐ Yes

10.6 How soon does it disappear?

☐ Within 10 minutes

☐ After more than 10 minutes

10.7 Where do you get this pain or discomfort?
You can give more than one answer.

☐ In the upper part of my chest

☐ In the lower part of my chest

☐ On the left side of my chest

☐ In my left arm

☐ Somewhere else namely:

……

10.8 Have you **ever** had severe pain across the front of your chest that lasted for half an hour or more?

☐ Yes

☐ No

11. Leg pain

11.1 Do you get pain in either leg when walking?

☐ No- **go to Question 12**

☐ Yes

11.2 Does this pain ever start when you're standing still or sitting?

☐ No

☐ Yes

11.3 In which part of your legs do you get this pain?
You can give more than one answer.

☐ In the calf

☐ Other location, namely:

……..

11.4 Do you get this pain when walking fast or climbing stairs?

- ☐ No
- ☐ Yes
- ☐ I never do this

11.5 Do you ever get this pain when you're just walking along the street at a normal pace?

- ☐ No
- ☐ Yes

11.6 Does the pain ever disappear while you're still walking?

- ☐ No
- ☐ Yes

11.7 What do you do if this happens while you're walking?

- ☐ Continue on the same pace
- ☐ Stop or slow down

11.8 What happens if you stop?

- ☐ The pain usually disappears within 10 minutes
- ☐ The pain usually disappears after more than 10 minutes
- ☐ The pain continues

12. Neurologic dysfunction

Neurologic dysfunction is a temporary loss of bodily function: sudden numbness or weakness in the face or other parts of the body (for example, having difficulty finding the right words, a partial or complete paralysis or 'drop' of your hand, arm, foot, leg, or face).

12.1 Have you **ever** had a loss of bodily function that lasted for **less than one day**?

- ☐ No- go to **Question 12.3**
- ☐ Yes

12.2 On this day, was this on just one side of your body or on both sides?

- ☐ On just one side
- ☐ On both sides(at the same time, or changing from left to right)

12.3 Have you **ever** had a stroke?

- ☐ No
- ☐ Yes

12.4 Have you **ever** fainted?

☐ No

☐ Yes

13. Family

By cardiovascular disease, we mean a heart attack, a dotter procedure (angioplasty) or bypass operation on the heart or legs, a TIA, or a stroke.

13.1 Has anyone in your immediate family (that is your parents, brothers, sisters, daughters or sons) ever been diagnosed with cardiovascular disease.

☐ No – **go to Question 13.3**

☐ Yes

☐ I don't know

13.2. Please indicate these family member(s), and how old they were when they were **first** diagnosed with cardiovascular disease *you can give more than one answer. If you aren't sure of their age, please try to estimate this.*

☐ Father, at the age of **|__|__|**

☐ Mother, at the age of **|__|__|**

☐ Brother, at the age of **|__|__|**
(if more than one brother, please put down the youngest age at occurrence)

☐ Sister, at the age of |__|__|
(if more than one sister, please put down the youngest age at occurrence)

☐ Son, at the age of |__|__|
(if more than one son, please put down the youngest age at occurrence)

☐ Daughter, at the age of |__|__|
(if more than one daughter, please put down the youngest age at occurrence)

13.3 Has anyone in your immediate family (that is, your parents, brothers, sisters, daughters or sons) **ever** suddenly died <u>when they were 60</u> <u>years old or younger</u> with no clear cause of death?

☐ No- **go to Question 14**

☐ Yes

☐ I don't know –**go to Question 14**

13.4 Could you please indicate these family member(s), and how old they were when they died suddenly? You can give more than one answer. If you aren't sure of their age, please try to estimate this

☐ Father, at the age of |__|__|

☐ Mother, at the age of |__|__|

☐ Brother, at the age of |__|__|
(if more than one brother, please put down the youngest age at occurrence)

☐ Sister, at the age of |__|__|
(if more than one sister, please put down the youngest age at occurrence)

☐ Son, at the age of |__|__|
(if more than one son, please put down the youngest age at occurrence)

☐ Daughter, at the age of |__|__|
(if more than one daughter, please put down the youngest age at occurrence)

Section D: Country of birth and lifestyle

14. Country of birth

14.1 What is your country of origin?

☐ South Africa

☐ Other: ……..

14.2 What is your ethnicity?

☐ Black

☐ White

☐ Mixed ancestry (Coloured)

☐ Asian

☐ Indian

14.3 How long have you been living in your area of residence?

Less than 6 Months ☐ Less than 1 Year ☐

1-5 Years ☐ 5 years and above ☐

14.4 What is your mother's country of birth?

☐ South African

☐ Other: ……

14.5 What is your mother's ethnicity?

☐ Black

☐ White

☐ Mixed ancestry (Coloured)

☐ Asian

☐ Indian

14.6 What is your father's country of birth?

☐ South African

☐ Other: ……

14.7 What is your father's ethnicity?

- ☐ Black
- ☐ White
- ☐ Mixed ancestry (Coloured)
- ☐ Asian
- ☐ Indian

15. Smoking

15.1 Do you smoke at all?

- ☐ Yes – **Go to question 15.4**
- ☐ No, I've never smoked- **go to question 15.6**
- ☐ No but I used to smoke **Go to question 15.4**

15.2 How long did you smoke? |__|__| years & |__|__| months

15.3 How long has it been since you quit? |__|__| years & |__|__| months

15.4 How many years have you smoked? If you aren't sure, please try to estimate this. |__|__| years

15.5 What did you smoke and how much? You can give more than one answer.
(After answering go to question 16)

☐ About |__| cigarettes from a pack of |___| or hand-rolled a day

☐ About |___| cigars a week

☐ About |__|__| package(s) of pipe tobacco (50 grams) a week

15.6 How many people in your household smoke |__|__|

15.7 For how many hours, on average each day, are you closely subjected to other people's tobacco smoke? |__|__|

16. Alcohol

16.1 Have you ever consumed any alcoholic drinks (Wine, Beer, and Spirits)?

☐ Yes

☐ No

16.2 Do you still consume alcoholic drinks?

☐ Yes

☐ No

16.3 If you consume or consumed alcohol, how old were you when you first started drinking? |__|___| years

16.4 If you stopped, how old were you when you stopped drinking? |__|___| years old

16.5 Which type of alcohol do you or did you drink?

☐ Beer

☐ Spirits

☐ Wine

☐ Other:_____

16.6 When you drink or drank alcoholic drinks, how many drinks or glasses do you or did you consume daily? Indicate the number |__|___|

16.7 How many days a week do you or did you consume alcohol? |__|___| day(s)

16.8 Have you or did you ever feel you should cut down your drinking?
☐ Yes
☐ No

16.9 Have people ever annoyed you by criticizing your drinking?
☐ Yes
☐ No

16.10 Have you ever felt bad about your drinking?
☐ Yes
☐ No

16.11 Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover (Eye Opener)
☐ Yes
☐ No

17. Physical activity

Next I am going to ask you about the time you spend doing different types of physical activity **in a typical week**. Please answer these questions even if you do not consider yourself to be a physically active person. Think first about the time you spend doing work. Think of work as the things that you have to do such as paid or unpaid work, study/training, household chores, harvesting food/crops, fishing or hunting for food, seeking employment. *[Insert other examples if needed].*

In answering the following questions 'vigorous-intensity activities' are activities that require hard physical effort and cause large increases in breathing or heart rate, 'moderate-intensity activities' are activities that require moderate physical effort and cause small increases in breathing or heart rate.

17.1 Work

Please describe your physical activity at work

a. Does your work involve vigorous-intensity activity that causes large increases in breathing or heart rate like (carrying or lifting ☐ Yes

heavy loads, digging or construction work) for **at least 10 minutes** continuously?

☐ No- **go to Question 17.1d**

b. In a **typical week**, on how many days do you do vigorous-intensity activities as part of your work?

|__| days

c. How much time do you spend doing vigorous-intensity activities at work on a **typical day**?

Hours |__| Minutes|__|

d. Does your work involve moderate-intensity activity that causes small increases in breathing or heart rate such as brisk walking (or carrying light loads) for **at least 10 minutes** continuously?

☐ Yes

☐ No- **go to Question 17.2**

e. In a **typical week**, on how many days do you do moderate-intensity activities as part of your work?

Number of days|__|

f. How much time do you spend doing moderate-intensity activities at work on a **typical day**?

Hours |__| Minutes|__|

17.2 Travel to and from places

The next questions exclude the physical activities at work that you have already mentioned. Now I would like to ask you about the usual way you travel to and from places. For example to work, for shopping, to market, to place of worship. [*Insert other examples if needed*].

a. Do you walk or use a bicycle (pedal cycle) for **at least 10 minutes continuously to get to and from places?**

☐ Yes

☐ **No- go to Question 17.3**

b. In a **typical week**, on how many days do you walk or cycle for **at least 10 minutes** to get to and from places?

|__| days

c. How much time do you spend walking or cycling for travel on a typical day?

Hours |__| Minutes|__|

17.3 Recreation al activities

The next questions exclude the work and transport activities that you have already mentioned. Now I would like to ask you about sports, fitness and recreational activities (leisure), [*Insert relevant terms*].

a. Do you do any vigorous-intensity sports, fitness or recreational (leisure) activities that cause large increases in breathing or heart rate (like running or football) for **at least 10 minutes** continuously?

☐ Yes

☐ No- **go to Question 17.3d**

b. In a **typical week**, on how many days do you do vigorous-intensity sports, fitness or recreational (leisure) activities?

|__| days

c. How much time do you spend doing vigorous-intensity sports, fitness or recreational activities on a **typical day**?

Hours |__| Minutes|__|

d. Do you do any moderate-intensity sports, fitness or recreational (leisure) activities that cause a small increase in breathing or heart rate such as brisk walking, (cycling, swimming, volleyball) for **at least 10 minutes** continuously?

☐ Yes

☐ No- **go to question 17.4**

e. In a **typical week**, on how many days do you do moderate-intensity sports, fitness or recreational (leisure) activities?

Number of days|__|

f. How much time do you spend doing moderate-intensity sports, fitness or recreational (leisure) activities on a **typical day**?

Hours |__| Minutes|__|

17.4 Sedentary behavior

The following question is about sitting or reclining at work, at home, getting to and from places, or with friends including time spent sitting at a desk, sitting with friends, traveling in car, bus, train, reading, playing cards or watching television, but **do not include time spent sleeping**. [*INSERT EXAMPLES*]

How much time do you usually spend sitting or reclining on a **typical day**?

Hours |__| Minutes|__|

Section E: Education and Employment

18. Education

18.1. What is the highest level of education you have completed?

This is the highest level of education you completed and for which you received a diploma or a certificate of proficiency.

☐ Primary School or less

☐ High School (Not Completed)

☐ High School graduate

☐ College Or Technical College (Not Completed)

☐ College or Technical College Graduate

☐ University or Technikon (Not Completed)

☐ University or Technikon graduate

18.2. Are you going to school at the moment?

☐ No - **go to Question 19**

☐ Yes, day classes

☐ Yes, evening classes

☐ Both day and evening classes

☐ Secondary education

18.3. Which course of study are you following right now?

☐ Secondary education

☐ College

☐ University/Technikon

☐ Other, namely:……

19. Employment

19.1. Which situation most applies to you?

☐ I have a paid job, and work 32 or more hours a week. - **Go to Question 19.4**

☐ I have a paid job, and work between 20 and 32 hours a week. -7 **Go to Question 19.4**

☐ I have a paid job, and work between 12 and 20 hours a week. -7 **Go to Question 19.4**

☐ I have a paid job, and work less than 12 hours a week. - **Go to Question 19.4**

☐ I'm retired.

☐ I'm unemployed and looking for work

☐ I'm unable to work I get social benefits

☐ I'm a full-time homemaker (male or female).

☐ I'm a student without part time work

☐ I'm a student with part time work

19.2. If you're **not** working **right now**, have you had a paid job in the past?

☐ Yes

☐ No- **Go to Question 20**

19.3. When (in what year) did you stop working?

Year |__|__|__|__|

19.4 . What is your job or profession now? Or if you're not working right now, what was your last job or profession?

Please describe this with as much detail as you can (for example, primary school teacher, manager of a software company, or worker in a cheese factory rather than teacher, manager, or factory worker).

……

……

……

……

……

19.5. Do you (or did you) have to work irregular hours, such as shift work or nights?

☐ Yes, namely |__|__| hours a week

☐ No

19.6 Which situation best describes (or described) you?

☐ Salaried job

☐ Self-employed

☐ Working in a family business

20. Household income

20.1. Which of these options add to the net income of your household?

This relates to the income of the whole household, not just your own (so you can give more than one answer

☐ Wages or salary

☐ Income from own company or activities

☐ Income from investments

☐ Pension

☐ State pension benefits

☐ Incapacity (sickness) benefits

☐ Unemployment benefits

☐ Social benefits

☐ Student grants and loans

☐ Others, namely:……

20.2 How many people in your household need to live from this income (including yourself)?

|__|__| people

20.3 Are there people outside your household who live wholly or partially from this income?

Think of children away at university, alimony for an ex-partner, etc

☐ Yes, namely: |__|__| person(s)

☐ No

20.4 During the past year, did you have problems managing your household income?

☐ No, no problems at all

☐ No problems, but I have to watch what I spend

☐ Yes, some problems

Yes, lots of problems

☐

Section F: Personality, Experiences and Well-being

21. Dealing with everyday problems

The following statements are about how you deal with everyday problems.
For each statement please indicate to what extent it applies to you.

	Totally disagree	Disagree	Neutral	Agree	Totally agree
a. I have little control about the things that happen to me.	☐	☐	☐	☐	☐
b. I can't seem to solve some of my problems at all.	☐	☐	☐	☐	☐
c. There isn't much I can do to change important things in my life.	☐	☐	☐	☐	☐
d. I often feel helpless in dealing with the problems of life.	☐	☐	☐	☐	☐
e. Sometimes I feel like a play ball of life.	☐	☐	☐	☐	☐

22. Diet

22.1. What day was it yesterday

Sunday |	Monday |	Tuesday |	Wednesday |	Thursday |	Friday |

22.2. Would you describe the food that you ate yesterday as typical of your usual food intake?

	Yes		No

I want to find out everything you ate or drank yesterday, including water or food you picked up from the veld. Please tell me everything you ate from the time you woke up to the time you went to sleep. I will also ask you where you ate the food and how much you ate.

Morning (up to 9:00am)			
Time	Place	Description of food ate	Amount

Mid-Morning(9am to 11.59am)			
Time	Place	Description of food ate	Amount

Afternoon (12:00pm to 3:00pm)			
Time	Place	Description of food ate	Amount

Mid-Afternoon (3pm to 5pm)			
Time	Place	Description of food ate	Amount

Evening (5pm to 9pm)			
Time	Place	Description of food ate	Amount

Before bed(9pm till late)			
Time	Place	Description of food ate	Amount

23. Recent experiences

23.1. We will now mention some events. Please indicate whether you've experienced these events in the **past 12 months**

a. You suffered from a serious illness or injury	☐ No	☐ Yes
b. A close relative had a serious illness or injury	☐ No	☐ Yes
c. Your parent, child, brother, sister, or spouse died	☐ No	☐ Yes
d. Another relative (such as an aunt, cousin, or grandparent) or close friend died.	☐ No	☐ Yes
e. You broke off a steady relationship	☐ No	☐ Yes
f. A long-term friendship with a good friend or family member was broken off	☐ No	☐ Yes
g. You had a serious problem with a good friend, family member, or neighbor	☐ No	☐ Yes
h. You were sacked from your job or became unemployed	☐ No	☐ Yes
i. You had a major financial crisis	☐ No	☐ Yes

23.2. **In the past 12 months**, have you felt stressed (feeling irritable or anxious or having trouble sleeping) because of the situation at work or place of study

- ☐ Never
- ☐ Some periods
- ☐ Several periods
- ☐ Constantly
- ☐ Doesn't apply

24. Recent well being

In the **past 2 weeks**, how often have you had the following problems?

	Never	On several days	On more than half of the days	Nearly every day
a. Little interest or pleasure in doing things	☐	☐	☐	☐ .
b. Feeling down, depressed, or hopeless	☐	☐	☐	☐ .
c. Trouble falling or staying asleep, or sleeping too much	☐	☐	☐	☐ .
d. Feeling tired or having little energy	☐	☐	☐	☐ .
e. Poor appetite or overeating	☐	☐	☐	☐ .
f. Feeling bad about yourself or feeling like a failure or like you've let yourself or your family down	☐	☐	☐	☐ .
g. Trouble concentrating on things, like reading the newspaper or watching television	☐	☐	☐	☐ .
h. Moving or speaking so slowly that other people might notice	☐	☐	☐	☐ .
i. Being so fidgety or restless that you move around more than usual	☐	☐	☐	☐ .
j. Thinking that you'd be better off dead, or thinking about hurting yourself in some way	☐	☐	☐	☐ .
k. Feeling stressed due to the financial or material demands of your family/friends/relatives	☐	☐	☐	☐ .
l. Feeling stressed due to the demands of the society (e.g. rules, fast way of living, bureaucratic system)	☐	☐	☐	☐

Section G: Body shape

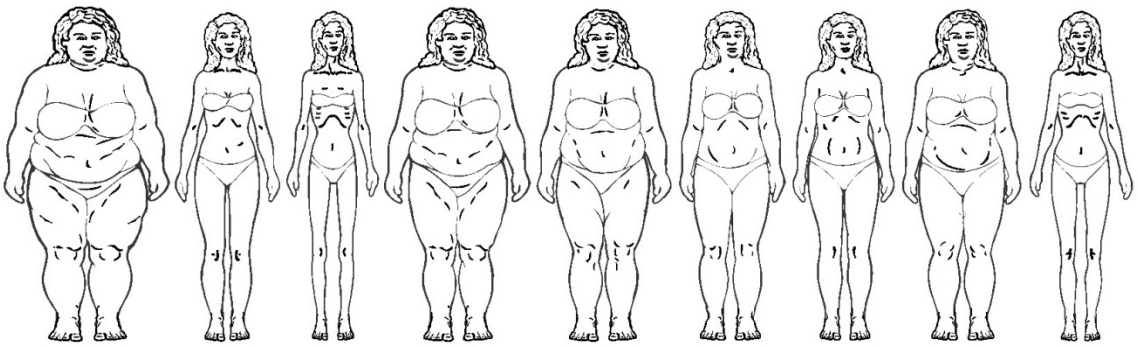
25. Body shape (Females only)

*If you are a **woman**, please answer the questions below. If you are a man, you can go directly to **Question 26**.*

Finally, we want to ask some questions about body shape. For the following questions, you can choose one of the pictures below. Under each picture is a number. Please use this number for your answer

	Please put an **X** under **one** of the numbers below								
	1	2	3	4	5	6	7	8	9
a. Which picture do you most look like right now?	☐	☐	☐	☐	☐	☐	☐	☐	☐
b. Which picture would you most prefer to look like?	☐	☐	☐	☐	☐	☐	☐	☐	☐
c. Which picture is most like other women your age?	☐	☐	☐	☐	☐	☐	☐	☐	☐
d. Which picture do you think most of the men around you would prefer women to look like?	☐	☐	☐	☐	☐	☐	☐	☐	☐

.

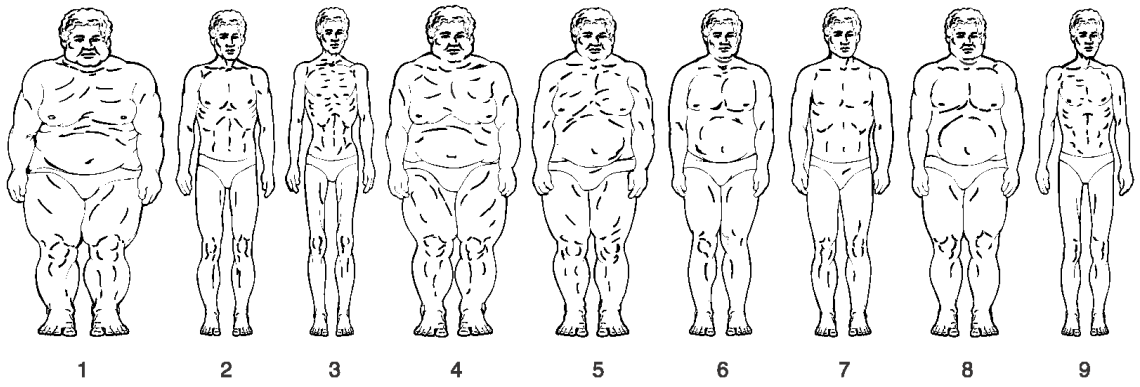


26. Body shape (Males only)

*If you are a **man**, please answer the questions below.*

Finally, we want to ask some questions about body shape. For the following questions, you can choose one of the pictures below. Under each picture is a number. Please use this number for your answer

	Please put an **X** under **one** of the numbers below								
	1	2	3	4	5	6	7	8	9
e. Which picture do you most look like right now?	☐	☐	☐	☐	☐	☐	☐	☐	☐
f. Which picture would you most prefer to look like?	☐	☐	☐	☐	☐	☐	☐	☐	☐
g. Which picture is most like other women your age?	☐	☐	☐	☐	☐	☐	☐	☐	☐
h. Which picture do you think most of the men around you would prefer women to look like?	☐	☐	☐	☐	☐	☐	☐	☐	☐



27. Feelings

Now I'm going to ask you questions about how you've been feeling over the past week.

Please tell me the best answer for ***how you have felt over the past week***:

Yes	No	Geriatric Depression Scale
[ ]	[ ]	1. Are you basically satisfied with your life?
[ ]	[ ]	2. Have you dropped many of your activities and interests?
[ ]	[ ]	3. Do you feel that your life is empty?
[ ]	[ ]	4. Do you often get bored?
[ ]	[ ]	5. Are you in good spirits most of the time?
[ ]	[ ]	6. Are you afraid that something bad is going to happen to you?
[ ]	[ ]	7. Do you feel happy most of the time?
[ ]	[ ]	8. Do you often feel helpless?
[ ]	[ ]	9. Do you prefer to stay at home, rather than going out and doing new things?
[ ]	[ ]	10. Do you feel that you have more problems with memory than most people?
[ ]	[ ]	11. Do you think it is wonderful to be alive now?
[ ]	[ ]	12. Do you feel pretty worthless the way you are now?

Yes	No	
[ ]	[ ]	13. Do you feel full of energy?
[ ]	[ ]	14. Do you feel that your situation is hopeless?
[ ]	[ ]	15. Do you think that most people are better off than you are?

28. Anxiety (Hopkins Symptom Checklist):

28.1 During the past week, have you felt nervous or shaky inside?

0=No
1=a little
2=sometimes

3=extremely
4=do not know

28.2 During the past week, did you have to avoid certain things, places or activities because they frighten you?

0=No
1=a little
2=sometimes
3=extremely
4=do not know

28.3 During the past week, have you felt tense?

0=No
1=a little
2=sometimes
3=extremely
4=do not know

28.4 During the past week, have you felt fearful?

0=No
1=a little
2=sometimes
3=extremely
4=do not know

29. Mastery:

Please tell me whether you agree or disagree with this statement: I can do just about anything I really set my mind to.

1=strongly agree
2=somewhat agree
3=somewhat disagree
4=strongly disagree

Section I: Clinical Measurements

30. Body Weight

30.1 What do you think of your body weight?

- ☐ I'm much too heavy
- ☐ I'm a little too heavy
- ☐ I'm just about right
- ☐ I'm a little too thin
- ☐ I'm much too thin

30.2 Are you trying to do something about your weight right now?

- ☐ No, nothing
- ☐ Yes, I'm trying to lose weight
- ☐ Yes, I'm trying to stay the same weight
- ☐ Yes, I'm trying to gain weight

31. Weight and height

Body Weight (kg)	☐☐☐	**Comment:** ……
Body height (cm)	☐☐☐	**Comment:** ……..
Visceral fat	☐☐☐	**Comment:** ……..
Body fat rate	☐☐☐	**Comment:** ……..
Muscle percentage	☐☐☐	**Comment:** ……..
RM	☐☐☐	**Comment:** ……..
BMI	☐☐☐	**Comment:** ……..

32. Circumference measurements

Waist Circumference 1 (cm)				Comment:
Waist Circumference 2 (cm)				
Waist Circumference 3 (cm)				

Hip Circumference (cm)				Comment:
Hip Circumference (cm)				
Hip Circumference (cm)				

33. Blood pressure measurements

Systolic Pressure 1 (mmHg)				Comment:
Systolic Pressure 2 (mmHg)				
Systolic Pressure 3 (mmHg)				

Diastolic Pressure 1 (mmHg)				Comment:
Diastolic Pressure 2 (mmHg)				
Diastolic Pressure 3 (mmHg)				

Pulse 1 (Beat per Minute)				Comment:
Pulse 2 (Beat per Minute)				
Pulse 3 (Beat per Minute)				

Section J: Blood collection

34. Fasting bloods

34.1 Are you a diabetic

- ☐ Yes, **Complete question 34.1 and skip question 35**
- ☐ No, **must complete question 34 &35**

34.2 Have you collected the following fasting bloods?

34.2.1 One(1) 5ml Green Top Tubes

- ☐ Yes
- ☐ No, why:_____

34.2.2 Two (2) 10ml or Three (3) 5ml Gold Top Tubes

- ☐ Yes,
- ☐ No, why:_____

34.2.3 One(1) 5ml Grey Top Tubes

- ☐ Yes
- ☐ No, why:_____

34.2.4 Two (2) 5ml Purple Top Tubes

- ☐ Yes
- ☐ No ,why:_____

34.2.4 One (1) RNA tube

- ☐ Yes
- ☐ No ,why:_____

35. Glucose bloods

Have you collected the following fasting bloods?

35.2.1 One (1) 10ml or Two(2) 5ml Gold Top Tubes

- ☐ Yes
- ☐ No, why:_____

35.2.2 One (1) 5ml Grey Top Tubes

- ☐ Yes
- ☐ No, why:_____

35.2.3 One(1) 5ml Purple Top Tubes

- ☐ Yes
- ☐ No, why:_____

35.2.4 One(1) 4ml Light Blue Top Tubes

- ☐ Yes
- ☐ No, why:_____

Appendix B: Consent form

PARTICIPANT INFORMATION AND INFORMED CONSENT FORM FOR RESEARCH INVOLVING GENETIC STUDIES

TITLE OF RESEARCH PROJECT: PROGRESSIVE RESEARCH ON RISK FACTORS OF TYPE 2 DIABETES AND CARDIOVASCULAR DISEASES IN SOUTH AFRICA

REFERENCE NUMBER:

PRINCIPAL INVESTIGATORS:	**Professor Tandi Matsha (Cape Peninsula University of Technology)** **Professor Rajiv Erasmus (Stellenbosch University)** **Professor Andre Kengne (SA Medical Research Council)**
Project manager:	**Dr Gloudina Maria Hon (Cape Peninsula University of Technology)**
ADDRESS:	**Obesity and chronic diseases of lifestyle Department of Biomedical Sciences Faculty of Health & Wellness Sciences** **Cape Peninsula University of Technology, Bellville**
CONTACT NUMBER:	**Prof T Matsha 021 959 6366 or email:matshat@cput.ac.za**
Ethics approval:	**Cape Peninsula University of Technology Ethics Reference number: CPUT/SW-REC 2015/H01** **University of Stellenbosch Ethics Reference number: N14/01/003**

We would like to invite you to participate in a research study that involves genetic analysis and possible long-term storage of blood or tissue specimens. Please take some time to read the information presented here which will explain the details of this project. Please ask the study staff or doctor any questions about any part of this project that you do not fully understand. It is very important that you are fully satisfied that you clearly understand what this research entails and how you could be involved. Also, your participation is ***entirely voluntary*** and you are free to decline to participate. If you say no, this will not affect you negatively in any way whatsoever. ***You are also free to withdraw from the study at any point, even if you do agree to take part initially.***

This research study has been approved by the ethics **Faculty of Health & Wellness Sciences of the Cape Peninsula University of Technology** and it will be conducted according to international and locally accepted ethical guidelines for research, namely the Declaration of Helsinki, and the SA Department of Health's 2004 Guidelines: *Ethics in Health Research: Principles, Structures and Processes.*

1. What is Genetic research?

Genetic material, also called DNA or RNA, is usually obtained from a small blood sample. Occasionally genetic material is obtained from other sources such as saliva or biopsy specimens. (A biopsy is a tiny piece of tissue that is cut out e.g. from the skin or from a lump, to help your doctor make a diagnosis.) Genes are found in every cell in the human body. Our genes determine what we look like and sometimes what kind of diseases we may be susceptible to. Worldwide, researchers in the field of genetics are continuously discovering new information that may be of great benefit to future generations and also that may benefit people today, who suffer from particular diseases or conditions.

2. What does this particular research study involve?

This research study seeks to address the increasing problem of diabetes and cardiovascular diseases such as heart attack and stroke amongst the mixed ancestry or coloured population of South Africa. In this study we shall identify people with diabetes and those at high risk of diabetes as well as investigate the environmental and genetic risk factors that predispose some individuals to the development of diabetes and cardiovascular diseases. Examples of environmental factors include body weight, diet, and physical activity. Additionally, this project aims to investigate whether oral health is a risk factor for diabetes and cardiovascular diseases. In this study we shall investigate whether some individuals have early cardiovascular diseases by using an ultrasound machine. This project also aims to collect genetic material (blood) to analyze for certain variants and to store excess material for future research. When a large group of patients with similar diseases has been collected, meaningful research into the disease processes may become possible.

3. Why have you been invited to participate?

Our research team has previously conducted a similar research study involving the coloured community and found out that more that 18 out of 100 individuals had diabetes but did not know. We also found that some of the risk factors associated with diabetes in other populations were not necessary the same as those affecting the coloured population of South Africa. You have therefore been invited to take part in this research study to assist in establishing the risk factors for diabetes and cardiovascular diseases affecting the coloured people of South Africa.

4. What procedures will be involved in this research?

A. You will be requested to provide information about your medical history, family history and information on eating, drinking and smoking habits. Completion of the questionnaire will take no longer than 30 minutes.

B. You shall be requested to provide a record of the medication you are currently taking, therefore if you are taking chronic medication, you shall be requested to provide this to the research team to record the medication.

C. Measurement such as weight, height, waist and hip will be done.

D. Fasting Venous Blood (20ml) will be collected thereafter you will be asked to drink a glucose solution (glucose content 75g). After two hours another venous blood (10ml) will be collected. The blood will be used to determine whether you have diabetes or you are at high risk for developing diabetes.

E. The other tests that will be determined from your blood sample are: Cholesterol, triglycerides, creatine levels to assess your kidney function, liver enzymes to assess your liver, and biochemical markers for inflammation.

F. A finger prick blood sample (a drop of blood), to be taken at the same time of the first venous blood sample, may also be required from you. The finger prick blood sample will be used to test for diabetes or the risk of developing diabetes on a point-of-care test instrument. Researchers will compare the finger prick point-of-care diabetes test with that of the send away venous blood laboratory test and would be able to establish whether the point-of-care test provides the same accurate results as that of the laboratory. Point-of-care testing may in the future be used to provide fast and accurate results without the need to send blood away to a laboratory for processing. This may be of benefit to people undergoing testing for diabetes as results would be available within a few minutes.

G. The remainder of the blood sample will be used for genetic and future research studies. The serum and DNA may be stored for several years until the technology for meaningful analysis becomes available. No pharmaceutical agent (medication) will be tested in the study.

H. For oral health, research study personnel will extract wooden toothpick, flocked brush, and mouthwash saliva samples from you to test for the presence of Porphyromonas gingivalis as an indicator for periodontal disease. Flocked brush and wood toothpick sampling will involve inserting devices in the subgingival crevice between the last upper premolar and the first upper molar. The device will sweep down the anterior surface of the first upper molar with the direction of motion away from the gum to minimize any potential discomfort. Mouthwash sampling will involve rinsing with 10 ml sterile saline solution for 20 seconds.

I. Early cardiovascular diseases will be performed by means of an ultrasound machine.

J. The research team will follow up on you on a yearly basis and some of these test may be repeated. The investigators wish to follow you up for your entire life. In the unfortunate event that you are deceased during the study period. The study team will review stats SA data and/or medical records to ascertain whether the cause of death was due to diabetes or cardiovascular diseases. ***. If you do not wish to be followed up on a yearly basis and your Statistics SA and/or medical records not to be accessed in the unfortunate event that you are deceased whilst being a participant in the study, you will have an opportunity to request that it be not accessed when you sign the consent form.***

K. Radio imaging techniques will be done on consenting subjects. These include **(i)** ultra sound to assess whether you have signs of early cardiovascular diseases, **(ii)** computed tomography scan (CT-scan) to accurately assess the fat content that is dangerous for cardiovascular diseases **(iii)** Dual-energy X-ray absorptiometry (DXA) devices will be used to study the morphology of the liver. These radio imaging techniques involve radiation which can be harmful if one is exposed excessively. For this study a low dose radiation will be used for acquisition of the images thereby minimizing radiation exposure to the participant. ***. If you do not wish to undergo any of these radio imaging techniques, you will have an opportunity to decline when you sign the consent form.***

L. An eye examination will be done to test your eye vision and any other abnormalities inthe eye. For this examination, drops placed in your eyes widen (dilate) your pupils to allow the doctor to better view inside your eyes. The drops may cause your close vision to blur for a short while.

5. Are there any risks involved in genetic research?

A slight bruising might occur after blood has been drawn from the arm but this will heal quickly. After the administration of the glucose solution, you may feel nauseous and dizzy in which case you must notify the medical personnel. A medical nurse or doctor will be present on all occasions. You may also learn that you have diabetes, in which case you will be referred to your health care giver with the results for further treatment and management. If during the study it is discovered that you have changes in your genes that may lead to a serious disease, a genetic counsellor at the expense of the principal investigators will counsel you. Radio imaging techniques such as the CT-scan involves radiation which can be harmful if one is exposed excessively. For this study a low dose radiation will be used for acquisition of the images thereby minimizing radiation exposure to the participant.

6. Are there any benefits to your taking part in this study and will you get told your results?

Your personal results will be made known to you ***only if they indicate*** that you may:

- Have diabetes, thereafter, you will be referred to your local health centre or general practitioner for further investigations and treatment.
- Have a condition or predisposition to developing diabetes that is treatable or avoidable
  e.g. by a lifestyle modification.
- Need genetic counselling.

➢ **However, participants with normal results who wish to know their results are free to contact the research team and their results will be given upon written request.**

7. How long will your blood be stored and where will it be stored?

The blood samples may be stored ***indefinitely*** to accommodate new technologies that may develop. In the event that a technology is not available in South Africa to analyse your blood sample, your blood specimen may be sent to another country with the technology either now or at a later date. However, if your specimen is to be sent to another country, permission to do so will be sought from relevant bodies. Your blood specimen will be stored at the Cape Peninsula University of Technology.

8. If your blood is to be stored is there a chance that it will be used for other research?

Your blood will only be used for genetic research that is directly related to Diabetes and cardiovascular diseases. Also if the researchers wish to use your stored blood for ***additional research in this field*** they will be required to apply for permission to do so from the ethics ***Faculty of Health & Wellness Sciences of the Cape Peninsula University of Technology. If you do not wish your blood specimen to be stored after this research study is completed you will have an opportunity to request that it be discarded when you sign the consent form.***

9. **How will your confidentiality be protected?**

Your identity will be recorded once and kept confidential throughout. This is to allow the principal investigators to convey information that may be beneficial to you. Access will be limited to the principal investigators by assigning a special study code to all your data and blood samples. This means that your sample will be identified with a special study code that will remain linked to your name and contact details. However, during the entire research study, your blood specimens will be anonymised and the research staff won't be able to associate it with your name and contact details. You shall also be supplied this code so that if at anytime the investigators need to contact you, you may only identify yourself using your special code. Any scientific publications, lectures or reports resulting from the study will not identify you.

Some insurance companies may mistakenly assume that taking part in research indicates a higher risk for disease. Thus no information about you or your family will be shared with such companies.

10. Will you or the researchers benefit financially from this research?

You will not be paid to take part in this study ***although your out-of-pocket expenses may be reimbursed.*** The expenses that will be covered by the research team are those that include transportation to a hospital radiography department should you consent to radio imaging.

Important information: In the unlikely event that this research leads to the development of a commercial application or patent, you or your family will not receive any profits or royalties, but profits will be reinvested into supporting the cause of further research which may bring benefits to you or your family and to the community, such as health screening, medical treatment, educational promotions, etc.

11. Is there anything else you should know or do?

You should inform your family practitioner or usual doctor that you are taking part in a research study. You can contact

Prof T Matsha at 021 959 6366 or matshat@cput.ac.za,

If you have any further queries or encounter any problems, you can also contact the Cape Peninsula University of Technology Health and Wellness Sciences Research Ethics Committee,

Chairperson: Prof Engel-hills at 0219596570 or EngelhillsP@cput.ac.za or

You will receive a copy of this information and consent form for your own records if it is requested.

12. Declaration by participant

By signing below, I ……..…….. agree to take part in a research project that includes genetic research study entitled *(***PROGRESSIVE RESEARCH ON RISK FACTORS OF TYPE 2 DIABETES AND CARDIOVASCULAR DISEASES IN SOUTH AFRICA***)*.

I declare that:

- I have read or had read to me this information and consent form and it is written in a language with which I am fluent and comfortable.
- I have had a chance to ask questions and all my questions have been adequately answered.
- I understand that taking part in this study is voluntary and I have not been pressurised to take part.
- I have received a signed duplicate copy of this consent form for my records.

13. Tick the option you choose:

☐ I agree that my blood or tissue sample can be stored ***indefinitely*** after the project is completed but that it is anonymised with all possible links to my identity removed, and that the researchers may then use it for additional research in this or a related field. Once my sample is anonymised, my rights to that sample are waivered. My sample may be shipped to another laboratory in SA or abroad to be used in other research projects in this or a related field

OR

☐ I agree that my blood or tissue sample can be stored ***indefinitely,*** but I can choose to request at any time that my stored sample be destroyed. My sample will be identified

with a special study code that will remain linked to my name and contact details. I have the right to receive confirmation that my request has been carried out.

OR

☐ Please destroy my blood sample as soon as the current research project has been completed.

14. Tick the option you choose:

☐ I ***consent*** that the research team may follow me up for yearly check-up ***AND*** in the unfortunate event that I am deceased whilst still part of the study, I ***consent*** that the team may access Statistics SA and/or my medical records to ascertain whether the cause of my death was due to diabetes or cardiovascular diseases.

OR

☐ I ***do not consent*** to follow me up for yearly check-up ***BUT*** in the unfortunate event that I am deceased whilst still part of the study, I ***consent*** that the team may access Statistics SA and/or my medical records to ascertain whether the cause of my death was due to diabetes or cardiovascular diseases.

OR

☐ I **do not consent** to follow me up for yearly check-up ***AND*** in the unfortunate event that I am deceased whilst still part of the study, I ***do not consent*** that the team accessing Statistics SA and/or my medical records to ascertain whether the cause of my death was due to diabetes or cardiovascular diseases.

15. Tick the option you choose: Radio Imaging

☐ I ***consent*** to ultra sound techniques to assess if I have early cardiovascular diseases

☐ I ***do not consent*** to ultra sound techniques that assess if I have early cardiovascular diseases

AND

☐ I ***consent*** computed tomography scan (CT-scan) to accurately assess the fat content that is dangerous for cardiovascular diseases

☐ I ***do not consent*** to computed tomography scan (CT-scan) that accurately assess the fat content that is dangerous for cardiovascular diseases

AND

☐ I ***consent*** to Dual-energy X-ray absorptiometry (DXA) used to study body composition.

☐ I ***do not consent*** Dual-energy X-ray absorptiometry (DXA) used to study body composition

Signed at (*place*) ..... on (*date*) .....

Finger print	.....	.....
	Signature of participant	**Signature of witness**

16. Declaration by investigator

I *(name)* …… declare that:

- I explained the information in this document to ……..……...
- I encouraged him/her to ask questions and took adequate time to answer them.
- I am satisfied that he/she adequately understands all aspects of the research as discussed above.
- I did/did not use a interpreter. (*If a interpreter is used then the interpreter must sign the declaration below.*

Signed at (*place*) .....…..…….. on (*date*) ……...…….. 2016.

.....	.....
Signature of investigator	**Signature of witness**

17. Declaration By Interpreter

I *(name)* …… declare that:

- I assisted the investigator (*name*) ……. to explain the information in this document to (*name of participant*) …….. Using the language medium of Afrikaans/Xhosa.
- We encouraged him/her to ask questions and took adequate time to answer them.
- I conveyed a factually correct version of what was related to me.
- I am satisfied that the participant fully understands the content of this informed consent document and has had all his/her question satisfactorily answered.

Signed at (*place*) .....…..…….. on (*date*) ……...…….. 2016.

.....	.....
Signature of interpreter	**Signature of witness**

epigenomes

MDPI

Article

Genome-Wide DNA Methylation and LncRNA-Associated DNA Methylation in Metformin-Treated and -Untreated Diabetes

Wendy L. Solomon [1], Stanton B. E. Hector [1], Shanel Raghubeer [1], Rajiv T. Erasmus [2], Andre P. Kengne [3,4] and Tandi E. Matsha [1,*]

1 SAMRC/CPUT/Cardiometabolic Health Research Unit, Department of Biomedical Sciences, Faculty of Health and Wellness Science, Cape Peninsula University of Technology, P.O. Box 1906, Bellville, Cape Town 7530, South Africa; solomonw@cput.ac.za (W.L.S.); hectors@cput.ac.za (S.B.E.H.); shanelraghubeer@gmail.com (S.R.)

2 Division of Chemical Pathology, Faculty of Health Sciences, National Health Laboratory Service (NHLS) and University of Stellenbosch, Cape Town 7505, South Africa; rte@sun.ac.za

3 Non-Communicable Diseases Research Unit, South African Medical Research Council, Cape Town 7505, South Africa; andre.kengne@mrc.ac.za

4 Department of Medicine, University of Cape Town, Cape Town 7700, South Africa

* Correspondence: matshat@cput.ac.za; Tel.: +27-21-959-6366; Fax: +27-21-959-6760

Received: 8 July 2020; Accepted: 14 August 2020; Published: 1 September 2020

Abstract: Metformin, which is used as a first line treatment for type 2 diabetes mellitus (T2DM), has been shown to affect epigenetic patterns. In this study, we investigated the DNA methylation and potential lncRNA modifications in metformin-treated and newly diagnosed adults with T2DM. Genome-wide DNA methylation and lncRNA analysis were performed from the peripheral blood of 12 screen-detected and 12 metformin-treated T2DM individuals followed by gene ontology (GO) and KEGG pathway analysis. Differentially methylated regions (DMRs) observed showed 22 hypermethylated and 11 hypomethylated DMRs between individuals on metformin compared to screen-detected subjects. Amongst the hypomethylated DMR regions were the SLC gene family, specifically, SLC25A35 and SLC28A1. Fifty-seven lncRNA-associated DNA methylation regions included the mitochondrial ATP synthase-coupling factor 6 (ATP5J). Functional gene mapping and pathway analysis identified regions in the axon initial segment (AIS), node of Ranvier, cell periphery, cleavage furrow, cell surface furrow, and stress fiber. In conclusion, our study has identified a number of DMRs and lncRNA-associated DNA methylation regions in metformin-treated T2DM that are potential targets for therapeutic monitoring in patients with diabetes.

Keywords: metformin; DNA methylation; lncRNA; diabetes mellitus; Africa

1. Introduction

DNA methylation, the most widely studied epigenetic mechanism, involves the covalent addition of a methyl group at the 5′ position of the cytosine ring within the 5′-CpG-3′ dinucleotides to create a 5-methylcytosine (5-mC). The target of DNA methylation, catalyzed by DNA methyltransferases (DNMTs) enzymes, are CpG nucleotides, which are usually unmethylated [1]. These CpG nucleotides occur at high-frequency in the promoter regions of genes and are frequently associated with hyper- or hypomethylation events [2]. Hypermethylation of promoter CpG islands can result in suppression of gene expression, whereas hypomethylation is associated with the transcriptional activation of affected genes [3]. Various studies suggest that these modifications may alter the transcriptional activity of genes and contribute to pathogenic conditions, such as the type 2 diabetes mellitus (T2DM)

phenotype [4,5]. Studies also indicate that response to anti-diabetic agents and occurrence of diabetes complications can result from the actions of DNA methylation [4,6].

Although progression of disease cannot solely be attributed to DNA methylation, the impact of long non-coding RNAs (lncRNAs) on biological and pathologic processes have also been linked to various conditions including cancers and metabolic diseases [7,8]. LncRNAs are transcription products greater than 200 nucleotides with limited protein coding function [9]. They have been implicated in the regulation gene expression at the epigenetic, transcriptional, and post-transcription level [10]. Studies show that lncRNAs may also play a role in the diagnosis and therapeutic management of diabetes due to their involvement in regulatory processes and complications of T2DM [11,12].

Metformin, a drug commonly used for the treatment of T2DM, is highly effective with minimal side effects [13]. It has the ability to promote the phosphorylation and activation of AMP-activated protein kinase (AMPK), which results in the inhibition of gluconeogenic genes. In addition to glucose metabolism, the activation of AMPK impacts other pathways, such as lipid metabolism, mitochondrial biogenesis, autophagy, cell growth, and circadian rhythm [14]. Once activated, AMPK phosphorylates epigenetic enzymes, such as DNA methyltransferases (DNMTs), resulting in their inhibition [15]. The effects of metformin on DNA methylation include both hypo- and hypermethylation at the promoters of different genes, which in turn, could act to enhance or suppress gene expression [16–19]. Metformin was shown to affect DNA methylation even in healthy individuals immediately 10 h after drug administration [19]. These alterations in DNA methylation has also been evident in cancer related studies, showing that DNA methylation plays a role in the antidiabetic and potential anti-cancer actions of metformin [15,20–22].

Despite recent advances in the role of DNA methylation and diabetes, data on its effect in those under treatment with metformin in Africa are lacking. We, therefore, aimed to characterize the DNA methylation modifications in newly diagnosed and metformin-treated South Africans with T2DM. The knowledge gained could be used as a basis for further studies to elucidate the role of DNA methylation in the monitoring and treatment of T2DM within a South African context.

2. Results

2.1. Clinical Characteristics of the Study Population

The general clinical characteristics of the study population are summarized in Table 1. The study sample comprised 24 participants—12 screen-detected and 12 metformin-treated T2DM. There were no significant differences between the two groups in all the clinical characteristics. The duration of disease in the metformin-treated T2DM ranged from 0.5 to 17 with an average of 5.2 years.

Table 1. Clinical characteristics of the study population.

Characteristics	Screen-Detected Diabetes Mellitus $n = 12$	Known Diabetes Mellitus $n = 12$	p-Value
	Mean ± SD	Mean ± SD	
Age (years)	54.8 ± 7.5	53.2 ± 9.6	0.658
Body mass index (kg/m^2)	33.5 ± 8.9	29.4 ± 5.0	0.174
Waist circumference (cm)	101.3 ± 19.7	91.7 ± 10.5	0.150
Hip circumference (cm)	109.4 ± 16.6	103.3 ± 11.9	0.311
Waist hip ratio	0.92 ± 0.07	0.89 ± 0.05	0.195
Systolic blood pressure (mmHg)	142.9 ± 32.9	136.3 ± 25.9	0.587
Diastolic blood pressure (mmHg)	94.2 ± 22.1	83.8 ± 11.5	0.165
Fasting plasma glucose (mmol/L)	9.1 ± 3.6	11.0 ± 5.8	0.352
Post 2-h plasma glucose (mmol/L)	16.5 ± 4.71	-	-
HbA1c (%)	7.97 ± 2.58	9.33 ± 3.04	0.254
Fasting serum insulin (mIU/L)	15.3 ± 10.6	11.3 ± 7.6	0.316

Table 1. *Cont.*

Characteristics	Screen-Detected Diabetes Mellitus $n = 12$	Known Diabetes Mellitus $n = 12$	p-Value
	Mean ± SD	Mean ± SD	
Triglycerides (mmol/L)	2.18 ± 1.35	2.02 ± 0.96	0.760
Total cholesterol (mmol/L)	6.36 ± 0.85	6.10 ± 1.47	0.607
Low-density lipoprotein-cholesterol (mmol/L)	4.17 ± 0.84	4.08 ± 1.22	0.831
High-density lipoprotein-cholesterol (mmol/L)	1.38 ± 0.58	1.23 ± 0.35	0.448
Ultrasensitive C-reactive protein (mg/L)	11.1 ± 12.3	14.4 ± 12.5	0.531
Serum cotinine (ng/mL)	127.4 ± 149.5	120.7 ± 179.7	0.921
Gamma-glutamyl transferase (IU/L)	64.5 ± 57.4	43.7 ± 27.2	0.287

2.2. Differentially Methylated Regions, LncRNA-Associated DNA Methylation, Gene Ontology (GO) and Pathway Analysis

A total of 33 differentially methylated regions (DMRs) were observed between individuals on metformin treatment compared to screen-detected subjects. Of these, 22 were hypermethylated, whilst another 11 were hypomethylated in participants treated with metformin, and these are summarized in Table 2. Lnc-associated DNA methylation peaks in the promoter regions are summarized in Table 3 showing that 36 were hypermethylated, and 21 were hypomethylated in individuals on metformin. KEGG pathway analysis revealed no enriched pathways. Based on GO analyses, we retrieved the biological process, cellular process, and molecular function of the DMRs, and these are presented in Figures 1 and 2. The top enrichment scores for cellular processes of hypermethylated DMRs in subjects on metformin were associated with the axon initial segment, node of Ranvier, cell periphery, cleavage furrow, cell surface furrow, and stress fiber (Figure 1), whilst the hypomethylated biological processes were associated with photoreceptor outer segment (Figure 2).

Table 2. Differentially methylated regions (DMRs) in T2DM on metformin versus newly diagnosed cases.

Hypermethylated DMRs					
Gene Name	**Genomic Coordinates**	**DMR Length**	**log_2FC**	**p-Value**	**q-Value**
XAGE1E	chrX:52260741-52261020	279	1.66	<0.001	0.001
XAGE1B	chrX:52260741-52261020	279	1.66	<0.001	0.001
KIAA1467	chr12:13198981-13199200	219	1.62	<0.001	0.001
ASB2	chr14:94442921-94443160	239	1.6	<0.001	0.001
GABPA	chr21:27105221-27105420	199	1.56	<0.001	0.004
ZNF346	chr5:176448161-176448360	199	1.47	<0.001	0.004
FKBP8	chr19:18655321-18655520	199	1.4	<0.001	0.001
CTAGE15	chr7:143268761-143269080	319	1.38	<0.001	0.001
VIPR1	chr3:42531141-42531360	219	1.37	<0.001	0.026
TMEM204	chr16:1584901-1585100	199	1.31	<0.001	0.002
RNF103-CHMP3	chr2:86948981-86949200	219	1.31	<0.001	0.010
PARVB	chr22:44394581-44394780	199	1.26	<0.001	0.004
POTED	chr21:14980641-14980880	239	1.23	<0.001	0.001
STAG2	chrX:123095761-123095980	219	1.21	<0.001	0.008
KCNQ3	chr8:133459461-133459680	219	1.19	<0.001	0.026
TBCE	chr1:235532261-235532520	259	1.16	<0.001	0.001
GAREML	chr2:26393901-26394160	259	1.14	<0.001	0.003
SEPT12	chr16:4838741-4839000	259	1.13	<0.001	0.003
OR6C3	chr12:55727101-55727440	339	1.11	<0.001	0.011
PPP1R32	chr11:61247661-61247920	259	1.1	<0.001	0.003
ZNF169	chr9:97023241-97023440	199	1.07	<0.001	0.024
TAS1R1	chr1:6616841-6617200	359	1.07	<0.001	0.001

Table 2. *Cont.*

Hypomethylated DMRs					
TPD52L2	chr20:62497561-62497920	359	−1	<0.001	0.003
GAGE7	chrX:49217161-49217580	419	−1.15	<0.001	0.011
NUDT10	chrX:51075781-51076040	259	−1.37	<0.001	0.001
OPN1MW2	chrX:153446941-153447140	199	−1.39	<0.001	0.008
OPN1MW	chrX:153446941-153447140	199	−1.39	<0.001	0.008
BRDT	chr1:92415321-92415520	199	−1.39	<0.001	0.023
ELAC2	chr17:12919641-12919840	199	−1.45	<0.001	0.012
SLC25A35	chr17:8196461-8196680	219	−1.52	<0.001	0.004
C18orf8	chr18:21081741-21081940	199	−1.55	<0.001	0.002
SLC28A1	chr15:85429461-85429660	199	−1.67	<0.001	0.008
FBXW8	chr12:117350081-117350320	239	−1.79	<0.001	0.008

Gene name refers to the name of the DMR-associated gene. Genomic coordinates refers to the genomic locus of the DMR. DMR Length refers to the length of the DMR. log2FC refers to the fold change of normalized tag counts between two groups (log2 transformed). The *p*-value refers to the *p*-value of the DMR, the smaller, the more significant. The *q*-value refers to the Benjamini-Hochberg False Discovery Rate (BH FDR) corrected *p*-value.

Table 3. LncRNA-associated DNA methylation peaks of known diabetes versus screen-detected diabetes.

Hypermethylated					
Gene Name	**Genomic Coordinates**	**DMR Length**	**log2FC**	***p*-Value**	***q*-Value**
SLC26A9	chr1:205895421-205895620	199	1.91	<0.001	0.001
FAM223A	chrX:153859601-153859800	199	1.82	<0.001	0.001
SDK2	chr17:71432461-71432680	219	1.69	<0.001	0.001
XAGE1B	chrX:52260741-52261020	279	1.66	<0.001	0.001
SCRIB	chr8:144877441-144877660	219	1.62	<0.001	0.001
KIAA1467	chr12:13198981-13199200	219	1.62	<0.001	0.001
AK092098	chr11:63591421-63591720	299	1.58	<0.001	0.001
ATP5J	chr21:27105221-27105420	199	1.56	<0.001	0.004
AF420437	chr1:146216561-146217120	559	1.49	<0.001	0.002
ZNF346	chr5:176448161-176448360	199	1.47	<0.001	0.004
AX747590	chr8:12435501-12435760	259	1.46	<0.001	0.001
AK128525	chr2:89160101-89160340	239	1.45	<0.001	0.001
XLOC_007349	chr9:38128521-38128740	219	1.4	<0.001	0.001
FKBP8	chr19:18655321-18655520	199	1.4	<0.001	0.001
LOC101927468	chr1:147717321-147717520	199	1.38	<0.001	0.001
CTAGE15	chr7:143268761-143269080	319	1.38	<0.001	0.001
AF258560	chr16:24930681-24930880	199	1.38	<0.001	0.005
LOXL2	chr8:23190561-23190780	219	1.35	<0.001	0.003
AC016644.1	chr5:56238121-56238320	199	1.35	<0.001	0.019
RP11-14N7.2	chr1:148934661-148934860	199	1.34	<0.001	0.006
AP001476.4	chr21:47470561-47470760	199	1.34	<0.001	0.020
RP3-399L15.2	chr6:114858501-114858700	199	1.31	<0.001	0.011
AK310441	chr1:148876821-148877060	239	1.28	<0.001	0.001
RP11-423O2.7	chr1:142958401-142958660	259	1.25	<0.001	0.017
LOC101928402	chrX:123095761-123095980	219	1.21	<0.001	0.008
LINC00521	chr14:94461821-94462080	259	1.18	<0.001	0.006
TBCE	chr1:235532261-235532520	259	1.16	<0.001	0.001
SMIM22	chr16:4838741-4839000	259	1.13	<0.001	0.003
XLOC_l2_000395	chr1:142839541-142839880	339	1.11	<0.001	0.026
SEMA4C	chr2:97531721-97531960	239	1.11	<0.001	0.012
LOC101929378	chr2:157111641-157111940	299	1.08	<0.001	0.010
ZNF169	chr9:97023241-97023440	199	1.07	<0.001	0.024
GNPTG	chr16:1409001-1409360	359	1.07	<0.001	0.004
TIMELESS	chr12:56816721-56817100	379	1.05	<0.001	0.002
RP13-638C3.3	chr17:80544641-80544940	299	1.02	<0.001	0.009
XLOC_009584	chr11:123084121-123084460	339	1	<0.001	0.006

Table 3. *Cont.*

Hypomethylated					
SSH1	chr12:109199901-109200160	259	−1.12	<0.001	0.004
XLOC_005639	chr6:21980601-21980860	259	−1.13	<0.001	0.025
RP11-458D21.1	chr1:145380441-145380780	339	−1.19	<0.001	0.011
LOC100506603	chr14:77252181-77252640	459	−1.19	<0.001	0.002
AK125727	chr14:77252181-77252640	459	−1.19	<0.001	0.002
AP001476.3	chr21:47477561-47477760	199	−1.24	<0.001	0.025
BC034416	chr3:180586661-180586880	219	−1.31	<0.001	0.007
RN7SL367P	chr16:1946361-1946700	339	−1.35	<0.001	0.004
RP11-586K12.4	chr16:32752701-32752900	199	−1.37	<0.001	0.005
EIF3B	chr7:2412041-2412260	219	−1.37	<0.001	0.004
ANKIB1	chr7:91999241-91999440	199	−1.37	<0.001	0.013
XLOC_010373	chr13:45618701-45618900	199	−1.38	<0.001	0.013
RP11-510M2.5	chr16:71577621-71577820	199	−1.38	<0.001	0.004
OPN1MW	chrX:153446941-153447140	199	−1.39	<0.001	0.008
ELAC2	chr17:12919641-12919840	199	−1.45	<0.001	0.012
SLC25A35	chr17:8196461-8196680	219	−1.52	<0.001	0.004
AK095057	chr5:179268841-179269080	239	−1.53	<0.001	0.024
C18orf8	chr18:21081741-21081940	199	−1.55	<0.001	0.002
RP11-168K11.3	chr9:116382121-116382460	339	−1.62	<0.001	0.005
LL22NC03-N27C7.1	chr22:24081461-24081680	219	−1.63	<0.001	0.009
CRAMP1L	chr16:1716841-1717040	199	−1.76	<0.001	0.014

Gene name refers to the name of the DMR-associated gene. Genomic coordinates refers to the genomic locus of the DMR. DMR Length refers to the length of the DMR. log2FC refers to the fold change of normalized tag counts between two groups (log2 transformed). The *p*-value refers to the *p*-value of the DMR, the smaller, the more significant. The *q*-value refers to the Benjamini-Hochberg False Discovery Rate (BH FDR) corrected *p*-value.

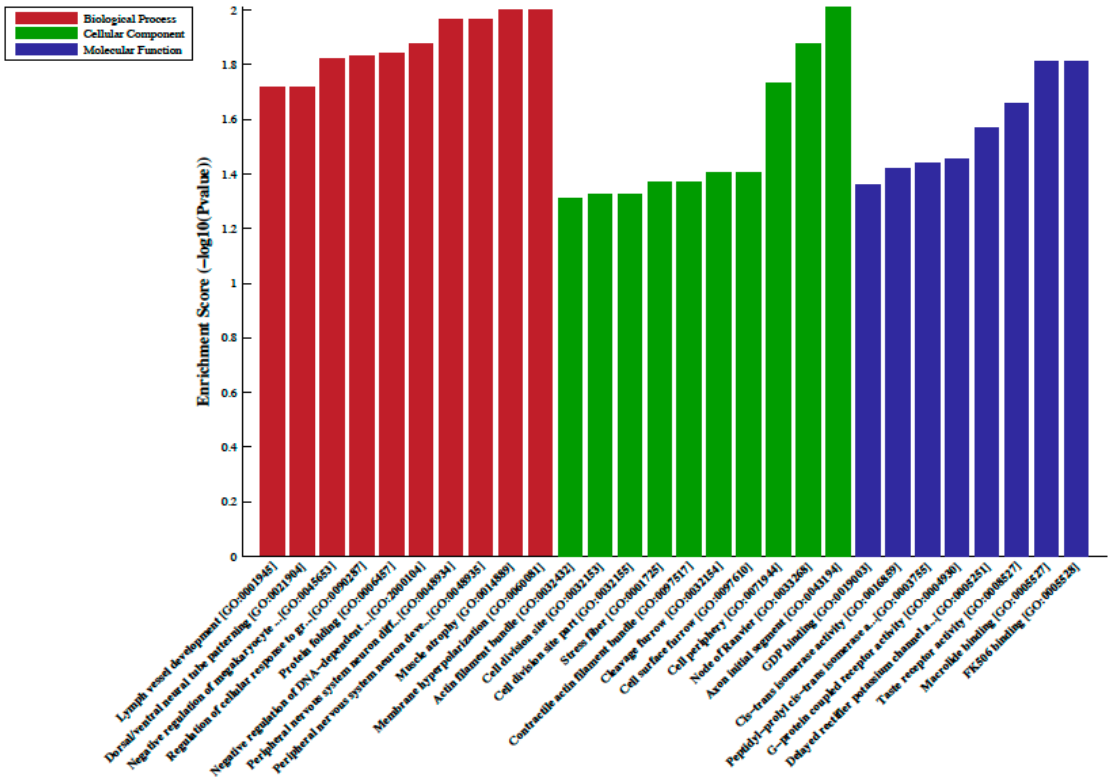


Figure 1. Gene ontology (GO) enrichment analysis of the differentially hypermethylated genes in metformin-treated diabetes. The bar plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components, and molecular function. Data are presented as enriched scores expressed as −log10 (*p* value).

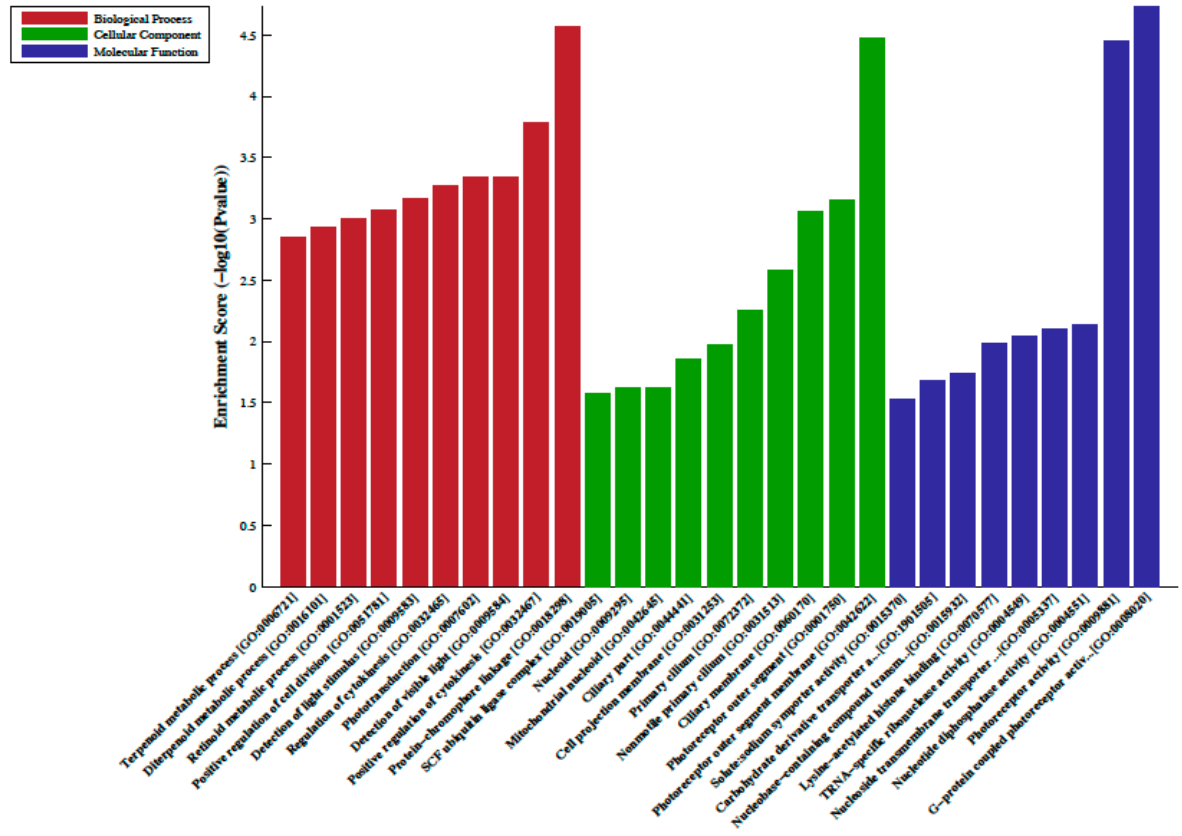


Figure 2. Gene ontology (GO) enrichment analysis of the differentially hypomethylated genes in metformin-treated diabetes. The bar plot shows the top ten enrichment score values of the significant enrichment terms. Enriched GO terms were categorized into biological processes, cellular components, and molecular function. Data are presented as enriched scores expressed as −log10 (*p* value).

3. Discussion

In this study, we measured DNA methylation in diabetic individuals on metformin treatment compared to newly diagnosed diabetes cases and found 33 differentially methylated regions (DMRs) of which 22 (67%) were hypermethylated in diabetes subjects on metformin therapy. Furthermore, 57 lncRNA-associated DNA Methylation regions (36 hypermethylated and 21 hypomethylated) were detected of which 63% were hypermethylated in metformin-treated subjects. Functional pathway analysis of these DMRs revealed that they affect gene expression in the axon initial segment (AIS), node of Ranvier, cell periphery, cleavage furrow, cell surface furrow, and stress fiber.

Amongst the hypomethylated DMRs found in this study were genes in the SLC family, specifically SLC25A35 and SLC28A1. The SLC family is known for its importance in drug development, and their proteins include passive transporters, symporters, and antiporters and are located in cellular and organelle membranes [23]. Transporters facilitate the movement of a specific substrate across the membrane with or against its concentration gradient and sequence analysis of SLC25A35 indicates that it likely functions as an oxaloacetate carrier, implying mitochondrial association [24]. On the other hand, SLC28A1, a high-affinity pyrimidine nucleoside transporter, plays a role in renal reabsorption and has been observed to be impaired during diabetes [25]. Metformin treatment has been associated with lower methylation levels in SLC transporter genes, as was shown in a study conducted on metformin transporter genes in liver tissue [18]. Mitochondrial dysfunction due to diabetes affects oxidative phosphorylation and decreases ATP production. As SLC proteins transport various solutes across the mitochondrial membrane in order to partake in a number of metabolic pathways [26,27], the decrease in methylation and subsequent increase in gene expression of SLC transporters could be indicative of the antidiabetic effect of metformin treatment. It is, therefore, likely that metformin in its

demethylation action of SLC mitochondrial carriers could possibly aid cell repair in these patients, however, this requires further investigation.

Functional pathway analysis observed in this study is consistent with the basic pathological abnormalities in Diabetic Peripheral Neuropathy (DPN), such as axonal degeneration and demyelination, lack of sensation, numbness, paresthesia, and allodynia experienced by diabetic individuals [28]. Cell death of nerves in DPN results from multifactorial metabolic imbalances associated with diabetes. The resulting mitochondrial dysfunction through a series of cascade effects involving AMP-activated protein kinase (AMPK), sirtuin (SIRT), and peroxisome proliferator-activated receptor-γ coactivator α (PGCα) suppresses mitochondrial oxidative phosphorylation, resulting in neuronal and axonal degeneration through increased oxidative injury [29,30].

Treatment with metformin was shown to decrease the incidence of DPN as was observed by the Bypass Angioplasty Revascularization Investigation 2 Diabetes trial [31]. Although metformin cannot reverse the nerve damage caused by diabetes, it could assist in managing blood glucose levels and improving the symptoms for patients.

In addition to DMRs, 57 lncRNA-associated DNA Methylation Peaks were detected when comparing known diabetic individuals to screen detected patients. Most recently the NONCODE database has updated the numbers of human lncRNAs to 167,150 with numbers still increasing [32]. Recent genome-wide association studies (GWAS) have shown positive correlation of some lncRNAs and diabetes [33]. In a related study, Sathishkumar et al. (2018) found increased levels of lncRNAs in T2DM patients, including HOTAIR, MEG3, LET, MALAT1, MIAT, CDKN2BAS1/ANRIL, XIST, PANDA, GAS5, Linc-p21, ENST00000550337.1, PLUTO, NBR2THRIL, and SALRNA1. The majority of these lncRNAs were involved in cell cycle regulation and senescence with their expression levels correlating to poor glycemic control, insulin resistance, and inflammation [11]. Similarly, HECTD4 and MBTPS1 were identified as the target genes for lncRNAs ENST00000364558 and ENST00000565382, respectively, with involvement in the development of T2DM by means of the lysosome and phagocytic signaling pathways [34]. Our findings indicate several novel lncRNA, including a lncRNA associated with the mitochondrial ATP synthase-coupling factor 6 (ATP5J) enzyme thought to be involved in the oxidative phosphorylation pathway [35]. Our data suggest higher methylation levels of this lncRNA in metformin-treated subjects, possibly pointing to suppression of this lncRNA allowing for ATP5J expression. Although little association was found between metformin and lncRNAs in our study, the significant novel lncRNA identified warrants further investigation to explore possible roles in type 2 diabetes.

The limitations of this study include the small sample size and the inclusion of women only; however, this allowed comparison and limited error that may result in statistical manipulation of a small sample size by sex. Furthermore, we used peripheral blood DNA to perform the genome-wide DNA methylation analysis. Epigenetic changes are believed to be organ specific; however, investigations on peripheral blood DNA have shown consistent methylation patterns with other organs [36,37]. Although the average (5.2 years) duration of disease in metformin-treated subjects was within the four to six years in which a person may have had the condition before clinical diagnosis [38], these findings should be interpreted with caution. In conclusion, our study has identified a number of DMRs and lncRNA-associated DNA methylation regions in metformin-treated T2DM that are potential targets for therapeutic monitoring in diabetes patients. However, these findings require further longitudinal study investigations that can clearly ascertain that these observations are not confounded by the duration and severity of diabetes.

4. Materials and Methods

4.1. Ethical Approval of the Study

This investigation used data from the Cape Town Vascular and Metabolic Health (VMH) study), which were approved by the Research Ethics Committees of the Cape Peninsula University of

Technology and Stellenbosch University (resp., NHREC: REC-230 408-014 and N14/01/003; approved date: 21 May 2018). The Code of Ethics of the World Medical Association (Declaration of Helsinki) was also applied to the study. Signed written consent was obtained from all participants after all procedures were explained in the language of their choice.

4.2. Study Procedures

In this case-control study, the participants were females matched for both age and body mass index. All study participants underwent a standardized interview, blood pressure, and anthropometric measurements. A 75 g oral glucose tolerance test (OGTT) was performed on participants with no previous diagnosis of diabetes mellitus. Participants who met the World Health Organisation (WHO) criteria for diabetes were termed as screen-detected or newly diagnosed diabetes. Biochemical parameters analyzed at an ISO 15189 accredited pathology practice (PathCare, Reference Laboratory, Cape Town, South Africa) included the following: plasma glucose, serum insulin, serum creatinine, total cholesterol (TC), high-density lipoprotein cholesterol (HDL-c), triglycerides (TG), low-density lipoprotein cholesterol (LDL), C-reactive protein (CRP), γ-glutamyl transferase (GGT), AST, ALT, and glycated hemoglobin (HbA1c), certified by the National Glycohemoglobin Standardization Program (NGSP). In addition, a full blood count was also done for all participants, and ethylenediaminetetraacetic acid (EDTA) treated blood samples were stored at −20 degrees Celsius for DNA extraction and analysis.

4.3. Genome-Wide DNA Methylation Sequencing

Genomic DNA was extracted from peripheral blood using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. At least 2 μg of DNA (concentrations ranging between 70 and 130 ng/μL) with A260/A280 and A260/A230 ratios ≥ 1.8 was shipped frozen on dry ice, as instructed by Arraystar Inc. (Rockville, MD, USA). Methylated DNA immunoprecipitation (MeDIP) was performed by Arraystar Inc. (Rockville, MD, USA) according to Down et al. [39], with minor modifications as follows. About 1 μg of fragmented DNA was prepared for Illumina HiSeq 4000 sequencing as the following steps: (1) end repair of DNA samples with T4 DNA polymerase, Klenow DNA polymerase, and T4 PNK; (2) a single 'A' base was added to the 3' ends with Klenow (exo minus) polymerase; (3) Illumina's genomic adapters were ligated to DNA fragments; (4) DNA fragments were immunoprecipitated by anti-5-methylcytosine antibody (Diagenode); (5) immunoprecipitated DNA fragments were amplified by PCR amplification; (6) ~300–600 bp DNA fragments were extracted by gel purification. The completed libraries were quantified by Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The libraries were denatured with 0.1 M NaOH to generate single-stranded DNA molecules, captured on Illumina flow cell, amplified in situ. The libraries were then sequenced on the Illumina HiSeq 4000 following the TruSeq SBS Kit v5 protocol. The enrichment of DNA immunoprecipitation was analyzed by qPCR using specific methylated sites at H19 locus and non-methylated sites at GAPDH.

4.4. MeDIP-Seq Data Analysis

The enrichment of DNA immunoprecipitation was analyzed by qPCR using specific methylated sites at H19 locus and non-methylated sites at GAPDH. Image analysis and base calling were performed using Off-Line Basecaller software (OLB V1.8). After passing a Solexa CHASTITY quality filter, the clean reads were aligned to the human genome (UCSC HG19) using HISAT2 software (V2.1.0). Briefly, individual bases generated from original image files have quality scores, which reflect the probability whether base calling is correct or not. The score is calculated by CHASTITY Formula. The CHASTITY (C) of each base in the short reads is determined by the intensity of four colors (IA, IC, IG, and IT here), and the formula means "the ratio of the highest (IC here) of the four (base type) intensities to the sum of highest two (IC and IG here)." The CHASTITY (C) should be no less than 0.6 in the first 25 bases. Statistically significant MeDIP-enriched regions (peaks) detected by MACS v2 were identified by comparison to input background, using a q-value threshold of 10^{-5}. The peaks in samples were

annotated by the nearest gene using the newest UCSC RefSeq database. Differentially methylated regions (DMRs) located within gene promoters (TSS − 2000 bp, TSS + 2000 bp) with statistical significance between the two groups were identified by diffReps (Cut-off: log2FC = 1.0, p-value = 10^{-4}).

4.5. Gene Ontology (GO) and KEGG Pathway Analysis

The ontology covers three domains, namely biological process, cellular component, and molecular function. Fisher's exact test was used to determine whether there was more overlap between the DE list and the GO annotation list than would be expected by chance. The p value denotes the significance of GO terms enrichment in the DE genes. The lower the p value, the more significant the GO term; a p value ≤ 0.05 was considered significant. Annotation was performed using standard workflow according to http://geneontology.org/. Pathway analysis was done using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The p value (EASE score, Fisher's p-value, or hypergeometric p-value) denotes the significance of the pathway correlated to the conditions. The lower the p value is, the more significant the pathway is; a p value ≤ 0.05 was considered significant.

Author Contributions: W.L.S.: wrote the first draft, experimental procedures, data analysis, and interpretation. S.B.E.H.: co-drafted the manuscript, statistical analysis, and interpretation of data. S.R.: interpretation of data, editing and revising it for intellectual content. R.T.E.: conception, interpretation of the data, revising it for intellectual content and final approval of the version to be published. A.P.K.: conception, interpretation of the data, revising it for intellectual content and final approval of the version to be published. T.E.M.: conception and design of the study, analysis and interpretation of the data, revising it for intellectual content and final approval of the version to be published. All authors have read and agreed to the published version of the manuscript.

Funding: This research project was supported by a grant from the South African Medical Research Council (SAMRC), with funds from National Treasury under its Economic Competitiveness and Support Package (MRC-RFA-UFSP-01-2013/VMH Study), South African National Research Foundation (SANRF) (Grant no. 115450). Any opinions, findings, conclusions, or recommendations expressed in this article are those of the author(s), and the SAMRC and/or SANRF do not accept any liability in this regard.

Acknowledgments: We thank the Bellville South community and their community Health Forum for supporting the study. T.E.M. takes full responsibility for the work as a whole, including the study design, access to data, and the decision to submit and publish the manuscript.

Conflicts of Interest: The authors declare no conflict of interests.

References

1. Deaton, A.; Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **2011**, *25*, 1010–1022. [CrossRef] [PubMed]
2. Du, Q.; Luu, P.L.; Stirzaker, C.; Clark, S.J. Methyl-CpG-binding domain proteins: Readers of the epigenome. *Epigenomics* **2015**, *7*, 1051–1073. [CrossRef] [PubMed]
3. Pasquier, J.; Hoarau-Vechot, J.; Fakhro, K.; Rafii, A.; Abi Khalil, C. Epigenetics and Cardiovascular Disease in Diabetes. *Curr. Diab. Rep.* **2015**, *15*, 1–12. [CrossRef] [PubMed]
4. Bansal, A.; Pinney, S.E. DNA methylation and its role in the pathogenesis of diabetes. *Pediatr. Diabetes* **2017**, *18*, 167–177. [CrossRef]
5. Muka, T.; Nano, J.; Voortman, T.; Braun, K.V.E.; Ligthart, S.; Stranges, S.; Bramer, W.M.; Troup, J.; Chowdhury, R.; Dehghan, A.; et al. The role of global and regional DNA methylation and histone modifications in glycemic traits and type 2 diabetes: A systematic review. *Nutr. Metab. Cardiovasc. Dis.* **2016**, *26*, 553–566. [CrossRef]
6. Ronn, T.; Ling, C. DNA methylation as a diagnostic and therapeutic target in the battle against Type 2 diabetes. *Epigenomics* **2015**, *7*, 451–460. [CrossRef]
7. Li, Y.; Xu, K.; Xu, K.; Chen, S.; Cao, Y.; Zhan, H. Roles of Identified Long Noncoding RNA in Diabetic Nephropathy. *J. Diabetes Res.* **2019**, *2019*, 1–8. [CrossRef]
8. Zhang, Y.; Jiang, X.; Wu, Z.; Hu, D.; Jia, J.; Guo, J.; Tang, T.; Yao, J.; Liu, H.; Tang, H. Long noncoding RNA LINC00467 promotes glioma progression through inhibiting p53 expression via binding to DNMT1. *J. Cancer* **2020**, *11*, 2935–2944. [CrossRef]

9. Huang, Y.; Li, J.; Chen, S.; Zhao, S.; Huang, J.; Zhou, J.; Xu, Y. Identification of Potential Therapeutic Targets and Pathways of Liraglutide Against Type 2 Diabetes Mellitus (T2DM) Based on Long Non-Coding RNA (lncRNA) Sequencing. *Med. Sci. Monit.* **2020**, *26*, e922210. [CrossRef]
10. Zhang, W.; Zheng, J.; Hu, X.; Chen, L. Dysregulated expression of long noncoding RNAs serves as diagnostic biomarkers of type 2 diabetes mellitus. *Endocrine* **2019**, *65*, 494–503. [CrossRef]
11. Sathishkumar, C.; Prabu, P.; Mohan, V.; Balasubramanyam, M. Linking a role of lncRNAs (long non-coding RNAs) with insulin resistance, accelerated senescence, and inflammation in patients with type 2 diabetes. *Hum. Genomics* **2018**, *12*, 1–9. [CrossRef] [PubMed]
12. Leti, F.; DiStefano, J.K. Long noncoding RNAs as diagnostic and therapeutic targets in type 2 diabetes and related complications. *Genes* **2017**, *8*, 207. [CrossRef] [PubMed]
13. Sanchez-Rangel, E.; Inzucchi, S.E. Metformin: Clinical use in type 2 diabetes. *Diabetologia* **2017**, *60*, 1586–1593. [CrossRef] [PubMed]
14. Priya, G.; Kalra, S. Metformin in the management of diabetes during pregnancy and lactation. *Drugs Context* **2018**, *7*, 1–21. [CrossRef] [PubMed]
15. Bridgeman, S.C.; Ellison, G.C.; Melton, P.E.; Newsholme, P.; Mamotte, C.D.S. Epigenetic effects of metformin: From molecular mechanisms to clinical implications. *Diabetes Obes. Metab.* **2018**, *20*, 1553–1562. [CrossRef] [PubMed]
16. Zhong, T.; Men, Y.; Lu, L.; Geng, T.; Zhou, J.; Mitsuhashi, A.; Shozu, M.; Maihle, N.J.; Carmichael, G.G.; Taylor, H.S.; et al. Metformin alters DNA methylation genome-wide via the H19/SAHH axis. *Oncogene* **2017**, *36*, 2345–2354. [CrossRef]
17. Ishikawa, K.; Tsunekawa, S.; Ikeniwa, M.; Izumoto, T.; Iida, A.; Ogata, H.; Uenishi, E.; Seino, Y.; Ozaki, N.; Sugimura, Y.; et al. Long-term pancreatic beta cell exposure to high levels of glucose but not palmitate induces DNA methylation within the insulin gene promoter and represses transcriptional activity. *PLoS ONE* **2015**, *10*, e0115350. [CrossRef]
18. García-Calzón, S.; Perfilyev, A.; Männistö, V.; de Mello, V.D.; Nilsson, E.; Pihlajamäki, J.; Ling, C. Diabetes medication associates with DNA methylation of metformin transporter genes in the human liver. *Clin. Epigenetics* **2017**, *9*, 1–9. [CrossRef]
19. Elbere, I.; Silamikelis, I.; Ustinova, M.; Kalnina, I.; Zaharenko, L.; Peculis, R.; Konrade, I.; Ciuculete, D.M.; Zhukovsky, C.; Gudra, D.; et al. Significantly altered peripheral blood cell DNA methylation profile as a result of immediate effect of metformin use in healthy individuals. *Clin. Epigenetics* **2018**, *10*, 156. [CrossRef]
20. Banerjee, P.; Surendran, H.; Chowdhury, D.R.; Prabhakar, K.; Pal, R. Metformin mediated reversal of epithelial to mesenchymal transition is triggered by epigenetic changes in E-cadherin promoter. *J. Mol. Med.* **2016**, *94*, 1397–1409. [CrossRef]
21. Yu, X.; Mao, W.; Zhai, Y.; Tong, C.; Liu, M.; Ma, L.; Yu, X.; Li, S. Anti-tumor activity of metformin: From metabolic and epigenetic perspectives. *Oncotarget* **2017**, *8*, 5619–5628. [CrossRef] [PubMed]
22. Cuyàs, E.; Fernández-Arroyo, S.; Verdura, S.; García, R.Á.F.; Stursa, J.; Werner, L.; Blanco-González, E.; Montes-Bayón, M.; Joven, J.; Viollet, B.; et al. Metformin regulates global DNA methylation via mitochondrial one-carbon metabolism. *Oncogene* **2018**, *37*, 963–970. [CrossRef] [PubMed]
23. Zhang, Y.; Zhang, Y.; Sun, K.; Meng, Z.; Chen, L. The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. *J. Mol. Cell Biol.* **2018**, *11*, 1–13. [CrossRef] [PubMed]
24. Haitina, T.; Lindblom, J.; Renström, T.; Fredriksson, R. Fourteen novel human members of mitochondrial solute carrier family 25 (SLC25) widely expressed in the central nervous system. *Genomics* **2006**, *88*, 779–790. [CrossRef]
25. Rodríguez-Mulero, S.; Errasti-Murugarren, E.; Ballarín, J.; Felipe, A.; Doucet, A.; Casado, F.J.; Pastor-Anglada, M. Expression of concentrative nucleoside transporters SLC28 (CNT1, CNT2, and CNT3) along the rat nephron: Effect of diabetes. *Kidney Int.* **2005**, *68*, 665–672. [CrossRef]
26. Palmieri, F. The mitochondrial transporter family SLC25: Identification, properties and physiopathology. *Mol. Aspects Med.* **2013**, *34*, 465–484. [CrossRef]
27. Palmieri, F.; Monné, M. Discoveries, metabolic roles and diseases of mitochondrial carriers: A review. *Biochim. Biophys. Acta—Mol. Cell Res.* **2016**, *1863*, 2362–2378. [CrossRef]
28. Guo, K.; Elzinga, S.; Eid, S.; Figueroa-Romero, C.; Hinder, L.M.; Pacut, C.; Feldman, E.L.; Hur, J. Genome-wide DNA methylation profiling of human diabetic peripheral neuropathy in subjects with type 2 diabetes mellitus. *Epigenetics* **2019**, *14*, 766–779. [CrossRef]

29. Fernyhough, P. Mitochondrial Dysfunction in Diabetic Neuropathy: A Series of Unfortunate Metabolic Events. *Curr. Diab. Rep.* **2015**, *15*, 24–27. [CrossRef]
30. Fujimaki, S.; Kuwabara, T. Diabetes-induced dysfunction of mitochondria and stem cells in skeletal muscle and the nervous system. *Int. J. Mol. Sci.* **2017**, *18*, 2147. [CrossRef]
31. Pop-Busui, R.; Lu, J.; Brooks, M.M.; Albert, S.; Althouse, A.D.; Escobedo, J.; Green, J.; Palumbo, P.; Perkins, B.A.; Whitehouse, F.; et al. Impact of glycemic control strategies ontheprogressionofdiabeticperipheral neuropathy in the bypass angioplasty revascularization investigation 2 diabetes (BARI 2D) Cohort. *Diabetes Care* **2013**, *36*, 3208–3215. [CrossRef] [PubMed]
32. Zhao, X.Y.; Lin, J.D. Long Noncoding RNAs: A New Regulatory Code in Metabolic Control. *Trends Biochem. Sci.* **2015**, *40*, 586–596. [CrossRef] [PubMed]
33. Mirza, A.H.; Kaur, S.; Brorsson, C.A.; Pociot, F. Effects of GWAS-associated genetic variants on lncRNAs within IBD and T1D candidate loci. *PLoS ONE* **2014**, *9*, e105723. [CrossRef] [PubMed]
34. Pengyu, Z.; Yan, Y.; Xiying, F.; Maoguang, Y.; Mo, L.; Yan, C.; Hong, S.; Lijuan, W.; Xiujuan, Z.; Hanqing, C. The Differential Expression of Long Noncoding RNAs in Type 2 Diabetes Mellitus and Latent Autoimmune Diabetes in Adults. *Int. J. Endocrinol.* **2020**, *2020*, 1–12. [CrossRef] [PubMed]
35. Huang, C.; Kim, Y.; Caramori, M.L.; Moore, J.H.; Rich, S.S.; Mychaleckyj, J.C.; Walker, P.C.; Mauer, M. Diabetic nephropathy is associated with gene expression levels of oxidative phosphorylation and related pathways. *Diabetes* **2006**, *55*, 1826–1831. [CrossRef] [PubMed]
36. Farré, P.; Jones, M.J.; Meaney, M.J.; Emberly, E.; Turecki, G.; Kobor, M.S. Concordant and discordant DNA methylation signatures of aging in human blood and brain. *Epigenetics Chromatin* **2015**, *8*, 1–17. [CrossRef]
37. Crujeiras, A.B.; Diaz-Lagares, A.; Sandoval, J.; Milagro, F.I.; Navas-Carretero, S.; Carreira, M.C.; Gomez, A.; Hervas, D.; Monteiro, M.P.; Casanueva, F.F.; et al. DNA methylation map in circulating leukocytes mirrors subcutaneous adipose tissue methylation pattern: A genome-wide analysis from non-obese and obese patients. *Sci. Rep.* **2017**, *7*, 41903. [CrossRef]
38. Porta, M.; Curletto, G.; Cipullo, D.; De la Longrais, R.R.; Trento, M.; Passera, P.; Taulaigo, A.V.; Di Miceli, S.; Cenci, A.; Dalmasso, P.; et al. Estimating the Delay Between Onset and Diagnosis of Type 2 Diabetes From the Time Course of Retinopathy Prevalence. *Diabetes Care* **2014**, *37*, 1668–1674. [CrossRef]
39. Down, T.A.; Rakyan, V.K.; Turner, D.J.; Flicek, P.; Li, H.; Kulesha, E.; Gräf, S.; Johnson, N.; Herrero, J.; Tomazou, E.M.; et al. A Bayesian deconvolution strategy for immunoprecipitation- based DNA methylome analysis. *Nat. Biotechnol.* **2008**, *26*, 779–785. [CrossRef]