MicroRNA expression profiling in peripheral blood mononuclear cells and serum of Type 2 diabetic, Pre-diabetic and Normo-glycaemic individuals

by
Stephanie Dias

Dissertation presented for the degree of *Master of Science in Molecular Biology in the Faculty of Medicine and Health Sciences at Stellenbosch University*

Supervisor: Dr Carmen Pheiffer
Co-supervisor: Dr Sian Hemmings

March 2016

Declaration

By submitting this thesis/dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2016

Abstract

MicroRNAs (miRNAs) are small non-coding RNAs that play a fundamental role in cellular function by regulating messenger RNA gene expression. Alterations in miRNA expression are implicated in metabolic dysregulation, with several studies reporting the involvement of miRNAs in the pathophysiology of Type 2 diabetes (T2D). Recently, circulating miRNAs have attracted considerable interest as biomarkers to identify individuals at risk for T2D, thus we hypothesised that circulating miRNA could be used as markers for T2D progression. The aim of this study was to determine whether miRNA expression profiles differ between diabetic, pre-diabetic and normo-glycaemic individuals.

Individuals were recruited from local communities and classified as diabetic, pre-diabetic or normo-glycaemic according to World Health Organization criteria, whereafter miRNAs were extracted from peripheral blood mononuclear cells (PBMCs) and serum of age-, gender-, ethnicity- and BMI-matched diabetic (n=4), pre-diabetic (n=4) and normo-glycaemic (n=4) individuals. MiRNAs extracted from PBMCs were sequenced using the Illumina HiSeq 2500 platform, and validated by quantitative real time PCR (qRT-PCR) in PBMCs and serum of these individuals. Moreover, bioinformatics was conducted using various target prediction programs (TargetScan, DIANA and PITA) and the DAVID functional gene annotation tool to assign biological significance to the differentially expressed miRNAs identified by sequencing.

Sequencing showed that 267 (pre-diabetics vs. normo-glycaemics), 277 (diabetics vs. normo-glycaemics) and 267 (pre-diabetics vs. diabetics) miRNAs were differentially expressed between groups. Of these, five differentially expressed miRNAs (miR-27b, miR-379, miR-21, miR-98 and miR-143) were selected for validation by qRT-PCR in PBMCs. Only miR-143 and miR-27b were significantly differentially expressed using qRT-PCR, although the results for miR-143 were different compared to the sequencing data. MiR-143 was upregulated in pre-diabetics compared to normo-glycaemic individuals (1.40-fold, $p \leq 0.01$), whereas sequencing showed upregulation of miR-143 in diabetics compared to pre-diabetics (1.75-fold, $p \leq 0.05$). The differential expression of miR-27b was consistent between qRT-PCR (1.55-fold; $p=0.07$) and sequencing (1.15-fold; $p<0.01$), where both methods

showed upregulation in pre-diabetics compared to normo-glycaemic individuals. The expression of miR-27b was similarly upregulated in serum of pre-diabetics compared to normo-glycaemic individuals (2.0-fold; $p \leq 0.05$). Furthermore, five novel miRNAs identified by sequencing were successfully validated in PBMCs of diabetic, pre-diabetic and normo-glycaemic individual.

Sequencing and qRT-PCR showed that miR-27b was upregulated in PBMCs and serum of pre-diabetics compared to normo-glycaemic individuals. Bioinformatics identified peroxisome proliferator-activated receptor gamma (*Pparg)* as a target for miR-27b. PPARG is an insulin sensitizing agent, thus we speculate that increased miR-27b expression in pre-diabetes suppresses *Pparg*, thereby inhibiting insulin signaling and subsequently decreasing glucose uptake. The increased insulin and glucose levels observed in the pre-diabetic individuals support this idea, although further work is required to confirm this hypothesis.

In conclusion, we showed that miRNA profiles differ during T2D progression, and are able to discriminate between diabetic, pre-diabetic and normo-glycaemic individuals. To our knowledge, this is the first study to report differential expression of miR-27b during T2D, suggesting its potential as a biomarker that could be incorporated into predictive models for the identification of high risk individuals. However, miRNA profiling in a larger sample size and prospective longitudinal studies are required to assess clinical applicability.

Words: 500

Uitreksel

MikroRNAs (miRNAs) is klein nie-koderende RNAs wat 'n fundamentele rol in sellulêre funksie speel deur regulering van boodskapperRNA geenuitdrukking. Verskeie studies ïmpliseer veranderings in miRNA ekspresie met metaboliese disregulering en in die patofisiologie van Tipe 2-diabetes (T2D). Onlangs het sirkulerende miRNAs groot belangstelling uitgelok as biomerkers om individue te identifiseer wat ‘n verhoogde risiko vir T2D het. Ons hipotese stel dus voor dat sirkulerende miRNA gebruik kan word as merkers vir T2D siekteprogressie. Die doel van hierdie studie was om vas te stel of miRNA geenuitdrukkings profiele verskil tussen diabete, prediabete en normoglisemiese individue.

Individue wat uit plaaslike gemeenskappe gewerf is, is volgens die Wêreld Gesondheid Organisasie riglyne geklassifiseer as diabete, pre diabete of normoglisemiese individue. Hierna is miRNAs uit die perifere bloed mononukleêreselle (PBMS) en serum van ouderdom, geslag, etniesiteit en liggaamsmassa-indeks vergelykbare diabete (n=4), prediabete (n=4) en normoglisemiese individue (n=4), geïsoleer. Die geenvolgordebepaling van die geïsoleerde miRNAs is bepaal deur ‘n Illumina HiSeq 2500 platform, en bevestig deur kwantitatiewe “real time PCR” (qRT-PCR). Verder, is bioinformatika uitgevoer met behulp van verskeie teikenvoorspellings programme (TargetScan, Diana en PITA) asook David se funksionele geenannotasie instrument om biologiese betekenis aan die differensieel uitgedrukte miRNAs, te koppel.

Geenvolgordebepaling het getoon dat 267 (prediabete vs. normoglisemies), 277 (diabete vs. normoglisemies) and 267 (prediabete vs. diabete) miRNAs differensieel uitgedruk word. Hiervan is vyf differensieel uitgedrukte miRNAs (miR-27b, miR-379, miR-21, miR-98 en miR-143) gekies vir bevestiging deur qRT-PCR in PBMS. MiR-143 en miR-27b differensiasie was deur qRT-PCR bevestig, hoewel die qRT-PCR resultate vir miR-143 verskil het met die geenvolgordebepaling data. Met qRT-PCR is miR-143 opgereguleer in die prediabete teenoor normoglisemiese individue (1,40-voudig, $p \leq 0.01$), terwyl met geenvolgordebepaling miR-143 in diabete teenoor prediabete (1,75-voudig, $p \leq 0.05$) opgereguleer was. Daar was ooreenstemming in die differensiële uitdrukking van miR-27b tussen die qRT-PCR (1,55-voudig; $p=0,07$) en geenvolgordebepaling (1,15-voudig; $p<0,01$), waar albei metodes

opregulering gewys het in die prediabete teenoor normoglisemiese individue. In die serum monsters was die uitdrukking van miR-27b soortgelyk opgereguleer in prediabete (2,0-voudig; p≤0.05). Verder is vyf unieke miRNAs geïdentifiseer deur geenvolgordebepaling wat suksesvol bevestig is in PBMS van diabete en prediabete.

Bioinformatika het *Pparg* geïdentifiseer as 'n teiken vir miR-27b. PPARG is 'n insuliensensiterings agent, dus spekuleer ons dat hoër miR-27b ekspresie, in prediabete *Pparg* onderdruk, wat die insuliensein demp en tot verlaagde glukose opname lei. Die verhoogde insulien en glukose vlakke wat in prediabete voorkom ondersteun hierdie idee, alhoewel verdere werk nodig is om hierdie hipotese te bevestig.

Ten slotte, het ons getoon dat miRNA profiele tydens die T2D siekteprogressie verskil, en in staat is om tussen diabete, prediabete en normoglisemiese individue te diskrimineer. Tot ons kennis, dit is die eerste studie wat differensiele uitgedrukking van miR-27b in T2D rapporteer, en die potensiële toepassing as 'n nie-indringende biomerker uitwys. Dit kan moontlik in voorspellende modelle geïnkorporeer kan word vir die identifisering van hoë risiko individue. Maar verdere studies met groter monster getalle en prospektiewe longitudinale studies is nodig om die kliniese toepaslikheid te evalueer.

Words: 488

Acknowledgments

I wish to express my sincere gratitude and appreciation to my supervisor and co-supervisor, Dr C. Pheiffer and Dr S. Hemmings for their outstanding supervision, motivation and guidance during the course of my study.

I am grateful to Carmen for providing me with the foundation and skills necessary to pursue new challenges. Her unequivocal support, patience and assistance were instrumental for this study to be a success.

I am grateful to Sian for always being available to assist me no matter the timing or situation.

I would like to express my gratitude to Dr J. Louw and Dr C. Muller for believing in me and for giving me the opportunity to complete my MSc, and pursue a career in science. Also, a special thanks to Dr C. Muller for the translation of my abstract.

Thank you to my fellow colleagues in the Biomedical Research and Innovation Platform for their continuous positive attitudes and willingness to help when needed.

Thank you to the South African Medical Research Council, Cannon Collins, the Ernst & Ethyl Eriksen Trust, Harry Crossley and Whitehead Scientific for the generous financial support.

Finally, a special thank you to my parents Charmaine and Stephen for their unwavering love, support and inspiration throughout this challenging, yet exciting year. I would not be where I am today if it wasn't for you. Also, to Jason, thank you for your unconditional love and understanding, and for being the most dependable person in my life. I am indebted mostly to you.

Table of Contents

List of Figures

Figure 4.1 Schematic representation of the insulin signaling pathway and a possible role of miR-27b.The peroxisome proliferator-activated receptor gamma (PPARG) stimulates the insulin signaling pathway, which leads to the translocation of glucose transporter 4 (GLUT4) vesicle to the plasma membrane, allowing glucose uptake into the cell. The negative effect of miR-27b on PPARG causes dysregulation of insulin signaling proteins (insulin receptor substrate 1 (IRS-1) and kinase/serine-threonine protein kinase-1 (PI3K/AKT)), thereby preventing glucose uptake.

List of Tables

List of equations

Nomenclature

ADA:	American Diabetes Association
Ago1-4:	Argonaute
AIDS:	Acquired Immune Deficiency Syndrome
AKT-1:	Serine-threonine protein kinase-1
ANOVA:	One-way analysis of variance
BMI:	Body mass index
Cacna1a:	Calcium channel voltage-dependent alpha 1A subunit
cDNA:	Complementary DNA
$C_{t:}$	Cycle threshold
CVD:	Cardiovascular disease
DAVID:	Database for Annotation, Visualization and Integrated Discovery
DIANA:	DNA Intelligent Analysis
DM:	Diabetes mellitus
DNA:	Deoxyribonucleic acid
ELISA:	Enzyme-linked immunosorbent assay
ETDA:	Ethylenediaminetetraacetic acid
Fgf1:	Fibroblast growth factor 1
Flt4:	Fms-related tyrosine kinase 4
FltT1:	Fms-related tyrosine kinase 1
FPG:	Fasting plasma glucose
GDM:	Gestational diabetes mellitus

GLUT4:	Glucose transporter type 4
HbA1c:	Glycated haemoglobin A1c
HDL:	High-density-lipoprotein
HIV:	Human Immunodeficiency Virus
HMDD:	Human miRNA Disease Database
HTS:	High throughput sequencing
IDDM:	Insulin dependent diabetes mellitus
IDF:	International Diabetes Federation
IFG:	Impaired fasting glucose
Igf1r:	Insulin-like growth factor 1 receptor
IGT:	Impaired glucose tolerance
IL:	Interleukin
Insr:	Insulin receptor
IR:	Insulin resistance
IRS1/2:	Insulin receptor substrate 1/2
IRS1:	Repression of insulin receptor substrate 1
Kdr:	Kinase insert domain receptor
KEGG:	Kyoto Encyclopedia of Genes and Genomes
KRAS:	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
LDL:	Low-density-lipoprotein
Map4k3:	Mitogen-activated protein kinase 3
MicroRNA:	MiRNA
MODY:	Maturity onset diabetes of young
mRNA:	Messenger RNA

mTOR:	Mechanistic target of rapamycin
NFk:	Nuclear factor kB
NFQ:	Non-fluorescent quencher
NGS:	Next generation sequencing
NIDDM:	Non-insulin dependent diabetes mellitus
NPM1:	Nucleophosphin 1
Nrp2:	Neuropilin 2
nt:	Nucleotide
OGTT:	Oral glucose tolerance
PBMCs:	Peripheral blood mononuclear cells
PI3K:	Phosphatidylinositol 3-kinase
PITA:	Probability of Interaction by Target Accessibility
PKC:	Protein kinase C
Pkrce:	PRKCE protein kinase C, epsilon
Pparg:	Peroxisome proliferator-activated receptor gamma
qRT-PCR:	Quantitative real-time polymerase chain reaction
Q-scores:	Phred quality scoring
RBP:	Ribonucleoproteins
RIN:	RNA integrity number
RNA:	Ribonucleic acid
RPG:	Random plasma glucose concentration
Scd:	Stearoyl-CoA desaturase (delta-9-desaturase)
SEM:	Standard error of the mean
SNP:	Single nucleotide polymorphism

STRING:	Search Tool for the Retrieval of Interacting Genes/Proteins
T1D:	Type 1 diabetes
T2D:	Type 2 diabetes
T2D-db:	Type 2 diabetes database
Taok2:	TAO kinase 2
TargetScan:	TargetScan-Human
TNFα:	Tumour necrosis factor alpha
TRBP:	Transactivation-responsive RNA binding protein
Tsc1:	Tuberous sclerosis 1
Ulk2:	Unc-51-like kinase
UTR:	Untranslated region
Vegfc :	Vascular endothelial growth factor C
WHO:	World Health Organization
α-cell:	Alpha-cell
β-cell:	Beta-cell

1. INTRODUCTION

1.1 The global burden of Diabetes mellitus

Diabetes mellitus (DM) is characterized by persistent hyperglycaemia (fasting blood glucose ≥7.0 mmol/L), caused by defects in insulin secretion by pancreatic beta (β)-cells and/or insulin action in peripheral tissue such as, skeletal muscle, liver and adipocytes (Fernandez-Valverde *et al.*, 2011; Tripathy and Chavez, 2010). According to the International Diabetes Federation (IDF), approximately 382 million people worldwide had diabetes in 2013 and this number is expected to increase to more than 592 million people by 2035 (Fig.1.1) (IDF, 2013). The prevalence of DM in the African region is projected to increase from 19.8 million cases in 2013 to 41.5 million cases in 2035, representing an approximately 109% increase in people afflicted with the disease (Fig.1.1). Conservative estimates indicate that South Africa currently has the fifth highest prevalence (9.3%) of DM in Africa (IDF, 2013), which vary between ethnic groups and regions (Erasmus *et al.*, 2012). In 2005, the South African Medical Research Council conducted a study on chronic diseases of lifestyle in South Africa between 1995 and 2005. This study revealed that the highest prevalence of DM was among the Indian community at 8.5% and 11.5%, followed by the mixed ancestry community with a prevalence of 3.1% and 5.8% for men and women, respectively (Goedecke *et al.*, 2005). The lowest prevalence of DM was observed among males and females in the Northwest province of South Africa at 0.9% and 1.1%, respectively (Goedecke *et al.*, 2005). However, recent studies have shown an increase in the prevalence of T2D in the mixed ancestry community in the Western Cape of South Africa (Erasmus *et al.*, 2012).

The major types of DM include Type 1 DM (T1D), Type 2 DM (T2D), Gestational DM and Maturity onset diabetes of young (MODY) (American Diabetes Association 2014). Type 2 diabetes (T2D), also called non-insulin dependent diabetes (NIDDM) or adult-onset diabetes, is the most common form of diabetes accounting for approximately 90% of all cases worldwide, while T1D, gestational DM and MODY make up the remaining 10% (ADA, 2014; Butt and Swaminathan, 2015). Type 1 diabetes mellitus is referred to as insulin-dependent diabetes mellitus (IDDM), and occurs due to the inability of the pancreas to secrete insulin due to β-cell destruction, thus, requires insulin to maintain normo-glycaemia. Gestational DM is defined as glucose intolerance first diagnosed during pregnancy, and it is estimated that approximately 5-10% of all pregnancies are complicated by hyperglycaemia (Gunderson *et*

al., 2014). Most cases of GDM resolve after delivery, however, these women have a greater than 70% lifetime risk of developing T2D (Ratner, 2007). Maturity onset diabetes of the young (MODY) occurs due to an autosomal dominant genetic mutation, and is characterized by impaired insulin secretion, with little or no defects in insulin action (Shields *et al.*, 2010).

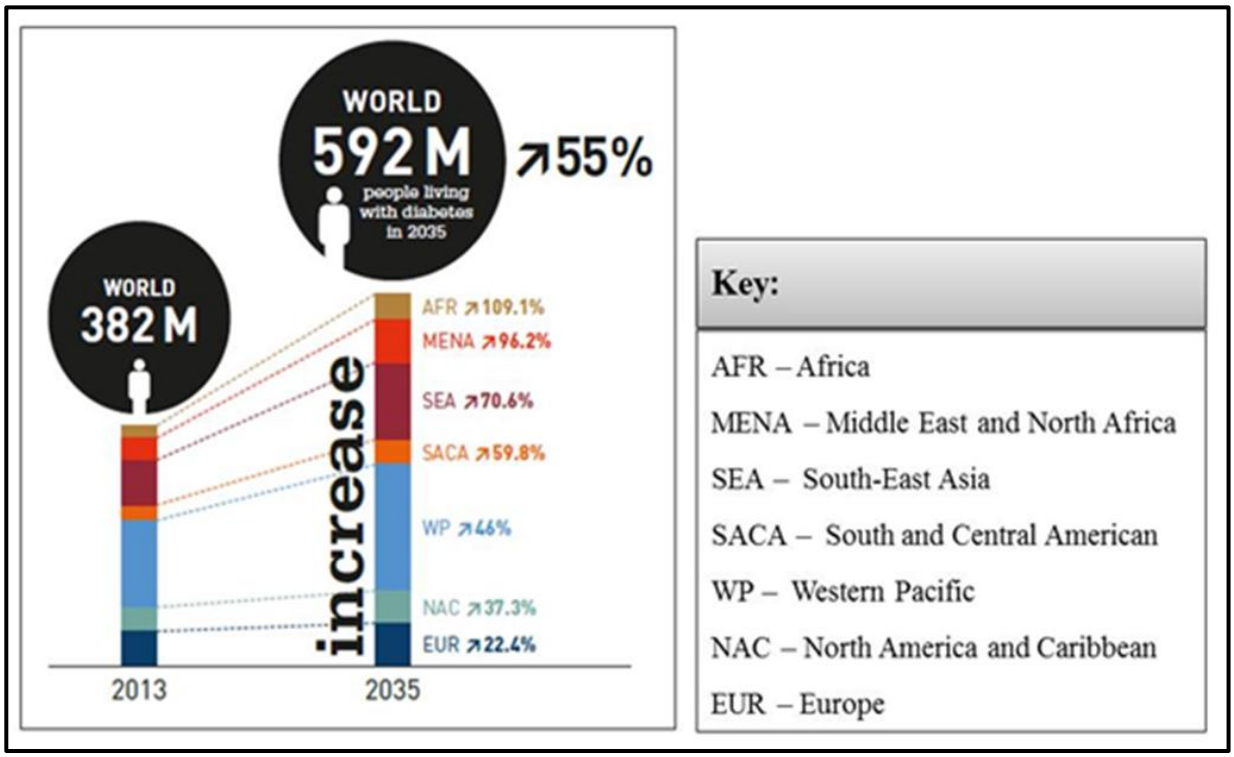


Figure 1.1 The global prevalence of diabetes in 2013 and the predicted global prevalence in 2035 (Adapted from: IDF, 2013).

1.2 Aetiology of Type 2 diabetes

Type 2 diabetes is a complex, multifactorial disease involving the interplay of many risk factors. These include, amongst others, genetics, lifestyle, nutrition and lack of physical activity (Fig. 1.2) (Hu, 2011; Wild *et al.*, 2004; Zimmet *et al.*, 2014).

Over 70 susceptibility loci have been identified for T2D (Sun *et al.*, 2014b), however, these account for only approximately 5-10% of all cases, suggesting that the increasing prevalence of T2D is not driven by genetic factors (McCarthy and Menzel, 2001). Indeed, several lines of evidence have suggested that an unhealthy diet, high in fats and sugars, together with a sedentary lifestyle are the main contributors of the current T2D pandemic (Wing *et al.*, 2001).

Reports that the increased rates of T2D in developing countries, are due to, amongst others, increased prosperity, urbanization and a shift towards a "westernized lifestyle" (Ostbye *et al.*, 1989), confirm the importance of diet and physical activity in the development of T2D. The "westernized lifestyle" is characterized by high caloric intake (Hu, 2011; Popkin, 1999; Popkin and Gordon-Larsen, 2004) that is associated with highly processed and refined foods, which contain high levels of salt, sugars and fats (Odermatt, 2011), a sedentary lifestyle (Popkin, 1999), smoking, and alcohol consumption, (Hu, 2011), among others.

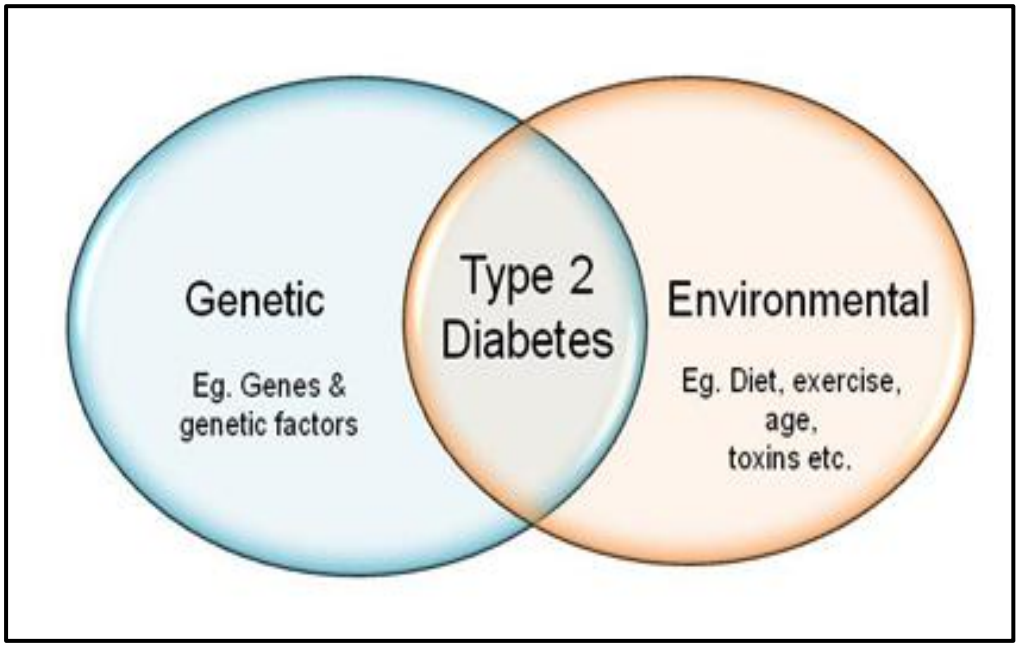


Figure 1.2 Interaction between genetic and environmental factors that contribute to the development of T2D.

1.3 Major metabolic mechanisms that characterize Type 2 diabetes

As discussed previously T2D is a complex disease underpinned by a number of biological mechanisms. These mechanisms primarily include impaired insulin action, β-cell dysfunction, increased endogenous glucose output and obesity. These will be discussed below.

1.3.1 Insulin resistance

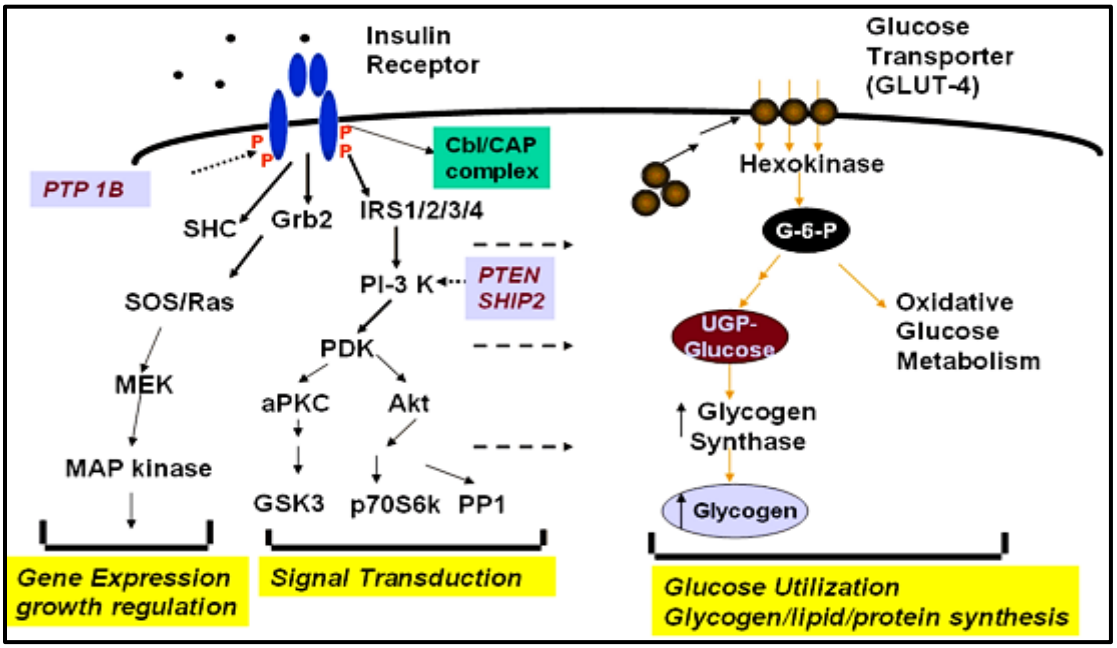
1.3.1.1 Insulin action

Insulin, a peptide hormone secreted by β-cells in the pancreas, is secreted in response to elevated blood glucose levels (Kahn, 1998). The effect of insulin on glucose metabolism varies in different tissue types. Insulin regulates the metabolism of carbohydrates and fats by promoting the absorption of glucose from the blood in skeletal muscle, and promoting fat storage in adipose tissue (Saltiel and Kahn, 2001). Insulin also inhibits hepatic glucose production by inhibiting gluconeogenesis (glucose production) (Claus and Pilkis, 1976) and glycogenolysis (glycogen breakdown) (Marks and Botelho, 1986).

1.3.1.2 Insulin signaling

Insulin is the primary mediator of glucose homeostasis (Leahy, 2005). During conditions of hyperglycaemia, β-cells in the pancreas increase their secretion of insulin to stimulate glucose uptake in insulin-responsive tissues such as the skeletal muscle so as to restore normo-glycaemia (Araujo *et al.*, 2013). Insulin initiates its biological action by binding to the tyrosine kinase insulin receptor located in the plasma membrane of insulin-responsive tissues. Phosphorylation of the insulin receptor results in the activation of a number of signaling cascades that regulates several biological processes including glucose uptake (Fig. 1.3). Activation of the insulin signaling cascade leads to the translocation of glucose transporter type 4 (GLUT4) to the cell membrane and the uptake of glucose from the circulation (Frosig *et al.* 2007). Skeletal muscle is considered the predominant site for insulin-mediated glucose disposal, and accounts for approximately 80% of peripheral glucose uptake in the postprandial state (Defronzo *et al.*, 1981; Defronzo and Tripathy,

2009). Dysregulation of any intracellular protein (protein kinase C (PKC), insulin receptor substrate 1/2 (IRS1/2) or phosphatidylinositol 3-kinase/serine-threonine protein kinase-1 (PI3K/AKT-1)) involved in the insulin signaling cascade leads to the development of insulin resistance (IR) (Kahn, 1998). The nature and extent of cellular IR depends on the tissue type, and vary according to the metabolic action of insulin within the tissue (Poornima *et al.*, 2006).



Key: protein tyrosine phosphatase-1B (PTP-1B), growth factor receptor bound protein 2 (Grb2), SHC-transforming protein (SHC), protein subfamily SOS/Ras, mitogen-activating protein kinase (MAPK/MEK), cbl-associated protein complex (cbl/CAP), insulin receptor substrate 1/2/3/4, phosphatidylinositol 3-kinase/serine-threonine protein kinase-1 (PI3K/Akt), PI3k- dependent serine/threonine kinase (PDK), atypical protein kinase C (aPKC), glycogen synthase kinase 3 (GSK3), (AKT), preproinsulin (PPI), p70 ribosomal subunit S6 kinase (p70S6k), glucose-6-phosphate (G-6-P), glucose transporter 4 (GLUT4).

Figure 1.3 Insulin signaling cascades (Adapted from: Li and Zhang, 2007).

1.3.1.3 Insulin resistance and fatty acids

Fundamental in the progression of T2D, are the deleterious effects of increased lipid accumulation and adipocyte hypertrophy and hyperplasia, which leads to the dysregulation of adipocyte control mechanisms and the recruitment of macrophages into adipose tissue, inflammation and the release of several factors that further exacerbate the IR state (Greenberg and Obin, 2006).

Several studies have suggested that increased levels of non-esterified fatty acids (NEFA) during disease progression alter insulin signaling pathways through different mechanisms (Poornima *et al.*, 2006). For example, Shulman *et al.* (2000) suggests that elevated levels of fatty acids inhibit the insulin signaling pathway, by activating PKC; an important intracellular insulin signaling protein (Shulman *et al.*, 2000). Activation of PKC leads to serine/threonine phosphorylation on the IRS-1/2, failure to activate PI3K, and decreased translocation of GLUT4 to the cell membrane, and subsequently decreased glucose uptake into the cell (Dresner *et al.* 1999; Shulman, 2000). Others have also reported that fatty acids can inhibit insulin signaling through PKC-independent pathways, where fatty acids induce IR in cells by attenuating insulin receptor gene expression (Bhattacharya *et al.*, 2007; Dey *et al.*, 2005). These examples confirm that NEFAs play a significant role in altering cellular insulin signaling pathways, thereby contributing to IR.

Furthermore, studies have suggested that elevated levels of glucose and fatty acids may impair β-cell function, and at a later stage, affect β-cell survival (Morgan, 2009; Purrello and Rabuazzo, 2000). It is generally agreed that both IR and β-cell dysfunction play important roles in the pathogenesis of T2D, although there is uncertainty about the relative contribution of these factors (Scheen, 2003; Kahn, 2003).

1.3.2 Beta-cell dysfunction

The primary function of β-cells is to synthesize and release insulin in response to increased blood glucose concentrations, thus restoring homeostasis. The process of insulin secretion is disrupted in dysfunctional β-cells as a result of irreversible damage to cellular components of insulin production over time (Stumvoll *et al.*, 2005). Several mechanisms including glucotoxicity, lipotoxicity, and amyloid formation have been proposed as a direct link to β-cell dysfunction (Biden *et al.*, 2014; Maedler, 2008; Stumvoll *et al.*, 2005). The glucotoxic condition (chronic hyperglycaemia exposure) is characterized by decreased insulin gene transcription, due to hyperglycaemia-induced loss of critical proteins that activate the insulin promoter (Kaiser *et al.*, 2003). The effect of hyperglycaemia on β-cells is often followed by a reduction in β-cell mass, as a result of β-cell apoptosis, without a compensatory increase in proliferation or neogenesis (cell renewal) (Bonner-Weir and O'Brien, 2008; Meier and Bonadonna, 2013).

The detrimental effects of excess glucose converge with the adverse consequences of lipotoxicity, both of which cause increased β-cell apoptosis. Lipotoxic conditions are induced in β-cells during chronic exposure to elevated levels of NEFAs, which is a characteristic of obesity and T2D (Kahn *et al.*, 2014). Chronically elevated NEFAs and the accumulation of long-chain acyl coenzyme A inhibit insulin secretion, as a result of pre-existing hyperglycaemia and glucose-induced fatty acid oxidation (Robertson *et al.*, 2003; Stumvoll *et al.*, 2005). In addition, elevated glucose concentrations increase the levels of reactive oxygen species, thereby inducing oxidative stress in β-cells over time. Pancreatic β-cells are particularly sensitive to oxidative stress due to their low intrinsic antioxidant capacity, thus oxidative stress may further exacerbate the impairment of β-cells during the development of T2D (Drews *et al.*, 2010).

1.3.3 Obesity

Obesity is currently a major health concern, affecting more than 475 million adults and 200 million school-aged children globally (World Obesity, 2012). The increasing prevalence of T2D is concurrent with the rising rates of obesity, and appears to reflect common environmental and genetic factors that underlie both conditions (Feero *et al.*, 2010; Hu, 2011). Indeed, overweight and obesity is widely considered to be the major driver of T2D, and studies have reported that 90% of adults with T2D are overweight or obese (Whitmore, 2010). These conditions are characterized by the excessive accumulation of body fat due to an imbalance between energy intake and expenditure, i.e. increased consumption of high fat, nutrient poor foods and decreased physical activity (Misra *et al.*, 2009). The body mass index (BMI), is a tool used to calculate overweight and obesity, and individuals with a BMI of ≥25 kg/m^2 or ≥30 kg/m^2 is defined as overweight or obese, respectively (Puoane *et al.*, 2002). Obesity is often associated with hypertension, low serum high-density-lipoprotein (HDL) cholesterol concentrations, and high serum low-density-lipoprotein (LDL) cholesterol, triglyceride and non-HDL cholesterol concentrations (Han *et al.*, 1998; Mooradian, 2009; Pradhan *et al.*, 2001). Together these factors are referred to as the metabolic syndrome, a risk factor for a number of chronic diseases, including T2D, and present major future challenges in reducing T2D mortality.

1.4 Progression of Type 2 diabetes

Type 2 diabetes is a chronic disease that progresses and worsens over time (Fig. 1.4). The disease is usually diagnosed during the later stages of disease progression, when insulin sensitivity and β-cell function is already significantly impaired. Together, the effects of increased IR and impaired β-cell function exacerbate hyperglycaemia, ultimately resulting in insulin deficiency and excess glucagon production (Fanelli *et al.*, 2006). Chronic hyperglycaemia leads to many long-term complications in the nerves, heart, kidney, eyes and blood vessels that cause irreversible tissue damage. Individuals with pre-diabetes, defined as having higher glucose levels than normal, but not high enough to be considered as T2D are also at risk of these micro- (nephropathy, neuropathy and retinopathy) and macro-vascular (coronary artery disease, stroke and peripheral arterial disease) complications (Fowler, 2008; Vinik and Flemmer, 2002). The risk of diabetic retinopathy and nephropathy, caused by progressive damage to the retina and kidney failure, respectively, is related to the severity of hyperglycaemia and the presence of hypertension in the pre-diabetic stage (Fowler, 2008). The risk for diabetic neuropathy is increased, depending on both the magnitude and duration of hyperglycaemia exposure, before the development of T2D. This is indicated by the presence of symptoms and/or signs of peripheral nerve dysfunction (Fowler, 2008). The central pathological mechanism in macrovascular disease is the process of atherosclerosis, which leads to the hardening and narrowing of arterial walls throughout the body. Atherosclerosis is thought to result from chronic inflammation and injury to the arterial wall, which leads to increased risk of developing cardiovascular disease (CVD) (Boyle, 2007). Conservative estimates indicate that more than 70% of patients with T2D die of cardiovascular causes (Laakso, 2010).

Taken together, these studies emphasize the need to detect T2D early, or to identify high risk individuals in the early stages, thereby, preventing or delaying T2D disease progression, and ultimately reducing mortality and morbidity worldwide. A number of studies have already reported that the benefits of the early detection and treatment of T2D can improve prognosis/management, and reduce T2D-related complication (Callejas *et al.*, 2013; Shamoon *et al.*, 1993; Tuomilehto *et al.*, 2001).

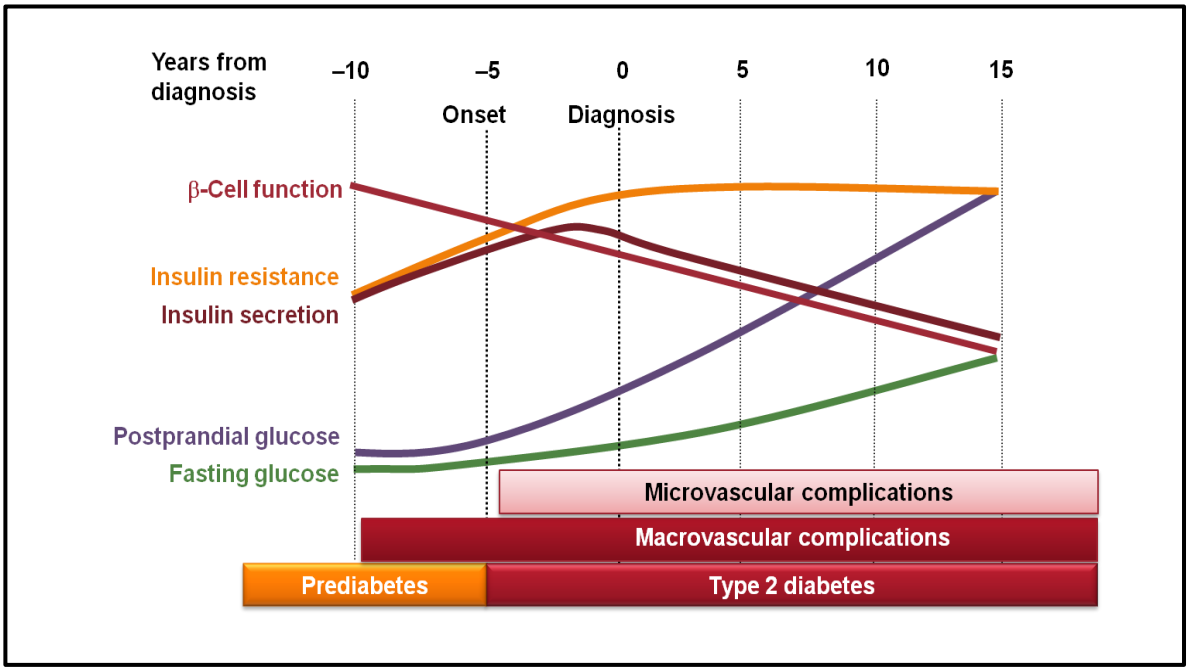


Figure 1.4 Type 2 diabetes disease progression (Adapted from: AACE diabetes resource centre, 2013).

1.5 Diagnosis of Type 2 diabetes

Diabetes is diagnosed according to the American Diabetes Association (ADA) or WHO criteria, by a fasting plasma glucose (FPG) concentration > 7.0 mmol/L, or a two-hour plasma glucose concentration during an oral glucose tolerance (OGTT) of > 11.0 mmol/L (Table 1.1). The FPG test measures glucose levels after fasting overnight for at least 8 hrs, while the OGTT test measures glucose tolerance after ingesting 75 g of glucose diluted in water. Glycated haemoglobin A1c (HbA1c) refers to the binding of glucose to haemoglobin, and due to the life-span of haemoglobin, reflects average glucose control over a three month period. The ADA has recently recommended that HbA1c > 6.5% can be used to diagnose diabetes (ADA, 2014).

Glucose measuring devices such as a glucometer (finger prick) may be used as a quick indicator of high blood glucose concentrations, but are not considered accurate enough for diagnosis. Additionally, a random plasma glucose concentration (RPG) > 11.1 mmol/L may be used to indicate possible T2D, although a confirmatory test is required (ADA, 2014).

Currently, the FPG and OGTT are the preferred tests for diagnosing diabetes. However, OGTT has been regarded as clinically impractical due to the 2 hour waiting period, and has led to the FPG being the most common test used to diagnose T2D globally. Literature suggests that more than one test should be used to accurately diagnose diabetes or hyperglycaemia (Barr *et al.*, 2002; Wang *et al.*, 2002), however, criteria for diagnosing T2D varies greatly throughout the world.

The various criteria and cut-off values used for the diagnosis of diabetes include FPG, HbA1c and OGTT, as presented in Table 1.1 (Malkani and DeSilva, 2012). The same tests are used for both screening and diagnosis, and are based on the 2015 ADA guidelines and the 2006 WHO addendum report (ADA, 2015; WHO, 2006). A number of European countries, as well as South Africa (Amod *et al.*, 2012) prefer to use the WHO diagnostic criteria, while others in America prefer to use the ADA diagnostic criteria (ADA, 2015; Deckers *et al.*, 2006).

Table 1.1 The current ADA and WHO diabetes diagnostic criteria (ADA, 2015; WHO, 2006).

		WHO criteria	**ADA criteria**
FPG	Normal:	< 6.1 mmol/L	< 5.6 mmol/L
	IFG:	6.1-6.9 mmol/L	5.6-6.9 mmol/L
	Diabetic:	≥ 7.0 mmol/L	≥ 7.0 mmol/L
OGTT (2hr plasma)	Normal:	< 7.8 mmol/L	< 7.8 mmol/L
	IGT:	7.8-11.0 mmol/L	7.8-11.0 mmol/L
	Diabetic:	≥ 11.1 mmol/L	≥ 11.1 mmol/L
HbA1C	Normal:	Not specified	< 5.7%
	Pre-diabetes:	Not specified	5.7-6.4%
	Diabetes:	≥ 6.5%	≥ 6.5%

1.5.1 Pre-diabetes

Pre-diabetes is widely considered to be an intermediate state of hyperglycaemia (Beagley *et al.*, 2014), and is often refered to as a state of impaired fasting glucose (IFG) or impaired glucose tolerance (IGT). Although controversy exists regarding the diagnostic criteria for pre-diabetes, many studies have reported that it is a very high predictor for the development of overt diabetes and related complications (WHO, 2006; Forouhi *et al.*, 2006). It is estimated that approximately 5%-10% of all pre-diabetes cases convert to T2D yearly. The WHO classifies IFG and IGT at 6.1-6.9 mmol/L (FPG) and 7.8-11.0 mmol/L (OGTT) cut-off values, respectively (WHO, 2006), while the ADA uses the same cut-off for IGT, but has a lower threshold for IFG (5.6-6.9 mmol/L) (ADA, 2015). Their rationale is based on data which showed that individuals with FPG concentrations between 5.6 mmol/L and 6.05 mmol/L were at increased risk of developing T2D and CVD, which would not be identified if a FPG cut-off threshold of 6.1-6.9 mmol/L was used (Gabir *et al.*, 2000; Shyong Tai *et al.*, 2004). However, Ferouhi *et al.* 2006 showed that the lowered threshold values (5.6-6.9 mmol/L) improved the sensitivity of IFG as a predictor of diabetes, but at the cost of specificity, thus misrepresenting the amount of individuals at risk of developing diabetes (Forouhi *et al.*, 2006; Sacks *et al.* 2011). Furthermore, the ADA has an additional HbA1c test at a cut-off value of 5.7-6.4% for detecting pre-diabetes, which is not defined in WHO (ADA, 2015).

1.6 Limitations and shortfalls of current diagnostic tests

In certain cases, the glycaemic status of patients may vary when different tests are used. Such variability may arise due to changes that occur over time, measurement variability, or because FPG, OGTT and HbA1c measure different physiological processes during the pathogenesis of T2D (Selvin *et al.*, 2007). In addition to these factors, evidence suggests that several technical challenges are associated with each individual diagnostic test, and could impede the diagnosis of T2D. Furthermore, current available diagnostic tests by FPG, HbA1c and OGTT are limited with regards to predicting diabetes; as it does not allow the identification of individuals who are susceptible to develop diabetes when glucose levels are still considered normal, thus increasing the risk of developing several health complications. However, in the absence of a more specific biological marker to define diabetes, plasma glucose estimation remains the basis of diagnostic criteria (Molleutze, 2006).

1.6.1 Oral glucose tolerance test and fasting plasma glucose test

The OGTT is currently the gold standard for the diagnosis of T2D due to its superiority in diagnosing diabetes in the clinical setting compared to the FPG test (Salmasi and Dancy, 2005). However, the OGTT requires very stringent conditions, such as overnight fasting, multiple blood sampling, consumption of 75 grams (g) of glucose diluted in water, and preventing the subject from movement for the duration of the test (Appajigol *et al.*, 2011). These conditions influence the test results, often making the test impractical. Moreover, the OGTT has greater inter-individual variability compared to the FPG test and HbA1c, and it is recommended that the test be confirmed by repeat testing on a different day (Selvin *et al.*, 2007; Waugh *et al.*, 2007). Therefore, due to these limitations, the ADA recommended FPG as the preferred glucose-based diagnostic test (International Expert Committee, 1997). However, FPG lacks sensitivity and results in lower disease prevalence compared to OGTT, and moreover, cannot identify subjects with IGT (Waugh *et al.*, 2007). A study investigating glycaemia in a black South African population illustrated that the prevalence of T2D would be lower if FPG was solely used for diagnosis compared to using both FPG and OGTT (Motala *et al.*, 2008). This illustrates the different underlying pathophysiologies of T2D, and suggests that more specific tests are required to accurately diagnose diabetes.

1.6.2 Glycated haemoglobin A1c

The use of HbA1c has a number of advantages compared to glucose measurements by FPG and OGTT. These advantages include no fasting, less variability, greater pre-analytical stability, and a greater index for overall glycaemic control. However, there are a number of factors that could lead to the misinterpretation of HbA1c diagnostic measurements. These factors include altered red blood cells in patients with haemoglobinopathies, and variations due to iron deficiency, aging, ethnicity and antiretroviral drugs (Church and Simmons, 2014; International Expert Committee, 2009; Kilpatrick and Winocour, 2010; Kirkman and Kendall, 2011; Saudek *et al.*, 2008; Topic, 2014). These factors hamper the use of HbA1c, especially in countries where the prevalence of these disorders are high. For example, South Africa is a multi-ethnic country that currently has the highest prevalence of Human Immunodeficiency Virus Infection/Acquired Immune Deficiency Syndrome (HIV/AIDS) globally, therefore decreasing the predictive value of HbA1c (HIV and AIDS in South Africa, 2014). Indeed, a

recent analysis has precluded its use for diagnosing diabetes in a South African setting (George, 2011). The limitations of the current diagnostic tests underscore the need to identify new methods to diagnose T2D more accurately.

1.7 Biomarkers

Biomarkers are indicators of normal biological processes that can be used as indicators of a particular disease state or other biological states of organisms. They are clinically useful because they can potentially predict or diagnose disease, give insight into the pathophysiology of disease, and can be used to monitor or predict clinical outcome (Sahu *et al.*, 2011). Although the term 'biomarker' is fairly new, it has been used for many years in clinical diagnosis and research. Examples of a few well-known biomarkers include pulse and blood pressure (hypertension), cholesterol (coronary and vascular disease), C-reactive proteins (inflammation), and HbA1c, FPG, RPG and OGTT (diabetes) (Kumar and Khanna, 2011; Sahu *et al.*, 2011). However, biomarkers need to fulfil several criteria to be clinically useful. Among others, these include:

- Tissue or pathology specificity,
- Easily accessible through minimally invasive methods to collect biofluids,
- Sensitivity to relevant changes in the disease,
- Early detection of disease before clinical symptoms appear,
- A long half-life within the sample,
- Optimal speed, accuracy and ease of analysis,
- Cost effective and reproducible, and
- The ability to differentiate between pathologies (Etheridge *et al.*, 2011; Sahu *et al.*, 2011)

Effective biomarkers are ones that are able to monitor and accurately identify individuals at the subclinical stage and enable preventative measures before the disease develops (Lyons and Basu, 2012).

1.7.1 Recent advances in biomarker discovery

A number of genomic, transcriptomic, proteomic, and metabolic markers currently exist, and have been correlated with T2D disease progression (Bain *et al.*, 2009; Galazis *et al.*, 2013; McKillop and Flatt, 2011). However, these biomarkers often lack sensitivity and/or specificity

and are associated with the irreversible stages of T2D (Galazis *et al.*, 2013). Identification of novel biomarkers that fulfill some of the criteria listed in section 1.7 would be of great clinical value and would have the potential to facilitate intervention strategies that can be tailored to the characteristics of an individual to prevent or modify disease progression (Lyons and Basu, 2012). Recent advancements in the field of molecular biology have led to the development of molecular biomarkers that are easily measured in biological samples such as plasma, serum, and whole blood (Chen *et al.*, 2008; Holland *et al.*, 2003).

1.8 Epigenetics

Epigenetics is defined as the heritable changes in gene expression or phenotype that occurs without changes in the underlying DNA sequence (Christensen and Marsit, 2011). Recent findings suggest that epigenetics underpins the crucial link between environmental factors and genetic predisposition in the pathogenesis of T2D. Environmental exposures such as nutrition, toxins, age, physical inactivity, etc. modify the epigenome, causing epigenetic dysregulation; a key mechanism underlying the development of metabolic diseases (Hamilton, 2011). Recently, epigenetic mechanisms have attracted considerable interest as potential biomarkers which would identify T2D. Moreover, due to their reversible nature, epigenetic changes may provide a window of opportunity for intervention strategies to prevent or delay the progression to T2D (Reddy *et al.*, 2013).

These epigenetic mechanisms include DNA methylation, loss of genomic imprinting, chromatin remodeling and non-coding RNA (Gibney and Nolan, 2010; Hirst and Marra, 2009). Non-coding RNAs include long non-coding and short-non-coding RNAs such as microRNAs that are able to positively and negatively regulate gene expression in a signaling cascade (Stefani and Slack, 2008; Wahlestedt, 2013).

1.9 MicroRNAs

MicroRNAs (miRNAs) are a class of small, highly conserved non-coding RNA molecules that have recently attracted considerable interest as epigenetic modulators of gene expression in a wide range of diseases, including T2D (Kong *et al.*, 2011; Olson, 2014; Schwarzenbach *et al.*, 2014). MiRNAs are single-stranded RNA species approximately 22 nucleotides (nt) in length, that are able to regulate gene expression by inducing repression of target messenger RNA (mRNA) through translational inhibition or initiating mRNA degradation (Brennecke *et*

al., 2005). Since their discovery in *C. elegans* in 1993 (Lee *et al.*, 1993), over 1000 miRNAs have been identified in humans, where they regulate a large proportion of genes in the human genome (Ardekani and Naeini, 2010). MiRNAs play a critical role in gene regulation and have been shown to be involved in highly regulated processes including differentiation, apoptosis, proliferation and metabolic processes (Du and Zamore, 2005). Furthermore, miRNAs have been extensively studied to understand the regulatory mechanisms involved in the pathogenesis of diseases, such as neurological disorders, CVD, obesity and T2D (Creemers *et al.*, 2012; Guay *et al.*, 2011; Kong *et al.*, 2011; Zampetaki *et al.*, 2010; Zampetaki *et al.*, 2012; Hamilton, 2011; Pinney and Simmons, 2010). These miRNA regulatory mechanisms may offer new opportunities for the early detection of T2D and associated disorders, which can be used in clinical diagnostics to identify points between exposure and disease.

1.9.1 MicroRNA nomenclature

For ease of understanding and identification of experimentally confirmed miRNAs, a nomenclature system has been adopted. Briefly, the numbering of newly identified miRNAs is sequential, and is attached to the prefix ‘mir’ followed by a dash (-). The uncapitalized ‘mir’ refers to the pre-miRNA and the capitalized ‘miR’ refers to the mature form. MiRNAs are named using the prefix ‘miR’, followed by the unique identifying number prescribed to each miRNA in the numerical order of discovery (for eg. miR-1, miR-2, miR-3 etc.) (Wright and Bruford, 2011). When two mature miRNAs originate from either the 3’ or 5’ arm of the same pri-miRNA, they are denoted with either a ‘-3p’ or ‘-5p’ suffix, respectively, at the end of each miRNA (for eg. miR-1-3p or miR-1-5p) (Issler and Chen, 2015). However, when the mature miRNA found from one arm is more abundant than that from the other, an asterisk (*) following the miRNA name is denoted for the miRNA with the lowest concentration (Bartel, 2004). For example, miR-1 and miR-1* share a pri-miRNA hairpin, but higher concentrations of miR-1 is found in the cell.

MiRNAs with nearly identical mature sequences are annotated with lower case letter to show their similar structure (for example, miR-1a and miR-1b). Distinct precursor sequences and genomic loci from different regions of the genome that express identical mature sequences, are distinguished with an additional number (for example, miR1a-1 and miR-1a-2) (Issler and Chen, 2015). MiRNAs are also annotated according to the species they are observed in, and

are designated with a three-letter prefix. For example, has-miR-1-3p is observed in *Homo sapiens* (human), while mmu-miR-1-3p is observed in *Mus musculus* (mouse) (Ambros *et al.*, 2003).

1.9.2 MicroRNA biogenesis and mechanism of action

The biogenesis of miRNAs begins in the nucleus of the cell, where several primary miRNA sequences (pri-miRNAs) are transcribed by either RNA polymerase II or RNA polymerase III depending on promoter and terminator sequences (Cai *et al.*, 2004), and then capped, spliced and polyadenelated. Thereafter, pri-miRNAs are processed by microprocessor complex Drosha (a nuclear RNase III enzyme) and DGCR8 (encoded in humans by DiGeorge critical region 8*)* into precursor miRNAs (pre-miRNAs) of ~60-70 nt long stem-loop structures (Lee *et al.*, 1993; Muhonen and Holthofer, 2009). These pre-miRNAs are then exported from the nucleus to the cytoplasm by Exportin-5 (xpo5) and RanGTP complex. In the cytoplasm, pre-miRNAs are further processed and cleaved into ~22 nt long mature miRNA duplexes by Dicer, a cytoplasmic RNase III enzyme and its interacting partner, a transactivation-responsive RNA binding protein (TRBP or TARBP2) (Liu *et al.*, 2008; Pandey *et al.*, 2009). Having lower thermodynamic stability, the 5' end of the miRNA duplex is selected by the miRNA-induced silencing complex (miRISC), and is subsequently bound to the Argonaute protein, which forms part of the risk effector complex. The bound mature miRNA guides the miRISC to the 3' untranslated region (UTR) binding site of the target mRNA, where they are able to downregulate gene expression (Richard *et al.*, 2005). This may be achieved by two posttranscriptional mechanisms, namely, mRNA cleavage/degradation or translational repression (Kumar and Khanna, 2011) (Fig.1.5).

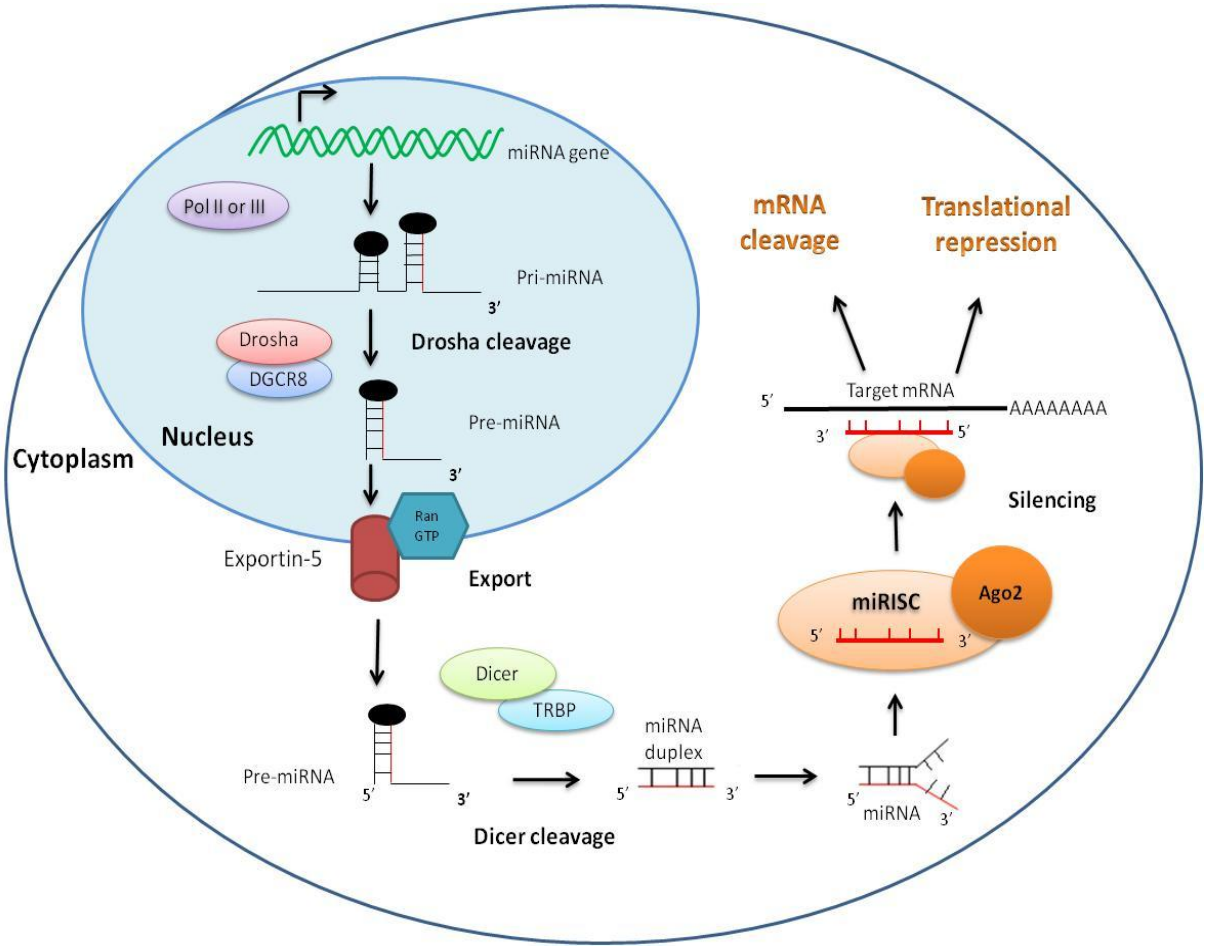


Figure 1.5 Schematic diagram representing the biogenesis of miRNA and its mechanism of action. Primary miRNA transcripts (pri-miRNA) transcribed by RNA Polymerase (pol II or III) and processed by microprocessor (Drosha and DGCR8), are exported as pre-miRNAs into the cytoplasm by Exportin-5 and RanGTP. Dicer and its interacting binding protein TRBP process pre-miRNAs into short miRNA duplexes. The 5' mature miRNA strand binds to the RNA-induced silencing complex (miRISC)-associated argonaut protein (Ago2), and induces silencing of their mRNA target sequences.

The basic difference between mRNA cleavage and translational repression is governed by the levels of complementarity between miRNAs and their target mRNA transcripts (Bartel, 2004). In plants and in a small class of eukaryotes, miRNAs bind to a single, generally perfect or near perfect complimentary site in either the coding or 3' UTR of the target mRNA, which results in target cleavage and degradation (Ghosh, 2011). In contrast, in most investigated animals (mammals), miRNAs bind to multiple, imperfect complementary sites in the 3' UTR target region, and directs the inhibition of protein accumulation through translational repression (Pillai *et al.*, 2007). The complementarity is usually restricted to the 5' end of the mature miRNA, at position 2 to 8, known as the miRNA 'seed region', a sequence that occurs when the nucleotide adenine (A) pairs with uracil (U) and guanidine

(G) pairs with cytosine (C) (Xu *et al.*, 2014) (Fig.1.6). However, when the nucleotide represents a G or a U, the pairing may be less specific, where two bases can be interchangeably recognised by the mRNA, more commonly known as wobble paring (Peterson *et al.*, 2014). A perfect seed match between the miRNA and mRNA target has no gaps in the alignment, and is regarded to be the most important feature for miRNA target recognition in mammals (Lewis *et al.*, 2003). Moreover, a number of miRNA target predicting analyses showed that the miRNA target sequence is highly conserved across species, and requires a number of base matches flanking the seed sequence to direct the specificity of miRNA:mRNA interactions (Lewis *et al.*, 2005; Peterson *et al.*, 2014). Although understanding of miRNA function is limited, enough evidence exists to illustrate that a given miRNA is able to regulate multiple mRNA targets in a signaling cascade (Friedman *et al.*, 2009; Lewis *et al.*, 2005), and is therefore involved in controlling many biological processes to maintain metabolic homeostasis (Felekkis *et al.*, 2010).

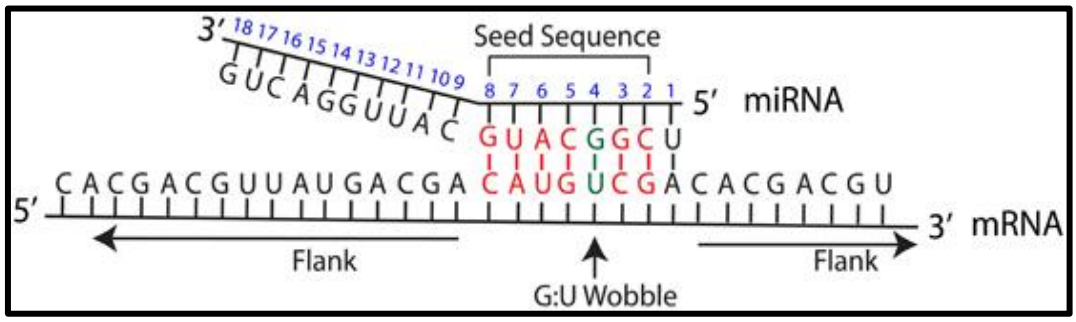


Figure 1.6 Schematic overview of the miRNA:mRNA target interaction. Watson-crick base pairing of the miRNA seed sequence and the mRNA target sequence is shown in red, and an example of a G-U wobble in the seed sequence is shown in green. Flank refers to the 5' or 3' mRNA sequence corresponding to the region on either side of the seed sequence (Adapted from: Peterson *et al.*, 2014).

1.9.3 The role of microRNAs in Type 2 diabetes

The majority of miRNAs is tissue and cell type specific, and plays a critical role in gene regulation, while others may be expressed ubiquitously depending on their function (Lagos-Quintana *et al.*, 2002; Lim *et al.*, 1999; Mao *et al.*, 2013). Recently, scientists have shown that miRNA expression is regulated by environmental factors, which contribute to the aberrant gene expression patterns seen in metabolic disorders (Rottiers and Näär, 2012). The dysfunction of miRNA regulation disrupts normal cellular activity which may lead to the development of various diseases, such as cancers, lymphomas, CVD complications, T2D

and neurological disorders such as Parkinson's and Alzheimer's disease (Cheng and Zhang, 2010; Cuk *et al.*, 2013; Maes *et al.*, 2009). Recently, there has been considerable interest in understanding the RNA regulatory phenomena and how miRNAs function in the development of T2D (Guay *et al.*, 2011).

In the context of T2D, miRNAs are widely expressed in blood, pancreas, liver, skeletal muscle and adipose tissue, and are able to regulate the expression of multiple genes (Rottiers and Näär, 2012). Together, these miRNAs regulate insulin sensitivity mainly by targeting insulin receptors and the components of insulin/protein kinase B (PKB) signaling pathways or GLUT4- mediated glucose uptake and metabolism (Tang *et al.*, 2008). Global miRNA profiling has provided valuable information that couples miRNA expression changes occurring in pancreatic β-cells and insulin target tissues, with the changes that occur in glucose levels during the pathogenesis of T2D (Guay *et al.*, 2011). For example, Herrara *et al.*, (2010) measured the expression of several miRNAs in the liver, skeletal muscle and adipose tissue in spontaneously diabetic (Goto-Kakizaki) and normo-glycaemic (Brown-Norway) inbred rats, and found that expression of several miRNAs correlated with the glycaemic status of the rats. Each miRNA showed a significant tissue-specific expression pattern that varied between the different strains of rats, and between normal and diabetic within each strain (Herrera *et al.*, 2010; Herrera *et al.*, 2009). Moreover, prolonged *in vitro* exposure of mouse β-cells (MIN6) to high glucose levels resulted in differential expression of a large number of miRNAs. Among these, miR-124a, miR-30d and miR-107 were upregulated, while miR-296, miR-484 and miR-690 were downregulated by high glucose treatment of MIN6 cells (Tang *et al.*, 2009).

Furthermore, several human studies have demonstrated that miRNAs are differentially expressed in multiple tissue types, and are able to regulate the expression of multiple genes in a signaling cascade. MiRNAs are expressed in several tissues that play a crucial role in insulin signaling, glucose metabolism and β-cell development. Among others, miRNAs expressed in pancreatic tissue include miR-375, miR-29a, miR-96, miR-124a, miR-376 and let-7, while three of these miRNAs (miR-29a, miR-375, and miR-96) are also expressed in other tissue types, such as muscle, liver and adipose, and regulate the expression of multiple genes involved in maintaining glucose homeostasis. Studies have reported that these miRNAs exert their action on different tissue types, and present a consistent regulatory role during the pathogenesis of T2D (Zhu and Leung, 2015). Although miRNA

research has skyrocketed over the last few years (Friedlander *et al.*, 2014), more studies are needed to improve the understanding of these miRNAs and their regulation during disease.

1.9.4 Circulating microRNAs

Majority of miRNAs are found intracellularly, while a significant number of miRNAs are found in the extracellular fluid outside the cell, and are particularly useful in biomarker discovery (Creemers *et al.*, 2012; Hanson *et al.*, 2009; Weber *et al.*, 2010). These extracellular miRNAs circulate in the bloodstream and compared to mRNA levels, have been found to be remarkably stable among individuals of the same species (Chen *et al.*, 2008; Hunter *et al.*, 2008). Circulating miRNAs are able to withstand unfavorable physiological conditions such as extreme pH variability, boiling, multiple freeze thaw cycles and extended storage, making them attractive, potential biomarkers for clinical research (Chen *et al.*, 2008).

It has been postulated, that the stability of these miRNAs may result from the formation of complexes between circulating miRNAs and specific proteins (Turchinovich *et al.*, 2012). These complexes include ribonucleoproteins (RBP) such as argonaute (Ago1-4) or nucleophosphin 1 (NPM1), exosomes, apoptotic bodies, microvesicles and lipoprotein complexes which serve to protect these miRNAs from nuclease degradation, and act as carriers to transport them to their target mRNAs (Creemers *et al.*, 2012) (Fig. 1.7). This suggests cell-to-cell communication, allowing miRNAs to influence gene expression in neighbouring cells when conditions become stressful (Creemers *et al.*, 2012). MiRNAs encapsulated in microvesicles (made up of exosomes and shedding vesicles) are released from the cell through blebbing (protrusion or bulge) of the plasma membrane, and can then be transferred to recipient cells where they trigger functional effects and modulate gene expression (Hunter *et al.*, 2008; Thery *et al.*, 2002; Valadi *et al.*, 2007).

The release of exosomes and microvesicles from cells are regulated by sphingomyelin, a major lipid in the lipid bilayer of the cell membrane. The biosynthesis of sphigomyelin is tightly controlled by neutral sphingolimyelinase 2, a hydrolase enzyme involved in lipid metabolism, and ceramide, a component of sphingomyelin, dependent secretary-machinery in response to cellular stresses (Cortez *et al.*, 2011; Kosaka *et al.*, 2010; Trajkovic *et al.*, 2008). Some miRNAs are found in apoptotic bodies and are released as byproducts from cells into the extracellular compartment during apoptosis. This could partly explain the

presence of tissue specific miRNAs in blood after toxicity in certain tissues (Laterza *et al.*, 2009; Zhang *et al.*, 2010). The majority (approximately 90-99%) of all circulating miRNAs appears to be vesicle-free, and is associated with RNA-binding proteins (eg. Ago, NPM1), which allow miRNAs to remain highly stable during extracellular circulation (Arroyo *et al.*, 2011; Turchinovich *et al.*, 2011). Additionally, a small portion of circulating miRNAs has also been found to be associated with high density lipoproteins (HDL) (Wagner *et al.*, 2013). High density lipoproteins are capable of delivering both exogenous and endogenous miRNAs to recipient cells, and mediating direct targeting of mRNA, which leads to gene expression alterations (Vickers *et al.*, 2011). However, in contrast to exosomes, miRNA associated lipoproteins are negatively regulated by neutral sphingolimyelinase 2, and delivery is dependent on the cell surface HDL receptors (SR-B1) (Chen *et al.*, 2012; Vickers *et al.*, 2011).

Although the mechanism of action for cell-to-cell communication is not completely understood, two theories currently exist regarding the export and biological function of extracellular miRNAs. One theory suggests that miRNAs are merely byproducts of microvesicle secretion and cell death, while the other theory suggests that miRNAs are specifically secreted and function in intercellular communication via paracrine and even endocrine signaling routes (Turchinovich *et al.*, 2012; Valadi *et al.*, 2007; Cortez *et al.*, 2011).

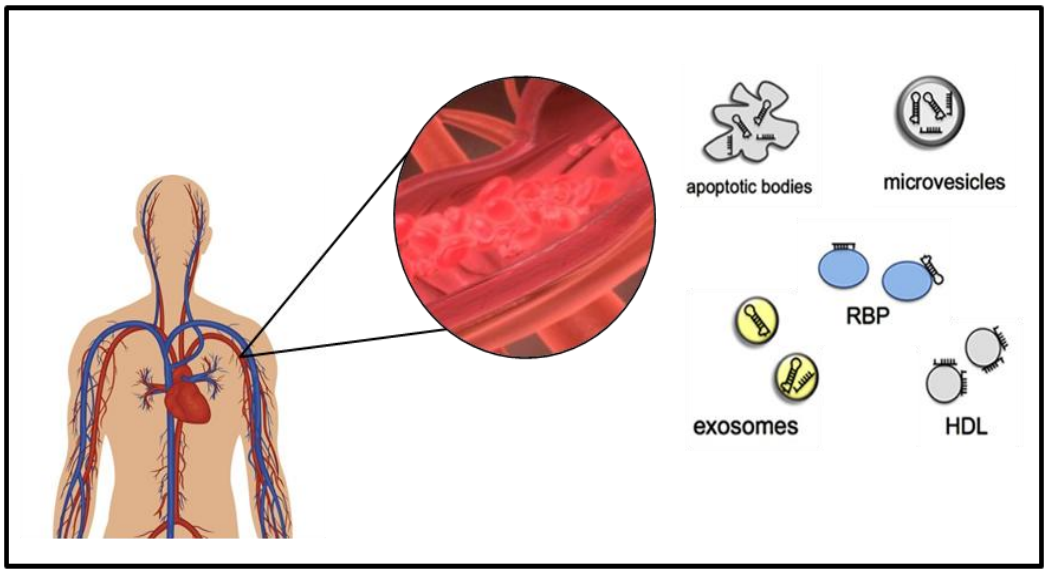


Figure 1.7 Circulating miRNAs and associated complexes, such as ribonuclear proteins (RBP), apoptotic bodies, microvesicles, exosomes and high density lipoproteins (HDL), found in the bloodstream (Adapted from Kinet *et al.*, 2013).

1.9.5 Circulating microRNAs as clinical biomarkers for Type 2 diabetes

Regardless of the mechanism regulating circulating miRNAs, recent studies suggest that different pathophysiological conditions cause inappropriate release of miRNAs into the blood, altering miRNA expression profiles, and mediating the repression of critical mRNA targets in recipient cells. This leads to altered gene expression levels that consequently cause metabolic dysregulation (Chen *et al.*, 2012; Santovito *et al.*, 2014). Thus, distinct modification in the miRNA profile in blood may reflect the development of various chronic diseases that can be detected several years before the disease and its complications manifests (Hydbring and Badalian-Very, 2013). Extracellular circulating miRNAs have signatures for various diseases, and could play a major role as potential biomarkers in health and disease (Chen *et al.*, 2008). The first link between circulating miRNAs and disease came from cancer studies in 2008 (Gilad *et al.*, 2008; Lawrie *et al.*, 2008; Mitchell *et al.*, 2008). These studies showed that tumour-derived miRNAs, were detected in serum or plasma, and could distinguish between cancer patients and healthy individuals (Lawrie *et al.*, 2008; Mitchell *et al.*, 2008). Subsequent studies have also showed that circulating miRNAs can identify and monitor the progression of CVD, liver injury and kidney disease (Ding *et al.*, 2012; Saal and Harvey, 2009; Van Rooij and Olson, 2012). These findings support the hypothesis that circulating miRNAs may be used as informative biomarkers to assess an individual's pathophysiological status, sparking interest in their use as high risk, diagnostic or prognostic biomarkers for disease.

In light of the positive associations between miRNAs and disease, a number of studies are exploring circulating miRNAs for their diagnostic and prognostic ability in T2D (Pescador *et al.*, 2013; Yang *et al.*, 2014; Zhang *et al.*, 2013). Indeed, a number of studies have shown that miRNAs that are dysregulated during glycaemia, can be present in whole blood, PBMCs, serum and plasma of diabetic, pre-diabetic and normo-glycaemic individuals (Table 1.2).

Studies have also investigated their mechanism of action, confirming their role in T2D pathogenesis. For example, Yan *et al.* (2014) showed that miR-199, a miRNA significantly upregulated in the plasma of T2D patients compared to normo-glycaemic individuals, binds to GLUT4 mRNA and represses its expression, and subsequently glucose uptake in L6 rat

myoblast cells (Yan *et al.*, 2014). The expression of MiR-144, another miRNA upregulated in T2D patients compared to normo-glycaemic individuals was upregulated in rat islets cultured under hyperglycaemic conditions (Karolina *et al.*, 2011). Furthermore, using the Dual Luciferase reporter assay and Western blot analysis, these authors showed that miR-144 induces repression of insulin receptor substrate 1 (IRS1) mRNA and protein, demonstrating its potential role in the insulin signaling pathway (Karolina *et al.*, 2011). Studies have shown that two T2D-associated miRNAs, miR-146a and miR-155, may mediate their pathogenic effects by stimulating inflammation (Balasubramanyam *et al.*, 2011; El-Ekiaby *et al.*, 2012; O'Connell *et al.*, 2010; Taganov *et al.*, 2006; Zhang *et al.*, 2010). The expression of these miRNAs was negatively associated with inflammation (nuclear factor kB (NFkB) mRNA levels and circulatory levels of tumour necrosis factor alpha (TNFα), and interleukin (IL-1β and IL-1)).

Interestingly, a number of studies have reported that circulating miRNAs can distinguish between T2D, pre-diabetes and normo-glycaemia. For example, according to Karolina *et al.* (2011), eight circulating miRNAs (miR-144, miR-146a, miR- 150, miR-182, miR-192, miR-29a, miR-30d and miR-320) were shown to be differentially expressed in PBMCs, and could serve as potential signatures to distinguish between IFG and T2D. Furthermore, Parrizas *et al.* (2014) showed that circulating levels of miR-192 and miR-193b were significantly upregulated in pre-diabetic compared to normo-glycaemic individuals, and interestingly, the increased levels of these miRNAs in pre-diabetics reverted back to normal following a therapeutic exercise intervention. Moreover, the authors showed that the abundance of these miRNAs was significantly correlated with metabolic parameters including OGTT, triglycerides and fatty liver index (Parrizas *et al.*, 2014). Similarly, Liu *et al.* (2014) showed that the decreased expression of miR-126 in pre-diabetic and diabetic individuals compared to normo-glycaemic individuals was returned to normal after glucose lowering treatment (Lui *et al.*, 2014). Moreover, Yang *et al.* (2014) explored the clinical significance of serum miRNAs (miR-23a, let-7i, miR-486, miR-96, miR-186, miR-191, miR-192, and miR-146a) in T2D patients. Of these miRNAs, miR-23a had significantly lower expression according to disease progression, i.e. diabetic < pre-diabetic < normo-glycaemic individuals (Yang *et al.*, 2014). Finally, a prospective, longitudinal study showed that the expression of five miRNAs, miR-15a, miR-29b, miR-126, miR-223 and miR-28-3p, changed in normo-glyceamic individuals during disease progression to T2D (Zampetaki *et al.*, 2010). Altered miRNA expression profiles were observed upon the manifestation of T2D, over a period of ten years.

Together, these studies support the hypothesis that circulating miRNAs could help to distinguish normo-glycaemic from diabetic and pre-diabetic individuals, and may have the potential as early predictive biomarkers of T2D.

Table 1.2 A summary of studies on circulating miRNAs in T2D (Adapted from Raffort *et al.*, 2015)

MicroRNA	Sample	Disease	Summary	Reference
miR-20b, -21,-24, -15a, -126, -191, -97, -223, -320, -486, -28-3p	Plasma	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Zampetaki *et al.*, 2010)
miR-15a, -29b, -126, -223, -28-3p	Plasma	T2D	Differential miRNA expression during disease progression	
miR-126	Plasma	T2D	Differential miRNA expression between T2D normo-glycaemic individuals, and between T2D susceptible and normo-glycaemic individuals	(Zhang *et al.*, 2015)
miR-146a	Plasma	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Rong *et al.*, 2013)
miR-24, -29b	Plasma	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Wang *et al.*, 2014)
miR-144	Plasma	T2D	Dysregulation associated with T2D in Swedish people, not in Iraqi people	

miR-375	Plasma	T2D	Differential miRNA expression between Kazak T2D and Han T2D, and differences in the methylation of promoter regions	(Chang *et al.*, 2014)
miR-140-5p, -142-3p, -222,-423-5p, -125b, -192, -195, -130b, -532-5p, -126	Plasma	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Ortega *et al.*, 2014)
miR-192, 140-5p, 222	Plasma	T2D	Modification of miRNA expression by metformin	
miR-222, 140-5p	Plasma	T2D	Modification of miRNA expression by molecules inducing IR	
miR-375	Plasma	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Sun *et al.*, 2014a)
miR-191, -200b	Plasma	T2D	Differential miRNA expression between T2D and normo-glycaemic individuals; and a correlation between miRNAs and C-reactive protein and cytokine levels in T2D patients	(Dangwal *et al.*, 2015)
miR-199	Plasma	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Yan *et al.*, 2014)

miR-326, -let-7a, -let-7f	Plasma	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Santovito *et al.*, 2014)
miR-9, -29a, -30d, -34a, -124a, -146a, -375	Serum	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals; No significant difference in miRNA expression between pre-diabetic and normal glucose tolerance group	(Kong *et al.*, 2011)
miR-126	Serum	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals; Differential miRNA expression between pre-diabetic patients and normo-glycaemic individuals; Modification of miRNA expression by glucose lowering treatment	(Liu *et al.*, 2014)
miR-23a, -let-7i, -486, -96, -186, -191, -192, -146a	Serum	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Yang *et al.*, 2014)
miR-23a			Differential miRNA expression between pre-diabetic and normo-glycaemic individuals	
miR-191, -139-5p, -21	Serum	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals	(Parrizas *et al.*, 2014)

miR-191, -15b, -128, -125a-5p, -50, -192, -193b	Serum	T2D	Differential miRNA expression between pre-diabetic and normo-glycaemic individuals	(Parrizas *et al.*, 2014)
miR-192, -193b	Serum	T2D	Differential miRNA expression in pre-diabetic but not in diabetic patients; Modification of miRNA expression by lifestyle intervention in pre-diabetics	
miR-138, -376a	Serum	T2D	Differential miRNA expression between normo-glycaemic individuals, diabetic and obese diabetic patients	(Pescador *et al.*, 2013)
miR-let-7a	Whole blood	T2D	Differential miRNA expression between T2D patients and normo-glycaemic individuals; Differential miRNA expression between T2D with diabetic nephropathy and T2D	(Zhou *et al.*, 2013)
miR-144, -146a, -150, -182, -192, -29a, -30d, -320	PBMC	T2D	Differential expression between patients with IFG and T2D	Karolina *et al.*, 2011
miR-146a, -155	PBMC	T2D	Differential miRNA expression between T2D patients and non-diabetic individuals; Correlation between miR-146a levels and fasting blood glucose, HbA1c and inflammatory signals	(Balasubramanyam *et al.*, 2011)

miR-103b	Platelets	T2D	Differential miRNA and gene target SFRP4 expression between pre-diabetic, non-complicated T2D, coronary heart disease T2D vs. healthy controls	(Luo *et al.*, 2015)

1.9.6 Techniques to study microRNA expression

MiRNA research has been facilitated by technological advancement that has spawned a multitude of platforms which enabled large scale epigenomic studies. At present, techniques including quantitative real-time polymerase chain reaction (qRT-PCR), microarray hybridization techniques such as oligonucleotide arrays and *in situ* hybridisation, and high throughput sequencing (HTS) of miRNAs are popular and widely used profiling methods (Thomas and Ansel, 2010; Git *et al.*, 2010; Pritchard *et al.*, 2012; Meyer *et al.*, 2010). Among these techniques, the most cost effective and frequently used method for miRNA profiling is qRT-PCR (Fu *et al.*, 2006; Kroh *et al.*, 2010). Quantitative RT-PCR is a highly reproducible, sensitive and specific method, which is used for quantifying miRNA expression levels. The target specificity is based on the design of primers and probes specific to the miRNA of interest (Benes and Castoldi, 2010). Furthermore, although more expensive, HTS allows global miRNA expression profiling for the detection of less abundant and novel miRNAs, which cannot be detected using qRT-PCR (Pritchard *et al.*, 2012).

1.10 Bioinformatics: messenger RNA target prediction analysis

One of the most interesting aspects of miRNA biology is the ability of a single miRNA to regulate multiple genes involved in specific cellular mechanisms and signaling cascades (He and Hannon, 2004). Experimental techniques to identify miRNA targets include Luciferase reporter assays, Western blot analysis and Enzyme-linked immunosorbent assay (ELISA) or Immuno-cytochemistry experiments (Kuhn *et al.*, 2008). The Luciferase reporter assay is a tool used to study gene expression at the transcriptional level, and can determine the direct binding ability of miRNAs on messenger RNA (mRNA) target sequence (Simon and Evin, 2008). Studies have shown that miRNAs are able to regulate the expression of targets at the protein level (Selbach *et al.*, 2008), thus, Western blot analysis, ELISA or Immuno-cytochemistry present techniques used to study protein expression of miRNA targets at the translational level (Kuhn *et al.*, 2008). However, these techniques are costly and time-consuming, focusing attention towards computational miRNA:mRNA target prediction (Min and Yoon, 2010; Witkos *et al.*, 2011). Bioinformatic analysis allows high-throughput target prediction, and is relatively easy to conduct. However, targets that were computationally

predicted, must be experimentally validated before they can be acknowledged as a miRNA target.

The four most common miRNA:mRNA bioinformatics target prediction algorithms are based on seed match complementarity (Fig.1.6), sequence conservation, free energy and binding site structural accessibility (Kertesz *et al.*, 2007; Lewis *et al.*, 2003; Lewis *et al.*, 2005; Oulas *et al.*, 2015; Witkos *et al.*, 2011). Although all these algorithms are critical for miRNA target prediction, their selection depends on the research question. Potential interactions are best identified by firstly using a program that considers seed match complementarity, since these are characterized by high sensitivity and precision. Thereafter, to increase specificity, additional features such as binding site structural accessibility and sequence conservation should be used (Witkos *et al.*, 2011).

1.11 Study motivation

The increasing prevalence of T2D globally places a huge financial and health burden on governments, particularly those in middle and low income countries (Whiting *et al.*, 2011). Individuals with diabetes are highly vulnerable to a multitude of complications, including CVD, retinopathy, peripheral arterial disease and infectious diseases, which increase the cost to treat the disease (Zhuo *et al.*, 2013). It is estimated that more than half of all people with diabetes, are unaware that they are living with the disease (IDF, 2013). Furthermore, the alarming increase of T2D in children and the high number of undiagnosed cases globally, suggests that T2D may become the world's greatest health challenge within the next 25 years.

Although a multitude of tests are currently available to diagnose T2D, they are all based on blood glucose measurements, and are unable to identify individuals in the asymptomatic stage (pre-diabetic), when glucose levels are still normal. Thus, these individuals remain undetected for many years, and have an increased risk of developing long term metabolic complications. Furthermore, despite the array of tools available for the management of T2D (drug therapies, advanced technology and improved education), the potential benefit of these tools are lost in people with undiagnosed diabetes (Waugh *et al.*, 2007).

Several lines of evidence suggest that early intervention is the most effective strategy to reduce the prevalence of T2D, manage the disease, and to prevent complications. For example, diagnosis of T2D is associated with micro- and macro-vascular complications (Fowler, 2008; Vinik and Flemmer, 2002). Thus, the detection of individuals with sub-clinical disease, for example, during pre-diabetes may facilitate intervention strategies to prevent or delay the progression to disease, ultimately reducing mortality and morbidity.

Research efforts to reduce the burden of T2D by early intervention and treatment are underpinned by the need for biomarkers to identify subclinical disease. Thus, the discovery of differentially expressed miRNAs between diabetic, pre-diabetic and normo-glycaemic individuals may have the potential as a pre-screening tool in high risk individuals. Moreover, identification of such miRNAs may provide insight into the pathophysiology of T2D and furthermore, may present novel therapeutic targets against diabetes.

1.11.1 Hypothesis

We hypothesize that miRNA expression patterns are associated with the development of diabetes and that their expression profiles change during disease progression. The identification of different miRNA expression profiles in diabetic, pre-diabetic and normo-glycaemic individuals could potentially represent early biomarkers for the development of T2D, and could facilitate intervention strategies to prevent, better manage or delay the development of T2D.

1.11.2 Aim

The aim of this study is to determine whether miRNA expression patterns differ in diabetic, pre-diabetic, and normo-glycaemic individuals, thereby offering the potential as biomarkers which could be used to identify individuals with a high risk of developing T2D.

1.11.3 Objectives

- Recruit subjects and obtain blood samples from diabetic, pre-diabetic and normo-glycaemic individuals,
- Measure metabolic parameters (FPG, OGTT, HbA1c, insulin, glucagon, C-peptides and lipid profile),
- Characterize participants according to WHO classification using FPG, OGTT and HbA1c test,
- Match participants according to age, gender, ethnicity and BMI,
- MiRNA expression profiling in PBMCs of diabetic, pre-diabetic and normo-glycaemic individuals using Illumina sequencing,
- Validation of selected differentially expressed miRNAs in PBMCs of diabetic, pre-diabetic and normo-glycaemic individuals by qRT-PCR,
- Validate the selected miRNAs in serum of diabetic, pre-diabetic and normo-glycaemic individuals by qRT-PCR, and
- Bioinformatic analysis to assess functional significance of miRNAs and their mRNA targets.

2. MATERIALS AND METHODS

2.1 Study design

A flow diagram illustrating the experimental protocol is shown in Figure 2.1. Briefly, 258 research participants were recruited to the study and were classified as diabetic (T2D), pre-diabetic or normo-glycaemic (control) according to WHO criteria (WHO, 2006). The preferred criteria used in the South African setting. These participants were matched according to age, gender, ethnicity and body mass index (BMI), and four individuals per group were selected for this study. Thereafter, miRNAs were profiled in peripheral blood mononuclear cells (PBMCs) and in serum samples of these individuals using next generation sequencing and qRT-PCR.

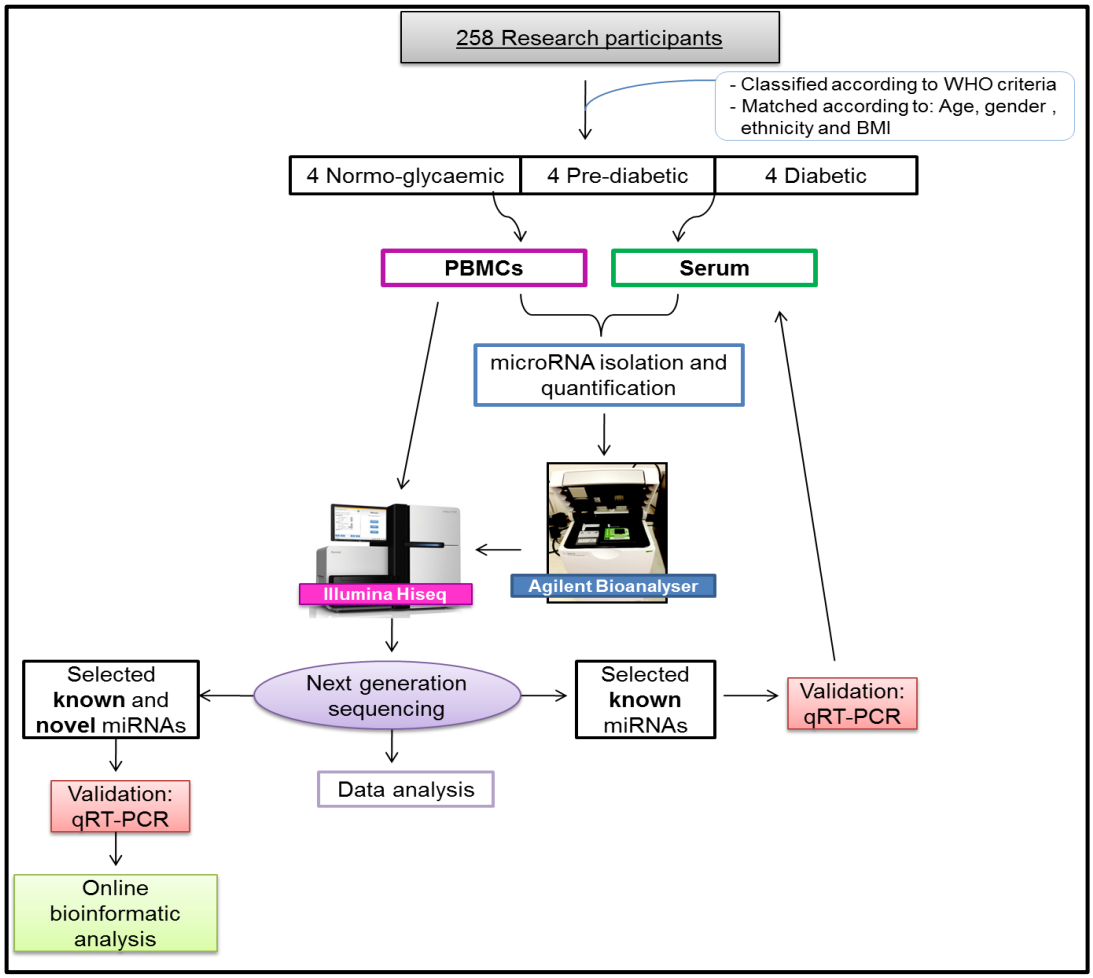


Figure 2.1 Study overview.

2.2 Subject recruitment

Relatives and friends of individuals from diabetes support groups or high risk individuals recommended by medical practitioners in the Cape Town Metropole, South Africa were invited to partake in the study. Individuals were predominantly from Malmesbury, while others were from Parow and Mitchells Plain. The inclusion criteria of the study were: 1) subjects that were never diagnosed with T2D, 2) subjects that were a first-degree blood relative of someone previously diagnosed with T2D i.e. a parent, child or full sister/brother, and 3) subjects that were a second-degree relative of someone previously diagnosed with diabetes i.e. a grandparent, grandchild, aunt/uncle, nephew/nieces or half-brother/sister. The exclusion criteria of the study were: 1) subjects that were on any form of medication for health related issues. The study aims and procedures were explained to subjects in a language they understood (Appendix 1). Subjects were recruited to the study after they provided written informed consent confirming that they understood and agreed to participate in the study (Appendix 2). All procedures were conducted according to Ethical Guidelines for research in South Africa and approval for the study was granted by the ethics committee of the South African Medical Research Council (EC010-5/2013) and the Human Research Ethics Committee of Stellenbosch University (S15/04/095). All participants were required to complete a questionnaire, whereafter bodyweight, height and blood pressure was measured with a calibrated scale (UC-321S digital, A&D Medical, San Jose, California) portable stadiometer (Seca 214, Seca GmbH & Co, Hamburg, Germany) and a blood pressure monitor (Omron, HiTech, Bryanston, South Africa), respectively. Staff who were involved in the study (conducted the interviews, anthropometric measurements and blood collection) were trained by the clinic site project manager at the start of the study.

2.3 Blood collection

Blood glucose concentrations were measured with a calibrated glucometer (OneTouch, Lifescan inc., Milpitas, California.) after an overnight fast of about 10 hours (hrs), using the second drop of blood obtained by finger prick. Thereafter, blood from the upper arm was collected by a trained nurse, phlebotomist or medical doctor. Twelve tubes of blood were collected from each participant; eight fasting and four during the oral glucose tolerance test (OGTT). Eight fasting blood samples were collected in one fluoride-containing tube (to reduce glycolysis) to measure glucose concentrations, four gel-containing tubes (induce coagulation and for collection of serum) to measure insulin, c-peptide, and lipid concentrations, and to determine protein markers, one Ethylenediaminetetraacetic acid

(EDTA) tube for HbA1c and DNA extraction, one EDTA tube containing trasylol for glucagon measurement, and finally in a PAXgene tube for the isolation of miRNA. Thereafter, subjects were required to drink 75 grams (g) of glucose in 100 ml of water, and blood was collected after 30 minutes (min) and 120 min in SST tubes and fluoride-containing tubes for measuring insulin and glucose concentrations. Samples were stored on ice, and were either sent to Pathcare (Pathcare, South Africa) or to the laboratory (BRIP, SAMRC) for processing using standard operating procedures.

For serum collection, BD Vacutainer® SST™ Tubes (Vacutainer, BD, Woodlands, South Africa) were stored at room temperature for 2 hrs prior to centrifugation (5810R, Eppendorf, Hamburg, Germany) at 4 Celsius (^{o}C) for 15 mins at 3220 g (relative centrifugal force). Each tube was coated with silicone and micronized silica particles to accelerate clotting of blood, and contains a gel that separates blood cells and plasma from serum. Once separated, the serum from each tube was aliquoted into cryotubes (250 µl) and stored at -80^{o}C for long term storage (Fig. 2.2A).

PAXgene® Blood RNA Tubes (PreAnalytix, Qiagen, Feldbachstrasse, Switzerland) contain 6.9 ml of RNA stabilising reagent that lyses peripheral blood mononuclear cells (PBMCs) and immediately stabilizes intracellular RNA by preventing *in vitro* RNA degradation and minimizing gene induction (Fig. 2.2B). Briefly, 2.5 ml of blood was collected into the tubes and processed according to the manufacturer's instructions. Tubes were gently inverted 10 times immediately after blood collection, and thereafter stored upright at room temperature (20ºC to 25ºC) for 2 hrs, transferred to 4ºC for 24 hrs, -20°C for 24 hrs and finally transferred to -80ºC for long-term storage upright in wire or metal racks.

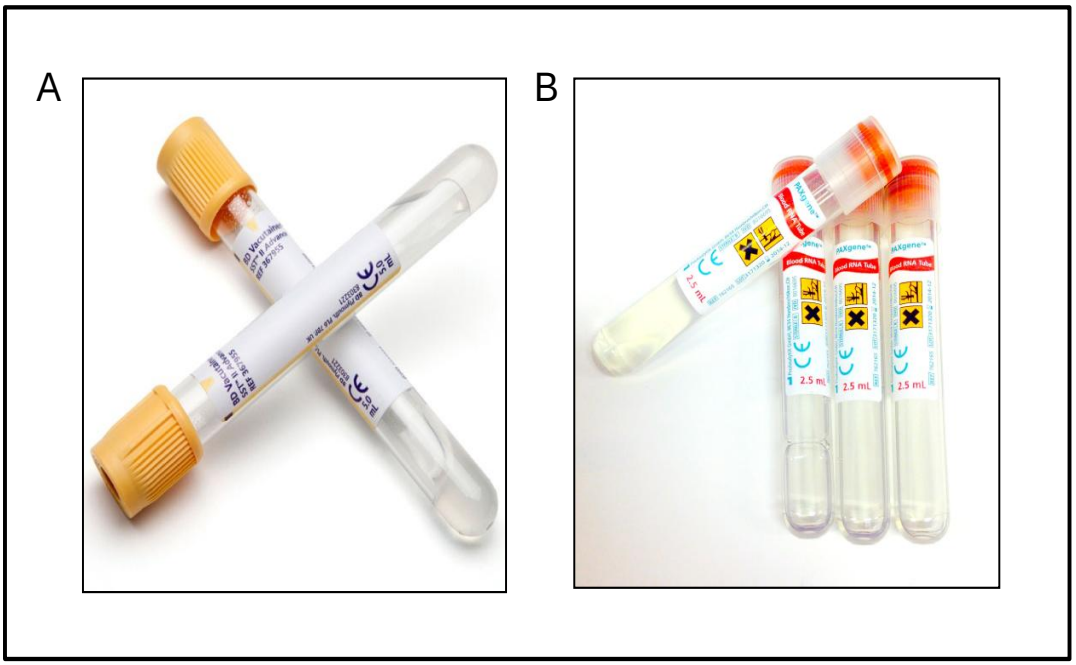


Figure 2.2 An illustration of A) BD vacutainer tubes and B) PAXgene blood RNA tubes for collecting blood.

2.4 Selection of subjects for microRNA profiling

Subjects were characterized as diabetic, pre-diabetic and normo-glycaemic individuals based on their OGTT, FPG and HbA1c concentrations, according to the WHO diagnostic classification criteria (Fig. 2.3, WHO, 2006). Subjects with OGTT, FPG and HbA1c levels ≥11.0 mmol/L, ≥7.0 mmol/L and ≥6.5 %, respectively, were classified as diabetic, while subjects with OGTT and FPG levels between 7.8 mmol/L-10.9 mmol/L and 6.1 mmol/L-6.9 mmol/L, respectively, were classified as pre-diabetic. Furthermore, subjects with OGTT and FPG levels <7.8 mmol/L and <6.1 mmol/L, respectively, were characterized as normo-glycaemic individuals. Thereafter, subjects in each group were matched according to age, gender, ethnicity and BMI for miRNA profiling of their blood samples.

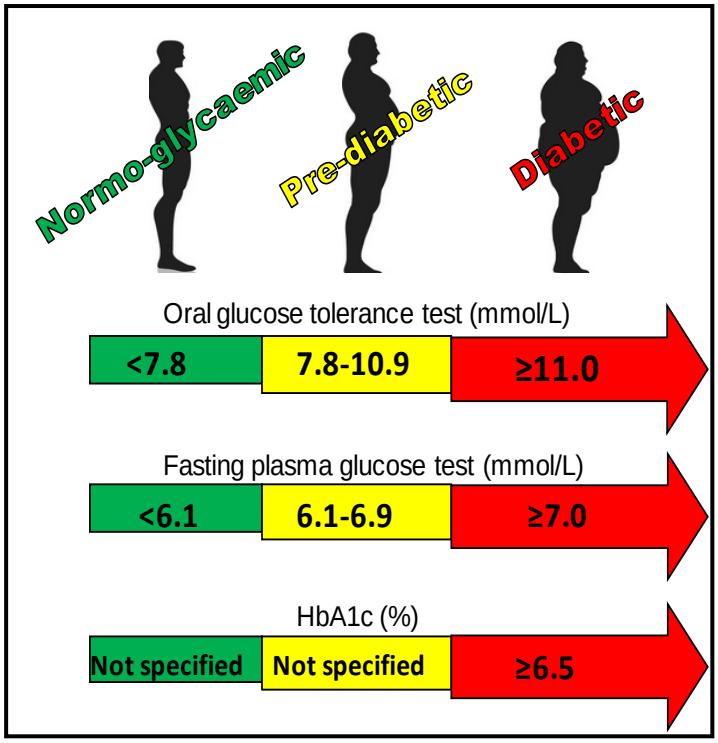


Figure 2.3 WHO criteria for classification of diabetic, pre-diabetic and normo-glycaemic (WHO, 2006).

2.5 RNA isolation and quantification:

RNA was isolated from the PBMCs and serum of diabetic, pre-diabetic and normo-glycaemic individuals, using the PAXgene blood miRNA kit (Qiagen, Hilden, Germany), and the miRNeasy Serum/Plasma kit (Qiagen, Hilden, Germany) respectively.

2.5.1 Peripheral blood mononuclear cells

PAXgene® Blood RNA tubes were thawed gradually according to the manufacturer's recommendations. Briefly, tubes were removed from -80°C and allowed to thaw by transferring them to -20°C for 24 hrs and thereafter to room temperature (20°C - 25°C) for 2 hrs. Once thawed, tubes were inverted 10 times whereafter RNA extraction was conducted using the PAXgene blood miRNA kit according to the manufacturer's instructions. Peripheral blood mononuclear cells were pelleted by centrifuging PAXgene blood RNA tubes at 3220 g for 10 mins using a swing-out rotor (5810R, Eppendorf, Hamburg, Germany). Thereafter, cells were washed with 4 ml of RNAse free water, vortexed for 10 seconds (s) until the pellet

was visibly dissolved, and re-centrifuged at 3220 g for 10 mins. The supernatant was removed by decanting and inverting the tube, the cell pellet was resuspended in 350 µl of buffer BM1 and vortexed for 10 mins until the pellet was visibly dissolved. Thereafter, samples were digested by adding 300 µl of buffer BM2 and 40 µl of proteinase K (20 mg/ml), vortexed for 5 s and incubated for 10 mins at 55°C in a shaking incubator at 1400 revolutions per minute (rpm) (Qiagen, Retsch, Germany).

After incubation, samples were homogenized by pipetting, transferred onto a PAXgene Shredder spin column in a 2 ml tube, and centrifuged for 3 mins at full speed (16100 g) (5415R, Eppendorf, Hamburg, Germany). Supernatants were transferred into new tubes, 700 µl of isoproponol was added and samples were mixed vigorously by vortexing. Thereafter, 700 µl of sample was pipetted onto a PAXgene spin column in a 2 ml tube, and centrifuged at 16100 g for 1 min. The spin column was placed in a new 2 ml tube, and the remaining sample pipetted into the PAXgene spin column, and centrifuged at 16100 g for 1 min. Thereafter, 350 µl of buffer BM3 was added to the spin column and centrifuged at 16100 g for 15 s. RNA was then subjected to on-column DNase digestion by pipetting 80 µl of DNase I incubation mix (10 µl DNase 1 stock solution added to 70 µl buffer RDD) directly onto the PAXgene spin column membrane, and incubating at room temperature (20°C - 25°C) for 15 mins. Samples were washed by adding 350 µl and 500 µl of buffer BM3 and BM4, respectively, to the spin columns, and centrifuged at 16100 g for 15 s, while discarding the flow through each time. Another 500 µl of buffer BM4 was added to the PAXgene spin column, and centrifuged at 16100 g for 2 mins. The PAXgene RNA spin column was placed in a new 2 ml tube, and centrifuged at 16100 g for 1 min to completely dry the spin column membrane. Thereafter, 40 µl of elution buffer BR5 was directly pipetted onto the spin column membrane, and centrifuged at 16100 g for 1 min to elute the RNA. RNA was denatured by incubating at 65^{o}C for 5 mins and immediately chilling it on ice.

RNA concentrations were determined by measuring the absorbance at a wavelength of 260 nm (A_{260}) using a nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, USA). RNA purity was assessed by using the $A_{260/280}$ (protein contamination) and $A_{260/230}$ (salt contamination) ratios, where a ratio of 1.8 - 2 indicates a high level of purity and is assumed suitable for gene expression measurements.

2.5.2 Serum

Frozen serum aliquots in cryotubes were removed from -80°C and allowed to thaw by incubating at 37°C in a water bath for 3 mins. Once samples were thawed, miRNA-enriched total RNA was extracted from 200 µl of serum using the miRNeasy Serum/Plasma kit, according to the manufacturer's instructions. One thousand microliters of QIAzol lysis reagent was added to each serum sample, mixed by vortexing for 5 s, and left on the benchtop at room temperature (15°C - 25°C) for 5 mins. Thereafter, 3.5 µl of MiRNeasy serum/plasma spike-in-control (*Caenorhabiditis elegans*-miR-39) (1.6 x 10^8 copies/µl working solution) was added as an exogenous synthetic miRNA to control for technical variation during RNA extraction, cDNA synthesis and qRT-PCR processing. Two hundred microliters of chloroform was then added to each sample and vortexed for 15 s for subsequent phase separation. Samples were left on the benchtop at room temperature (15°C - 25°C) for 2 to 3 mins, and then centrifuged (5415R, Eppendorf, Hamburg, Germany) at 4°C for 15 mins at 16100 g. After centrifugation, the upper colourless aqueous phase containing approximately 600 µl of RNA was transferred to a new collection tube, and mixed thoroughly with 900 µl of 100% ethanol. Thereafter, 700 µl of sample was pipetted into an RNeasy MinElute spin column, and centrifuged at 11200 g for 15 s at room temperature (15°C - 25°C). This step was repeated, and the flow through was discarded both times. Samples were then washed with 700 µl and 500 µl of buffer RWT and RPE, respectively, and centrifuged at 11200 g for 15 s. For a final wash, 500 µl of 80% ethanol was added into the RNeasy MinElute spin column, and centrifuged for 2 mins at 11200 g at room temperature (15°C - 25°C). The spin column was placed in a new 2 ml processing tube, and centrifuged at 16100 g for 5 mins to completely dry the spin column membrane. After drying, the spin column was placed in a 1.5 ml tube, and 14 µl of RNase-free water was directly pipetted onto the spin column membrane, and centrifuged at 16100 g for 1 min to elute the RNA.

Total RNA concentration was measured using a nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, USA) as mentioned previously.

2.6 Assessment of RNA integrity

RNA quality is known to affect the efficiency and accuracy of downstream experiments and high quality RNA is essential to obtain reliable gene expression data (Fleige and Pfaffl, 2006). In mammalians, the 28S/18S ratio is used to assess RNA integrity, where a ratio of 2:1 is generally representative of good quality RNA, and is considered optimal for downstream applications (Wieczorek *et al.*, 2012). RNA quality has historically been assessed using agarose or acrylamide gel electrophoresis to quantitatively and qualitatively measure RNA concentrations. However, newer methods such as the Agilent 2100 bioanalyser (Agilent Technologies, CA, USA) have revolutionized RNA integrity determination since they require less RNA sample and time. The bioanalyser is an automated device, utilizing microfluidic technology to analyze the quality of RNA on a sample-specific chip, using only 1 µl of RNA sample. The analysis can be completed in approximately 30 to 40 mins.

2.6.1 Total RNA

To assess the quality of the total RNA extracted using the PAXgene blood miRNA kit, the Agilent RNA 6000 Nano kit (Agilent Technologies, CA, USA) was used, according to the manufacturer's instructions. Firstly, 9 µl of fluorescent gel-dye mixture was pipetted into the nanochip, and primed using a plunger to form micro-channels consisting of interconnected networks (Fig. 2.4A). The plunger was held down for exactly 30 s before releasing the clip, and then slowly pulled back to the 1 ml position after 5 s. Another 9 µl gel-dye mixture was added to a different well in the nanochip. For calibration, 5 µl of RNA marker was added into all the sample wells, as well as to the RNA ladder well, which was included as a positive control to align all plots. The samples and RNA ladder were denatured for 2 mins at 70^{o}C and loaded into individual wells of the chip. Thereafter the chip was vortexed at 1000 rpm for 1 min (Fig. 2.4B), and immediately loaded onto the Agilent 2100 Bioanalyser (Fig. 2.4C)

.

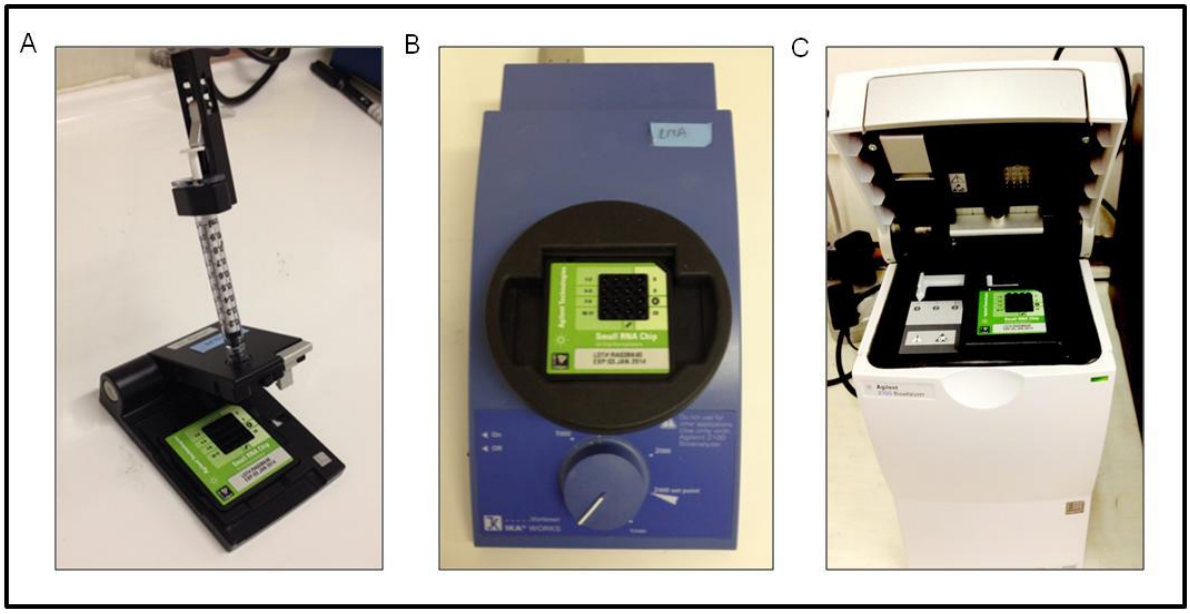


Figure 2.4 Representative image of an Agilent LabChip assay with A) priming station and plunger, B) the vortex used for the Agilent chip, and C) an Agilent 2100 Bioanalyser used to determined RNA quality.

In the bioanalyser, 16-pin electrodes are arranged to fit perfectly into each well of the chip and are connected to an independent power supply that provides maximum control and flexibility. Charged RNA biomolecules are electrophoretically driven by a voltage gradient of constant mass-to-charge ratio that separates these molecules according to size in the presence of the sieving polymer matrix. During the chip run, the dye intercalates directly with the RNA, and is detected via laser-induced fluorescence detection. The amount of measured fluorescence correlates directly with the amount of RNA of a given size (Mariam Ayabi, 2007). A RNA ladder containing a range of different size fragments is used as a reference for data analysis. The data is then visualized as an electropherogram (Fig. 2.5A), or translated into a virtual gel image (Fig. 2.5B). The 28S:18S ribosomal ratio is calculated by assessing the bands or peaks of each sample, and represents the quality of the RNA. The integrity of each sample is represented by an RNA integrity number (RIN) ranging between 1 and 10, where 10 represents high quality intact RNA and 1 represents a completely degraded RNA sample (Mueller *et al.*, 2004). All samples with a RIN greater than 7 were used in this study.

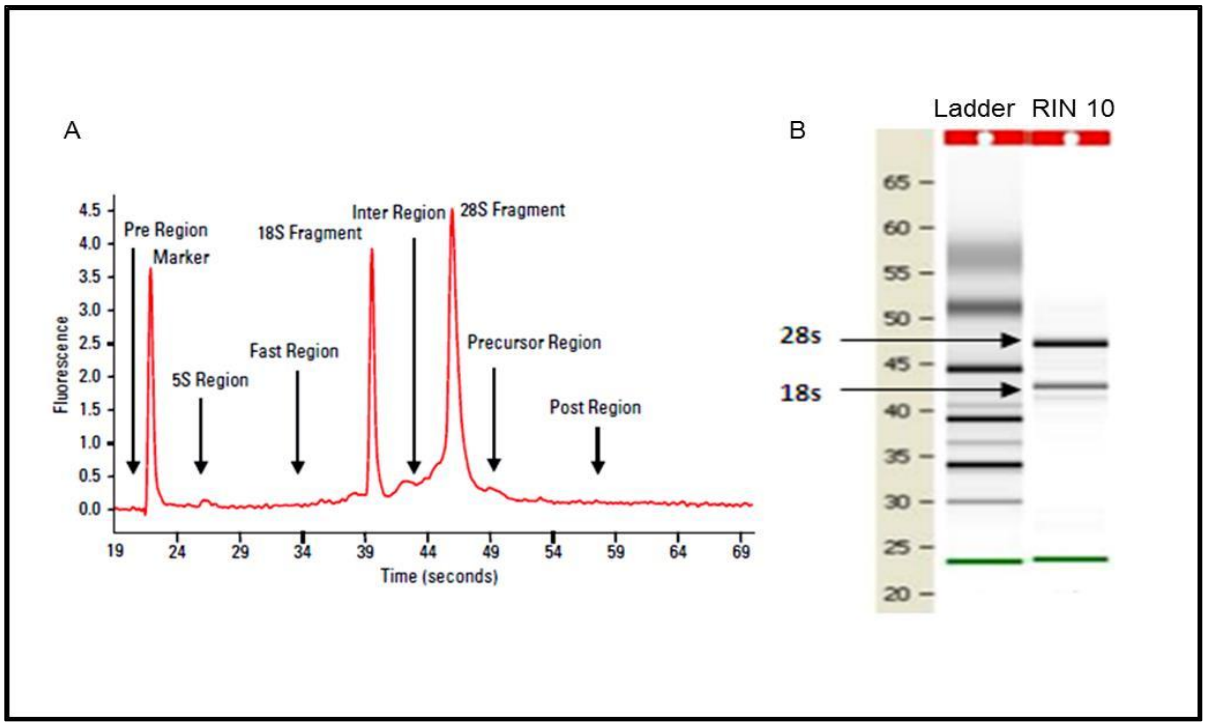


Figure 2.5 Example of a bioanalyser electropherogram and a virtual gel image detailing the regions that are indicative of good quality RNA. A) An electropherogram illustrating peaks representing variable rRNA sizes. B) A virtual gel illustrating two clear ribosomal bands illustrating 28S and 18S, with a RIN of 10 (Adapted from Mueller *et al.* 2004).

2.6.2 MicroRNA

MiRNAs are a unique class of small sized nucleic acids that can be accurately examined with the same technologies as total RNA. MicroRNA quality and its percentage relative to total RNA, was assessed using the Agilent small RNA (Pico) kit (Agilent Technologies, CA, USA), according to manufacturer's instructions. This kit enables the detection and analysis of small RNAs (<200 nt) and miRNAs (15 - 40 nt) within a small RNA fraction (Fig. 2.6). The concentration of miRNA is calculated as an absolute amount (pg/µl) and as a percentage (%) of small RNA in the total RNA sample. Briefly, the same procedures for miRNA analysis was followed as described previously for total RNA (section 2.6.1), except that: 1) the plunger in the priming station (which forms micro-channels) was held down for 60 s instead of 30 s and 2) the syringe clip was adjusted from the highest to the lowest position.

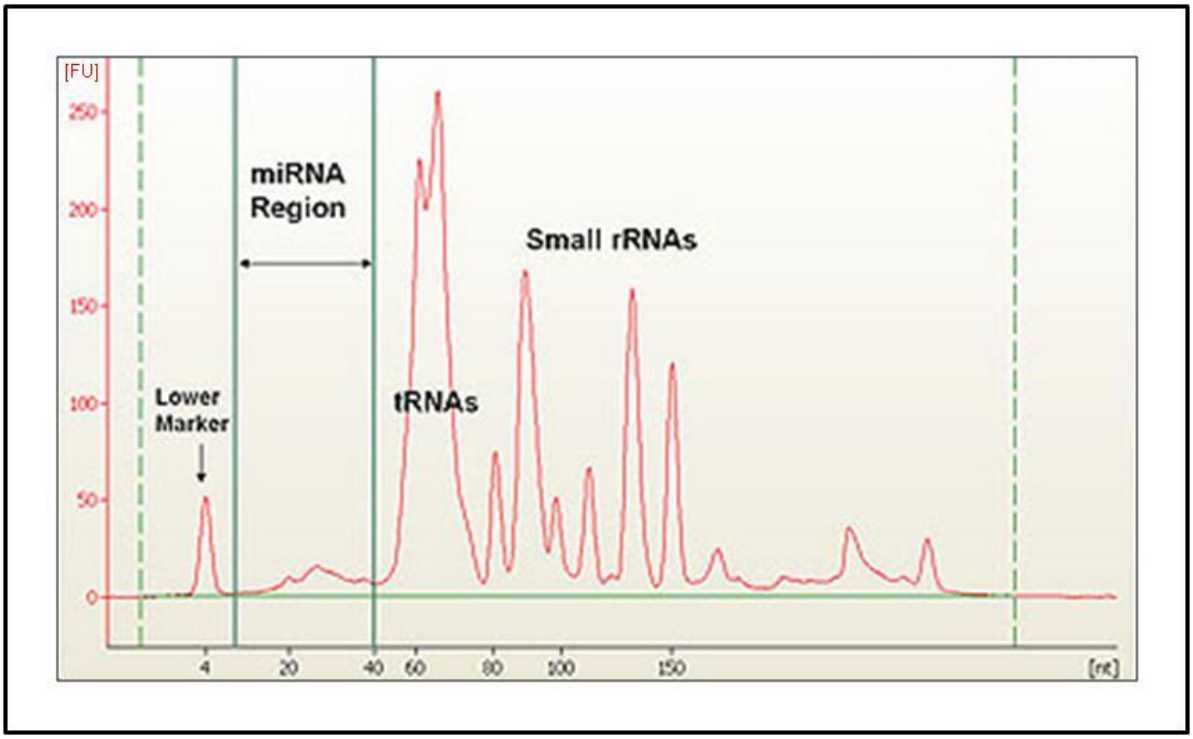


Figure 2.6 Representative image of an electropherogram for small RNA analysis, showing the content of small RNAs (miRNAs, tRNAs, and small rRNAs) that form part of the total RNA sample (Adapted from: Agilent Technologies, 2007).

2.7 MicroRNA sequencing

To gain insight into miRNA regulation during T2D progression, miRNAs were isolated from PBMCs of diabetic, pre-diabetic and normo-glycaemic individuals and subjected to high throughput miRNA sequencing (Arraystar, Rockville, USA). MiRNA sequencing has a wide dynamic range and is a highly sensitive and accurate technique that can distinguish between very similar miRNA sequences and isoforms, and is able to discover novel uncharacterized miRNAs (Motameny *et al.*, 2010).

Approximately 2 µg of RNA from age, gender, BMI and ethnicity matched diabetic (n=4), pre-diabetic (n=4) and normo-glycaemic individuals (n=4) were lyophilized by freeze-drying according to the manufacturer's instructions (VirTis Benchtop K, United Scientific), using 100 mTorr optimum vacuum pressure at -113^{o}C (condenser temperature) for approximately 90 mins. Thereafter, samples were stored at -80^{o}C for 24 hrs until shipping on dry ice to the USA, as recommended by Arraystar. MicroRNA sequencing was performed on an Illumina

Hiseq 2000 instrument (Illumina Inc., San Diego, USA), using TruSeq Rapid SBS kits (Illumina), according to the manufacturer's instructions. The procedures followed by Arraystar are described below.

Firstly, lyophilized samples were resuspended in water, and the concentrations of the RNA samples were measured using the NanoDrop ND-1000 instrument. The A_{260}/A_{280} ratios were used to assess purity, and each sample was subjected to miRNA sequencing. A miRNA sequencing library was prepared as follows: firstly, total RNA was subjected to polyacrylamide gel electrophoresis and bands corresponding to the size of the miRNAs (~15 - 35 nt) were excised from the gel for further processing (Fig. 2.7B). Thereafter, 3' and 5' small RNA adapters were sequentially ligated to the size selected RNA molecules, using T4 RNA ligase 2 (Fig. 2.7C). Adapters act as binding sites for primers, allowing reverse transcription and complementary DNA (cDNA) synthesis to form a cDNA library. Thereafter, the cDNA library was run on a polyacrylamide gel, and the band corresponding to the molecular size of miRNA fragments with ligated adapters (~135 - 155 bp) was excised from the gel for subsequent sequencing (Fig. 2.7D). The cDNA libraries were quantified by the Agilent 2100 Bioanalyser, using the Agilent DNA 1000 chip kit (Agilent Technologies), and denatured with 0.1M NaOH. The samples were diluted to a final concentration of 8 pM and cluster generation was performed on the Illumina cBot using the TruSeq Rapid SR cluster kit, according to manufacturer's instructions. Single stranded DNA molecules were captured on Illumina flow cells, amplified *in situ a*s clusters and sequenced for 36 cycles on the Illumina Hiseq 2000, according to the manufacturer's instructions. Several basic data processing steps were then performed to extract the relevant information from the raw data as described below.

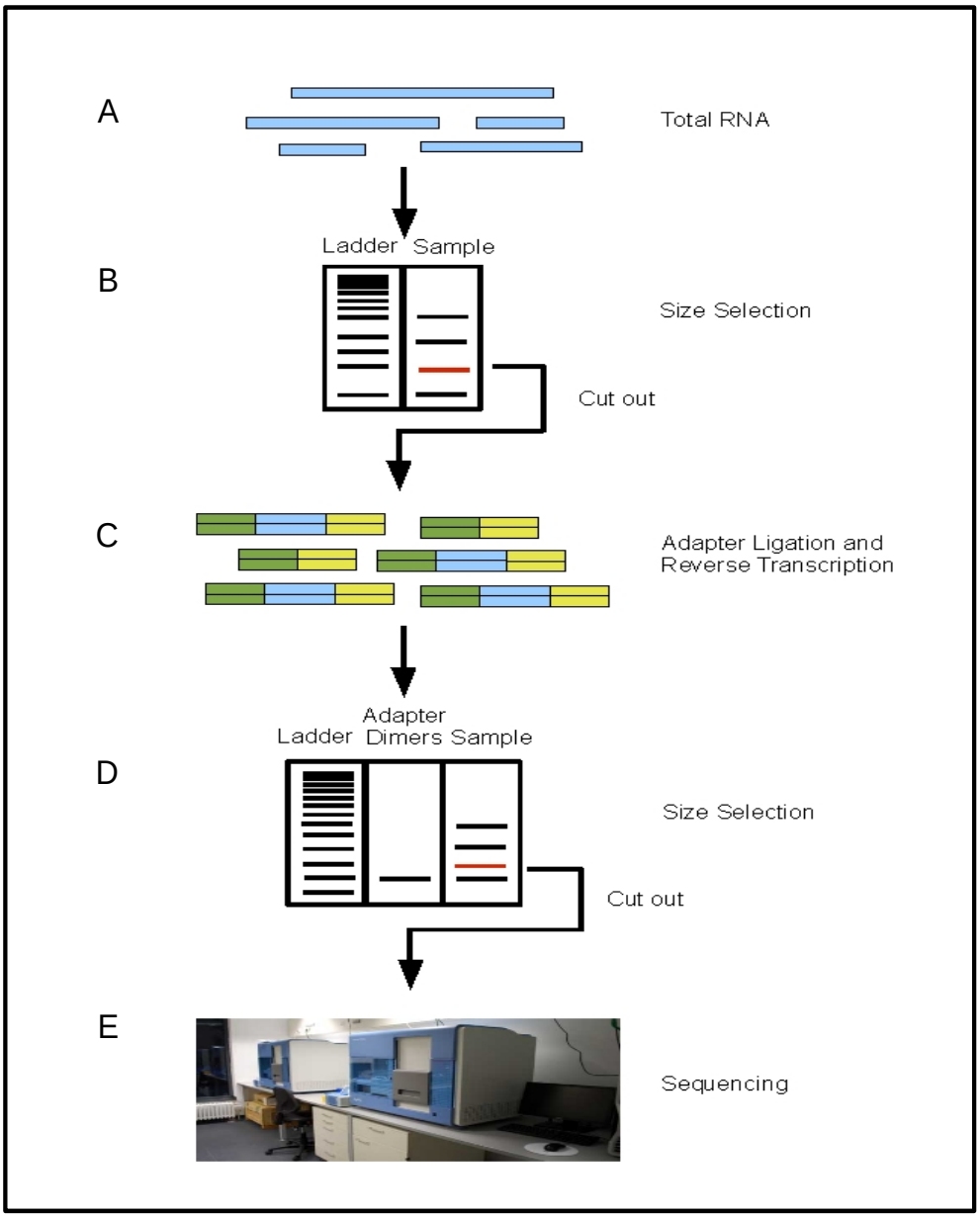


Figure 2.7 A Schematic representation of the miRNA sequencing procedure (Adapted from Motameny *et al.* 2010).

The raw data (also called reads) were generated by aligning the sequences of expressed miRNAs to the reference human genome in miRBase version (v) 21, to produce clean reads. Clean reads were provided as text files in FASTQ format as illustrated in Figure 2.8. These files contain four lines per read: the first line is the title line which contains the '@' symbol, followed by a (unique) read identifier; the second line contains the read's nucleotide sequence; the third line contains the read identifier again, preceded by a '+' symbol to signal the start of the quality string and the end of the sequence lines; and the fourth line contains

quality encoding scores (Cock *et al.*, 2010). The missing nucleotides in the read sequence were denoted with an 'N' character, underlined in red (Fig. 2.8). All sequence reads start at the first base after the 5' sequence adapter, and typically ends after 36 bp, thus excluding the 5' adapter sequence from the reads. However, all reads contained part of the 3' adapter sequence, and to eliminate ambiguity of sequence detection, these 3' adapter sequences were removed from the mature miRNA sequence (trimmed reads). Trimmed reads were recorded in FASTA format.

	Raw_Data (36bp)
	Sample_sequence.fastq (FASTQ format)
Line 1	@HWUSI-EAS1522_111013:1:1:6701:1030#0/1
Line 2	ANGAGCAGCATTGTATAGGGCTCTGTAGGCACCATC
Line 3	+HWUSI-EAS1522_111013:1:1:6701:1030#0/1
Line 4	[DFMRTSXSUWZVMOHZW]]][[]]_aaacb[abac
Line 1	@HWUSI-EAS1522_111013:1:1:9117:1032#0/1
Line 2	ANGAATGACACGATCACTCCCGTTGACTGTAGGCAC
Line 3	+HWUSI-EAS1522_111013:1:1:9117:1032#0/1
Line 4	^DX]VG_\W_[aYS\Hdddaaadadf[acaf[dafc

Figure 2.8 An extract of the sequence data in FASTQ format illustrating the four lines per read.

All trimmed reads greater than or equal to 15 nt in length, with zero or one base pair mismatch were aligned to the reference human genome in mirBase, to identify known miRNAs from previously annotated miRNAs, using Novoalign software, v2.07.11. All other reads were discarded, and the alignment results were saved in text files. The miRNA isoforms with the highest read counts were used to estimate the expression levels of each miRNA. The expression of each unique sequence (s) occurring among the reads in the sample were normalised (as reads per million) against the total number of reads produced for the sample, to estimate the relative expression level of each miRNA (Equation 2.1) (Motameny *et al.* 2010), and to compare expression patterns between groups (normo-glycaemic vs. pre-diabetic; normo-glycaemic vs. diabetic and pre-diabetic vs. diabetic). Significantly differentially expressed miRNAs were classified as those having fold changes (FC) $\geq$ 1.1 and *p*-values $\leq$ 0.05 (threshold values set by ArrayStar). Furthermore, novel miRNAs were predicted from the miRNA sequencing data using miRDeep algorithms. This uses a score compatibility model of the position and frequency of the mature miRNA sequence with the secondary structure of the miRNA precursor (Friedlander *et al.*, 2008).

$$reads\ per\ million = \frac{\text{number of reads in the sample that are equal to s}}{\text{total number of reads in the sample}} \times 1{,}000{,}000$$

Equation 2.1 Calculation used to determine reads in each sample.

2.8 Quantitative real time PCR

Quantitative real-time PCR (qRT-PCR) is a sensitive, accurate and reliable method to measure miRNA expression, and is often used to validate miRNA sequencing data. This method enables the detection and measurement of the exact amount (relative or quantitative) of DNA or RNA product, by determining the number of copies of DNA or RNA in each sample. When RNA is used as a template, reverse transcription before amplification of the mRNA is necessary in order to quantify gene expression (Livak and Schimittgen, 2001).

2.8.1 Reverse transcription

Reverse transcription refers to the process whereby miRNA is reverse transcribed into cDNA, using reverse transcriptase. The cDNA is then used as a template for exponential amplification, using PCR. In this study, the TaqMan® MicroRNA Reverse Transcription (RT) kit (Life Technologies, Carlsbad, USA) and the miScript II RT kit (Qiagen, Hilden, Germany) were used, according to the manufacturer's instructions. The Taqman RT kit uses highly stable, small miRNA-specific stem-loop RT primers which lengthen the target cDNA, thus reverse transcribing mature miRNAs. On the other hand, the miScript II RT kit uses random hexamers or oligo-(dT) primers that randomly binds to and reverse transcribes large fragments of mRNA, to form cDNA.

2.8.2 PCR

The two most common technologies for the detection of PCR products in qRT-PCR are Taqman probes and SYBR Green. Taqman probes are sequence specific oligonucleotides

that are labelled with a fluorescent reporter, such as FAM™ and a non-fluorescent quencher (NFQ), which is covalently attached to the 5' and 3' end of the probe, respectively. When the fluorescent reporter dye and quencher molecule are in close proximity, emission of the reporter dye is prevented. Furthermore, *Taq* polymerase extends the amplicon, making new complementary DNA before it reaches the 5' end of the TaqMan probe, where upon the 5'-3' exonuclease activity of *Taq* polymerase hydrolyses the probe base by base releasing the fluorescent reporter dye. Once the fluorescent reporter dye and quencher dye are no longer in close proximity, fluorescence is emitted and a signal is collected (Benes and Castoldi, 2010).

Using specific optical system software, fluorescence is measured at each cycle for each product and the signal is collected at a specific cycle threshold (C_t) on the Applied Biosystem 7500 instrument (Thermo Scientific™, MA, USA). The fluorescence is directly proportional to the amount of PCR product generated, and therefore permits a specific, quantitative measure of miRNA expression levels in each sample. SYBR Green is a fluorescent dye that intercalates with any double-stranded DNA resulting in light emission. SYBR green I dye will detect all double-stranded cDNA, including non-specific reaction products, while Taqman probes detects only specific amplification products (Bustin and Mueller 2005).

2.8.2.1 Taqman probes

MiRNA sequencing data was validated using Taqman probes. Briefly, 10 ng of miRNA-enriched total RNA was reverse transcribed into cDNA with small RNA-specific, stem-loop RT primers specific to the microRNAs of interest (Table 2.1), and to custom designed primers for novel miRNA detection (Table 2.2), using the TaqMan® MicroRNA Reverse Transcription Kit according to the manufacturer's instructions. Thereafter, 15 µl of cDNA was amplified with TaqMan probes specific to the miRNAs of interest (Table 2.1 and 2.2), and TaqMan® Universal PCR Master Mix II.

RNU6b, miR-454 and miR-425 were used as endogenous reference miRNAs (Table 2.1). Normfinder; a mathematical model of gene expression, that uses an algorithm to estimate overall variation of reference genes and variation between subgroups of the same sample set was used to identify the best endogenous reference miRNA (Andersen *et al.* 2004). The delta delta Ct ($2^{-\Delta\Delta CT}$) method was used to quantify relative expression using equation 2.2.

$$\Delta Ct = Ct(\text{target gene}) - Ct(\text{reference gene}) \quad [\text{Test sample}] \quad (a)$$

$$\Delta CT = Ct(\text{target gene}) - Ct(\text{reference gene}) \quad [\text{Calibration sample}] \quad (b)$$

$$\Delta\Delta CT = \Delta CT[\text{Test sample}] - \Delta CT[\text{Calibration sample}] \quad (c)$$

$$\text{Relative Quatification (RQ)} = 2^{-\Delta\Delta CT} \quad (d)$$

Equation 2.2 Calculation used to determine relative miRNA expression.

Table 2.1 MiRNA primer and probe assays used for validation analysis.

Assay name†	miRbase ID/ NCBI acc. #	Assay ID	Species	Mature microRNA sequence
hsa-miR-27b	hsa-miR-27b-3p	000409	Human	UUCACAGUGGCUAAGUUCUGC*±
hsa-miR-98	hsa-miR-98-5p	000577	Human	UGAGGUAGUAAGUUGUAUUGUU*±
hsa-miR-143	hsa-miR-143-3p	002249	Human	UGAGAUGAAGCACUGUAGCUC*±
hsa-miR-21	hsa-miR-21-3p	002438	Human	CAACACCAGUCGAUGGGCUGU*±
hsa-miR-379	hsa-miR-379-5p	001138	Human	UGGUAGACUAUGGAACGUAGG*±
hsa-miR-425	hsa-miR-425-5p	001516	Human	AAUGACACGAUCACUCCCGUUGA±
hsa-miR-454	hsa-miR-454-3p	002323	Human	UAGUGCAAUAUUGCUUAUAGGGU±
RNU6B	NR_002752 #	001093	Human	CGCAAGGAUGACACGCAAAUUCGU GAAGCGUUCCAUAUUUUU±

†miR refers to the mature form of the miRNA, while the first three letters (hsa) signifies the organism (human in this case) that the miRNA is expressed in. The number following 'miR' refers to the order in which they were discovered and published, and the 3p or 5p signifies the miRNA precursor arm (Wright and Bruford, 2011).

#NCBI miRNA accession number

*Mature miRNA sequence selected from the sequencing data

±Mature miRNA sequence selected from Applied Biosystem's database

Table 2.2 Custom designed primer and probe assays for validation of novel miRNAs.

Assay name[#]	Mature ID[†]	Assay ID	Species	Mature miRNA sequence
MYNO59	hsa-miR-novel-chr6_11861	CSCSU6	human	UACUUGACCUUGACUCUCCCU[±]
MYNO8	hsa-miR-novel-chr9_17319	CSD1TDE	human	AAGUUUCUCUGAACGUGUAGAGC[±]
MYNO66	hsa-miR-novel-chr22_33839	CSGJPPU	human	GAUGCCUGGGAGUUGCGAUCUG[±]
MYNO95	hsa-miR-novel-chr12_22392	CS39QZ3	human	AAAGCAAAUGUUGGGUGAACGG[±]
MYNO22	hsa-miR-novel-chr15_25713	CS5IO6B	human	UCCCUGUCCUCCAGGAGCU[±]

[#]Temporary primer assay name

[†]Mature novel miRNA ID illustrating its chromosomal location

[±]Mature sequence selected from sequencing data

2.8.2.2 SYBR Green

MiRNA detection in serum samples was conducted using the miScript PCR SYBR Green detection kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Briefly, 90 ng of total RNA in a 20 µl reaction was reverse transcribed into cDNA, using the miScript Hispec Buffer, nucleic acid and reverse transcriptase mix from the miScript II RT kit, according to the manufacturer's recommendations.

Complimentary DNA was diluted with 200 µl of RNAse free water and 1 µl was used as a template for real-time PCR analysis, using miRNA-specific miScript primer assays (Table 2.1) as forward primers and the miScript universal primer as the reverse primer and miScript SYBR Green PCR master mix, in a total volume of 10 µl. Relative expression was quantified using the $2^{-\Delta\Delta CT}$ method as described in section 2.8.2.1.

2.9 Messenger RNA target prediction analysis

Bioinformatic target prediction analysis was performed to predict the mRNA targets of miR-27b, miR-143, miR-98, miR-21 and miR-379. Messenger RNA targets are predicted according to seed match complementarity between mature miRNAs and the 3' untranslated region (UTR) of their mRNA target, sequence conservation, free energy and target site accessibility. Table 2.3 lists three mRNA target prediction programs (TargetScan-Human (TargetScan), Probability of Interaction by Target Accessibility (PITA) and DNA Intelligent Analysis (DIANA)-microT-CDS) that use different algorithms for target prediction (Fig. 2.9).

The TargetScan algorithm integrates thermodynamic based modeling of miRNA:mRNA interaction and searches for perfect complementarity and sequence conservation to predict miRNA targets conserved across multiple genomes. Binding of the 6 nt miRNA seed region forms perfect base pairing complementarity to the 3' UTR target mRNA region, and then extends to binding long fragments outside the seed region to filter out false positives (Friedman *et al.*, 2009; Lewis *et al.*, 2005). Thus, defining the score for miRNA:mRNA target interactions includes features such as 3' complementary pairing, local adenine and uracil content and position contribution of miRNA:mRNA binding (Lewis *et al.*, 2005; Grimson *et al.*, 2007). Compared to TargetScan, the DIANA algorithm uses a much longer frame (38 nt) for scanning complementarity, and focuses on coding regions of target mRNAs to increase specificity and sensitivity (Min and Yoon, 2010). It also calculates the minimum binding energy between miRNAs and sequences in the mRNA 3' UTR target region, which generates a scoring system to help evaluate the significance of the predicted targets (Kiriakidou *et al.*, 2004; Maragkakis *et al.*, 2009). Additionally, DIANA allows mismatches and weak binding at the 5' seed region, involving six consecutively paired nucleotides within the seed region or guanine:uracil (G:U) wobble pairs if additional base pairing between the miRNA 3' end and target region exists (Min and Yoon, 2010). Furthermore, the PITA algorithm predicts targets based on free energy and target site accessibility, by calculating the free energy gained from the miRNA:mRNA duplex and the energy cost of unpairing the target to make it accessible to the miRNA (Witkos *et al.*, 2011).

Table 2.3 Computational methods used for miRNA target prediction analysis.

Target prediction Program	Algorithm used for target prediction	Advantages	Disadvantages	URL	References
TargetScan-Human	Seed match; Sequence conservation; 3' complementarity; local AU content and position contribution	Parameters including in target scoring; Final score correlates with protein downregulation	Sites with poor seed paring are omitted	http://targetscan.org/	(Lewis *et al.*, 2005)
DIANA-microT-CDS	Free energy binding and complementarity	Signal-to-noise ratio and probability given to each target site-possibility of using own miRNA sequence as input	Some miRNA with multiple target sites may be omited	http://diana.imis.athena-innovation.gr/DianaTools/index.php	(Kariakidou *et al.*, 2004)
PITA	Target site accessibility; Free energy	The secondary structure of 3'UTR is considered for miRNA interaction	Low efficiency compared to other algorithms	http://genie.weizmann.ac.il/pubs/mir07/mir07_dyn_data.html	(Kertesz *et al.*, 2007)

UTRs = untranslated regions; PITA = Probability of Interaction by Target Accessibility; DIANA = DNA Intelligent Analysis; AU = adenine/uracil; URL = Uniform Resource Locator (website address)

The common mRNA target gene lists produced by each program, were filtered using an interactive venn diagram tool called Venny v2.0.2 (http://bioinfogp.cnb.csic.es/tools/venny_old/index.html) (Oliveros, 2007). Common mRNA targets were further analyzed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) to identify Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Fig. 2.9). Thereafter, genes potentially linked to T2D were manually selected and confirmed using an online Type 2 diabetes-database (T2D-db), and their interacting proteins were determined using a Search Tool for the Retrieval of Interacting Genes/Proteins (STRING). In addition, experimentally validated miRNA:mRNA interactions were identified for each miRNA using DIANA-TarBase and MiR-TarBase.

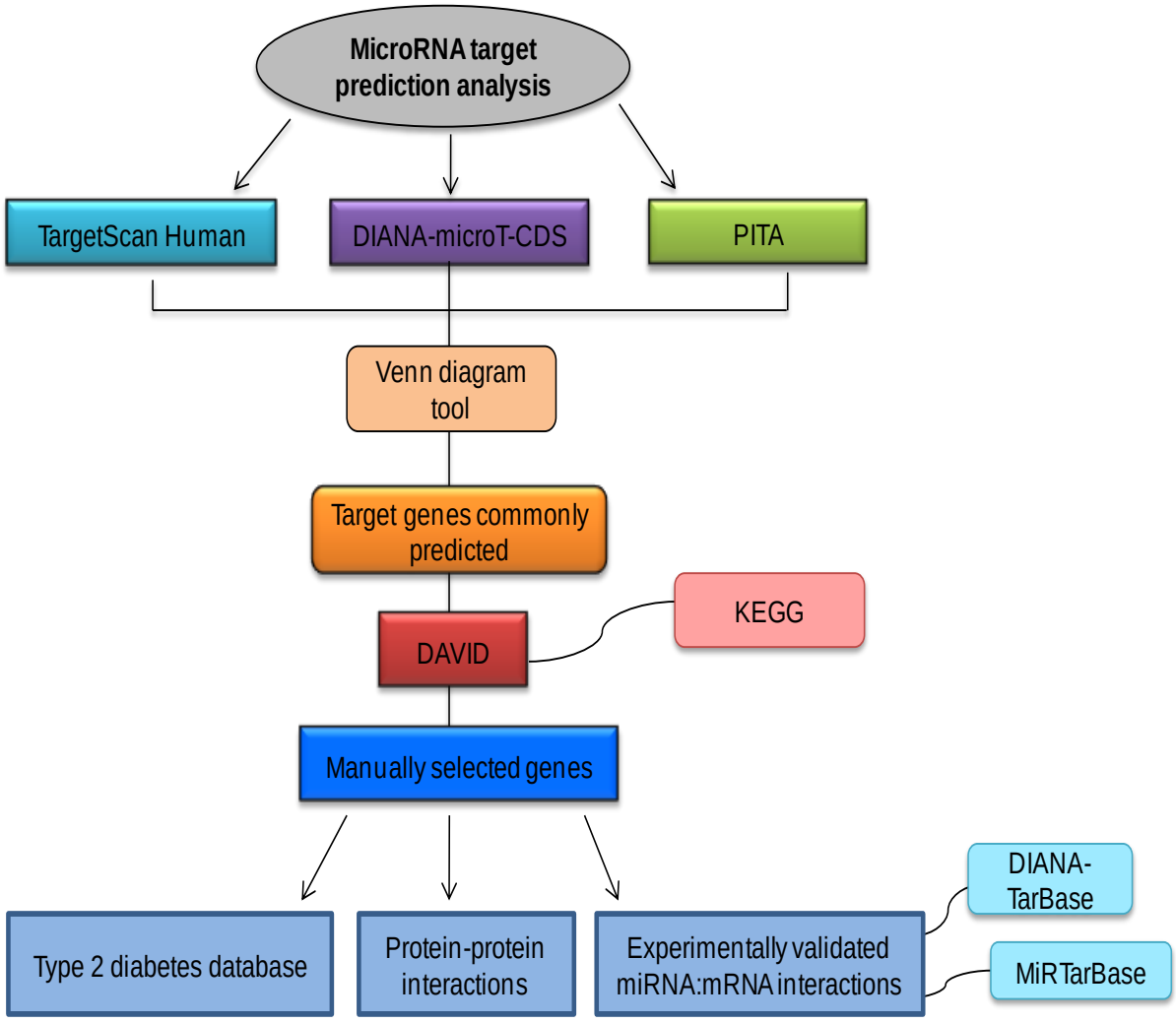


Figure 2.9 Schematic representation of microRNA target prediction workflow.

2.9.1 Functional analysis of predicted targets

The functional role of commonly predicted mRNAs targets for each miRNA was investigated using the DAVID gene functional annotation tool, which classifies mRNA targets according to biological pathways, functional biological modules, and inter-relationships between groups of genes (Huang *et al.*, 2008).

Commonly predicted target gene lists were uploaded, using the official gene symbol identifier (gene ID) for each gene. These gene IDs were then converted to gene names using the DAVID gene name batch viewer. The genes were functionally annotated to the *Homo sapien* genome. Thereafter, a more comprehensive functional analysis was performed by selecting or deselecting functional annotation categories of interest. By selecting the functional annotation summary and table, all data including biological processes, metabolic function and functional analysis of proteins were retrieved for all genes in the list. All genes mapped to the KEGG pathway database were selected. Genes potentially linked to T2D were selected based on their regulatory role in metabolic pathways previously associated with the development of the disease. Thereafter, the roles of these genes in T2D were confirmed using the T2D-db v2.0. The T2D-db is an integrated platform to study the molecular basis of T2D (http://t2ddb.ibab.ac.in/home.shtml) which provides information on candidate genes, gene expression data, gene/pathway interactions, protein-protein interactions and SNP markers and risk factors/complications from a combination of online content and published literature previously reported to be associated with T2D (Agrawal *et al.*, 2008) (Fig. 2.9).

MiRNA:gene interactions potentially linked to T2D were further investigated using experimentally validated miRNA:mRNA target prediction databases such as DIANA-TarBase v7.0 (http;//www.microrna.gr/tarbase) and miRTarbase v6.0 (http://miRTarBase.mbc.nctu.edu.tw/thout). These databases include miRNA:mRNA interactions that have been manually curated from information found in thousands of previously published studies. For DIANA-TarBase, miRNAs of interest were inserted into the search box, and filters including, *Homo sapien*, normal and high throughput method types, all experimental methods and all source types were selected and applied to the search. Further details regarding the miRNAs, target genes and interaction information was available by selecting the information links. For miRTarbase, standard default settings were used.

These databases provides a detailed description of involved genes and miRNAs, a list of publications supporting each interaction, the experimental methods used for validation along with their outcomes, and links related to KEGG pathways and other external databases (Fig. 2.9) (Elton and Yalowich, 2015; Hsu *et al.*, 2014).

2.9.2 Protein-protein interactions

To identify known and predicted protein-protein interactions of miRNA target gene/proteins, STRING v10.0 was used (http://string-db.org/). The interactions include direct (physical) and indirect (functional) associations that were derived from three sources including, genomic context, high throughput experiments, and previously published literature. Interactions are scored by integrating active predictive mechanisms (neighbourhood, gene fusion, co-occurrence, co-expression, experiments, databases and text mining of scientific texts), resulting in large comprehensive protein networks (Szklarczyk *et al.*, 2014).

On the home page, the name of a predicted miRNA target gene/protein was inserted into the search box, and the species *Homo sapien* was selected to restrict the search. After identification of proteins, the protein network was inspected for interactive evidence, and the confidence score cut-off was set at 0.95 or greater to filter out false positives. Thereafter, a limit was set to show no more than 10 interacting proteins.

2.9.3 Functional analysis of microRNAs

To determine the functional role of miRNAs in disease pathogenesis, the Human miRNA Disease Database (HMDD) v2.0 was used (http://www.cuilab.cn/hmdd). The HMDD is a collection of experimentally supported miRNA-disease associations, which provides data from genetics, epigenetics, circulating miRNAs and miRNA-target interactions. This database allows users to browse, search and download datasets, as well as submit novel data into the database. It is a valuable resource for investigating the roles of miRNAs in human disease, and currently contains over 10 000 experimentally supported miRNA-disease association entries, including 600 miRNA genes and 400 human diseases from more than 3000 articles (Li *et al.*, 2014).

For each miRNA searched, information about disease association, experimental evidence, dysfunction, genetics and epigenetics, presence in circulating biofluids, and the publication PubMed reference is given.

2.9.4 Statistical analysis

To assess the accuracy of the sequencing data, Chastity scores were calculated using the Solexa CHASTITY quality filter. Similar to the Phred quality scoring (Q-scores) system, individual bases are assigned a quality score that reflect the probability that the base call is correct. This score is quantified by a CHASTITY formula, which excludes clusters with a low signal to noise ratio. The default cut-off for filtering is no less than 0.6 in the first 25 bases. Statistical significance of differentially expressed miRNAs from the sequencing data was determined using the student t-test.

The data selected from the sequencing results were validated using qRT-PCR. Data was analysed using Microsoft Excel (2010) and relevant software packages, and statistical analysis was performed using GraphPad Prism v5.02. Thereafter, statistical analysis between diabetic, pre-diabetic and normo-glycaemic groups were performed using the One-way analysis of variance (ANOVA), followed by Bonferroni or Tukey *post hoc* and an unpaired t-test where appropriate. All values were presented as the mean ± standard error of the mean (SEM), and P values <0.05 were accepted as significant.

3. RESULTS

3.1 Clinical characteristics of participants

Participants were classified as diabetic (T2D), pre-diabetic or normo-glycaemic (controls) according to WHO criteria (World Health Organisation, 2006). Of the 258 study participants, 12 age-, gender-, ethnicity- and BMI-matched subjects representing diabetics (n=4), pre-diabetics (n=4) and normo-glycaemics (n=4) were selected for this part of the study.

The clinical parameters of the subjects are shown in Table 3.1. Subjects were matched for age and BMI, thus as expected, no significant differences in age (46.8 ± 6.6 vs. 47.3 ± 6.9 vs. 46.3 ± 5.7 years) and BMI (36.7 ± 3.1 vs. 38.6 ± 3.7 vs. 33.7 ± 1.8 kg/m^2) were observed between diabetic, pre-diabetic and normo-glycaemic individuals. Blood glucose concentrations varied significantly between groups. Fasting plasma glucose concentrations were 6.7 ± 0.5 vs. 5.4 ± 0.3 vs. 5.1 ± 0.1 mmol/L for diabetic, pre-diabetic and normo-glycaemic individuals, respectively (Fig. 3.1A). Two hour glucose concentrations obtained from the OGTT were 13.4 ± 0.7 vs. 8.9 ± 0.3 vs. 5.6 ± 0.3 mmol/L (Fig. 3.1B), while the percentage of HbA1c was 6.5 ± 0.2 vs. 5.6 ± 0.2 vs. 5.4 ± 0.07 (Fig. 3.1C) for diabetic, pre-diabetic and normo-glycaemic individuals, respectively. Statistical significance was determined using the p-value of ANOVA, followed by Tukey *post hoc* and an unpaired t-test.

Table 3.1 Clinical characteristics of participants.

	Diabetic	**Pre-diabetic**	**Normo-glycaemic**
N	4	4	4
Age (years)	46.8 ± 6.6	47.3 ± 6.9	46.3 ± 5.7
Gender	Female	Female	Female
Ethnicity	Mixed ancestry	Mixed ancestry	Mixed ancestry
BMI (kg/m^2)	36.7 ± 3.1	38.6 ± 3.7	33.7 ± 1.8
Blood pressure (mmHg):			
Systolic-	113.7 ± 10.8	131.3 ± 5.9	118.5 ± 4.0
Diastolic-	72.3 ± 8.7	83.3 ± 4.4	82.0 ± 3.4
FPG (mmol/L)	**6.7 ± 0.5***	5.4 ± 0.3	**5.1 ± 0.1***
OGTT (2 hr Glucose) (mmol/L)	**13.4 ± 0.7**$^{\gamma\dagger}$	**8.9 ± 0.3**$^{\gamma\Psi}$	**5.6 ± 0.3**$^{\dagger\Psi}$
HbA1c (%)	**6.5 ± 0.2**$^{*\gamma}$	**5.6 ± 0.2***	**5.4 ± 0.07**$^{\gamma}$
Fasting Insulin (ng/ml)	23.6 ± 6.7	34.8 ± 6.3	21.6 ± 3.1
2 hr Insulin (ng/ml)	97.8 ± 28.4	264.4 ± 87.7	92.5 ± 23.7
Fasting C-peptide (ng/ml)	2.7 ± 0.6	2.9 ± 0.4	2.7 ± 0.2
2 hr C-peptide (ng/ml)	7.9 ± 1.8	11.1 ± 1.7	9.1 ± 0.8
Glucagon (pg/ml)	91.0 ± 9.5	114.2 ± 16.9	79.6 ± 6.6
Cholesterol (mmol/L)	5.5 ± 0.7	5.3 ± 0.3	5.4 ± 0.8
HDL (mmol/L)	1.2 ± 0.1	1.2 ± 0.1	1.3 ± 0.03
LDL (mmol/L)	3.8 ± 0.6	3.6 ± 0.3	3.6 ± 0.7
Triglycerides (mmol/L)	1.2 ± 0.3	1.7 ± 0.3	1.1 ± 0.2

N = number of participants; BMI = body mass index; FPG = fasting plasma glucose; OGTT = oral glucose tolerance test; HbA1c = glycated haemoglobin A1c; 2 hr refers to a test performed 2 hrs after an oral glucose challenge. HDL = high-density lipoprotein; LDL = low-density lipoprotein.

Blood pressure was measured as systolic/diastolic.

Boldface values indicate statistical significance. $^{*}p\leq0.05$; $^{\gamma}p\leq0.01$; $^{\dagger\Psi}p\leq0.001$.

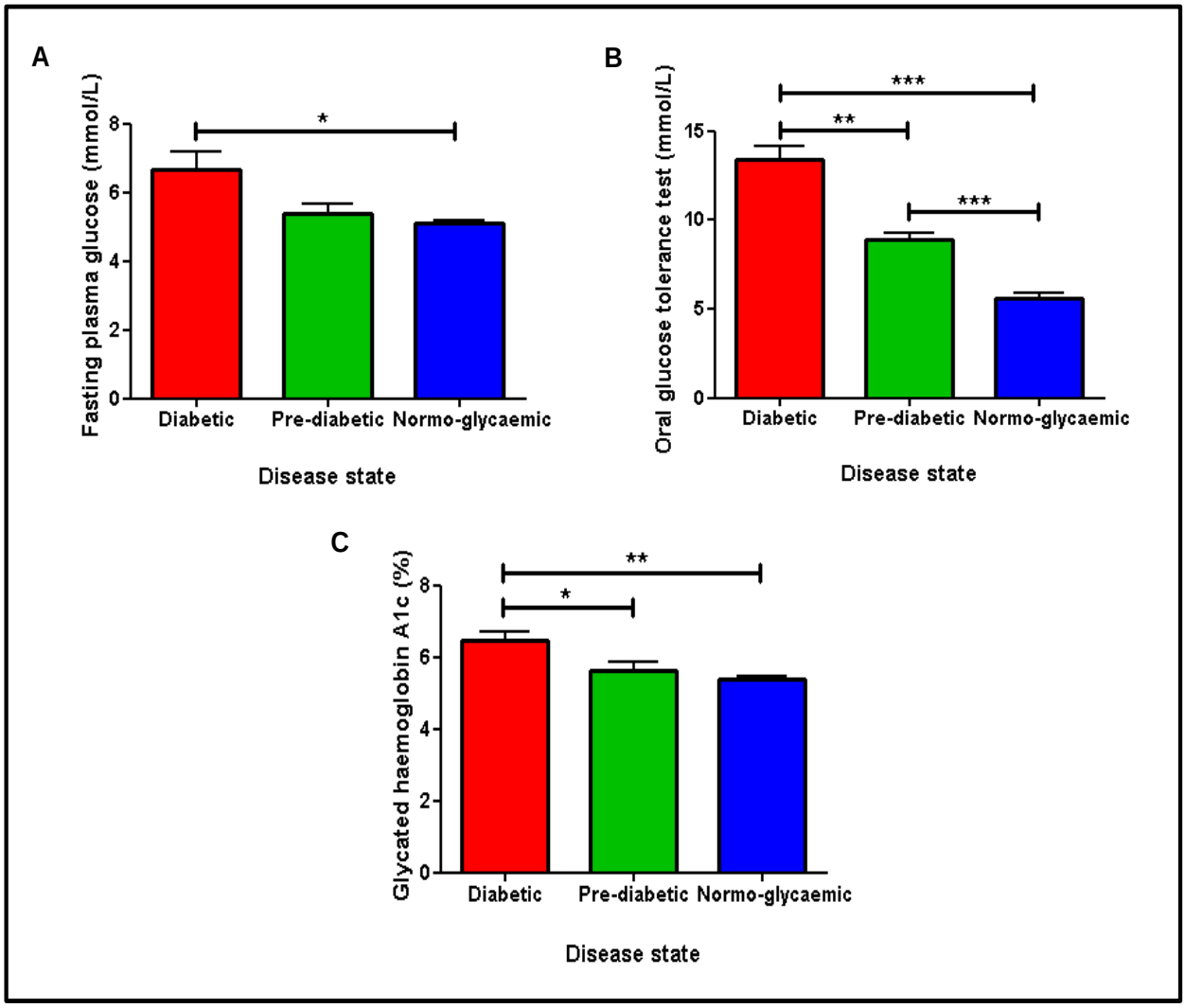


Figure 3.1 Blood glucose concentrations in diabetic, pre-diabetic and normo-glycaemic individuals. A) Fasting plasma glucose, B) 2 hr OGTT glucose, and C) HbA1c levels in diabetics (n=4), pre-diabetics (n=4) and normo-glycaemics (n=4). *p<0.05, **p<0.01, ***p<0.001.

Levels of insulin, the primary hormone responsible for maintaining glucose homeostasis, was measured in the fasting state, and 30 mins and two hrs after a glucose challenge (OGTT). Fasting insulin concentrations were 23.6 ± 6.7 vs. 34.8 ± 6.3 vs. 21.6 ± 3.1 ng/ml for diabetic, pre-diabetic and normo-glycaemic individuals, respectively (Fig. 3.2A). After the glucose challenge, insulin levels increased in all three groups until 30 mins, whereafter, levels decreased in normo-glycaemics, but continued to rise in diabetics and pre-diabetics, signifying abnormal glucose metabolism in these groups (Fig. 3.2B). The increase observed in pre-diabetics was higher than that observed in diabetics (2.85-fold).

No significant difference was observed when measuring additional parameters, including glucagon (91.0 ± 9.5 vs. 114.2 ± 16.9 vs. 79.6 ± 6.6 pg/ml), C-peptide (2.7 ± 0.6 vs. 2.9 ± 0.4 vs. 2.7 ± 0.2 ng/ml), cholesterol (5.5 ± 0.7 vs. 5.3 ± 0.3 vs. 5.4 ± 0.8 mmol/L), HDL (1.2 ± 0.1 vs. 1.2 ± 0.1 vs. 1.3 ± 0.03 mmol/L), LDL (3.8 ± 0.6 vs. 3.6 ± 0.3 vs. 3.6 ± 0.7 mmol/L) and triglycerides (1.2 ± 0.3 vs. 1.7 ± 0.3 vs. 1.1 ± 0.2 mmol/L), between diabetic, pre-diabetic and normo-glycaemic individuals, respectively (Table 3.1).

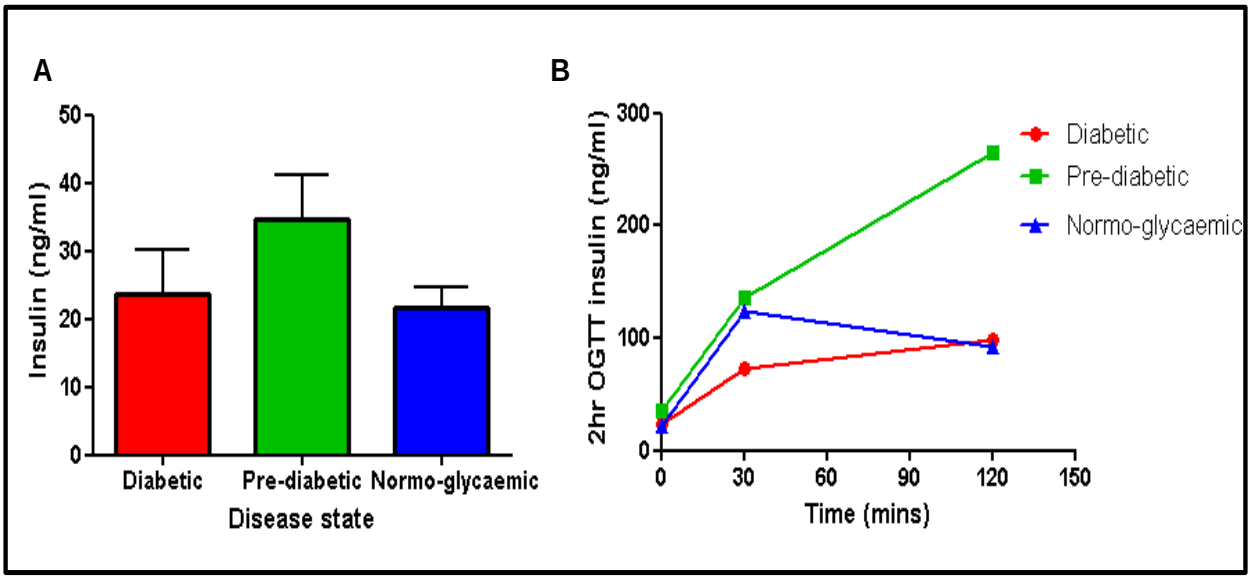


Figure 3.2 Fasting and 2hr OGTT insulin concentrations in diabetic, pre-diabetic and normo-glycaemic individuals. A) Fasting insulin and B) OGTT insulin concentrations in diabetic (n=4), pre-diabetic (n=4) and normo-glycaemic individuals (n=4)

.

3.2 MicroRNA expression analysis in peripheral blood mononuclear cells

3.2.1 RNA concentration and yield

Nanodrop spectrophotometry was used to determine the concentration, purity and yield of total RNA from PBMCs. The concentration and yield of RNA varied between 31.61 - 124.33 ng/µl and 1264.40 - 4973.00 ng, respectively (Table 3.2). The A_{260}/A_{280} and A_{260}/A_{230} ratios are indicative of protein and solvent/salt/other contamination, which varied between 2.03 - 2.10 and 0.68 - 2.06, respectively, indicating that most of the samples (~83.3 %) contained good quality RNA.

Table 3.2 Total RNA concentration, purity and yield of samples (n=12).

Sample ID	Concentration (ng/µl)*	Yield (ng)†	A_{260}/A_{280} (nm)ϒ	A_{260}/A_{230} (nm)Ψ
1	123.47	4938.60	2.08	2.05
2	56.45	2258.00	2.08	1.21
3	119.45	4777.80	2.08	1.94
4	76.15	3045.80	2.07	2.06
5	93.16	3726.20	2.09	1.78
6	124.33	4973.00	2.03	2.02
7	49.87	1994.80	2.10	1.59
8	75.85	3034.00	2.08	1.54
9	100.67	4026.80	2.08	1.87
10	31.61	1264.40	2.07	0.68
11	75.69	3027.60	2.08	1.57
12	108.05	4321.80	2.08	1.06

*RNA concentration was measured at A_{260} nm

†RNA yield was calculated by multiplying RNA concentration (ng/µl) by the sample volume (40 µl in this case)

ϒPresence of protein contaminants were measured at A_{260}/A_{280} nm

ΨPresence of solvent/salt/other contaminants were measured at A_{260}/A_{230} nm

3.2.1.1 RNA quality control

The success of downstream applications is dependent on the quality and integrity of RNA. The quality of the RNA isolated from PBMCs was assessed using the bioanalyser. A gel image representing the 12 samples using the RNA Nano Assay and 11 samples using the RNA small assay, and an electropherogram representing total and small RNA (miRNA) of one of the samples generated by the bioanalyser is indicated in Figure 3.3. The low visibility and signal in the miRNA region (Fig. 3.3C-D) may be due to low miRNA content (10%) after purification, and could be related to low RNA concentration.

The RNA integrity number (RIN), miRNA as a percentage (%) of total RNA, and the concentration (pg/µl) of miRNAs for all 12 samples are presented in Table 3.3. The RIN values ranged between 7.30 - 8.50 (7.91 ± 0.12), demonstrating intact and good quality RNA for all samples. The percentage and concentration of miRNAs within the samples ranged between 7 - 16 % (9.6 ± 0.81) and 230.60 - 897.90 pg/µl (374.9 ± 37.55), respectively.

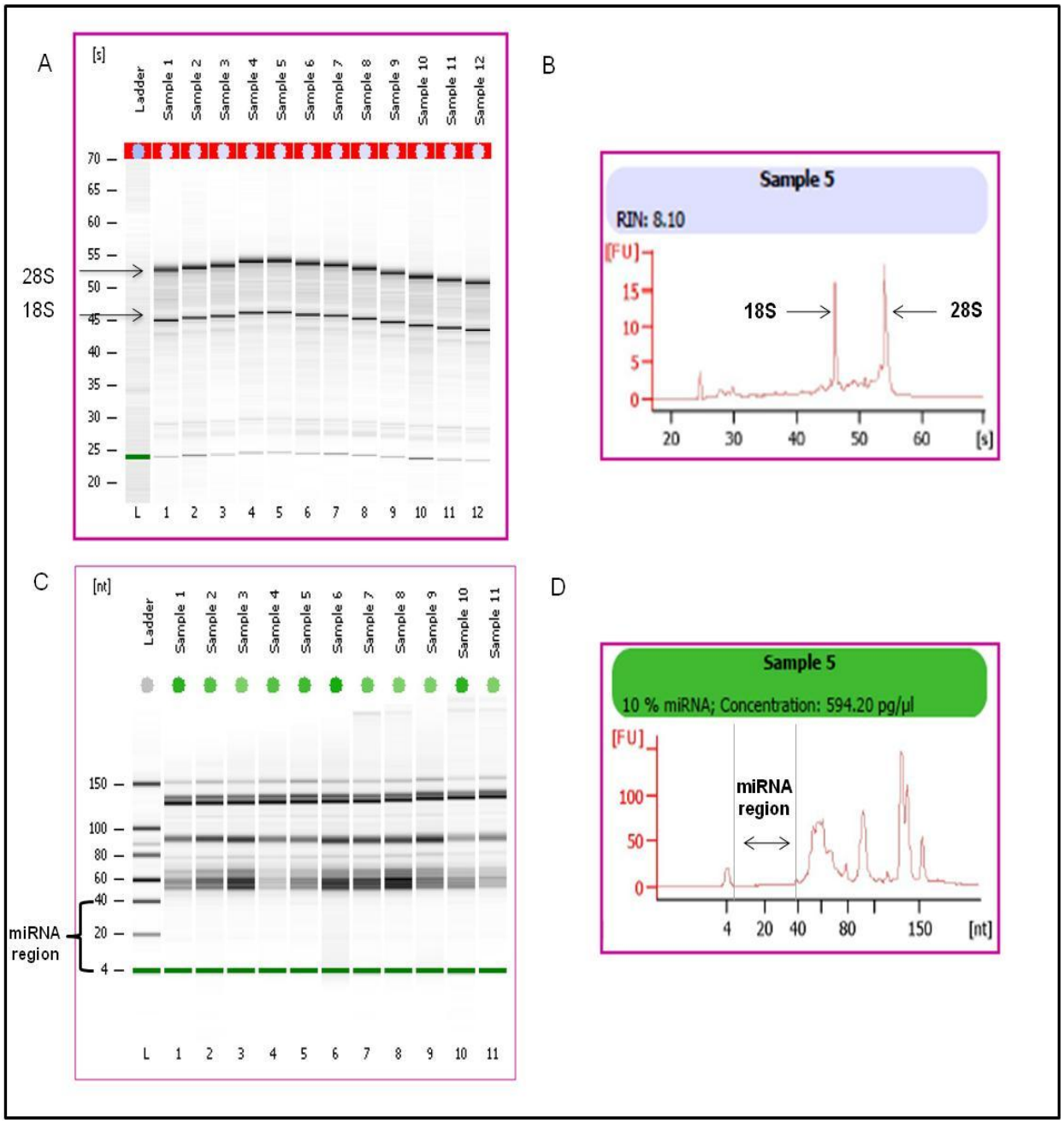


Figure 3.3 A gel image of total RNA and an electropherogram indicating RNA quality. A) Gel image representing the 12 RNA samples in lanes 1-12, showing bands representative of intact 28S and 18S rRNA molecules. Lane L represents the RNA ladder with the marker for comparative analysis indicated in green. B) Electropherogram of total RNA with two main peaks indicating 28S and 18S rRNA. C) Gel image representing 11 RNA samples in lanes 1-11, showing the region representative of the miRNA length (nt), and an D) electropherogram of small RNA, indicating the region of miRNAs between 6 and 40 nt.

Table 3.3 Integrity of total RNA, and the percentage and concentration of miRNAs within each sample (n=12).

Sample	RIN no.	% miRNA	miRNA concentration (pg/µl)
1	8.2	11.0	512.9
2	7.8	9.0	341.6
3	7.7	7.0	408.7
4	7.5	9.0	371.0
5	8.1	10.0	594.2
6	7.6	16.0	431.1
7	7.4	8.0	260.8
8	8.5	7.0	366.0
9	8.4	7.0	231.9
10	7.3	11.0	230.6
11	8.5	7.0	416.8
12	8.0	13.0	897.9

RIN = RNA integrity number

3.2.2 MicroRNA sequencing

To identify miRNAs that are dysregulated during disease progression, miRNA-enriched total RNA from diabetic, pre-diabetic and normo-glycaemic individuals were subjected to next generation sequencing on the Illumina HiSeq 2500. All the twelve RNA samples that were sent to Arraystar passed their quality control test, which included total RNA quantification and quality assurance by NanoDrop ND-1000 spectrophotometry. Thereafter, a sequencing library of the RNA samples were prepared as described in section 2.7, and the quality thereof assessed with the Agilent 2100 Bioanalyser. All the samples passed the quality control, and were thus subjected to miRNA sequencing. Details of the quality control test and the quality assessment of the sequencing library can be found in Appendix 3.

Reads generated by the Illumina HiSeq 2500 instrument were cleaned, adapters trimmed and aligned to a known reference human genome in the miRBase v21 database (http://www.mirbase.org/), to determine the number of reads generated, as described in

section 2.7. The total number of clean reads, adapter trimmed reads, reads aligned to known miRNA sequences, and the percentage of reads aligned to the human genome, are presented in Table 3.4.

The size distribution of adapter trimmed reads aligned to human reference miRNAs (Table 3.4, column 3) is illustrated in Figure 3.4. The majority of sequencing read peaks occurred around 22 nt as expected.

Table 3.4 The read counts at the different data processing stages.

Sample Name	Clean Reads	Adapter-trimmed Reads (length >= 15nt)	Reads aligned to known human pre-miRNA in miRBase 21	Percentage of aligned reads
1	6,124,069	6,061,197	5,676,739	93%
2	3,458,762	3,442,273	3,248,139	94%
3	5,404,671	5,356,925	5,014,891	94%
4	4,082,938	4,057,478	3,817,335	94%
5	4,872,306	4,853,250	4,577,000	94%
6	3,352,766	3,329,626	3,133,999	94%
7	4,050,955	4,015,379	3,737,566	93%
8	5,803,200	5,754,741	5,372,612	93%
9	7,611,759	7,525,970	6,950,059	92%
10	6,865,665	6,823,551	6,445,241	94%
11	6,433,218	6,353,760	5,880,953	93%
12	5,301,440	5,260,559	4,932,301	93%

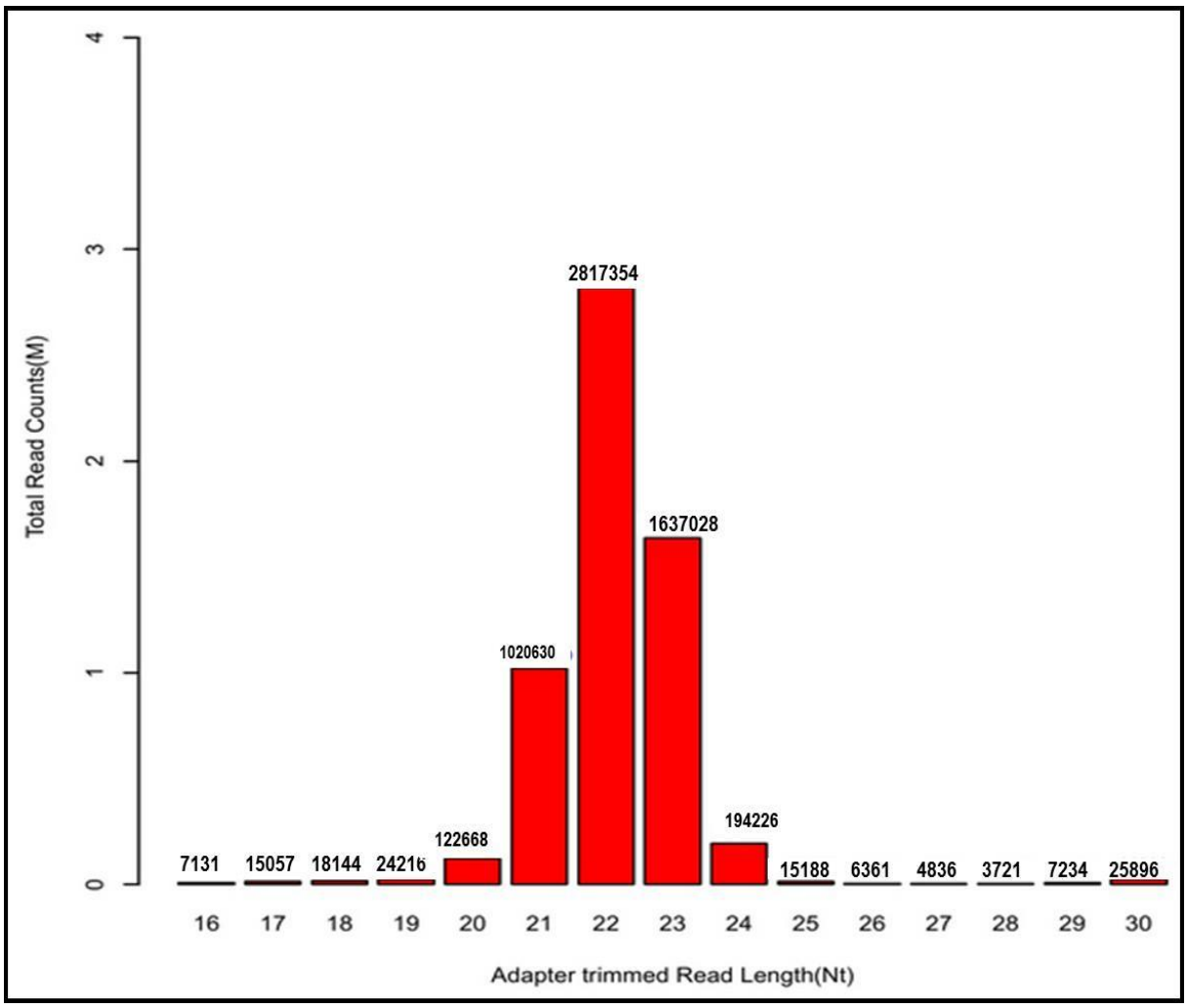


Figure 3.4 Size distribution of sequencing reads. The size distribution of sequence reads of a representative miRNA sample is indicated by total read counts (y-axis) and adapter trimmed read length (x-axis).

3.2.3 Differential expression of microRNAs during Type 2 diabetes progression

The expression of miRNAs in each sample was determined by comparing the expression of each unique sequence to the total number of reads per sample, using equation 2.1 (section 2.7). A total of 294 miRNAs were found to be expressed among all samples. Using Novoalign, 277, 267 and 267 miRNAs were differentially expressed between diabetics vs. normo-glycaemics, pre-diabetics vs. normo-glycaemics, and pre-diabetics vs. diabetics, respectively (Table 3.5). A number of novel miRNAs that have not previously been annotated in the miRNA database were identified; 35, 33 and 33 novel miRNAs were differentially expressed between diabetics vs. normo-glycaemics, pre-diabetics vs. normo-glycaemics and pre-diabetics vs. diabetics, respectively. Using an arbitrary cut-off of ≥1.1 expression fold difference and statistical significance set at $p \leq 0.05$, five miRNAs (miR-27b, miR-98, miR-143, miR-21 and miR-379) were selected for validation by qRT-PCR (Table 3.6).

MiR-143, miR-21 and miR-379 were upregulated in diabetics compared to pre-diabetics (1.75-fold; $p \leq 0.05$, 1.22-fold; $p < 0.01$, 1.12-fold; $p < 0.05$, respectively), while miR-27b was upregulated in pre-diabetic compared to normo-glycaemic individuals (1.15-fold; $p < 0.01$). In contrast, miR-98 was downregulated in diabetics compared to pre-diabetics (1.25-fold; $p < 0.05$) and in pre-diabetics compared to normo-glycaemic individuals (1.23-fold; $p < 0.05$) (Table 3.6).

Table 3.5 Total number of differentially expressed and novel miRNAs among all samples.

	Diabetic vs. Normo-glycaemic	**Pre-diabetic vs. Normo-glycaemic**	**Pre-diabetic vs. Diabetic**
No. of differentially expressed miRNAs	277	267	267
Novel miRNAs	35	33	33

Table 3.6 MiRNAs selected for qRT-PCR.

Gene symbol	Mature miRNA ID	Disease group	Fold change	P-value
MI0000100	hsa-miR-98	D vs. PD	-1.25	0.04
MI0000459	hsa-miR-143	D vs. PD	1.75	0.05
MI0000077	hsa-miR-21	D vs. PD	1.22	0.008
MI0000787	hsa-miR-379	D vs. PD	1.12	0.02
MI0000440	hsa-miR-27b	PD vs. C	1.15	0.003
MI0000100	hsa-miR-98	PD vs. C	-1.23	0.02

Positive values indicate miRNAs that were upregulated between groups, while negative values indicate miRNAs that were downregulated between groups. D= Diabetics; PD= Pre-diabetics; C= Normo-glycaemics.

The hierarchally clustered heat map that was generated to illustrate the differential expression of miRNAs between diabetic, pre-diabetic and normo-glycaemic samples is shown in Figure 3.5.

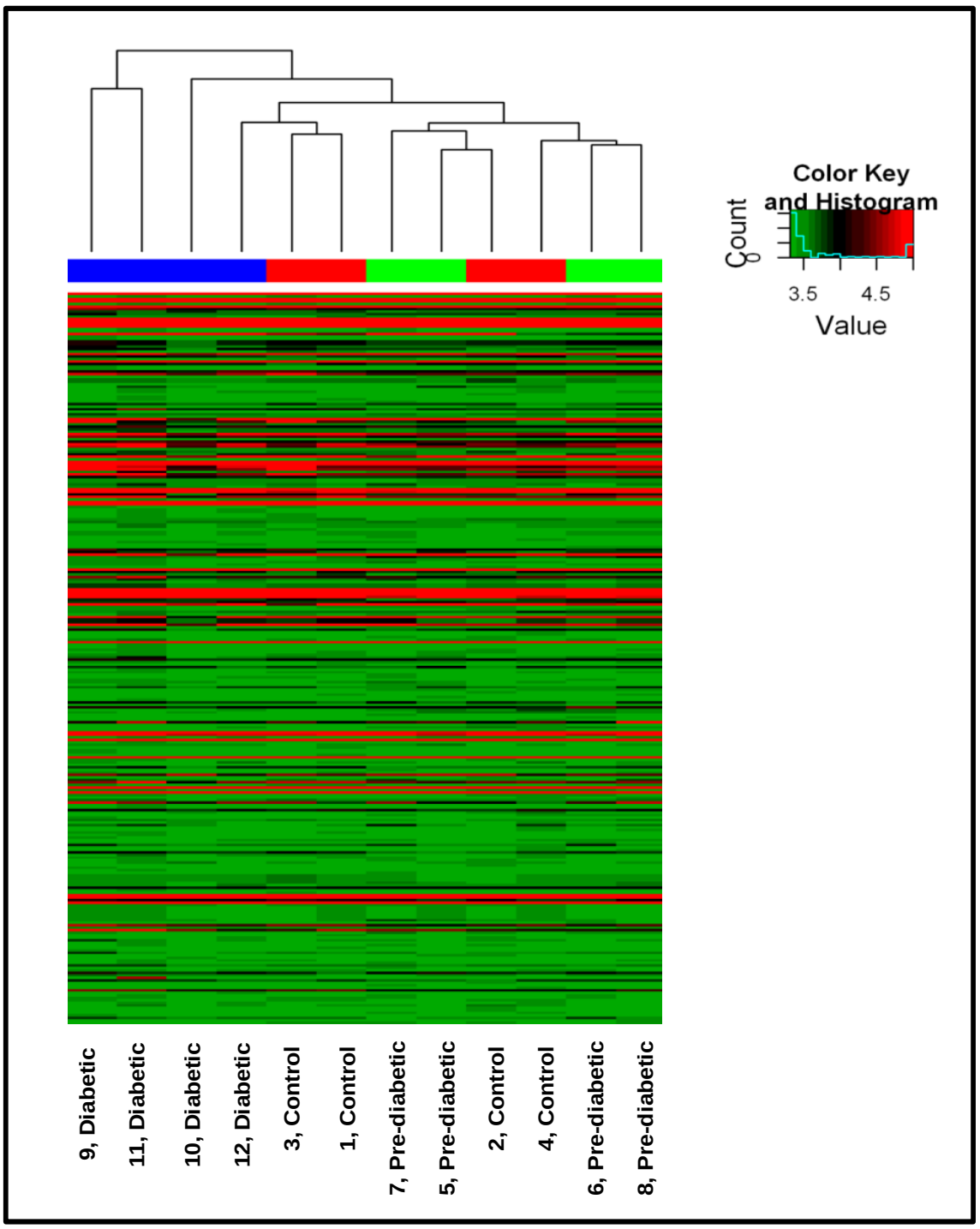


Figure 3.5 A hierarchal clustered heat map of differentially expressed miRNAs between diabetic, pre-diabetic and normo-glycaemic individuals. Heat map colours represent relative miRNA expression as indicated in the colour key. High and low miRNA expression levels are indicated by red, and green, respectively.

3.2.4 Quantitative real time PCR validation of differentially expressed microRNAs

Quantitative real time PCR was used to validate the differential expression of miR-27b, miR-98, miR-143, miR-21 and miR-379, identified by miRNA sequencing. Expression was calculated using the ΔΔCt method where RNU6B, miR-454 and miR-425 were used as endogenous reference genes, and the average expression in normo-glycaemic individuals (n=4) were used as a calibrator. The relative expression of each miRNA was calculated using the equation 2.2 in section 2.8.2.1.

Endogenous reference genes were stably expressed in all samples when combined; no significant difference was observed between diabetic, pre-diabetic and normo-glycaemic groups (Fig. 3.6A), thus making these genes ideal endogenous controls for miRNA expression analysis. The expression of miR-143 was upregulated in pre-diabetics compared to normo-glycaemic individuals (1.4-fold, $p \leq 0.01$) (Fig. 3.6B), while no significant changes were observed for miR-21, miR-27b, miR-379 and miR-98 (Fig. 3.6C-F). However, for miR-27b, a trend towards significance (1.55-fold; $p=0.07$) was observed between pre-diabetic and normo-glycaemic individuals (Fig. 3.6C), confirming sequencing results which showed significant upregulation of miR-27b in pre-diabetics compared to normo-glycaemic individuals (Table 3.6).

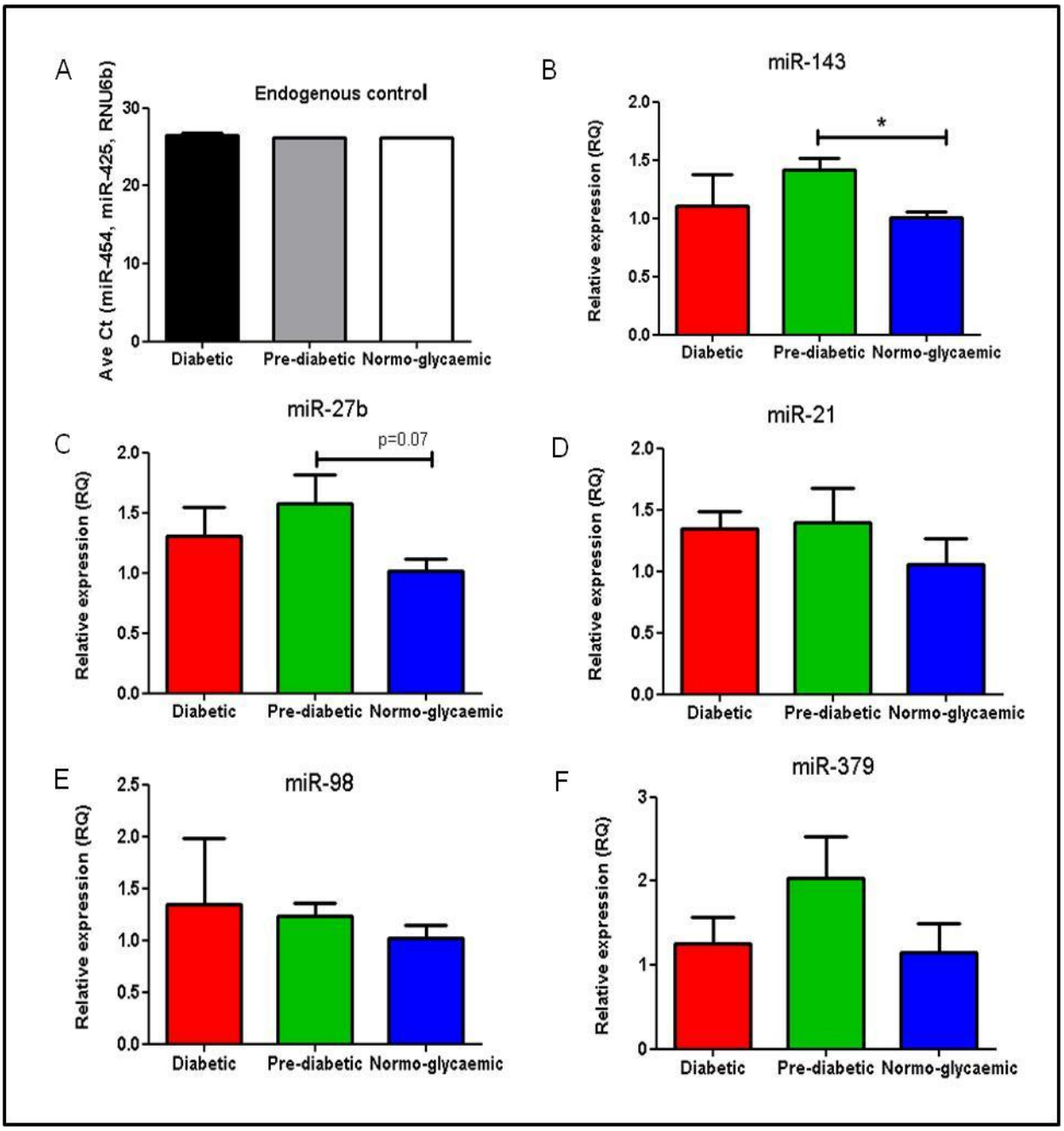


Figure 3.6 MiRNA expression analysis using real-time quantitative PCR. A) Average Ct of endogenous control, B) relative expression changes of miR-143, C) miR-27b, D) miR-21, E) miR-98, and F) miR-379 between diabetic (n=4), pre-diabetic (n=4) and normo-glycaemic individuals (n=4), *****p<0.05.

3.3 MicroRNA expression analysis in serum

The differentially expressed miRNAs in PBMCs were investigated in serum samples of the same individuals.

3.3.1 Total RNA concentration and yield

Nanodrop spectrophotometry was used to determine the concentration, purity and yield of total RNA-enriched miRNAs from serum samples. The concentration and yield of total RNA varied between 11.15 - 19.05 ng/µl and 227.10 - 380.90 ng, respectively (Table 3.7), indicating lower total RNA amounts in serum compared to PBMCs, as expected. The A_{260}/A_{280} and A_{260}/A_{230} ratios varied between 1.26 - 1.49 nm and 0.29 - 0.44 nm, respectively. These low ratios may be indicative of possible protein-bound miRNA detection or cellular contamination from miRNA origin.

Table 3.7 Total RNA concentration, purity and yield of serum samples (n=12).

Sample ID	Concentration (ng/µl)*	Yield (ng)†	A_{260}/A_{280} (nm)$^{\gamma}$	A_{260}/A_{230} (nm)$^{\Psi}$
1	17.03	340.60	1.49	0.29
2	17.16	343.20	1.34	0.27
3	14.51	290.10	1.30	0.28
4	12.11	242.10	1.33	0.28
5	13.70	273.90	1.27	0.31
6	19.05	380.90	1.26	0.41
7	12.78	255.50	1.29	0.44
8	17.96	359.20	1.44	0.21
9	14.80	296.00	1.39	0.21
10	14.03	280.50	1.28	0.42
11	11.58	231.60	1.31	0.17
12	11.36	227.10	1.42	0.07

*****RNA concentration was measured at A_{260} nm

†RNA yield was calculated using total RNA concentration (ng/µl) in 14 µl of total sample volume

ϒPresence of protein contaminants were measured at A_{260}/A_{280} nm

ΨPresence of solvent/salt/other contaminants were measured at A_{260}/A_{230} nm

3.3.2 MicroRNA quality control

The quality of the RNA-enriched miRNAs isolated from serum samples was assessed using the small Agilent bioanalyser for total miRNA detection. The percentage and concentration of miRNAs within the 12 samples ranged between 50 - 89 % (67.4 ± 9.4) and 134.40 – 1017.70 pg/µl (336.0 ± 249.7), respectively, and is presented in Table 3.8.

Table 3.8 Percentage and concentration of miRNAs in each serum sample (n=12).

Sample	% of miRNA	miRNA concentration (pg/µl)
1	74	330.10
2	71	318.40
3	69	134.40
4	89	260.70
5	67	161.00
6	69	331.50
7	65	191.80
8	50	117.50
9	69	1017.70
10	57	179.90
11	66	509.50
12	63	479.90

3.3.3 Differential expression of microRNAs in serum

The expression of miR-27b, miR-98, miR-143, miR-21 and miR-379 in serum of diabetic (n=4), pre-diabetic (n=4) and normo-glycaemic individuals (n=4) was calculated using the ΔΔCt methods and using miR-425 as the endogeneous control (Fig. 3.7).

Of the three endogenous controls investigated, Normfinder showed that miR-425 was the most stably expressed among all samples, with no significant difference between diabetic, pre-diabetic and normo-glycaemic groups (Fig. 3.7A). In addition, the exogenous synthetic spike in control (miR-39); used to control for technical variation, was stably expressed in all samples (Fig. 3.7B), suggesting efficiency of RNA extraction, cDNA synthesis and PCR amplification.

MiR-27b was upregulated (2.0-fold; $p \leq 0.05$) in pre-diabetic compared to normo-glycaemic individuals (Fig. 3.7C). Although miR-21, miR-379, miR-98 and miR-143 were also detected in serum, none of them were differentially expressed between the groups (Fig. 3.7D-G).

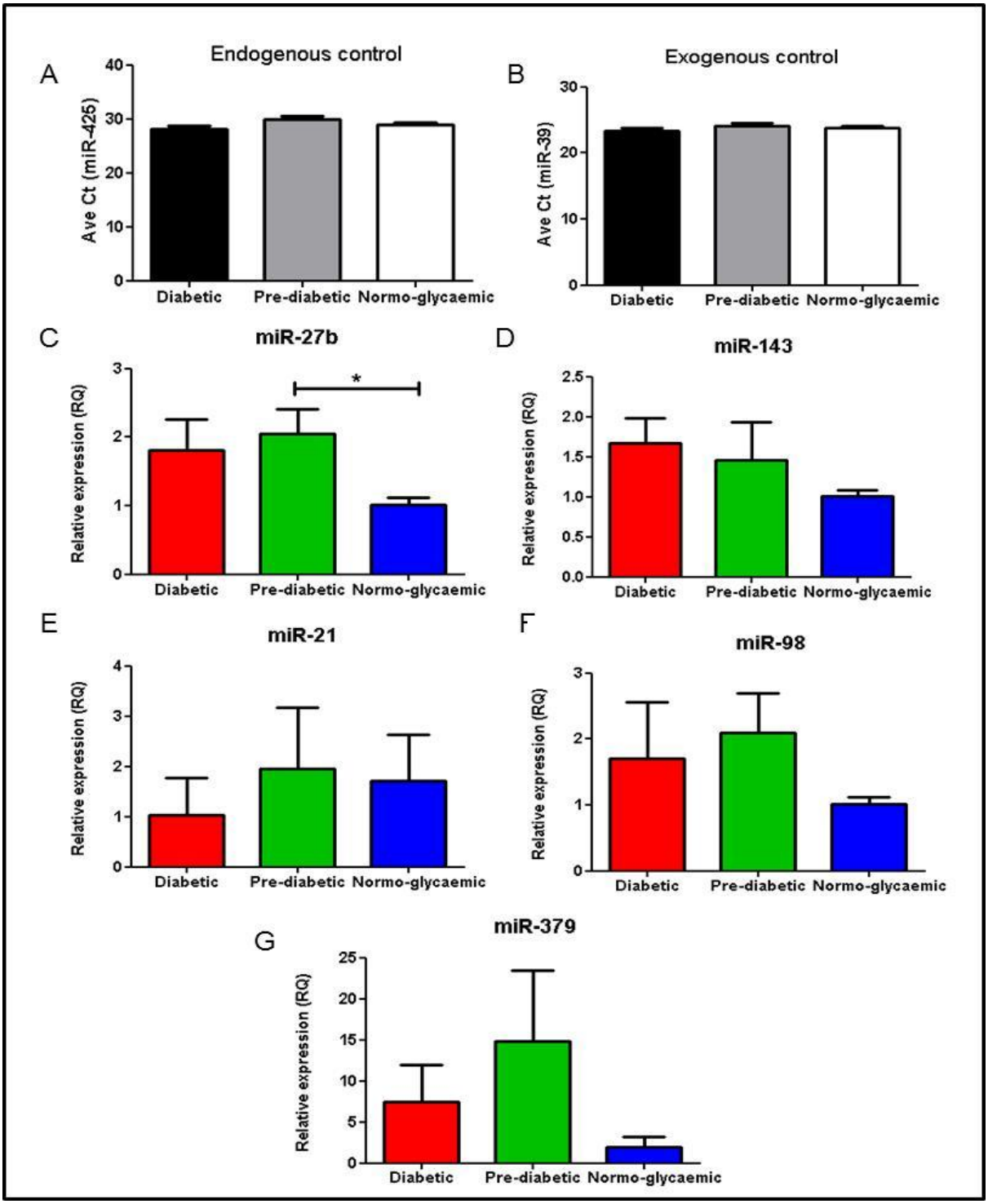


Figure 3.7 Expression of miRNAs in the serum of diabetic, pre-diabetic and normo-glycaemic individuals, using real-time quantitative PCR. A) Average Ct of endogenous control, B) average Ct of exogenous control, relative expression changes of C) miR-27b, D) miR-143, E) miR-21, F) miR-98 and G) miR-379 between diabetic (n=4), pre-diabetic (n=4) and normo-glycaemic individuals (n=4), *p<0.05.

3.4 Novel microRNAs

A total of 151 potential novel miRNA precursors (both 3p and 5p) were identified by next generation miRNA sequencing. MiRDeep algorithms showed that only 130 novel miRNA precursor structures were predicted to be significant across the genome. Of these, 33, 35 and 33 novel miRNAs were shown to be differentially expressed in pre-diabetic compared to normo-glycaemic individuals, diabetic compared to normo-glycaemic individuals, and pre-diabetics compared to diabetics, respectively. Comparative analysis of the differentially expressed miRNAs in each of the groups identified 23 novel miRNAs that were commonly differentially expressed in all 3 groups (pre-diabetic vs. normo-glycaemic, diabetic vs. normo-glycaemic and pre-diabetic vs. diabetic) (Fig. 3.8). Six miRNAs were differentially expressed in both diabetic and pre-diabetics compared to normo-glycaemic individuals, six miRNAs were differentially expressed in diabetics compared to both pre-diabetic and normo-glycaemic individuals, and four miRNAs were differentially expressed in pre-diabetics compared to diabetics and normo-glycaemic individuals (Fig. 3.8).

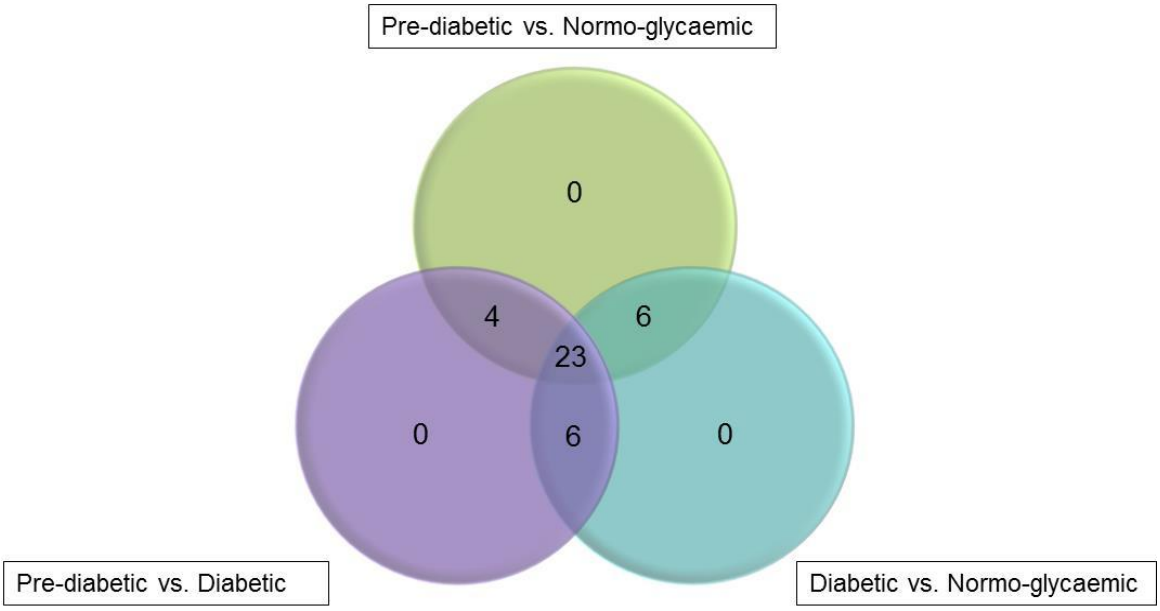


Figure 3.8 Venn diagram showing novel miRNAs commonly expressed between diabetic, pre-diabetic and normo-glycaemic groups.

3.4.1 Quantitative real time PCR validation of novel microRNAs

Five novel miRNAs (MYNO22, MYNO59, MYNO66, MYNO8, MYNO95) were selected for validation using qRT-PCR. Validation was performed in the same RNA samples used for miRNA sequencing; diabetic (n=3), pre-diabetic (n=4) and normo-glycaemic (n=4) samples. One diabetic sample did not have sufficient RNA for analysis. Quantitative real-time PCR confirmed the expression of MYNO22, MYNO59, MYNO66, MYNO8, MYNO95 in all samples (Fig. 3.9).

The relative expression levels of these miRNAs was calculated using the average Ct of RNU6B, miR-454 and miR-425 as endogenous reference genes for normalization (Fig. 3.9A). MYNO59, MYNO95 and MYNO66 was downregulated in diabetics compared to normo-glycaemic individuals (-0.4-fold; $p \leq 0.05$, -0.5-fold; $p \leq 0.05$, -0.5-fold; $p \leq 0.01$) (Fig. 3.9B-D). In addition, MYNO66 and MYNO8 was downregulated in pre-diabetics compared to normo-glycaemic individuals (-0.5-fold; $p \leq 0.05$), and in diabetics compared to pre-diabetics (-0.8-fold; $p \leq 0.05$), respectively (Fig. 3.9D-E). No significant differential expression was observed for MYNO22 (Fig. 3.9F).

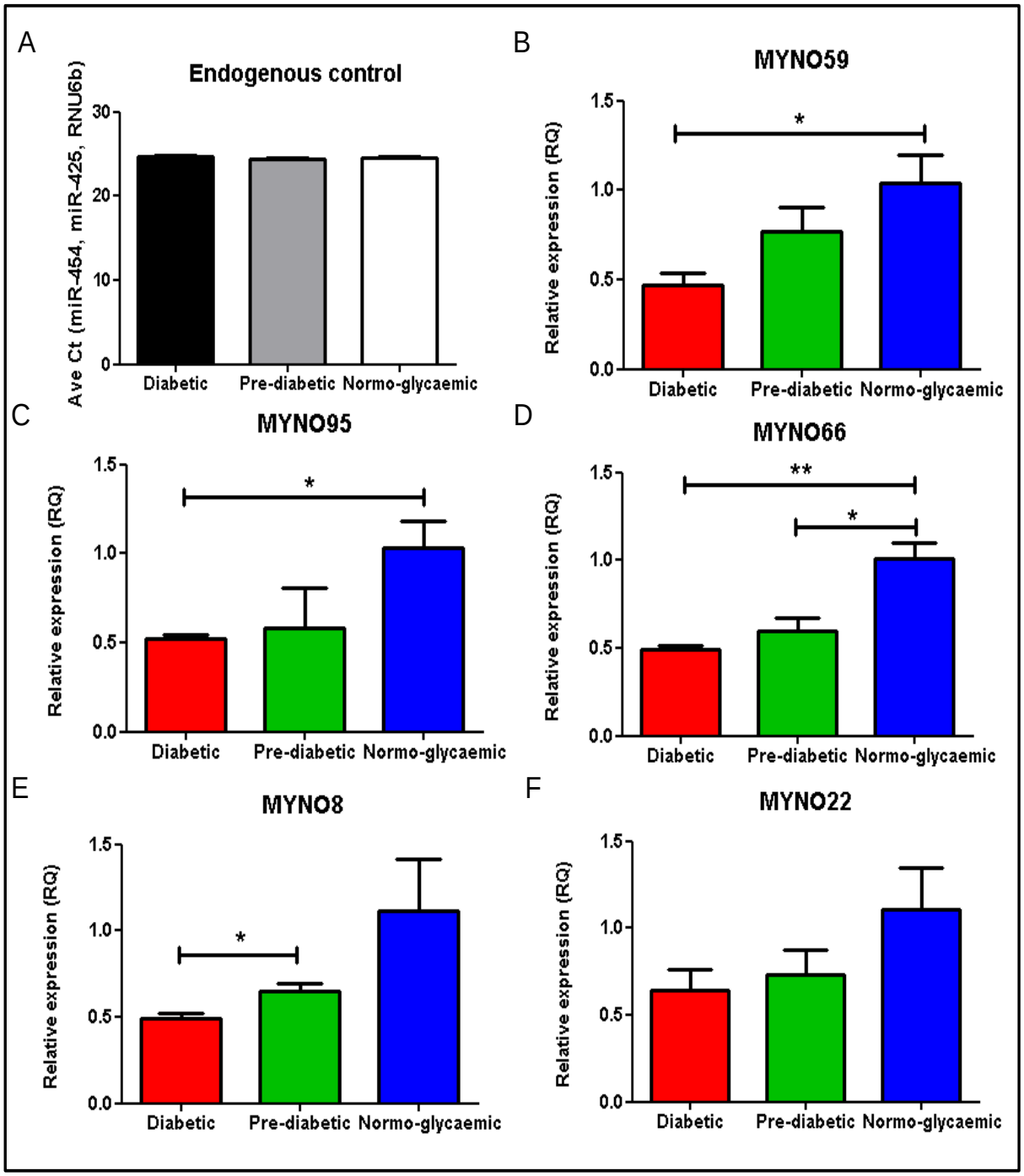


Figure 3.9 Validation of novel miRNAs using quantitative real-time PCR. A) Average Ct of endogenous control, relative expression changes of B) MYNO59, C) MYNO95, D) MYNO66, E) MYNO8, and F) MYNO22 between diabetic (n=3), pre-diabetic (n=4) and normo-glycaemic individuals (n=4), *p<0.05, **p<0.01.

3.5 Bioinformatics

3.5.1 Messenger RNA target prediction analysis

Messenger RNA (mRNA) target prediction for each of the five selected differentially expressed miRNAs (section 3.2.3) was conducted using three target prediction programs (TargetScan, DIANA and PITA) that differ in the algorithms they use for prediction (Table 3.9). The number of mRNA targets predicted for miRNAs varied according to the target prediction program used. For example, for miR-27b, TargetScan, DIANA and PITA predicted 71, 2107 and 1553 mRNA targets, respectively. Of these 39% of targets predicted by TargetScan were commonly predicted by all three programs, while 1.3% and 1.8% of targets predicted by DIANA and PITA, respectively, were commonly predicted by all three programs. TargetScan had the highest percentage of true predicted targets compared to DIANA and PITA, based on the assumption that mRNA targets predicted by all three programs are 'true' targets.

Table 3.9 Number of commonly predicted targets between TargetScan, DIANA and PITA.

	Number of targets predicted			
Mature miRNA	**TargetScan†**	**DIANAΨ**	**PITAγ**	**Common***
miR-27b	71 (39%)	2107 (1.3%)	1553 (1.8%)	28
miR-379	38 (18%)	377 (1.8%)	665 (1.5%)	7
miR-21	164 (7.9%)	554 (2.3%)	975 (1.3%)	13
miR-98	46 (26%)	991 (1.2%)	1291 (0.9%)	12
miR-143	279 (26.1%)	1062 (6.8%)	289 (25%)	73

†TargetScan predicts targets based on seed match complementarity and sequence conservation.

ΨDIANA predicts targets based on free energy binding and complementarity.

γPITA predicts targets based on target site accessibility and free energy.

*Commonly predicted targets determined using the venny tool (Oliveros 2007)

Percentages (%) were determined by dividing the commonly predicted targets by the total number of targets predicted by the program, and multiplying by 100.

3.5.2 Functional analysis of predicted targets

Pathway analysis of the commonly predicted mRNA target genes was conducted using the DAVID functional annotation tool. Commonly predicted mRNA target gene lists for the five miRNAs (Table 3.9) were copied into DAVID for pathway analysis. Results of the individual miRNAs and their predicted mRNA targets are indicated in Table 3.10. A total of 40 predicted mRNA target genes were mapped to KEGG pathways (Table 3.10, column 3). Of these, 13 predicted target genes were selected based on their potential role in T2D disease progression, as described in section 2.9.1 (Table 3.10, column 4). Thereafter, the role of these targets in T2D was confirmed using the T2D-db. Three of the 13 predicted target genes (peroxisome proliferator-activated receptor gamma *(Ppagr),* the insulin receptor *(Insr)* and v-Ki-ras2 kirsten rat sarcoma viral oncogene homolog *(Kras)* were identified in this database.

For miR-27b, 12 of the 28 mRNA targets commonly predicted by TargetScan, DIANA and PITA mapped to KEGG pathways. Of these, two genes; peroxisome proliferator-activated receptor gamma (*Pparg)* and vascular endothelial growth factor C *(Vegf-c),* which plays a role in the PPARG and *mTOR* signaling pathways, respectively, were selected as genes linked to T2D.

For miR-379, two of the seven mRNA targets commonly predicted by TargetScan, DIANA and PITA mapped to KEGG pathways. Of these, the insulin receptor (*Insr)* gene which plays a role in the development of T2D through its effect on the insulin signaling pathway was selected.

Four of the 13 mRNA target genes commonly predicted by TargetScan, DIANA and PITA for miR-21 mapped to KEGG pathways. Of these, one target gene; tuberous sclerosis 1 *(Tsc1),* was selected due to its role in T2D through effects on the mTOR and insulin signaling pathway.

For miR-98, four of the 12 mRNA target genes commonly predicted by TargetScan, DIANA and PITA mapped to KEGG pathways. Of these, three target genes; mitogen-activated

protein kinase 3 *(Map4k3)*, Unc-51-like kinase *(Ulk2)* and stearoyl-CoA desaturase (delta-9-desaturase) *(Scd)* were potentially linked to T2D, and is predicted to play a role in *MAPK, mTOR and* PPARG signaling pathways, respectively. Additionally, *Scd* may also play a potential role in the biosynthesis of unsaturated fatty acids.

For miR-143, 18 of the 73 mRNA targets commonly predicted by TargetScan, DIANA and PITA mapped to KEGG pathways. Of these, six target genes; fibroblast growth factor 1 *(Fgf1)*, calcium channel voltage-dependent alpha 1A subunit *(Cacna1a)*, protein kinase C, epsilon *(Prkce)*, tao kinase 2 *(Taok2),* v-ki-ras2 Kirsten rat sarcoma viral oncogene homolog *(Kras)*, and an independent companion of *mTOR* complex 2 *(Rictor)* may play a role in T2D disease progression. These genes are predicted to be involved in *mTOR*, *MAPK*, insulin and chemokine signaling.

Table 3.10 Significantly differentially expressed miRNAs and their mRNA targets. Genes listed were identified to play a potential role in T2D-related metabolic pathways, identified using the DAVID/KEGG pathway database.

MiRNA	No. of predicted mRNA targets	No. of linked to KEGG pathway	Genes linked to T2D*	Role of genes in T2D related metabolic pathways
miR-27b	28	12	*Pparg*	PPARG signaling
			Vegf-c	mTOR signaling
miR-379	7	2	*Insr*	Insulin signaling and T2D disease pathogenesis
miR-21	13	4	*Tsc1*	mTOR and insulin signaling
miR-98	12	4	*Map4k3*	MAPK signaling
			Ulk2	mTOR signaling
			Scd	PPARG signaling and the biosynthesis of fatty acids
miR-143	73	18	*Rictor*	mTOR signaling
			Taok2	MAPK signaling
			Cacna1a	T2D disease pathogenesis
			Fgf1	MAPK signaling
			Prkce	T2D disease pathogenesis
			Kras	MAPK, chemokine and insulin signaling

***Manually selected genes that have previously been reported to play a role in metabolic pathways linked to T2D.

3.5.3 Protein-protein interactions regulated by differentially expressed microRNAs

Protein-protein interactions possibly regulated by the differentially expressed miRNAs were investigated using the protein interaction database called STRING. Predicted mRNA target gene/proteins were investigated, and related interacting proteins are illustrated in Appendix 4. For example, miRNA-27b target: vascular endothelial growth factor C *(VEGFC)* interacts with fms-related tyrosine kinase 4 *(FLT4)*, kinase insert domain receptor *(KDR)*, neuropilin 2 *(NRP2)*, fms-related tyrosine kinase 1 *(FLT1)*, and insulin-like growth factor 1 receptor *(IGF1R)*. The estimated likelihood that the predicted proteins interact in the same metabolic map in the KEGG database was indicated by a 95% confidence score generated by the active prediction methods in the STRING tool.

3.5.4 Functional analysis of microRNAs

In addition to functional analysis of mRNA target genes, biological activity of the differentially expressed miRNAs was also explored. Using HMDD, the biofluid and disease association of the differentially expressed miRNAs were investigated using the miRNA database, HMDD. MiR-27b, miR-143 and miR-21 were previously identified in biofluids such as serum, plasma, urine and faeces, and were mostly found to be associated with neoplasms and different types of cancers (Table 3.11). Circulating miR-21 have also been found to be associated with T2D (Table 3.11). In contrast, miR-98 and miR-379 has not yet been detected in its circulatory form.

Table 3.11 Functional analysis of miRNAs using the HMDD (Li *et al.*, 2014).

MiRNA ID	Bio-fluid detection	Disease association
hsa-miR-27b	urine	urinary bladder neoplasms
miR-379	none	n/a
miR-21	plasma and serum	fatty liver non-alcoholic disease, neoplasm, hepatitis C, different cancer types, and T2D
miR-98	none	n/a
miR-143	faeces	pancreatic neoplasms

3.5.5 Experimental validation of microRNA:mRNA target interactions

The databases DIANA-TarBase and miRTarBase were used to investigate whether the interactions between differentially expressed miRNAs (miR-27b and miR-143) and their predicted mRNA targets were experimentally validated. MiRNA:mRNA interactions were validated using a range of experimental methods including Reporter assays, Western blot, qRT-PCR, Microarrays and others (Table 3.8). The predictive scores ranged from 0 - 1, with 1 representing the highest likelihood of an interaction, and 0 representing the lowest likelihood of a true interaction (Table 3.8). Both DIANA-TarBase and miRTarBase gave high predictive scores for the interaction between miR-27b and PPARG, and between MiR-143 and KRAS.

Table 3.12 Experimental validation of miRNA:mRNA target interactions.

		Prediction program		Validation method						
miRNA	**Target gene**	**DIANA-Tarbase**	**miRTar Base**	**Reporter assay**	**Western blot**	**qRT-PCR**	**Microarrays**	**Other**	**Predictive score**	**References**
miR-27b	*Pparg*	yes	yes	✓	✓	✓		✓	1	3
miR-143	*Kras*	yes	yes	✓	✓	✓	✓	✓	0.994	5

4. DISCUSSION

The prevalence of T2D continues to increase worldwide, placing a huge financial and health burden on governments, particularly those in middle and low income countries (Whiting *et al.*, 2011). The costs to treat T2D are worsened by the multitude of associated morbidities such as CVD, stroke and peripheral vascular disease (Fowler, 2008). Without effective intervention, the WHO predicts that T2D may become the world's greatest health challenge within the next 25 years, a prediction supported by the high numbers of undiagnosed cases globally, and the alarming increase of T2D in children (WHO, 2006).

A major priority in T2D research is the prevention, improved diagnosis and management of the disease. It is widely believed that the identification of individuals during the pre-diabetic, sub-clinical stage of disease will facilitate intervention strategies which may prevent or delay their disease progression (Sherwin *et al.*, 2004). Such intervention strategies could include lifestyle modifications such as diet and exercise, or therapeutic interventions that could protect against insulin resistance, β-cell dysfunction and obesity. The early identification of individuals who are at risk of T2D, before the clinical onset of symptoms, will decrease the burden of T2D and its micro- and macro-vascular complications.

Research efforts to reduce the burden of T2D by early intervention are underpinned by the need for biomarkers to identify subclinical disease. Despite the availability of the OGTT, FPG and HbA1c, limitations of these screening and diagnostic tests prevent the identification of high risk individuals or the early identification of T2D before the manifestation of micro- and macro-vascular complications. Recently a number of studies have reported that miRNAs play an important role in various diseases, including T2D (Guay *et al.*, 2011). Although miRNAs were first identified in tissue and play important regulator roles in cells, they also occur in PBMCs, or as free circulating miRNAs in serum and plasma, suggesting that they could be used as potential, minimally invasive blood-based biomarkers of disease (Lagos-Quintana *et al.*, 2002; Mao *et al.*, 2013). The ultimate goal would be to develop a reliable, affordable, straightforward point-of-care device that can analyze a patient's serum/plasma, urine or saliva in a clinical setting, allowing on-site miRNA detection and disease diagnosis.

In this study we hypothesized that circulating miRNAs differ during disease progression, and that these epigenetic marks could be used as biomarkers which will differentiate between normo-glycaemic, pre-diabetic and diabetic (T2D) individuals.

4.1 Clinical characterization of participants

Study participants were classified as diabetic, pre-diabetic or normo-glycaemic according to their FPG, OGTT and HbA1c levels. Higher fasting insulin and 2 hr OGTT insulin concentrations were observed in pre-diabetic compared to normo-glycaemic individuals, consistent with studies that have reported that pre-diabetes is characterized by insulin resistance and hyperinsulinaemia (Beagley *et al.*, 2014). During insulin resistance, insulin sensitive cells fail to respond to physiological concentrations of insulin, causing β-cells to compensate by increasing insulin secretion, resulting in hyperinsulinaemia and hyperglycaemia, as observed in the pre-diabetic individuals in this study. Similarly, fasting and 2 hr C-peptide concentrations were higher in pre-diabetic compared to diabetic and normo-glycaemic individuals. C-peptide is the cleavage product produced during the processing of insulin from the proinsulin molecule to mature insulin, and is secreted by β-cells in equimolar concentrations to insulin (Akuri, 2014). It is widely believed that quantification of C-peptides are better markers of β-cell secretory activity, since they are more stable in blood over longer time periods compared to insulin (~6-fold greater half-life) (Palmer *et al.*, 2004). Insulin concentrations in diabetics were decreased compared to normo-glycaemic individuals, illustrating possible β-cell dysfunction and failure to secrete insulin (Kahn, 2003).

Both insulin and glucagon concentrations were increased in pre-diabetics and diabetics compared to normo-glycaemic individuals, consistent with the hypothesis that dysglycaemia causes dysregulation of regulatory mechanisms in pancreatic cells, such as the failure to inhibit glucagon secretion from alpha (α) cells during hyperglycaemia (Moon and Won, 2015). The glucagon-like peptide hormone (GLP-1) is an incretin hormone secreted by the ileum in response to nutrients. Activation of GLP-1 stimulates insulin secretion and inhibits glucagon secretion in a glucose-dependent manner under normal fed conditions. During T2D, defects in α-cells cause impaired glucose sensing, thus preventing glucagon inhibition (Dunning *et al.*, 2005). It has also been reported that α-cells become resistant to insulin under conditions of impaired glucose sensing. For example, α-cell specific insulin receptor

knockout mice, exhibited mild glucose intolerance, hyperglycaemia and hyperglucagonaemia compared to control mice in the fed state, suggesting dysregulation of insulin and glucose sensing mechanisms (Kawamori *et al.*, 2009).

Despite differences observed in insulin, C-peptide and glucagon concentrations, these were not statistically significant. The failure to attain statistical significance could be due to the small sample size and heterogeneity between the individuals studied. Individuals of mixed ancestry originate from a combination of different ethnic backgrounds including Khoisan, African, Caucasian and Asian, thus represent a diverse genetic group (de Wit *et al.*, 2010). A number of studies have reported the complexities of studying mixed ancestry individuals. However we selected this population group due to their high prevalence in the Western Cape of South Africa where the study was based, and because of the prevalence of T2D within mixed ancestry communities (Erasmus *et al.*, 2012). Erasmus *et al.* (2012) reported that the prevalence of T2D was 28.2% in a mixed ancestry cohort in Bellville South, a suburb in the Western Cape. A previous study conducted by Levitt and colleagues in a mixed ancestry community in Mamre, another suburb in the Western Cape, reported a prevalence of 10.8% (Levitt *et al.*, 1993). The same authors (Levitt *et al.*, 1993) reported that the prevalence of T2D was 8.0% amongst black Africans living in African residential areas in Cape Town. These findings support the selection of mixed ancestry individuals to investigate the association between miRNAs and T2D progression.

Diabetes is often characterized by hyperlipidaemia (Krauss, 2004). In this study, no difference in the lipid profile (HDL, LDL, cholesterol and triglycerides) between diabetic, pre-diabetic and normo-glycaemic individuals was observed. Study participants in each group were matched according to BMI, to reduce confounding factors due to bodyweight, and could explain the failure to observe differences in the lipid profile between these groups. Several studies have shown the association between lipid composition and BMI (Shamai *et al.*, 2011; Szczygielska *et al.*, 2002). Szczygielska *et al.* (2002) reported that triglyceride and total cholesterol concentrations were higher in obese individuals compared to normal weight subjects (BMI<25 kg/m^2), while lower HDL cholesterol concentrations were observed in obese individuals compared to normal subjects. Another study showed a positive relationship between BMI and triglycerides in individuals with a wide range of BMI values (Shamai *et al.* 2011). All participants in this study were either overweight (BMI>25 kg/m^2) or obese (BMI>30 kg/m^2) with a mean BMI greater than 33 kg/m^2.

A number of factors including ethnicity, age, gender, BMI, socioeconomic status, nutrition, toxins, physical inactivity, alcohol, smoking, etc. are known to modify the epigenome and cause epigenetic dysregulation (Hamilton, 2011). Unfortunately, many of these factors, such as alcohol consumption, smoking, physical activity and diet were not known for the participants of this study. Participants were recruited from communities with similar socio-economic statuses, so as to minimize the effects of environmental influences. Moreover, individuals were matched for ethnicity, age, gender and BMI to further minimize confounding and bias.

4.2 MicroRNA sequencing

Next generation sequencing has gained popularity for assessing RNA abundance in biological samples, and as a high throughput platform for detecting differentially expressed known and novel miRNA expression profiles (David, 2013; Tam *et al.*, 2014). The majority of sequencing reads obtained with NGS varied between 19 nt and 24 nt in length, correlating with the average size of miRNAs in the human genome (Felekkis *et al.*, 2010). An average of 4.9 million reads were mapped to small RNA species, with a total of 294 miRNAs expressed among the 12 samples. Illumina technology can detect between 3.4 million and 3 billion reads per run (Moldoval *et al.* 2014). Total read counts vary according to sample size, sample type and disease association (Lopez *et al.*, 2015).

Our study identified 277, 267 and 267 differentially expressed miRNAs between diabetics vs. normo-glycaemics, pre-diabetics vs. normo-glycaemics, and pre-diabetics vs. diabetics, respectively. The presence of such a large population of differentially expressed miRNAs between diabetic, pre-diabetic and normo-glycaemic individuals provide support for their involvement in the regulation of genes important in the development of T2D (Karolina *et al.*, 2011). For example, a study conducted in HEK293T cells using the Dual-Luciferase assay reported that high expression levels of miR-199 as seen in T2D, represses GLUT4 expression (Yan *et al.*, 2014). Overexpression of miR-199 reduces glucose uptake in rat L6 muscle cells, as repression of GLUT4 inhibits insulin-stimulated glucose uptake (Yan *et al.*, 2014). Karolina *et al.* (2011) showed significant upregulation of miR-144 in PBMCs of T2D patients compared to healthy controls. Furthermore, a significant downregulation of IRS1 was observed when miR-144 was overexpressed, and a significant upregulation was

observed when miR-144 was inhibited, thus, illustrating the potential role of miR-144 in the insulin signaling pathway (Karolina *et al.*, 2011).

Although NGS provides a high throughput platform that is highly sensitive and specific for miRNA detection (Motameny *et al.*, 2010), the high cost of the technique limits the amount of samples that can be investigated. Due to inter-individual heterogeneity, as observed in our study, large sample sizes are required to attain statistical and clinical significance.

4.3 Validation of microRNA sequencing by quantitative real time PCR

The differential expression of five miRNAs identified by NGS was validated by qRT-PCR in the same PBMC samples that were previously used. Using the arbitrary cut-off threshold values set by ArrayStar; five differentially expressed miRNAs (miR-143, miR-98, miR-21, miR-27b and miR-397) were selected for further analysis. Of the five miRNAs tested, qRT-PCR results for miR-27b confirmed sequencing results, showing a trend towards significance in pre-diabetics compared to normo-glycaemic individuals. None of the expression differences of the other four miRNAs identified by NGS were confirmed by qRT-PCR. However, miR-143 was significantly upregulated in pre-diabetics compared to normo-glycaemics in contrast to sequencing, where a significant downregulation was observed in pre-diabetics compared to diabetics. No expression differences for the other three miRNAs were observed between disease groups.

The discrepancies between qRT-PCR and sequencing results could be due to technical differences of the techniques. For qRT-PCR, technical aspects that could influence the results include template quality, poor assay efficiency, primer–dimers (Klein, 2002), subjectivity in data analysis, and the selection of reference genes (Bustin and Nolan, 2004). Quantitative real time PCR produces expression data relative to stably expressed reference genes, whereas miRNA sequencing data generates millions of reads per sample, providing a more absolute representation of miRNA expression. This enables the highly sensitive sequencing technique to detect both high and low miRNA expression levels mapped to the entire human reference genome. Although qRT-PCR provides an easy, rapid and cost effective platform for quantifying miRNA expression profiles, NGS is considered more sensitive than qRT-PCR, and can detect novel miRNAs (Motameny *et al.*, 2010).

Furthermore, a lack of consensus exists regarding the selection of reference genes used for normalizing miRNA expression profiles in qRT-PCR (Marabita *et al.*, 2015). The most common choice of reference genes are endogenous miRNAs that are stably expressed among all samples. However, the selection of these genes may be arbitrary in a small sample size when based entirely on the variability of raw data, without further computational analysis. In our study NormFinder was used to confirm the appropriate selection of reference genes for normalisation, while some other studies use BestKeeper or eNorm, to identify the most suitable reference genes (Liu *et al.*, 2014). A number of published articles select reference genes randomly without validation, which could explain the lack of comparability in miRNA expression patterns in these studies (Tiberio *et al.*, 2015). Thus, although qRT-PCR is routinely used to validate NGS data, it has a number of limitations, which should be taken into account during data analysis.

4.4 Functional analysis of predicted messenger RNA targets

To gain insight into the biological function of these miRNAs, mRNA targets were predicted using three target prediction software programmes; TargetScan, PITA and DIANA. These programs differ in the algorithms they use for target prediction. TargetScan predicts mRNA targets based on sequence complementarity and conservation (Friedman *et al.*, 2009; Lewis *et al.*, 2005), DIANA focuses on the minimum binding energy between miRNA and sequences in the mRNA target region (Kiriakidou *et al.*, 2004; Maragkakis *et al.*, 2009), and PITA focuses on free energy and target site accessibility (Witkos *et al.*, 2011). Since target prediction programs are prone to predicting false-positives (Kertesz *et al.*, 2007; Kiriakidou *et al.*, 2004; Lewis *et al.*, 2005), only targets that were commonly predicted by all three programs were selected, thereby reducing the chance of identifying false-positive targets. In our study, TargetScan had a higher likelihood of predicting a true mRNA target compared to DIANA and PITA, based on the assumption that the commonly predicted targets were true-positives. In recently published literature, TargetScan has been considered one of the best target prediction programs, as it incorporates evolutionary conservation i.e. miRNA and mRNA sequences highly conserved across a wide range of species, as well as strict complementarity between sequences (Lewis *et al.*, 2005). Algorithms which consider complementarity along with conserved nature and/or free energy have been shown to perform better than methods that consider a single characteristic only (Yue *et al.*, 2009).

The functional role of mRNA targets were investigated using the DAVID gene functional annotation tool, which classifies mRNA targets according to their inter-relationships between gene groups and biological terms, specific biological pathways, and functional significance in biological modules (Huang *et al.*, 2007). A common problem encountered with DAVID is that the output of data is often large and overwhelming (Huang *et al.*, 2007; Huang *et al.*, 2008). Therefore, we focused our analysis on mRNA targets that mapped to KEGG pathways. Messenger RNA genes that were associated with metabolic pathways linked to the development of T2D were selected based on personal knowledge.

The personal selection of genes linked to T2D may have missed important genes due to the limitations inherent in personal knowledge. The selected T2D-related genes were confirmed using the T2D-db. The T2D-db provides information of all molecular factors known to be involved in the pathogenesis of T2D in humans (Agrawal *et al.*, 2008). Three of the selected genes (*Pparg, Kras* and *Insr)* known to be involved in T2D disease development were associated with the database. The lack of association between other genes may be due to missing information of newly discovered T2D-related genes, since the database was last updated in 2009.

To identify signaling cascades that may be affected by miRNA regulation, interacting proteins of predicted miRNA target gene/proteins were evaluated and identified using STRING. Analyzing gene targets in the context of protein regulation provides insight into how the dynamic biological system is controlled (Szklarczyk *et al.*, 2014). The interactions between proteins identified in our study illustrates various metabolic pathways implicated in the pathogenesis of T2D, suggesting that miRNA dysregulation forms part of a wider network of complex regulatory systems (Liang and Li, 2007).

4.5 Functional analysis of microRNAs

In addition to conducting functional analysis of target genes, we also investigated the biological role of miRNAs in circulation using HMDD. This database provides information about the association of miRNAs with disease. For example, miR-27b was previously investigated in a study using miRNAs as urinary biomarkers in cancer patients (Miah *et al.*, 2012). MiR-143 was expressed at high levels in faecal samples of pancreatic cancer patients and healthy individuals (Link *et al.*, 2012). Interestingly, miR-21 was decreased in plasma of

T2D patients compared to normo-glycaemic individuals (Zampetaki *et al.*, 2010). MiR-98, and miR-379 have not yet been discovered in its circulatory form.

4.6 Computational validation of microRNA:mRNA target interactions

In this study, we investigated the relationship between differentially expressed miRNAs and their T2D-related target genes, using DIANA-TarBase and miRTarBase, with a special focus on miR-27b and miR-143, as these showed significance in the validation analysis. These databases provide information for miRNA:mRNA target interactions that have previously been experimentally validated, providing a list of publications for each interaction (Elton and Yalowich, 2015; Hsu *et al.*, 2014). The analysis showed high predictive scores for previously experimentally validated miRNA:mRNA target interactions between miR-27b and *Pparg,* and between miR-143 and *Kras,* by three or more techniques. For example miR-27b:*Pparg* interaction was validated using a Luciferase reporter assay, Western blot and qRT-PCR (Karbiener *et al.*, 2009; Jennewein *et al.*, 2010; Lee *et al.*, 2012), while miR-143:*Kras* was validated using a Luciferase reporter assay, Western blot, qRT-PCR and Microarrays (Chen *et al.*, 2009; Lin *et al.*, 2009; Pagliuca *et al.*, 2012; Ni *et al.*, 2012), illustrating direct binding of miRNAs on predicted target sites.

4.7 Novel microRNA

Since their initial discovery in *C. elegans* in 1993 (Lee *et al.*, 1993), over 30,000 mature miRNA sequences have been identified in more than 150 species (miRBase v21). Conservative estimates indicate that over 1000 miRNAs are found in the human genome, where they regulate a large number of genes (Ardekani and Naeini, 2010). Our sequencing data identified a total of 151 novel miRNAs among all samples. Of these, 35, 33 and 33 novel miRNAs were differentially expressed in diabetic vs. normo-glycaemic, pre-diabetic vs. normo-glycaemic and pre-diabetic vs. diabetics, respectively. Validation of novel miRNAs identified by sequencing is required before they can be regarded as a 'true' finding (Juan *et al.*, 2014). In this study, the expression of five novel miRNAs (MYNO22, MYNO59, MYNO66, MYNO8 and MYNO95) was confirmed using qRT-PCR. The sequences of these novel miRNAs will be submitted into the online miRNA registry for the assignment of an appropriate miRNA gene name, after a manuscript describing experimental validation has been accepted (Griffiths-Jones *et al.*, 2006).

Both sequencing and qRT-PCR showed that three of the five novel miRNAs, MYNO59, MYNO95 and MYNO66 were downregulated in diabetics compared to normo-glycaemic individuals. Moreover, MYNO66 was significantly downregulated with disease progression, i.e. decreased expression in normo-glycaemics > pre-diabetics > diabetics. These results suggest that these novel miRNAs, particularly MYNO66, are associated with disease development, warranting further investigation of their role in T2D. Investigating the expression of these miRNAs under different metabolic conditions, and then identifying their mRNA targets, may provide a better understanding of their role in the development of T2D.

4.8 MicroRNA profiling in serum

Circulating miRNAs that occur cell-free in plasma and serum have attracted considerable interest as biomarkers for numerous diseases due to their highly stable nature (Chen *et al.*, 2008). The expression of miRNAs that were shown to be differentially expressed in PBMCs was investigated in serum samples of the same individuals. All five miRNAs were detected in serum of diabetic, pre-diabetic and normo-glycaemic individuals. Of these, only miR-27b was significantly upregulated in pre-diabetics compared to normo-glycaemic individuals. These results were consistent with results observed for sequencing and qRT-PCR in PBMCs. No difference in the expression of the other four miRNAs was observed.

The discrepancies of miRNA expression in PBMCs and serum samples could be due to the different qRT-PCR methods used. In serum samples, miRNAs were detected using the miScript PCR SYBR Green detection kit. This method uses SYBR Green I dye that intercalates into double-stranded cDNA, as well as into non-specific reaction products, and emits a fluorescent signal that is directly proportional to the amount of PCR product generated (Rajeevan *et al.*, 2001). On the other hand, miRNAs in PBMCs were detected using TaqMan primers and specific probes to the miRNAs of interest. This method detects only specific amplification products (Ponchel *et al.*, 2003). Thus differences of miRNA expression profiles in PBMCs and serum could be due to variable sensitivity and specificity of the validation platforms used (Bustin and Mueller, 2005).

Another factor to consider is the miRNA expression profile in different biological fluids. Wang *et al.* (2012) investigated the expression of miRNAs in serum, plasma and blood components (red blood cells (RBC), white blood cells (WBC) and platelets) of six healthy

individuals. Their study showed that more miRNAs were detected in blood cells, ranging from 280 and 477 miRNAs compared to serum and plasma, where only 118 miRNAs and 106 miRNAs were detected, respectively. The most abundantly expressed miRNAs (miR-126, miR-16 and miR-150) in blood cells were equally expressed in serum and plasma samples, while the most abundantly expressed miRNA in serum was observed at much lower levels in RBC (Wang *et al.*, 2012). Therefore, although biofluids share similar miRNA profiles, their patterns differ according to tissue and cell type, which may be due to differential uptake or release of miRNAs from different circulating cell types in the blood (Maldovan *et al.*, 2014).

Despite these discrepancies, to our knowledge, this study is the first to demonstrate the differential expression of miR-27b in PBMCs and serum of pre-diabetic and normo-glycaemic individuals, thus, illustrating its potential as a biomarker of pre-diabetes. Other studies have shown the association of miR-27b with a wide range of cancers, CVD, adipogenesis, atherosclerosis, T1D and diabetic nephropathy (Bi *et al.*, 2015; Goto *et al.*, 2014; Karbiener *et al.*, 2009; Kong *et al.*, 2014; Lehmann *et al.*, 2015; Veliceasa *et al.*, 2015; Zampetaki *et al.*, 2015). Using genome wide microarray analysis and qRT-PCR, Lin *et al.* (2009a) showed significant downregulation of miR-27b during adipogenesis, consistent with other literature (Karbiener *et al.*, 2009). Over expression of miR-27b inhibits adipocyte formation, without affecting myogenic differentiation, suggesting its role in the pathological development of obesity (Lin *et al.*, 2009a). Moreover, miR-27b was one of 12 miRNAs investigated in serum of diabetic, obese-diabetic, and healthy individuals (Pescador *et al.*, 2013). However, no significance in miRNA expression was observed between these groups. MiR-27b was significantly upregulated in serum of T1D patients compared to healthy individuals, illustrating a possible role in β-cell apoptosis and β-cell regulatory networks (Nielsen *et al.*, 2012). Li *et al.* (2010) showed the significant upregulation of three miRNAs, including miR-27b in the serum and tissue of patients with atherosclerosis compared to healthy controls, as evidence for a possible risk marker in the early stages of atherosclerosis (Li *et al.*, 2011). Results of these studies illustrate a possible role of miR-27b in T2D-related complications. However, data directly linking miR-27b to pre-diabetes and T2D is still lacking.

4.9 MiR-27b as a potential biomarker for Pre-diabetes

Both sequencing and qRT-PCR showed a significant increase in the expression of miR-27b in PBMCs and serum of pre-diabetics compared normo-glycaemic individuals. These results suggest that miR-27b could be used as a potential biomarker for pre-diabetes; the early stages of T2D development that precedes overt diabetes. Individuals with pre-diabetes have a 3 to 10–fold increased risk of developing T2D (Garber *et al.*, 2008), and the identification of these individuals in the early stages could delay or prevent disease progression through lifestyle interventions or therapeutic strategies to reverse progression. However, due the limitations of this study and the extensive heterogeneity and non-specificity of miRNAs (Turchinovich *et al.*, 2012), validation of miR-27b will be required in a larger cohort and in different populations to confirm its potential as a biomarker to identify those at risk for T2D.

It is currently suggested that a biomarker should be involved in disease progression. MiR-27b was not increased in diabetics, compared to pre-diabetics and normo-glycaemics, suggesting an alteration associated with a specific physiological state in the disease. Unlike intermediate biomarkers, which represent direct steps in the causal pathway of disease, circulating biomarkers represent changes in biological factors that lead to the disease (Mayeux, 2004), and are used as indicators of disease susceptibility. Thus, these biomarkers need not always be causal, and may merely reflect the subclinical stage of disease to be considered useful for screening and prevention.

4.9.1 Molecular mechanism of miR-27b

Both computational prediction programs and experimental studies have identified *Pparg* as a target for miR-27b. The positive effect of PPARG on insulin signaling has been widely reported (Leonardini *et al.*, 2009). Our results suggest that upregulation of miR-27b in pre-diabetics suppress PPARG activity by inhibiting *Pparg* with negative effects on insulin signaling and subsequent glucose uptake. The effects of insulin on glucose uptake is mediated by phosphorylation of the tyrosine kinase insulin receptor, which leads to increased IRS-1–associated PI3K activity, and subsequent translocation of GLUT4 from cytoplasmic vesicles to the cell membrane, and uptake of glucose into the cell, thereby reducing overall blood glucose concentrations (Chakraborty *et al.*, 2014). Dysregulation of intracellular proteins involved in the insulin signaling cascade leads to the development of

IR, characterized by high insulin and glucose concentrations. Indeed, insulin and glucose levels were increased in pre-diabetics compared to normo-glycaemics (Fig 4.1).

Using a Cre-loxP system for the deletion of *Pparg* in mouse skeletal muscle, Hevener *et al.* (2003) showed that mice with a *Pparg* deletion exhibited glucose intolerance and progressive insulin resistance after four months. Furthermore, using a hyperinsulinaemic-euglycaemic clamp, these authors showed a 80% reduction in insulin-stimulated glucose uptake *in vivo*, thus, confirming the crucial role that PPARG plays in skeletal muscle insulin action (Hevener *et al.*, 2003). Verna *et al.* (2004) showed that inhibition of *Pparg* in C2C12 skeletal muscle cells induced insulin resistance under normal conditions, while overexpression of *Pparg* in the presence of insulin and/or Pioglitazone (anti-diabetic drug) stimulation increased glucose uptake under both normal and insulin resistant conditions (Verma *et al.*, 2004).

Consistent with our findings, other studies showed the regulatory effect of miR-27b on *Pparg*. For example, transient transfection of miR-27b in human multipotent adipose-derived stem cells decreased the expression of adipogenic markers including *Pparg* (Karbeiner *et al.*, 2009). Using the Dual-Luciferase assay, Karbeiner *et al.* (2009) demonstrated that miR-27b directly binds to the 3' UTR of human *Pparg*, thereby decreasing PPARG protein levels, and inhibiting adipogenesis. Similarly, the expression of miR-27b in 3T3-L1 pre-adipocytes, inhibit both PPARG and CCAAT-enhancer-binding protein α (*C/EBPα)*, thus, confirming its role as an adipogenic inhibitor (Lin *et al.*, 2009). Moreover, miR-27b also targets *Pparg* in neuroblastoma cells, which leads to neuroblastoma growth inhibition, associated with decreased *Pparg* target gene expression, and reduced inflammatory responses. In the absence of miR-27b, *Pparg* promotes tumour formation in neuroblastomas (Lee *et al.*, 2012). Together, these studies suggest that disease progression associated with impaired PPARG signaling is often accompanied by the ability of miR-27b to bind to *Pparg* and decrease mRNA and protein expression.

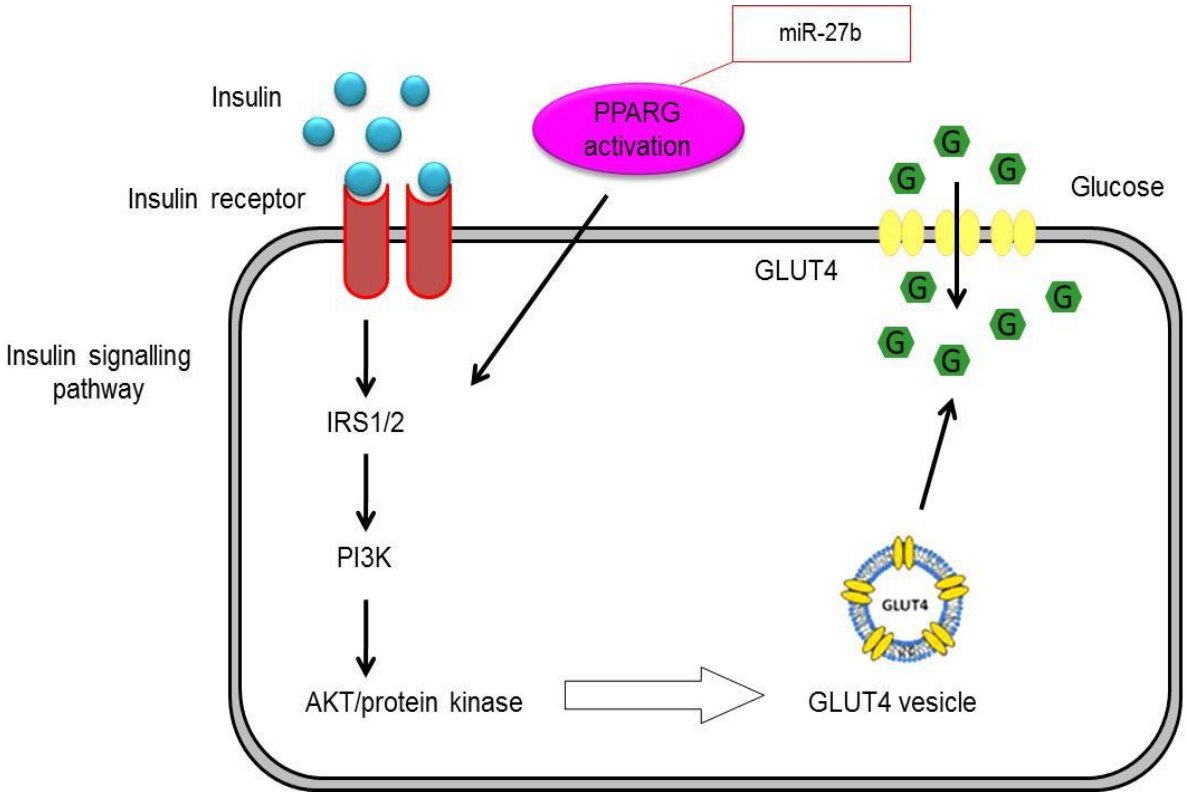


Figure 4.1 Schematic representation of the insulin signaling pathway and a possible role of miR-27b.The peroxisome proliferator-activated receptor gamma (PPARG) stimulates the insulin signaling pathway, which leads to the translocation of glucose transporter 4 (GLUT4) vesicle to the plasma membrane, allowing glucose uptake into the cell. The negative effect of miR-27b on PPARG causes dysregulation of insulin signaling proteins (insulin receptor substrate 1 (IRS-1) and kinase/serine-threonine protein kinase-1 (PI3K/AKT)), thereby preventing glucose uptake.

4.10 Limitations

Our study has a number of limitations. These include the small sample size and the absence of data for environmental factors that are known to affect miRNA expression. The small sample size was due to a number of factors including 1) the challenge to recruit diabetic individuals that were not on treatment, 2) the requirement of matching participants for ethnicity, age, gender and BMI, and 3) the high cost of sequencing. The primary goal of diabetic treatment is to restore glucose homeostasis. However, changes in glucose concentrations are known to alter miRNA expression patterns. For example, increased miRNA (miR-192 and miR-193b) expression levels in serum of pre-diabetics compared to T2D individuals was decreased after chronic exercise, illustrating the effects of glucose control on miRNA expression (Parrizas *et al.*, 2014). Matching participants in each group was necessary to reduce variability between individuals, and to further minimize confounding and bias. The requirement to match individuals for ethnicity, age, gender and BMI limited the number of individuals that could be studied. The high cost of sequencing prevented the analysis of a larger sample set. Large sample sizes are required to account for inter-individual heterogeneity between samples, and to attain clinical significance.

In addition, our study was limited by the lack of data regarding diet, physical activity, and behavior such as smoking and alcohol consumption, as these factors are well known to contribute to changes seen in miRNA expression patterns.

4.11 Future work

The results of this study will form the basis for future work to explore the role of miRNAs in the pathogenesis of T2D. Studies should be conducted in larger cohorts and in different populations to investigate the potential of miRNAs such as miR-27b as biomarkers to identify those at high risk of developing T2D, and to facilitate intervention strategies that can prevent, delay or reverse disease progression. However, due to the extensive heterogeneity of miRNAs, and their ability to bind multiple targets, future investigations should focus on a panel of T2D-specific miRNAs to identify high risk individuals.

The biological role and cell-type specificity of miR-27b and the novel miRNAs could be investigated by conducting gain-of-function and loss-of-function experiments in the four major cell-types associated with T2D, i.e., pancreatic β-cells, skeletal muscle cells, hepatocytes and adipocytes. For example, miRNA mimics and inhibitors could be transfected into pancreatic β-cells to investigate their effects on insulin secretion. Using a similar strategy, Tang *et al.* (2009) reported decreased glucose-stimulated insulin gene transcription in β-cells (MIN6) after miR-30d inhibition (Tang *et al.*, 2009). A number of miRNA target genes are involved in insulin signaling (Granjon *et al.*, 2009), thus their biological role could be investigated by transfection studies in skeletal muscle cells to measure the effectes of glucose metabolism under normal and insulin resistant conditions. The role of miRNAs in hepatocytes (Herrara *et al.*, 2010) and adipocytes (Hilton *et al.*, 2013) have been described, thus transfection experiments in hepatocytes and adipocytes could elucidate the molecular mechanism of these miRNAs.

4.12 Conclusion

In conclusion, we showed that miRNA profiles differ during T2D progression, and are able to discriminate between diabetic, pre-diabetic and normo-glycaemic individuals. Despite the limitations of the study, as alluded to previously, this study showed that miR-27b is differentially expressed in both PBMCs and serum of pre-diabetic compared to normo-glycaemic individuals. These findings demonstrate the importance of miR-27b in the pathophysiology of T2D and suggest its potential as a biomarker that could be incorporated into predictive models for the identification of high risk individuals with glucose dysregulation. Such strategies could facilitate interventions to prevent or better manage T2D, ultimately reducing the disease burden. These predictive models could include miRNA expression patterns and other known risk factors for pre-diabetes or T2D. However, miRNA profiling in a larger sample size and prospective longitudinal studies are required to assess clinical applicability.

References

Agarwal, V., Bell, G.W., & Bartel, D.P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *eLife, 4,* e05005

Agilent Technologies. (2007) [Cited 20 March 2015]; Available from: http://www.chem.agilent.com/Library/technicaloverviews/Public/5989-7002EN.pdf

Agrawal, S., Dimitrova, N., Nathan, P., Udayakumar, K., Lakshmi, S.S., Sriram, S., Manjusha, N., & Sengupta, U. (2008) T2D-Db: an integrated platform to study the molecular basis of Type 2 diabetes. *BMC Genomics*, 9 (1), 320-332

Akuri, S.R. (2014) C-peptide is the better marker to rule out prediabetes in chronic pancreatic disorder patients. *Asian Journal of Pharmaceutical and Clinical Research,* 7 (4), 137-140

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S.A.M., & Marshall, M. (2003) A uniform system for microRNA annotation. *RNA*, 9 (3), 277-279

American Association of clinical endocrinology (AACE) diabetes resource centre. (2013) [Cited 20 July 2015]; Available from: http://outpatient.aace.com/type-2-diabetes/clinical-presentation-of-type-2-diabetes-mellitus

American Diabetes Association. (2015) Classification and diagnosis of diabetes. *Diabetes Care*, 38 (Suppl1), S8-S16

American Diabetes Association. (2014) Diagnosis and classification of diabetes mellitus. *Diabetes Care*, 37 (Suppl1), S81-S90

Amod, A., Motala, A., Levitt, N., Berg, J., Young, M., Grobler, N., Heilbrunn, A., Distiller, L., Pirie, F., & Dave, J. (2012) The SEMDSA Guideline for the Management of type 2 Diabetes. *Journal of Endocrinology Metabolism & Diabetes of South Africa,* 17 (1), S1-S94

Andersen, C.L., Jensen, J.L., & Orntoft, T.F. (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*, 64 (15), 5245-5250

Appajigol, J., Somannavar, M., & Araganji, R. (2011) Feasibility of oral glucose tolerance test as a diagnostic tool for diabetes prevalence study in a rural community. *Recent Research in Science & Technology*, 3 (9), 53-56

Araujo, T.G., Oliveira, A.G., & Saad, M.J. (2013) Insulin-resistance-associated compensatory mechanisms of pancreatic beta cells: a current opinion. *Frontiers in Endocrinology*, 4 (49), 1-2

Ardekani, A.M. & Naeini, M.M. (2010) The role of microRNAs in human diseases. *Avicenna Journal of Medical Biotechnology*, 2 (4), 161-179

Arroyo, J.D., Chevillet, J.R., Kroh, E.M., Ruf, I.K., Pritchard, C.C., Gibson, D.F., Mitchell, P.S., Bennett, C.F., Pogosova-Agadjanyan, E.L., & Stirewalt, D.L. (2011) Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proceedings of the National Academy of Sciences*, 108 (12), 5003-5008

Bain, J.R., Stevens, R.D., Wenner, B.R., Ilkayeva, O., Muoio, D.M., & Newgard, C.B. (2009) Metabolomics applied to diabetes research moving from information to knowledge. *Diabetes*, 58 (11), 2429-2443

Balasubramanyam, M., Aravind, S., Gokulakrishnan, K., Prabu, P., Sathishkumar, C., Ranjani, H., & Mohan, V. (2011) Impaired miR-146a expression links subclinical inflammation and insulin resistance in Type 2 diabetes. *Molecular and Cellular Biochemistry*, 351 (1-2), 197-205

Baroukh, N., Ravier, M.A., Loder, M.K., Hill, E.V., Bounacer, A., Scharfmann, R., Rutter, G.A., & Van Obberghen, E. (2007) MicroRNA-124a regulates Foxa2 expression and intracellular signaling in pancreatic beta-cell lines. *Journal of Biological Chemistry*, 282 (27), 19575-19588

Barr, R.G., Nathan, D.M., Meigs, J.B., & Singer, D.E. (2002) Tests of glycemia for the diagnosis of type 2 diabetes mellitus. *Annals of Internal Medicine*, 137 (4), 263-272

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116 (2), 281-297

Beagley, J., Guariguata, L., Weil, C., & Motala, A.A. (2014) Global estimates of undiagnosed diabetes in adults. *Diabetes Research and Clinical Practice*, 103 (2), 150-160

Benes, V. & Castoldi, M. (2010) Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods*, 50 (4), 244-249

Bennett, C.M., Guo, M., & Dharmage, S.C. (2007) HbA1c as a screening tool for detection of type 2 diabetes: a systematic review. *Diabetic Medicine*, 24 (4), 333-343

Bhattacharya, S., Dey, D., & Roy, S.S. (2007) Molecular mechanism of insulin resistance. *Journal of Biosciences*, 32 (2), 405-413

Bi, R., Bao, C., Jiang, L., Liu, H., Yang, Y., Mei, J., & Ding, F. (2015) MicroRNA-27b plays a role in pulmonary arterial hypertension by modulating peroxisome proliferator-activated receptor γ dependent Hsp90-eNOS signaling and nitric oxide production. *Biochemical and Biophysical Research Communications*, 460 (2), 469-475

Biden, T.J., Boslem, E., Chu, K.Y., & Sue, N. (2014) Lipotoxic endoplasmic reticulum stress, β-cell failure, and type 2 diabetes mellitus. *Trends in Endocrinology & Metabolism*, 25 (8), 389-398

Bonner-Weir, S., & O'Brien, T.D. (2008). Islets in type 2 diabetes: in honor of Dr. Robert C. Turner. *Diabetes*, 57 (11), 2899-2904

Boyle, P.J. (2007) Diabetes mellitus and macrovascular disease: mechanisms and mediators. *The American Journal of Medicine*, 120 (Suppl 9), S12-S17

Brennecke, J., Stark, A., Russell, R.B., & Cohen, S.M. (2005) Principles of microRNA target recognition. *PLoS Biology*, 3 (3), e85

Bustin, S.A., & Mueller, R. (2005) Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis. *Clinical Science*, 109 (4), 365-379

Bustin, S.A., & Nolan, T. (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of Biomolecular Techniques*, 15 (3), 155-166

Butt, A., & Swaminathan, R. (2015) Circulating Nucleic Acids and Diabetes Mellitus. *Annals N Y Acad Sc*, 5 (1), 258-270

Cai, X., Hagedorn, C.H., & Cullen, B.R. (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, 10 (12), 1957-1966

Callejas, D., Mann, C.J., Ayuso, E., Lage, R., Grifoll, I., Roca, C., Andaluz, A., Ruiz-de Gopegui, R., Montané, J., & Muñoz, S. (2013) Treatment of diabetes and long-term survival following insulin and glucokinase gene therapy. *Diabetes*, 62 (5), 1718-29

Chakraborty, C., Doss, C., Bandyopadhyay, S., & Agoramoorthy, G. (2014) Influence of miRNA in insulin signaling pathway and insulin resistance: micro-molecules with a major role in type 2 diabetes. *Wiley Interdisciplinary Reviews: RNA*, 5 (5), 697-712

Chang, X., Li, S., Li, J., Yin, L., Zhou, T., Zhang, C., Chen, X., & Sun, K. (2014) Ethnic differences in microRNA-375 expression level and DNA methylation status in type 2 diabetes of Han and Kazak populations. *Journal of Diabetes Research;* doi: 10.1155/2014/761938

Chang-Chen, K.J., Mullur, R., & Bernal-Mizrachi, E. (2008) Beta cell failure as a complication of diabetes. *Reviews in Endocrine and Metabolic Disorders*, 9 (4), 329-343

Chen, X., Ba, Y., Ma, L., Cai, X., Yin, Y., Wang, K., Guo, J., Zhang, Y., Chen, J., & Guo, X. (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Research*, 18 (10), 997-1006

Chen, X., Guo, X., Zhang, H., Xiang, Y., Chen, J., Yin, Y., Cai, X., Wang, K., Wang, G., & Ba, Y. (2009) Role of miR-143 targeting KRAS in colorectal tumorigenesis. *Oncogene*, 28 (10), 1385-1392

Chen, X., Liang, H., Zhang, J., Zen, K., & Zhang, C.Y. (2012) Secreted microRNAs: a new form of intercellular communication. *Trends in Cell Biology*, 22 (3), 125-132

Cheng, Y., & Zhang, C. (2010) MicroRNA-21 in cardiovascular disease. *Journal of Cardiovascular Translational Research*, 3 (3), 251-255

Christensen, B.C., & Marsit, C.J. (2011) Epigenomics in environmental health. *Frontiers in Genetics*, 2 (84), 1-10

Church, D., & Simmons, D. (2014) More evidence of the problems of using HbA1c for diagnosing diabetes? The known knowns, the known unknowns and the unknown unknowns. *Journal of Internal Medicine*, 276 (2), 171-173

Claus, T.H., & Pilkis, S.J. (1976) Regulation by insulin of gluconeogenesis in isolated rat hepatocytes. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 421 (2), 246-262

Cock, P.J., Fields, C.J., Goto, N., Heuer, M.L., & Rice, P.M. (2010) The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, 38 (6), 1767-1771

Cortez, M.A., Bueso-Ramos, C., Ferdin, J., Lopez-Berestein, G., Sood, A.K., & Calin, G.A. (2011) MicroRNAs in body fluids: the mix of hormones and biomarkers. *Nature reviews Clinical Oncology*, 8 (8), 467-477

Creemers, E.E., Tijsen, A.J., & Pinto, Y.M. (2012) Circulating microRNAs novel biomarkers and extracellular communicators in cardiovascular disease? *Circulation Research*, 110 (3), 483-495

Cristina, L., Roberto, L., & Stefano, D.P. (2008) Beta-cell Failure in Type 2 Diabetes Mellitus. *Current Diabetes Reports,* 8 (3), 179-184

Cuk, K., Zucknick, M., Heil, J., Madhavan, D., Schott, S., Turchinovich, A., Arlt, D., Rath, M., Sohn, C., & Benner, A. (2013) Circulating microRNAs in plasma as early detection markers for breast cancer. *International Journal of Cancer*, 132 (7), 1602-1612

Dangwal, S., Stratmann, B., Bang, C., Lorenzen, J.M., Kumarswamy, R., Fiedler, J., Falk, C.S., Scholz, C.J., Thum, T., & Tschoepe, D. (2015) Impairment of Wound Healing in Patients With Type 2 Diabetes Mellitus Influences Circulating MicroRNA Patterns via Inflammatory Cytokines. *Arteriosclerosis, Thrombosis & Vascular Biology*, 35 (6), 1480-1488

de Wit, E., Delport, W., Rugamika, C.E., Meintjes, A., Möller, M., van Helden, P.D., Seoighe, C., & Hoal, E.G. (2010) Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape. *Human Genetics*, 128 (2), 145-153

Deckers, J.G., Schellevis, F.G., & Fleming, D.M. (2006) WHO diagnostic criteria as a validation tool for the diagnosis of diabetes mellitus: a study in five European countries. *The European Journal of General Practice*, 12 (3), 108-113

Defronzo, R.A., & Tripathy, D. (2009) Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. *Diabetes Care*, 32 (suppl 2), S157-S163

Defronzo, R.A., Ferrannini, E.L.E.U., Sato, Y.U.Z.O., Felig, P.H.I.L., & Wahren, J. (1981) Synergistic interaction between exercise and insulin on peripheral glucose uptake. *Journal of Clinical Investigation*, 68 (6), 1468-1474

Dey, D., Mukherjee, M., Basu, D., Datta, M., Roy, S.S., Bandyopadhyay, A., & Bhattacharya, S. (2005) Inhibition of insulin receptor gene expression and insulin signaling by fatty acid: interplay of PKC isoforms therein. *Cellular Physiology & Biochemistry*, 16 (4-6), 217-228

Ding, X., Ding, J., Ning, J., Yi, F., Chen, J., Zhao, D., Zheng, J., Liang, Z., Hu, Z., & Du, Q. (2012) Circulating microRNA-122 as a potential biomarker for liver injury. *Molecular Medicine Reports*, 5 (6), 1428-1432

Dresner, A., Laurent, D., Marcucci, M., Griffin, M.E., Dufour, S., Cline, G.W., Slezak, L.A., Andersen, D.K., Hundal, R.S., & Rothman, D.L. (1999) Effects of free fatty acids on glucose transport and IRS-1 associated phosphatidylinositol 3-kinase activity. *Journal of Clinical Investigation*, 103 (2), 253-259

Drews, G., Krippeit-Drews, P., & Defer, M. (2010) Oxidative stress and beta-cell dysfunction. *Archiv-European Journal of Physiology*, 460 (4), 703-718

Du, T., & Zamore, P.D. (2005) microPrimer: the biogenesis and function of microRNA. *Development*, 132 (21), 4645-4652

Dunning, B.E., Foley, J.E., & Ahrén, B. (2005) Alpha cell function in health and disease: influence of glucagon-like peptide-1. *Diabetologia*, 48 (9), 1700-1713

Edgerton, D.S., Lautz, M., Scott, M., Everett, C.A., Stettler, K.M., Neal, D.W., Chu, C.A., & Cherrington, A.D. (2006) Insulin's direct effects on the liver dominate the control of hepatic glucose production. *Journal of Clinical Investigation*, 116 (2), 521-527

El-Ekiaby, N., Hamdi, N., Negm, M., Ahmed, R., Zekri, A.R., Esmat, G., & Abdelaziz, A.I. (2012) Repressed induction of interferon-related microRNAs miR-146a and miR-155 in peripheral blood mononuclear cells infected with HCV genotype 4. *FEBS open Bio*, 2 (1), 179-186

Elton, T.S., & Yalowich, J.C. (2015) Experimental procedures to identify and validate specific mRNA targets of miRNAs. *EXCLI Journal*, 14 (5), 758-790

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. *SAMJ: South African Medical Journal*, 102 (11), 841-844

Etheridge, A., Lee, I., Hood, L., Galas, D., & Wang, K. (2011) Extracellular microRNA: a new source of biomarkers. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 717 (1), 85-90

Fanelli, C.G., Porcellati, F., Rossetti, P., & Bolli, G.B. (2006) Glucagon: the effects of its excess and deficiency on insulin action. *Nutrition, Metabolism & Cardiovascular Diseases*, 16 (Suppl 1), S28-S34

Feero, W.G., Guttmacher, A.E., & McCarthy, M.I. (2010) Genomics, type 2 diabetes, and obesity. *New England Journal of Medicine*, 363 (24), 2339-2350

Felekkis, K., Touvana, E., Stefanou, C., & Deltas, C. (2010) microRNAs: a newly described class of encoded molecules that play a role in health and disease. *Hippokratia*, 14 (4), 236-240

Fernandez-Valverde, S.L., Taft, R.J., & Mattick, J.S. (2011) MicroRNAs in β-cell biology, insulin resistance, diabetes and its complications. *Diabetes*, 60 (7), 1825-1831

Fleige, S., & Pfaffl, M.W. (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine*, 27 (2), 126-139

Forouhi, N.G., Balkau, B., Borch-Johnsen, K., Dekker, J., Glumer, C., Qiao, Q., Spijkerman, A., Stolk, R., Tabac, A., & Wareham, N.J. (2006) The threshold for diagnosing impaired fasting glucose: a position statement by the European Diabetes Epidemiology Group. *Diabetologia*, 49 (5), 822-827

Fowler, M.J. (2008) Microvascular and macrovascular complications of diabetes. *Clinical Diabetes*, 26 (2), 77-82

Friedlander, M.R., Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., & Rajewsky, N. (2008) Discovering microRNAs from deep sequencing data using miRDeep. *Nature Biotechnology*, 26 (4), 407-415

Friedlander, M.R., Lizano, E., Houben, A.J., Bezdan, D., Banez-Coronel, M., Kudla, G., Mateu-Huertas, E., Kagerbauer, B., Gonzalez, J., & Chen, K.C. (2014) Evidence for the biogenesis of more than 1,000 novel human microRNAs. *Genome Biol*, 15 (4), 57-62

Friedman, R.C., Farh, K.K.-H., Burge, C.B., & Bartel, D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 19 (1), 92-105

Frosig, C., Rose, A.J., Treebak, J.T., Kiens, B., Richter, E.A., & Wojtaszewski, J.F. (2007) Effects of endurance exercise training on insulin signaling in human skeletal muscle interactions at the level of phosphatidylinositol 3-kinase, Akt, and AS160. *Diabetes*, 56 (8), 2093-2102

Fu, H.J., Zhu, J., Yang, M., Zhang, Z.Y., Tie, Y., Jiang, H., Sun, Z.X., & Zheng, X.F. (2006) A novel method to monitor the expression of microRNAs. *Molecular Biotechnology*, 32 (3), 197-204

Gabir, M.M., Hanson, R.L., Dabelea, D., Imperatore, G.I.U.S., Roumain, J.A.N.I., Bennett, P.H., & Knowler, W.C. (2000). The 1997 American Diabetes Association and 1999 World Health Organisation criteria for hyperglycemia in the diagnosis and prediction of diabetes. *Diabetes Care*, 23 (8), 1108-1112

Galazis, N., Afxentiou, T., Xenophontos, M., Diamanti-Kandarakis, E., & Atiomo, W. (2013) Proteomic biomarkers of type 2 diabetes mellitus risk in women with polycystic ovary syndrome. *European Journal of Endocrinology*, 168 (2), 33-43

Garber, A., Handelsman, Y., Einhorn, D., Bergman, D., Bloomgarden, Z., Fonseca, V., Timothy Garvey, W., Gavin III, J., Grunberger, G., & Horton, E. (2008) Diagnosis and management of prediabetes in the continuum of hyperglycemia: when do the risks of diabetes begin? A consensus statement from the American College of Endocrinology and the American Association of Clinical Endocrinologists. *Endocrine Practice*, 14 (7), 933-946

George, J.A. (2011) Should haemoglobin A1c be used for the diagnosis of diabetes mellitus in South Africa? *Journal of Endocrinology, Metabolism & Diabetes of South Africa*, 16 (3), 122-127

Ghosh, S. (2011) Micro RNA-Biogenesis, Mechanism of Action and Applications-A Review. *Journal of Biotechnology & Biochemistry,* 1 (1), 11-36

Gibney, E.R., & Nolan, C.M. (2010) Epigenetics and gene expression. *Heredity*, 105 (1), 4-13

Gilad, S., Meiri, E., Yogev, Y., Benjamin, S., Lebanony, D., Yerushalmi, N., Benjamin, H., Kushnir, M., & Cholakh, H. (2008) Serum microRNAs are promising novel biomarkers. *PloS One,* 3 (9), e3148

Git, A., Dvinge, H., Salmon-Divon, M., Osborne, M., Kutter, C., Hadfield, J., Bertone, P., & Caldas, C. (2010) Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA*, 16 (5), 991-1006

Goedecke, J.H., Jennings, C., & Lambertc, E.V. (2006) In South Africa. [Cited 20 November 2015]; Available from: http://www.mrc.ac.za/chronic/cdl1995-2005.htm

Goto, Y., Kojima, S., Nishikawa, R., Enokida, H., Chiyomaru, T., Kinoshita, T., Nakagawa, M., Naya, Y., Ichikawa, T., & Seki, N. (2014) The microRNA-23b/27b/24-1 cluster is a disease progression marker and tumor suppressor in prostate cancer. *Oncotarget*, 5 (17), 7748-7759

Granjon, A., Gustin, M.P., Rieusset, J., Lefai, E., Meugnier, E., Guller, I., Cerutti, C., Paultre, C., Disse, E., & Rabasa-Lhoret, R. (2009) The microRNA Signature in Response to Insulin Reveals Its Implication in the Transcriptional Action of Insulin in Human Skeletal Muscle and

the Role of a Sterol Regulatory Element-Binding Protein-1c/Myocyte Enhancer Factor 2C Pathway. *Diabetes*, 58 (11), 2555-2564

Greenberg, A.S., & Obin, M.S. (2006) Obesity and the role of adipose tissue in inflammation and metabolism. *The American Journal of Clinical Nutrition*, 83 (Suppl 2), 461S-465S

Griffiths-Jones, S., Grocock, R.J., Van Dongen, S., Bateman, A., & Enright, A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, 34 (suppl 1), D140-D144

Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P., & Bartel, D.P. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Molecular Cell*, 27 (1), 91-105

Guay, C., Roggli, E., Nesca, V., Jacovetti, C., & Regazzi, R. (2011) Diabetes mellitus, a microRNA related disease? *Transl Res,*157 (4), 253-264

Gunderson, E.P., Chiang, V., Pletcher, M.J., Jacobs, D.R., Quesenberry, C.P., Sidney, S., & Lewis, C.E. (2014) History of Gestational Diabetes Mellitus and Future Risk of Atherosclerosis in Mid-life: The Coronary Artery Risk Development in Young Adults Study. *Journal of the American Heart Association*, 3 (2), e000490

Hamilton, J.P. (2011) Epigenetics: principles and practice. *Digestive Diseases*, 29 (2), 130-135

Han, T.S., Feskens, E.J.M., Lean, M.E.J., & Seidell, J.C. (1998) Associations of body composition with type 2 diabetes mellitus. *Diabetic Medicine*, 15 (2), 129-135

Hanson, E.K., Lubenow, H., & Ballantyne, J. (2009) Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. *Analytical Biochemistry*, 387 (2), 303-314

He, L., & Hannon, G.J. (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nature Reviews Genetics*, 5 (7), 522-531

Herrera, B.M., Lockstone, H.E., Taylor, J.M., Ria, M., Barrett, A., Collins, S., Kaisaki, P., Argoud, K., Fernandez, C., & Travers, M.E. (2010) Global microRNA expression profiles in

insulin target tissues in a spontaneous rat model of type 2 diabetes. *Diabetologia*, 53 (6), 1099-1109

Herrera, B.M., Lockstone, H.E., Taylor, J.M., Wills, Q.F., Kaisaki, P.J., Barrett, A., Camps, C., Fernandez, C., Ragoussis, J., & Gauguier, D. (2009) MicroRNA-125a is over-expressed in insulin target tissues in a spontaneous rat model of Type 2 Diabetes. *BMC Medical Genomics*, 2 (1), 54-62

Hevener, A.L., He, W., Barak, Y., Le, J., Bandyopadhyay, G., Olson, P., Wilkes, J., Evans, R.M., & Olefsky, J. (2003) Muscle-specific PPARG deletion causes insulin resistance. *Nature Medicine*, 9 (12), 1491-1497

Hilton, C., Neville, M.J., & Karpe, F. (2013) MicroRNAs in adipose tissue: their role in adipogenesis and obesity. *International Journal of Obesity*, 37 (3), 325-332

Hirst, M., & Marra, M.A. (2009) Epigenetics and human disease. *The International Journal of Biochemistry & Cell Biology*, 41 (1), 136-146

HIV and AIDS in South Africa. (2014) [Cited 23 February 2015]; Available from: http://www.avert.org/professionals/hiv-around-world/sub-saharan-africa/overview

Holland, N.T., Smith, M.T., Eskenazi, B., & Bastaki, M. (2003) Biological sample collection and processing for molecular epidemiological studies. *Mutation Research/Reviews in Mutation Research*, 543 (3), 217-234

Hsu, S.D., Tseng, Y.T., Shrestha, S., Lin, Y.L., Khaleel, A., Chou, C.H., Chu, C.F., Huang, H.Y., Lin, C.M., & Ho, S.Y. 2014. miRTarBase update (2014) an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Research*, 42 (Suppl D1), D78-D85

Hu, F.B. (2011) Globalization of Diabetes The role of diet, lifestyle, and genes. *Diabetes Care*, 34 (6), 1249-1257

Huang, D.W., Sherman, B.T., & Lempicki, R.A. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4 (1), 44-57

Huang, D.W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., & Lane, H.C. (2007) DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Research*, 35 (suppl 2), W169-W175

Hunter, M.P., Ismail, N., Zhang, X., Aguda, B.D., Lee, E.J., Yu, L., Xiao, T., Schafer, J., Lee, M.L.T., & Schmittgen, T.D. (2008) Detection of microRNA expression in human peripheral blood microvesicles. *PloS One*, 3 (11), e3694

Hydbring, P., & Badalian-Very, G. (2013) Clinical applications of microRNAs. *F1000 Research*, 3 (2), 132

International Diabetes Federation. (2013) [Cited 24 February 2014]; Available from: http://www.idf.org/diabetesatlas

International Expert Committee. (2009) International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. *Diabetes Care*, 32 (7), 1327-1334

Issler, O., & Chen, A. (2015) Determining the role of microRNAs in psychiatric disorders. *Nature Reviews Neuroscience*, 16 (4), 201-212

Jennewein, C., von Knethen, A., Schmid, T., & Brune, B. (2010) MicroRNA-27b contributes to lipopolysaccharide-mediated peroxisome proliferator-activated receptor gamma (PPARG) mRNA destabilization. *Journal of Biological Chemistry*, 285 (16), 11846-11853

Juan, L., Tong, H.l., Zhang, P., Guo, G., Wang, Z., Wen, X., Dong, Z., & Tian, Y.P. (2014) Identification and characterization of novel serum microRNA candidates from deep sequencing in cervical cancer patients. *Scientific Reports*, 4 (6277), 1-9

Kahn, B.B. (1998) Type 2 diabetes: when insulin secretion fails to compensate for insulin resistance. *Cell*, 92 (5), 593-596

Kahn, S.E. (2003) The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of type 2 diabetes. *Diabetologia*, 46 (1), 3-19

Kahn, S.E., Cooper, M.E., & Del Prato, S. (2014) Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. *The Lancet*, 383 (9922), 1068-1083

Kaiser, N., Leibowitz, G., & Nesher, R. (2003) Glucotoxicity and beta-cell failure in type 2 diabetes mellitus. *Journal of Pediatric Endocrinology and Metabolism*, 16 (1), 5-22

Karbiener, M., Fischer, C., Nowitsch, S., Opriessnig, P., Papak, C., Ailhaud, G., Dani, C., Amri, E.Z., & Scheideler, M. (2009) MicroRNA miR-27b impairs human adipocyte differentiation and targets PPARG. *Biochemical and Biophysical Research Communications*, 390 (2), 247-251

Karolina, D.S., Armugam, A., Tavintharan, S., Wong, M.T., Lim, S.C., Sum, C.F., & Jeyaseelan, K. (2011) MicroRNA 144 impairs insulin signaling by inhibiting the expression of insulin receptor substrate 1 in type 2 diabetes mellitus. *PloS One*, 6 (8), e22839

Kawamori, D., Kurpad, A.J., Hu, J., Liew, C.W., Shih, J.L., Ford, E.L., Herrera, P.L., Polonsky, K.S., McGuinness, O.P., & Kulkarni, R.N. (2009) Insulin signaling in α-cells modulates glucagon secretion in vivo. *Cell Metabolism*, 9 (4), 350-361

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., & Segal, E. (2007) The role of site accessibility in microRNA target recognition. *Nature Genetics*, 39 (10), 1278-1284

Kilpatrick, E.S., & Winocour, P.H. (2010) ABCD position statement on haemoglobin A1c for the diagnosis of diabetes. *Practical Diabetes International*, 27 (7), 306-310

Kinet, V., Halkein, J., Dirkx, E., & De Windt, L.J. (2013) Cardiovascular extracellular microRNAs: emerging diagnostic markers and mechanisms of cell-to-cell RNA communication. *Frontiers in Genetics*, 4, 214

Kiriakidou, M., Nelson, P.T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z., & Hatzigeorgiou, A. (2004) A combined computational-experimental approach predicts human microRNA targets. *Genes & Development*, 18 (10), 1165-1178

Kirigia, J.M., Sambo, H.B., Sambo, L.G., & Barry, S.P. (2009) Economic burden of diabetes mellitus in the WHO African region. *BMC International Health & Human Rights*, 9 (1), 6

Kirkman, M.S., & Kendall, D.M. (2011) Hemoglobin A1c to diagnose diabetes: why the controversy over adding a new tool? *Clinical Chemistry*, 57 (2), 255-257

Klein, D. (2002) Quantification using real-time PCR technology: applications and limitations. *Trends in Molecular Medicine*, 8 (6), 257-260

Kong, L., Zhu, J., Han, W., Jiang, X., Xu, M., Zhao, Y., Dong, Q., Pang, Z., Guan, Q., & Gao, L. (2011) Significance of serum microRNAs in pre-diabetes and newly diagnosed type 2 diabetes: a clinical study. *Acta Diabetologica*, 48 (1), 61-69

Kong, X., Yu, J., Bi, J., Qi, H., Di, W., Wu, L., Wang, L., Zha, J., Lv, S., & Zhang, F. (2014) Glucocorticoids transcriptionally regulate miR-27b expression promoting body fat accumulation via suppressing the browning of white adipose tissue. *Diabetes*, 64 (2), 393-404

Kosaka, N., Iguchi, H., Yoshioka, Y., Takeshita, F., Matsuki, Y., & Ochiya, T. (2010) Secretory mechanisms and intercellular transfer of microRNAs in living cells. *Journal of Biological Chemistry*, 285 (23), 17442-17452

Krauss, R.M. (2004) Lipids and lipoproteins in patients with type 2 diabetes. *Diabetes Care*, 27 (6), 1496-1504

Kroh, E.M., Parkin, R.K., Mitchell, P.S., & Tewari, M. (2010) Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods*, 50 (4), 298-301

Kuhn, D.E., Martin, M.M., Feldman, D.S., Terry, A.V., Nuovo, G.J., & Elton, T.S. (2008) Experimental validation of miRNA targets. *Methods*, 44 (1), 47-54

Kumar, A., & Khanna, R. C. (2011) Biomarkers: Its Novel Application. *The Pharma Journal*, 1 (1), 22-28

Laakso, M. (2010) Cardiovascular Disease in Type 2 Diabetes From Population to Man to Mechanisms The Kelly West Award Lecture 2008. *Diabetes Care*, 33 (2), 442-449

Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., & Tuschl, T. (2002) Identification of tissue-specific microRNAs from mouse. *Current Biology*, 12 (9), 735-739

Laterza, O.F., Lim, L., Garrett-Engele, P.W., Vlasakova, K., Muniappa, N., Tanaka, W.K., Johnson, J.M., Sina, J.F., Fare, T.L., & Sistare, F.D. (2009) Plasma MicroRNAs as sensitive and specific biomarkers of tissue injury. *Clinical Chemistry*, 55 (11), 1977-1983

Lawrie, C.H., Gal, S., Dunlop, H.M., Pushkaran, B., Liggins, A.P., Pulford, K., Banham, A.H., Pezzella, F., Boultwood, J., & Wainscoat, J.S. (2008) Detection of elevated levels of tumour associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *British Journal of Haematology*, 141 (5), 672-675

Leahy, J.L. (2005) Pathogenesis of type 2 diabetes mellitus. *Archives of Medical Research*, 36 (3), 197-209

Lee, J.J., Drakaki, A., Iliopoulos, D., & Struhl, K. (2012) MiR-27b targets PPARG to inhibit growth, tumor progression and the inflammatory response in neuroblastoma cells. *Oncogene*, 31 (33), 3818-3825

Lee, R.C., Feinbaum, R.L., & Ambros, V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75 (5), 843-854

Lehmann, T.P., Korski, K., Gryczka, R., Ibbs, M., Thieleman, A., Grodecka-Gazdecka, S., & Jagodzinski, P.P. (2015) Relative levels of let-7a, miR-17, miR-27b, miR-125a, miR-125b and miR-206 as potential molecular markers to evaluate grade, receptor status and molecular type in breast cancer. *Molecular Medicine Reports*, 12 (3), 4692-4702

Leonardini, A., Laviola, L., Perrini, S., Natalicchio, A., & Giorgino, F. (2009) Cross-Talk between PPAR and Insulin Signaling and Modulation of Insulin Sensitivity. *PPAR Research*; doi: 10.1155/2009/818945

Levitt, N.S., Katzenellenbogen, J.M., Bradshaw, D., Hoffman, M.N., & Bonnici, F. (1993) The prevalence and identification of risk factors for NIDDM in urban Africans in Cape Town, South Africa. *Diabetes Care*, 16 (4), 601-607

Lewis, B.P., Burge, C.B., & Bartel, D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120 (1), 15-20

Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., & Burge, C.B. (2003) Prediction of mammalian microRNA targets. *Cell*, 115 (7), 787-798

Li, C., & Zhang, B.B. (2007) Insulin signaling and action: glucose, lipids, protein. [Cited 13 October 2015]; Available from: http://www.endotext.org/Diabetes/diabetes4/diabetesframe4.htm.

Li, T., Cao, H., Zhuang, J., Wan, J., Guan, M., Yu, B., Li, X., & Zhang, W. (2011) Identification of miR-130a, miR-27b and miR-210 as serum biomarkers for atherosclerosis obliterans. *Clinica Chimica Acta*, 412 (1), 66-70

Li, Y., Qiu, C., Tu, J., Geng, B., Yang, J., Jiang, T., & Cui, Q. (2014) HMDD v2. 0: a database for experimentally supported human microRNA and disease associations. [Cited 14 August 2015]; Available from: http://www.cuilab.cn/hmdd

Liang, H., & Li, W.H. (2007) MicroRNA regulation of human protein-protein interaction network. *RNA*, 13 (9), 1402-1408

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., & Johnson, J.M. (1999) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Mol.Cell.Biol*, 19 (7), 6379-6395

Lin, Q., Gao, Z., Alarcon, R.M., Ye, J., & Yun, Z. (2009a) A role of miR-27 in the regulation of adipogenesis. *Febs Journal*, 276 (8), 2348-2358

Lin, T., Dong, W., Huang, J., Pan, Q., Fan, X., Zhang, C., & Huang, L. (2009b) MicroRNA-143 as a tumor suppressor for bladder cancer. *The Journal of urology*, 181 (3), 1372-1380

Link, A., Becker, V., Goel, A., Wex, T., & Malfertheiner, P. (2012) Feasibility of fecal microRNAs as novel biomarkers for pancreatic cancer. *PLoS One*, 7 (8), e42933

Liu, C., Xin, N., Zhai, Y., Jiang, L., Zhai, J., Zhang, Q., & Qi, J. (2014) Reference gene selection for quantitative real-time RT-PCR normalization in the half-smooth tongue sole (Cynoglossus semilaevis) at different developmental stages, in various tissue types and on exposure to chemicals. *PLoS One*, 9 (3), e91715

Liu, R., Wang, X., Aihara, K., & Chen, L. (2014) Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers. *Medicinal Research Reviews*, 34 (3), 455-478

Livak, K.J., & Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta}C_T$ method. *Methods*, 25 (4), 402-408

Lopez, J.P., Cruceanu, C., Fiori, L.M., Laboissiere, S., Guillet, I., Fontaine, J., Ragoussis, J., Benes, V., Turecki, G., & Ernst, C. (2015) Biomarker discovery: quantification of microRNAs and other small non-coding RNAs using next generation sequencing. *BMC Medical Genomics*, 8 (1), 35

Luo, M., Li, R., Deng, X., Ren, M., Chen, N., Zeng, M., Yan, K., Xia, J., Liu, F., & Ma, W. (2015) Platelet-derived miR-103b as a novel biomarker for the early diagnosis of type 2 diabetes. *Acta diabetologica*, 52 (5), 943-949

Lyons, T.J., & Basu, A. (2012) Biomarkers in diabetes: hemoglobin A1c, vascular and tissue markers. *Translational Research*, 159 (4), 303-312

Maedler, K., (2008) Beta cells in type 2 diabetes: a crucial contribution to pathogenesis. *Diabetes, Obesity & Metabolism*, 10 (5), 408-420

Maes, O.C., Chertkow, H.M., Wang, E., & Schipper, H.M. (2009) MicroRNA: implications for Alzheimer disease and other human CNS disorders. *Current Genomics*, 10 (3), 154-168

Malkani, S., & DeSilva, T. (2012) Controversies on how diabetes is diagnosed. *Current Opinion in Endocrinology, Diabetes & Obesity*, 19 (2), 97-103

Mao, Y., Mohan, R., Zhang, S., & Tang, X. (2013) MicroRNAs as pharmacological targets in diabetes. *Pharmacological Research*, 75 (5), 37-47

Marabita, F., de Candia, P., Torri, A., Tegnér, J., Abrignani, S., & Rossi, R.L. (2015) Normalization of circulating microRNA expression data obtained by quantitative real-time RT-PCR. *Briefings in Bioinformatics:* doi: 10.1093/bib/bbv056

Maragkakis, M., Reczko, M., Simossis, V.A., Alexiou, P., Papadopoulos, G.L., Dalamagas, T., Giannopoulos, G., Goumas, G., Koukis, E., & Kourtis, K. (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. [Cited 21 July 2015]; Available from: http://diana.imis.athena-innovation.gr/DianaTools/index.php

Mariam, A. (2007) Quality assessment of total RNA. *Biomedical Genomics*. [Cited 20 March 2015]; Available from: http://www.biomedicalgenomics.org/How_does_Agilent_2100_Bioanalyzer_work.html

Marks, J.S., & Botelho, L.H. (1986) Synergistic inhibition of glucagon-induced effects on hepatic glucose metabolism in the presence of insulin and a cAMP antagonist. *Journal of Biological Chemistry*, 261 (34), 15895-15899

Mayeux, R. (2004) Biomarkers: potential uses and limitations. *NeuroRx*, 1 (2), 182-188

McCarthy, M., & Menzel, S. (2001) The genetics of type 2 diabetes. *British Journal of Clinical Pharmacology*, 51 (3), 195-199

McKillop, A.M., & Flatt, P.R. (2011) Emerging applications of metabolomic and genomic profiling in diabetic clinical medicine. *Diabetes Care*, 34 (12), 2624-2630

Meier, J.J., & Bonadonna, R.C. (2013) Role of reduced beta-cell mass versus impaired beta-cell function in the pathogenesis of type 2 diabetes. *Diabetes Care*, 36 (Suppl 2), S113-S119

Meyer, S.U., Pfaffl, M.W., & Ulbrich, S.E. (2010) Normalization strategies for microRNA profiling experiments: a normal way to a hidden layer of complexity? *Biotechnology Letters*, 32 (12), 1777-1788

Miah, S., Dudziec, E., Drayton, R.M., Zlotta, A.R., Morgan, S.L., Rosario, D.J., Hamdy, F.C., & Catto, J.W.F. (2012) An evaluation of urinary microRNA reveals a high sensitivity for bladder cancer. *British journal of Cancer*, 107 (1), 123-128

Min, H., & Yoon, S. (2010) Got target?: computational methods for microRNA target prediction and their extension. *Experimental and Molecular Medicine*, 42 (4), 233-244

Misra, A., Khurana, L., Isharwal, S., & Bhardwaj, S. (2009) South Asian diets and insulin resistance. *British Journal of Nutrition*, 101 (4), 465-473

Mitchell, P.S., Parkin, R.K., Kroh, E.M., Fritz, B.R., Wyman, S.K., Pogosova-Agadjanyan, E.L., Peterson, A., Noteboom, J., O'Briant, K.C., & Allen, A. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proceedings of the National Academy of Sciences*, 105 (30), 10513-10518

Moldovan, L., Batte, K.E., Trgovcich, J., Wisler, J., Marsh, C.B., & Piper, M. (2014) Methodological challenges in utilizing miRNAs as circulating biomarkers. *Journal of cellular and molecular medicine*, 18 (3), 371-390

Molleutze, L. (2006) Diabetes mellitus and impaired glucose tolerance. [Cited 26 September 2015]; Available From: http://www.mrc.ac.za/chronic/cdl1995-2005.htm.

Moon, J.S., & Won, K.C. (2015) Pancreatic Beta-Cell Dysfunction in Type 2 Diabetes: Old Kids on the Block. *Diabetes & Metabolism Journal*, 39 (1), 1-9

Mooradian, A.D. (2009) Dyslipidemia in type 2 diabetes mellitus. *Nature Cinical Practice Endocrinology & Metabolism*, 5 (3), 150-159

Morgan, N.G. (2009) Fatty acids and beta-cell toxicity. *Current Opinion in Clinical Nutrition & Metabolic Care*, 12 (2), 117-122

Motala, A.A., Esterhuizen, T., Gouws, E., Pirie, F.J., & Omar, M.A. (2008) Diabetes and other disorders of glycemia in a rural South African community prevalence and associated risk factors. *Diabetes Care*, 31 (9), 1783-1788

Motameny, S., Wolters, S., Nurnberg, P., & Schumacher, B. (2010) Next generation sequencing of miRNAs - Strategies, Resources and Methods. *Genes*, 1 (1), 70-84

Mueller, O., Lightfoot, S., & Schroeder, A. (2004) RNA integrity number (RIN)-standardization of RNA quality control. *Agilent Application Note,* 1-8

Muhonen, P., & Holthofer, H. (2009) Epigenetic and microRNA-mediated regulation in diabetes. *Nephrology Dialysis Transplantation*, 24 (4), 1088-1096

Ni, Y., Meng, L., Wang, L., Dong, W., Shen, H., Wang, G., Liu, Q., & Du, J. (2013) MicroRNA-143 functions as a tumor suppressor in human esophageal squamous cell carcinoma. *Gene*, 517 (2), 197-204

Nielsen, L.B., Wang, C., Sørensen, K., Bang-Berthelsen, C.H., Hansen, L., Andersen, M.L., Hougaard, P., Juul, A., Zhang, C.Y., & Pociot, F. (2012) Circulating levels of microRNA from children with newly diagnosed type 1 diabetes and healthy controls: evidence that miR-25 associates to residual beta-cell function and glycaemic control during disease progression. *Experimental Diabetes Research:* doi: 10.1155/2012/896362

O'Connell, R.M., Kahn, D., Gibson, W.S., Round, J.L., Scholz, R.L., Chaudhuri, A.A., Kahn, M.E., Rao, D.S., & Baltimore, D. (2010) MicroRNA-155 promotes autoimmune inflammation by enhancing inflammatory T cell development. *Immunity*, 33 (4), 607-619

Odermatt, A. (2011) The Western-style diet: a major risk factor for impaired kidney function and chronic kidney disease. *American Journal of Physiology-Renal Physiology*, 301 (5), 919-931

Oliveros, J.C. (2007) VENNY. An interactive tool for comparing list with Venn Diagrams [Cited 22 July 2015]; Available from: http://bioinfogp.cnb.csic.es/tools/venny_old/venny.php

Olson, D.E., Rhee, M.K., Herrick, K., Ziemer, D.C., Twombly, J.G., & Phillips, L.S. (2010) Screening for diabetes and pre-diabetes with proposed A1C-based diagnostic criteria. *Diabetes Care*, 33 (10), 2184-2189

Olson, E.N. (2014) MicroRNAs as therapeutic targets and biomarkers of cardiovascular disease. *Science Translational Medicine*, 6 (239), 239-242

Ortega, F.J., Mercader, J.M., Moreno-Navarrete, J.M.a., Rovira, O., Guerra, E., Esteve, E., Xifra, G., Martenez, C., Ricart, W., & Rieusset, J. (2014) Profiling of circulating microRNAs reveals common microRNAs linked to type 2 diabetes that change with insulin sensitization. *Diabetes Care*, 37 (5), 1375-1383

Ostbye, T., Welby, T.J., Prior, I.A.M., Salmond, C.E., & Stokes, Y.M. (1989) Type 2 (non-insulin-dependent) diabetes mellitus, migration and westernisation: the Tokelau Island Migrant Study. *Diabetologia*, 32 (8), 585-590

Oulas, A., Karathanasis, N., Louloupi, A., Pavlopoulos, G. A., Poirazi, P., Kalantidis, K., & Iliopoulos, I. (2015) Prediction of miRNA targets. *RNA Bioinformatics*, 1268, (5) 207-229

Pagliuca, A., Valvo, C., Fabrizi, E., di Martino, S., Biffoni, M., Runci, D., Forte, S., De Maria, R., & Ricci-Vitiani, L. (2013) Analysis of the combined action of miR-143 and miR-145 on oncogenic pathways in colorectal cancer cells reveals a coordinate program of gene repression. *Oncogene*, 32 (40), 4806-4813

Palmer, J.P., Fleming, G.A., Greenbaum, C.J., Herold, K.C., Jansa, L.D., Kolb, H., Lachin, J.M., Polonsky, K.S., Pozzilli, P., & Skyler, J.S. (2004) C-peptide is the appropriate outcome measure for type 1 diabetes clinical trials to preserve β-cell. *Diabetes*, 53 (1), 250-264

Pandey, A., Agarwal, P., Kaur, K., & Datta, M. (2009) MicroRNAs in Diabetes: Tiny Players in Big. *Cell Physiol Biochem*, 23 (4-6), 221-232

Parrizas, M., Brugnara, L., Esteban, Y., Gonzollez-Franquesa, A., Canivell, S., Murillo, S., Gordillo-Bastidas, E., Cusso, R., Cadefau, J.A., & Garcia-Roves, P.M. (2014) Circulating miR-192 and miR-193b are markers of prediabetes and are modulated by an exercise intervention. *The Journal of Clinical Endocrinology & Metabolism,* 100 (3), 149-153

Pescador, N., Perez-Barba, M., Ibarra, J.M., Corbaton, A., Martinez-Larrad, M.T., & Serrano-Roos, M. (2013) Serum circulating microRNA profiling for identification of potential type 2 diabetes and obesity biomarkers. *PLoS One*, 8 (10), e77251

Peterson, S.M., Thompson, J.A., Ufkin, M.L., Sathyanarayana, P., Liaw, L., & Congdon, C.B. (2014) Common features of microRNA target prediction tools. *Frontiers in Genetics*, 18 (5), e23

Pinney, S.E., & Simmons, R.A. (2010) Epigenetic mechanisms in the development of type 2 diabetes. *Trends in Endocrinology & Metabolism*, 21 (4), 223-229

Ponchel, F., Toomes, C., Bransfield, K., Leong, F.T., Douglas, S.H., Field, S.L., Bell, S.M., Combaret, V., Puisieux, A., & Mighell, A.J. (2003) Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *BMC Biotechnology*, 3 (1), 18

Poornima, I.G., Parikh, P., & Shannon, R.P. (2006) Diabetic cardiomyopathy the search for a unifying hypothesis. *Circulation Research*, 98 (5), 596-605

Popkin, B.M., & Gordon-Larsen, P. (2004) The nutrition transition: worldwide obesity dynamics and their determinants. *International Journal of Obesity*, 28 (Suppl), S2-S9

Popkin, B.M. (1999) Urbanization, lifestyle changes and the nutrition transition. *World Development*, 27 (11), 1905-1916

Pradhan, A.D., Manson, J.E., Rifai, N., Buring, J.E., & Ridker, P.M. (2001) C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *Jama*, 286 (3), 327-334

Pritchard, C.C., Cheng, H.H., & Tewari, M. (2012) MicroRNA profiling: approaches and considerations. *Nature Reviews Genetics*, 13 (5), 358-369

Puoane, T., Steyn, K., Bradshaw, D., Laubscher, R., Fourie, J., Lambert, V., & Mbananga, N. (2002) Obesity in South Africa: the South African demographic and health survey. *Obesity Research*, 10 (10), 1038-1048

Purrello, F., & Rabuazzo, A.M. (2000) Metabolic factors that affect beta-cell function and survival. *Diabetes, Nutrition & Metabolism*, 13 (2), 84-91

Qi, L. (2014) Personalized nutrition and obesity. *Annals of Medicine*, 46 (5), 247-252

Raffort, J., Hinault, C., Dumortier, O., & Van Obberghen, E. (2015) Circulating microRNAs and diabetes: potential applications in medical practice. *Diabetologia*, 58 (9), 1978-1992

Rajeevan, M.S., Ranamukhaarachchi, D.G., Vernon, S.D., & Unger, E.R. (2001) Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. *Methods*, 25 (4), 443-451

Ratner, R.E. (2007) Prevention of type 2 diabetes in women with previous gestational diabetes. *Diabetes Care*, 30 (Suppl 2), S242-S245

Reddy, M. A., Tak Park, J., & Natarajan, R. (2013) Epigenetic Modifications in the Pathogenesis of Diabetic Nephropathy. *Semin Nephrol,* 33 (*4*), *341–353*

Robertson, R.P., Harmon, J., Tran, P.O., Tanaka, Y., & Takahashi, H. (2003) Glucose toxicity in beta-cells: type 2 diabetes, good radicals gone bad, and the glutathione connection. *Diabetes*, 52 (3), 581-587

Rong, Y., Bao, W., Shan, Z., Liu, J., Yu, X., Xia, S., Gao, H., Wang, X., Yao, P., & Hu, F.B. (2013) Increased microRNA-146a levels in plasma of patients with newly diagnosed type 2 diabetes mellitus. PLoS One, 8 (9), e73272

Rottiers, V., & Näär, A.M. (2012) MicroRNAs in metabolism and metabolic disorders. *Nature Reviews Molecular Cell Biology*, 13 (4), 239-250

Saal, S., & Harvey, S.J. (2009) MicroRNAs and the kidney: coming of age. *Current Opinion in Nephrology & Hypertension*, 18 (4), 317-323

Sacks, D.B., Arnold, M., Bakris, G.L., Bruns, D.E., Horvath, A.R., Kirkman, M.S., Lernmark, A., Metzger, B.E., & Nathan, D.M. (2011) Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Diabetes Care*, 34 (6), e61-e99

Sahu, P., Pinkalwar, N., Dubey, R.D., Paroha, S., Chatterjee, S., & Chatterjee, T. (2011) Biomarkers: an emerging tool for diagnosis of a disease and drug development. *Asian Journal of Research in Pharmaceutical Science*, 1 (1), 9-16

Salmasi, A.M. & Dancy, M. (2005) The Glucose Tolerance Test, But Not HbA1c, Remains the Gold Standard in Identifying Unrecognized Diabetes Mellitus and Impaired Glucose Tolerance in Hypertensive Subjects. *Angiology*, 56 (5), 571-579

Saltiel, A.R., & Kahn, C.R. (2001) Insulin signalling and the regulation of glucose and lipid metabolism. *Nature*, 414 (6865), 799-806

Santovito, D., De Nardis, V., Marcantonio, P., Mandolini, C., Paganelli, C., Vitale, E., Buttitta, F., Bucci, M., Mezzetti, A., & Consoli, A. (2014) Plasma exosome microRNA profiling unravels a new potential modulator of adiponectin pathway in diabetes: effect of glycemic control. *The Journal of Clinical Endocrinology & Metabolism*, 99 (9), 1681-1685

Saudek, C.D., Herman, W.H., Sacks, D.B., Bergenstal, R.M., Edelman, D., & Davidson, M.B. (2008) A new look at screening and diagnosing diabetes mellitus. *The Journal of Clinical Endocrinology & Metabolism*, 93 (7), 2447-2453

Scheen, A.J. (2003) Pathophysiology of type 2 diabetes. *Acta Clinica Belgica*, 58 (6), 335-341

Schwarzenbach, H., Nishida, N., Calin, G.A., & Pantel, K. (2014) Clinical relevance of circulating cell-free microRNAs in cancer. *Nature reviews Clinical Oncology*, 11 (3), 145-156

Selbach, M., Schwanhñusser, B.r., Thierfelder, N., Fang, Z., Khanin, R., & Rajewsky, N. (2008) Widespread changes in protein synthesis induced by microRNAs. *Nature*, 455 (7209), 58-63

Selvin, E., Crainiceanu, C.M., Brancati, F.L., & Coresh, J. (2007) Short-term variability in measures of glycemia and implications for the classification of diabetes. *Archives of Internal Medicine*, 167 (14), 1545-1551

Shamai, L., Lurix, E., Shen, M., Novaro, G.M., Szomstein, S., Rosenthal, R., Hernandez, A.V., & Asher, C.R. (2011) Association of body mass index and lipid profiles: evaluation of a broad spectrum of body mass index patients including the morbidly obese. *Obesity Surgery*, 21 (1), 42-47

Shamoon, H., Duffy, H., Fleischer, N., Engel, S., Saenger, P., Strelzyn, M., Litwak, M., Wylierosett, J., Farkash, A., & Geiger, D. (1993) The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes-mellitus. *New England Journal of Medicine*, 329 (14), 977-986

Sherwin, R.S., Anderson, R.M., Buse, J.B., Chin, M.H., Eddy, D., Fradkin, J., Ganiats, T.G., Ginsberg, H.N., Kahn, R., & Nwankwo, R. (2004) Prevention or delay of type 2 diabetes. *Diabetes Care*, 27 (Suppl), S47

Shields, B.M., Hicks, S., Shepherd, M.H., Colclough, K., Hattersley, A.T., & Ellard, S. (2010) Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia*, 53 (12), 2504-2508

Shulman, G.I. (2000) Cellular mechanisms of insulin resistance. *Journal of Clinical Investigation*, 106 (2), 171-176

Shyong Tai, E., Goh, S.Y., Lee, J.J., Wong, M.S., Heng, D., Hughes, K., Suok, K.C., Cutter, J., Chew, W., & Gu, K. (2004) Lowering the criterion for impaired fasting glucose: Impact on disease prevalence and associated risk of diabetes and ischemic heart disease. *Diabetes care*, 27 (7), 1728-1734

Simon, A., & Evin, K. (2008) Luciferase reporter assays: powerful, adaptable tools for cell biology research. *Promega*, 1 (21), 9-16

Stefani, G., & Slack, F.J. (2008) Small non-coding RNAs in animal development. *Nature reviews Molecular Cell Biology*, 9 (3), 219-230

Stumvoll, M., Goldstein, B.J., & van Haeften, T.W. (2005) Type 2 diabetes: principles of pathogenesis and therapy. *The Lancet*, 365 (9467), 1333-1346

Sun, K., Chang, X., Yin, L., Li, J., Zhou, T., Zhang, C., & Chen, X. (2014a) Expression and DNA methylation status of microRNA-375 in patients with type 2 diabetes mellitus. *Molecular Medicine Reports*, 9 (3), 967-972

Sun, X., Yu, W., & Hu, C. (2014b) Genetics of type 2 diabetes: insights into the pathogenesis and its clinical application. *BioMed Research International*

Szczygielska, A., Widomska, S., Jaraszkiewicz, M., Knera, P., & Muc, K. (2002) Blood lipids profile in obese or overweight patients. *Medicina*, 58 (2), 343-349

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., & Tsafou, K.P. (2014) STRING v10: protein-protein interaction networks, integrated over the tree of life. [Cited 19 August 2015]; Available from: http://string-db.org

Taganov, K.D., Boldin, M.P., Chang, K.J., & Baltimore, D. (2006) NF-kB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proceedings of the National Academy of Sciences*, 103 (33), 12481-12486

Tam, S., de Borja, R., Tsao, M.S., & McPherson, J.D. (2014) Robust global microRNA expression profiling using next-generation sequencing technologies. *Laboratory Investigation*, 94 (3), 350-358

Tang, X., Muniappan, L., Tang, G., & Ozcan, S. (2009) Identification of glucose-regulated miRNAs from pancreatic beta cells reveals a role for miR-30d in insulin transcription. *RNA*, 15 (2), 287-293

Tang, X., Tang, G., & Őzcan, S. (2008) Role of microRNAs in diabetes. *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms*, 1779 (11), 697-701

Thery, C., Zitvogel, L., & Amigorena, S. (2002) Exosomes: composition, biogenesis and function. *Nature Reviews Immunology*, 2 (8), 569-579

Thomas, M. F., & Ansel, K. M. (2010) "Construction of small RNA cDNA libraries for deep sequencing. *Methods in Molecular Biology*, 667 (9), 93-111

Tiberio, P., Callari, M., Angeloni, V., Daidone, M.G., & Appierto, V. (2015) Challenges in using circulating miRNAs as cancer biomarkers. *BioMed Res Int*, 2015, 731479

Topic, E. (2014) Guidelines and recommendations for testing in diagnosis of diabetes mellitus: The role of HbA1c. *Biochemia Medica*, 24 (Suppl 1), S1-S78

Trajkovic, K., Hsu, C., Chiantia, S., Rajendran, L., Wenzel, D., Wieland, F., Schwille, P., Brűgger, B., & Simons, M. (2008) Ceramide triggers budding of exosome vesicles into multivesicular endosomes. *Science*, 319 (5867) ,1244-1247

Tuomilehto, J., Lindstrőm, J., Eriksson, J.G., Valle, T.T., Hämäläinen, H., Ilanne-Parikka, P., Keinäñen-Kiukaanniemi, S., Laakso, M., Louheranta, A., & Rastas, M. (2001) Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *New England Journal of Medicine*, 344 (18), 1343-1350

Turchinovich, A., Weiz, L., & Burwinkel, B. (2012) Extracellular miRNAs: the mystery of their origin and function. *Trends in Biochemical Sciences*, 37 (11), 460-465

Turchinovich, A., Weiz, L., Langheinz, A., & Burwinkel, B. (2011) Characterization of extracellular circulating microRNA. *Nucleic Acids Research*, 39 (16), 7223-33

Valadi, H., Ekstrőm, K., Bossios, A., Sjostrand, M., Lee, J.J., & Lőtvall, J.O. (2007) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nature Cell Biology*, 9 (6), 654-659

Van Rooij, E., & Olson, E.N. (2012) MicroRNA therapeutics for cardiovascular disease: opportunities and obstacles. *Nature Reviews Drug Discovery*, 11 (11), 860-872

Veliceasa, D., Biyashev, D., Qin, G., Misener, S., Mackie, A.R., Kishore, R., & Volpert, O.V. (2015) Therapeutic manipulation of angiogenesis with miR-27b. *Vascular Cell*, 7 (1), 1-13

Verma, N.K., Singh, J., & Dey, C.S. (2004) PPARG expression modulates insulin sensitivity in C2C12 skeletal muscle cells. *British Journal of Pharmacology*, 143 (8), 1006-1013

Vickers, K.C., Palmisano, B.T., Shoucri, B.M., Shamburek, R.D., & Remaley, A.T. (2011) MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nature Cell Biology*, 13 (4), 423-433

Vinik, A., & Flemmer, M. (2002) Diabetes and macrovascular disease. *Journal of Diabetes & its Complications*, 16 (3), 235-245

Wagner, J., Riwanto, M., Besler, C., Knau, A., Fichtlscherer, S., Röxe, T., Zeiher, A.M., Landmesser, U., & Dimmeler, S. (2013) Characterization of levels and cellular transfer of circulating lipoprotein-bound microRNAs. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 33 (6), 1392-1400

Wahlestedt, C. (2013) Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nature Reviews Drug Discovery*, 12 (6), 433-446

Wang, K., Yuan, Y., Cho, J.H., McClarty, S., Baxter, D., & Galas, D.J. (2012) Comparing the MicroRNA spectrum between serum and plasma. *PLoS One*, 7 (7), e41561

Wang, W., Lee, E.T., Fabsitz, R., Welty, T.K., & Howard, B.V. (2002) Using HbA1c to Improve Efficacy of the American Diabetes Association Fasting Plasma Glucose Criterion in Screening for New Type 2 Diabetes in American Indians The Strong Heart Study. *Diabetes Care*, 25 (8), 1365-1370

Wang, X., Sundquist, J., Zöller, B., Memon, A.A., Palmér, K., Sundquist, K., & Bennet, L. (2014) Determination of 14 circulating microRNAs in Swedes and Iraqis with and without diabetes mellitus type 2. *PloS One*, 9 (1), e86792

Waugh, N., Scotland, G., McNamee, P., Gillett, M., Brennan, A., Goyder, E., Williams, R., & John, A. (2007) Screening for type 2 diabetes: literature review and economic modelling. *British Journal of Clinical Governance,* 12 (4), doi: 10.1108/cgij.2007.24812dae.003

Weber, J.A., Baxter, D.H., Zhang, S., Huang, D.Y., Huang, K.H., Lee, M.J., Galas, D.J., & Wang, K. (2010) The microRNA spectrum in 12 body fluids. *Clinical chemistry*, 56 (11), 1733-1741

Wellen, K.E., & Hotamisligil, G.S. (2005) Inflammation, stress, and diabetes. *Journal of Clinical Investigation*, 115 (5), 1111-1119

Whiting, D.R., Guariguata, L., Weil, C., & Shaw, J. (2011) IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Research and Clinical Practice*, 94 (3), 311-321

Wieczorek, D., Delauriere, L., & Schagat, T. (2012) Methods of RNA Quality Assessment. [Cited 6 July 2015]; Available from: https://worldwide.promega.com/resources/pubhub/methods-of-rna-quality-assessment/

Wild, S., Roglic, G., Green, A., Sicree, R., & King, H. (2004) Global prevalence of diabetes estimates for the year 2000 and projections for 2030. *Diabetes Care*, 27 (5), 1047-1053

Wing, R.R., Goldstein, M.G., Acton, K.J., Birch, L.L., Jakicic, J.M., Sallis, J.F., Smith-West, D., Jeffery, R.W., & Surwit, R.S. (2001) Behavioral science research in diabetes lifestyle changes related to obesity, eating behavior, and physical activity. *Diabetes Care*, 24 (1), 117-123

Witkos, T.M., Koscianska, E., & Krzyzosiak, W.J. (2011) Practical aspects of microRNA target prediction. *Current Molecular Medicine*, 11 (2), 93-109

World Health Organisation. (2006) International Diabetes Federation. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia. *Report of WHO/IDF consultation*

World Obesity. (2012) [Cited 11 February 2015]; Available from: http://www.worldobesity.org/aboutobesity/

Wright, M.W. & Bruford, E.A. (2011) Naming 'junk': human non-protein coding RNA (ncRNA) gene nomenclature. *Hum Genomics*, 5 (2), 90-98

Xu, W., Wang, Z., & Liu, Y. (2014) The characterization of microRNA-mediated gene regulation as impacted by both target site location and seed match type. *PLoS One*, 9 (9), e108260

Yan, S.T., Li, C.L., Tian, H., Li, J., Pei, Y., Liu, Y., Gong, Y.P., Fang, F.S., & Sun, B.R. (2014) MiR-199a is overexpressed in plasma of type 2 diabetes patients which contributes to type 2 diabetes by targeting GLUT4. *Molecular and Cellular Biochemistry*, 397 (1-2), 45-51

Yang, Z., Chen, H., Si, H., Li, X., Ding, X., Sheng, Q., Chen, P., & Zhang, H. (2014) Serum miR-23a, a potential biomarker for diagnosis of pre-diabetes and type 2 diabetes. *Acta Diabetologica*, 51 (5), 823-831

Yue, D., Liu, H., & Huang, Y. (2009) Survey of computational algorithms for MicroRNA target prediction. *Current Genomics*, 10 (7), 478-492

Zampetaki, A., Kiechl, S., Drozdov, I., Willeit, P., Mayr, U., Prokopi, M., Mayr, A., Weger, S., Oberhollenzer, F., & Bonora, E. (2010) Plasma microRNA profiling reveals loss of endothelial miR-126 and other microRNAs in type 2 diabetes. *Circulation Research*, 107 (6), 810-817

Zampetaki, A., Willeit, P., Burr, S., Yin, X., Langley, S.R., Kiechl, S., Klein, R., Rossing, P., Chaturvedi, N., & Mayr, M. (2015) Angiogenic MicroRNAs Linked to Incidence and Progression of Diabetic Retinopathy in Type 1 Diabetes. *Diabete,* db150389

Zampetaki, A., Willeit, P., Drozdov, I., Kiechl, S., & Mayr, M. (2012) Profiling of circulating microRNAs: from single biomarkers to re-wired networks. *Cardiovascular Research*, 93 (4), 555-562

Zhang, J., Xiao, X., & Ji, L. (2013) Epigenetic disorder plays a fundamental role in the etiology of type 2 diabetes. *Wuhan University Journal of Natural Sciences*, 18 (1), 9-19

Zhang, T., Li, L., Shang, Q., Lv, C., Wang, C., & Su, B. (2015) Circulating miR-126 is a potential biomarker to predict the onset of type 2 diabetes mellitus in susceptible individuals. *Biochemical and Biophysical Research Communications,* 463 (1-2), 60-63

Zhang, Y., Jia, Y., Zheng, R., Guo, Y., Wang, Y., Guo, H., Fei, M., & Sun, S. (2010) Plasma microRNA-122 as a biomarker for viral-, alcohol-, and chemical-related hepatic diseases. *Clinical Chemistry*, 56 (12), 1830-1838

Zhou, J., Peng, R., Li, T., Luo, X., Peng, H., Zha, H., Yin, P., Wen, L., & Zhang, Z. (2013) A potentially functional polymorphism in the regulatory region of let-7a-2 is associated with an increased risk for diabetic nephropathy. *Gene*, 527 (2), 456-461

Zhu, H., & Leung, S.W. (2015) Identification of microRNA biomarkers in type 2 diabetes: a meta-analysis of controlled profiling studies. *Diabetologia*, 58 (5), 900-11

Zhuo, X., Zhang, P., & Hoerger, T.J. (2013) Lifetime direct medical costs of treating type 2 diabetes and diabetic complications. *American Journal of Preventive Medicine*, 45 (3), 253-261

Zimmet, P.Z., Magliano, D.J., Herman, W.H., & Shaw, J.E. (2014) Diabetes: a 21st century challenge. *The Lancet Diabetes & Endocrinology*, 2 (1), 56-64

Appendices

Appendix 1: Information sheet

What will you gain?

1. At the screening session, you will receive the following:
   - Refreshments after the last blood sample has been collected.
   - Individualised feedback by a nutritionist and a registered nurse on weight and height (BMI) status and blood pressure status.
   - Immediate referral to a clinician if the results of the fasting blood glucose or blood pressure results require medical attention.
2. Although you will not be paid for your participation in this study you will greatly assist us in developing a way to identify people at risk before they develop diabetes.
3. You will also receive at no cost to yourself, a medical and laboratory assessment of your current medical condition

How will you get feedback?

In the event of the discovery of any adverse laboratory results of your blood samples a member of the research team will contact you directly by telephone and advise you or refer you for further medical assistance. A full report of your blood results will be posted to you as soon as it gets available.

What are your rights in the study?

Your participation in this research study is voluntary, so you may withdraw at any given time. All your information will be confidential and treated with respect. You may also refuse to answer any questions you do not feel comfortable answering and you may stop during the interview and not continue. Your name will not be linked to the information collected at any time of the research and will not appear in any report or publication. The only people with access linking your name to information collected will be the researchers, who will respect the confidentiality of the information. Your information will be stored in locked files and protected computer files ensuring your confidentiality.

This study has been approved by the Research Ethics Committee of the Medical Research Council, Cape Town and will be carried out according to the ethical guidelines and principles of the International Declaration of Helsinki. If you have questions about your rights as someone who took part in the study, you are welcome to contact the Chairperson of the Research Ethics Committee, Prof Danie Du Toit, at the Medical Research Council, P.O. Box 19070, Tygerberg. 7505. Contact telephone number: 021 938 0341; email: adri.labuschagne@mrc.ac.za.

You will receive a copy of this information and consent form for your own records. If you have any questions or concerns about the research, please feel free to contact Sr Debbie Jonathan 021 938 0802, Erica April 938 0463 or Dr Johan Louw at the Medical Research Council, Tygerberg (Cape Town), Tel: 021 938 0322.

If you do **not suffer from diabetes** and are happy to participate in this study please supply the following information and send this form with your family member to the next wellness group meeting.

Family member's name and surname: _____
Own name and surname: _____
Contact details Telephone nr 1 _____
Telephone nr 2 _____

Male ☐ Female ☐ Age ☐

I can attend the screening session on a normal week day ☐
I can attend during the week if arrangements are made with my employer ☐
I can only attend on a Saturday ☐

Appendix 2: Consent form

CONSENT FORM: Early markers and determinants of diabetes, as well as health cognition and behaviour of relatives of type 2 diabetics attending diabetes support groups in the Cape Town Metropole

I declare that:

I have read or had read to me this information and it is written in a language with which I am fluent and comfortable.

I am between the ages of 18 and 65.

I have had a chance to ask questions and all my questions have been adequately answered.

I know that taking part in this study is voluntary and I have not been forced to take part. I may choose to leave the study at any time without any problems.

Signed at ..... on (date) ..... 2013.

....

Name of participant Signature of participant

....

Name of witness Signature of witness

Appendix 3: RNA quality control and assessment

Table A3.1 Total RNA quantification and quality assurance by NanoDrop ND-1000 spectrophotometer

Sample Name	OD260/280 Ratio	Conc. (ng/µl)	Volume (µl)	Total Amount (ng)
1	1.81	121.96	10	1219.6
2	1.73	123.19	10	1231.9
3	1.81	133.72	10	1337.2
4	1.78	175.11	10	1751.1
5	1.78	150.92	10	1509.2
6	1.87	124.07	10	1240.7
7	1.76	88.53	10	885.3
8	1.78	162.49	10	1624.9
9	1.78	138.24	10	1382.4
10	1.71	64.37	10	643.7
11	1.80	130.06	10	1300.6
12	1.79	147.66	10	1476.6

Table A3.2 Quality assessment of the sequencing library, determined by Agilent 2100 Bioanalyser using the Agilent DNA 1000 chip kit (Agilent technologies).

Sample Name	Size (bp)	Conc. (nmol/L)	Conc. (ng/µl)	Volume (µl)**	Total Amount (ng)
1	142	54.2	5.08	20	101.6
2	140	54.5	5.03	20	100.6
3	141	53.3	4.95	20	99.0
4	140	59.4	5.49	20	109.8
5	140	66.1	6.12	20	122.4
6	141	52.8	4.93	20	98.6
7	141	39.4	3.67	20	73.4
8	142	53.8	5.04	20	100.8
9	141	34.5	3.22	20	64.4
10	142	89.6	8.38	20	167.6
11	141	58.0	5.41	20	108.2
12	141	63.2	5.86	20	117.2

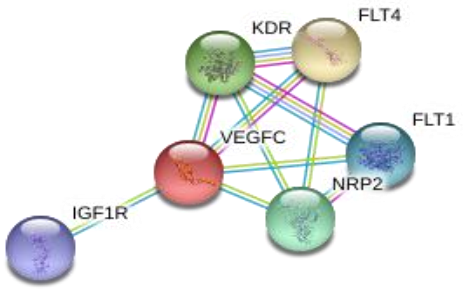
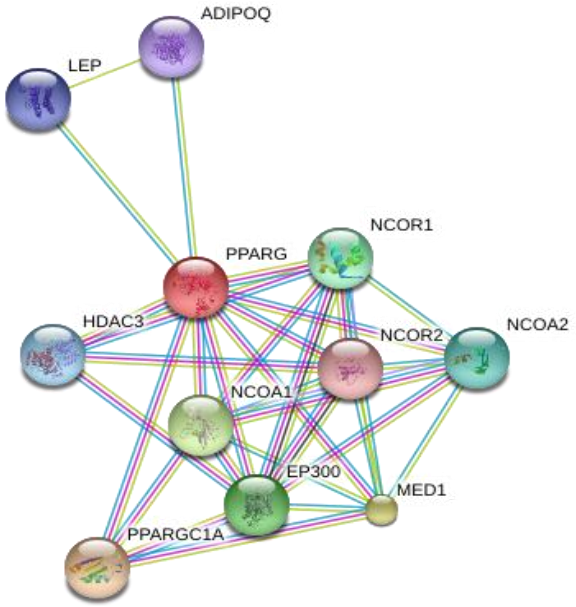
**The concentration of each library was adjusted to 10nM before cluster generation

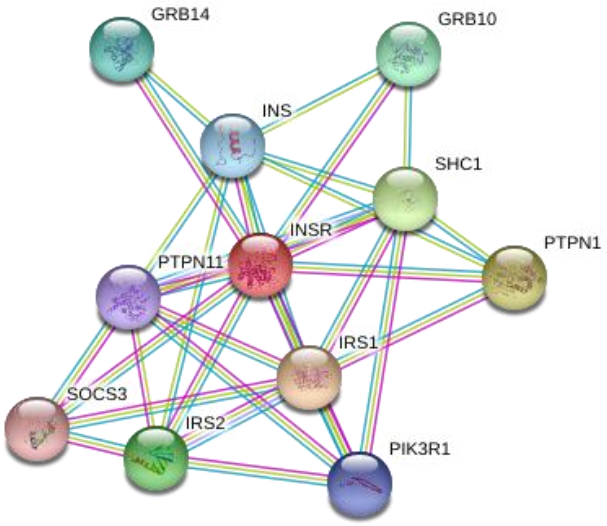
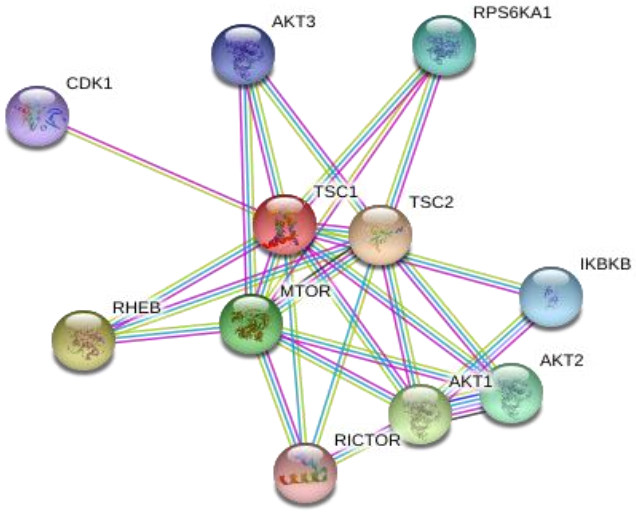
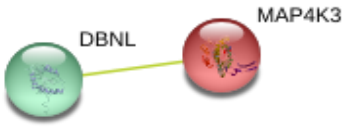
.

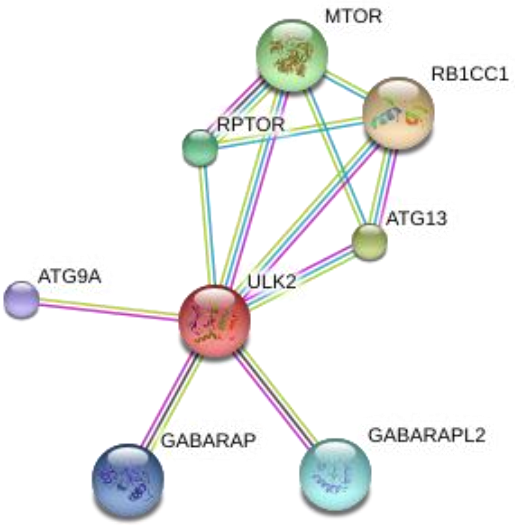
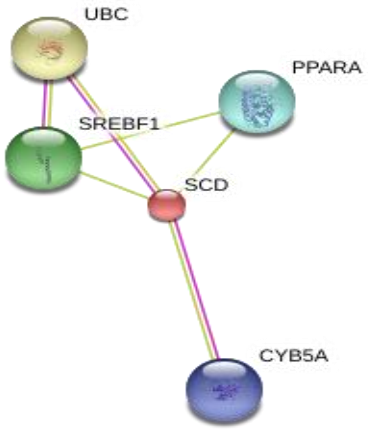
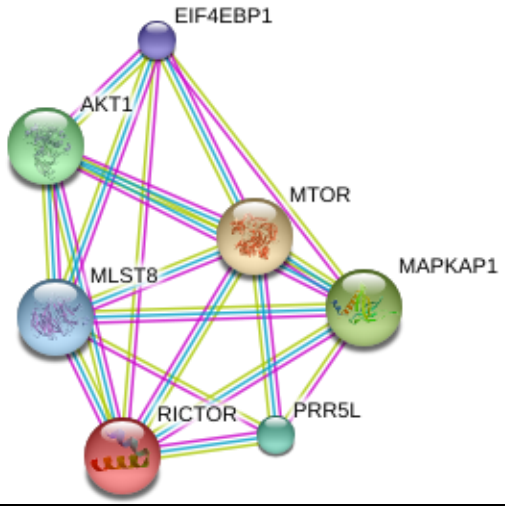
Appendix 4: Protein-protein interactions

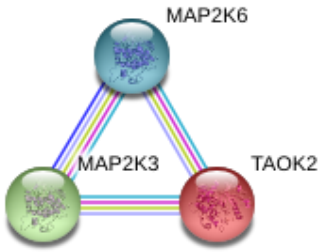
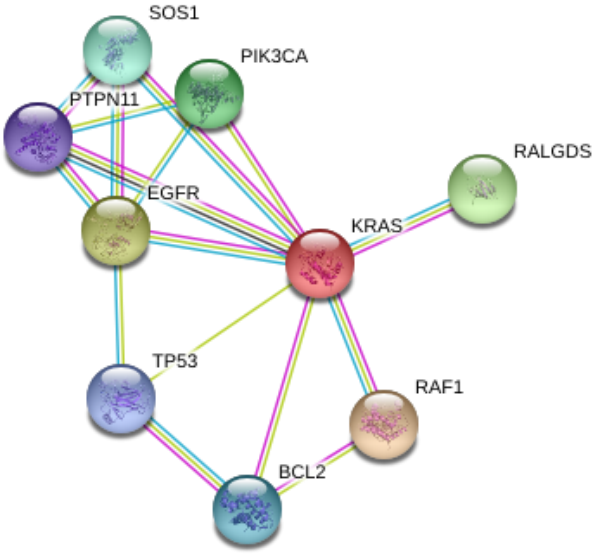
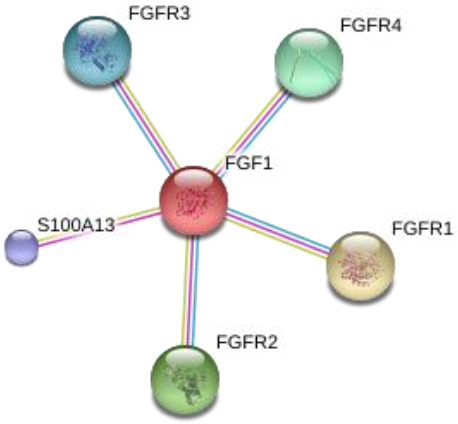
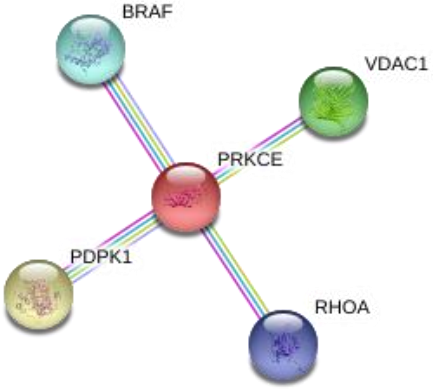
Protein-protein interactions possibly regulated by the differentially expressed miRNAs were investigated using the protein interaction database called STRING. The red nodes represent the target protein of interest, which interact with a number of functional proteins predicted using different sources of evidence. The different colour lines indicate different prediction methods such as, fusion evidence (red), neighbourhood (green), co-occurrence (blue), experimental evidence (purple), text-mining (yellow), database evidence (light blue) and co-expression (black).

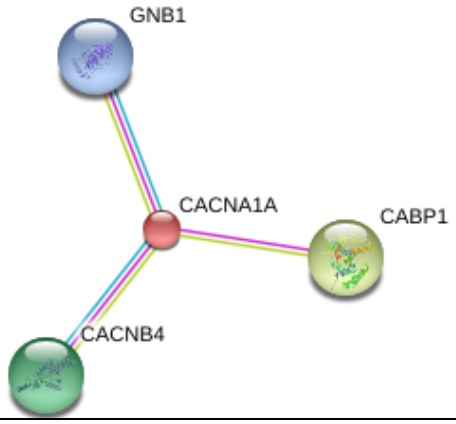
Table A4 Functional protein-protein interactions determined by STRING.

MicroRNA target gene/protein	**Functional protein-protein interaction network**
Vascular endothelial growth factor C **(VEGFC)**	
Peroxisome proliferator-activated receptor gamma **(PPARG)**	

Insulin receptor **(INSR)**	
Tuberous sclerosis 1 **(TSC1)**	
Mitogen-activated protein 4 kinase 3 **(MAP4K3)**	

Unc-51-like kinase **(ULK2)**	
Stearoyl-CoA desaturase **(SCD)**	
RPTOR independent companion of mTOR **(RICTOR)**	

TAO kinase 2 **(TAOK2)**	
v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog **(KRAS)**	
Fibroblast growth factor 1 **(FGF1)**	
Protein kinase C, epsilon **(PRKCE)**	

Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit **(CACNA1A)**	

The names of proteins that interact with predicted miRNA targets are as follows:

- **VEGFC**; fms-related tyrosine kinase 4 (FLT4), kinase insert domain receptor (KDR), neuropilin 2 (NRP2), fms-related tyrosine kinase 1 (FLT1) and insulin-like growth factor 1 receptor (IGF1R).
- **PPARG**; peroxisome proliferator-activated receptor gamma (PPARGC1A), mediator complex subunit 1 (MED1), nuclear receptor coactivator 1 (NCOA1), E1A binding protein p300 (EP300), nuclear receptor corepressor 1 (NCOR1), nuclear receptor coactivator 2 (NCOA2), histone deacetylase 3 (HDAC3), leptin (LEP), adiponectin (ADIPOQ) and nuclear receptor corepressor 2 (NCOR2).
- **INSR**; insulin receptor substrate 1 (INS1), protein tyrosine phosphatase non-receptor type 1 (PTPN) , SHC (Src homology 2 domain containing) transforming protein 1 (SHC1), insulin receptor substrate 2 (IRS2), growth factor receptor-bound protein 10 (GRB10), growth factor receptor-bound protein 14 (GRB14), insulin (INS), phosphoinositide-3-kinase regulatory subunit 1 (alpha) (PIK3R1), protein tyrosine phosphatase non-receptor type 11 (PTPN11) and suppressor of cytokine signaling 3 (SOCS3).
- **TSC1**; tuberous sclerosis 2 (TSC2), Ras homolog enriched in brain (RHEB), v-akt murine thymoma viral oncogene homolog 1 (AKT1), mechanistic target of rapamycin (mTOR), v-akt murine thymoma viral oncogene homolog 2 (AKT2), ribosomal protein S6 kinase, 90kDa, polypeptide 1 (RPS6KA1), inhibitor of kappa light polypeptide gene enhancer in B-cells (IKBKB), v-akt murine thymoma viral oncogene homolog 3 (AKT3), cyclin-dependent kinase 1 (CDK1) and RPTOR independent companion of mTOR (RICTOR).
- **MAP4K3**; debrin-like (DBNL)
- **ULK2**; RB1-inducible coiled-coil 1 (RB1CC1), autophagy related 13 (ATG13), mechanistic target of rapamycin (mTOR), regulatory associated protein of mTOR,

complex 1 (RPTOR), GABA(A) receptor-associated protein-like 2 (GABARAPC2), GABA(A) receptor-associated protein (GABARAP) and autophagy related 9A (ATG9A)

- **SCD**; ubiquitin C (UBC), sterol regulatory element binding transcription factor 1 (SREBF1), peroxisome proliferator-activated receptor alpha (PPARA) and cytochrome b5 type A (CYB5A)
- **RICTOR**; mechanistic target of rapamycin (mTOR), mitogen-activated protein kinase associated protein 1 (MAPKAP1), v-akt murine thymoma viral oncogene homolog 1 (AKT1), proline rich 5 like (PRR5L), mTOR associated protein (MLST8) and eukaryotic translation initiation factor 4 binding protein 1 (EIF4BP1)
- **TAOK2**; mitogen-activated protein kinase 3(MAP2K3) and mitogen-activated protein kinase 6 (MAP2K6)
- **KRAS**: v-raf-1 murine leukemia viral oncogene homologue 1 (RAF1), epidermal growth factor receptor (EGFR), ral guanidine nucleotide dissociation stimulator (RALGDS), phosphatidylinositol-4-5-bisphosphate 3-kinase (PIK3CA), Son of sevenless homolog 1 (SOS1), B-cell CLL/lymphoma 2 (BCL2), protein tyrosine phosphatase, non-receptor type 11 (PTPN11) and tumor protein p53 (TP53).
- **FGF1**; fibroblast growth factor receptor 1 (FGFR1), fibroblast growth factor receptor 2 (FGFR2) , fibroblast growth factor receptor 4 (FGFR4), fibroblast growth factor receptor 3 (FGFR3), S100 calcium binding protein A13 (S100A13) and fribroblast growth factor 23 (FGF23)
- **PRKCE**; 3-phosphoinositide dependent protein kinase-1 (PDPK1), voltage-dependent anion channel 1 (VDAC1), v-raf murine sarcoma viral oncogene homolog B1 (BRAF) and ras homolog family member A (RHOA)
- **CACNA1A**; calcium binding protein 1 (CABP1), calcium channel voltage dependent beta-4 subunit (CACNB4) and guanine nucleotide binding protein (GNB1)