Serum circulating microRNA profiling for identification of potential markers of diabetic nephropathy in black African South Africans with type 2 diabetes mellitus

Mr. Stefan Drikus Valentin

717417

MSc (Med) by Dissertation

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science in Medicine.

Johannesburg, 2020

Declaration

I, Stefan Drikus Valentin, declare that the project titled “Serum circulating microRNA profiling for identification of potential markers of diabetic nephropathy in black African South Africans with type 2 diabetes mellitus” is my own work. It is being submitted in fulfilment for the degree of Master of Science in Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Stefan Drikus Valentin **27 July 2020**

Conference presentations

1. Third Pathology Research and Development Congress (PathRed) 2019 (Johannesburg)
   Valentin, S.D., Prigge, K., Cave, E., Ambele, M., and Padoa, C.J. (2019). Differentially expressed serum microRNAs in South African black type 2 diabetic patients with diabetic nephropathy.
   Presentation format: Poster Presentation and Speed Presentation.
   Awarded the Best Poster within the Human Genetics Division.

2. South African Society of Human Genetics (SASHG) Congress 2019 (Cape Town).
   Valentin, S.D., Prigge, K., Cave, E., Ambele, M., and Padoa, C.J. (2019). Differentially expressed serum microRNAs in South African black type 2 diabetic patients with diabetic nephropathy.
   Presentation format: Flash oral presentation.

Abstract

Introduction: Type 2 diabetes mellitus (T2DM) is a metabolic disorder characterised by insulin resistance. The prevalence of T2DM is increasing with an estimated 14.2 million adults with T2DM expected in Africa by the year 2040. Poor glycaemic control, over a prolonged period of time, leads to diabetic micro- and macrovascular complications. Of the microvascular complications, diabetic nephropathy (DN) has been implicated in causing the majority of deaths associated with T2DM. DN is characterised by progressive decline of renal function, resulting in end-stage renal disease. With early intervention, kidney damage can be prevented or even reversed, however no test currently exists for early detection of DN. Currently, urine albumin-to-creatinine ratios (UACR) are used to monitor kidney function and diagnose DN. UACR as a marker is unspecific and unreliable, and there is a need for a more sensitive and specific marker. Serum circulating microRNA (miRNA) have shown promise as potential novel biomarkers as they are stable and often differentially expressed in different pathological states. This study, therefore, aimed to detect differentially expressed serum miRNAs in black South African participants with T2DM with and without microalbuminuria (MA) to identify potential biomarkers for early detection of DN. In addition, the prevalence of polymorphisms in the *HMGA2* and *TGF-β1* gene and their associations with markers of kidney function were investigated. Further, the use of serum cystatin C (SCC) as an alternative to serum creatinine for estimation of glomerular filtration rate (GFR; used to measure kidney function and to classify the stage of kidney disease) was assessed.

Methods: Black South African participants with T2DM (n = 238) were recruited from diabetes clinics at Charlotte Maxeke Johannesburg Academic Hospital and Chris Hani Baragwanath Hospital. A subset of participants (n = 14) were classified into two groups based on the absence (UACR < 2 mg/mmol; controls, n = 7) or presence (UACR 2-20 mg/mmol; cases, n = 7) of MA. The participants were matched for age, medication use, and anthropometry. miRNA expression profiles were analysed using the Affymetrix® GeneChip® 4.0 assay and transcriptome analysis console (TAC) software. Two differentially expressed miRNAs were selected for validation through real-time quantification polymerase chain reaction (PCR) in a subset of 40 participants (20 participants with the lowest UACR, and 20 participants with the highest UACR). Two genes that were targeted by the differentially expressed miRNA were selected for polymorphism analysis. The total cohort was genotyped for *HMGA2* rs1114167319, and *TGF-β1*

rs1800417 polymorphism by amplification refractory mutation systems (ARMS) PCR. In addition, participants were genotyped for the *HMGA2* rs1114167320 polymorphism by PCR restriction fragment length polymorphism (RFLP).

SCC concentrations were measured using a R&D enzyme-linked immunosorbent assay (ELISA). Estimated GFR (eGFR) were calculated through the use of two creatinine-based equations (modification of diet in renal disease study (MDRD) equation without ethnic factor correction and the chronic kidney disease epidemiology collaboration (CKD-EPI) equation), and three SCC-based equations. Passing-Bablok linear regression and Band-Altman plots were generated to evaluate correlation and method agreement between the eGFR equations, using the MDRD equation as the reference method.

Results: The study population consisted of a 238 black South Africans with T2DM. The cohort had a mean age of 56.6 ± 9.88 years and a mean duration of disease of 11.5 ± 8.82 years. The majority of the participants were female (61.8 %), and they had poor glycaemic control (HbA1c of 8.68 ± 2.73 %).

Differential miRNA expression profiling (> 2-fold difference between cases and controls) identified nine miRNAs. miR-455-3p and let-7b-5p were upregulated, and miR-4740-3p, miR-7704, miR-101-5p, miR-5189-3p, miR-1273g-3p, miR-16-5p and miR-6880-5p were downregulated in cases relative to controls. Real-time quantitative PCR validation of expression levels of miR-455-3p and let-7b-5p found that expression levels of both miRNAs were higher in the group with the lowest UACR compared to the group with the highest UACR (miR-455-3p: 40.5 [39.1; 42.8] vs 42.5 [41.3; 45.3] C_T value, p = 0.043; and let-7b-5p: 26.3 ± 1.50 vs 27.6 ± 2.03 C_T value, p = 0.022).

The *HMGA2* rs111416319 and rs111416320 major alleles were the C (0.51) and A allele (1.0), respectively. The major allele for the *TGF-β1* rs1800471 polymorphism was the G allele (0.64). None of the polymorphisms were in Hardy-Weinberg equilibrium (HWE). Despite not being in HWE, we did perform additional analyses with rs111416319 and rs1800471. rs111416319 was not associated with any markers of kidney function or DM. Similarly, rs1800471 was not associated with markers of kidney function, however, participants with the GG genotype had significantly lower glucose concentrations compared to participants with the GC/CC genotype (6.60 [4.80; 8.80] vs. 8.10 [5.70; 11.3] mmol/l; p = 0.011).

The classification of stages of kidney disease using the five different eGFR equations resulted in significantly different clinical classifications of kidney function (p < 0.001). The MDRD

and CKD-EPI equation correlated strongly (r = 0.945; p < 0.001) however there was poor agreement between the equations with a bias of -39.7 %. There was good agreement between SCC equation 2 and equation 3 with the MDRD equation with a bias of -7.71 % and -7.10 %, respectively.

Conclusion: Let-7b-5p and miR-455-3p protect against the development of DN. Let-7b-5p binds to the *TGF-β1* receptor 1 inhibiting the TGF-β1 signalling pathway. This pathway has been shown to play a role in DN by enhancing glucose-induced cell hypertrophy and increased expression of extracellular matrix (ECM) genes (e.g. collagen and fibronectin). In addition, let-7b-5p targets the 3' untranslated region of the *HMGA2* gene, that has been implicated in activation of the TGF-β1 signalling pathway. Thus, increased expression of let-7b-5p leads to decreased renal fibrosis via inhibiting *HMGA2* and *TGF-β1* receptor expression. miR-455-3p overexpression decreases ECM synthesis and inflammatory cytokines resulting in protection from renal fibrosis. Furthermore, TGF-β1 inhibits the expression of miR-455-3p, thus activating renal fibrotic pathways, resulting in reduced kidney function and DN development.

Further studies are required in the South African black population to confirm the role of these miRNAs in DN. Longitudinal studies need to be conducted to confirm the role of let-7b-5p and miR-455-3p as possible early markers of DN. The eGFR equations showed significant variation in terms of classification of kidney disease and cannot be used interchangeably.

Acknowledgements

The researchers would like to acknowledge the following individuals and organisations for their contribution to the project:

- The National Health Laboratory Services (NHLS), Medical Research Council (MRC), and the University of the Witwatersrand Faculty Research Council (FRC) for grant funding, and the University of the Witwatersrand for the Postgraduate Merit Award.
- The University of Witwatersrand and the Department of Chemical Pathology for providing the infrastructure to conduct the research.
- Special thanks to the following individuals for their contribution to the project in the following ways:
    - Dr. Derryn Legg E'Silva, and Ms. Katherine Prigge from the Department of Chemical Pathology (University of the Witwatersrand), for administration of tests within the NHLS laboratories.
    - Dr. Marketa Toman from the Department of Chemical Pathology (University of the Witwatersrand) for access, interpretation, and aid with the ELISA equipment and results.
    - Sr. Nomsa Ramela for her assistance with patient recruitment and sample collection in the diabetic clinics.
    - Dr. Pareen Patal from the Department of Molecular Medicine and Haematology (University of the Witwatersrand) for assistance with and access to the Agilent Bioanalyzer.
    - Mr. Dean Harris and Dr. Michelle Bronze from DiagnosTech, and Ms. Michaella Hully from the Department of Human Genetics (University of the Witwatersrand), for their assistance with and access to the Agilent Bioanalyzer.
    - Prof. Michael S. Pepper and Dr. Melvin Ambele from the University of Pretoria, Faculty of Health Sciences, School of Pathology, Department of Immunology, for their assistance with and access to the Agilent machinery and interpretation of the expression profiling results.
    - Dr. Eleanor Cave for her assistance with quantitative PCR design and interpretation of results.

□ Ms. Katherine Prigge, Dr. Eleanor Cave, and Dr. Carolyn Padoa, for their endless assistance, support, and supervision throughout the project.

Table of contents

List of abbreviations

A	Adenosine
ACE	Angiotensin converting enzyme
ADM	Adrenomedullin
AGE	Advanced glycation end
ANOVA	Analysis of variance
ARMS	Amplification refractory mutation systems
ATP	Adenosine 5'-triphosphate
BMI	Body mass index
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
cDNA	Complimentary deoxyribonucleic acid
CHBH	Chris Hani Baragwanath Hospital
CKD	Chronic kidney disease
CKD-EPI	Chronic kidney disease epidemiology collaboration
cm	Centimetre
CMJAH	Charlotte Makexe Johannesburg Academic Hospital
CSF-1	Colony stimulating factor one
C_T	Cycle number at which the qPCR amplification curve rose above the threshold
dBP	Diastolic blood pressure
del	Deletion
DKD	Diabetic kidney disease
DM	Diabetes mellitus
DMSO	Dimethyl sulfoxide
DN	Diabetic nephropathy
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid

e.g.	Exempli gratia (example)
eGFR	Estimated glomerular filtration rate
ELISA	Enzyme-linked immunosorbent assay
ELOSA	Enzyme-linked oligosorbent assay
EMT	Epithelial cell to mesenchymal cell transition
eq	Equation
ESRD	End stage renal disease
EtBr	Ethidium bromide
EVA	Extra-cellular vesicle associated
FABP	Fatty acid-binding protein
FRC	Faculty Research Council
g	Grams
G	Guanine
GBM	Glomerular basement membrane
GFR	Glomerular filtration rate
HbA1c	Glycated haemoglobin A1c
HMGA2	High mobility group AT-hook two
HWE	Hardy-Weinberg equilibrium
IL-6	Interleukin six
INT$	International dollar
JNK	c-Jun N-terminal kinase
kDa	Kilodalton
KDIGO	Kidney disease improving global outcomes
kg	Kilogram
kg/m^2	Kilogram per square meter
l	Litre
LD	Linkage disequilibrium
m	Metre
MA	Microalbuminuria
MAPKs	Mitogen-activated protein kinase
MDRD	Modification of diet in renal disease
mg/day	Milligram per day

mg/l	Milligram per litre
mg/ml	Milligram per millilitre
mg/mmol	Milligram per millimole
MGB	Minor grove binding
min	Minutes
miRNA	Micro-ribonucleic acid
ml	Millilitre
ml/min/1.73m^2	Millilitre per minute per 1.73 meter squared
mM	Millimolar
mmHg	Millimetre mercury
mmol/l	Millimole per litre
$MnCl_2$	Manganese dichloride
MRC	Medical Research Council
mRNA	Messenger ribonucleic acid
n	Sample size
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B-cells
ng	Nanogram
ng/µl	Nanogram per microlitre
ng/ml	Nanogram per millilitre
NHLS	National Health Laboratory Services
nm	Nanometre
nM	Nanomolar
NO	Nitric oxide
NTC	Non-template control
NVA	Non-vesicle associated
OGTT	Oral glucose tolerance test
PAP	Phosphatidate phosphatase
PathRed	Pathology Research and Development Congress
PBS	Phosphate buffered saline
PCA	Principle component analysis
PCR	Polymerase chain reaction
pg	Picogram

pg/µl	Picogram per microlitre
PKC-β	Protein kinase C-beta
Pre-miRNA	Precursor micro-ribonucleic acid
Pri-miRNA	Primary micro-ribonucleic acid
QC	Quality control
qPCR	Quantitative polymerase chain reaction
RFLP	Restriction fragment length polymorphism
RIN	Ribonucleic acid integrity number
RISC	Ribonucleic acid-induced silencing complex
RNA	Ribonucleic acid
ROCK2	Rho-associated coiled coil-containing protein two
ROS	Reactive oxygen species
rpm	Revolutions per minute
RPSAP52	Ribosome protein SA pseudogene 52
rRNA	Ribosomal ribonucleic acid
RSS	Russel-Silver syndrome
RT	Room temperature
SA-HRP	Streptavidin horseradish peroxide
SASHG	South African Society of Human Genetics
sBP	Systolic blood pressure
SCC	Serum cystatin C
S_{cr}	Serum creatinine
SD	Standard deviation
SDS	Sodium dodecyl sulphate
sec	Seconds
SEMDSA	Society of endocrine, metabolism and diabetes of South Africa
SMAD	Protein homologous to the Drosophila MAD (mothers against decapentaplegic) and Caenorhabditis elegans SMA (small worm phenotype) proteins
SNP	Single nucleotide polymorphism
SSC	Saline sodium citrate
SST	Serum separating tube

SUMMIT	Surrogate markers for micro- and macro-vascular hard endpoint for innovative diabetes tools
T	Thymine
T1DM	Type one diabetes mellitus
T2DM	Type two diabetes mellitus
TAC	Transcriptome analysis console
TBE	Trisaminomethane, boric acid, ethylenediaminetetraacetic acid solution
TGF-β	Transforming growth factor beta
TGF-βR	Transforming growth factor beta receptor
THP	Tamm-Horsfall urinary protein
Tris	Trisaminomethane
U/µl	Units per microlitre
U/ml	Units per millilitre
UACR	Urine albumin-to-creatinine ratio
UK	United Kingdom
UNG	Uracil-N-glycosylase
USA	United States of America
UTR	Untranslated region
V	Volts
VDR	Vitamin D receptor
VEGF	Vascular endothelial growth factor
vs	Versus
WHR	Waist-to-hip ratio
x g	Times earth's gravity
β	Beta
χ^2	Chi-square
µl	Microlitre
µmol/l	Micromole per litre
%	Percentage
“	Inch
°C	Degree Celsius

24h	24 hours
3’	Three-prime
5’	Five-prime

List of Figures

List of Tables

1. Introduction

1.1. Diabetes mellitus

Diabetes mellitus (DM) is a metabolic disorder, whereby the body becomes incapable of insulin production, utilization or both. In general, the disease results in chronic hyperglycaemia and disturbs the metabolism of carbohydrates, fats, and proteins and often results in long-term damage to bodily systems (Alberti and Zimmet, 1998; American Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013).

Typical early-stage symptoms of DM include polyuria (excessive urination), polydipsia (excessive thirst), blurred vision, weight loss and polyphagia (excessive hunger) (Alberti and Zimmet, 1998). A more severe manifestation, namely ketoacidosis or non-ketotic hyperosmolar states can occur, which can lead to stupors or coma and death, if left untreated (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017).

1.1.1. DM diagnosis

A diagnosis of DM in individuals with symptoms of hyperglycaemia or metabolic decompensation can be confirmed by a single random glucose measurement ≥ 11.1 mmol/l, a fasting glucose measurement ≥ 7.0 mmol/l, or a haemoglobin A1c (HbA1c) result ≥ 6.5 %. For asymptomatic patients the fasting glucose and HbA1c tests, in addition to a two-hour oral glucose tolerance test (OGTT), need to be repeated on separate days within a two-week period to confirm the diagnosis. The two-hour OGTT glucose result should be ≥ 11.1 mmol/l for a DM diagnosis. In instances where the HbA1c result is close to the cut off, a glucose-based measurement is required to confirm the diagnosis (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017).

1.1.2. DM classification

The most widely used classification of DM is based on disease aetiology. Four types of DM have been described.

Type 1 DM (T1DM) results from destruction of the pancreatic beta (β)-cells (American Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013). This leads to insulin deficiency, and affected individuals are prone to ketoacidosis. T1DM can be divided into immune-mediated (Ali, 2013) or idiopathic (American

Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013).

Type 2 DM (T2DM) results from the progressive loss of insulin secretion and the development of insulin resistance. T2DM is diagnosed once all other causes of DM have been excluded (American Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013).

The third type, termed "other specific types of diabetes," has multiple aetiologies, primarily of genetic origin or associated with other diseases or drugs (American Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013).

Gestational diabetes is the fourth type of DM and is diagnosed when hyperglycaemia persists in the second or third trimester of pregnancy, and is not connected to other types of DM (American Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013).

1.2. Type 2 diabetes mellitus (T2DM)

T2DM is the most common of the four types of DM, estimated to account for 90-95 % of all DM cases, including undiagnosed DM cases (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). An estimated national prevalence of T2DM of 7.0 % has been reported in the South African adult population (aged 20-79 years), however, it is estimated that 61.1 % of diabetic cases remain undiagnosed (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). Globally, the number of adults with DM is expected to increase by 48 % by the year 2045, with the greatest increase of 156 % predicted in Africa (International Diabetes Federation, 2017).

T2DM occurs due to progressive loss of insulin secretion, with the key characteristic being the development of insulin resistance (American Diabetes Association, 2010). Insulin is a peptide hormone produced and secreted by the β-cells of the pancreatic islets of Langerhans. Insulin functions to maintain the normal levels of blood glucose through managing cellular glucose uptake and regulating carbohydrate, lipid and fat metabolism (Wilcox, 2005).

Insulin resistance is an attenuated biological response in the presence of normal or elevated insulin production. In most cases this refers to impaired sensitivity to insulin mediated glucose disposal (Wilcox, 2005), often as a result of increased lipid-induced alteration of the glucose uptake process (Samuel and Shulman, 2012). This results in compensatory

hyperinsulinaemia to account for the rising levels of glucose (Wilcox, 2005). While the specific aetiology of T2DM is not completely understood, it is thought to be due to a combination of environmental factors, genetic predisposition and dysregulation (Wilcox, 2005; Ali, 2013), and epigenetic factors (Dayeh *et al.*, 2013).

1.2.1. Risk factors for T2DM

Many risk factors for T2DM have been linked to susceptibility to and development of insulin resistance. Some of the factors that impact an individual's risk of developing T2DM include age, ethnicity, family history, high blood pressure, obesity and a sedentary lifestyle (Lim, 2014; Wilcox, 2005). In addition, population, family, and twin-based studies have shown that genetic contributions account for 20-80 % of the risk of developing T2DM, depending on the mode of inheritance (Ali, 2013).

1.2.2. Complications of T2DM

With prolonged disease duration, complications may arise due to the natural progression of DM (Alberti and Zimmet, 1998). These can be classified into two categories, namely macrovascular and microvascular complications (Alberti and Zimmet, 1998; SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017).

Macrovascular complications arise in the major circulatory network and include coronary artery disease, peripheral arterial disease, stroke and other forms of heart-failure (Alberti and Zimmet, 1998). Microvascular complications relate to the smaller circulatory network, and lead to the dysfunction of other organ systems; namely the eyes (retinopathy), the nervous system (neuropathy), and the kidneys (nephropathy) (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017).

1.3. Diabetic nephropathy

Diabetic nephropathy (DN), also referred to as diabetic kidney disease (DKD) (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017), is the leading cause of end stage renal disease (ESRD) and renal replacement therapy, in the western world (American Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013), and represents a major health problem (Gilbertson *et al.*, 2005). If left untreated, it can lead to

chronic kidney disease (CKD), and death (Alberti and Zimmet, 1998; American Diabetes Association, 2010; Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, 2013). Approximately 40 % of individuals with DM will develop CKD in the United States of America (USA) (Rotchford and Rotchford, 2002; Palmer *et al.*, 2008; Pezzolesi and Krolewski, 2013; Noubiap, Naidoo and Kengne, 2015).

Genetic research into T2DM and DN has historically been focused on people of European ancestry (Knerr, Wayman and Bonham, 2011). Thus, there is limited data on DN in the African population, with DN prevalence in sub-Saharan Africa ranging between 31-36.7 % (Levitt *et al.*, 1997; Choukem *et al.*, 2012). In South Africa, the prevalence of DN is estimated to be approximately 14-16 % (Naicker, 2009), however underreporting is suspected (F Raal, personal communication, April 2014).

Renal fibrosis is the primary cause of mortality and morbidity rates linked to DN (Wu *et al.*, 2018). A study conducted in a cohort of T2DM patients from Cape Town showed that renal failure was a major cause of death (28.8 %), particularly in non-Caucasian patients (Keeton, van Zyl Smit and Bryer, 2004).

1.3.1. Pathology of DN

DN is defined as the fibrosis and progressive failure of the kidneys due to high glucose levels. Kidney fibrosis is characterised by excessive extracellular matrix (ECM) synthesis and accumulation resulting in glomerular and tubular damage (Wang *et al.*, 2014). This leads to progressive decline of renal function (Wu *et al.*, 2018).

The process of DN development is initiated by metabolic, haemodynamic, inflammatory, and fibrotic changes that occur in the presence of hyperglyceamia and dyslipidaemia as a result of DM (Figure 1.1). These pathways overlap and interact with one another enhancing the physiological effects on the kidney (Wolf, 2004).

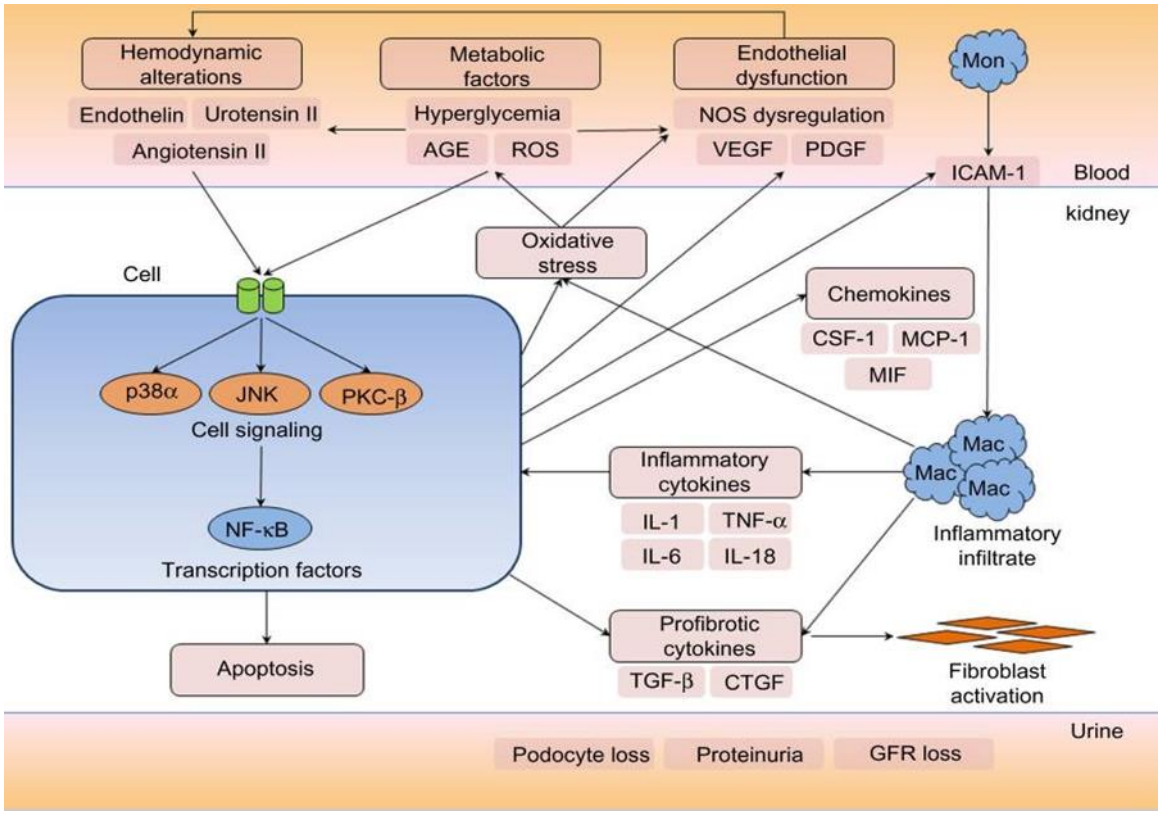


Figure 1.1: Schematic representation of the pathways (and their dysregulation) that lead to DN in the presence of DM

Interactions between metabolic and haemodynamic factors trigger multiple signalling pathways, such as MAPKs and PKC-β, resulting in transcription factor activation (NF-κB). The renal cells respond by producing profibrotic and inflammatory cytokines, growth factors, and chemokines. Collectively and interactively, these pathways result in podocyte loss, GFR decline, and proteinuria (Lim, 2014). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (as amended).

1.3.1.1. Metabolic and haemodynamic changes

Hyperglycaemia leads to the accumulation of various metabolites and by-products, such as advanced glycation end (AGE) products and reactive oxygen species (ROS), via glucose dependent pathways (Figure 1.1). In addition, high glucose levels stimulate vasoactive hormone pathways (renin-angiotensin system) leading to angiotensin II synthesis (Wolf, Butzmann and Wenzel, 2003; Chawla, Sharma and Singh, 2010). These metabolic and haemodynamic factors together induce diabetic renal pathology via the activation of intracellular second messengers such as mitogen-activated protein kinases (MAPK*s*; e.g. p38 and c-Jun N-terminal kinase (JNK)) and protein kinase C-β (PKC-β). These signalling cascades lead to the activation of key transcription factors (nuclear factor kappa-light-chain-enhancer of activated B-cells (NF-κB)), which are responsible for mediating a cellular response (Lim, 2014). In response to these signals, renal cells

produce inflammatory (e.g. interleukin (IL)-6) and profibrotic (e.g. transforming growth factor (TGF)-β) cytokines, chemokines (e.g. colony stimulating factor one (CSF-1)) and growth factors (e.g. vascular endothelial growth factors (VEGF)). These pathways overlap and interact with one another resulting in the structural and functional changes in the kidneys that affect overall renal function (Wolf, 2004; Lim, 2014; Alicic, Rooney and Tuttle, 2017).

Haemodynamic changes, such as hyperfiltration and hyperperfusion, within the glomeruli, occur early in DN (Hostetter, 2003). This is a result of increased glomerular capillary pressure due to enhancement of the trans-capillary hydraulic pressure gradient, and glomerular plasma flow (Hostetter, 2003). The tubular hypothesis proposes that hyperfiltration is the result of increased sodium reabsorption in the proximal tubules or loops of Henle, and subsequent reduction in proximal tubular pressure, leading to an increased glomerular filtration rate (GFR) (Thomson, Vallon and Blantz, 2004) (Figure 1.2).

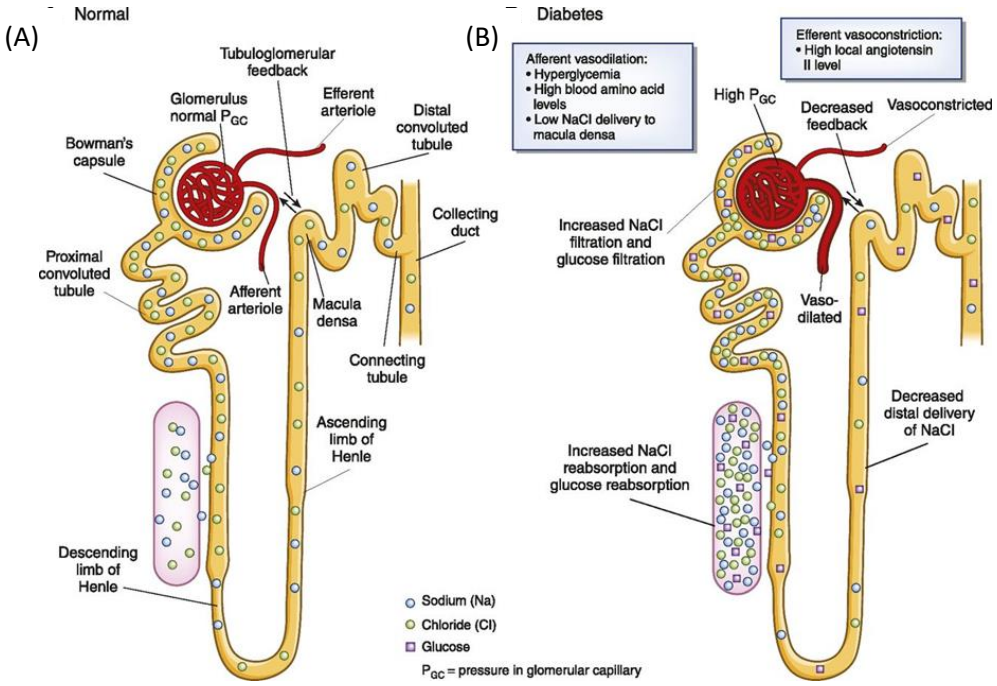


Figure 1.2: Diagram illustrating the differences between a normal (A) and diabetic (B) nephron due to renal haemodynamic changes

In the diabetic kidney (B) increased levels of sodium reabsorption result in increased GFR, affecting functioning of the kidneys, as compared to a normal functioning kidney (A) (Alicic, Rooney and Tuttle, 2017). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (as amended).

1.3.1.2. Structural changes in the glomerulus

The diabetic-induced hypertrophy of tubules mediating sodium chloride reabsorption, together with elevated intraglomerular pressure results in structural changes in the kidney (Chawla, Sharma and Singh, 2010) (Figure 1.3). The earliest structural changes are mesangial cell matrix expansion and thickening of the glomerular basement membrane (GBM) (Mason and Wahab, 2003; Wolf, Butzmann and Wenzel, 2003; Wolf, 2004). This is due to an increase in the ECM and mesangial cell hypertrophy. This increase in mesangial cell matrix is connected to a decrease in GFR (Mauer *et al.*, 1984; Wolf, 2004; Nabrdalik *et al.*, 2013). The expansion of mesangial cells is a result of arrestment of the cells in the G1 cell cycle phase in the presence of high glucose levels (Wolf *et al.*, 1997; Wolf, 2004).

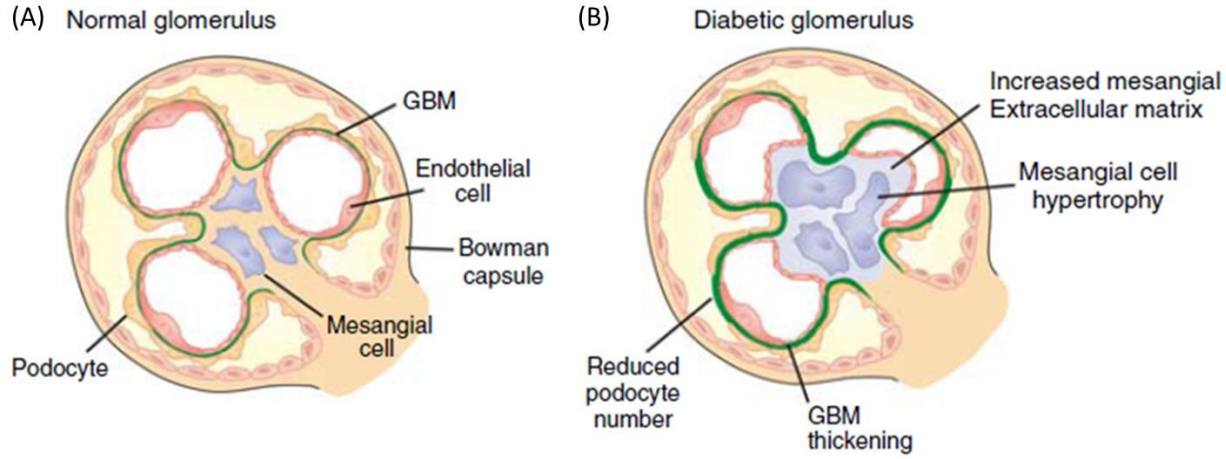


Figure 1.3: Diagram illustrating the structural changes that occur in the kidney glomerulus in the presence of DM that lead to DN

An affected kidney (B) presents with thickened GBM, a decreased number of podocytes and increased mesangial matrix expansion, as compared to a health kidney structure (A) (Jefferson, Shankland and Pichler, 2008). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (as amended).

1.3.1.3. Podocyte loss

The glomerular epithelial cells (podocytes) line the urinary side of the GBM and function to allow size-selective filtering of molecules under normal conditions (Mason and Wahab, 2003; Li *et al.*, 2007). Structural and functional changes of the podocytes occur early in DN progression (Mason and Wahab, 2003). Podocytes attach to the GBM via integrin and dystroglycans (Shankland, 2006; Garg, 2018).

Hyperglycaemia results in a dysregulation of integrin expression, leading to detachment of podocytes from the GBM (Tsilibary, 2003; Shankland, 2006) and decreased renal function (Hostetter, 2003; Chawla, Sharma and Singh, 2010). The remaining podocytes increase in size, similar to mesangial cell hypertrophy, and this mechanical stretch increases the intraglomerular capillary pressure (Petermann *et al.*, 2005; Xu *et al.*, 2005).

The continued loss of podocytes over time (Wolf, 2004) results in additional thickening of the GBM due to the change of metabolism within the ECM (Shankland, 2006), initiating the development of glomerulosclerosis (Kriz, Gretz and Lemley, 1998), protein leakage across the bare GBM (Kriz, Gretz and Lemley, 1998; Tryggvason and Wartiovaara, 2001), and later tubulointerstitial fibrosis (Cooper, Gilbert and Epstein, 1998).

1.3.2. Diagnosis and monitoring of DN

The presence of albumin in urine is commonly used to assess renal function (Li *et al.*, 2007). Microalbuminuria (MA; or moderately increased albuminuria) is the earliest clinically detectable marker of kidney damage (Lee and Lam, 2015). An urine albumin excretion rate greater than 30 mg/day is necessary for a diagnosis of DN (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017).

Urine albumin excretion can either be expressed in per day (over 24-hours; 24h) measurements (mg/day), or as an urinary albumin-to-creatinine ratio (UACR; mg/mmol) (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). The level of albumin excretion measured is then classified as either normoalbuminuria (< 30 mg/day or UACR < 2 mg/mmol), MA (30-300 mg/day or UACR 2-20 mg/mmol), or macroalbuminuria (> 300 mg/day or UACR > 20 mg/mmol) (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). The 24h method is regarded as the gold standard; however, the urine collection process is laborious, time consuming, and prone to inaccuracies. The UACR measurement is based on spot urine sample and has been proven to be an accurate and time-effective alternative to the 24h method. Thus, the UACR is often the preferred method to monitor DN development in the clinical setting (Teimoury *et al.*, 2014). To confirm the diagnosis and monitor disease progression, it is recommended that the tests take place on two separate occasions within a three to six month period to account for variability in albumin excretion (Gross *et al.*, 2005). A disadvantage of using albumin as a tool for measuring kidney function is that it is influenced by the hydration status of the patient, exercise, blood pressure, infection, time of sampling, and the use of certain

medications, such as angiotensin converting enzyme (ACE) inhibitors and angiotensin receptor blockers (Ma, 2016). In addition, albumin measurements have a low sensitivity and specificity, therefore, MA can present without the presence of kidney damage (Lee and Lam, 2015). MA can also be completely absent in the presence of DN or other kidney diseases (Campion, Sanchez-Ferras and Batchu, 2017). It has been estimated that only about 30 % of patients with MA will develop DN after a 10 year follow-up period (Rossing, Hougaard and Parving, 2005; Retnakaran *et al.*, 2006; Motawi *et al.*, 2018).

An alternative screening method for kidney damage is the measurement of serum creatinine for the estimation of GFR (eGFR) to quantify renal function. A number of equations have been developed to calculate eGFR. The National Health Laboratory Service (NHLS) Chemical Pathology laboratory at the Charlotte Makexe Johannesburg Academic Hospital (CMJAH) use two equations when reporting eGFR; namely, the modification of diet in renal disease (MDRD) formula (Rule *et al.*, 2006) and the chronic kidney disease epidemiology collaboration (CKD-EPI) equation (Silveiro *et al.*, 2011). Reporting of eGFR using the CKD-EPI equation was introduced in March 2020. Both methods rely on age, sex, serum creatinine measurement, and race (Rule *et al.*, 2006), however the CKD-EPI method is often used to quantify eGFR values greater than 60 ml/min/1.73m^2 (Silveiro *et al.*, 2011). The MDRD equation was only validated in individuals with impaired kidney function (< 60 ml/min/1.73m^2), whereas the CKD-EPI equation was validated in a cohort with normal and impaired kidney function, resulting in the latter equation being more accurate at eGFR > 60 ml/min/1.73m^2 (Levey *et al.*, 2009).

An eGFR below 60 ml/min/1.73m^2 using the MDRD equation is required for a diagnosis of DN (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). This method is limited however, as the true filtration rate may be altered by the use of medications or lifestyle differences that are not taken into account in the calculated estimation (Poortmans *et al.*, 2013). eGFR is only used for prolonged monitoring of kidney function, but does not add any advantage for extremely compromised, advanced kidney damage, or dehydration (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). Furthermore, the CKD-EPI equation has been found to underestimate the GFR in patients with T2DM (Silveiro *et al.*, 2011).

Early detection of DN is critical for effective treatment, as well as for possible reversal of kidney injury and prevention of further injury (Jia *et al.*, 2016). The current screening methods, however, are not robust, and therefore not useful as predictive markers of DN as glomerular damage has often occurred by the time a clinical diagnosis is made (Jia *et al.*, 2016). Thus,

alternative biomarkers have been investigated in attempts to identify novel markers of early kidney disease.

1.3.3. Alternative biomarkers

Multiple biomarkers of early kidney damage in serum and urine have been investigated. The classification of these markers is based on the different mechanisms or structural damage they represent (Colhoun and Marcovecchio, 2018); such as biomarkers of glomerular damage, tubular dysfunction, oxidative stress, inflammation, and endothelial damage.

1.3.3.1. Serum cystatin C

A potential marker of DN is serum cystatin C (SCC). Cystatin C is a protein produced in all nucleated cells, which is unrestrictedly filtered within the glomeruli, and completely metabolised by the proximal renal tubes (Lee and Lam, 2015). SCC measurements allow monitoring of GFR without the need to account for weight, gender, and protein intake (Takir *et al.*, 2016), as are needed for serum creatinine based eGFR equations (Krolewski *et al.*, 2012). SCC thus allows a more accurate picture of the eGFR and renal function. However, raised levels of SCC do not precede the pathophysiological manifestation of DN (Takir *et al.*, 2016). In addition, while elevated SCC levels have been recorded before the onset of detectable MA, SCC concentration increases with age, which may affect the accurate diagnosis of DN (Takir *et al.*, 2016).

Although SCC may be a more sensitive detection method for early kidney disease, it may also be elevated in other DM complications, lowering its specificity for DN (Krolewski *et al.*, 2012).

1.3.3.2. Biomarker panels

In addition to studies that have investigated the utility of individual biomarkers, researchers are also investigating the ability of multiple biomarkers, in a single assay, to increase prediction of renal function decline or other kidney disease-related phenotypes (Colhoun and Marcovecchio, 2018). The “surrogate markers for micro- and macro-vascular hard endpoints for innovative diabetes tools” (SUMMIT) study series is an example of a study looking at a panel of DN markers to improve prediction (Looker *et al.*, 2015; Colhoun and Marcovecchio, 2018). SUMMIT attempts to identify biomarkers of strong and weak association that could lead to improved detection of kidney function in DM (Looker *et al.*, 2015). However, many identified biomarkers still require further analysis and validation before implementation can be expected

(Colhoun and Marcovecchio, 2018). Some of the biomarkers in the panel include adrenomedullin (ADM), fatty acid-binding protein (FABP), Tamm–Horsfall urinary protein (THP) and TGF-β1 (Looker *et al.*, 2015).

1.3.3.3. TGF-β

TGF-β is a is a 28 kilodalton (kDa) dimeric protein, found in many cells, that regulates a number of biological processes, including cell differentiation and proliferation, immune regulation, apoptosis and inflammation (Qi *et al.*, 2006). TGF-β exists in three isoforms in mammalian cells; namely, TGF-β1, TGF-β2, and TGF-β3. TGF-β1 is the only isoform that circulates in the blood and is the most highly expressed isoform in the kidney (Qi *et al.*, 2006).

TGF-β1 plays a pivotal role in DN progression. Studies have shown that this transcription factor mediates the increased expression of collagen IV, one of the main components of the GBM, thus leading to GBM thickening (Li *et al*., 2007). TGF-β1 may also mediate the dilation of the afferent arteriole by inhibiting calcium absorption, resulting in increased glomerular blood flow which contributes to glomerular hypertrophy and shear-induced vessel damage (Chawla, Sharma and Singh, 2010). Furthermore, TGF-β1 increases nitric oxide (NO) production, further affecting vascular dysfunction (Sharma *et al.*, 2003) by increasing vascular dilation, and increasing intravascular pressure, resulting in vessel damage (Hostetter, 2003; Chawla, Sharma and Singh, 2010). Finally, TGF-β1 has been shown to suppress the glomerular expression of integrin resulting in decreased podocyte adhesion to the GBM and increased apoptosis (Soon Lee, 2013).

As a potential biomarker, increased TGF-β1 expression has been associated with decreased kidney function and increased proteinuria (Do Nascimento *et al.*, 2013; Verhave *et al.*, 2013; Wu *et al.*, 2014a).

A study conducted in 60 participants with T2DM (20 with normoalbuminuria, 20 with MA and 20 with macroalbuminuria) and 20 healthy controls found that serum TGF-β1 levels were significantly higher in participants with macro- and microalbuminuria than the normoalbuminuria and control groups (Shaker and Sadik, 2013). In addition, TGF-β1 was positively correlated with UACR and negatively with eGFR. The authors, therefore, concluded that TGF-β1 could be used as a marker for DN progression (Shaker and Sadik, 2013).

1.3.3.4. miRNAs as alternative biomarkers for DN

An alternative biomarker that has been the subject of increasing attention is micro-ribonucleic acids (miRNAs). miRNAs are a class of small (18-25 base pairs (bp) in length), non-coding ribonucleic acids (RNAs), involved in regulation of cellular processes (Etheridge *et al.*, 2011; Ichii and Horino, 2018), such as cell differentiation, communication and survival (Jia *et al.*, 2016).

miRNAs contain a seed region consisting of 2-8 nucleotides in the five-prime (5‘) end of the miRNA (Ichii and Horino, 2018). The seed region is highly conserved across species and individuals (Etheridge *et al.*, 2011) and is important for recognition of target messenger RNAs (mRNA). One miRNA may have multiple targets (Großhans and Filipowicz, 2008; Suzuki, 2018).

miRNAs function by either regulating mRNA translation post-transcriptionally by binding to the three-prime (3’)-untranslated region (UTR) and preventing transcription, or by initiating degradation of the mRNA molecule (Großhans and Filipowicz, 2008). In addition, miRNAs may bind to the deoxyribonucleic acid (DNA) sequences directly as epigenetic regulatory elements to either promote or repress transcription (Liu *et al.*, 2013; Ichii and Horino, 2018).

1.3.3.4.1. miRNA biogenesis

miRNAs are tissue specific, however, biogenesis does not always take place at the site of action (Großhans and Filipowicz, 2008; Etheridge *et al.*, 2011). miRNA biogenesis (Figure 1.4) starts with transcription of non-coding DNA regions by RNA polymerase II, producing a primary miRNA (pri-miRNA), which forms a double-stranded hair-pin structure shortly after transcription. The pri-miRNA is cleaved by a RNase III Drosha enzyme, to produce a precursor miRNA (pre-miRNA). The double stranded pre-miRNA is further cleaved (removing the loop joining the 3' and 5' arms) within the cytoplasm by RNase III Dicer enzyme into a double-stranded miRNA duplex. The double-stranded miRNA is then incorporated into a protein complex called the RNA-induced silencing complex (RISC), where the guide strand (specific to the gene of interest) is retained and the passenger strand is degraded (Ichii and Horino, 2018; Suzuki, 2018).

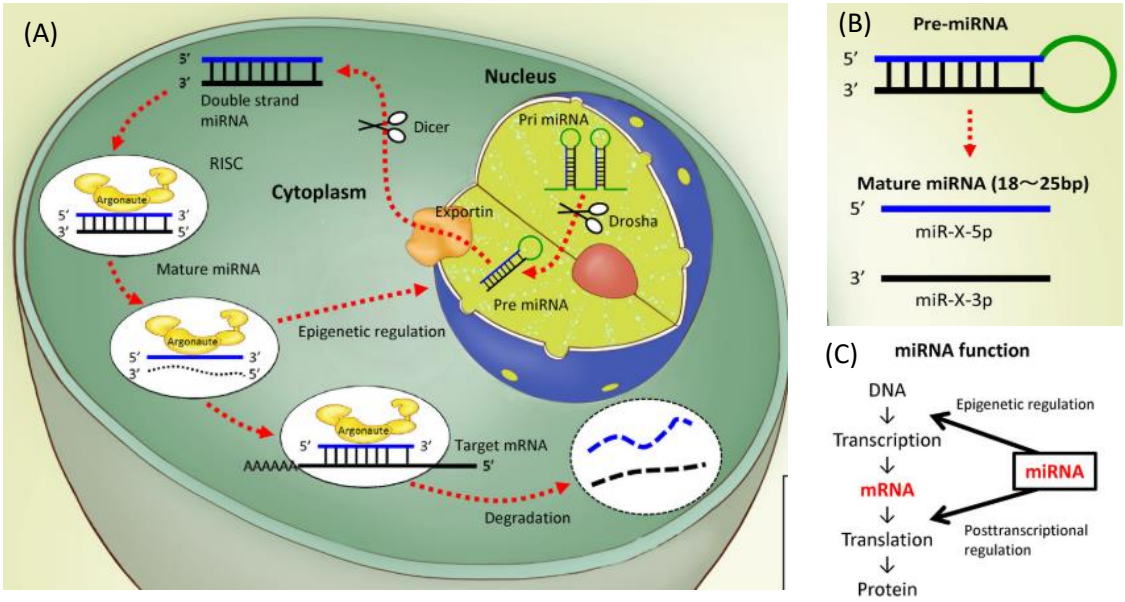


Figure 1.4: Biogenesis and function of miRNAs

(A) Pri-miRNAs are transcribed from non-coding DNA, and cleaved by the Drosha enzyme to produce pre-miRNA. The pre-miRNA is exported to the cytoplasm and cleaved by the Dicer enzyme to produce a double-stranded miRNA complex. The complex is incorporated into the RISC protein, and the mature miRNA is cleaved from the passenger strand. The passenger strand is degraded. (B) The mature miRNA is named according to which arm (the 5' or 3') of the double-stranded complex is the functional miRNA. (C) The mature miRNAs function by either inhibiting transcription (epigenetic regulation) or by preventing translation and initiating degradation of the targeted mRNA (post-transcriptional regulation) (Ichii and Horino, 2018). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (as amended).

Mature miRNAs can function either within the cell where they are transcribed (Großhans and Filipowicz, 2008) or they can circulate in the blood as extra-cellular vesicle associated (EVA) or non-vesicle associated (NVA) miRNAs. Circulating miRNAs are hypothesised to be involved in cell-to-cell communication (Jia *et al.*, 2016), and are more stable and better protected against degradation than non-circulating miRNAs (Simpson *et al.*, 2016).

Circulating miRNAs can be isolated from various body fluids, including urine, plasma, serum, milk, and saliva; making them ideal biomarkers for investigation (Yang *et al.*, 2013).

1.3.3.4.2. miRNAs in DN

miRNAs may be differentially expressed in disease states, and these changes may be detected before physiopathological manifestation of a disease occurs (Etheridge *et al.*, 2011; Blondal *et al.*, 2013; Jia *et al.*, 2016). miRNA expression profiles also vary depending on the sample type and stage of disease (Etheridge *et al.*, 2011; Blondal *et al.*, 2013; Jia *et al.*, 2016).

Urinary miRNA studies in individuals with T1DM and DN have shown that miR-130 and miR-145 were upregulated in patients with MA while miR-155 and miR-424 were down regulated (Barutta *et al.*, 2013). Increased levels of miR-145 were found in a mouse model of DN as well as in cultured mesangial cells when exposed to increasing concentrations of glucose. This points to miR-145 being a possible novel marker of diabetes complications (Barutta *et al.*, 2013).

Three families of urinary miRNAs (miR-21, -200, and -29) were found to have altered expression levels in DN patients (Kato *et al.*, 2011; Wang *et al.*, 2012; Lin *et al.*, 2014). These miRNAs are regulated by TGF-β1 in renal cells and studies have shown that when their expression is normalised, fibrogenesis is reversed (McClelland and Kantharidis, 2014). In addition, urinary miR-29a is disease specific as studies have shown that miRNA-29a expression levels correlated with albuminuria in patients with T2DM (Peng *et al.*, 2013), but was decreased in patients with CKD (Lv *et al.*, 2013). Interestingly, expression levels of urinary EVA miRNAs have been shown to be increased in T2DM patients with MA, but decreased in patients with macroalbuminuria (Jia *et al.*, 2016).

miRNA expression profiles in urine do not correlate with those observed in blood samples (McClelland and Kantharidis, 2014). Additionally, larger amounts of circulating miRNAs are found in blood, specifically serum, than urine (Yang *et al.*, 2013). Thus, the probability of finding novel, robust biomarkers may be greater in serum.

Investigations in serum of T2DM patients revealed reduced levels of miR-23a, -486, -96, -186, -191, -192, and -146a, and let-7i, when compared to pre-diabetic and healthy controls (Yang *et al.*, 2014). When T2DM patients with and without microvascular complications were compared with each other and healthy controls, five serum miRNAs (miR-571, -661, -770-5p, -892b, and -1301) were found to be increased in T2DM patients and were significantly higher in T2DM patients with microvascular complications (Wang *et al.*, 2016a). In addition, a study carried out on human mesangial and proximal tubule epithelial cells treated with high concentrations of glucose or TGF-β1 resulted in downregulation of miR-455. This lead to overexpression of its

target, Rho-associated coiled coil-containing protein kinase two (ROCK2), and subsequent renal fibrosis (Wu *et al.*, 2018).

Studies investigating the role of miRNAs as biomarkers of renal pathologies are promising and provide hope for future treatment options for DN.

1.4. Diabetes mellitus within the South African health system

Due to the growing prevalence of T2DM, there is a large financial burden on both the individual and the health care system (Adeghate, Schattner and Dunn, 2006; Tucci and Akey, 2019). Almost 50 % of the individuals with T2DM fall into the economically productive age group of 40-59 years. In addition, the greatest proportion of deaths due to diabetes occurs in individuals less than 60 years of age (International Diabetes Federation, 2017). With such a large disease burden in adults of working age, enormous economic challenges are created for South Africa. The estimated health care cost to the country is INT$ 1 143 per diabetes case per year (Bertram *et al.*, 2013) which is not sustainable.

Research and development of simple, cost-effective, and early detection methods for DN may reduce the economic burden and improve the quality of life of patients living with T2DM. This can be achieved through early interventions that can delay/prevent progression to later, irreversible stages of the disease. Novel biomarkers may also give insights into novel drug targets and subsequent drug development.

Furthermore, the African genome is known to have greater variability than the genomes of European populations. It is thus of considerable value to investigate the genetic background of diseases in groups of African ancestry in order to gain further insight into the disease aetiology (Tucci and Akey, 2019).

2. Aim and study objectives

The aim of this study was to identify differentially expressed circulating miRNAs that may be used as potential biomarkers for early detection of DN in black South African individuals with T2DM and to investigate genetic variation in their target genes. Furthermore, the study aimed to investigate the role of SCC in estimating GFR for diagnosis and staging of DN.

The objectives for the study were:

1. To compare serum circulating miRNA expression profiles between black South Africans with T2DM with and without MA using microarray analysis.
2. To validate microarray findings through real-time quantitative polymerase chain reaction (qPCR).
3. To determine the genotypic and allelic frequencies of polymorphisms within the target genes of differentially expressed miRNAs.
4. To measure SCC levels via enzyme-linked immunosorbent assays (ELISA).
5. To compare the utility of serum creatinine eGFR equations (MDRD and CKD-EPI) and SCC eGFR equations in diagnosing and staging of DN.

3. Methods

3.1. Study cohort

Black African South African individuals with T2DM (n = 254) were recruited from diabetes clinics at CMJAH and Chris Hani Baragwanath Hospital (CHBH). However, 16 participants were excluded from the study as urine samples for UACR calculations were not available for ten participants, four participants had no serum samples, and buffy coats were missing for two participants. Thus, the final cohort consisted of 238 participants.

Participants were provided with an information sheet and required to sign informed consent (Appendix A) before being enrolled into the study. A questionnaire (Appendix B) regarding the participant's diabetes and complications was completed and anthropometric measurements (height, weight, hip and waist circumference) were taken. In addition, blood pressure (mmHg), glucose (mmol/l), HbA1c (%), serum creatinine (mg/l), and eGFR (ml/min/1.73m^2; if available) results were recorded. Participants with T2DM, regardless of age, duration of disease or presence of co-morbidities were included in this study. Participants with diabetes other than T2DM and participants with overt infections, as noted by the nurse/researcher, were excluded from the study. Ethics was obtained from the Human Research Ethics Committee (Medical) of the University of the Witwatersrand (ethics clearance certificate numbers M140626 and M180647; Appendix C).

3.1.1. Sample processing and storage

Three tubes of blood (one ethylenediaminetetraacetic acid (EDTA) tube, and two serum-separating tubes (SST); BD Vacutainer®, Plymouth, United Kingdom (UK)) and a urine sample were collected from each participant. The three blood sample tubes were spun at 1 200 x g for 10 minutes (min) at room temperature (RT) in a Beckman Coulter Allegra X-22R centrifuge (California, USA). The plasma and buffy coat were removed from the EDTA tube and stored separately at -80 °C in 2 ml cryovials. Serum from the SST tubes was aliquoted into two 2 ml cryovials (Corning, New York, USA). These were snap-frozen in a dry ice and acetone (Merck, Gauteng, South Africa) bath and stored at -80 °C.

The urine samples were tested for overt infection with a urine dipstick (ACON Laboratories Inc., San Diego, USA). Urine (1.5 ml) was aliquoted into a 2 ml cryovial and stored at -80 °C. A

further urine aliquot was sent for urine creatinine and albumin measurement in the Department of Chemical Pathology, NHLS CMJAH Diagnostic laboratory. The samples were run on the Cobas c 502 module auto-analyser (Roche Diagnostics, Illinois, USA).

3.1.1.1. Measurement of urine creatine concentrations

The creatinine measurements were performed using a kinetic colorimetric assay, which is based on the Jaffé method (Roche Diagnostics, Illinois, USA). This method makes use of the measurement of the rate of yellow-orange complex formation when creatinine reacts with picrate in an alkaline solution. “Rate-blanking” was performed to minimise interference from bilirubin, and the absorbance of the solution was read at a wavelength of 570/505 nm on the Cobas c 520 module auto-analyser (Roche Diagnostics, Illinois, USA), and the concentration reported in mmol/l.

3.1.1.2. Measurement of urine albumin concentrations

The albumin concentration (mg/l) was measured using an immunoturbidimetric assay (Roche Diagnostics, Illinois, USA). This assay is based on the turbidimetric measurement (at 340 nm) of a precipitate that forms the urine when albumin interacts with a specific antiserum (polyclonal anti-human albumin antibodies (sheep)) (Roche Diagnostics, Illinois, USA).

The UACR (mg/mmol) was then calculated by dividing the creatinine concentration by the albumin concentration.

3.2. Laboratory practices

The bench and equipment were thoroughly cleaned with nuclease-free 70 % ethanol and RNaseZap® (Ambion, Inc., Foster City, California, USA), before and after experimentation, to avoid any contamination and degradation of samples due to RNases.

All the glassware and microcentrifuge tubes (Eppendorf, Hamburg, Germany) used for miRNA-related work were washed with nuclease-free 70 % ethanol and RNaseZap® and rinsed with nuclease-free water (HyClone™ HyPure™ Molecular Biology Grade Water (GE Healthcare Life Sciences, Utah, USA)). Glassware was baked at 200 °C for two hours and allowed to cool before use. Eppendorf tubes were allowed to dry and any excess water found was removed

manually. Serum miRNA-related laboratory practices made use of PCR clean purity grade, unfiltered, non-autoclaved pipette tips (Eppendorf®, Hamburg, Germany).

3.3. miRNA expression studies

3.3.1. Sample size

A subset of 14 individuals with a disease duration greater than 10 years, were selected for microarray expression analysis. The 10-year disease duration was selected as the cut-off for disease duration as the clinical experience of endocrinologists is that individuals who have not developed kidney disease within 10 years of DM diagnosis, have a low probability of developing any kidney disease in the future (F Raal, personal communication, April 2014).

The 14 samples were divided into two groups. Group 1 (controls; n = 7) included participants without MA (UACR < 1.8 mg/mmol) and group 2 (cases; n = 7) included participants with MA (UACR between 2.2-19.8 mg/mmol measured on two consecutive occasions at least three to six months apart) (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). The cut-offs used were more stringent than the proposed society of endocrine, metabolism and diabetes of South Africa (SEMDSA) guidelines of UACR < 2 mg/mmol for normoalbuminuria, and 2-20 mg/mmol for microalbuminuria, to ensure patients clearly fell within the designated categories and were not on the border of either category. The two groups were matched, as far as possible, in five-year bands for age, in five kg/m^2 bands for BMI, medication use (for a particular complaint), known diabetes complications, and sex. It has been shown that 10-15 biological replicates are sufficient to observe consistent gene expression patterns, with little benefit observed in using more than 15 samples (Pavlidis, Li and Noble, 2003).

3.3.2. miRNA extraction

The miRNA extraction was performed using the Qiagen miRNeasy kit (Qiagen, Germantown, Maryland, USA) adapted from the manufacturer's instructions. All centrifugation steps were carried out in the MIKRO 22R desktop centrifuge (Hettich Zentrifugen, Tuttling, Germany) at RT, unless stated otherwise.

Serum samples were thawed using a heating block at 37 °C for 5 min or until no crystals were visible. An aliquot (200 µl) of serum was transferred to a nuclease-free 1.5 ml microcentrifuge tube placed on ice. All aliquots were kept on ice during the procedure.

QIAzol Lysis Reagent (1000 µl) was added to each sample, mixed by pipetting and incubated at RT for 5 min. Chloroform (200 µl; Minema Chemicals, Gauteng, South Africa) was added and the tubes mixed by vortexing for 15 seconds (sec), followed by centrifugation for 15 min at 12 000 x g at 4 °C. The upper aqueous phase was transferred to a 2 ml collection tube and 900 µl 100 % ethanol (Sigma-Aldrich, Misouri, USA) was added and the tubes mixed by pipetting. Samples (700 µl at a time) were transferred into RNeasy MinElute spin columns placed in 2 ml collection tubes. The tubes were centrifuged at 12 000 x g for 15 sec. The flow through was discarded and the step was repeated with any remaining sample. Kit buffer RWT (700 µl) was added to the spin columns and centrifuged for 15 sec at 12 000 x g. The flow through was discarded and 500 µl Kit buffer RPE added. The tubes were centrifuged for 15 sec at 12 000 x g. The flow through was discarded and 500 µl 80 % ethanol pipetted into the spin column followed by centrifugation for 2 min at 12 000 x g. The spin columns were placed into new 2 ml collection tubes, and centrifuged with open lids for 5 min at full speed (21 910 x g). The spin columns were placed into new 1.5 ml collection tubes. Samples were incubated at RT for 10 min after the addition of 14 µl RNase-free water to the column. Samples were centrifuged for 1 min at full speed. The eluted miRNA was stored at -80 °C until required.

3.3.3. miRNA quantification

RNA integrity and quantity were determined using the Agilent RNA 6000 Pico Kit (Agilent, Santa Clara, California, USA) on the Agilent 2100 Bioanalyzer (Santa Clara, California, USA) according to the manufacturer's instructions.

Ladder preparation: The provided ladder was spun down and denatured in a heating block for 2 min at 70 °C, before being immediately placed on ice. RNase-free water (90 µl) was added to the ladder, mixed thoroughly, aliquoted into 0.5 ml RNase-free vials (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA) and stored at -80 °C. The ladder was thawed on ice before use.

Gel preparation: All samples were equilibrated to RT before use and the dye concentrate protected from light. The gel was prepared before use by pipetting 550 µl RNA gel matrix into a spin filter. The spin filter was centrifuged at 15 000 x g for 10 min, and 65 µl aliquots of filtered gel were placed into 0.5 ml RNase-free microcentrifuge tubes. Tubes were stored at 2-8 °C for up to four weeks.

Preparation of the gel-dye mix: The RNA 6000 Pico dye concentrate was vortexed for 10 sec, spun down, and 1 µl of dye added to a 65 µl gel aliquot. The solution was vortexed and centrifuged at 13 040 x g for 10 min.

Setting up the chip priming station: The Chip Priming Station (Agilent, California, USA) was inspected to ensure the syringe was correctly placed into the luer lock adapter and tightly screwed to the station with the syringe clip in the top position, the base plate set to position "C" and the screws tightly fastened.

Loading the gel-dye mix, conditioning solution and marker: A new pico-chip was placed into the chip priming station, and 9 µl of the prepared gel-dye mix pipetted into the bottom of the well indicated in Figure 3.1 (A). The syringe plunger was set at 1 ml, and the station closed. The plunger was slowly pressed down until it was held by the clip to pressurise the chip. The plunger was released after 30 sec. After a further 5 sec, the plunger was slowly pulled back to the 1 ml position. The priming station was opened and 9 µl of gel-dye mix added to the wells shown in Figure 3.1 (B). The remaining gel-dye mix was discarded. RNA conditioning solution (9 µl) was added to the well shown in Figure 3.1 (C). RNA marker (5 µl) was pipetted into the 11 sample wells and the ladder-well (Figure 3.1 (C)). In addition, 1 µl of the thawed, previously prepared, ladder was added to the ladder-well (Figure 3.1 (D)). Each sample well was filled with 1 µl of sample (Figure 3.1 (E)). The chip was vortexed for 1 min at 2400 revolutions per minute (rpm) on the IKA MS 3 vortexer with a MS 3.4 microtilter adapter (IKA®; Staufen, Germany).

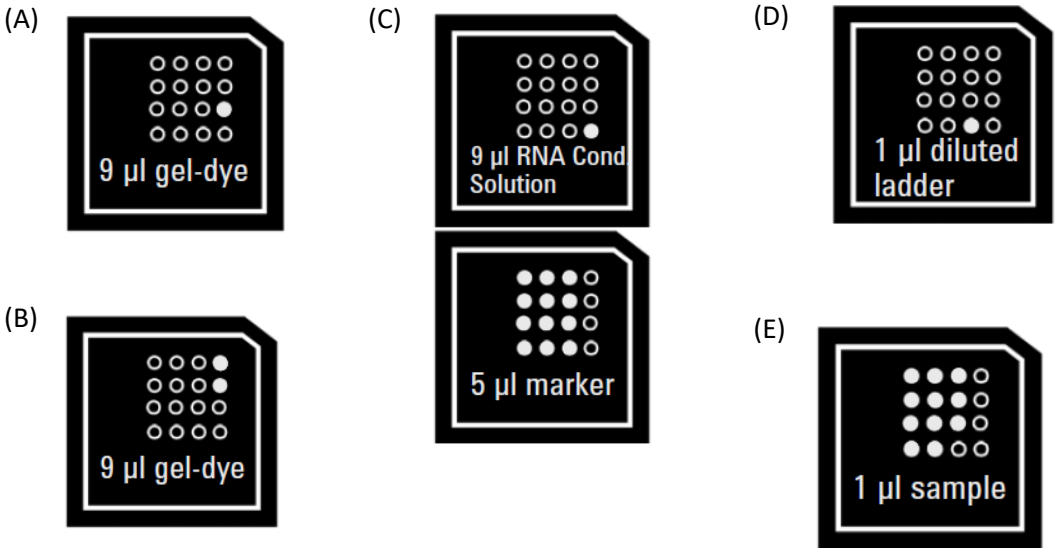


Figure 3.1: The Agilent pico-chip showing the well locations for the different reagents used for miRNA quantification

(A) and (B) indicate the positions for the gel-dye mix before and after the chips is pressurised respectively. (C) The well position for the RNA conditioning solution and the RNA marker. (D) The well position for the diluted ladder, and (E) the well position for each sample (Agilent Technologies, 2016). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (as amended).

The chip was placed into the Agilent 2100 Bioanalyzer instrument and samples run using the "Eukaryotic Total RNA Pico series II" assay. The results of the run were viewed on the 2100 Expert Software (Version B.02.10.S1764, Agilent Technologies).

3.3.4. Labelling of miRNA samples for GeneChip® 4.0 miRNA array analysis

miRNA samples were labelled using the FlashTag™ Biotin HSR RNA Labelling Kit (Affymetrix®, Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA), according to the manufacturer's instructions, with minor modifications.

3.3.4.1. Addition of the poly adenosine (A)-tail

RNA sample (8 µl; 130-1 000 ng) was transferred to a RNase-free tube (on ice) and 2 µl RNA Spike Control Oligos added. The adenosine 5'-triphosphate (ATP) mix was diluted in 1 millimolar (mM) trisaminomethane buffer (Tris; Merck, Darmstadt, Germany; Appendix D) according to the formula: "$Dilution\ factor\ = 5000\ \div input\ RNA\ concentration\ (pg)$".

A poly (A) master mix was prepared (for each sample) in a nuclease-free tube as shown in Table 3.1. Master mix (5 µl) was added to the 10 µl RNA/Spike Control Oligos mixture of each sample. The mixture was mixed gently by pipetting, microfuged, and incubated at 37 °C for 15 min in a heating block.

Table 3.1: Reagents required for the poly (A) master mix per miRNA array

Reagents	Volume (µl)
10 x Reaction buffer	1.5
25 mM Manganese dichloride ($MnCl_2$)	1.5
Diluted ATP mix	1
Phosphatidate phosphatase (PAP) enzyme	1

3.3.4.2. FlashTag™ Biotin HSR ligation

The 15 µl poly (A) tail mixture was spun briefly and placed on ice for the duration of the ligation procedure. FlashTag Biotin HSR Mix (5 x; 4 µl) was added to each sample followed by the addition of 2 µl T4 DNA ligase. The mixture was gently mixed by pipetting, microcentrifuged, and incubated at RT for 30 min. The reaction was stopped by adding 2.5 µl HSR Stop Solution, mixed gently, and microcentrifuged. At this stage, 2 µl of the mixture was removed for use in the enzyme-linked oligosorbent assays (ELOSA) quality control (QC) assay to evaluate the labelling process. The remaining 21.5 µl of the samples were stored on ice prior to hybridisation on the GeneChip® 4.0 miRNA arrays.

3.3.4.3. The ELOSA QC assay

The wells of an ELISA plate (Thermo Fisher Scientific, Massachusetts, USA) were coated with the Affymetrix® ELOSA Spotting Oligos, by diluting 1.5 µl ELOSA Spotting Oligos in 73.5 µl 1 x phosphate buffered saline (PBS; GE Healthcare Life Sciences, Utah, USA) for each well. The 75 µl dilution was added to each well, the plate covered and incubated at 2-8 °C overnight. The following morning, the ELOSA Spotting Oligos were removed by expelling the liquid into a sink. The plate was washed twice with 1 x PBS, 0.02 % Tween-20 mixture (Appendix D) and blotted dry. Following the addition of 150 µl 5 % bovine serum albumin (BSA) in 1 x PBS (Appendix D) to

each well, the plate was covered and incubated at RT for one hour. The BSA blocking solution was expelled from the wells by decanting the liquid into a sink and blotting the plate dry.

A hybridisation master mix was made by mixing 48 µl 5 x saline-sodium citrate (SSC), 0.05 % sodium dodecyl sulphate (SDS), 0.0005 % BSA mixture (Appendix D) with 2.5 µl 25 % dextran sulphate (Appendix D) for each sample. The mixture was vortexed until completely dissolved. 2 µl for each sample was added to 50.5 µl hybridisation master mix. For the positive control, 2 µl of the ELOSA Positive Control was added to 50.5 µl master mix, and for the negative control 2 µl of nuclease-free water was added to 50.5 µl master mix.

The entire hybridisation solution (52.5 µl) was added to a designated well. The plate was covered and incubated for one hour at RT. The hybridisation solution was expelled into the sink and the wells were washed four times with 1 x PBS, 0.02 % Tween-20 mixture, and the plate blotted dry.

Streptavidin horseradish peroxidase (SA-HRP; Thermo Fisher Scientific, Massachusetts, USA) was diluted 1:4 000 in 5 % BSA in 1 x PBS, and 75 µl of the diluted SA-HRP was added to each well, the plate covered and incubated at RT for 30 min. The SA-HRP was expelled into the sink and the plate washed three times with 1 x PBS, 0.02 % Tween-20, and blotted dry. TMB substrate (100 µl) was added to each well. The plate was covered and incubated at RT for 15 min in the dark. A blue colour change in the substrate indicated a positive result.

3.3.4.4. GeneChip® 4.0 miRNA array procedure

The successfully biotin-labelled samples (21.5 µl left after ELOSA QC) were used for the Affymetrix® GeneChip® miRNA array procedure. The Affymetrix® GeneChip® 4.0 miRNA array is a high-density comprehensive tool, that contains full coverage of all miRNAs recorded in miRBase version 20, with probes specific to human, mouse, and rat miRNAs, both mature and pre-mature (Affymetrix, 2017). Before starting, the Affymetrix Hybridisation Oven 645 (Santa Clara, USA) was switched on and set to 48 °C and 60 rpm, the arrays were unwrapped and allowed to equilibrate to RT, and a 200 µl pipette tip (unfiltered) was placed in the upper right septum of the chip to allow for venting when injecting the hybridisation cocktail.

3.3.4.4.1. Hybridisation

All reagents were allowed to equilibrate to RT before use. The 20 x Eukaryotic Hybridisation Controls were thawed and then heated for 5 min at 65 °C in a heating block. A hybridisation cocktail was prepared by adding the reagents listed in Table 3.2 to the 21.5 µl biotin-labelled sample. Samples were incubated at 99 °C for 5 min in a heating block followed by 45 °C for 5 min in a water bath. A total of 130 µl was slowly injected into the array to prevent the formation of air bubbles. The ventilation pipette tip was removed and both septa were covered with ½" Tough-Spots (Star Lab Group, Hamburg, Germany) to minimize evaporation. The arrays were placed into the oven and incubated at 48 °C and 60 rpm for approximately 17 hours.

Table 3.2: Hybridisation cocktail for GeneChip® assay hybridisation for a single chip

Reagents	Volume (µl)
2 x Hybridisation mix	66
27.5 % Formamide	19.2
Dimethyl sulfoxide (DMSO)	12.8
20 x Hybridisation controls	6.6
Control Oligo B2, 3 nM	2.2
Nuclease-free water	3.7
Total volume	110.5

3.3.4.4.2. Washing and staining

After hybridisation, the arrays were removed from the oven and the tough spots removed. The hybridisation cocktail was extracted from the arrays. Each array was slowly filled with 140 µl Array Holding Buffer and equilibrated to RT. The arrays were placed on ice, and transported to the Department of Immunology, School of Pathology, Faculty of Health Sciences, University of Pretoria where the Fluidics Station 450 and the GeneChip® Scanner 3000 7G instruments were located.

Prior to washing the arrays, Wash Buffer A and B were connected to the Fluidics Station 450 (Affymetrix®, Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA), and the priming protocol run.

Once primed, one amber vial containing 600 µl Stain Cocktail 1 was placed into sample holder one. One clear vial containing 600 µl Stain Cocktail 2 was placed in sample holder two and

one clear vial containing 800 µl Array Holding Buffer was placed in sample holder three. The fluidics station pre-loaded protocol was checked to match the desired protocol in Table 3.3. The Tough-Spots were removed from the array and each array was placed into one of the prepared fluidics station modules. The appropriate protocol was selected (Table 3.3) and the wash and stain process was initiated.

Table 3.3: Wash and stain protocol for GeneChip® 4.0 miRNA assay

Steps	Protocol
Post hybridisation wash #1	10 cycles of 2 cycles with Wash Buffer A at 30 °C
Post hybridisation wash #2	6 cycles of 15 cycles with Wash Buffer B at 50 °C
1st Stain	5 min with Stain Cocktail 1 at 35 °C
Post stain wash	10 cycles of 4 cycles with Wash Buffer A at 30 °C
2nd Stain	5 min with Stain Cocktail 2 at 35 °C
3rd Stain	5 min with Stain Cocktail 1 at 35 °C
Final wash	15 cycles of 4 cycles with Wash Buffer A at 35 °C
Array Holding Buffer	Fill with Array Holding Buffer

3.3.4.4.3. Scanning

The arrays were checked for air bubbles and, if present, filled with additional Array Holding Buffer to remove the air bubbles. The septa were covered with $^{3}/_{8}$" Tough-Spots (Diversified Biotech, Denham, USA) and the array glass surface inspected to ensure it was clear of any dust and/or particles. The arrays were then placed into the GeneChip® Scanner 3000 7G (Affymetrix®, Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA) and the scan initiated on the Command Console.

To prevent batch effect, the matched case and control sample chips were washed and stained at the same time.

3.3.4.4.4. Analysis

The Affymetrix® Transcriptome Analysis Console (TAC) 2.0 Software (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA) was used for data analysis, normalisation and QC. The "RMA + DABG (Human Only)" algorithm was used for "Expression (Gene)" analysis as per the manufacturer's instructions.

The algorithm was modified by changing the "Gene-level Fold Change" of interest to be "≤ 2 or ≥ 2". To calculate the expression level differences between the cases and controls, a two-way analysis of variance (ANOVA) was performed.

QC was performed by investigating the Principle Component Analysis (PCA) cluster graph and signal box plots. The PCA graph consists of three components (PCA1, PCA2, and PCA3), with the majority of the variation captured in the PCA1 component (Ambele *et al.*, 2016).

Before data normalisation, the CEL-file PCA and signal box plots were investigated. The CEL-files address performance of the assay probes by expressing probe-level intensity. Any outliers were identified and removed before normalisation was initiated. The PCA cluster graphs indicate variations between sample probe intensity profiles, while the box plots identify any differences in probe intensity measurements across the assays.

Following data normalisation, QC was performed using the CHP-files. The sample ID and classification (case or control), were used as the basis for comparison and normalisation. The PCA and signal box blots of the CHP-files express gene-level information, and any outliers can be removed from further analysis if the gene-level information is significantly different. Further QC includes inspection of the 3' and 5' hybridisation, and the spike-in controls. In order to pass the QC, a signal presence of > 70 % is expected. An assay is also only deemed usable if the 3' and 5' hybridisation controls have passed.

3.4. Real-time qPCR validation

Selected miRNAs with a greater than two-fold difference in expression levels between cases and controls were validated in a subset of 40 participants using real-time qPCR. Samples were selected based on those in the cohort with the highest UACRs (≥ 18.3 mg/mmol; n = 20) and those with the lowest UACRs (≤ 4.0 mg/mmol; n = 20). Two miRNAs (let-7b-5p and miR-455-3p) were selected for real-time qPCR validation based on their target genes and the pathways affected. Target genes were identified using TargetScan (Agarwal *et al.*, 2015), the miRNA data bases and search engines, miRDB (mirdb.org: Liu and Wang, 2019; Chen and Wang, 2020), and miRBase (mirbase.org: Griffiths-Jones, 2004; Griffiths-Jones *et al.*, 2006; Griffiths-Jones *et al.*, 2008; Kozomara and Griffiths-Jones, 2011; Kozomara and Griffiths-Jones, 2014; Kozomara, Birgaoanu and Griffiths-Jones, 2019), and from a review of the literature (Alkayyali *et al.*, 2013; Ma, 2016; Mou *et al.*, 2016).

3.4.1. miRNA extraction

miRNA was extracted from serum using the Qiagen miRNeasy kit as described in section 3.3.2.

3.4.2. cDNA synthesis

Complementary DNA (cDNA) synthesis was performed using the TaqMan® MicroRNA Reverse Transcription kit (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA) according to the manufacturer's instructions. The kit reagents were thawed on ice and a master mix prepared with the reagents listed in Table 3.4. The master mix was gently mixed and centrifuged and placed on ice until use. Master mix (7 µl) was added to a 0.2 ml polypropylene tube containing 5 µl RNA. The mixture was gently mixed, microcentrifuged, and placed on ice.

Table 3.4: Reagents and the corresponding volumes required for cDNA synthesis

Reagent	**Volume (µl)**
dNTP mix (100 mM total)	0.15
Multiscribe™ RT enzyme (50 U/µl)	1
10 x RT Buffer	1.5
RNase inhibitor (20 U/µl)	0.19
Nuclease-free water	4.16
Total volume	7

Custom made primers designed by Applied Biosystems (Massachusetts, USA), specific for the miRNA of interest (let-7b-5p and miR-455-3p), were thawed on ice and briefly centrifuged before transferring 3 µl to each sample tube. Separate reactions were prepared for each miRNA of interest. Tubes were gently mixed and centrifuged. The final 15 µl reaction was placed on ice and incubated for 5 min. A negative reverse transcriptase control was included to control for genomic DNA contamination from the RNA extraction. Samples were then transferred to a T100™ Thermocycler PCR machine (Bio-Rad; California, USA), and amplified using the conditions listed in Table 3.5.

Table 3.5: Thermocycling conditions for cDNA synthesis

Step	Time (min)	Temperature (°C)
Hold	30	16
Hold	30	42
Hold	5	85
Hold	∞	4

3.4.3. Real-time qPCR

Relative quantitation of miRNA expression levels was performed using the TaqMan® Universal PCR Master Mix II, with no uracil-N-glycoslyase (UNG) and the TaqMan® Custom TM Assay (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA) on the 7500 Real-time PCR system (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA). The custom assay consists of a mixture of the forward primer, the reverse primer and a 5' FAM dye-labelled TaqMan® minor grove binding (MGB) probes specific to let-7b-5p and miR-455-3p target sequences, respectively. A master mix was prepared, for each miRNA of interest, with the reagents listed in Table 3.6. The master mix was mixed gently and placed on ice.

Table 3.6: Master mix reagents for the real-time qPCR reaction

Reagents	Volume (µl)
TaqMan® Universal PCR Master Mix II, no UNG	10
TaqMan® Custom TM Assay (20 x)	1
Total volume	11

cDNA (9 µl) was added to 11 µl master mix aliquoted into 0.2 ml PCR tube-strips (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA). Samples were run on the 7500 Real-time PCR system using the following conditions: initial denaturation for 10 min at 95 °C, and 40 cycles of denaturation for 15 sec at 95 °C and annealing/extension for 60 sec at 60°C. Signal measurements were acquired at the annealing/extension step. Results were analysed using the 7500 Real-time PCR system software (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA).

3.4.4. Real-time qPCR using EvaGreen

Despite using custom designed assays for the miRNA expression analysis, no amplification was detected for the miRNAs of interest (Figure 4.6) despite the presence of PCR products when examined by agarose gel electrophoresis (Figure 4.7), which were confirmed to be the correct fragment by sanger sequencing. It was, therefore, determined that there was probe failure. To compensate for this probe failure, the qPCR was repeated with the addition of 0.5 µl 50 x EvaGreen (Bioline, Memphis, USA). The cycling conditions were kept the same but the number of cycles were increased to 60 for let-7b-5p and miR-455-3p. Samples were run in the Corbett Rotor-Gene™ HRM real-time PCR 6000 equipped with a 72 tube-rotor (Corbett Research, Sydney, Australia), and 0.1 ml Rotor-Gene™ tubes (Celtic Molecular Diagnostics, Cape Town, South Africa) were used. The programme was set to acquire in the "Green" channel (530 nm) at the end of the annealing/extension step. This was followed by a melt step where the samples were gradually heated from 77 °C to 90 °C in 0.2 °C increments (Appendix E).

A non-template control (NTC) was included in each miRNA qPCR run. The threshold was manually set above the background noise levels and the C_T (the cycle number at which the amplification curve rose above the threshold) was determined for both the let-7b-5p and miR-455-3p PCR reactions for each participant. As the input serum volume for all reactions was 200 µl, this was used as the normalisation criteria and the relative miRNA levels were therefore expressed as C_T per 200 µl serum.

Samples were run on a 2.5 % agarose gel to verify amplification of the miRNAs of interest.

3.5. Genetic studies

The association of single nucleotide polymorphisms (SNPs) in the high mobility AT-hook 2 (*HMGA2*; rs1114167319 and rs1114167320) and *TGF-β1* (rs1800471) genes with kidney function was investigated.

3.5.1. DNA extraction

DNA was extracted from stored buffy coats using the QIAamp® Blood Mini Kit (Qiagen, Germantown, Maryland, USA) according to the manufacturer's instructions.

The Qiagen kit was prepared by adding 130 ml 100 % ethanol to buffer AW1 and 160 ml 100 % ethanol to buffer AW2. The lyophilized protease was resuspended in 5.5 ml Protease

Solvent. Buffy coat (200 µl) was added to 20 µl protease in a 1.5 ml microcentrifuge tube followed by the addition of 200 µl buffer AL. The tubes were mixed by vortexing and incubated at 56 °C for 10 min in a heating block. After incubation, the tubes were briefly centrifuged and 200 µl ethanol added. The mixture was vortexed and briefly centrifuged. The sample (620 µl) was pipetted into the QIAamp Mini spin column in a 2 ml collection tube and centrifuged at full speed (21 910 x g) for 1 min. The flow throw was discarded and 500 µl Buffer AW1 was added. This was centrifuged at 6 000 x g for 1 min and the flow through was discarded. The spin column was placed into a new collection tube, and 500 µl buffer AW2 added. The column was centrifuged at 21 910 x g for 3 min, and the flow through discarded. The column was centrifuged for another minute at full speed to eliminate any buffer AW2 carry over. The flow through and collection tube were discarded and the spin column placed into a new 1.5 ml microcentrifuge tube. Buffer AE (200 µl) was added to the spin column, incubated at RT for 1 min and then centrifuged at 6 000 x g for 1 min to elute the DNA.

DNA samples were stored at -80 °C until required.

3.5.2. Genotyping participants for polymorphisms in the *HMGA2* gene

The genotypes for the two *HMGA2* polymorphisms, rs1114167319 and rs1114167320, were determined by amplification refractory mutation system (ARMS) polymerase chain reaction (PCR) and PCR restriction fragment length polymorphism (RFLP), respectively.

3.5.2.1. ARMS PCR for *HMGA2* rs1114167319 genotyping

A single reverse primer and two forward primers (specific for the cytosine (C) or thymine (T) allele at the 3' end) were designed using Primer3 software (Koressaar and Remm, 2007; Untergasser *et al.*, 2012) (Table 3.7).

Table 3.7: Primer sequences for the *HMGA2* rs1114167319 polymorphism ARMS PCR

Primer	Primer sequence (5'-3')	Manufacturer
Reverse primer	TGCTTACAAAAAAGGTTAACTCGC	Inqaba Biotec™, Pretoria, South Africa
Forward primer (C-allele)	AGAGTCCCTCTAAAGCAGCTC	Inqaba Biotec™, Pretoria, South Africa
Forward primer (T-allele)	AGAGTCCCTCTAAAGCAGCTT	Inqaba Biotec™, Pretoria, South Africa

Two PCR reactions (one for each forward primer) were set up for each sample using the reagents listed in Table 3.8. A NTC was included in every reaction to check for contamination. PCR reactions were set up in a final volume of 20 µl.

Table 3.8: PCR reagents for the amplification of *HMGA2* rs1114167319

Reagent	Volume (µl)	Final concentration
Nuclease-free water	6	-
KapaTaq Ready Mix (2 x; KAPA Biosystems, Indianapolis, USA)	10	1 x
Reverse primer (10 mM)	1	0.5 mM
Forward primer (10 mM)	1	0.5 mM
DNA	2	~ 50 ng/µl
Total volume	20	-

The PCR reaction was carried out in a T100™ thermocycler PCR machine (Bio-Rad, California, USA) according to the conditions described in Table 3.9.

Table 3.9: Thermocycling conditions for the amplification of *HMGA2* rs1114197319

Stage	Temperature (°C)	Time	Cycles
Initial denaturation	95	3 min	1
Denaturation	95	30 sec	
Annealing	55.9	30 sec	35
Extension	72	30 sec	
Final extension	72	1 min	1
Hold	4	∞	1

3.5.2.2. Agarose gel electrophoresis

A 1 % agarose gel was prepared by dissolving 1 g of agarose (SeaKem®, Lonza, Maine, USA) in 100 ml 1 x Tris, boric acid, EDTA mixture (TBE; Appendix F). The mixture was microwaved until completely dissolved and the gel cooled. Ethidium bromide (2 µl EtBr; 10 mg/ml; Sigma-Aldrich, Missouri, USA) was added to the gel, poured into a casting tray with combs, and allowed to set. PCR product (2 µl) was mixed with 0.4 µl Gel Loading Dye Purple, no SDS (6 x; BioLabs®, Massachusetts, USA) and loaded into the wells. A 100 bp ladder (0.5 µl; BioLabs®, Massachusetts, USA) was added to one well to correctly size the PCR products. The gel was run at 120 volts (V) for 50 min in a gel-tank (Cleaver Scientific, Warwickshire, UK) filled with 1 x TBE buffer, connected to a nanoPAC-300 power supply (Cleaver Scientific, Warwickshire, UK). The gels were visualised using the ChemiDoc™ MP Imaging System (Bio-Rad, California, USA). If an individual was homozygous for the C allele, a band of 199 bp would be seen in the PCR reaction using the C allele primer. The PCR reaction with the T allele primer would have no amplification product. Similarly, an individual homozygous for the T allele would have a PCR product in the reaction with the T allele primer, and no product in the PCR reaction with the C allele primer. If an individual was heterozygous for the rs1114197319 polymorphism, a band of 199 bp would be seen in both PCR reactions with the C and T allele forward primer.

3.5.2.3. PCR RFLP for the *HMGA2* rs1114197320 genotyping

A PCR RFLP was designed to genotype participants for the *HMGA2* rs1114167320 polymorphism. Primers were designed using Primer3 software to amplify a 181 bp region of the gene containing the rs1114197320 polymorphism (Table 3.10).

Table 3.10: Primer sequences for the amplification of the *HMGA2* rs1114167320 polymorphism

Primer	Primer sequence (5'-3')	Manufacturer
Reverse primer	TGGCTGCGGAGTTAGAATGG	Inqaba Biotec™, Pretoria, South Africa
Forward primer	TTAGGAACCAACCGGTGAGC	Inqaba Biotec™, Pretoria, South Africa

PCR reactions were set up in 0.2 ml microcentrifuge tubes in a final volume of 20 µl. The reagents used in the PCR reaction are listed in Table 3.11. Each run included a NTC to ensure that there was no contamination.

Table 3.11: PCR reagents for the amplification of *HMGA2* rs1114167320

Reagent	Volume (µl)	Final concentration
Nuclease-free water	6	-
KapaTaq Ready Mix (2 x; KAPA Biosystems, Indianapolis, USA)	10	1 x
Reverse primer (10 mM)	1	0.5 mM
Forward primer (10 mM)	1	0.5 mM
DNA	2	~ 50 ng/µl
Total volume	20	-

The PCR reactions were carried out under the conditions listed in Table 3.12 in the T100™ thermocycler.

Table 3.12: Thermocycling conditions for the amplification of *HMGA2* rs1114197320 polymorphism

Stage	Temperature (°C)	Time	Cycles
Initial denaturation	95	3 min	1
Denaturation	95	30 sec	
Annealing	57.4	30 sec	35
Extension	72	30 sec	
Final extension	72	1 min	1
Hold	4	∞	1

PCR products were run on a 1 % agarose gel and visualised on the ChemiDoc™ MP Imaging System as described in Section 3.5.2.2.

3.5.2.4. Restriction digest of the *HMGA2* rs1114197320 PCR product

Successfully amplified PCR products were digested with AluI (New England Biolabs, Massachusetts, USA) in a final volume of 25 µl (Table 3.13). Samples were digested for 20 min at 37 °C followed by 20 min at 80 °C to inactivate the enzyme.

Table 3.13: Reagents for the digestion of *HMGA2* rs1114167320 PCR products

Reagents	Volume (µl)	Fina concentration
Nuclease-free water	15	-
10x NEBuffer	2.5	1 x
AluI (10 000 U/ml)	0.5	200 U/ml
PCR product	7	-
Total volume	25	-

Digested PCR products (20 µl) were mixed with 6 x Gel Loading Dye Purple (3 µl; BioLabs®, Massachusetts, USA) and run on a 2 % agarose gel at 120 V for 50 min. Gels were visualised on the ChemiDoc™ MP Imaging System. A 50 bp ladder (0.5 µl; BioLabs® Inc., Massachusetts, USA) was used to correctly size the digested PCR products. Participants were genotyped for the rs1114197320 polymorphism based on the fragment sizes produced (Table 3.14).

Table 3.14: Possible genotypes and their corresponding fragment sizes for the *HMGA2* rs1114197320 polymorphism

Polymorphism	Genotype	Fragment size (bp)
rs1114197320 (A>C)	AA	181
	AC	181; 98; 84
	CC	98; 84

3.5.3. Genotyping participants for the *TGF-β1* gene rs1800471 polymorphism

Participants were genotyped for the rs1800471 polymorphism by ARMS PCR using a single reverse primer and two forward primers (specific for the C or guanine (G) allele at the 3' end). Primers were designed using Primer3 software (Table 3.15).

Table 3.15: Primer sequences for the amplification of the *TGF-β1* rs1800471 polymorphism by ARMS PCR

Primers	Primer sequence (5'-3')	Manufacturer
Reverse primer	CGGTCGCGGGTGCTGTTGTACAGG	Inqaba Biotec™, Pretoria, South Africa
Forward primer (C-allele)	GCTACTGGTGCTGACGCCTGGCCC	Inqaba Biotec™, Pretoria, South Africa
Forward primer (G-allele)	GCTACTGGTGCTGACGCCTGGCCG	Inqaba Biotec™, Pretoria, South Africa

Two PCR reactions (one for each forward primer) were set up for each participant using the reagents listed in Table 3.16. Each run included a NTC to ensure that there was no contamination.

Table 3.16: PCR reagents for the amplification of *TGF-β1* rs1800471

Reagents	Volume (µl)	Final concentration
Nuclease-free water	4	-
KapaTaq Ready Mix	7.5	1 x
Reverse primer (10 mM)	0.75	0.5 mM
Forward primer (10 mM)	0.75	0.5 mM
DNA	2	~ 50 ng/µl
Total volume	15	-

All PCR reactions were run in a T100™ thermocycler according to the conditions listed in Table 3.17. PCR products (210 bp) for the C and G allele PCR reactions were run on a 1 % agarose gel and were visualised on the ChemiDoc™ MP Imaging System.

Table 3.17: Thermocycling conditions for the amplification of *TGF-β1* rs1800471

Stage	Temperature (°C)	Time	Cycles
Initial denaturation	95	3 min	1
Denaturation	95	30 sec	
Annealing	66	30 sec	35
Extension	72	30 sec	
Final extension	72	1 min	1
Hold	4	∞	1

3.6. Measurement of serum cystatin C

SCC concentrations (ng/ml) were measured using the R&D Systems Quantikine® ELISA Human Cystatin C kit (Minneapolis, USA) as per the manufacturer's instructions. The serum samples were thawed before use and 20 µl of sample diluted in 580 µl Calibrator Diluent RD 5-24 (30 times dilution).

The Human Cystatin C Standard (200 ng/ml) was reconstituted by the addition of 1 ml deionised water, mixed well, and allowed to stand for at least 30 min with gentle agitation (300 rpm) on the ThermoStar DMG LabTech desktop shaker (Istanbul, Turkey). The human cystatin C

standard was serially diluted (Figure 3.2) with 300 µl of Calibrator Diluent to create concentrations of 100, 50, 25, 12.5, 6.25, and 3.13 ng/ml to be used for the standard curve. The blank consisted of 300 µl diluted Calibrator Diluent RD5-24.

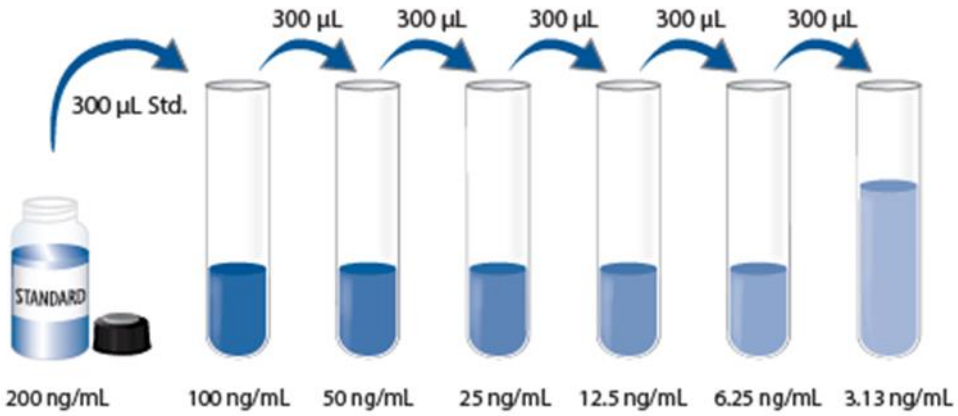


Figure 3.2: Diagram illustrating the serial dilution of the human cystatin C standard

Standard (300 µl) was added to a tube containing 300 µl Calibrator Diluent RD5-24 to give a concentration of 100 ng/ml. This was repeated to generate the six standards for the standard curve (100, 50, 25, 12.5, 6.25, and 3.13 ng/ml) (R&D Systems, 2017). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (as amended).

With the exception of the conjugate which was kept at 2-8 °C, all remaining reagents were brought to RT before use. The microplate was removed from the foil pouch and 100 µl Assay Diluent RD 1-43 was added to each well.

The Quantikine® Immunoassay Control Group 8 sample controls (R&D Systems, Minneapolis, USA) used in the ELISA consisted of three vials of high, medium, and low human cystatin C concentrations (Table 3.18). The controls were reconstituted in 4 ml deionised water and were diluted 30 times with kit diluent, as were the study samples.

Table 3.18: Quantikine® immunoassay control sample concentration ranges

Control	Concentration rage (ng/ml)
High	42.5-84.9
Medium	24.2-39.6
Low	8.02-13.1

Standards, blank, positive controls, or sample (50 µl) were added to their designated wells. The microplate was sealed with an adhesive strip and incubated for three hours at 2-8 °C. The plate was aspirated and washed four times with 400 µl Wash Buffer. Washing took place in the BioTek ELx50 auto washer (Agilent, Vermont, USA). After the wash steps were completed, the plate was inverted and blotted dry against paper towels. Cold (2-8 °C) human cystatin C conjugate was added to each well. The plate was covered and incubated for one hour at 2-8 °C.

After incubation, the wash steps were repeated as stated above. Substrate Solution (200 µl; made from equal parts of Colour Reagent A and B) was added to each well. The plate was sealed with an adhesive cover and incubated at RT, in the dark, for 30 min.

Stop Solution (50 µl) was added to each well and the plate was placed into the Biotek Synergy HT microplate reader (Agilent, Vermont, USA). The plate was read at a test wavelength of 450 nm and a correction wavelength of 540 nm to account for optical imperfections.

For each plate reading a four-parameter logistic standard curve and line-of-best-fit was generated and concentrations were calculated using the Gen5 1.11 software (Agilent, VT, USA). Each recorded concentration was multiplied by the dilution factor to obtain the corrected concentration for each sample.

3.6.1. Calculating eGFR from serum cystatin C concentrations

The SCC based eGFR (ml/min/1.73m^2) was calculated using three previously described equations (eq) listed in Table 3.19.

Table 3.19: Equations to calculate eGFR using SCC concentrations

Method	Equation
SCC eq1 (Bevc *et al.*, 2012)	$eGFR = \frac{100}{SCC\ (mg/l)}$
SCC eq2 (Stevens *et al.*, 2008)	$eGFR = 76.7 \times SCC^{-1.19}$
SCC eq3 (Stevens *et al.*, 2008)	$eGFR = 127.7 \times SCC^{-1.17} \times age^{-0.13} \times (0.91\ if\ female) \times (1.06\ if\ black)$

3.6.2. Calculating eGFR from serum creatinine concentrations

The eGFR (ml/min/1.73m^2) was calculated from serum creatinine concentrations using the modified MDRD and CKD-EPI equations (Table 3.20). The modified MDRD equation does not include the ethnic correction factor of 1.212 estimated for African Americans, because a study by van Deventer and colleagues (2008) showed that the bias was significantly lower when comparing MDRD eGFR to measured GFR with the omitted ethnic factor omitted. These two equations were chosen for the study as they are used for reporting eGFR by the NHLS. Serum creatinine results were available for 185 participants, thus eGFR results were only calculated for these participants, and not the entire cohort.

Table 3.20: MDRD and CKP-EPI equations to calculate eGFR using serum creatinine concentrations

Method	**Equation**
MDRD (Levey *et al.*, 1999)	$eGFR = 175 \times {S_{cr}}^{-1.154} \times age^{-0.203} \times (0.742\ if\ female)$
CKP-EPI (Levey *et al.*, 2009)	$eGFR = 141 \times min(S_{cr}/\kappa, 1)^{\alpha} \times \max(S_{cr}/\kappa, 1)^{-1.209} \times 0.993^{age} \times (1.018\ if\ female)$ κ = 61.9 for females and 79.6 for males α = -0.329 for females and -0.411 for males, min = minimum of S_{cr}/κ or 1 max = maximum of S_{cr}/κ or 1

S_{cr} = serum creatinine in µmol/l

3.6.3. Categorisation of kidney function stage

Based on their eGFR results, participants were categorised into one of five stages of kidney function according to the kidney disease improving global outcomes (KDIGO) guidelines (Table 3.21) (KDIGO, 2013).

Table 3.21: KDIGO stages of kidney function based on eGFR values

Stage	eGFR (ml/min/1.73m^2)	Description
1	> 90	Normal kidney function but urine findings or structural abnormalities or genetic trait point to kidney disease
2	60-89	Mildly reduced kidney function, and other findings (as for stage 1) point to kidney disease
3A	45-59	Moderately reduced kidney function
3B	30-44	
4	15-29	Severely reduced kidney function
5	< 15 or on dialysis	Very severe, or end-stage kidney failure

3.7. Statistical Analysis

Statistical analysis was conducted on Statistica software version 13.5.0.17 (TIBCO Software Inc., California, USA) using R version 3.6.3 (R Core Team, 2018). Normality was determined using skewness values between 1 and -1 (West, Finch and Curran, 1995). Non normally distributed variables were log transformed to reach normality. Continuous variables that were normally distributed were represented as mean ± standard deviation (SD). Continuous variables that were not normally distributed were represented as median [lower; and upper quartiles]. Normally distributed continuous variables were compared using the Student's t-tests (two groups), or by ANOVA (three groups). Variables (SCC) that could not be log transformed to normality were analysed by non-parametric Mann Whitney U tests. Categorical variables were analysed by the Chi-squared ($\chi 2$) test and values presented as frequencies or percentages. Hardy-Weinberg equilibrium (HWE) was determined for all polymorphisms studied. Where genotype frequencies were too low for statistical analysis, the two lowest frequencies were combined. Passing-Bablok linear regression and Bland-Altman plots were generated to evaluate correlation and method agreement, using the MDRD eGFR method as the reference method. For all analyses, $p < 0.05$ was considered statistically significant.

4. Results

4.1. Characteristics of the study population

In total, 238 black African South African participants with T2DM were recruited for the study. The anthropometric and clinical characteristics of the study participants are summarised in Table 4.1. The mean age of the T2DM participants was 56.6 ± 9.88 years with a median [lower and upper quartile] glucose concentration of 7.60 [5.60; 10.5] mmol/l. The majority of participants were female (61.8 %, 147 of 238) and the mean BMI was 32.0 ± 6.14 kg/m^2. The median UACR for the study cohort was 1.89 [0.59; 7.24] mg/mmol with a mean MDRD eGFR of 70.9 ± 24.1 ml/min/1.73m^2. SCC concentrations ranged between 0.001 and 3.69 mg/l with a median concentration of 1.04 [0.81; 1.33] mg/l.

Table 4.1: Anthropometric and clinical characteristics of the study participants

Variable	Total cohort (n = 238)
Age (years)	56.6 ± 9.88
Sex	
% male (n)	38.2 (91)
% female (n)	61.8 (147)
Disease duration (years)	11.5 ± 8.82
BMI (kg/m^2)	32.0 ± 6.14
WHR	0.92 ± 0.08
Glucose (mmol/l)	7.60 [5.60; 10.5]
HbA1c (%)	8.68 ± 2.73
sBP (mmHg)	144.3 ± 21.9
dBP (mmHg)	79.6 ± 13.0
UACR (mg/mmol)	1.89 [0.59; 7.24]
eGFR (ml/min/1.73m^2)	70.9 ± 24.1
SCC (mg/l)	1.04 [0.81; 1.33]

Results are expressed as mean ± standard deviation or median [lower quartile; upper quartile]; BMI = body mass index, WHR = waist-to-hip ratio, HbA1c = glycated haemoglobin, sBP = systolic blood pressure, dBP = diastolic blood pressure, UACR = urine albumin-to-creatinine ratio, eGFR = estimated glomerular filtration rate using the MDRD equation, SCC = serum cystatin C

4.1.1. Classification of participants based on their UACR results

Participants were further classified into normo-, micro-, or macroalbuminuria (based on their UACR results) and their anthropometric and clinical characteristics were compared (Table 4.2). The majority of participants were classified as normoalbuminuria (50.4 %, 120 of 238), with 39.5 % (94 of 238) classified as MA and the remaining 10.1 % (24 of 238) falling into the macroalbuminuria category. Participants with normo-, micro- or macroalbuminuria did not differ significantly by age (56.5 ± 10.3 vs. 56.0 ± 9.94 vs. 59.4 ± 6.70 years, respectively; $p = 0.323$) and BMI (32.6 ± 6.25 vs. 31.4 ± 5.98 vs. 31.5 ± 6.12 kg/m^2, respectively; $p = 0.319$). Participants with macroalbuminuria had a significantly longer duration of disease than the normo- and microalbuminuria groups (16.0 ± 9.02 vs. 11.9 ± 9.27 vs. 10.4 ± 8.21 years, respectively; $p = 0.022$). sBP (139.4 ± 20.8 vs. 146.9 ± 21.5 vs. 158.5 ± 21.5 mmHg, respectively; $p < 0.001$) and SCC concentrations (1.00 [0.80; 1.27] vs. 1.05 [0.80; 1.33] vs. 1.34 [0.99; 1.68] mg/l, respectively; $p = 0.005$) increased significantly across each group. eGFR levels were significantly higher in the normo- and microalbuminuria groups compared to the macroalbuminuria group (74.3 ± 21.6 vs. 72.7 ± 24.1 vs. 46.2 ± 23.6 ml/min/1.73m^2, respectively; $p < 0.001$).

Table 4.2: Anthropometric and clinical characteristics of the study cohort based on their albuminuria classifications

Variable	Classification			p-value
	Normoalbuminuria (UACR < 2 mg/mmol) (n = 120)	**Microalbuminuria (UACR 2-20 mg/mmol) (n = 94)**	**Macroalbuminuria (UACR > 20 mg/mmol) (n = 24)**	
Age (years)	56.5 ± 10.3	56.0 ± 9.94	59.4 ± 6.70	0.323
Sex				
% male (n)	38.3 (46)	38.3 (36)	37.5 (9)	0.888
% female (n)	61.7 (74)	61.7 (58)	62.5 (15)	
Disease duration (years)	10.4 ± 8.21	11.9 ± 9.27	16.0 ± 9.02	0.022
BMI (kg/m^2)	32.6 ± 6.25	31.4 ± 5.98	31.5 ± 6.12	0.319
WHR	0.91 ± 0.09	0.92 ± 0.08	0.92 ± 0.09	0.513
Glucose (mmol/l)	7.45 [5.50; 9.80]	7.50 [5.70; 11.3]	8.45 [6.00; 13.1]	0.270
HbA1c (%)	8.27 ± 2.62	9.03 ± 2.96	9.30 ± 2.01	0.089
sBP (mmHg)	139.4 ± 20.8	146.9 ± 21.5	158.5 ± 21.5	< 0.001
dBP (mmHg)	78.17 ± 12.78	80.6 ± 13.3	83.2 ± 12.0	0.152
UACR (mg/mmol)	0.60 [0.37; 0.96]	5.31 [3.05; 10.4]	35.7 [25.3; 78.3]	< 0.001
eGFR ($ml/min/1.73m^2$)	74.3 ± 21.6	72.7 ± 24.1	46.2 ± 23.6	< 0.001
SCC (mg/l)*	1.00 [0.80; 1.27]	1.05 [0.80; 1.33]	1.34 [0.99; 1.68]	0.005

Results are expressed as mean ± standard deviation or median [lower quartile; upper quartile]; BMI = body mass index, WHR = waist-to-hip ratio, HbA1c = glycated haemoglobin, sBP = systolic blood pressure, dBP = diastolic blood pressure, UACR = urine albumin-to-creatinine ratio, eGFR = estimated glomerular filtration rate using the MDRD equation, SCC = serum cystatin C; *SCC could not be log-transformed thus result were analysed using non-parametric Mann Whitney U-test; remaining p-values calculated using ANOVA and considered significant at $p < 0.05$

4.2. miRNA expression profiling

A subset of participants with (n = 7; defined as cases) and without (n = 7; defined as controls) MA were selected from the larger cohort for miRNA expression profiling. The anthropometric and clinical characteristics of the participants are summarised in Table 4.3. Cases and controls were matched for age (58.3 ± 10.6 vs. 56.4 ± 11.1 years, respectively; p = 0.755), gender (% male: 28.6 vs. 28.6, respectively; p = 1.00), and BMI (35.6 ± 5.13 vs. 32.0 ± 5.27; p = 0.221). The only significant difference observed between cases and controls was, as expected, for UACR (6.37 [5.33; 12.3] vs. 0.65 [0.47; 1.17] mg/mmol, respectively; p < 0.001).

Table 4.3: Anthropometric and clinical characteristics of the 14 participants used for miRNA expression analysis

Variable	Cases (n = 7)	Controls (n = 7)	p-value
Age (years)	58.3 ± 10.6	56.4 ± 11.1	0.755
Sex			
% male (n)	28.6 (2)	28.6 (2)	1.000
% female (n)	71.4 (5)	71.4 (5)	
Disease duration (years)	13.0 ± 4.32	16.1 ±5.90	0.278
BMI (kg/m^2)	35.6 ± 5.13	32.0 ± 5.27	0.221
WHR	0.92 ± 0.11	0.94 ± 0.05	0.545
Glucose (mmol/l)	7.50 [4.70; 16.2]	4.60 [4.10; 10.1]	0.158
HbA1c (%)	10.6 ± 4.22	9.67 ± 3.38	0.668
sBP (mmHG)	151.7 ±31.0	141.6 ± 12.8	0.446
dBP (mmHg)	81.3 ± 15.4	86.6 ± 12.1	0.506
UACR (mg/mmol)	6.37 [5.33; 12.3]	0.65 [0.47; 1.17]	< 0.001
eGFR (ml/min/1.73m^2)	75.8 ± 27.3	77.8 ± 18.3	0.908
SCC (mg/l)	0.89 [0.60; 1.64]	0.93 [0.65; 1.09]	0.877

Results are expressed as mean ± standard deviation or median [lower quartile; upper quartile]; BMI = body mass index, WHR = waist-to-hip ratio, HbA1c = glycated haemoglobin, sBP = systolic blood pressure, dBP = diastolic blood pressure, UACR = urine albumin-to-creatinine ratio, eGFR = estimated glomerular filtration rate using the MDRD equation, SCC = serum cystatin C; p-value calculated using Student t-test and is significant at $p < 0.05$

4.2.1. miRNA quantification

Serum miRNA extractions yielded total miRNA concentrations between 46 and 553 pg/µl. The Agilent RNA 6000 Pico Kit RNA marker peak was present in all the sample runs, however the 18S and 28S ribosomal RNA (rRNA) peaks, also used for QC, were not present in any of the samples' electropherograms (Figure 4.1) or virtual gels (Figure 4.2). The absence of the 18S and 28S rRNA molecules is due to the fact that they are mostly cell-bound (Li *et al.*, 2014) and RNA was extracted from serum samples. Furthermore, RNA extraction was specific for miRNAs, and thus rRNAs would not be present, as they would have been excluded based on their size. RNA quality would normally be determined by the RNA integrity number (RIN) which is calculated by the Agilent Bioanalyzer 2100 using an algorithm that takes the 28S/18S ratio and other characteristics of the electropherogram into consideration. As rRNA molecules were not present in the electropherograms, RIN could not be calculated. Circulating miRNA are relatively stable and protected against degradation (Simpson *et al.*, 2016). The unlikelihood of degradation was supported by the lack of multiple peaks on the electropherogram.

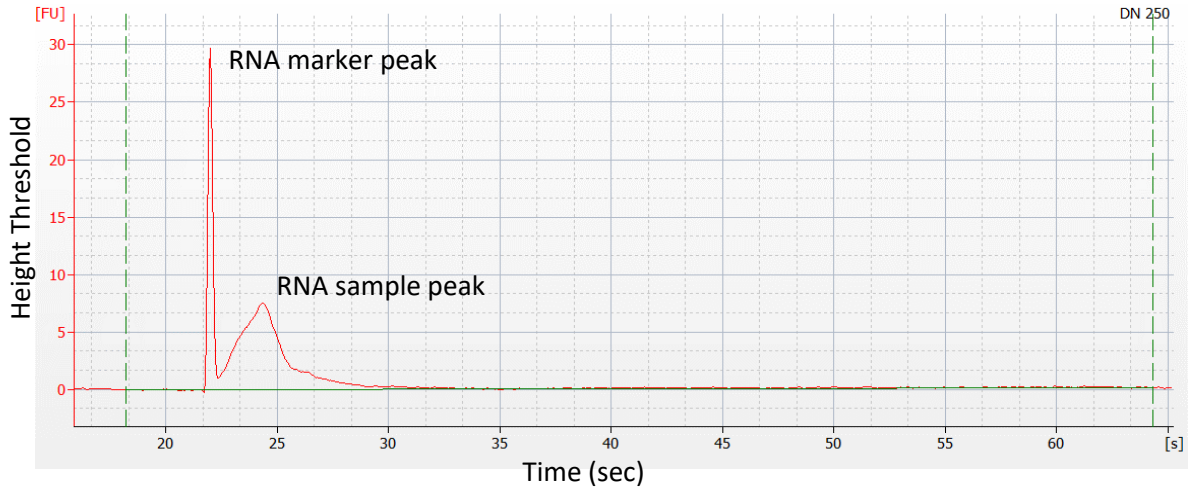


Figure 4.1: Representative Agilent Bioanalyzer electropherogram of a miRNA-extracted sample

The sample presents with the initial RNA marker peak, as well as the sample peak. No peaks were observed for the 18S and 28S rRNAs.

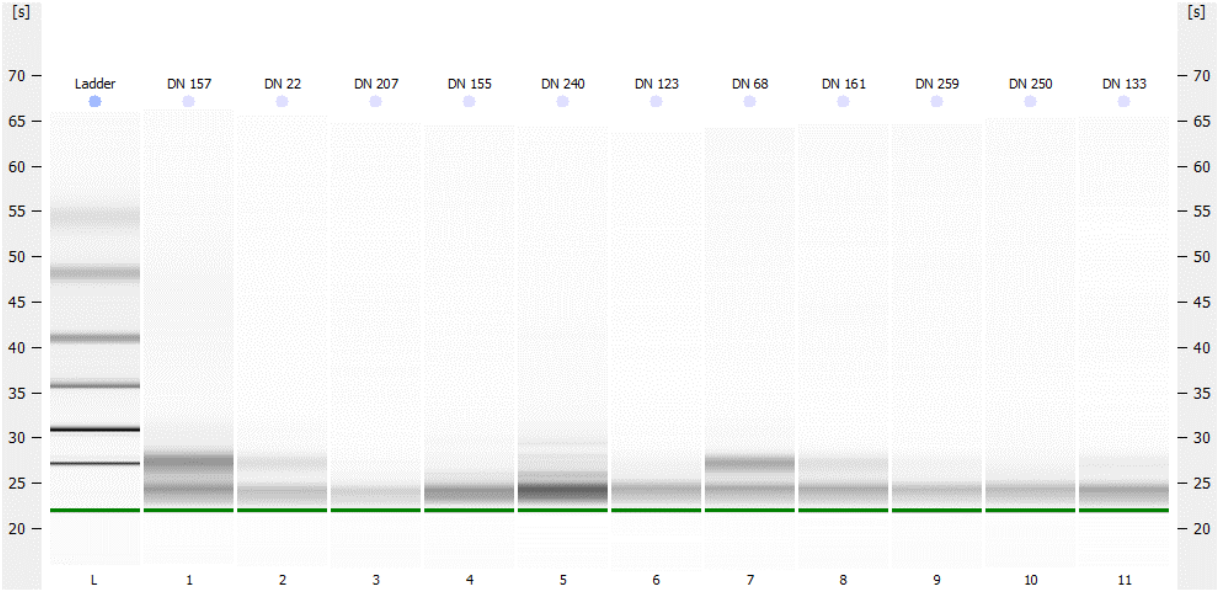


Figure 4.2: Representative Agilent Bioanalyzer virtual gel of miRNA-extracted samples

Lane L: Ladder for size comparison in 5 bp increments.
Lane 1-11: samples run on the chip. Bands on the lower end of the ladder sizing are indicative of small RNA. Bands indicting larger RNAs, such as rRNAs are not present. The green band indicates the presence of the RNA marker.

4.2.2. miRNA profiling QC

Once microarrays were run, QC parameters were investigated in order to determine if the data was valid. This was done using the TAC software both before and after data normalisation. Two graphs were generated for QC analysis, namely the PCA cluster graph and signal box plots.

4.2.2.1. QC before data normalisation

The matched case-control pair PCA cluster graph before normalisation showed some distinction between the case and control samples as evidenced by the clusters formed in the PAC1 plane (Figure 4.3 (A)). Cluster 1 consisted of three control samples, while cluster 2 consisted of two case samples. The majority of the variation can be seen in PCA1 (67.9 %).

When looking at the signal box plots before normalisation (Figure 4.3 (B)), the case and control samples had similar probe-level intensity performance.

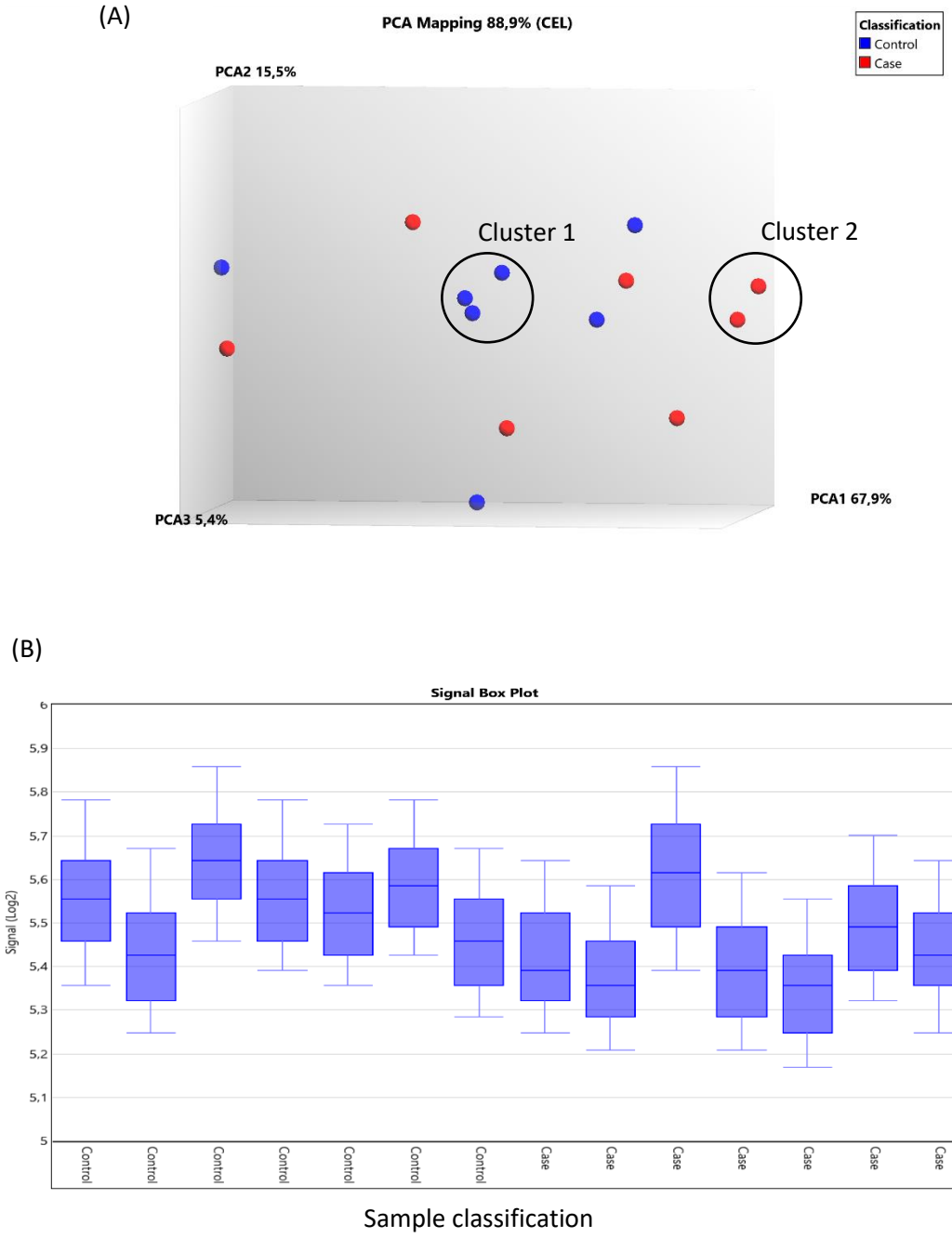


Figure 4.3: The PCA cluster graph and signal box plot QC before data normalisation of the matched pair samples used in microarray analysis

(A) The PCA cluster graph indicates some clustering of controls (cluster 1) and cases (cluster 2). (B) The Signal box blots indicate that there was not much variation in probe signal intensity between case and control samples.

4.2.2.2. QC after data normalisation

The case-control matched pair QC graphs after data normalisation are shown in Figure 4.4. The PCA cluster graph (Figure 4.4 (A)) identified a cluster of matched pairs (cluster 1) suggesting no distinction in gene expression between the case and control samples. The similarities between cases and controls resulting in this clustering could be due to shared expression profiles as a result of their shared DM status, and the differences within the cluster and across the selected pairs could be due to expression differences due to differences between DN status. The majority of variation was present in the PCA1 component (29.3 %).

When addressing the CHP-file signal box blots (Figure 4.4 (B)) very minimal variation could be seen in the gene-level information performance. This indicated that the information from the assays was comparable across the assays.

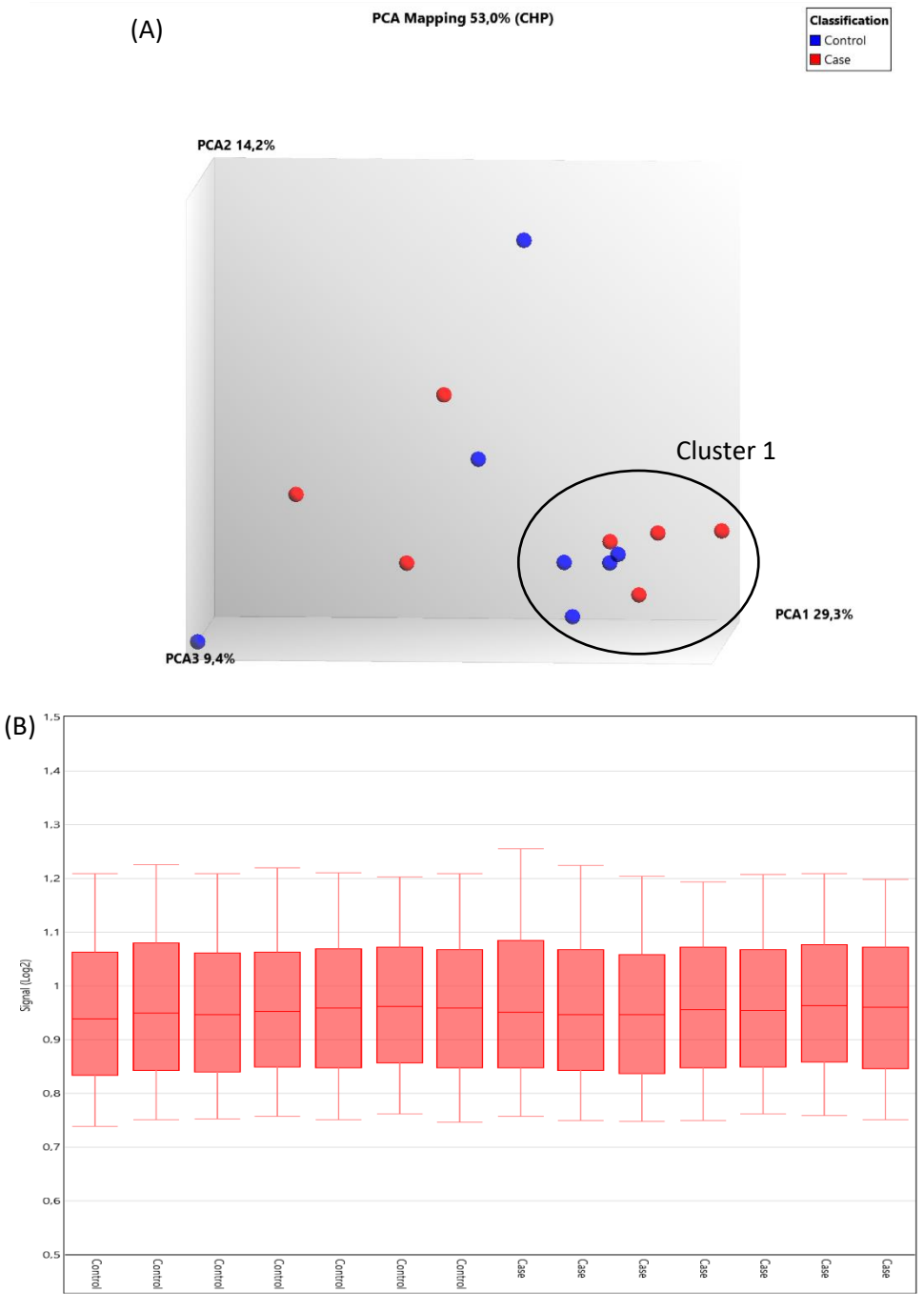


Figure 4.4: The PCA cluster graph and signal box plot QC after data normalisation of the matched pair samples used in microarray analysis

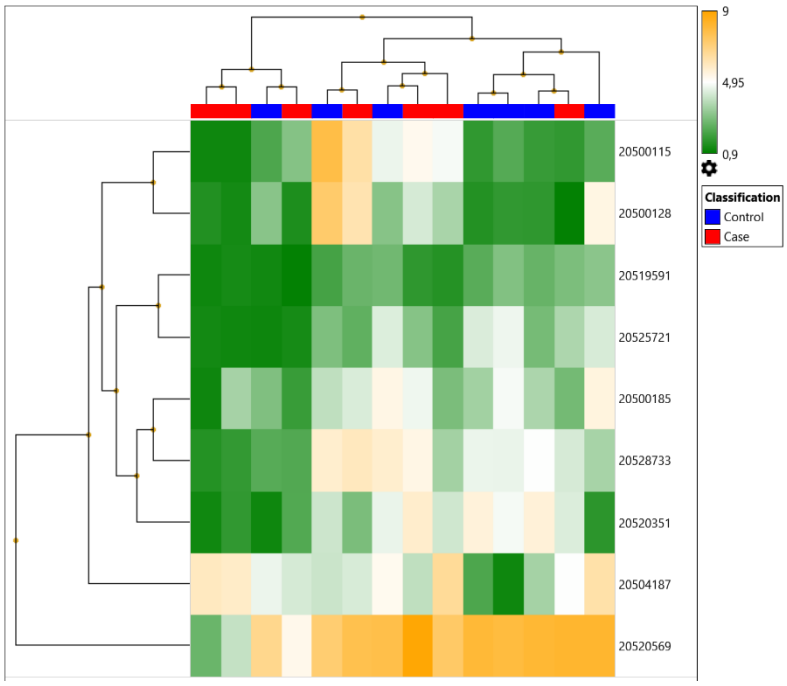
(A) The PCA cluster graph indicated a cluster of matched case-controls (cluster 1). (B) The signal box blot showed no variation in gene level probe intensity between case and control samples.

4.2.3. Identification of differentially expressed miRNA

Once the QC was satisfactory, the miRNA expression profiles were visualised in a heatmap (hierarchical clustering) and a scatter plot (Figure 4.5). The hierarchical clustering indicates sample clusters based on miRNA expression profiles, while the scatter plot is used to identify individual miRNAs with different expression profiles (Ibberson *et al.*, 2009).

Nine miRNAs were differentially expressed between cases and controls (Table 4.4). Negative fold change indicates down regulation of miRNA expression in cases compared to controls, whereas positive fold change indicate upregulation in cases compared to controls.

(A)



(B)

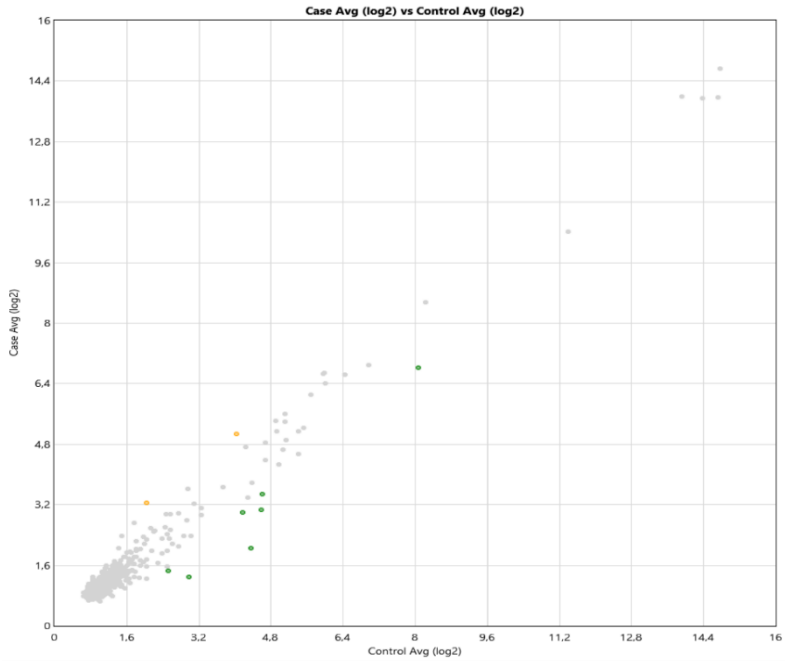


Figure 4.5: miRNA expression profiles of the matched case-control samples

(A) A heatmap representing the miRNA expression profiles in seven participants with normoalbuminuria (blue) and seven participants with microalbuminuria (red). Orange blocks indicate upregulation and green blocks, downregulation of miRNA expression. The intensity (brightness) of the colour indicates the magnitude of the expression difference. White blocks indicate that there is no difference in miRNA expression levels. Each column represents one participant (case or control) and each row represents a different miRNA. (B) A scatterplot indicating changes in expression of the miRNAs of interest (orange = upregulated, green = downregulated, and grey = non-differentially expressed miRNAs).

Table 4.4: miRNAs that were differentially expressed between cases and controls

miRNA ID	miRNA name	Fold change
20500115	let-7b-5p	+2.3
20504187	miR-455-3p	+2.06
20519591	miR-4740-3p	-2.1
20528733	miR-7704	-2.17
20500185	miR-101-5p	-2.25
20520569	miR-5189-3p	-2.35
20520351	miR-1273g-3p	-2.85
20500128	miR-16-5p	-3.24
20525721	miR-6880-5p	-4.95

4.3. Validation of differentially expressed miRNAs

4.3.1. Real-time qPCR validation

Let-7b-5p and miR-455-3p were selected for validation based on their predicted genomic targets (*HMGA2* targeted by let-7b-5p (Appendix G)) and/or due to previous studies showing an association of these miRNAs with DN and/or T2DM (Alkayyali *et al.*, 2013; Ma, 2016; Mou *et al.*, 2016). The expression levels of these two miRNAs were validated in a subset of the study population: 20 T2DM participants with the lowest UACR (≤ 4 mg/mmol) and 20 T2DM participants with the highest UACR (≥ 18.3 mg/mmol). No amplification signal was observed for the amplification of the let-7b-5p and miR-455-3p miRNAs using TaqMan® Custom Assays (Figure 4.6) despite seeing a band of the correct size when the PCR product was run on an agarose gel (Figure 4.7).

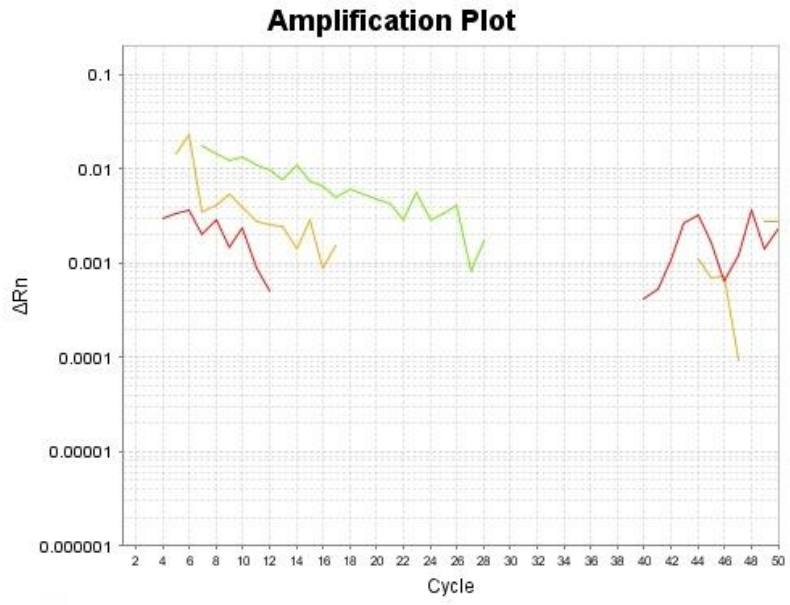


Figure 4.6: The amplification plot of let-7b-5p using the TaqMan® Custom Assay on the 7500 real-time PCR system

Initial qPCR attempts resulted in no signal from the probe-dyes, indicating no amplification of the miRNA in the samples. This is a representation of what was seen following optimisation and troubleshooting of three samples represented by the red, yellow, and green lines.

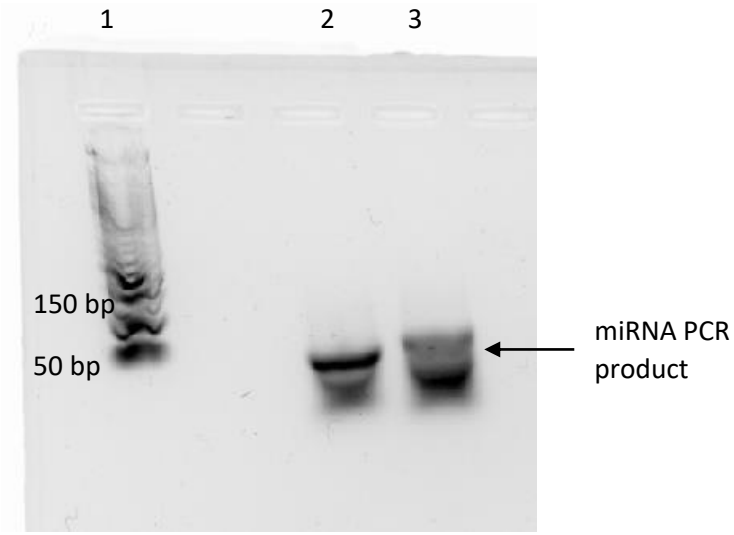


Figure 4.7: qPCR amplification products of let-7b-5p and miR-455-3p run on a 2.5 % agarose gel

Amplification of the target miRNAs were validated by running the samples from initial qPCR attempts on an agarose gel.
Lane 1: 50 bp DNA ladder. Lane 2: let-7b-5p PCR product. Lane 3: miR-455-3p PCR product.

EvaGreen was subsequently used to overcome the probe malfunction. Representative amplification curves, and melt curves of the real time assay PCR products for let-7b-5p and miR-455-3p can be found in Appendix E. A summary of the clinical and phenotypic characteristics of the participants used and the results from the real-time qPCR analysis can be found in Table 4.5. C_T values were used to compare expression levels, the lower the C_T value the higher the expression level. Let-7b-5p relative miRNA expression levels were significantly higher in T2DM participants with low UACR compared to participants with high UACR (26.3 ± 1.50 vs. 27.6 ± 2.03; p = 0.022). Similarly, miR-455-3p miRNA expression levels were significantly higher in T2DM individuals with low UACR compared to individuals with high UACR (40.5 [39.1; 42.8] vs. 42.5 [41.3; 45.3]; p = 0.043). In addition, the participants with the lowest UACR had a significantly shorter disease duration than those with the highest UACR (8.10 ± 6.90 vs. 17.1 ± 9.47 years; p = 0.002).

Table 4.5: Clinical and phenotypic characteristics of study participants and levels of miRNA expression

Variable	Lowest UACR (UACR ≤ 4 mg/mmol) (n = 20)	Highest UACR (UACR ≥ 18.3 mg/mmol) (n = 20)	p-value
Age (years)	54.7 ± 10.5	59.2 ± 6.13	0.106
Sex			
% male (n)	45.0 (9)	25.0 (5)	0.320
% female (n)	55.0 (11)	75.0 (15)	
Disease duration (years)	8.10 ± 6.90	17.1 ± 9.47	0.002
BMI (kg/m^2)	32.5 ± 6.11	30.1 ± 5.47	0.218
WHR	0.91 ± 0.06	0.91 ± 0.09	0.924
Glucose (mmol/l)	8.25 [7.20; 9.80]	8.20 [6.20; 13.0]	0.757
HbA1c (%)	8.77 ± 2.98	9.66 ± 1.65	0.277
sBP (mmHg)	135.8 ± 20.3	151.3 ± 21.8	0.025
dBP (mmHg)	75.7 ± 10.3	79.6 ± 14.4	0.344
UACR (mg/mmol)	0.32 [0.25; 0.36]	27.3 [22.1; 63.3]	< 0.001
eGFR (ml/min/1.73m^2)	76.2 ± 18.7	50.5 ± 23.8	0.004
SCC (mg/l)	0.98 [0.79; 1.19]	1.25 [0.95; 1.60]	0.017
C_T values (let-7b-3p)	26.3 ± 1.50	27.6 ± 2.03	0.022
C_T values (miR-455-3p)	40.5 [39.1; 42.8][a]	42.5 [41.3; 45.3][b]	0.043

Results are expressed as mean ± standard deviation or median [lower quartile; upper quartile]; BMI = body mass index, WHR = waist-to-hip ratio, HbA1c = glycated haemoglobin, sBP = systolic blood pressure, dBP = diastolic blood pressure, UACR = urine albumin-to-creatinine ratio, eGFR = estimated glomerular filtration rate using the MDRD equation, SCC = serum cystatin C; Missing results for C_T-values (miR-455-3p) for [a]lowest UACR n = 10, and [b]highest UACR n = 6; p-value calculated using Student t-test and is significant at $p < 0.05$

4.4. Genotyping participants for the *HMGA2* polymorphisms

Participants were genotyped for the *HMGA2* rs1114167319 and rs1114167320 polymorphism using ARMS PCR and PCR RFLP, respectively. The 199 bp PCR products obtained for the amplification of the rs1114167319 C and T alleles are shown in Figure 4.8. The 181 bp *HMGA2* rs1114167320 PCR product and fragments generated following restriction enzyme digestion with AluI are shown in Figure 4.9.

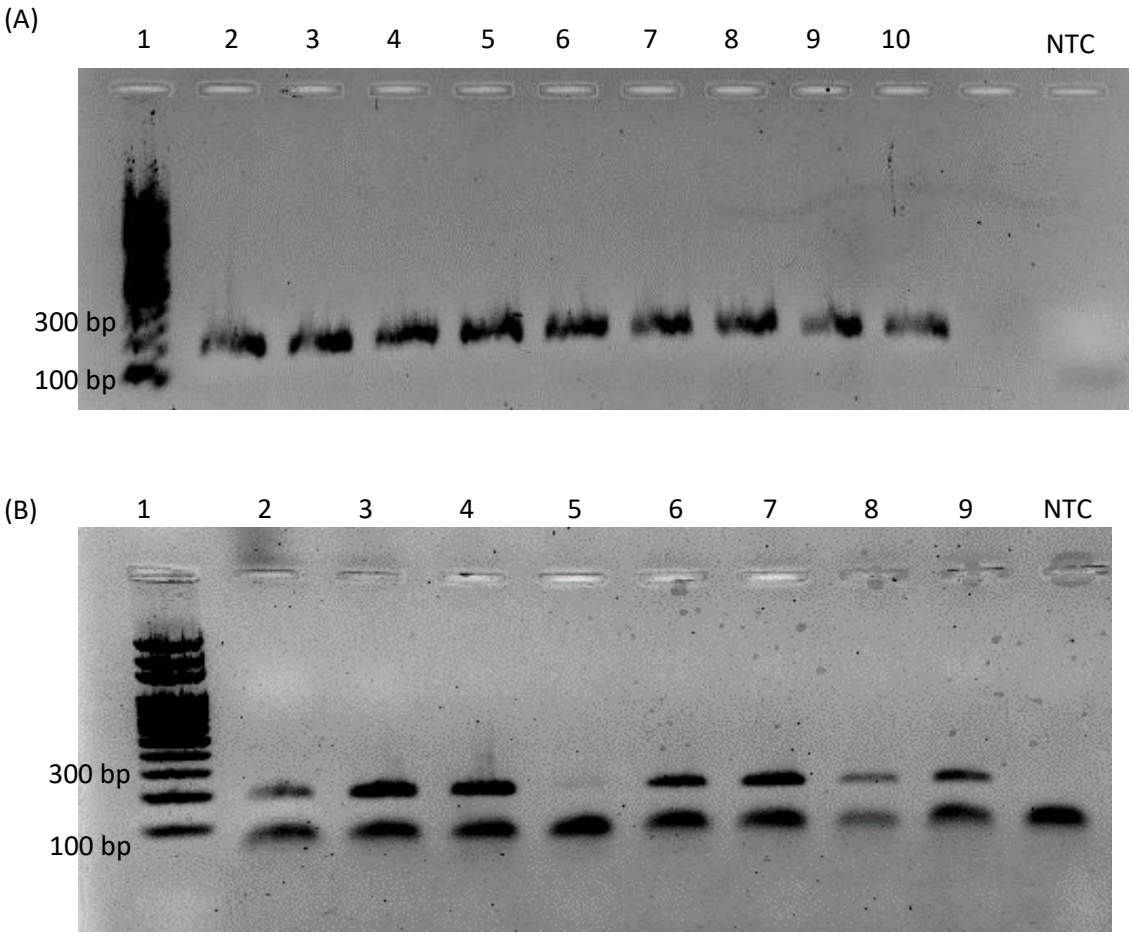


Figure 4.8: PCR products for the *HMGA2* rs1114167319 polymorphism using the C allele primer (A) and T allele primer (B)

(A) Lane 1: 100 bp ladder, Lane 2-10: 199 bp amplicon for the C allele, NTC = non-template control.
(B) Lane 1: 100 bp ladder, Lane 2-9: 199 bp amplicon for the T allele, NTC = non-template control. The smaller second band below the 100 bp mark indicates primer dimer formation.
All participants presented are heterozygous CT for the rs1114167319 polymorphism.

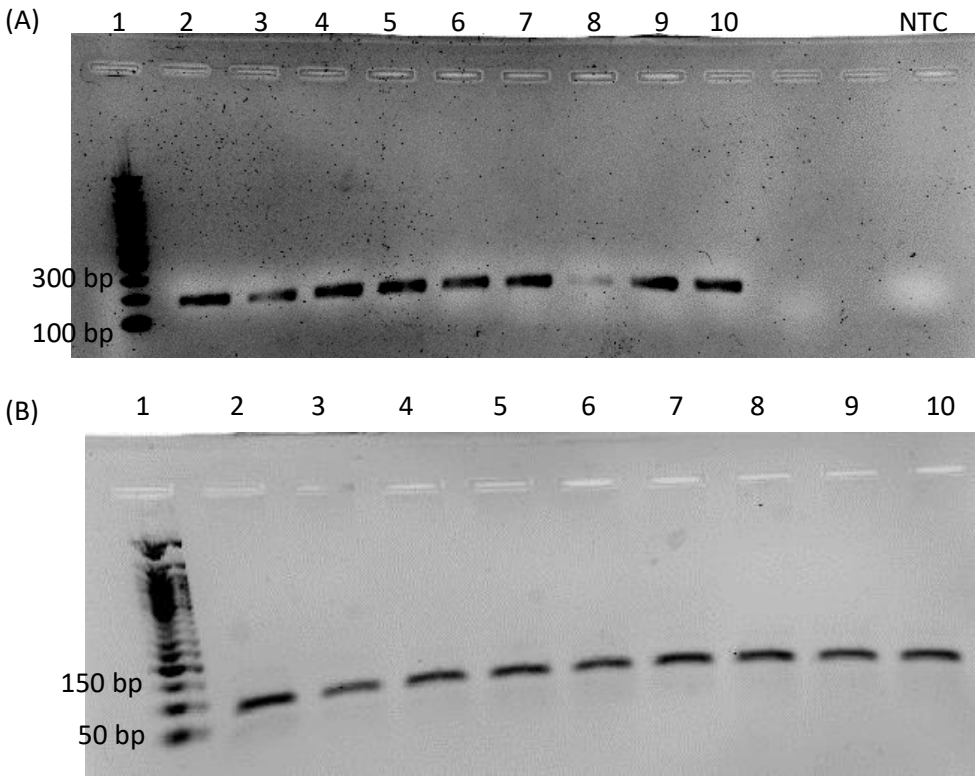


Figure 4.9: PCR products (A) and restriction enzyme digestion profiles (B) for the *HMGA2* rs1114167320 polymorphism run on a 1% and 2% agarose gel, respectively

(A) Lane 1: 100 bp DNA ladder; lane 2-10: 181 bp PCR product; NTC = non-template control.
(B) Lane 1: 50 bp DNA ladder; lane 2-10: participants homozygous for the A allele (98 and 84 bp).

4.4.1. Sanger sequencing of the *HMGA2* polymorphisms

Samples representative of the different rs1114167319 and rs1114167320 polymorphism genotypes were sent for sequencing to confirm the results obtained by PCR ARMS and PCR RFLP. The sequencing results for the *HMGA2* polymorphisms are shown in Figure 4.10.

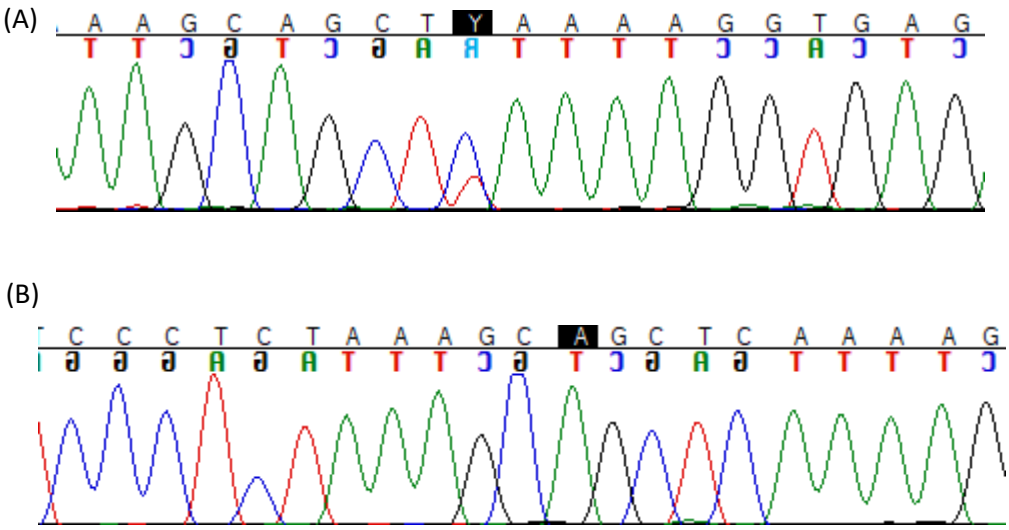


Figure 4.10: Chromatograms obtained from the amplification of the *HMGA2* gene containing the rs1114167319 (A) and rs114167320 (B) polymorphism

(A) Chromatogram of a participant heterozygous CT for the rs1114167319 polymorphism.
(B) Chromatogram of a participant homozygous AA for the rs1114167320 polymorphism.
The black highlighted box indicates the position of the polymorphism in the chromatogram.

4.4.2. Genotypic and allelic frequencies for the *HMGA2* polymorphisms

The study cohort was not in HWE for the rs1114167319 polymorphism (χ^2 = 215.2, p < 0.001; Appendix H) and HWE could not be calculated for the rs1114167320 polymorphism as only one genotype (AA) was identified in the study population. However, for the purpose of this dissertation the genotypic data was still analysed, where possible.

The genotypic and allelic frequencies for both the *HMGA2* polymorphisms are summarised in Table 4.6. The rs1114167319 C allele was the major allele in our population with a frequency of 0.51. The A allele was the major allele for the rs1114167320 polymorphism with a frequency of 1.0.

Table 4.6: Genotypic and allelic frequencies of *HMGA2* gene polymorphisms in black South Africans with T2DM

Polymorphism	Genotype	Frequency (n)	Allele	Frequency (n)
	CC	0.03 (6)	C	0.51 (244)
rs1114167319	CT	0.97 (232)	T	0.49 (232)
	TT	0.0 (0)		
	AA	1.00 (238)	A	1.00 (476)
rs1114167320	Adel	0.0 (0)	del	0.0 (0)
	deldel	0.0 (0)		

del = deletion

4.4.3. The association of *HMGA2* rs1114167319 genotypes with clinicopathological variables in South African Black participants with T2DM

The *HMGA2* rs1114167319 polymorphism genotypes and their association with different clinicopathological variables are summarised in Table 4.7. As no participants had the TT genotype, the comparison was done between participants with the CC genotype and those with the heterozygous CT genotype. No significant differences were seen between the rs1114167319 genotypes and clinicopathological variables and markers of kidney function.

Table 4.7: Associations of the *HMGA2* rs1114167319 polymorphism with clinicopathological variables in the South African black population with T2DM

Variable	*HMGA2* rs1114167319 genotype		p-value
	CC (n = 6)	CT + TT (n = 232)	
Age (years)	59.8 ± 5.13	56.5 ± 9.97	0.442
Sex			
% male (n)	33.3 (2)	38.4 (89)	1.000
% female (n)	66.7 (4)	61.6 (143)	
Disease duration (years)	13.2 ± 10.2	11.2 ± 8.80	0.645
BMI (kg/m^2)	29.4 ± 4.74	32.1 ± 6.16	0.295
WHR	0.93 ± 0.03	0.92 ± .0.009	0.625
Glucose (mmol/l)	7.05 [4.20; 10.1]	7.60 [5.70; 10.5]	0.674
HbA1c (%)	8.15 ± 2.10	8.69 ± 2.74	0.695
sBP (mmHg)	151.0 ± 25.3	144.1 ± 21.8	0.447
dBP (mmHg)	84.8 ± 7.81	79.5 ± 13.0	0.318
UACR (mg/mmol)	0.87 [0.00; 1.00]	2.03 [0.61; 7.75]	0.323
SCC (mg/l)	1.07 [0.92; 1.91]	1.04 [0.80; 1.33]	0.222
eGFR (ml/min/1.73m^2)	70.8 ± 14.7	70.9 ± 24.4[a]	0.986

Results are expressed as mean ± standard deviation or median [lower quartile; upper quartile]; BMI = body mass index, WHR = waist-to-hip ratio, HbA1c = glycated haemoglobin, sBP = systolic blood pressure, dBP = diastolic blood pressure, UACR = urine albumin-to-creatinine ratio, eGFR = estimated glomerular filtration rate using the MDRD equation, SCC = serum cystatin C; [a]Missing results for eGFR for the CT + TT genotype n = 53; p-value calculated using Student t-test and is significant at $p < 0.05$

As all participants were homozygous AA for the rs1114167320 polymorphism, no statistical analysis could be performed.

4.5. Genotyping participants for the *TGF-β1* polymorphism

Participants were genotyped for the *TGF-β1* rs1800471 polymorphism using ARMS PCR. The 210 bp PCR products obtained for the amplification of the rs1800471 C and G alleles are shown in Figure 4.11.

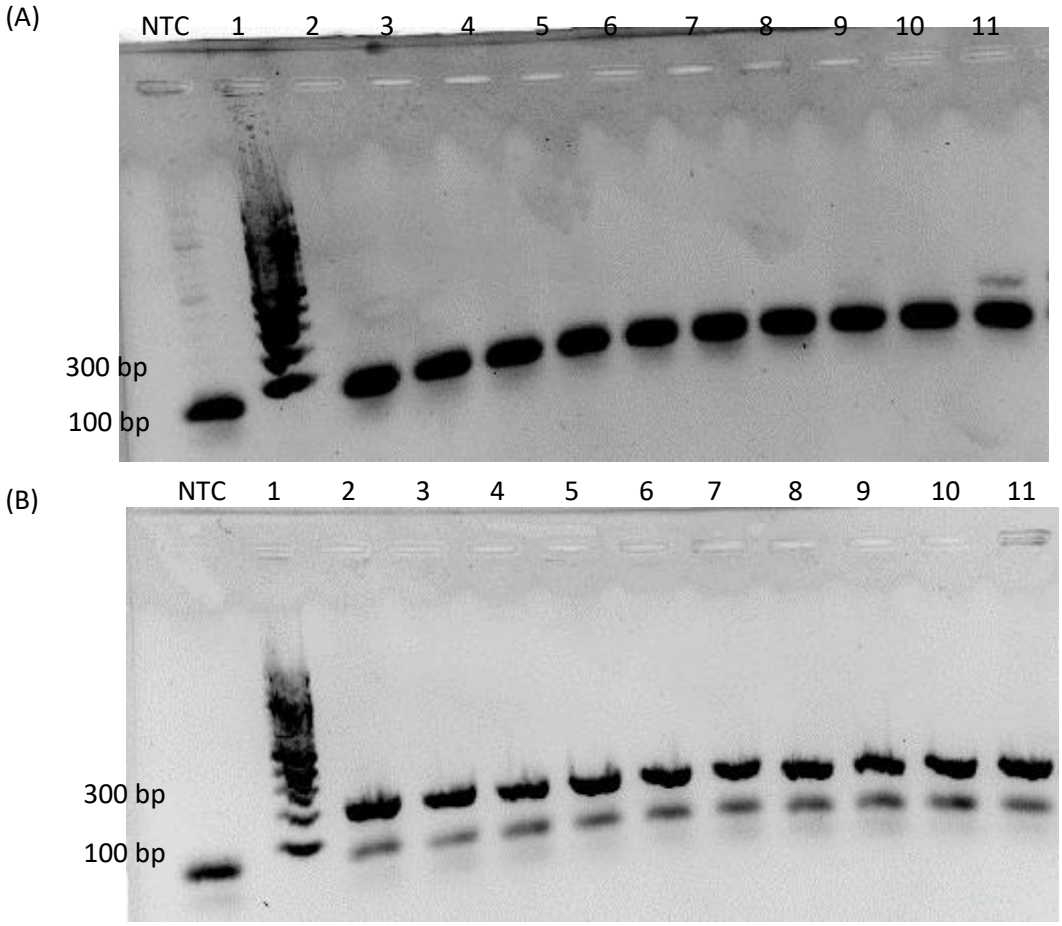


Figure 4.11: *TGF-β1* rs1800471 polymorphism PCR products for the C allele (A) and G allele primers (B) run on a 1 % agarose gel

(A) NTC= non-template control; Lane 1: 100 bp DNA ladder; Lane 2-8 and 10 no amplification for the C allele, Lane 9 and 11: 210 bp amplicon for the C allele.
(B) NTC= non-template control; Lane 1: 100 bp DNA ladder; Lane 2-11: 210 bp amplicon for the G allele. The participants in lane 9 and 11 is heterozygous CG, and the remaining participants are homozygous GG for the rs1800471 polymorphism. Primer dimers were present as a band under the 100 bp DNA ladder mark.

4.5.1. Sanger sequencing of the *TGF-β1* rs1800471 polymorphism PCR product

Samples representative of the different *TGF-β1* rs1800471 polymorphism genotypes were sent for sequencing to confirm the results obtained by ARMS PCR. A representative chromatogram from the sequencing of the *TGF-β1* rs1800471 polymorphism PCR product is shown in Figure 4.12.

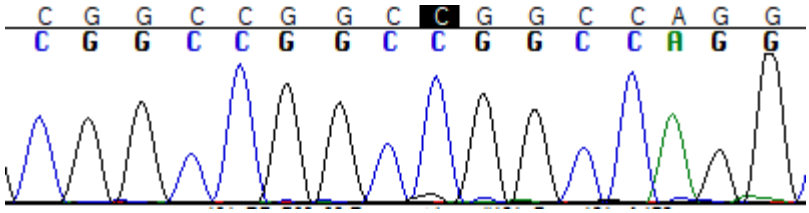


Figure 4.12: Chromatogram obtained from the amplification of a portion of the *TGF-β1* gene containing the rs1800471 polymorphism

Chromatogram of a participant homozygous CC for the rs1800471 polymorphism. The box indicates the position of the polymorphism in the chromatogram.

4.5.2. Genotypic and allelic frequencies for the *TGF-β1* rs1800471 polymorphism

The genotypic frequencies of *TGF-β1* rs1800471 were not in HWE (χ^2 = 63.4, p < 0.001, Appendix H). The genotypic and allelic frequencies were calculated and are summarised in Table 4.8. The G-allele was present in the majority of the participants with a frequency of 0.64.

Table 4.8: Genotypic and allelic frequencies of the *TGF-β1* gene rs1800471 polymorphism in black South Africans with T2DM

Polymorphism	Genotype	Frequency (n)	Allele	Frequency (n)
rs1800471	GG	0.29 (69)	G	0.64 (302)
	GC	0.70 (164)	C	0.36 (168)
	CC	0.01 (2)		

4.5.3. The association of *TGF-β1* rs1800471 genotypes with clinicopathological variables in South African black participants with T2DM

The *TGF-β1* rs1800471 polymorphism genotypes and their association with different clinicopathological variables are summarised in Table 4.9. The heterozygous GC and homozygous

CC genotypes were combined for statistical analysis due to the small number of participants with the CC genotype. No significant differences were seen between the rs1800471 genotypes and markers of kidney function. Participants with the GG genotype had significantly lower glucose levels than those with the GC and CC genotypes (6.60 [4.80; 8.80] vs. 8.10 [5.70; 11.3] mmol/l; p = 0.011).

Table 4.9: Associations of the *TGF-β1* rs1800471 polymorphism with clinicopathological variables in the South African black population with T2DM

Variable	***TGF-β1* rs1800471 genotype**		**p-value**
	CC + GC (n = 166)	**GG (n = 69)**	
Age (years)	56.2 ± 9.50	57.6 ± 10.6	0.338
Sex			
% male (n)	38.6 (64)	36.2 (25)	0.770
% female (n)	61.4 (102)	63.8 (44)	
Disease duration (years)	11.4 ± 8.91	12.3 ± 8.63	0.486
BMI (kg/m^2)	32.1 ± 6.37	31.9 ± 5.37	0.795
WHR	0.91 ± 0.09	0.93 ± 0.08	0.256
Glucose (mmol/l)	8.10 [5.70; 11.3]	6.60 [4.80; 8.80]	0.011
HbA1c (%)	8.66 ± 2.73	8.68 ± 2.67	0.971
sBP (mmHg)	142.9 ± 22.2	148.0 ± 20.7	0.126
dBP (mmHg)	79.8 ± 13.0	79.0 ± 12.7	0.680
UACR (mg/mmol)	1.81 [0.53; 7.12]	2.31 [0.83; 7.37]	0.201
SCC (mg/l)	1.07 [0.81; 1.31]	1.01 [0.78; 1.35]	0.411
eGFR (ml/min/1.73m^2)	71.2 ± 24.1[a]	70.2 ± 24.6[b]	0.777

Results are expressed as mean ± standard deviation or median [lower quartile; upper quartile]; BMI = body mass index, WHR = waist-to-hip ratio, HbA1c = glycated haemoglobin, sBP = systolic blood pressure, dBP = diastolic blood pressure, UACR = urine albumin-to-creatinine ratio, eGFR = estimated glomerular filtration rate using the MDRD equation, SCC = serum cystatin C; [a]Missing results for eGFR CC + GC genotype n = 45; [b]Missing results for eGFR GG genotype n = 8; p-value calculated using the Student t-test and is significant at $p < 0.05$

4.6. Calculations of eGFR

4.6.1. Calculations of eGFR using SCC concentrations

eGFR was calculated from SCC measurements using three different equations and compared to the creatinine based eGFR results calculated using the MDRD and CKD-EPI equations. Table 4.10 summarises the eGFR values obtained for the different equations. A significant difference in the number of participants classified with different stages of kidney function was observed based on the eGFR equation used ($p < 0.001$).

As CKD-EPI equation performs better at eGFR > 60 ml/min/1.73m^2, the eGFR results were split into two groups (< 60 and ≥ 60 ml/min/1.73m^2) and all the eGFR equations compared to the MDRD eGFR results (Table 4.11). Significantly fewer participants were classified as having reduced kidney function when using the CKD-EPI and SCC eq1 compared to the MDRD eGFR equation (6.5 vs 28.1 %; $p < 0.001$ and 12.5 vs. 28.1 %; $p < 0.001$, respectively).

Table 4.10: Comparison of eGFR from equations based on creatinine and SCC concentrations

eGFR (ml/min/1.73m^2)	MDRD (n = 185)	CKD-EPI (n = 185)	SCC eq1 (n = 184)	SCC eq2 (n = 184)	SCC eq3 (n= 181)	p value
Mean	70.9 ± 24.1	110.7 ± 31.4	95.3 [75.1; 119.6]	72.5 [54.5; 94.9]	70.5 [534; 93.9]	
Range	2.77 - 152.5	3.88 - 168.5	27.1 - 278.2	16.2 - 259.2	16.2 - 244.6	
< 15; % (n)	0.5 (1)	0.5 (1)	0.0 (0)	0.0 (0)	0.0 (0)	
15-29; % (n)	3.2 (6)	1.1 (2)	0.5 (1)	4.9 (9)	5.5 (10)	< 0.001
30-44; % (n)	7.0 (13)	2.2 (4)	3.8 (7)	10.9 (20)	12.2 (22)	
45-59; % (n)	17.3 (32)	2.7 (5)	8.2 (15)	16.8 (31)	13.3 (24)	
60-89; % (n)	52.4 (97)	14.6 (27)	29.9 (55)	38.6 (71)	40.3 (73)	
> 90; % (n)	19.5 (36)	78.9 (146)	57.6 (106)	28.8 (53)	28.7 (52)	

p-value calculated using the Chi-squared test, and is significant at $p < 0.05$

Table 4.11: Comparison of MDRD eGFR levels less than and greater than 60 ml/min/1.73m^2 with those obtained from CKD-EPI and SCC equations

eGFR (ml/min/1.73m^2)	MDRD (n = 185)	CKD-EPI (n = 185)	SCC eq1 (n = 184)	SCC eq2 (n = 184)	SCC eq3 (n= 181)
< 60; % (n)	28.1 (52)	6.5 (12)	12.5 (23)	32.6 (60)	30.9 (56)
≥ 60; % (n)	71.9 (133)	93.5 (173)	87.5 (161)	67.4 (124)	69.1 (125)
p value*	-	< 0.001	< 0.001	0.347	0.553

*p-value for comparison of kidney stage classifications based on eGFR from CKD-EPI, SCC eq1, SCC eq2, and SCC eq3 with MDRD, calculated using the Chi-squared test

4.6.3. Comparison of the UACR with eGFR for the diagnosis of DN

To determine whether eGFR-based stages of kidney disease correlated with UACR results, participants were split into three groups according to their UACR (i.e. normo-, micro-, and macroalbuminuria). Within each group participants were categorised into the different stages of kidney disease based on their eGFR results obtained from the five different equations used (Table 4.12-4.14).

Table 4.12 summarises the stages of kidney disease for 18 participants with macroalbuminuria (UACR > 20 mg/mmol). Only the SCC eq2 and eq3 correctly classified all participants as having kidney dysfunction (stage 2 and higher). The CKD-EPI and SCC eq1 classified over 65 % of participants as having normal to mildly reduced kidney function (stage 1 and 2), where all the participants should have moderate to severely reduced kidney function (stage 3 or higher) based on their UACR classification.

Table 4.12: Comparison of UACR classification vs eGFR in 18 individuals with a UACR > 20 mg/mmol

Kidney stage	**eGFR (ml/min/1.73m^2)**				
	MDRD (n = 18)	**CKD-EPI (n = 18)**	**SCC eq1 (n = 18)**	**SCC eq2 (n = 18)**	**SCC eq3 (n = 18)**
Stage 1 (> 90); n (%)	1 (5.6)	8 (44.4)	2 (11.1)	0 (0.0)	0 (0.0)
Stage 2 (60-89); n (%)	4 (22.2)	4 (22.2)	11 (61.1)	5 (27.8)	5 (27.8)
Stage 3a (45-59); n (%)	5 (27.8)	2 (11.1)	2 (11.1)	7 (38.9)	3 (16.7)
Stage 3b (30-44); n (%)	4 (22.2)	1 (5.6)	3 (16.7)	3 (16.7)	7 (38.9)
Stage 4 (15-29); n (%)	3 (16.7)	2 (11.1)	0 (0.0)	3 (16.7)	3 (16.7)
Stage 5 (< 15); n (%)	1 (5.6)	1 (5.6)	0 (0.0)	0 (0.0)	0 (0.0)

eGFR = estimated glomerular filtration rate, MDRD = modification of diet and renal disease equation, CKD-EPI = chronic kidney disease epidemiology collaboration equation, SCC = serum cystatin C, eq = equation

In Table 4.13, participants with MA (UACR 2-20 mg/mmol) were classified according to their eGFR results. All the participants in this category should have some degree of kidney damage, as MA is indicative of the presence of DN, according to the diagnostic criteria. The MDRD equation appeared to perform the best for this group, classifying most participants as stage 2 (58.3 %) with mildly reduced kidney function. The CKD-EPI and SCC eq1 performed the worst with the majority of participants (80.5 and 61.1 %, respectively) classified as stage 1 (normal kidney function).

Table 4.13: Comparison of UACR classification vs eGFR in 72 individuals with a UACR 2-20 mg/mmol

Kidney stage	**eGFR (ml/min/1.73m^2)**				
	MDRD (n = 72)	**CKD-EPI (n = 72)**	**SCC eq1 (n = 72)**	**SCC eq2 (n = 72)**	**SCC eq3 (n = 72)**
Stage 1 (> 90); n (%)	14 (19.4)	58 (80.6)	44 (61.1)	25 (34.7)	25 (34.7)
Stage 2 (60-89); n (%)	42 (58.3)	10 (13.9)	18 (25.0)	27 (37.5)	28 (38.9)
Stage 3a (45-59); n (%)	8 (11.1)	1 (1.4)	6 (8.3)	7 (9.7)	7 (9.7)
Stage 3b (30-44); n (%)	5 (6.9)	3 (4.2)	3 (4.2)	9 (12.5)	7 (9.7)
Stage 4 (15-29); n (%)	3 (4.2)	0 (0.0)	1 (1.4)	4 (5.6)	5 (6.9)
Stage 5 (< 15); n (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

eGFR = estimated glomerular filtration rate, MDRD = modification of diet and renal disease equation, CKD-EPI = chronic kidney disease epidemiology collaboration equation, SCC = serum cystatin C, eq = equation

Table 4.14 shows the kidney stage classification of participants with normoalbuminuria (UACR < 2 mg/mmol) and thus having normal kidney function. The CKD-EPI equation performed the best, classifying the majority (84.2 %) of participants as stage 1 (normal kidney function). The MDRD, SCC eq2, and SCC eq3 methods performed the worst with over 60 % of participants classified with mild to severely reduced kidney function.

Table 4.14: Comparison of UACR classification vs eGFR in 95 individuals with a UACR < 2 mg/mmol

Kidney stage	eGFR (ml/min/1.73m^2)				
	MDRD (n = 95)	**CKD-EPI (n = 95)**	**SCC eq1 (n = 94)**	**SCC eq2 (n = 94)**	**SCC eq3 (n = 91)**
Stage 1 (> 90); n (%)	21 (22.1)	80 (84.2)	60 (63.8)	28 (29.8)	27 (29.7)
Stage 2 (60-89); n (%)	51 (53.7)	13 (13.7)	26 (27.7)	39 (41.5)	40 (44.0)
Stage 3a (45-59); n (%)	19 (20.0)	2 (2.1)	7 (7.4)	17 (18.1)	14 (15.4)
Stage 3b (30-44); n (%)	4 (4.2)	0 (0.0)	1 (1.1)	8 (8.5)	8 (8.8)
Stage 4 (15-29); n (%)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.1)	2 (2.2)
Stage 5 (< 15); n (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

eGFR = estimated glomerular filtration rate, MDRD = modification of diet and renal disease equation, CKD-EPI = chronic kidney disease epidemiology collaboration equation, SCC = serum cystatin C, eq = equation

4.6.4. Comparison of eGFR equations

Passing-Bablok regression and Bland-Altman plots were used to determine the correlation and agreement of the CKD-EPI and SCC eGFR equations with the MDRD equation. The MDRD eGFR was chosen as the reference method as this is the equation that has been used by the NHLS up until March 2020 to report kidney function.

The MDRD and CKD-EPI equations correlated strongly ($r = 0.945$, $p < 0.001$; Figure 4.13), however, this correlation was weaker ($r = 0.890$, $p < 0.001$; Appendix I) when only looking at eGFR results > 60 ml/min/1.73m^2. Despite the strong correlation seen, there was a large negative bias (-39.7 %) which was more evident for the eGFR values > 60 ml/min/1.73m^2. That is, MDRD eGFR results were lower than those obtained for the CKP-EPI equation and this difference was exaggerated for eGFR values > 60 ml/min/1.73m^2. The SCC eq2 and eq3 eGFR results were in strongest agreement with the MDRD equation with a negative bias of -7.71 % and -7.10 %, respectively (Figure 4.13). The SCC eq3 showed poor agreement with the MDRD equation with a negative bias of -29.7 % (Figure 4.13).

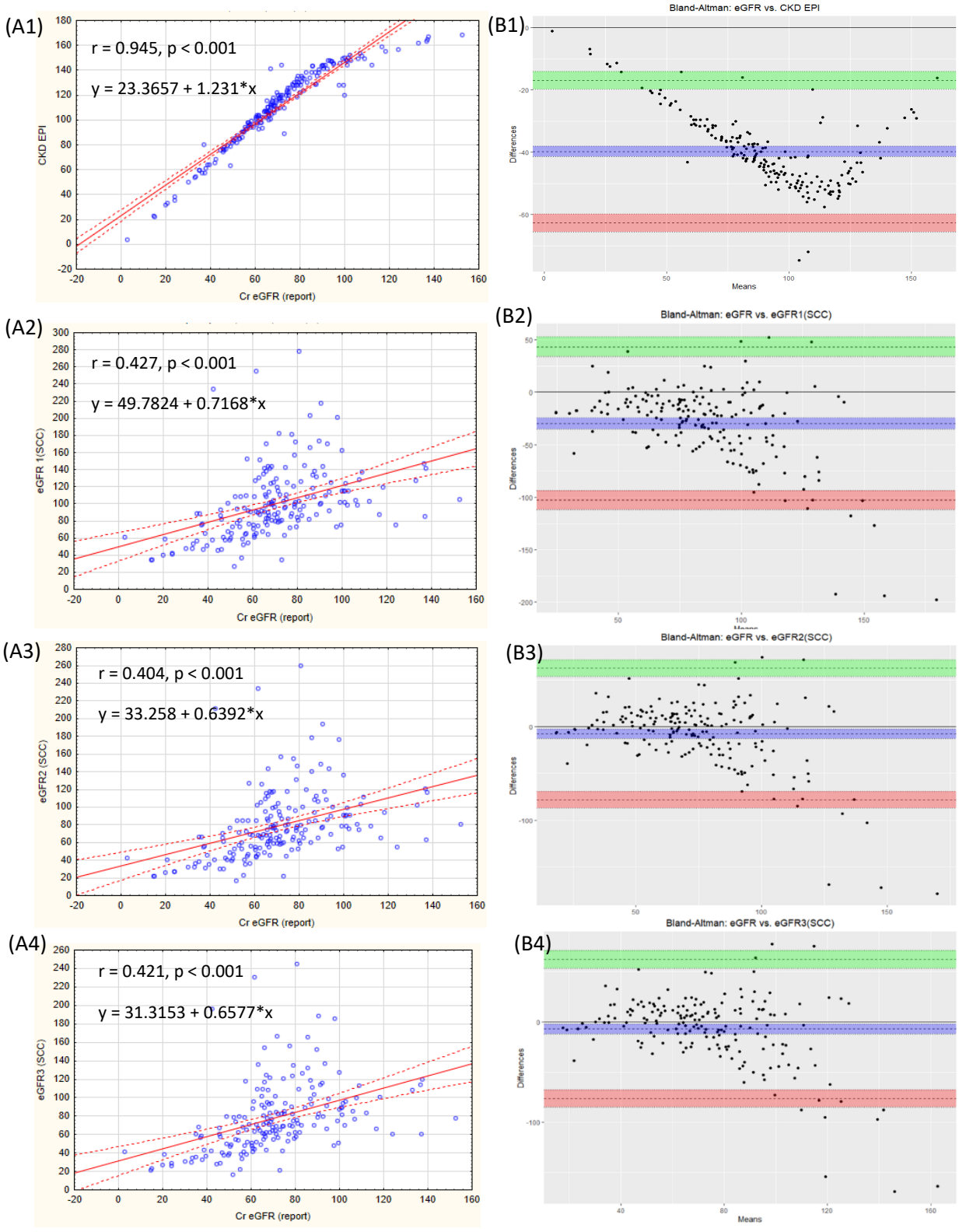


Figure 4.13: Passing-Bablok regression curves (A) and Bland-Altman plots (B) comparing eGFR results from the four eGFR equations with those from the MDRD eGFR equation

eGFR results using (1) CKD-EPI, (2) SCC eq1, (3) SCC eq2 and (4) SCC eq3. The MDRD eGFR method was used as the reference method.

5. Discussion

Approximately 35 million South Africans are living with T2DM (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). An estimated 9.5 billion USA dollars are spent on diabetes related healthcare interventions per year in Africa (International Diabetes Federation, 2019). In South Africa, almost a quarter (23 %) of the total healthcare budget is spent on diabetes (International Diabetes Federation, 2019). T2DM is, therefore, a national health priority. Poorly controlled diabetes can lead to the development of T2DM-related complications, such as DN, the leading cause of ESRD. The identification of biomarkers for early diagnosis and treatment of individuals at increased risk of developing renal disease would provide some relief of the economic burden placed on the healthcare systems and individuals (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017).

The biomarker commonly used for identification and monitoring of kidney function is urine MA, however it lacks both specificity and sensitivity (Lee and Lam, 2015). Thus, the aim of this study was to identify possible new biomarkers – serum circulating miRNAs – for early detection of DN. Identification of DN before the onset of clinical signs and symptoms, may offer an opportunity for early intervention, and ultimately reversal of kidney damage or prevention of further disease development.

Serum miRNA expression profiles identified two miRNAs (let-7b-5p and miR-455-3p) that were differentially expressed between participants with T2DM with and without MA. Validation of the miRNAs in a larger cohort, by real-time qPCR, indicated that both miRNAs were upregulated in patients with T2DM with normoalbuminuria and low MA, as compared to those with macroalbuminuria. Investigations of three polymorphisms (rs1114167319, rs1114167320, and rs1800471) in two genes identified as targets of let-7b-5p and miR-455-3p (namely, *HMGA2* and *TGF-β1*) showed no associations between the genotypes and markers of kidney function. The only association seen was between participants with the *TGF-β1* GG genotype and lower glucose levels.

The clinical significance of SCC as a marker of kidney function was evaluated by calculating the eGFR of participants using three SCC eGFR equations and the CKD-EPI equation and comparing them to MDRD eGFR equation results. The MDRD and CKD-EPI equations showed the strongest correlation, however, a large negative bias was observed, with significantly lower eGFR values obtained with the MDRD equation. The SCC equations correlated weakly with the MDRD

equation, but good agreement was seen between the SCC eq2 and SCC eq3 eGFR results with the MDRD eGFR, as evidenced by the smaller negative biases of -7.71 % and -7.10 %, respectively. The equations do not appear to be interchangeable within a clinical setting.

5.1. Characteristics of the study participants

The T2DM study participants had a mean age of 56.6 ± 9.88 years, were predominantly female (61.8 %), obese (32.0 ± 6.14 kg/m^2), hypertensive (sBP: 144.3 ± 21.9 mmHg), and had poor glycaemic control (HbA1c: 8.68 ± 2.73 %). The mean age of the population is in keeping with the known fact that T2DM is a disease that occurs in adults > 45 years of age (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). It should be noted, however, that due to the obesity epidemic, an increasing number of children are being diagnosed with T2DM (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). The mean age observed in our cohort was similar to that reported in other studies from South Africa, where the mean ages were 59.4 ± 9.9 years (Bulbulia, Variava and Bayat, 2020) and 59.4 ± 10.4 years (Rotchford and Rotchford, 2002) in T2DM participants where the majority of the participants were black.

Excess body weight is a known risk factor for T2DM development. In South Africa, it is estimated that about 87 % of DM cases are attributed to obesity (Groenewald *et al.*, 2007), and approximately 70 % of South African women are overweight or obese (Stats SA, 2008). Although, men are more likely to be diagnosed with DM, and at lower BMIs than women (Harreiter and Kautzky-Willer, 2018), women are more likely to seek medical intervention for weight problems and other medical interventions (Kanter and Caballero, 2012). This could account for why more females were recruited in our study, and why the majority of the cohort were obese. The predominant female presence was also observed in other studies of South African populations with T2DM (Hilawe *et al.*, 2013; Bulbulia, Variava and Bayat, 2020).

The risk of hypertension is twice as high in T2DM individuals as compared to healthy individuals (The Emerging Risk Factors Collaboration, 2010). In addition, the prevalence of coexisting T2DM and hypertension is increasing (Song *et al.*, 2016), thus it is not surprising that a raised systolic blood pressure (144.3 ± 21.9 mmHg) was noticed in the South African black population with T2DM in this study.

In T2DM, raised glucose levels are due to the loss of insulin action and/or secretion which ultimately leads to insulin resistance and hence hyperglycaemia (American Diabetes Association,

2010). An HbA1c result ≥ 6.5 % can be used to diagnose T2DM (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). Intensive glucose control (HbA1c ≤ 7.0 %) is advocated to reduce the risk of diabetes related complications (SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017). Based on the mean HbA1c results (8.68 ± 2.73 %) in this cohort, the participants are not attaining the desired level of glucose control and at increased risk of micro- and macrovascular complications.

5.1.1. The frequency of DN in the black South African T2DM population

The presence of albumin in the urine (MA; 2-20 mg/mmol) is one of the earliest clinical signs of kidney dysfunction. MA is sometimes referred to as incipient nephropathy. As the disease progresses, albumin levels in the urine increase until overt nephropathy is reached (> 20 mg/mmol). When this cohort was divided into normo-, micro-, and macroalbuminuria, based on the measured UACR, 10.1 % of participants were classified as having overt DN. A further 39.5 % of participants had MA. It is possible that some of the MA participants had DN, however, as we were unable to repeat the MA results within 3-6 months (as required for DN diagnosis) this could not be confirmed. In addition, it has been reported that 20-40 % of patients with MA will develop overt nephropathy if left untreated (American Diabetes Association, 2004). Thus, the 10.1 % frequency of DN in this population is likely an underestimation of the true frequency of this diabetic complication. Based on the percentage of individuals with MA that go on to progress to DN, the actual frequency in the study population could range between 18.1-26.6 %. In South Africa, the prevalence of DN is between 14 and 16 % (Naicker, 2009). In a sub-Saharan African meta-analysis of 27 studies, a 31 % prevalence of DN was reported (Wagnew *et al.*, 2018), and across African countries a MA prevalence ranging from 21 to 45 % has been reported (Noubiap, Naidoo and Kengne, 2015).

Furthermore, it is believed that the frequency of DN in developing countries with T2DM are higher than those in developed countries (Noubiap, Naidoo and Kengne, 2015). It is believed that the accurate diagnosis of DN in African countries is hindered as a result of delayed diagnosis of T2DM and limited availability of resources for screening and early diagnosis of MA (Swanepoel, Wearne and Okpechi, 2013). This supports our suggestion that the 10.1 % frequency of DN in this population is likely an underestimation of the true frequency of this diabetic complication.

5.1.2. Decreased renal function is associated with a longer disease duration and hypertension

Decreased renal function (expressed as a measure of albumin in the urine) was associated with a significantly longer duration of disease ($p = 0.022$) and higher systolic blood pressure ($p < 0.001$). Kidney function declines as a natural consequence of age, further supporting the association of macroalbuminuria with increased disease duration (Denic, Glassock and Rule, 2016). In addition, two of the major risk factors of kidney damage and the progression of MA to macroalbuminuria are duration of disease and systemic hypertension (Defarrari *et al.*, 1998). Studies have shown a correlation between increasing MA and an increase in T2DM disease duration (Kundu *et al.*, 2013). This association is most likely due to the prolonged exposure of cells to high glucose levels.

Hyperglycaemia has been associated with podocyte loss through dysregulation of integrin expression (Tsilibary, 2003; Shankland, 2006). As a result, the remaining podocytes increase in size, and are said to be hypertrophic (Petermann *et al.*, 2005; Xu *et al.*, 2005). The larger podocytes enhance protein leakage in the kidneys and MA increases (Kriz, Gretz and Lemley, 1998; Tryggvason and Wartiovaara, 2001). In addition, hyperglycaemia activates vasoactive hormone pathways (Wolf, Butzmann and Wenzel, 2003), which in turn activates transcription factors that result in a profibrotic response resulting in renal fibrosis and ultimately DN (Wolf, 2004; Lim, 2014; Alicic, Rooney and Tuttle, 2017).

In addition to hyperglycaemia, arterial hypertension has been shown to enhance the progression of renal damage (Chockalingam, Campbell and Fodor, 2006; Wagnew *et al.*, 2018). As described in the introduction, hypertension results in increased glomerular capillary pressure, that enhances the glomerular plasma flow, and subsequently the GFR (Thomson, Vallon and Blantz, 2004). In combination with DM-induced hypertrophy, the kidney undergoes GBM thickening, to further reduce GFR (Chawla, Sharma and Singh, 2010), which results in further podocyte loss (Wolf, 2004). Podocyte loss is greatly enhanced through hyperglycaemia and poor glycaemic control (Tsilibary, 2003; Shankland, 2006). The association of increase blood pressure and increasing albuminuria is supported in a number of studies conducted in participants with T2DM from different populations (Gall *et al.*, 1991; Bruno *et al.*, 1996; Kalk, Raal and Joffe, 2010; Moosa *et al.*, 2015).

As expected, macroalbuminuria was associated with a low GFR (MDRD; $p < 0.001$), and increased SCC concentrations ($p = 0.005$). Cystatin C has been identified as a promising marker of renal function. Cystatin C is produced at a constant rate and is freely filtered in the renal

glomeruli and metabolised in the proximal tube (Lee and Lam, 2015). The stepwise increase of SCC with increasing albuminuria was supported by a Korean study (SCC: 0.91 ± 0.26 vs 1.05 ± 0.38 vs 2.04 ± 1.19 mg/l; $p < 0.001$), which showed a decrease in eGFR with increasing albuminuria (MDRD eGFR: 84.6 ± 23.3 vs 76.0 ± 27.8 vs 44.9 ± 26.6 ml/min/7.73m^2; $p < 0.001$) (Jeon *et al.*, 2011).

5.2. Nine miRNAs were differentially expressed in participants with and without MA

Nine differentially expressed miRNAs were identified by GeneChip® microarray analysis.

Of the nine miRNAs, seven (miR-4740-3p, -7704, -101-5p, 5189-3p, -1273g-3p, -16-5p, and -6880-5p) were downregulated in cases vs controls. The possible targets and pathways of these miRNAs were investigated in order to identify the miRNAs specifically involved with DN and/or DM.

miR-7704, miR-4740-3p, and miR-5189-5p have been associated with different types of cancer (Schotte *et al.*, 2011; Bibi *et al.*, 2016; Wang *et al.*, 2017; Martin-Guerrero *et al.*, 2018). To the best of our knowledge, miR-7704, miR-4740-3p and miR-5189-3p have not previously been associated with either T2DM or DN.

miRNAs identified that may have some connection to DM include miR-101-5p, -1273g-3p, and -6880-5p.

miR-101-5p has been described as a potential hepatocellular carcinoma marker (Yang *et al.*, 2018), as well as having associations with apoptosis and pancreatic β-cell dysfunction (Grieco *et al.*, 2017), and T2DM disease progression (Higuchi *et al.*, 2015; Catanzaro *et al.*, 2018).

miR-1273g-3p and miR-16-5p have been associated with the development pathway of the DM complication retinopathy (Wu *et al.*, 2014b; Elmasry *et al.*, 2018; miRDB: Wong and Wang, 2015; Liu and Wang, 2019). Furthermore, miR-16-5p has been shown to be upregulated in gestational DM (Zhu *et al.*, 2015; Guarino *et al.*, 2018), and in the plasma of patients with T1DM and DN (Assmann *et al.*, 2019). An increase in miR-16-5p has also been described in urine samples from DM-induced rats suffering from DN, and is predicted to target the *IL-6* gene (Ghai *et al.*, 2018). IL-6 has been implicated in the inflammation process, which is involved in DN development and kidney fibrosis (Alicic, Rooney and Tuttle, 2017). IL-6 activity is reduced in DM-induced cells, indicating a protective mechanism against fibrosis for this miRNA (Mohan *et al.*, 2016). Another

study, however, found no differential expression of miR-16-5p in the urine of T1DM patients with DN (Ghai *et al.*, 2018).

The final downregulated miRNA identified in this study, miR-6880-5p, has been connected to vitamin D3 receptor formation through control of the vitamin D receptor (*VDR*) gene (miRDB: Wong and Wang, 2015; Liu and Wang, 2019). *VDR* is expressed in pancreatic β-cells, and the VDR protein is involved in regulating insulin secretion (Zeitz *et al.*, 2003; Bland *et al.*, 2004). Various polymorphisms in the *VDR* gene have been described and associated with vitamin D metabolism, glucose intolerance, and insulin sensitivity, often resulting in increased susceptibility to T2DM development (Bland *et al.*, 2004; Palomer *et al.*, 2008). This supports the suggestion that vitamin D plays a role in the pathogenesis of T2DM (Palomer *et al.*, 2008; Manchanda and Bid, 2012).

Two miRNAs (let-7b-5p and miR-455-3p) that were upregulated in T2DM participants with MA as compared to those with normoalbuminuria, were selected for further validation studies by real-time qPCR. These miRNAs were selected based on the genes they target and previous literature that showed an association of one, or both miRNAs, with kidney dysfunction.

Let-7 is a family of miRNAs involved in regulating the body's response to inflammation. The dysregulation of this family can give rise to multiple diseases, and has been associated with the development of liver and lung fibrosis, cancer, diabetic-associated atherosclerosis, and renal fibrosis (Brennan *et al.*, 2017).

Let-7d targets the *HMGA2* gene (Wang *et al.*, 2016b), and as such, is involved in the regulation of fibrotic processes, alongside another let-7 family member, let-7b, which target the 3' UTR of the *HMGA2* gene. By inhibiting the expression of *HMGA2*, *TGF-β1*-mediated epithelial cell to mesenchymal cell transition (EMT) is supressed. The expression of let-7d has been shown to be decreased in the presence of increased expression of *HMGA2* (Wang *et al.*, 2016b).

In addition, let-7b also targets the expression of the *TGF-β receptor 1* (*TGF-βR1*) gene, which has been shown to be associated with renal fibrosis in DN (Wang *et al*., 2016b). Let-7b inhibits *TGF-βR1* expression and reduced *TGF-βR1* expression leads to decreased signalling via the TGF-β/SMAD (protein homologous to the Drosophila MAD (mothers against decapentaplegic) and *Caenorhabditis elegans* SMA (small worm phenotype) proteins) signalling pathway (Wu and Derynck, 2009), thereby supressing TGF-β1's profibrotic effects. TGF-β binds to TGF-βR2, which recruits TGF-βR1 forming a hetero-tetrameric complex with the ligand, leading to phosphorylation of the TGF-βR1 serine residues by TGF-β2, activating TGF-βR1. SMAD3 then binds to the activated receptor, which results in phosphorylation of the serine residue in SMAD3,

causing a conformational change. SMAD3 dissociates from the receptor and the phosphorylated SMAD enters the nucleus where it modulates expression of ECM genes, such as collagen I, IV, and fibronectin, which lead to renal fibrosis (Wu and Derynck, 2009). Additionally, it has been shown that let-7b expression is inhibited by the expression of TGF-β1 within the proximal tubular cells of the kidneys (Wang *et al.*, 2014; Wang *et al.*, 2016b).

The second miRNA of interest, miR-455-3p, has been found to be downregulated in DN cell-lines (Wu *et al.*, 2018), and expression levels noted to be regulated by increased TGF-β1 and glucose levels in the proximal tubular epithelial cells in DM (Wu *et al.*, 2018). When overexpression of this miRNA was induced, it was found that it reversed pathological changes of renal fibrosis through suppression of the ROCK2 pathway. The ROCK2/Rho-associated signalling pathway has been implicated as a key contributor to diabetic kidney pathologies, specifically related to kidney fibrosis (Wu *et al.*, 2018). *ROCK2* was found to be the direct target of miR-455-3p, and reduces the inhibitory actions that miR-455-3p has on cell proliferation, ECM formation, and EMT (Wu *et al.*, 2018). miR-455-3p has also been linked to repression of renal fibrosis in DN by repressing the TGF-β1 pathway (Wu *et al.*, 2018). The downregulation of miR-455-3p can result in loss of protection against renal fibrosis, resulting in worsening kidney function, as the profibrotic effects of TGF-β1 are no longer inhibited (Wu *et al.*, 2018).

5.2.1. Let-7b-5p and miR-455-3p were upregulated in participants with a UACR ≤ 4.0 mg/mmol supporting a protective role in kidney dysfunction

Comparison of the qPCR results between groups with different UACR (≤ 4.0 mg/mmol) showed significantly lower C_T-values for both let-7b-5p and miR-455-3p in the low UACR group indicating upregulation of these miRNAs in this group, compared to those with the highest UACR (≥ 18.3 mg/mmol). Studies in cell lines have shown that overexpression of let-7b-5p and miR-455-3p provide protection from renal fibrosis (as described in the previous section) (Wang *et al.*, 2016b; Wu *et al.*, 2018) and thus support our findings that let-7b-5p and miR-455-3p provide protection from renal fibrosis. However, these findings from the larger validation study are contradictory to the results seen in the microarray studies, which found these miRNAs to be upregulated in patients with MA compared to those with normoalbuminuria.

The microarray assays were performed on normoalbuminuria vs MA participants, whereas the qPCR was performed on a group of mixed normoalbuminuria and low-level MA participants, and compared to a group of participants with high-levels of MA and macroalbuminuria. The

difference in expression levels between the microarray and qPCR data may indicate a change in levels of the miRNAs over the course of DN progression, with higher levels of the miRNAs at the start of the disease, followed by a decrease in expression as the disease progresses.

Some biomarkers have been described as having different expression profiles in individuals with MA as compared to those with macroalbuminuria (Jia *et al.*, 2016). Similarly, different biomarker expression profiles have been found in different kidney diseases (DN vs CKD) (Lv *et al.*, 2013; Peng *et al.*, 2013). As only participants with DM were used in this study, it may be assumed that the difference between the microarray and qPCR results may possibly be due to the stage of kidney disease of the individuals tested, rather than due to different kidney diseases.

It should be noted that, due to time and budget constraints, the qPCR validation in this study was performed only on a small subset of patients. In order to determine if the differences described above are, in fact, due to varying kidney disease stages, the qPCR validation would have to be repeated in a larger set of patients with normoalbuminuria, MA and macroalbuminuria.

5.3. Gene targets of let-7b-5p and miR-455-3p

The gene targets of the upregulated miRNAs (let-7b-5p and miR-455-3p) that may be involved in DN and/or DM were further investigated in order to identify whether genetic variation in the genes themselves may be implicated in the development of DN. The two genes investigated were *HMGA2* and *TGF-β1*. *HMGA2* has been found to be regulated by the let-7 miRNA family, through a gene-gene interaction pathway with the *TGF-β1* gene, which regulates miR-455-3p expression.

5.3.1. The *HMGA2* gene

HMGA proteins are involved in stem-cell development and proliferation, and have been found in high levels in human embryonic stem cells (Alkayyali *et al.*, 2013). In addition, these proteins bind to DNA in the minor groove altering the DNA conformation, which affects the accessibility of regulatory factors involved in controlling gene expression (Vignali and Marracci, 2020). HMGA2 is a transcription regulator, with increased expression seen in cells that have undergone EMT (Wang *et al.*, 2016b).

HMGA2 has been identified as playing a role in the proliferation and differentiation of adipose tissue stem cells, and is often rearranged in tumours of mesenchymal origins. The rearrangement often results in an increase in *HMGA2* expression. This increase of expression is not always linked to tumour-like formations, but can also be triggered through other stimuli, such as white adipose tissue formation through diet-induced hyperplasia and hypertrophy (Markowski *et al.*, 2013).

A temporal increase in adipose tissue leads to increased cell senescence, and with it an increase in insulin resistance in response to adipose tissue inflammation (Markowski *et al.*, 2013).

Overexpression of *HMGA2* has been associated with the development of micro-polycystic kidneys, suggesting an involvement in kidney development (Fedele *et al.*, 2002). If kidney development is affected, an increase in later kidney-related disease can be expected. This may translate to an increased risk for DN in T2DM, should it later develop, however further research is required to confirm this hypothesised association (Alkayyali *et al.*, 2013).

Linkage and association studies relating to the *HMAG2* gene region have described two pathogenic SNPs, upstream of the *HMGA2* gene, that may affect the gene function, as well as potentially being involved in the development of T2DM and susceptibility to DN (Alkayyali *et al.*, 2013; Dayeh *et al.*, 2013).

The first SNP, rs1531343 (Dayeh *et al.*, 2013), is an ancestral guanine to cytosine change (G>C) located upstream of the *HMGA2* gene, within the ribosome protein SA pseudogene 52 (*RPSAP52*) gene (Billings and Florez, 2010). This SNP was found to be in strong linkage disequilibrium (LD) with the *HMGA2* gene. The G allele has been associated with susceptibility to T2DM, increased risk of DN, and reduced GFR in European diabetic individuals as compared to non-diabetic individuals (Alkayyali *et al.*, 2013).

The rs1531343 polymorphism was only investigated in people of European ancestry (Alkayyali *et al.*, 2013). In individuals of African ancestry, however, a different SNP, rs2358944, was associated with similar phenotypic results as seen for the rs1531343 polymorphism (McDonough *et al.*, 2011; Alkayyali *et al.*, 2013). This SNP results in an ancestral guanine to adenine change (G>A) upstream of rs1531343 (Billings and Florez, 2010), and has been strongly associated with a reduced incidence of ESRD in diabetic individuals, but not in non-diabetic individuals (McDonough *et al.*, 2011; Alkayyali *et al.*, 2013). It has been proposed that the possible mechanism for this is epigenetic changes in DNA methylation patterns across the genome due to CpG-island SNP variants. The change in methylation can in turn affect either

repression or expression of closely located genes, or the function or downstream processing of transcriptional products (Dayeh *et al.*, 2013).

No known polymorphisms in the *HMGA2* gene itself have been linked to T2DM or DN, however, two SNPs (rs1114167319 and rs1114167320) in the gene have been flagged as being possibly pathogenic (Habib *et al.*, 2018).

The first polymorphism, rs1114167319, is a stop-gain mutation caused by an ancestral cytosine to thiamine change (C>T), and the second polymorphism, rs1114167320, is a frameshift mutation caused by an ancestral adenine deletion (A>del) (Habib *et al.*, 2018).

Both polymorphisms are suspected to be associated with Russel-Silver Syndrome (RSS) development. RSS is a rare growth disorder (Shiraishi, Mishima and Umeda, 2019), characterised by reduced kidney function and development of kidney disease (Ahn *et al.*, 2009). The mechanism of the *HMGA2* gene and these variants in RSS development is unknown, but it is speculated that the variants could affect kidney function, given the role the gene plays in the fibrotic process.

5.3.2. Polymorphisms in the *HMGA2* gene are not associated with markers of kidney function in black South Africans with T2DM

The major allele for the *HMGA2* rs1114167319 SNP was the C allele with a frequency of 0.51, and the major allele for the rs1114167320 SNP was the A allele with a frequency of 1.0. No population data was available from the 1000 Genome project for comparison of allelic frequencies of either polymorphism with other population groups.

No associations were found between the *HMAG2* polymorphisms and markers of kidney function or DM. To our knowledge, no previous investigations have been conducted regarding these *HMGA2* polymorphisms and their associations with DM or DN.

Additionally, the polymorphisms were not in HWE, indicating that the frequencies found are not representative of the greater population. A larger sample in HWE is necessary before any conclusions can be made regarding the role of these polymorphisms in DN.

5.3.3. The *TGF-β1* gene

The TGF-β family is a group of proteins known to be involved in multiple cellular functions, including cell differentiation, cell growth, adhesion, migration, and death. The signalling pathway of the TGF-β family has been linked to the pathogenesis of multiple diseases, such as fibrotic diseases, cancer, and cardiovascular diseases (Suzuki, 2018).

Serum TGF-β1 has been investigated as a possible biomarker for DN (Mou *et al.*, 2016). TGF-β1 induces apoptosis and increased serum concentrations of TGF-β1 can be seen before podocyte depletion, urine albuminuria and mesangial matrix expansion (Susztak *et al.*, 2006; Li *et al.*, 2007), making it an attractive biomarker. Elevated serum levels of TGF-β1 have been associated with the presence of DN in T2DM patients (Mou *et al.*, 2016). It has also been speculated that TGF-β1-induced kidney fibrosis is mediated by HMGA2 protein expression (Alkayyali *et al.*, 2013). Suppression of *HMGA2* mRNA overexpression has been identified as being under the control of the let-7 miRNA family (Yong and Dutta, 2007). Furthermore, let-7 expression is regulated by TGF-β1 (Wang *et al.*, 2016b).

In the *TGF-β1* gene itself, rs1800471 (Wallace and Wilcos, 2004), a missense guanine to cytosine polymorphism (G>C) (McKnight *et al.*, 2007) has been studied in relation to kidney disease. An association between increased plasma *TGF-β1* levels and the G allele has been noted in European individuals (Nabrdalik *et al.*, 2013). In addition, non-diabetic European individuals with CKD of mixed origins homozygous for the G allele showed a more aggressive development of CKD compared to individuals with the C allele (Nabrdalik *et al.*, 2013).

These associations, however, differ depending on the population group studied. In a Mexican population with T2DM, the CC and GC genotypes were associated with greater odds of developing DN than T2DM patients with the GG genotype ($p = 0.08$, OR = 4.073) (Valladares-Salgado *et al.*, 2010). However, in an Egyptian population with T2DM, no significant association was found between the *TGF-β1* rs1800471 SNP and DN (El-Sherbini *et al.*, 2013; Wei *et al.*, 2018). Furthermore, in CKD of diabetic origin, particularly T1DM, no association was found between the polymorphism and kidney disease in Caucasian populations (Nabrdalik *et al.*, 2013).

The ethnic variations observed, and the unknown mechanism through which the polymorphism increases susceptibility to DN requires further research in larger and diverse population groups (Nabrdalik *et al.*, 2013).

The *HMGA2* and *TGF-β1* polymorphisms were therefore genotyped in our study cohort in order to investigate their possible involvement in DN in the South African black population with T2DM.

5.3.4. The *TGF-β1* C allele is associated with higher glucose levels in black South Africans with T2DM

In this study, the *TGF-β1* rs1800471 G allele was the major allele with a frequency of 0.64. The frequency of the rs1800471 G allele in the South African black population was significantly lower than that seen in the 1000 Genome project African population (0.64 vs 0.94; $p < 0.001$) (The 1000 Genomes Project Consortium, 2015).

Participants with the rs1800471 CC/GC genotypes had significantly higher glucose concentrations compared to participants with the GG genotype (8.10 [5.70; 11.3] vs 6.60 [4.80; 8.80]; $p = 0.011$). Hyperglycaemia increases protein synthesis and cell size, promoting hypertrophy (Sartorelli and Fulco, 2004), and has been associated with increased *TGF-β1* expression (Wu and Derynck, 2009). Increased *TGF-β1* expression has also been shown to increase globular blood flow, which further contributes to hypertrophy and sheer-induced vessel damage in the kidneys (Chawla, Sharma and Singh, 2010), and leads to increased apoptosis and podocyte loss (Soon Lee, 2013). TGF-β1 also increases the expression of ECM proteins, contributing to increased GBM thickening (Li *et al.*, 2007). Thus, over time, hyperglycaemia leads to decreased renal function (Wu and Derynck, 2009). In addition, the G allele has previously been shown to be associated with increased levels of TGF-β1. Therefore, we would expect to see an association between the G allele and higher glucose levels which is contradictory the association we found; namely, the association of the C allele with higher glucose levels.

The polymorphism was not in HWE, however, indicating that the cohort was not representative of the larger population. Therefore, to confirm the association identified with genotype and glucose levels, the study needs to be repeated in a larger cohort.

5.4. SCC as a marker of kidney function

Five eGFR equations were evaluated for use to monitor kidney function. Two equations were based on serum creatinine levels (MDRD and CKD-EPI), and three on SCC concentrations. The true gold standard for measurement of GFR is determined from the urinary clearance of a bolus infusion of inulin (KDIGO, 2013). This method, together with silver methods that measure clearance of exogenous substances (e.g. 125-I-iothalamate and iohexol) are time-consuming, expensive, have risks of ionising radiation, and are not readily available in South Africa (KDIGO, 2013). Creatinine and SCC are endogenous substances that have more cost effective and widely

available tests than the exogeneous methods (Lee and Lam, 2015). Due to lack of standardisation of SCC methods, serum creatinine-based methods are more commonly used (Ebert *et al.*, 2016). The MDRD equation was the method in use by the NHLS during the time this study was performed, and therefore, for the purposes of this study was used as the "gold standard" or for comparison purposes.

5.4.1. The CKD-EPI equation classifies fewer participants with kidney dysfunction than the MDRD equation

The CKD-EPI equation is an expansion of the MDRD equation, which attempts to obtain a more accurate estimation of the GFR, specifically at values greater than 60 ml/min/1.73m^2 (Silveiro *et al.*, 2011). Comparing the CKD-EPI equation with the MDRD equation, a strong positive correlation was seen ($r = 0.945$; $p < 0.001$), however the correlation decreased slightly when comparing eGFRs > 60 ml/min/1.73m^2 ($r = 0.890$; $p < 0.001$). This finding confirms previous reports that have shown reduced accuracy of the MDRD equation in the > 60 ml/min/1.73m^2 eGFR range (Silveiro *et al.*, 2011). The reduced accuracy of the MDRD equation at eGFR > 60 ml/min/1.73m^2 is due to the fact that this equation was validated in a cohort of individuals with CKD whereas the CKD-EPI equation was validated in a cohort of participants with and without kidney disease (Levey *et al.*, 1999; Levey *et al.*, 2009). As a result, when eGFRs derived from the MDRD equation are reported, results are only reported as an exact value until 60 ml/min/1.73m^2, after which they are reported as > 60 ml/min/1.73m^2.

Despite a strong correlation, there was a poor agreement between the two equations, with a large negative bias (-39.7 %) in this study, with CKD-EPI eGFR values significantly higher than those from the MDRD equation. This finding is supported by a comparative study of the CKD-EPI and MDRD equations, in a large UK population, which found a significantly higher median CKD-EPI eGFR than median MDRD eGFR values (82 vs 76 ml/min/1.73m^2; $p < 0.001$) even though they found a small bias (5.0 %) between equations (Carter *et al.*, 2011). In addition, a meta-analysis study consisting of 40 countries, and mixed ethnicities, found higher eGFR results from the CKD-EPI equation compared to the MDRD equation in both healthy participants and participants with CKD (Matsushita *et al.*, 2012). Thus, the CKD-EPI equation classified fewer individuals as having CKD, which is consistent with our results where 93.5 % of participants were classified with an eGFR ≥ 60 ml/min/1.73m^2 using the CKD-EPI equation compared to the 71.9 % ($p < 0.001$) for the

MDRD equation. A study conducted in a Brazilian population, however, found that the CKD-EPI equations performed similarly in participants with T2DM (Camargo *et al.*, 2011).

5.4.2. The SCC equations cannot be used interchangeably for classification of kidney function

The SCC based equations have been described as more accurate methods for GFR estimations (Krolewski *et al.*, 2012), however they are not widely used. Comparisons between the SCC equations and the creatinine eGFR equations indicated a significant difference in kidney disease staging ($p < 0.001$). However, when the SCC methods were compared separately to the MDRD method, across the cohort, significant, but weak, correlations were found (SCC eq1 $r = 0.427$, SCC eq2 $r = 0.404$, SCC eq3 $r = 0.421$; $p < 0.001$). SCC eq2 and eq3 showed good agreement with the MDRD equation with biases of -7.71 % and -7.10 % respectively. In addition, no significant differences in kidney disease classifications were observed between the SCC eq2 and eq3 with the MDRD equation when participants were grouped according to eGFR < or ≥ 60 ml/min/1.73m^2.

A study in T2DM patients with CKD reported more accurate eGFR results from SCC based equations than the MDRD equation (bias: -2.4 % vs -27.1 %, respectively), when compared to a reference isotopic GFR measurement (MacIsaac *et al.*, 2007; MacIsaac, Premaratne and Jerums, 2011). Similarly, a smaller bias was seen for the SCC eq3 than the MDRD equation when compared to a reference isotopic GFR measurement in Australian T2DM participants (-8.5 % vs -27.1 %, respectively) (MacIsaac, Premaratne and Jerums, 2011). All these equations are widely used in the research setting, however limited studies are available on the different SCC-based equations and their clinical applications, and further research is required. In addition, SCC measurements need to be standardised and the effect of non-renal factors investigated before this method is likely to be used in a clinical setting.

The MDRD equation is known to be affected by age, race, sex, and lifestyle differences, making it highly variable and a less reliable method than the gold standard of eGFR measurement by urinary inulin clearance (Rule *et al.*, 2006). The MDRD equation has the advantage, however, of being an easily accessible and cheaper alternative to the gold standard (KDIGO, 2013). The variations observed in our study between the different eGFR equations can thus be due to the lack of accuracy of the MDRD equation. To obtain true comparative values between the different GFR estimation equations, the results should be compared to a true gold standard method for GFR measurement, namely inulin clearance.

Inconsistencies were also noted in classifications of the stage of the renal disease, when comparing the two creatinine and three SCC based methods. This further supports our conclusion that the different equations cannot be used interchangeably in the clinical setting of T2DM, for estimation of GFR, classifying stages of kidney disease or for monitoring kidney function.

6. Limitations of the study

miRNA expression assays were limited by the lack of markers to validate the integrity of the extracted miRNAs. The Agilent RNA 6000 Pico kit that was used for quantification of miRNA was intended for total RNA extractions. When total RNA is extracted, rRNAs (18S and 28S) will be present and can thus be used to validate the integrity of the extracted samples. Since, only miRNAs were extracted with the Qiagen miRNeasy kit, the rRNAs needed for integrity checks were absent.

The miRNA validation and SCC measurements were performed on the respective sections of the cohort singly, based on availability of kits and budget. Duplicate/triplicate measurements are recommended to ensure result reliability (Monach, 2012). The lack of repeat measurements can limit the application of the results (Monach, 2012), as outliers/errors in measurements are less likely to be detected.

The miRNA qPCR validation was limited as an internal control was not used. House-keeping genes are often used as internal controls in RNA expression studies (Tsotetsi *et al.*, 2018), however minimal research has been done on identifying circulating miRNAs as normalisation controls (Yang *et al.*, 2014). To account for this, the study normalised the measurements through serum sample volume used for extraction. This normalisation does not account for intra-individual variations, nor miRNA extraction performance.

In some analyses, only a portion of the full cohort was used or could be used. The reduced sample size in the miRNA validation and calculatable UACR could affect trends seen in the results. To account for this, larger study cohorts should be used to improve the quality and statistical reliability of the results.

A further limitation of this study is that the performance of the eGFR results were not compared to the gold standard of measured GFR, namely inulin or the accepted alternative iohexol.

Lastly, none of the polymorphisms investigated were in HWE, and although analysis was performed, these results need to be repeated in a larger cohort where the polymorphism reach HWE before conclusions can be made with respect to their associations with DM and/or DN.

7. Future studies

Further validation studies can be performed through comparisons of miRNA expression profiles between patients with normoalbuminuria, MA, and macroalbuminuria to validate expression levels in different stages of kidney function. This work should be carried out in a larger cohort. Additional studies would also include investigation into the seven differentially downregulated miRNAs. In addition, all the miRNAs could be investigated in different racial populations of South Africa to determine if there are differences in disease aetiology in different racial groups.

Further validation of the real-time qPCR findings can also be done. Validation can include being done on a larger population group, and of mixed ancestral backgrounds. The miRNA expression profiles in serum can also be compared with those in plasma and urine. Part of this would also include identifying, using, and validating a possible internal circulating serum miRNA control to normalise the results to. Clinical validation and longitudinal studies of let-7b-5p and miR-455-3p as a biomarker for DN in patients with T2DM can be encouraged.

Possible future work to continue from this project includes further analysis of the *HMGA2* polymorphisms, in particular the *HMGA2* rs1114167320 polymorphism, where no homozygous or heterozygous participants for the deletions were noted. Functional studies of the polymorphism could give strength to the hypothesis that *HMGA2* rs1114167320 might have pathogenic influences if the deletion is present.

The comparisons between methods to calculate eGFR can be improved by incorporating the true gold standard inulin, or a more affordable and acceptable alternative such as iohexol, as the basis of comparison. If the project were to continue further, it may be of some value to include a group of participants without DM, and a group of participants with CKD without DM. This could assist in validating any findings as specific to DM and/or DN.

8. Conclusion

Development of new, efficient and cost-effective biomarkers for DN could result in improved quality of life for individuals with T2DM, as well as significant cost-savings, particularly in the public healthcare sector.

Two miRNAs were identified and validated using microarray expression profiling assays and real-time qPCR, namely let-7b-5p and miR-455-3p. The miRNAs were upregulated in participants with a low UACR compared to those with a higher UACR. Increased expression of these miRNAs has been shown to attenuate profibrotic signalling pathways and therefore protect against renal fibrosis and DN. Thus, these miRNAs appear to be valuable markers for the diagnosis of DN and further studies to confirm this association are warranted.

The miRNAs were investigated for target genes that are related to DN and DM, and two genes were identified, namely *HMGA2* and *TGF-β1*. No associations were found between the polymorphisms and DN, or markers of kidney function. The only association noticed was between increased glucose levels and the *TGF-β1* rs1800471 polymorphism, however the population was not in HWE. Further work is therefore needed to confirm this association.

More studies need to be conducted comparing SCC eGFR equations with the gold standard to determine their usefulness in determining kidney function. From this study, we can conclude that these equations cannot be used interchangeably in a clinical setting.

9. References

Adeghate, E., Schattner, P. and Dunn, E. (2006) 'An update on the etiology and epidemiology of diabetes mellitus', *Annals of the New York Academy of Sciences*, 1084, pp. 1–29.

Affymetrix (2017) Data Sheet GeneChip miRNA 4.0 Array. USA.

Agarwal, V. *et al.* (2015) 'Predicting effective microRNA target sites in mammalian mRNAs', *eLife*, 4(e5005), pp. 1–38.

Agilent Technologies (2016) *Agilent RNA 6000 pico kit guide*. Germany.

Ahn, Y. H. *et al.* (2009) 'Chronic renal failure in Russell-Silver Syndrome', *Journal of the Korean Society of Pediatric Nephrology*, 13(2), p. 256.

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

Ali, O. (2013) 'Genetics of type 2 diabetes', *World Journal of Diabetes*, 4(4), pp. 114–123.

Alicic, R. Z., Rooney, M. T. and Tuttle, K. R. (2017) 'Diabetic kidney kisease: Challenges, progress, and possibilities', *Clinical journal of the American Society of Nephrology*, 12, pp. 2032–2045.

Alkayyali, S. *et al.* (2013) 'Common variant in the HMGA2 gene increases susceptibility to nephropathy in patients with type 2 diabetes', *Diabetologia*, 56, pp. 323–329.

Ambele, M. A. *et al.* (2016) 'Genome-wide analysis of gene expression during adipogenesis in human adipose-derived stromal cells reveals novel patterns of gene expression during adipocyte differentiation', *Stem Cell Research*. Elsevier B.V., 16, pp. 725–734.

American Diabetes Association (2004) 'Nephropathy in Diabetes', *Diabetes Care*, 27(Supplement 1), pp. S79–S83.

American Diabetes Association (2010) 'Diagnosis and classification of diabetes mellitus', *Diabetes Care*, 33(Supplement 1), pp. S62-69.

Assmann, T. S. *et al.* (2019) 'Circulating miRNAs in diabetic kidney disease: Case-control study and in silico analyses', *Acta Diabetologica*. Springer Milan, 56(1), pp. 55–65.

Barutta, F. *et al.* (2013) 'Urinary exosomal microRNAs in incipient diabetic nephropathy', *PLoS ONE*, 8(11), pp. 1–8.

Bertram, M. Y. *et al.* (2013) 'The non-fatal disease burden caused by type 2 diabetes in South Africa, 2009', *Global Health Action*, 6, p. 19244.

Bevc, S. *et al.* (2012) 'Simple cystatin C formula for estimation of glomerular filtration rate in overweight patients with diabetes mellitus type 2 and chronic kidney disease', *Experimental Diabetes Research*, 2012, pp. 1–8.

Bibi, F. *et al.* (2016) 'microRNA analysis of gastric cancer patients from Saudi Arabian population', *BMC Genomics*. BMC Genomics, 17(Supplement 9), pp. 51–87.

Bikandi, J. *et al.* (2004) 'In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction', *Bioinformatics*, 20(5), pp. 798–799.

Billings, L. K. and Florez, J. C. (2010) 'The genetics of type 2 diabetes: What have we learned from GWAS?', *Annals of the New York Academy of Sciences*, 1212, pp. 59–77.

Bland, R. *et al.* (2004) 'Expression of 25-hydroxyvitamin D3-1α-hydroxylase in pancreatic islets', *The Journal of Steroid Biochemistry and Molecular Biology*, 89–90, pp. 121–125.

Blondal, T. *et al.* (2013) 'Assessing sample and miRNA profile quality in serum and plasma or other biofluids', *Methods*. Elsevier Inc., 59, pp. S1–S6.

Brennan, E. *et al.* (2017) 'Protective effect of let-7 miRNA family in regulating inflammation in diabetes-associated atherosclerosis', *Diabetes*, 66(8), pp. 2266–2277.

Bruno, G. *et al.* (1996) 'Prevalence and risk factors for micro- and macroalbuminuria in an italian population-based cohort of NIDDM subjects', *Diabetes Care*, 19(1), pp. 43–47.

Bulbulia, S., Variava, F. and Bayat, Z. (2020) 'Are type 2 diabetic patients meeting targets? A Helen Joseph Hospital Diabetic Clinic audit', *Journal of Endocrinology, Metabolism and Diabetes of South Africa*, 25(1), pp. 12–17.

Camargo, E. G. *et al.* (2011) 'The chronic kidney disease epidemiology collaboration (CKD-EPI) equation is less accurate in patients with type 2 diabetes when compared with healthy individuals', *Diabetic Medicine*, 28(1), pp. 90–95.

Campion, C. G., Sanchez-Ferras, O. and Batchu, S. N. (2017) 'Potential role of serum and urinary biomarkers in diagnosis and prognosis of diabetic nephropathy', *Canadian Journal of Kidney Health and Disease*, 4, pp. 1–18.

Canadian Diabetes Association Clinical Practice Guidelines Expert Committee (2013) 'Clinical practice guidelines for the prevention and management of diabetes in Canada', *Canadian Journal of Diabetes*, 37, pp. S1–S3.

Carter, J. L. *et al.* (2011) 'Estimating glomerular filtration rate: Comparison of the CKD-EPI and MDRD equations in a large UK cohort with particular emphasis on the effect of age', *Quarterly Journal of Medicine*, 104, pp. 839–847.

Catanzaro, G. *et al.* (2018) 'Circulating microRNAs in elderly type 2 diabetic patients', *International Journal of Endocrinology*. Hindawi, 2018, pp. 1–11.

Chawla, T., Sharma, D. and Singh, A. (2010) 'Role of the renin angiotensin system in diabetic nephropathy', *World Journal of Diabetes*, 1(5), pp. 141–145.

Chen, Y. and Wang, X. (2020) 'miRDB: An online database for prediction of functional microRNA targets', *Nucleic Acids Research*. Oxford University Press, 48(Database issue), pp. D127–D131.

Chockalingam, A., Campbell, N. R. and Fodor, J. G. (2006) 'Worldwide epidemic of hypertension', *Canadian Journal of Cardiology*, 22(7), pp. 553–555.

Choukem, S. P. *et al.* (2012) 'Comparison of different blood pressure indices for the prediction of prevalent diabetic nephropathy in a sub-Saharan African population with type 2 diabetes', *Pan African Medical Journal*, 11, p. 67.

Colhoun, H. M. and Marcovecchio, M. L. (2018) 'Biomarkers of diabetic kidney disease', *Diabetologia*. Diabetologia, 61, pp. 996–1011.

Cooper, M. E., Gilbert, R. E. and Epstein, M. (1998) 'Pathophysiology of diabetic nephropathy', *Metabolism*, 47(1), pp. 3–6.

Dayeh, T. A. *et al.* (2013) 'Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets', *Diabetologia*, 56, pp. 1036–1046.

Defarrari, G. *et al.* (1998) 'Diabetic nephropathy: From micro- to macroalbuminuria', *Nephrology Dialysis Transplantation*, 13(Supplement 8), pp. 11–15.

Denic, A., Glassock, R. J. and Rule, A. D. (2016) 'Structural and functional changes with the aging kidney', *Advances in Chronic Kidney Disease*, 23(1), pp. 19–28.

van Deventer, H. E. *et al.* (2008) 'Estimating glomerular filtration rate in black South Africans by use of the modification of diet in renal disease and cCockcroft-Gault equations', *Clinical Chemistry*, 54(7), pp. 1197–1202.

Ebert, N. *et al.* (2016) 'Cystatin C standardization decreases assay variation and improves assessment of glomerular filtration rate', *Clinica Chimica Acta*. Elsevier B.V., 456, pp. 115–121.

El-Sherbini, S. M. *et al.* (2013) 'Gene polymorphism of transforming growth factor-β1 in Egyptian patients with type 2 diabetes and diabetic nephropathy', *Acta Biochimica et Biophysica Sinica*, 45(4), pp. 330–338.

Elmasry, K. *et al.* (2018) 'Epigenetic modifications in hyperhomocysteinemia: Potential role in diabetic retinopathy and age-related macular degeneration', *Oncotarget*, 9(16), pp. 12562–12590.

Etheridge, A. *et al.* (2011) 'Extracellular microRNA: A new source of biomarkers', *Mutation Research*, 717(1–2), pp. 85–90.

Fedele, M. *et al.* (2002) 'Overexpression of the HMGA2 gene in transgenic mice leads to the onset of pituitary adenomas', *Oncogene*, 21, pp. 3190–3198.

Gall, M. *et al.* (1991) 'Prevalence of micro- and macroalbuminuria, arterial hypertension, retinopathy and large vessel disease in European Type 2 (non-insulin-dependent) diabetic patients', *Diabetetology*, 34, pp. 655–661.

Garg, P. (2018) 'A review of podocyte biology', *American Journal of Nephrology*, 47(supplement 1), pp. 3–13.

Ghai, V. *et al.* (2018) 'Genome-wide profiling of urinary extracellular vesicle microRNAs associated with diabetic nephropathy in type 1 diabetes', *Kidney International Reports*. Elsevier Inc, 3, pp. 555–572.

Gilbertson, D. T. *et al.* (2005) 'Projecting the number of patients with end-stage renal disease in the United States to the year 2015', *Journal of the American Society of Nephrology*, 16, pp. 3736–3741.

Grieco, F. A. *et al.* (2017) 'MicroRNAs miR-23a-3p, miR-23b-3p, and miR-149-5p regulate the expression of proapoptotic bh3-only proteins DP5 and PUMA in human pancreatic β-cells', *Diabetes*, 66, pp. 100–112.

Griffiths-Jones, S. (2004) 'The microRNA registry', *Nucleic Acids Research*, 32(Database issue), pp. 109–111.

Griffiths-Jones, S. *et al.* (2006) 'miRBase: microRNA sequences , targets and gene nomenclature', *Nucleic Acids Research*, 34(Database issue), pp. 140–144.

Griffiths-Jones, S. *et al.* (2008) 'miRBase: Tools for microRNA genomics', *Nucleic Acids Research*, 36(Database issue), pp. 154–158.

Groenewald, P. *et al.* (2007) 'Estimating the burden of disease attributable to smoking in South Africa in 2000', *South African Medical Journal*, 97(8), pp. 674–681.

Gross, J. L. *et al.* (2005) 'Diabetic nephropathy: Diagnosis, prevention, and treatment', *Diabetes Care*, 28(1), pp. 176–188.

Großhans, H. and Filipowicz, W. (2008) 'The expanding world of small RNAs', *Nature*, 451, pp.

414–416.

Guarino, E. *et al.* (2018) 'Circulating microRNAs as biomarkers of gestational diabetes mellitus: Updates and perspectives', *International Journal of Endocrinology*. Hindawi, 2018, pp. 1–11.

Habib, W. A. *et al.* (2018) 'Genetic disruption of the oncogenic HMGA2-PLAG1-IGF2 pathway causes fetal growth restriction', *Genetics in Medicine*. Nature Publishing Group, 20(2), pp. 250–258.

Harreiter, J. and Kautzky-Willer, A. (2018) 'Sex and gender differences in prevention of type 2 diabetes', *Frontiers in Endocrinology*, 9, pp. 1–15.

Higuchi, C. *et al.* (2015) 'Identification of circulating miR-101, miR-375 and miR-802 as biomarkers for type 2 diabetes', *Metabolism: Clinical and Experimental*. Elsevier Inc., 64, pp. 489–497.

Hilawe, E. H. *et al.* (2013) 'Differences by sex in the prevalence of diabetes mellitus, impaired fasting glycaemia and impaired glucose tolerance in sub-Saharan Africa: a systematic review and meta-analysis', *Bulletin of the World Health Organization*, 91(9), pp. 671–682.

Hostetter, T. H. (2003) 'Hyperfiltration and glomerulosclerosis', *Seminars in Nephrology*, 23(2), pp. 194–199.

Ibberson, D. *et al.* (2009) 'RNA degradation compromises the reliability of microRNA expression profiling', *BMC Biotechnology*, 9, p. 102.

Ichii, O. and Horino, T. (2018) 'MicroRNAs associated with the development of kidney diseases in humans and animals', *Journal of Toxicologic Pathology*, 31, pp. 23–34.

International Diabetes Federation (2017) *IDF diabetes atlas eighth edition 2017*. Eighth. Edited by S. Karuranga et al. Belgium.

International Diabetes Federation (2019) *IDF Diabetes Atlas Ninth edition 2019*. Ninth, *International Diabetes Federation*. Ninth. Edited by B. Malanda et al. Belgium: International Diabetes Federation.

Jefferson, J. A., Shankland, S. J. and Pichler, R. H. (2008) 'Proteinuria in diabetic kidney disease: A mechanistic viewpoint', *Kidney International*. Elsevier Masson SAS, 74(1), pp. 22–36. Available at: http://dx.doi.org/10.1038/ki.2008.128.

Jeon, Y. K. *et al.* (2011) 'Cystatin C as an early biomarker of nephropathy in patients with type 2 diabetes', *Journal of Korean Medical Science*, 26, p. 258.

Jia, Y. *et al.* (2016) 'miRNAs in urine extracellular vesicles as predictors of early-stage diabetic

nephropathy', *Journal of Diabetes Research*, 2016, pp. 1–10.

Kalk, W. J., Raal, F. J. and Joffe, B. I. (2010) 'The prevalence and incidence of and risk factors for, microalbuminuria among urban Africans with type 1 diabetes in South Africa: An inter-ethnic study', *International Journal of Diabetes Mellitus*. International Journal of Diabetes Mellitus, 2(3), pp. 148–153.

Kanter, R. and Caballero, B. (2012) 'Global gender disparities in obesity: A review', *Advances in Nutrition*, 3, pp. 491–498.

Kato, M. *et al.* (2011) 'A microRNA circuit mediates transforming growth factor-β1 autoregulation in renal glomerular mesangial cells', *Kidney International*. Elsevier Masson SAS, 80, pp. 358–368.

KDIGO (2013) 'KDIGO 2012 Clinical practice guidelines for the evaluation and management of chronic kidney disease', *Official Journal of the Internalional Society of Nephorpathy*, 3(1), pp. 1–150.

Keeton, G. R., van Zyl Smit, R. and Bryer, A. (2004) 'Renal outcome of type 2 diabetes in South Africa - A 12-year follow-up study', *Journal of Endocrinology, Metabolism and Diabetes of South Africa*, 9(3), pp. 84–88.

Knerr, S., Wayman, D. and Bonham, V. L. (2011) 'Inclusion of racial and ethnic minorities in genetic research: Advance the spirit by changing the rules?', *Journal of Law, Medicine and Ethics*, 39(3), pp. 502–512.

Koressaar, T. and Remm, M. (2007) 'Enhancements and modifications of primer design program Primer3', *Bioinformatics*, 23(10), pp. 1289–1291.

Kozomara, A., Birgaoanu, M. and Griffiths-Jones, S. (2019) 'miRBase: From microRNA sequences to function', *Nucleic Acids Research*. Oxford University Press, 47(Database issue), pp. D155–D162.

Kozomara, A. and Griffiths-Jones, S. (2011) 'miRBase: Integrating microRNA annotation and deep-sequencing data', *Nucleic Acids Research*, 39(Database issue), pp. 152–157.

Kozomara, A. and Griffiths-Jones, S. (2014) 'miRBase: Annotating high confidence microRNAs using deep dequencing data', *Nucleic Acids Research*, 42(Database issue), pp. 68–73.

Kriz, W., Gretz, N. and Lemley, K. V. (1998) 'Progression of glomerular diseases: Is the podocyte the culprit?', *Kidney International*, 54, pp. 687–697.

Krolewski, A. S. *et al.* (2012) 'Serum concentration of cystatin C and risk of end-stage renal disease in diabetes', *Diabetes Care*, 35, pp. 2311–2316.

Kundu, D. *et al.* (2013) 'Relation of microalbuminuria to glycosylated hemoglobin and duration of type 2 diabetes', *Nigerian Journal of Clinical Practice*, 16(2), pp. 216–220.

Lee, C.-H. and Lam, K. S. L. (2015) 'Biomarkers of progression in diabetic nephropathy: The past, present and future', *Journal of Diabetes Investigation*, 6(3), pp. 247–249.

Levey, A. S. *et al.* (1999) 'A more accurate method to estimate glomerular filtration rate from serum creatinine: A new prediction equation', *Annals of Internal Medicine*, 130(6), pp. 461–470.

Levey, A. S. *et al.* (2009) 'A new equation to estimate glomerular filtration rate', *Annals of Internal Medicine*, 150(9), pp. 604–612.

Levitt, N. S. *et al.* (1997) 'Audit of public sector primary diabetes care in Cape Town, South Africa: High prevalence of complications, uncontrolled hyperglycaemia, and hypertension', *Diabetic Medicine*, 14, pp. 1073–1077.

Li, J. J. *et al.* (2007) 'Podocyte biology in diabetic nephropathy', *Kidney International*, 72, pp. 36–42.

Li, M. *et al.* (2014) 'Analysis of the RNA content of the exosomes derived from blood serum and urine and its potential as biomarkers', *Philosophical transactions of the Royal Society*, 369, p. 20130502.

Lim, A. K. H. (2014) 'Diabetic nephropathy – complications and treatment', *International Journal of Nephrology and Renovascular Disease*, 7, pp. 361–381.

Lin, C. L. *et al.* (2014) 'MicroRNA-29a promotion of nephrin acetylation ameliorates hyperglycemia-induced podocyte dysfunction', *Journal of the American Society of Nephrology*, 25, pp. 1698–1709.

Liu, W. and Wang, X. (2019) 'Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data', *Genome Biology*. Genome Biology, 20(1), pp. 1–10.

Liu, X. *et al.* (2013) 'Regulation of microRNAs by epigenetics and their interplay involved in cancer', *Journal of Experimental and Clinical Cancer Research*, 32, p. 96.

Looker, H. C. *et al.* (2015) 'Biomarkers of rapid chronic kidney disease progression in type 2 diabetes', *Kidney International*. Elsevier Masson SAS, 88, pp. 888–896.

Lv, L.-L. *et al.* (2013) 'MicroRNA-29c in urinary exosome/microvesicle as a biomarker of renal fibrosis', *American Journal of Renal Physiology*, 305, pp. F1220–F1227.

Ma, R. C. (2016) 'Genetics of cardiovascular and renal complications in diabetes', *Journal of*

Diabetes Investigation, 7(2), pp. 139–154.

MacIsaac, R. J. *et al.* (2007) 'The accuracy of cystatin C and commonly used creatinine-based methods for detecting moderate and mild chronic kidney disease in diabetes', *Diabetic Medicine*, 24, pp. 443–448.

MacIsaac, R. J., Premaratne, E. and Jerums, G. (2011) 'Estimating glomerular filtration rate in diabetes using serum cystatin C', *Clinical Biochemist Reviews*, 32, pp. 61–67.

Manchanda, P. and Bid, H. (2012) 'Vitamin D receptor and type 2 diabetes mellitus: Growing therapeutic opportunities', *Indian Journal of Human Genetics*, 18(3), pp. 274–275.

Markowski, D. N. *et al.* (2013) 'HMGA2 expression in white adipose tissue linking cellular senescence with diabetes', *Genes and Nutrition*, 8, pp. 449–456.

Martin-Guerrero, I. *et al.* (2018) 'Variants in the 14q32 miRNA cluster are associated with osteosarcoma risk in the Spanish population', *Scientific Reports*, 8, p. 15414.

Mason, R. M. and Wahab, N. A. (2003) 'Extracellular matrix metabolism in diabetic nephropathy', *Journal of the American Society of Nephrology*, 14, pp. 1358–1373.

Matsushita, K. *et al.* (2012) 'Comparison of risk prediction using the CKD-EPI equation and the MDRD study equation for estimated glomerular filtration rate', *Journal of the American Medical Association*, 307(18), pp. 1941–1951.

Mauer, S. M. *et al.* (1984) 'Structural-functional relationships in diabetic nephropathy', *Journal of Clinical Investigation*, 74, pp. 1143–1155.

McClelland, A. D. and Kantharidis, P. (2014) 'MicroRNA in the development of diabetic complications', *Clinical Science*, 126, pp. 95–110.

McDonough, C. W. *et al.* (2011) 'A genome-wide association study for diabetic nephropathy genes in African Americans', *Kidney International*, 79(5), pp. 563–572.

McKnight, A. J. *et al.* (2007) 'Resequencing of genes for transforming growth factor β1 (TGFB1) type 1 and 2 receptors (TGFBR1, TGFBR2), and association analysis of variants with diabetic nephropathy', *BMC Medical Genetics*, 8(5).

Mohan, A. *et al.* (2016) 'Urinary exosomal microRNA-451-5p is a potential early biomarker of diabetic nephropathy in rats', *PLoS ONE*, 11(4), pp. 1–18.

Monach, P. A. (2012) 'Repeating tests: Different roles in research studies and clinical medicine', *Biomarkers in Medicine*, 6(5), pp. 691–703.

Moosa, M. R. *et al.* (2015) 'Important causes of chronic kidney disease in South Africa', *South African Medical Journal*, 105(4), pp. 1–8.

Motawi, T. K. *et al.* (2018) 'Potential serum biomarkers for early detection of diabetic nephropathy', *Diabetes Research and Clinical Practice*. Elsevier B.V., 136, pp. 150–158.

Mou, X. *et al.* (2016) 'Serum TGF-β1 as a biomarker for type 2 diabetic nephropathy: A meta-analysis of randomized controlled trials', *PLoS ONE*, 11(2), pp. 1–9.

Nabrdalik, K. *et al.* (2013) 'Association of rs1800471 polymorphism of TGFB1 gene with chronic kidney disease occurrence and progression and hypertension appearance', *Archives of Medical Science*, 9(2), pp. 230–237.

Naicker, S. (2009) 'End-stage renal disease in sub-Saharan Africa', *Ethnicity and Disease*, 19, pp. S1-13.

Do Nascimento, J. F. *et al.* (2013) 'Messenger RNA levels of podocyte-associated proteins in subjects with different degrees of glucose tolerance with or without nephropathy', *BMC Nephrology*, 14(214), pp. 1–10.

Noubiap, J. J. N., Naidoo, J. and Kengne, A. P. (2015) 'Diabetic nephropathy in Africa: A systematic review', *World Journal of Diabetes*, 6(5), pp. 759–773.

Palmer, A. J. *et al.* (2008) 'A health economic analysis of screening and optimal treatment of nephropathy in patients with type 2 diabetes and hypertension in the USA', *Nephrology Dialysis Transplantation*, 23, pp. 1216–1223.

Palomer, X. *et al.* (2008) 'Role of vitamin D in the pathogenesis of type 2 diabetes mellitus', *Diabetes, Obesity and Metabolism*, 10, pp. 185–197.

Pavlidis, P., Li, Q. and Noble, W. S. (2003) 'The effect of replication on gene expression microarray experiments', *Bioinformatics*, 19(13), pp. 1620–1627.

Peng, H. *et al.* (2013) 'Urinary miR-29 correlates with albuminuria and carotid intima-media thickness in type 2 diabetes patients', *PLoS ONE*, 8(12), p. e82607.

Petermann, A. T. *et al.* (2005) 'Mechanical stretch induces podocyte hypertrophy in vitro', *Kidney International*, 67, pp. 157–166.

Pezzolesi, M. G. and Krolewski, A. S. (2013) 'The genetic risk of kidney disease in type 2 diabetes', *Medical Clinics of North America*, 97(1), pp. 91–107.

Poortmans, J. R. *et al.* (2013) 'Limitations of serum values to estimate glomerular filtration rate during exercise', *British Journal of Sports Medicine*, 47(18), pp. 1166–1170.

Qi, W. *et al.* (2006) 'TGF-β1 induces IL-8 and MCP-1 through a connective tissue growth factor-independent pathway', *American Journal of Physiology-Renal Physiology*, 290, pp. F703–F709.

R&D Systems (2017) *Quantikine ELISA Human Cystatin C Immunoassay*. USA.

R Core Team (2018) 'R: A Language and Environment for Statistical Computing', *R Foundation for Statistical Computing*, Vienna, Au. Available at: https://www.r-project.org/.

Retnakaran, R. *et al.* (2006) 'Risk factors for renal dysfunction in type 2 diabetes: UK Prospective Diabetes Study 74', *Diabetes*, 55(6), pp. 1832–1839.

Rossing, P., Hougaard, P. and Parving, H. H. (2005) 'Progression of microalbuminuria in type 1 diabetes: Ten-year prospective observational study', *Kidney International*, 68, pp. 1446–1450.

Rotchford, A. P. and Rotchford, K. M. (2002) 'Diabetes in rural South Africa--an assessment of care and complications.', *South African Medical Journal*, 92(7), pp. 536–41.

Rule, A. D. *et al.* (2006) 'Limitations of estimating glomerular filtration rate from serum creatinine in the general population', *Mayo Clinic Proceedings*, 81(11), pp. 1427–1434.

Samuel, V. T. and Shulman, G. I. (2012) 'Integrating mechanisms for insulin resistance', *Cell*, 148(5), pp. 852–871.

Sartorelli, V. and Fulco, M. (2004) 'Molecular and cellular determinants of skeletal muscle atrophy and hypertrophy', *Science's signal transduction knowledge environment*, (244), pp. 1–10.

Schotte, D. *et al.* (2011) 'Discovery of new microRNAs by small RNAome deep sequencing in childhood acute lymphoblastic leukemia', *Leukemia*. Nature Publishing Group, 25(9), pp. 1389–1399.

SEMDSA Type 2 Diabetes Guidelines Expert Committee (2017) 'The SEMDSA 2017 guidelines for the management of type 2 diabetes mellitus', in Amond, A. et al. (eds) *Journal of Endocrinology, Metabolism and Diabetes of South Africa*. South Africa, pp. S1–S196.

Shaker, O. G. and Sadik, N. A. H. (2013) 'Transforming growth factor beta 1 and monocyte chemoattractant protein-1 as prognostic markers of diabetic nephropathy', *Human and Experimental Toxicology*, 32, pp. 1089–1096.

Shankland, S. J. (2006) 'The podocyte's response to injury: Role in proteinuria and glomerulosclerosis', *Kidney International*, 69(12), pp. 2131–2147.

Sharma, K. *et al.* (2003) 'Involvement of transforming growth factor-β in regulation of calcium transients in diabetic vascular smooth muscle cells', *American Journal of Physiology-Renal Physiology*, 285(6), pp. F1258–F1270.

Shiraishi, M., Mishima, K. and Umeda, H. (2019) 'Russell-Silver syndrome with cleft palate: A

case report', *Oral and Maxillofacial Surgery*. Oral and Maxillofacial Surgery, 23, pp. 113–117.

Silveiro, S. P. *et al.* (2011) 'Chronic kidney disease epidemiology collaboration (CKD-EPI) equation pronouncedly underestimates glomerular filtration rate in type 2 diabetes', *Diabetes Care*, 34(11), pp. 2353–2355.

Simpson, K. *et al.* (2016) 'MicroRNAs in diabetic nephropathy: From biomarkers to therapy', *Current Diabetes Reports*, 16(35), pp. 1–7.

Song, Y. *et al.* (2016) 'Increasing trend of diabetes combined with hypertension or hypercholesterolemia: NHANES data analysis 1999–2012', *Scientific Reports*, 6(1), p. 36093.

Soon Lee, H. (2013) 'Pathogenic role of TGF-β in diabetic nephropathy', *Journal of Diabetes and Metabolism*, 01(S9), pp. 1–7.

Stats SA (2008) *Mid-year population estimates 2018*, *Statistical Release*. Pretoria, South Africca.

Stevens, L. A. *et al.* (2008) 'Estimating GFR using serum cystatin C alone and in combination with serum creatinine: A pooled analysis of 3,418 individuals with CKD', *American Journal of Kidney Diseases*, 51(3), pp. 395–406.

Susztak, K. *et al.* (2006) 'Glucose-induced reactive oxygen species cause apoptosis of podocytes and podocyte depletion at the onset of diabetic nephropathy', *Diabetes*, 55(1), pp. 225–233.

Suzuki, H. I. (2018) 'MicroRNA control of TGF-β signaling', *International Journal of Molecular Sciences*, 19(7), p. 1901.

Swanepoel, C. R., Wearne, N. and Okpechi, I. G. (2013) 'Nephrology in Africa - not yet uhuru', *Nature Reviews Nephrology*. Nature Publishing Group, 9(10), pp. 610–622.

Takir, M. *et al.* (2016) 'Cystatin-C and TGF-β levels in patients with diabetic nephropathy', *Nefrología*. Sociedad Española de Nefrología, 36(6), pp. 653–659.

Teimoury, A. *et al.* (2014) 'Why 24-h urine albumin excretion rate method still is used for screening of diabetic nephropathy in Isfahan Laboratories?', *International Journal of Preventive Medicine*, 5(3), pp. 341–7.

The 1000 Genomes Project Consortium (2015) 'A global reference for human genetic variation', *Nature*, 526(7571), pp. 68–74.

The Emerging Risk Factors Collaboration (2010) 'Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102

prospective studies', *The Lancet*. Elsevier Ltd, 375(9733), pp. 2215–2222.

Thomson, S. C., Vallon, V. and Blantz, R. C. (2004) 'Kidney function in early diabetes: The tubular hypothesis of glomerular filtration', *American Journal of Physiology - Renal Physiology*, 286, pp. 8–15.

Tryggvason, K. and Wartiovaara, J. (2001) 'Molecular basis of glomerular permselectivity', *Current Opinion in Nephrology and Hypertension*, 10(4), pp. 543–549.

Tsilibary, E. C. (2003) 'Microvascular basement membranes in diabetes mellitus', *Journal of Pathology*, 200(4), pp. 537–546.

Tsotetsi, T. N. *et al.* (2018) 'Selection and evaluation of housekeeping genes as endogenous controls for quantification of mRNA transcripts in Theileria parva using quantitative real-time polymerase chain reaction (qPCR)', *PLoS ONE*, 13(5), pp. 1–15.

Tucci, S. and Akey, J. M. (2019) 'The long walk to African genomics', *Genome Biology*. Genome Biology, 20(130), pp. 19–21.

Untergasser, A. *et al.* (2012) 'Primer3 - new capabilities and interfaces', *Nucleic Acids Research*, 40(15), pp. e115–e115.

Valladares-Salgado, A. *et al.* (2010) 'Association of polymorphisms within the transforming growth factor-β1 gene with diabetic nephropathy and serum cholesterol and triglyceride concentrations', *Nephrology*, 15, pp. 644–648.

Verhave, J. C. *et al.* (2013) 'Clinical value of inflammatory urinary biomarkers in overt diabetic nephropathy: A prospective study', *Diabetes Research and Clinical Practice*, 101, pp. 333–340.

Vignali, R. and Marracci, S. (2020) 'HMGA genes and proteins in development and evolution', *International Journal of Molecular Sciences*, 21(2), p. 654.

Wagnew, F. *et al.* (2018) 'Diabetic nephropathy and hypertension in diabetes patients of sub-Saharan countries: A systematic review and meta-analysis', *BMC Research Notes*. BioMed Central, 11(1), pp. 1–7.

Wallace, S. and Wilcos, W. (2004) *Camurati–Engelmann disease*, *GeneReviews*. Edited by M. Adam et al. Seattle: University of Washington.

Wang, B. *et al.* (2014) 'Transforming growth factor-β1-mediated renal fibrosis is dependent on the regulation of transforming growth factor receptor 1 expression by let-7b', *Kidney International*. Nature Publishing Group, 85(2), pp. 352–361.

Wang, C. *et al.* (2016) 'Increased serum microRNAs are closely associated with the presence of

microvascular complications in type 2 diabetes mellitus', *Scientific Reports*. Nature Publishing Group, 6(20032), pp. 1–9.

Wang, G. *et al.* (2012) 'Urinary miR-21, miR-29, and miR-93: Novel biomarkers of fibrosis', *American Journal of Nephrology*, 36(5), pp. 412–418.

Wang, X. *et al.* (2017) 'Differential expression profile analysis of miRNAs with HER-2 overexpression and intervention in breast cancer cells', *International Journal of Clinical and Experimental Pathology*, 10(5), pp. 5039–5062.

Wang, Y. *et al.* (2016) 'Let-7d miRNA prevents TGF-β1-induced EMT and renal fibrogenesis through regulation of HMGA2 expression', *Biochemical and Biophysical Research Communications*, 479(4), pp. 676–682.

Wei, L. *et al.* (2018) 'The susceptibility genes in diabetic nephropathy', *Kidney Diseases*, 4, pp. 226–237.

West, S. G., Finch, J. F. and Curran, P. J. (1995) 'Structural equation models with nonnormal variables: Problems and Remedies', in Hoyle, R. H. (ed.) *Structural equation modeling: Concepts, issues and applications.* Newbery Park, CA: Sage, pp. 56–75.

Wilcox, G. (2005) 'Insulin and insulin resistance', *The Clinical Biochemist Reviews*, 26, pp. 19–39.

Wolf, G. *et al.* (1997) 'High glucose stimulates expression of p27(Kip1) in cultured mouse mesangial cells: Relationship to hypertrophy', *American Journal of Physiology - Renal Physiology*, 273, pp. 348–356.

Wolf, G. (2004) 'New insights into the pathophysiology of diabetic nephrophathy: From haemodynamics to molecular pathology', *European Journal of Clinical Investigation*, 34(12), pp. 785–796.

Wolf, G., Butzmann, U. and Wenzel, U. O. (2003) 'The renin-angiotensin system and progression of renal disease: From hemodynamics to cell biology', *Nephron Physiology*, 93, pp. 3–13.

Wong, N. and Wang, X. (2015) 'miRDB: An online resource for microRNA target prediction and functional annotations', *Nucleic Acids Research*, 43(Database issue), pp. D146–D152.

Wu, C. *et al.* (2014) 'The changes of serum sKlotho and NGAL levels and their correlation in type 2 diabetes mellitus patients with different stages of urinary albumin', *Diabetes Research and Clinical Practice*. Elsevier Ireland Ltd, 106(2), pp. 343–350.

Wu, H. *et al.* (2014) 'The Role of microRNAs in diabetic nephropathy', *Journal of Diabetes Research*, 2014, pp. 1–12.

Wu, J. *et al.* (2018) 'MiR-455-3p suppresses renal fibrosis through repression of ROCK2

expression in diabetic nephropathy', *Biochemical and Biophysical Research Communications*, 503(2), pp. 977–983.

Wu, L. and Derynck, R. (2009) 'Essential role of TGF-β signaling in glucose-induced cell hypertrophy', *Developmental Cell*, 17(1), pp. 35–48.

Xu, Z. G. *et al.* (2005) 'Angiotensin II receptor blocker inhibits p27 Kip1 expression in glucose-stimulated podocytes and in diabetic glomeruli', *Kidney International*, 67(3), pp. 944–952.

Yang, X. *et al.* (2018) 'Diagnostic value of strand-specific miRNA-101-3p and miRNA-101-5p for hepatocellular carcinoma and a bioinformatic analysis of their possible mechanism of action', *FEBS Open Bio*, 8(1), pp. 64–84.

Yang, Y. *et al.* (2013) 'Urine miRNAs: Potential biomarkers for monitoring progresion of early stages of diabetic nephropathy', *Medical Hypotheses*, 81(2), pp. 274–278.

Yang, Z. *et al.* (2014) 'Serum miR-23a, a potential biomarker for diagnosis of pre-diabetes and yype 2 diabetes', *Acta Diabetologica*, 51, pp. 823–831.

Yong, S. L. and Dutta, A. (2007) 'The tumor suppressor microRNA let-7 represses the HMGA2 oncogene', *Genes and Development*, 21(9), pp. 1025–1030.

Zeitz, U. *et al.* (2003) 'Impaired insulin secretory capacity in mice lacking a functional vitamin D receptor', *Federation of American Societies for Experimental Biology*, 17(3), pp. 509–511.

Zhu, Y. *et al.* (2015) 'Profiling maternal plasma microRNA expression in early pregnancy to predict gestational diabetes mellitus', *International Journal of Gynecology and Obstetrics*. Elsevier B.V., 130(1), pp. 49–53.

Appendices

Appendix A: Information sheets and consent forms

INFORMATION SHEET AND INFORMED CONSENT FORM FOR ADULTS

1. STUDY TITLE

Serum circulating miRNA profiling for identification of potential markers of diabetic nephropathy in black African type 2 diabetic South Africans

2. THE NATURE AND PURPOSE OF THIS STUDY

Hello. My name is Stefan D. Valentin, and I am doing a research study on diabetes and its complications. In this study we want to look for markers which show that people with diabetes might be more likely to get complications, such as problems with their kidneys.

Diabetes affects all age groups of men and women, and sometimes people with diabetes will go on to get complications. Sometimes, when we start to get sick, even before we feel ill, our cells produce different molecules. These molecules can get into the blood, and by looking at your blood we can tell if you are getting sick. If we can tell early on that your diabetes is causing your kidneys to stop working properly, we can give treatments sooner and therefore, delay or even prevent you from developing kidney problems.

We would like to find molecules (markers) that are seen more often in diabetic people with kidney problems than in diabetic patients who don't get kidney problems. This might allow us to identify people before the complications begin by just looking at your blood. Treatment can then be started earlier, or new medicines could be made, so that patients do not get these complications.

Some people have small changes in their genes that are involved in how your kidneys function. We would like to see if these changes make you more likely to develop kidney problems as a result of your diabetes.

This information leaflet is to help you decide if you would like to participate in our study. You should fully understand what is involved before you agree to take part in this study. If you have any questions, please don't hesitate to ask me.

3. EXPLANATION OF PROCEDURES TO BE FOLLOWED

This study involves answering some questions about your diabetes, taking measurements such as weight, height, hip and waist, a urine sample and a small amount of blood from your arm. This will only take about 10 minutes and will all be done on the same day as you are visiting the clinic. Three blood tubes (about 3 teaspoons) will be taken. Some blood will be analysed and some blood samples will be stored for testing at a later stage. Any tests done on these blood samples will only be related to diabetes and its complications. If we find that you do not currently have any kidney problems, when you come back to the clinic for another check-up (in 6-12 months' time), we may ask you for another urine sample and two tubes of blood. If at this point you no longer wish to take

part in the study, this will not affect the treatment and care that you receive at the clinic. Participating today does not mean that you have to participate in the future.

4. RISK AND DISCOMFORT INVOLVED

The only risk and discomfort involved is the taking of blood from a vein which may leave a small bruise. The procedure will be performed under sterile conditions.

5. POSSIBLE BENEFITS OF THIS STUDY

The findings of a study such as this will not benefit you personally but will lead to a better understanding of how diabetes, and its complications, is caused. It may also lead to new treatments or earlier detection and intervention in the future.

6. I UNDERSTAND THAT IF I DO NOT WANT TO PARTICIPATE IN THIS STUDY, I WILL STILL RECEIVE STANDARD TREATMENT FOR MY ILLNESS

7. I MAY AT ANY TIME WITHDRAW FROM THIS STUDY

8. FINANCIAL ARRANGEMENTS

You will not be required to pay for any of the study related procedures.

9. REIMBURSEMENT FOR STUDY PARTICIPATION

If you are selected for additional sampling outside of your normal clinic visit times, we will offer a reimbursement for travel costs, up to R150.

10. ETHICAL APPROVAL

This study has been approved by the Human Ethics Committee of the University of the Witwatersrand, clearance number M180647.

Please feel free to contact the Chair of the Human Research Ethics Committee with any questions or complaints: Prof Penny Tel: 011 717 1234.

11. INFORMATION

If you have any questions concerning this study, please contact: Dr C Padoa, Tel : 011 489 8514.

12. CONFIDENTIALITY

All records obtained whilst in this study will be regarded as confidential. Results will be published or presented in such a fashion that patients remain unidentifiable.

CONSENT TO PARTICIPATE IN THIS STUDY

I have read or had read to me in a language that I understand the above information before signing this consent form. The content and meaning of this information have been explained to me. I have been given the opportunity to ask questions and am satisfied that they have been answered satisfactorily. I understand that if I do not participate it will not alter my management in any way. I hereby volunteer to take part in this study.

....

Patient's name Patient's signature

Date.....

....

Investigator's name Investigator's signature

Date.....

VERBAL PATIENT INFORMED CONSENT (applicable when patients cannot read or write)

I, the undersigned, ……, have read and have explained fully to the patient, named …….. and/or his/her relative, the patient information leaflet, which has indicated the nature and purpose of the study in which I have asked the patient to participate. The explanation I have given has mentioned both the possible risks and benefits of the study. The patient indicated that he/she understands that he/she will be free to withdraw from the study at any time for any reason and without jeopardising his/her further treatment.

I hereby certify that the patient has agreed to participate in this trial.

....

Independent literate witness' name Independent literate witness' signature

Date.....

....

Investigator's Name Investigator's Signature

Date.....

INFORMATION SHEET AND INFORMED CONSENT FORM FOR ADULTS – GENETIC STUDIES

1. STUDY TITLE

Serum circulating miRNA profiling for identification of potential markers of diabetic nephropathy in black African type 2 diabetic South Africans

2. GENETIC STUDIES

We would like to invite you to participate in our study on diabetes and its complications, particularly kidney problems. In order to look at markers of kidney disease (nephropathy) in your blood, we will need to do some genetic tests. That means we will look at some genes that this, or previous, studies identifiy to be important for/ linked to kidney disease and diabetes. Your DNA will not be used for any purpose other than that which has been approved by the Human Research Ethics Committee (HREC) of the University of the Witwatersrand. Any changes, or the identification of new genes for study, will first be approved by the HREC.

DNA will be stored in a -80˚C freezer in the Department of Chemical Pathology for a maximum of 10 years. If at any time you choose to withdraw from the study, your DNA will be removed from the freezer and destroyed. If you choose to withdraw from the study, at any point, it will not alter the management/treatment that you receive at the diabetes clinic in any way. If you have any questions, please don't hesitate to ask me.

3. ETHICAL APPROVAL

This study has been approved by the Human Ethics Committee of the University of the Witwatersrand, clearance number M180647.

Please feel free to contact the Chair of the Human Research Ethics Committee with any questions or complaints: Prof Penny Tel: 011 717 1234.

4. INFORMATION

If you have any questions concerning this study, please contact: Dr C Padoa, Tel : 011 489 8514.

CONSENT TO DO GENETIC STUDIES IN THE DEPARTMENT OF CHEMICAL PATHOLOGY IN THE UNIVERSITY OF THE WITWATERSRAND

I have read or had read to me in a language that I understand the above information before signing this consent form. The content and meaning of this information have been explained to me. I have been given the opportunity to ask questions and am satisfied that they have been answered satisfactorily. I understand that my DNA will be used for genetic studies relating to kidney disease and diabetes. I understand that if I do not participate in the study it will not alter my treatment in any way.

....

Patient's name Patient's signature

Date.....

....

Investigator's name Investigator's signature

Date.....

VERBAL PATIENT INFORMED CONSENT (applicable when patients cannot read or write)

I, the undersigned, ……, have read and have explained fully to the patient, named …… and/or his/her relative, the patient information leaflet, which explains the need for doing genetic studies and using the patient's DNA. The patient indicated that he/she understands that he/she will be free to withdraw from the study at any time for any reason and without jeopardising his/her further treatment and that upon withdrawal from the study the patient's DNA will be destroyed immediately.

I hereby certify that the patient has given his/her verbal consent for the use of his/her DNA for genetic studies.

....

Independent literate witness' name Independent literate witness' signature

Date.....

....

Investigator's Name Investigator's Signature

Date.....

INFORMATION SHEET AND INFORMED CONSENT FORM FOR ADULTS – STORAGE OF BLOOD

1. STUDY TITLE

Serum circulating miRNA profiling for identification of potential markers of diabetic nephropathy in black African type 2 diabetic South Africans

2. STORAGE OF BLOOD

In order to look at markers or differences in genes associated with kidney disease (nephropathy) in your blood, we will need to store your blood in a freezer in our laboratory in the Department of Chemical Pathology. It will be stored for a maximum time of 10 years, and will not be used for any purpose other than that which has been approved by the Human Research Ethics Committee of the University of the Witwatersrand.

If at any time you choose to withdraw from the study, your blood will be removed from the freezer and destroyed. If you choose to withdraw from the study, at any point, it will not alter the management/treatment that you receive at the diabetes clinic in any way. If you have any questions, please don't hesitate to ask me.

3. ETHICAL APPROVAL

This study has been approved by the Human Ethics Committee of the University of the Witwatersrand, clearance number M180647.

Please feel free to contact the Chair of the Human Research Ethics Committee with any questions or complaints: Prof Penny Tel: 011 717 1234.

4. INFORMATION

If you have any questions concerning this study, please contact: Dr C Padoa, Tel: 011 489 8514.

CONSENT TO STORE MY BLOOD IN THE DEPARTMENT OF CHEMICAL PATHOLOGY IN THE UNIVERSITY OF THE WITWATERSRAND

I have read or had read to me in a language that I understand the above information before signing this consent form. The content and meaning of this information have been explained to me. I have been given the opportunity to ask questions and am satisfied that they have been answered satisfactorily. I understand that my blood will be stored for up to ten years, after which it will be disposed of. I understand that if I do not participate in the study it will not alter my treatment in any way.

....

Patient's name Patient's signature

Date.....

....

Investigator's name Investigator's signature

Date.....

VERBAL PATIENT INFORMED CONSENT (applicable when patients cannot read or write)

I, the undersigned, ……, have read and have explained fully to the patient, named …….. and/or his/her relative, the patient information leaflet, which has indicated where the patient's blood will be stored. The explanation I have given has mentioned that the patient's blood will be stored for up to ten years. The patient indicated that he/she understands that he/she will be free to withdraw from the study at any time for any reason and without jeopardising his/her further treatment, and that upon withdrawal from the study the patient's blood will be destroyed immediately.

I hereby certify that the patient has given his/her verbal consent for the storage of his/her blood.

....

Independent literate witness' name Independent literate witness' signature

Date.....

....

Investigator's Name Investigator's Signature

Date.....

Appendix B: Questionnaire and datasheet

Demographics

Study number:
Date of visit (dd/mm/yyyy):

Biogram

Date of birth (dd/mm/yyyy):
Gender: M / F
Race:
Hospital/clinic number:
Telephone number of self or best possible contact:

Family history of diabetes

On Mothers side:
On Fathers side:
Of Siblings:

Risk factors

Smoking:	Y / N / Ex / U	Comments: (ex > 1yr stopped)
Random Capillary Glucose (from file):		
HbA1c within last 4 months:		
Blood pressure (mmHg):		
Hypertension on treatment:	Y / N / U	Year of diagnosis:
Weight (kg):		
Height (cm):		
Waist circumference (cm):		
Hip circumference (cm):		

Diabetes type

Clinical judgement on type of DM:	1 / 2	Year of diagnosis:

Complications

Macrovascular:

MI	Y	N
Stroke	Y	N
Revascularization	Y	N
Amputation	Y	N

Microvascular:

Lasered / retinopathy	Y	N
Neuropathy	Y	N
Nephropathy (microalbuminuria)	Y	N

Medication

Insulin
Oral medication
Ace inhibitors
Statins
Others

Other diseases

1.
2.
3.
4.
5.

Appendix C: Ethics certificates

R14/49 Dr Carolyn Padoa

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M140626

NAME: (Principal Investigator)	Dr Carolyn Padoa
DEPARTMENT:	Chemical Pathology Charlotte Maxeke Johannesburg Academic Hospital, National Health Laboratory Service
PROJECT TITLE:	Serum Circulating Mirna Profiling for Identification of Potential Markers of Diabetic Nepropathy in Black Type 2 Diabetic South Africans
DATE CONSIDERED:	27/06/2014
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Nigel Crowther
APPROVED BY:	Professor P Cleaton-Jones, Chairperson, HREC (Medical)
DATE OF APPROVAL:	27/03/2015

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**.

Principal Investigator Signature Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

R14/49 Mr Stefan Drikus Valentin et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M180647

NAME: (Principal Investigator)	Mr Stefan Drikus Valentin et al
DEPARTMENT:	Chemical Pathology Charlotte Maxeke Johannesburg Academic Hospital Department of Chemical Pathology Research Laboratory 3Q07
PROJECT TITLE:	Serum circulating microRNA profiling for identification of potential markers of diabetic nephropathy in black African type 2 diabetic South Africans
DATE CONSIDERED:	29/06/2018
DECISION:	Approved unconditionally
CONDITIONS:	Lab study
SUPERVISOR:	Dr Carolyn Padoa
APPROVED BY:	CB Penny Professor CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	20/08/2018

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the Third Floor, Faculty of Health Sciences, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in **June** and will therefore be due in the month of **June** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

Date 22 Aug 2018

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix D: miRNA expression profile reagent preparations

a) 1 mM Tris

Trisaminomethane (Tris) buffer (1 mM) was prepared by mixing 50 µl 1M Tris buffer in 50 ml nuclease-free water (from which 50 µl water had been removed). The solution was kept at room temperature (RT).

1 M Tris buffer was prepared by mixing 12.11 g Tris powder (Merck, Darmstadt, Germany) with nuclease free water (volume made up to a 100 ml). The pH was adjusted to 8.0 using the Starter 2100 bench pH meter (New Jersey, USA).

b) 1 x PBS, 0.02 % Tween-20

One litre (l) of 1 x phosphate buffered saline (PBS; GE Healthcare Life Sciences, Utah, USA) was used with the addition of 200 µl Tween-20 (Merck, Darmstadt, Germany), and stored at RT.

c) 5 % BSA in 1 x PBS

Bovine serum albumin (BSA) powder (2 g; Sigma-Aldrich, Missouri, USA) was transferred to a 50 ml conical vial and 40 ml 1 x PBS was slowly added. The mixture was shake to mix, and eight aliquots were made in 5 ml vials. The aliquots were stored at -20 °C, and thawed when used. The thawed solutions were kept at 4 °C for up to one week before a new aliquot was used.

d) 5 x SSC, 0.05 % SDS, 0.0005 % BSA

In a 10 ml conical vial, 2.5 ml 20 x saline sodium citrate (SSC; Sigma-Aldrich, Missouri, USA), 50 µl 10 % sodium dodecyl (SDS; Invitrogen Life Technologies, Thermo Fisher Scientific, Massachusetts, USA), and 10 µl 5 % BSA in 1 x PBS were added, and nuclease-free water added to a final volume of 10 ml. After mixing well, 10 aliquots were made of 1 ml each and stored at -20 °C and thawed when needed. Each thawed aliquot was kept at 2-4 °C for up to one week before a new aliquot was used.

e) 25 % Dextran sulphate

The dextran sulphate solution was prepared by slowly pouring 5 ml 50 % dextran sulphate (Sigma-Aldrich, Missouri, USA) into a 15 ml conical tube. Nuclease free water (5 ml) was added, and the solution was thoroughly vortexed until the dextran sulphate was completely dissolved. The solution was kept at RT.

Appendix E: Real-time qPCR QC

After addition of EvaGreen (Bioline, Memphis, Tennessee, USA) to the reaction mix, a signal for amplification was obtained on the Corbett Rotor-Gene™ HRM real-time PCR 6000 instrument (Corbett Research, Sydney, Australia) (Figure AE1). A melt curve was performed to validate that the amplification of the target miRNAs was, in fact, the desired targets and not primer dimers (Figure AE2). A product melt was observed around 82 °C for let-7b-5p and 84 °C for miR-455-3p. The calculated melting temperature for both miRNAs are 64 °C (Bikandi *et al.*, 2004).

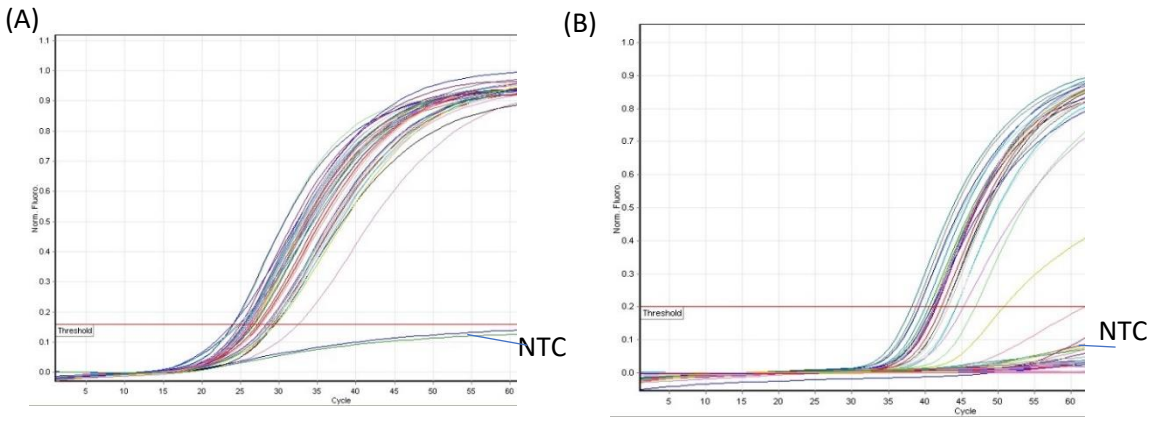
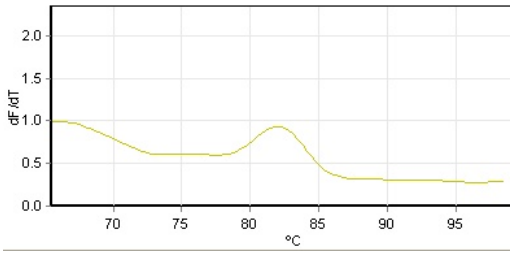


Figure AE1: Amplification curve from the qPCR of let-7b-5p (A) and miR-455-3p (B) with EvaGreen

Following the addition of EvaGreen to the reaction mix, an amplification signal was detected on the Rotor-Gene™ for let-7b-5p (A) and miR-455-3p (B). The red line indicates the position of the C_T threshold.

(A)



(B)

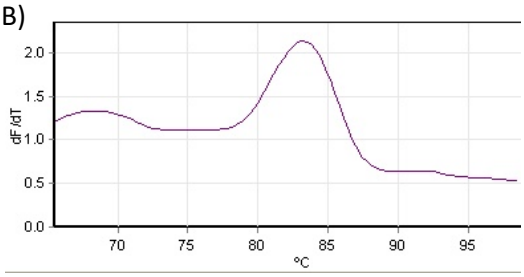


Figure AE2: Melt curve validation of qPCR target amplification

The melt curves indicated product formation around 82 °C for let-7b-5p (A), and 84 °C for miR-455-3p (B), corresponding to the melt-temperatures of the corresponding miRNAs.

Appendix F: Agarose gel reagent

a) TBE solution

10 x Tris, Boric Acid, Ethylenediaminetetraacetic acid (TBE) solution was prepared by dissolving 109 g Tris base powder,, 55.6 g Boric acid powder (Sigma-Aldrich, Misouri, USA), and 7.4 g Ethylenediaminetetraacetic acid (EDTA; Merk, Cape Town, South Africa), in nuclease-free water (volume made up to 1 l), and then mixture was the autoclaved.

The 10 X TBE solution (100 ml) was diluted in 900 ml nuclease-free water, to produce a 1 x TBE working solution. The solutions were kept at RT.

Appendix G: let-7b-5p binding sites in the *HMGA2* gene

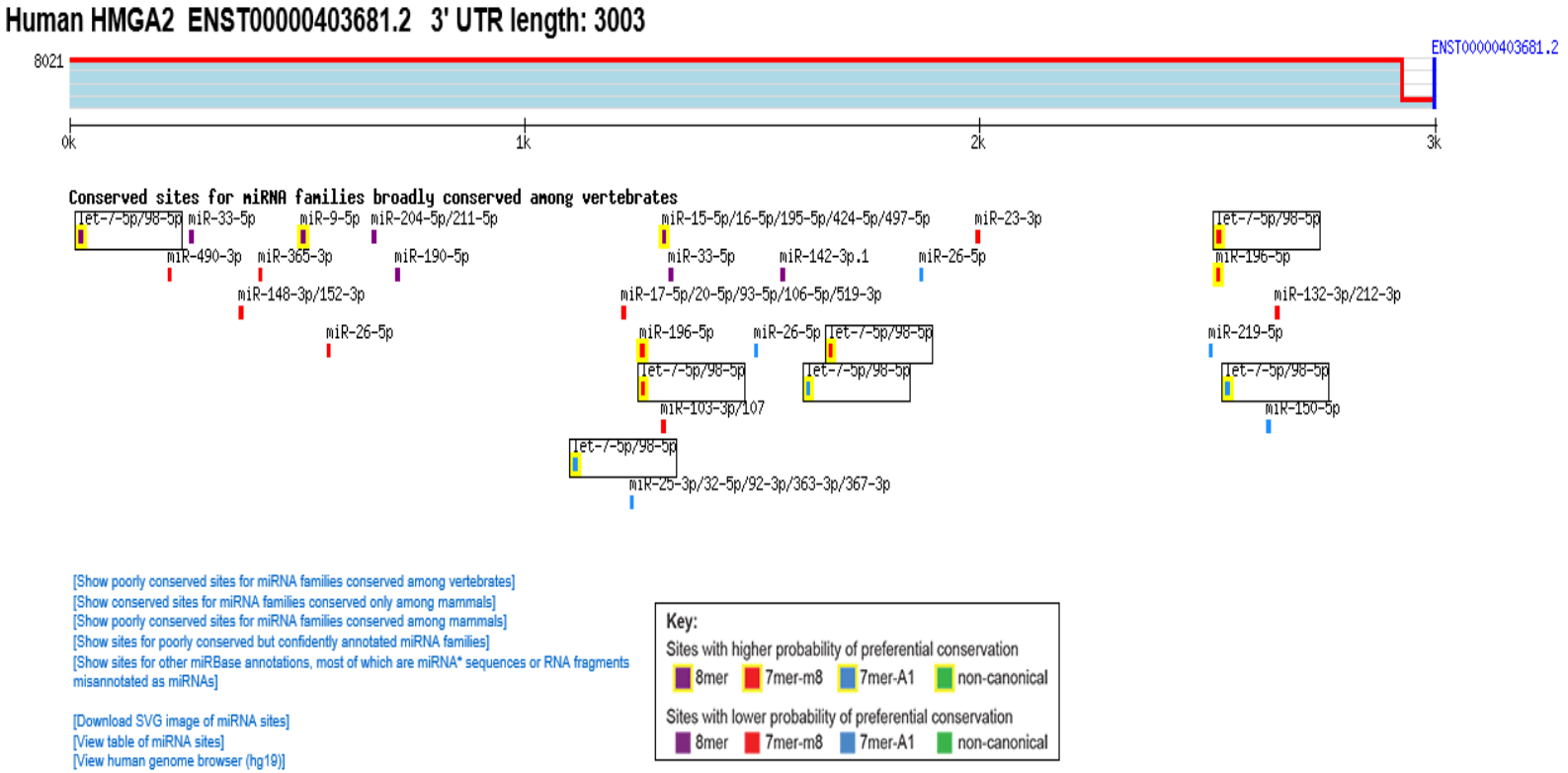


Figure AG1: Schematic representation of the 3' UTR of the *HMGA2* gene illustrating the binding sites of miRNA let-7b-5p (Agarwal *et al.*, 2015)

The position of the let-7b-5p miRNA binding sites is shown by a box in the 3'-UTR of the gene. This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (as amended).

Appendix H: HWE results for *HMGA2* and *TGF-β1* gene polymorphisms

Table AH1: HWE calculations for the *HMGA2* gene polymorphisms in the South African black population with T2DM

Polymorphism	Genotype	Participants % (n)	χ^2	p-value
rs1114167319	CC	0.03 (6)		
(C>T)	CT	0.97 (232)	215.2	< 0.001
	TT	0.0 (0)		
rs1114167320	AA	1 (238)		
(A>del)	Adel	0 (0)	UNC	UNC
	deldel	0 (0)		

del = deletion; UNC = unable to calculate; p-values of < 0.05 are not consisted with HWE

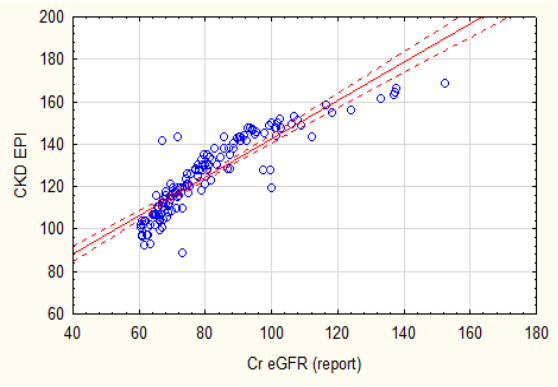
Table AH2: HWE calculations for the *TGF-β1* gene polymorphism in the South African black population with T2DM

Polymorphism	Genotype	Participants % (n)	χ^2	p-value
rs1800471(G>C)	CC	0.01 (2)		
	CG	0.70 (164)	63.4	< 0.001
	GG	0.29 (69)		

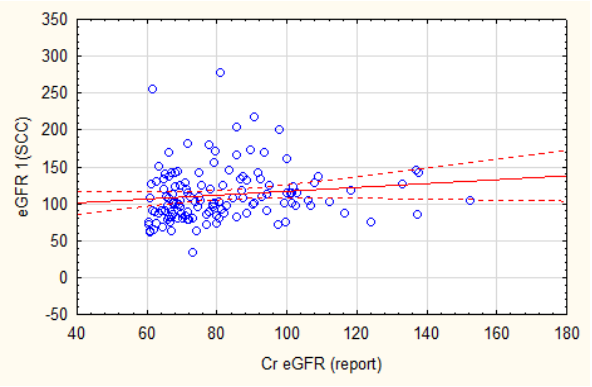
p-values of < 0.05 are not consisted with HWE

Appendix I: Correlations between eGFR calculation methods

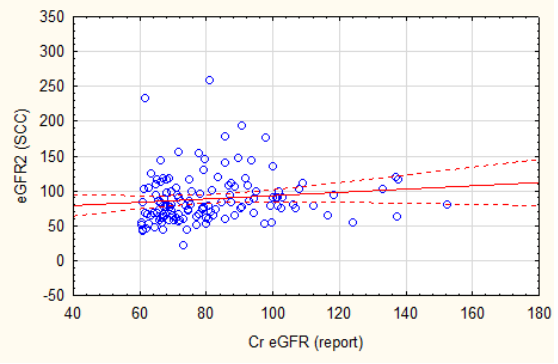
(A)
r = 0.890, p < 0.001
y = 52.2 + 0.904*x



(B)
r = 0.131, p = 0.135
y = 90.7 + 0.262*x



(C)
r = 0.122, p = 0.165
y = 69.7 + 0.236*x



(D)
r = 0.129, p = 0.145
y = 68.6 + 0.247*x

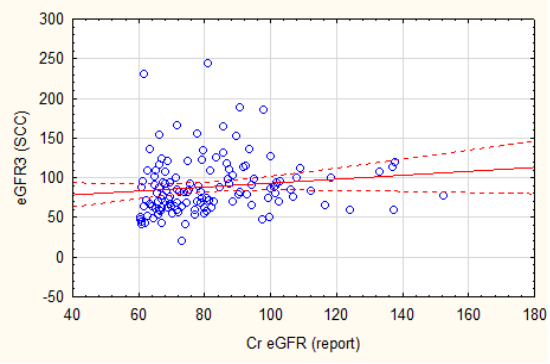
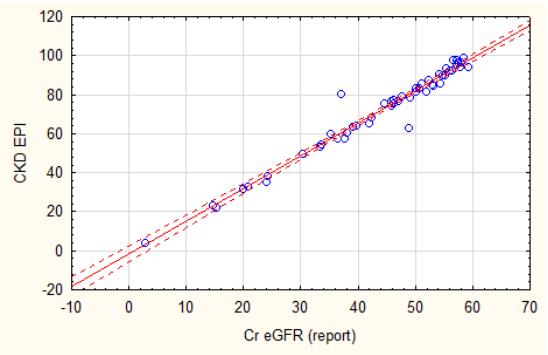


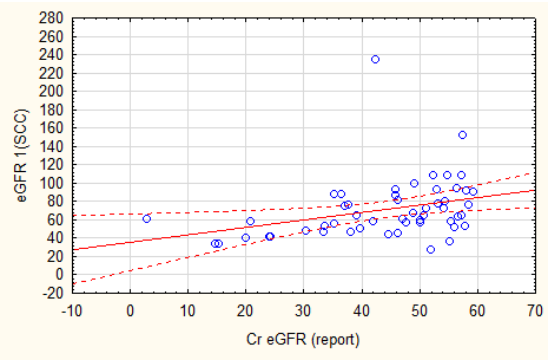
Figure AI1: Passing-Bablok regression curves comparing eGFR results from the four eGFR equations with those from the MDRD eGFR equation in black South Africans with a MDRD eGFR > 60 ml/min/1.73m^2

eGFR results using (A) CKD-EPI, (B) SCC eq1, (C) SCC eq2, and (D) SCC eq3, of participants with MDRD eGFR results > 60 ml/min/1.73m^2. The MDRD eGFR method was used as the reference method, as this method is used by the NHLS to report kidney function.

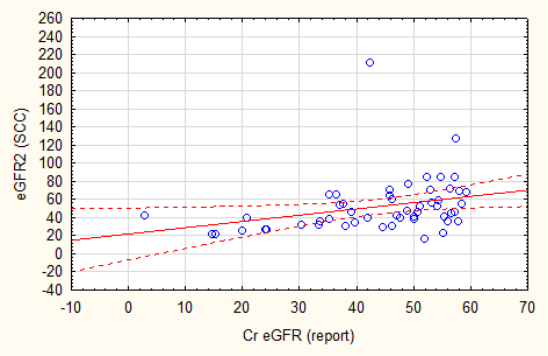
(A)
r = 0.983, p < 0.001
y = -1.57 + 1.167*x



(B)
r = 0.325, p = 0.019
y = 35.5 + 0.813*x



(C)
r = 0.301, p = 0.030
y = 21.9 + 0.691*x



(D)
r = 0.304, p = 0.029
y = 22.3 + 0.649*x

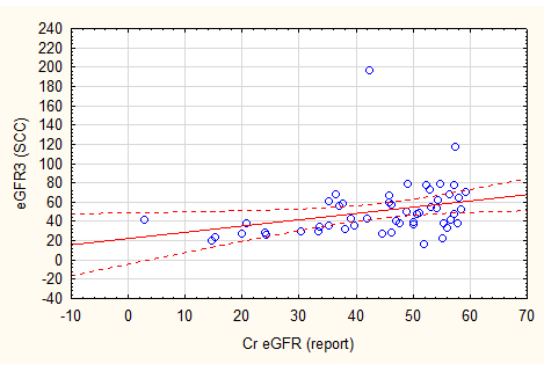


Figure AI2: Passing-Bablok regression curves comparing eGFR results from the four eGFR equations with those from the MDRD eGFR equation in black South Africans with a MDRD eGFR < 60 ml/min/1.73m^2

eGFR results using (A) CKD-EPI, (B) SCC eq1, (C) SCC eq2, and (D) SCC eq3, of participants with MDRD eGFR results < 60 ml/min/1.73m^2. The MDRD eGFR method was used as the reference method, as this method is used by the NHLS to report kidney function.

Appendix J: Turnitin receipt and report

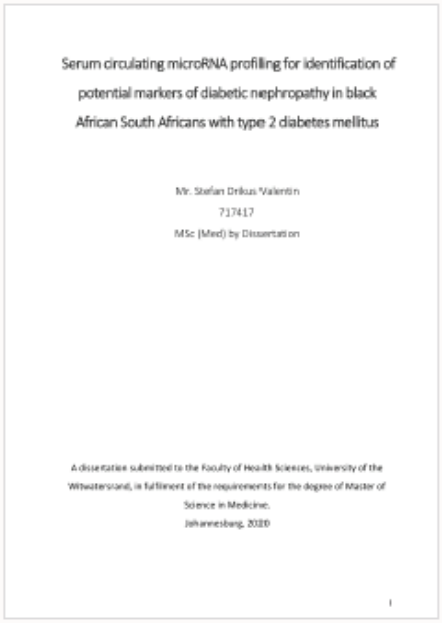
turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Stefan Valentin
Assignment title:	dissertation new
Submission title:	717417:717417_SDV_MScMed_Dis…
File name:	3-4aea-9f58-9b8ae1bc69ab_71741…
File size:	3.18M
Page count:	148
Word count:	37,788
Character count:	196,375
Submission date:	19-Apr-2020 07:55PM (UTC+0200)
Submission ID:	1301646379



Serum circulating microRNA profiling for identification of potential markers of diabetic nephropathy in black African South Africans with type 2 diabetes mellitus

Mr. Stefan Drikus Valentin

717417

MSc (Med) by Dissertation

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science in Medicine.

Johannesburg, 2020

717417:717417_SDV_MScMed_Dissertation_2020.pdf

ORIGINALITY REPORT

2%	2%	4%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	"Renal Fibrosis: Mechanisms and Therapies", Springer Science and Business Media LLC, 2019 Publication	1%
2	repository.up.ac.za Internet Source	1%
3	Submitted to Universiti Sains Malaysia Student Paper	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		

I hereby verify that I have seen the above turn it in report and the percentage similarity of 2% is acceptable.

Dr. C Padoa 20/04/2020