Investigating the functionality of African-specific variants in the *TGFB1* regulatory region and their potential role in HIV-associated kidney complications

JOY-MARI BUYS
BYSJOY001

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN
In fulfilment of the requirements for the degree

MSc(Med) Cell Biology

Faculty of Health Sciences
UNIVERSITY OF CAPE TOWN

Supervised by Prof J.M. Heckmann
Co-supervised by Prof S. Prince and
Dr M. Nel

8 September 2016

Department of Human Biology
Faculty of Health Sciences
University of Cape Town, Observatory, Cape Town, South Africa

DECLARATION

I, Joy-Mari Buys, hereby declare that the work on which this thesis is based is my original work (except where acknowledgments indicate otherwise) and that neither the whole work or any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the University of Cape Town to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed by candidate

Date: 08 September 2016

Acknowledgements

I would like to express my gratitude and thanks to the following individuals:

My supervisor **Associate Professor Jeannine Heckmann** for her support, guidance and patience throughout my MSc. I must also give thanks for her help in writing this thesis and for keeping me motivated throughout. Your enthusiasm for science has inspired me and I am grateful for all you have taught me.

My co-supervisor **Professor Sharon Prince** for her guidance and support, especially with laboratory work, and for motivating me to share my work with my peers.

My co-supervisor **Dr Melissa Nel** who is a wonderful and patient teacher. Thank you for all of your advice and for always making the time to see me when I was in crisis.

The members of the **T-Box lab** for always being willing to give their advice and help in times of need.

My parents who still don't fully understand what it is that I do but support me anyway. Thank you for being proud of me and for always being there when I needed you.

To my **brother and sister** who supported me but also provided much needed distractions in the form of exercise or movie marathons.

To my **family and friends** for their understanding and support. A special thank you so my aunt **Susan** who welcomed me into her home and who has taken care of me throughout the course of my postgraduate studies. To **Megan, Edyth, Melissa, Sally and Romney** for their help editing my thesis and for all of their motivational messages. I love all of you dearly.

The **University of Cape Town** and the **Medical Research Council of South Africa** for their financial support.

Table of Contents

List of Figures and Tables

Figures

Introduction

Abstract

Transforming growth factor beta-1 (TGF-β1) is a cytokine involved in immune modulation and tissue regeneration and has a polymorphic *TGFB1* promoter. African-specific single nucleotide polymorphisms (SNPs) have been identified in the *TGFB1* promoter and they occur at higher frequencies in South Africans with African-genetic ancestry (≈17%) compared with West and East African genomes (<5%). However, the functional significance of these SNPs has only been partially explored. In this study it was hypothesized that higher frequencies of immune-mediated disease complications in South Africans, such as HIV-associated nephropathy (HIVAN), may be influenced by functional genetic variations in *TGFB1*. To address this, the African-specific *TGFB1* haplotypes containing singular, or combinations of, *-1287 G>A* (rs11466314), *-1154 C>T* (rs35318502), *-387 C>T* (rs11466316) and *-14 G>A* (rs9282871) were investigated for their effect on *TGFB1* promoter activity. Briefly, an extended *TGFB1* regulatory region driving a luciferase reporter was used as a template to generate six *TGFB1* promoter haplotypes (referred to as H-1 through H-6 in this thesis) by site-directed mutagenesis and luciferase activity was used to measure promoter activity. The functional *TGFB1 -1347 C>T* variant was also investigated (H-5 and H-6 containing *-1347 C* and *T* alleles, respectively) because all four of the African variants more frequently co-occurred with the previously reported "lower expressing" *TGFB1 -1347 C* variant (H-1 through H-4). Transient transfection of the *TGFB1* promoter reporter constructs in two renal cell lines (RCC4+VHL and Caki-2) showed no difference between *-1347 C* (H-5) and *T* (H-6) basal promoter activity. Having at least one African variant resulted in ~5-fold loss of basal *TGFB1* promoter activity in renal cells when compared to the most common haplotype (H-5) ($p<0.05$). The repressive effect is mainly attributed to the *-387 T* variant (H-1) as the addition of other African variants on this haplotype showed no additional *TGFB1* promoter repression. The repressive effect of the *TGFB1 -387 T* variant was maintained even after *in-vitro* treatment of transfected renal cells with recombinant human TGF-β1 (rhTGF-β1).

To determine whether the African-specific *TGFB1* promoter haplotypes (H-AFR), containing *TGFB1 -387 T* and tested in luciferase assays, also impact on endogenous TGF-β1 protein levels, western blot analysis was performed on human dermal fibroblasts from patients who had been genotyped. Similar to the promoter studies, basal TGF-β1 protein levels in cells with the H-AFR were ~47% lower compared to cells without ($p=0.04$) and no difference was seen in response to hrTGF-β1. Interestingly, when western blots were screened for phosphorylated Smad3 (pSmad3) protein levels, as an indicator of the activated TGF-β1 canonical pathway, similar pSmad3 levels were observed under basal and hrTGF-β1 stimulated conditions for all haplotypes ($p>0.05$).

The possible interaction of HIV tat on the African *TGFB1* promoter variants was also assessed. Luciferase activity was measured after co-transfecting renal RCC4+VHL and HT1080 fibroblast cells with a HIV Tat expression vector and the H-6, H-5 and H-AFR promoter luciferase constructs. Results showed that the promoter activity for all *TGFB1* haplotypes was upregulated in the renal cells (≥1.6-fold; $p<0.001$). The same result was, however, only observed for *TGFB1* haplotypes H-5 and H-AFR in the fibroblast cells (≥1.4-

fold; $p<0.01$). This is interesting because no difference was seen between *TGFB1* H-5 and H-6 basal promoter activity.

To investigate whether histopathological severity of HIVAN correlated with TGF-β1 staining patterns, immunohistochemistry was performed on 20 renal biopsies from patients with HIVAN. A semi-quantitative "histo" score (H-score) was calculated by multiplying the percentage of positive cells with the intensity of the stain before comparing the scores with control biopsies (HIV-negative, n=3 and HIV-positive without HIVAN, n=3 biopsies). Compared to the HIV-positive controls the kidney tubules of HIVAN biopsies had higher H-scores. Strikingly, the interstitium of HIVAN samples stained much more prominently (17/18) than both control groups suggesting that TGF-β1 staining in the renal interstitium appear to be specific for HIVAN.

In conclusion this study shows that the African-specific haplotypes effect basal *TGFB1* promoter activity and TGF-β1 protein levels. However, they do not seem to affect the cytoplasmic TGF-β1/pSmad3 protein levels in response to rhTGF-β1 autocrine stimulation. Immunohistochemistry results suggest that TGF-β1 pathway may be prominently dysregulated in the renal interstitium of HIVAN cases.

1. Introduction

1.1 TGF-β1 as a cytokine

TGF-β1 plays an important role in immune signalling, growth and developmental processes as well as contributing to the pathogenesis of fibrotic disease in many tissues including the kidney (Olson et al. 1986; Leask & Abraham 2004). TGF-β1 protein levels are highly regulated, requiring activation from the inactive precursor present in the extracellular matrix in order to become active latent complexes (Nunes et al. 1997). As a latent complex TGF-β1 can interact with many cell types, such as muscle and kidney cells, by initiating signalling across the plasma membrane into the cell (Heldin et al. 1997). The signalling happens through induced heteromeric complexes of type I and type II receptors with serine/threonine kinase activity (Wrana et al. 1994).

TGF-β1 is usually increased by an inflammatory response such as tissue injury. In response to an increase in TGF-β1 levels, myoblasts and fibroblasts are induced to express more TGF-β1 and initiate the production of fibrogenic proteins which convert them to a myofibroblast phenotype (Li et al. 2004; Böttinger 2007). Myofibroblasts are key cellular components in fibrotic lesions and cause the accumulation of fibrogenic extracellular matrix which result in impaired function (Lehto et al. 1986; Montesano & Orci 1988; Leask & Abraham 2004). TGF-

β1 activates a host of fibrotic cytokines that help in the remodelling and repair of damaged tissue, but when overexpressed may lead to unfavourable scarring and altered healing of kidney cells (Kovacs & DiPietro 1994; Epstein et al. 1994; Wynn 2010). It is pertinent to note that TGF-β1 has been shown to have both pro-fibrotic as well as anti-inflammatory roles often occurring in different cellular contexts. Interestingly, depending on the duration of renal injury, TGF-β1 signalling is able to switch from an anti-inflammatory role in acute kidney damage to a pro-fibrotic role in chronic kidney damage. This dual functionality suggests that TGF-β1 levels and functional effects are both complex and tightly regulated in a renal context (Roberts et al. 1986; Border & Noble 1994; Wang et al. 2005).

1.2 TGF-β1 signalling

Upon ligand-induced activation of TGF-β1 receptors the receptor-complexes are able to signal to the cell nucleus via the Smad (signalling mother against decapentaplegic peptide) family of proteins (Heldin et al. 1997). Smad proteins act as messengers of gene regulation at the transcript level (Kim et al. 1989; Heldin et al. 1997; Shi & Massague 2003). Figure 1.1 illustrates the binding of TGF-β1 to the serine-threonine kinase receptors (TβRI and TβRII) on the cell surface. The TβRI-kinase then phosphorylates the receptor-regulated Smads (R-Smads) such as Smad2 and Smad3 which then forms a transcriptional complex with a common-partner Smad (Co-Smads) such as Smad4 (Lan 2011). The R-Smad/Co-Smad complex then enters the nucleus, where it binds to promoters and transcription co-factors in order to activate the transcription of genes (Leask & Abraham 2004; Lan 2011). The R-Smads (Smad2/3) and Co-Smad (Smad4) complexes are regulated by the inhibitory Smads (I-Smads) such as Smad7 (Figure 1.1) (Lan 2011).

TGF-β1 is expressed in the proximal renal tubular epithelial cells (referred to as renal tubular epithelial cells in this thesis) located in the renal cortex (Figure 1.2). The renal cortex is the outer layer of the kidney and the medulla the inner layer; the nephron is a tube-like structure spanning both of these layers (Figure 1.2). The nephron is the functional unit of the kidney where ultrafiltration occurs; podocytes, glomeruli and renal tubules are located in this area. In the kidney TGF-β1 has complex pro-fibrotic and anti-inflammatory roles and is able to switch between these two roles depending on the environmental stress as well as the mechanism by which the Smad pathway is activated or inhibited (TGF-β1-dependent and -independent signalling). In the context of kidney disease it has been shown that the downstream canonical

Smad2/3 pathway can be activated by TGF-β1 and mitogen-activated protein kinase (MAPK) signalling (also known as the ERK/p38 MAP kinase-dependent mechanism) (Wang et al. 2005; Lan 2011; X. Cheng et al. 2013).

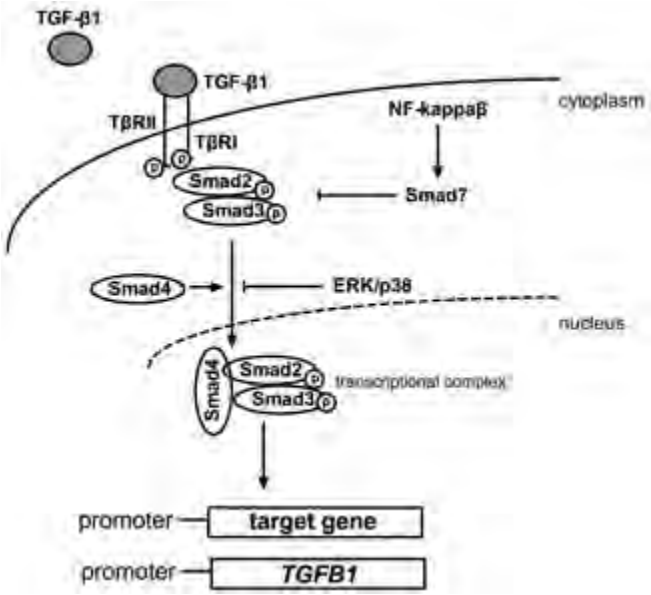


Figure 1.1: Schematic showing TGF-β1 activated Smad pathway and how this signalling pathway regulates gene expression. The involvement of NF-*K*β and the ERK/p38/MAP kinase pathways are also shown together with the inhibitory Smad7. TβR-I and TβR-II refer to type 1 and type 2 receptor and 'p' indicates phosphorylation.

Chronic kidney injury and the subsequent increase in TGF-β1 protein levels induce epithelial to mesenchymal transition (EMT). During EMT the TGF-β/Smads (Smad2 and Smad3) and BMP/Smads (Smad1, Smad5 and Smad8) counter regulate each other in order to maintain balance between the two pathways whereby one promotes inflammation and fibrosis and the other inhibits those effects, respectively (Bonniaud et al. 2005; Wang et al. 2005; Lan 2011). Previously, Smad2 knockout studies in a mouse model showed that Smad2 not only enhances the phosphorylation of Smad3 and its binding to various promoters that play a role in collagen matrix production, such as the ColA1 promoter, but may also be able to competitively inhibit the phosphorylation of Smad3 in response to TGF-β1 (Bonniaud et al. 2005; Lan 2012). The interaction between Smad2/3 and Smad4 (formation of the transcription complex) can also have an effect on the nuclear translocation of phosphorylated Smad3 and its subsequent binding to target genes. Smad7, an inhibitor of the Smad2/3 complex, can promote nuclear factor kappa-light-chain-enhancer of activated B cells–dependent (NF-*K*β) renal inflammation (Lan 2012). NF-*K*β is a protein complex that controls the transcription of genes including

cytokines and is activated during kidney disease where it regulates renal inflammation (Lan 2011; Lan 2012). It has been demonstrated that TGF-β1 can induce Smad7 mRNA expression during chronic kidney injury which leads to the activation of Smurf proteins (E3 ubiquitin-protein ligase) which in turn leads to Smad7 degradation and enhanced activation of Smad2/3 and NF-*K*β-dependent inflammatory pathways (Lan 2011). This highlights the complexity of TGF-β1 signalling during kidney disease where inhibitors, such as Smad7, can now also activate the Smad2/3 pathway.

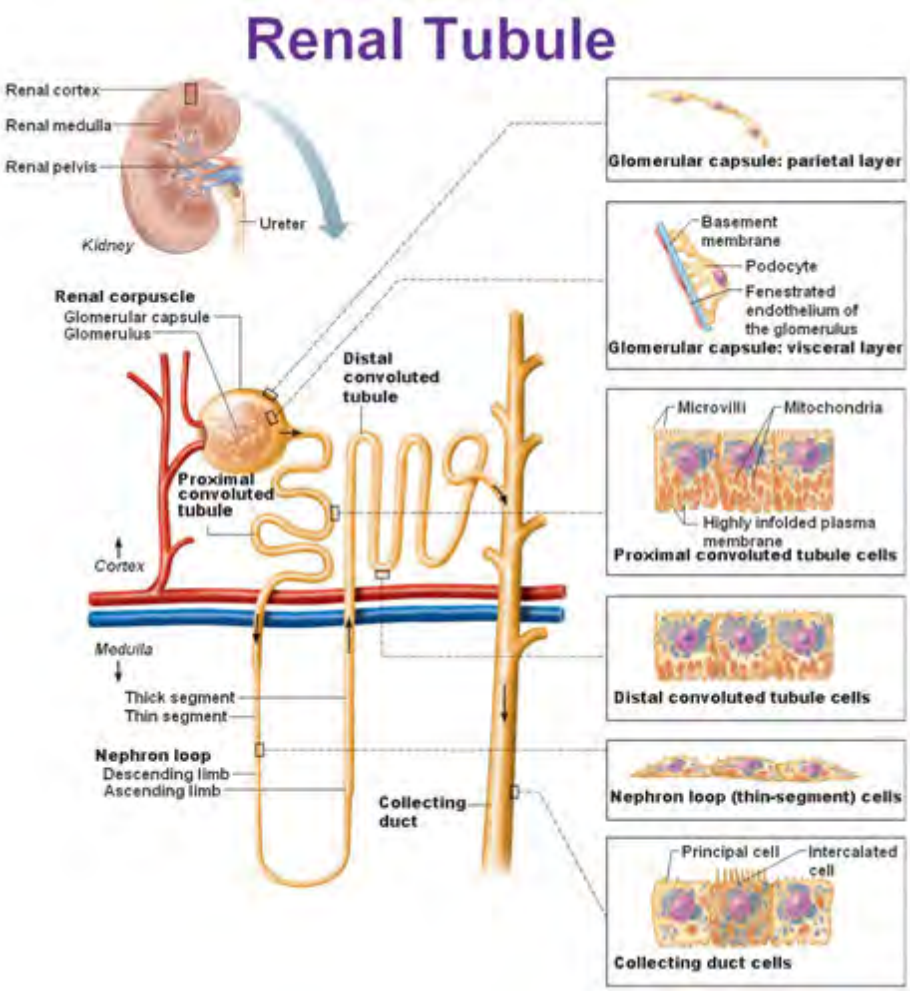


Figure 1.2: Schematic showing the anatomy of the kidney. The renal cortex contains the tubular epithelial cells where TGF-β1 is expressed during acute and chronic kidney injury. Image sourced form http://antranik.org/the-urinary-system-kidneys/

Renal cells, and others, are able to respond to autocrine and paracrine TGF-β1 signalling (Bódi et al. 1997; Nunes et al. 1997; Böttinger 2007). Auto-induction of TGF-β1 in the kidney can lead to prolonged fibrosis through the MAP kinase pathway and is also responsible for increased production of the extracellular matrix (ECM) through activation of connective tissue

growth factors and collagen proteins (Roberts et al. 1986; Bonniaud et al. 2005; Zhang et al. 2006). When there is chronic kidney injury, such as HIV-infection, prolonged TGF-β1 signalling occurs which can lead to severe fibrosis and scaring.

1.3 TGFB1 transcriptional control and promoter polymorphisms

The *TGFB1* gene has two major start sites regulated by a complex and extensive promoter which controls TGFB1 mRNA levels (Kim et al. 1989). Comprehensive analysis of the different regulatory elements, -2665bp upstream of the ATG translation start site, has revealed various functional promoter regions (shown in Figure 1.3) that act as either enhancers or negative regulators of gene transcription (Kim et al. 1989). It is known that mutations in the 5' or 3' regulatory regions of cytokines can have a significant effect on transcription and protein expression, as they have the potential to influence mRNA splicing, mRNA stability (3'UTR) and levels of gene expression (5'UTR) (Bidwell et al. 1999). The *TGFB1* promoter (5'UTR) is highly polymorphic and many of these single nucleotide polymorphisms (SNPs) have been shown to be functional as they alter binding affinity of nuclear factors/transcription factors in electrophoretic mobility shift assays (EMSA) and influence reporter gene expression in promoter studies (Shah, Rahaman, et al. 2006; Healy et al. 2009; Mohy & Fouad 2014). Since *TGFB1* is a key player in signalling networks related to immune regulation and tissue injury, functional SNPs which alter the level of this cytokine, could understandably have detrimental outcomes to tissue healing responses (Lehto et al. 1986; Roberts et al. 1986). Indeed, many studies have implicated *TGFB1* genetic variants in a variety of disease complications though few have interrogated African-specific *TGFB1* SNPs specifically (Shah, Rahaman, et al. 2006; Healy et al. 2009; Nel et al. 2015).

Nel *et al.* (2015) previously studied a tissue specific complication observed in autoimmune myasthenia gravis using an extended *TGFB1* promoter construct (Figure 1.3A), focusing on the effect of the African-specific *TGFB1 -387 C>T* variant and found that it associated with the development of the complication. Their study showed that this SNP is functional and significantly represses *TGFB1* promoter activity in reporter gene assays and patient specific fibroblasts with *-387 T* were shown to have significantly lower TGF-β1 protein levels than those with the *-387 C* allele (3). Therefore, there is evidence that the *TGFB1 -387 C>T* variant may impact *TGFB1* expression levels and its effect may be the same in cells of kidney origin. Nel *et al.* (2015) also investigated the functionality of the African-specific *TGFB1 -14 G>A*

SNP in fibroblast and myoblast cells and found it not to be functional. The functionality of the two other African-specific SNPs, *-1287 G>A* and *-1154 C>T*, have not been investigated using the extended *TGFB1* promoter and were therefore included in this study. Nel *et al.* (2015) also investigated the *TGFB1 -1347 C>T* SNP site, also known as the *-509 C>T*, and found that the *T* allele increased *TGFB1* promoter activity in fibroblast and myoblast cells. Two shorter *TGFB1* promoter constructs have also been used in previous studies to investigate the functionality of the *TGFB1 -1347 C>T* (*-509 C>T*), *-1287 G>A* and *-1154 C>T* variants, but they lacked some of the regulatory regions and the functional *-387 C>T* variant (Figure 1.3B and C).

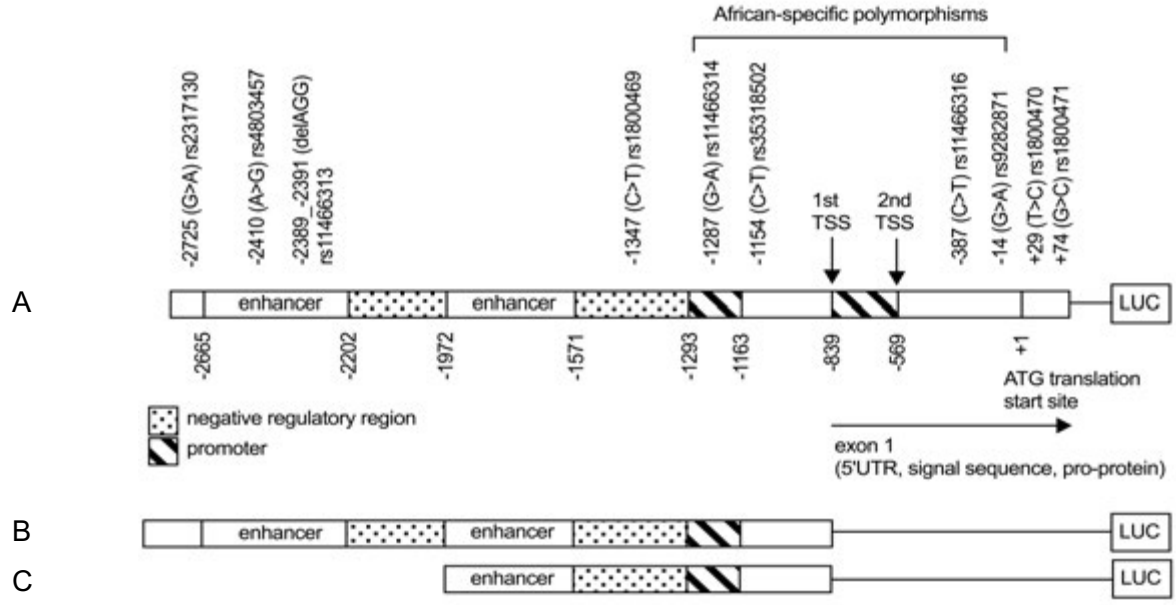


Figure 1.3: Schematic of the human extended *TGFB1* 5'UTR cloned into a luciferase reporter construct with 10 polymorphic sites, four of which are African-specific. B and C are two shorter/incomplete promoter constructs previously used in functional studies. Figure sourced from Nel et al. (2015). Regulatory regions and polymorphisms indicated relative to the ATG translation start site.

A brief synopsis of TGFB1 and renal complications

The kidney is particularly susceptible to fibrotic disease as the production of extracellular matrix; appearance of myofibroblasts and inflammatory cells are triggered in response to renal insults (Montesano & Orci 1988; Böttinger 2007). The role of TGF-β1 in the pathogenesis of renal fibrosis is well established and it is of interest to note that there is research to suggest that both genetic-ancestry and HIV status may predispose to the development of renal

complications such as hypertensive kidney disease and HIV-associated nephropathy (HIVAN) (Suthanthiran et al. 1998; Eiser 2010). Several conditions related to TGF-β1 appear to be more common in African-American individuals, suggesting that there is a genetic component that contributes to the development of adverse disease complications upon immune insult. In addition, nephrologists in South Africa (SA) have reported that the incidence of HIVAN appears to be higher in individuals with African-genetic ancestry (Wearne 2012; Kasembeli 2015). The fact that TGF-β1 signalling is so critical in renal homeostasis leads one to speculate that *TGFB1* promoter variation may also be involved in modulating adverse kidney disease outcomes in Africans with HIV.

Briefly, HIVAN is characterised by histopathological features such as collapsing glomeruli, interstitial inflammation, interstitial fibrosis and tubular abnormalities (microcyst formation) within the renal cortex. These pathological features can reduce the excretory renal function and lead to end-stage renal failure (Liu 2006; Wyatt & Klotman 2009; Medapalli et al. 2011). TGF-β1 induces fibroblast proliferation and the secretion of extracellular matrix components in the case of acute injury or chronic/persistent kidney injury (August et al. 2009; Wynn 2010). In acute injury, the epithelial cells nearest the site of injury will produce low levels of TGF-β1 which stimulate fibroblast proliferation but this modest increase in TGF-β1 levels does not perturb fibroblasts from their resting state (Figure 1.4). However, the epithelial cells will enter the cell cycle and start to proliferate and will repair the site of injury (Figure 1.4) (Yang et al. 2010). When there is chronic kidney injury the majority of epithelial cells will arrest between G2 and M phases of the cell cycle and stimulate increased production of TGF-β1, in an autocrine/paracrine fashion (Overall et al. 1989; Wynn 2010). The increase in TGF-β1 simultaneously results in a decrease of proteins such as matrix metalloproteinases, which degrade the extracellular matrix (ECM), and directs the activated fibroblasts to differentiate into myofibroblasts (referred to as EMT) (Wynn 2010). These myofibroblasts will then activate the production of extracellular matrix proteins, such as collagen I, which ultimately lead to a pathogenic fibrotic response instead of normal kidney repair (Figure 1.4) (Liu 2006; Lan 2011).

Increased levels of TGF-β1 in the kidney have been associated with HIVAN where fibrosis is activated by changes in cytokine release during HIV infection (Abbott et al. 2001; X. Cheng et al. 2013). Studies on African-American patients with HIVAN have shown differential

expression of *TGFB1* in their kidneys compared to Caucasians (Abbott et al. 2001; Eiser 2010). Thus, this study hypothesised that excessive upregulation of the TGF-β1 signalling pathway may be implicated in Africans who develop HIVAN and in turn may implicate functional African-specific *TGFB1* promoter polymorphisms. Therefore, this study investigated the African *TGFB1* promoter SNPs (*-1287 G>A*, *-1154 C>T* , *-387 C>T* and *-14 G>A*) using an extended *TGFB1* promoter, which has not been previously used to examine functionality of these variants in kidney cells.

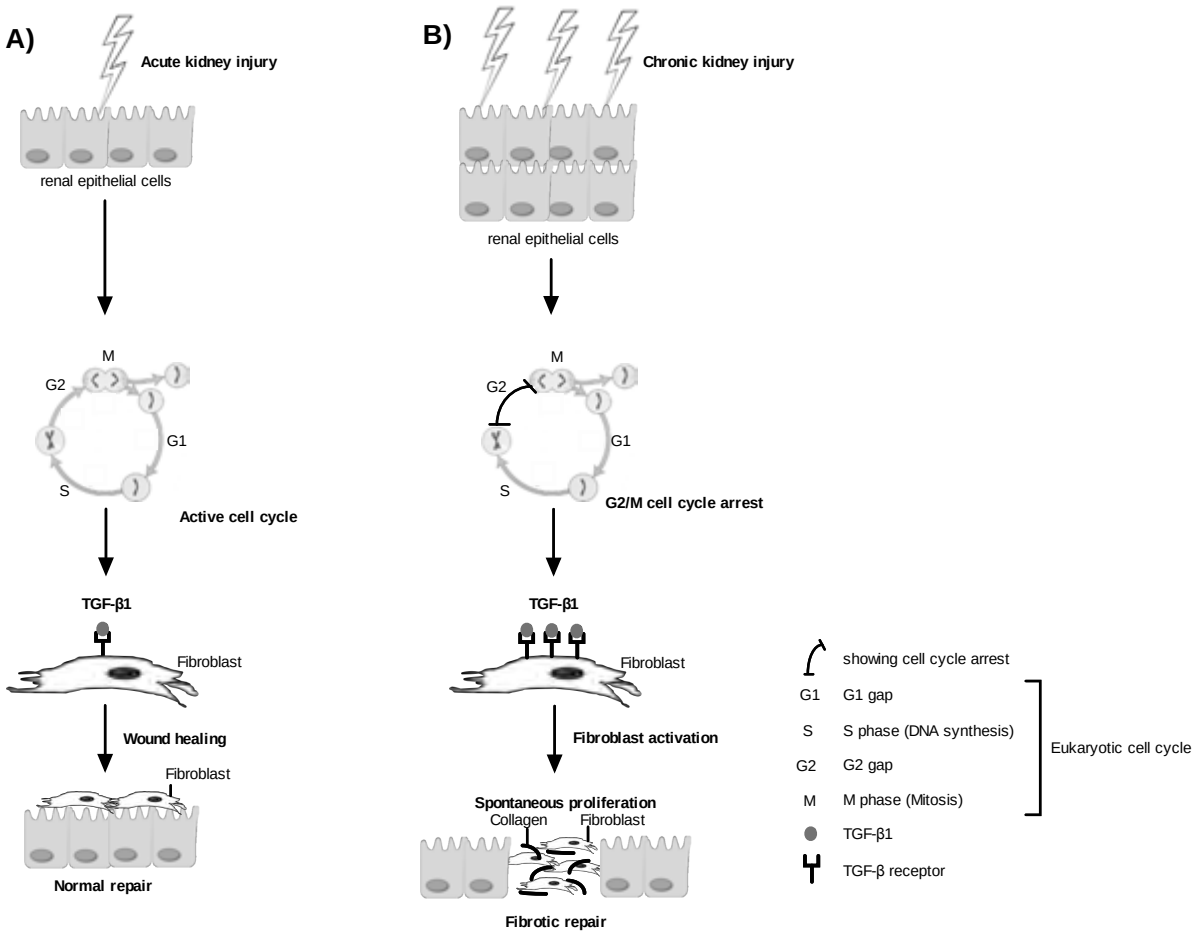


Figure 1.4: Schematic showing TGF-β1 signalling in the kidney during A) acute and B) chronic kidney injury. Figure adapted from Wynn *et al.* (2010).

HIVAN risk alleles in the Apoloprotein-1 gene

A region on chromosome 22q12 encoding Apoliprotein-1 (*APOL1*) has been associated with increased risk of HIVAN in patients with African-ancestry (Genovese et al. 2010). Previous genome wide association studies (GWAS) revealed that the increased risk is due to three variants in the Apoloprotein-1 gene (*APOL1*) which are not expressed in Europeans (Kopp et

al. 2011; Tayo et al. 2013; Kasembeli et al. 2015). Two tightly linked G1 risk alleles, termed $G1^{S342G}$ and $G1^{I384M}$, and one 6 base pair deletion termed G2 ($G2^{\Delta 6}$) have been identified as the HIVAN risk alleles (Genovese et al. 2010; Kasembeli et al. 2015). The G1 and G2 alleles are mutually exclusive, never occurring on the same allele. The G1 risk variants have been identified as the main contributors to increased susceptibility of HIVAN in patients with African-ancestry (Kasembeli et al. 2015). Kasembeli *et al.* (2015) showed that the $G1^{GM}$ (containing both G1 risk alleles but not the G2 allele) haplotype is the most common occurring haplotype in a Johannesburg HIVAN group with a haplotype frequency of 0.53 compared to a haplotype frequency of 0.06 in population controls (Kasembeli et al. 2015). Therefore this study included the *APOL1* risk alleles in the genotyping of a Cape Town HIVAN sample with African-ancestry comprising 20 individuals. In this work the minor allele frequencies and haplotype frequencies of the *APOL1* risk alleles will be determined for a Cape Town HIVAN group.

Apoloprotein-1 (APOL1) can be expressed extracellularly, where it provides resistance against trypanosomes, or intracellularly where it functions as a crucial component in the human host defence as well as cellular homeostasis (Hu et al. 2012). Intracellular expression of APOL1 in renal tubular epithelial cells and podocytes can induce low levels of autophagy which act to maintain cellular homeostasis in the kidney (Hu et al. 2012). Variants in the *APOL1* gene can lead to an upregulation of APOL1 in renal tubular epithelial cells which in turn can induce apoptotic or proliferative changes in these cells (Hu et al. 2012).

microRNAs in HIVAN

In addition to variants in the 5'UTR there are also mechanisms in the 3'UTR that can influence TGF-β1 levels in the kidney. MicroRNA-200 (miRNA-200) has recently been shown to play a role in the pathogenesis of HIVAN (Kato et al. 2011; K. Cheng et al. 2013). Its downregulation in podocytes leads to the collapse of the glomerulus, more commonly known as collapsing focal segmental glomerulosclerosis (FSGS), and is also able to induce EMT of the tubular epithelial cells (K. Cheng et al. 2013).

HIV Tat increases TGFB1 mRNA expression

HIV-1-trans-activating (Tat) protein is encoded by the Tat gene and functions as an early RNA binding protein enhancing the efficiency of viral transcription (Lotz et al. 1994). HIV

Tat has been predicted to have a trans-activation response element (TAR-like) binding sequence on the *TGFB1* regulatory region, but seems unlikely to be able to bind due to the absence of a stem-loop structure (Thatikunta et al. 1997). Therefore it is speculated that the enhanced TGFB1 mRNA expression seen in response to HIV Tat dosage may be due to the induction or promoter binding proteins. However, functional studies on the effect of HIV Tat on the *TGFB1* promoter activity are lacking. Yamamoto et al. 1999 showed that there was an increase in TGF-β1 staining levels in renal biopsies from HIV-positive patients with African-ancestry when compared with HIV-positive normal and HIV-negative kidneys, identifying TGF-β1 as a mediating factor in the pathogenesis of HIVAN and end-stage renal disease (Yamamoto et al. 1999). In order to investigate whether HIV Tat has an effect on the *TGFB1* regulatory region, and specifically the African-specific variants, this study set out to perform luciferase reporter assays using co-transfections of the *TGFB1* African haplotypes as well as an HIV Tat expression vector (provided by Dr. Mowla from the University of Cape Town) in renal and fibroblast cells (representing the renal tubular epithelial and interstitial cells, respectively).

2. Hypothesis

Higher frequencies of immune-mediated kidney disease complications in South Africans, such as HIV-associated nephropathy (HIVAN), may be influenced by functional genetic variations in *TGFB1*.

Aims and objectives

Aim 1: To determine whether the African-specific variants in the regulatory region of TGFB1 are functional

Objectives

- Perform haplotype analysis on the minor alleles of the *TGFB1* promoter
- Generate six African-specific *TGFB1* haplotypes using site-directed mutagenesis
- Transiently transfect the six *TGFB1* haplotypes in two renal cell lines using a reporter gene assay to measure basal and stimulated promoter activity
- Perform western blot analysis on primary dermal fibroblasts with known genotypes and compare TGF-β1 and pSmad3 protein levels between the haplotypes at basal and stimulated (rhTGF-β1) conditions

- Co-transfect renal and fibroblast cell lines with the *TGFB1* haplotypes and a HIV-tat expression construct and compare its effect on basal *TGFB1* promoter activity using a reporter gene assay

Aim 2: To investigate whether the African variants have a potential role in HIV-associated kidney complications

Objectives

- Perform immunohistochemistry on 20 biopsies from patients with HIV-associated nephropathy (HIVAN) and compare TGF-β1 positive staining patterns of HIV-positive and –negative control biopsies without HIVAN
- Perform genotyping for the functional African-specific *-387 C>T TGFB1* variant as well as the *APOL1* risk variants for the 20 individuals with HIVAN

3. Materials and Methods

3.1. *Patients*

DNA was extracted from peripheral blood samples of HIV-positive South Africans with African-ancestry who presented with HIV-associated nephropathy (HIVAN) at Groote Schuur hospital (n=20 patients). HIVAN was previously characterised by clinicians at Groote Schuur hospital and histopathological features were defined as a constellation of glomerular, interstitial and tubular abnormalities (Wearne et al. 2012). These blood samples were genotyped for the variants in the upstream regulatory region of *TGFB1* (5'UTR) (as previously described by Nel *et al.,* 2015) as well as G1 and G2 risk alleles in *APOL1*.

This study was approved by the University of Cape Town Health Sciences Faculty Research Ethics committee (HREC#591/2014 for TGF-β1 and #491/2008 for HIVAN). Informed consent was obtained for all participants.

3.2. *Sequencing of the APOL1 variants*

Sequencing primers were designed to sequence the end of the *APOL1* gene (exon 6, chromosome 22) based on previously used primer sets (forward and reverse primers) from Canaud *et al.* (2013). Unfortunately, these primers yielded "background noise" or odd peaks on the chromatogram, which suggests that there was non-specific binding of the primers.

Therefore, a new primer set was designed by Inqaba Biotechnical Industries (Pty) Ltd (South Africa) and the Sanger sequencing approach was used to sequence the HIVAN samples for *APOL1*. Chromatograms were analysed using the CLC Bio Main workbench version primer design tool.

3.3. *Haplotype analysis*

Four African-specific single nucleotide polymorphisms (SNPs) have been identified in the *TGFB1* regulatory region (*TGFB1 -1287 G>A, -1154 C>T, -387 C>T* and *-14G>A*) are shown in figure 1.3. Haplotype analysis of the unphased genotyping data from the 1000 Genomes Project Phase 3 European and African super-populations was performed in Haploview (Broad Institute, MIT and Harvard) using the confidence interval (nearest neighbour) algorithm described by (Gabriel et al. 2002; Barrett et al. 2005). Additionally, in this thesis the default confidence bound cut-offs were customized to only include *TGFB1 -1347C>T, -1298G>A, -1154*C>T, *-387C>T* and *-14G>A* polymorphic sites. The linkage disequilibrium plot for these sites are shown in figure 4.1B.

The genotyping data from South Africans with African genetic ancestry (n=344 chromosomes) previously genotyped and reported (Nel *et al.*, 2015) were used in *TGFB1* haplotype analysis using Haploview software (Broad Institute, MIT and Harvard, accessible at www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview).

3.4. *TGFB1 promoter-reporter gene constructs*

As described in Nel *et al.* (2015) a 2856 base-pair fragment of the *TGFB1* regulatory region was cloned upstream of a luciferase reporter gene in the pGL3 Basic Firefly Luciferase Reporter Vector (Promega, Madison, WI, USA). The cloned *TGFB1* promoter fragment was amplified from genomic DNA and corresponds with base pairs -2750 to +106 numbered relative to the ATG translation start site and is identical to that of the NCBI reference sequence NC_000019.10 from nucleotide positions 41352723 to 41355915 (chromosome 19, GRCh38.p2 primary assembly 12-03-15). This promoter drives expression of the firefly (*Photinus pyralis*) and renilla reporter (*Renilla reniformis*) enzymes by addition of Luciferase Assay Reagent II (LARII) to generate a stable luminescent signal (firefly) which is quantified before addition of the Stop & Glo® reagent to measure the renilla reaction (Promega Dual-Luciferase Reporter Assay System).

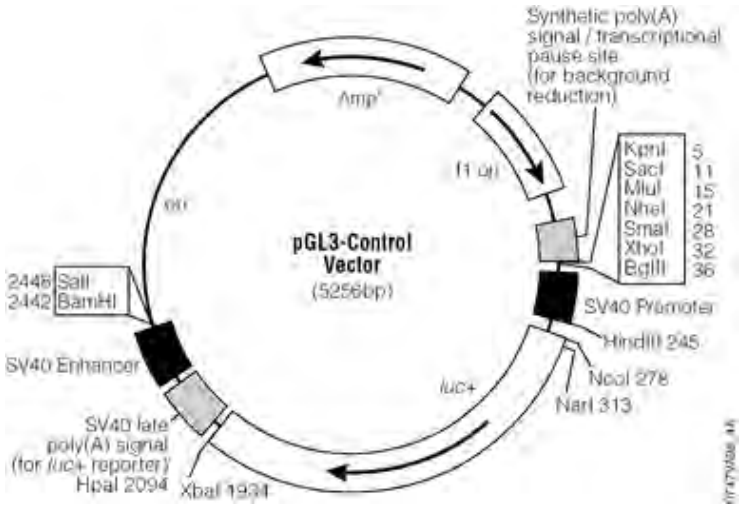


Figure 3.1: pGL3 Basic Firefly reporter vector construct. The diagram depicts the position of the Ampicillin resistance gene (Amp^r), multiple cloning site (upstream of SV40 promoter), and the luciferase gene (luc+). The luciferase gene lies downstream of the SV40 promoter and it is in this region where the *TGFB1* promoter sequence was cloned. Image sourced from: https://worldwide.promega.com/products/reporter-assays-and-transfection/reporter-vectors-and-cell-lines/pgl3-luciferase-reporter-vectors/.

3.5. *Plasmid constructs*

The pRL-TK vector/renilla luciferase vector (Promega, Madison, WI, USA) was used as an internal control for transfection efficiency. To create the various *TGFB1* promoter haplotype constructs, the *-1347 C>T, -1287 G>A, -1154 C>T, -387 C>T* and *-14 G>A* promoter polymorphisms were introduced into the cloned *TGFB1* promoter reporter construct template by site-directed mutagenesis using the QuickChange kit (Stratagene, La Jolla, CA, USA). This was done sequentially to generate *TGFB1* promoter haplotypes H-1 to H-6 as shown in figure 4.1C.

3.5.1. *Polymerase chain reaction (PCR) and Dpn1 digestion*

Point mutations were introduced into the *TGFB1* promoter reporter construct using site-directed mutagenesis, appropriate primer pairs were used to create the desired mutations. In the site-directed mutagenesis approach, two products are produced during a polymerase chain reaction (PCR), one of which is the mutated plasmid achieved by addition of site-directed mutagenesis primers that contain the desired point mutation. The second product is the parental template DNA, which will be digested by the endonuclease Dpn1. Primers were

designed using the QuickChange Primer Design Tool accessible here: www.genomics.agilent.com/primerDesignProgram.jsp

Table 3.1: Site-directed mutagenesis primers used to introduce each of the *TGFB1* SNPs to create the *TGFB1* promoter haplotypes

Site and direction	*Primer sequence 5' to 3'*
-1347 C>T forward	GAGGACTGGGAAGGTAGGAAGTCCACAGGACAACGGG
-1347 C>T reverse	CTCCTGACCCTTCCATCCTTCAGGTGTCCTGTTGCCC
-1287 G>A forward	AGAAGAGGGTTTGTCAACATGGGGGCCTCC
-1287 G>A reverse	GGAGGCCCCCATGTTGACAAACCCTCTTCT
-1154 C>T forward	GCGGGCGGTATGGGTCACCAGAGAAAG
-1154 C>T reverse	CTTTCTCTGGTGACCCATACCGCCCGC
-387 C>T forward	TCGCCGACCCGGCTTCCCGCAAAG
-387 C>T reverse	CTTTGCGGGAAGCCGGGTCGGCGA
-14 G>A forward	GGAGGCGGCGTCCCCCGGCAC
-14 G>A reverse	GTGCCGGGGGACGCCGCCTCC

Capital letters indicate nucleobases where A represents Adenine, G represents Guanine, C represents Cytosine and T represents Thymine.

The standard PCR mix contained: 2x KAPA HiFi Hotstart Ready Mix, consisting of DNA polymerase (Kapa Biosystems, Cape Town, South Africa), 10ng DNA template, 10µM forward primer and 10µM reverse primer, made up in a total volume of 25µl. The parameters for the PCR amplification was: 2 minutes (min) at 95°C for 1 cycle, 20 seconds (sec) at 98°C, 15 sec at 65°C, 4 min at 72°C, for 16 cycles. PCR was performed using an Applied Biosystems 2720 Thermal cycler. The PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The PCR product was first assessed by agarose gel electrophoresis (appendix) to confirm successful amplification and then purified using 1µl Dpn1 endonuclease (10U/µl) for 1 hour at 37°C. This digestion of the parental DNA was done in order to allow for selection of the DNA containing the mutation. The digest was followed by heat inactivation of the

endonuclease Dpn1 at 80°C for 20min. Samples were stored at 4°C and used for transformation experiments on the same day.

3.5.2. *Sequence verification*

After the Dpn1 digest, a small sample of the *TGFB1* promoter constructs generated by SDM were genotyped by for the *TGFB1* promoter variants. The Sanger sequencing approach was used by Inqaba Biotechnical Industries (Pty) Ltd (South Africa) in order to verify that the point mutations were successfully introduced during SDM. Sequencing primers were designed based on the principles of primer walking, where primers are designed consecutively to extend along an unknown sequence. Primers were designed using the CLC Bio Main workbench version primer design tool. To account for dye blobs and mismatches that occur during the starting point of annealing, the primers were designed to anneal 100 bp upstream of the region of interest. Primer parameters were as follows: the length of the primers were between 16-21 base pairs with a melting temperature (Tm) of 51-59°C where the DNA:primer complex will dissociate and become single stranded, guanine-cytosine content (GC content) of 40-60% which determines the stability of the primer and a guanine or a cytosine nucleic acid base on both ends.

Table 3.2: Sequencing primers for the *TGFB1* regulatory region

Name	Sequence 5'-3'	Tm (°C)	GC content	Special modifications	Coverage site
JoySeq3F	CTATCGCCTGCACACA	54.3	56%		*-1347 C>T, -1287 G>A* and *-1154 C>T*
JoySeq4F	CTGGTCCTCTTTCTCTG	53.23	53%	*	*-387 C>T*
JoySeq2R	AAGTCTTTGCCGGGAGG	56.69	56%	**	*-14 G>A*

Table 3.1 shows the specific parameters and sequences for each of the primers. The JoySeq4F primer reaction needed treatment with dimethyl sulphoxide (DMSO) due to the GC rich region (referenced by *). The JoySeq2R primer was modified (referenced **) to be degenerative at the site marked in red to enable recognition of either a C or a T base, as it falls over a single polymorphic site (*TGFB1 -387 C>T*).

3.5.3. *DH5α competent bacterial cells*

Competent bacterial cells were generated by inoculation of Luria Broth (LB) (appendix) with the DH5α strain of *E. coli.* The culture was grown overnight, with shaking at 37°C. A small amount of the overnight culture was then used to inoculate 100ml of LB and incubated for 2-3 hours until the bacterial cells reached log growth phase (defined as an OD_{595} reading of 0.5-0.8). The culture was then centrifuged at 3000rpm for 10min and the supernatant discarded before the pellet was resuspended in ice-cold 100mM calcium chloride ($CaCl_2$). The resuspended cells were kept on ice for 1 hour before centrifugation at 3000rpm for 10min at 4°C, after which the supernatant was discarded. The pellet was resuspended in a final volume of 1ml 100mM $CaCl_2$. Competent cells were kept overnight at 4°C before immediate use or stored in 50% glycerol (final volume 15%), snap-frozen in liquid nitrogen and kept at -80°C until use.

3.5.4. *Transformation*

20ng DNA was added to a thawed glycerol stock of competent DH5α *E. coli* cells and incubated on ice for 15 min. The cells and DNA were then subjected to heat shock at 42°C for 45 seconds before incubating on ice for 5min. This rendered the bacterial cell membrane permeable, allowing the DNA to move into the cells were it would later be expressed. LB broth was added to the competent cell-DNA mixture and incubated at 37°C at 200rpm for 30min. The resuspension was then plated onto a LB agar plate containing ampicillin (100µg/ml) using aseptic spreading technique.

The plates were inverted and incubated for 14-16 hours at 37°C. Antibiotics were added to allow for antibiotic selection in order to achieve single colony formation and for the selection of cells expressing the transformed plasmid. Negative controls containing untransformed bacterial cells were included to test whether the competent cells were not inherently antibiotic resistant. The single transformed colonies were used in small- and large-scale DNA preparations.

3.5.5. *Small-scale (mini) preparation of plasmid DNA*

Small-scale preparations were used to purify the plasmid DNA for further use in experiments and verification via Sanger sequencing. Following transformation, a single transformed colony was picked by sterile pipette tip or toothpick and used to inoculate a small volume of LB culture containing 1% ampicillin. The culture was grown at 200rmp and 37°C overnight

before being purified using the QIAprep® Spin Miniprep Kit (Qiagen, Hilden, Germany). Elutions were performed using the elution buffer (buffer EB, Qiagen, Hilden, Germany).

The samples were run on a 1% agarose gel at 70V for 45 min in order to visualise the integrity of the DNA and to compare the plasmid to the appropriate positive control. Positive controls were used to evaluate specificity of primer binding. Verification of the point mutations were performed using the Sanger sequencing approach (Inqaba Biotechnical Industries (Pty) Ltd, Pretoria, South Africa). Unused DNA was stored at -20°C.

3.5.6. *Large-scale (maxi) preparation of plasmid DNA*

The large-scale DNA preparations were performed on selected mutant constructs for further use in luciferase assays. The Promega PureYield™ Plasmid Maxiprep system (Promega, Madison, WI, USA) was used according to the manufacturer's instruction. Briefly: a single transformed colony from a LB ampicillin plate was picked with a sterile pipette tip/toothpick and used to inoculate a small volume of LB containing 1% ampicillin. The culture was grown for 8 hours at 200rmp and 37°C before 1ml was used to inoculate a bigger volume of LB with 1% ampicillin. The culture was grown overnight at 200rpm and 37°C before it underwent centrifugation at 5000xg for 10 min at room temperature (RT) in a Beckman coulter Acanti J-E centrifuge with a JA-10 rotor. The supernatant was discarded after centrifugation and the pellet was resuspended in Cell Resuspension Solution. Next, Cell Lysis Solution was added in order to lyse the cells and denature the chromosomal DNA. The tube was inverted and allowed to incubate for 3 min at RT before the addition of Neutralising Solution in order to neutralise the cell lysate and to allow for the chromosomal DNA to re-aggregate with cellular proteins. The tube was inverted several times to ensure precipitation of cellular debris before centrifugation at 7000xg for 30min at RT. This separated the chromosomal DNA, protein and cell debris from the rest of the solution containing plasmid DNA. The DNA was then purified by passing it through a PureYield™ Maxi Binding Column attached to a PureYield™ Maxi Clearing Column using a vacuum manifold and performing several washes until the DNA was successfully bound to the membrane. The membrane was dried before elutions were performed using pre-warmed nuclease free water and centrifugation at 2000xg for 5min at RT.

DNA concentration and integrity was calculated and analysed by nanodrop (ND-1000 v3.5.2 software, Inqaba Biotechnical Industries, Pretoria, South Africa) before maxipreps were

stored at -20°C and aliquots were prepared to minimize the impact of multiple freeze thaw cycles. The purified plasmid DNA samples were run on a 1% agarose gel at 70V for 45 min to visualize the integrity of the DNA and to compare it to an appropriate positive control.

3.6. *Cell culture*

RCC4+VHL (human renal cell carcinoma stably transfected with pcDNA3-VHL), Caki-2 (human clear cell renal cell carcinoma), HT1080 (human fibrosarcoma) and dermal fibroblast cells were cultured in Dulbecco's Modified Eagle's medium (DMEM) (Whitehead Scientific, Cape Town, South Africa) supplemented with 10% fetal bovine serum (FBS) (Scientific Group, Johannesburg, South Africa), 100U/ml penicillin and streptomycin (Pen/Strep) (Sigma-Aldrich, St Louis, MO, USA). The FBS was added to provide the cells with nourishment and growth factors in order to proliferate. The antibiotics penicillin and streptomycin were added in order to prevent infection of the cell cultures. All cells were maintained in an incubator at 37°C (5% CO_2 and 65% humidity). The medium was replaced every 2-3 days and sterile technique was practised when culturing the cells. Regular tests for mycoplasma infection were performed.

3.6.1. *Mycoplasma test*

Cells were grown on a coverslip for 3-5 days in antibiotic-free medium and then fixed using a 1:3 mixture of glacial acetic acid and methanol for 5 sec. Following fixation, the cells were briefly rinsed with water before air-drying at RT for 5-10 min. The cells were then stained with a working solution of Hoechst (0.5µg/ml) for 10sec. The excess stain was washed off with water before mounting the coverslip onto a glass slide using mounting fluid (see appendix). The cells were then viewed by fluorescence microscopy using the Axiovert fluorescent microscope and software for analysis (Zeiss, Oberkochen, Germany).

3.7. *Luciferase assays*

3.7.1. *Transfection*

RCC4+VHL cells were plated at 5 $x10^4$ cells per well and Caki-2 and HT1080 cells were plated at 7.5 $x10^4$ cells per well of a 12-well plate one day prior to transfection. Sixteen hours after plating cells were transiently transfected using the XtremeGENE HP DNA transfection reagent according to manufacturer's instructions (Roche, Basel, Switzerland). The transfection reagent is comprised of lipids and acts by forming a complex with the transfected DNA, which is then transported into the cells. Dose curves were performed in all of the cell

lines with varying doses of *TGFB1* promoter constructs in order to establish the optimal DNA concentrations for transfections. Each transfection and subsequent luciferase assay was performed in duplicate, where the pGL3-empty reporter plasmid and *TGFB1* promoter constructs were transfected into the different cells.

For co-transfection experiments, XtremeGene HP DNA transfection reagent was also used to transfect the relevant *TGFB1* promoter haplotypes with 350ng pcDNA3.1-HIV tat expression construct made up to a total of 550ng DNA with the pcDNA3.1 empty vector. For all transfection experiments 50ng of the pRL-TK vector (Promega, Madison, WI, USA), which contains the HSV-thymidine kinase promoter driving the expression of a renilla reporter (*Renilla reniformis*), was used as an internal control. Unfortunately, the renilla reporter was independently regulated by the HIV tat construct and we therefor did not include the renilla luciferase values in our analysis. Instead β-galactosidase assays were used to account for transfection efficiency (see section 3.7.4)

3.7.2. *In vitro treatment of TGFB1 promoters with human recombinant TGF-β1*

RCC4+VHL and HT1080 cells were plated in 12-well plates as previously described (section 3.7.1.) and transient transfection performed with H-ANC and H-AFR *TGFB1* promoter haplotypes. The cells were serum-starved in medium containing 0,5% FBS and 1% Pen/Strep twenty-four hours before treatment with 5µg/ml recombinant human TGF-β1 (rhTGF-β1) (R&D Systems, MN, USA). Cells were subsequently treated with rhTGF-β1 for one or six hours, before harvesting with Passive Lysis Buffer (Promega, WI, USA). pRL-TK renilla reporter was used as an internal control and an empty vector control was also added in order to compare *TGFB1* promoter activity responses. *TGFB1* promoter activity was measured using a luminometer (Promega, Madison, WI, USA).

3.7.3. *Luciferase assays*

Luciferase assays were performed in duplicate using the Promega DUAL-Luciferase Reporter Assay System (Promega, WI, USA). The cells were harvested 30 hours post-transfection and washed with ice-cold Phosphate Buffer Saline (PBS) (see appendix) before harvesting the cell lysates in Passive Lysis Buffer (PLB) (Promega, WI, USA). The PLB was prepared in a 1:5 dilution with nuclease free water, and after addition of the lysis buffer, the cells were scraped down using a syringe as explained in the Promega DUAL-Luciferase Reporter Assay System manufacturer instructions. Cell lysates were stored at -80°C for 14-16

hours and freeze-thawed to effectively lyse the cells. The reporter activity of the lysates was assayed and measured for both the firefly luciferase as well as renilla luciferase activity.

Reporter activity is measured by the reaction of firefly luciferase enzyme catalysing the conversion of luciferin to oxyluciferin, which emits light that is measured by the luminometer. Renilla luciferase is then measured, which converts the co-elenterate-luciferin (coelenterazine) to coelenteramide in the presence of oxygen. The dual-luciferase reporter assay utilizes these reporter enzymes to simultaneously measure both the reporter gene activity of the transfected *TGFB1* promoter construct as well as the renilla internal control.

The lysates were centrifuged at 10 000xg for 2min to pellet the cell debris prior to use in the assay. The samples were then added to a 96-well luminometer plate before adding LARII (Promega, Madison, WI, USA). The light emitted by the firefly luciferase reaction was then measured and quantified by the luminometer (Thermo LabSystems, Luminoskan Ascent). 1x Stop & Glo buffer was then added to each sample to simultaneously quench the firefly luciferase reaction and initiate the renilla luciferase reaction. These values were also quantified and recorded by the luminometer using the Ascent Software 2.6 (Thermo Fisher Scientific, Waltham, MA, USA) as well as Microsoft Excel (Windows Vista). Relative luciferase activity values were normalised to the renilla activity to account for variation in transfection efficiency before making comparisons between the activity of the different *TGFB1* promoter haplotypes. This normalisation was not performed for the HIV-tat co-transfection experiments as the internal control pRL-TK was regulated by the pcDNA3.1-HIV tat construct.

3.7.4. *β-galactosidase vector as internal control for transfection efficiency*

The pCMV β-galactosidase vector (Clontech Laboratories, Mountain View, CA, USA) was used as an internal control for transfection efficiency in the HIV tat co-transfection experiments. The pCMV β-galactosidase vector contains the cytomegalovirus promoter which drives the expression of a *lacZ* reporter plasmid. Cells were transfected as previously stated in section 3.7.1 and cultured for 48 hours. The cells were subsequently washed with 1xPBS and fixed with 4% paraformaldehyde for 15min, followed by another wash before addition of the β-gal stain (see appendix) and incubated for 2-4 hours. The blue-white staining was monitored and recorded by light microscope. Images were taken 2-4 hours after

the stain was applied. The cells transfected with the β-galactosidase vector (presenting as blue from the reaction with β-gal stain) was quantified by counting 5-10 field of views before and after transfection and calculating the average amount of cells per view. Transfection efficiency was expressed as the percentage of cells that were positive/blue out of the total number of cells per field of view for both the RCC4+VHL and HT1080 cell lines.

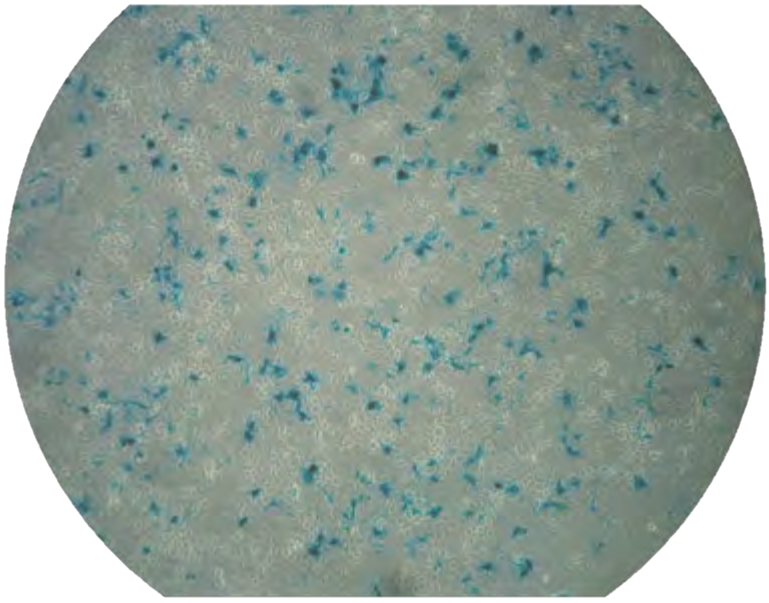


Figure 3.2: Light microscopy image of RCC4+VHL cells after two hours of incubation with the β-galactosidase stain. Figure is representative of what was observed in 2 independent experiments for both RCC4+VHL and HT1080 cells where ~30% of cells were transfected with the *TGFB1* promoter constructs and β-galactosidase expression vector.

3.8. *Western blot analysis*

3.8.1. *Protein harvesting*

Western blot analysis was performed to investigate the downstream effects of the African-specific *TGFB1* promoter variants on TGF-β1 and phosphorylated SMAD3 protein levels. To this end dermal fibroblasts from patients with known *TGFB1* promoter haplotypes were plated at $2.5x10^4$ cells per well of a 6-well plate and allowed to grow for 16-24 hours before treatment with 5μg/ml human recombinant TGF-β1 protein for 60 min. An untreated well of dermal fibroblasts was also included in the experiments to represent basal TGF-β1 and phosphor-SMAD3 protein expression for each of the *TGFB1* promoter haplotypes. After treatment, the cells were harvested using trypsin/EDTA and then centrifuged at 4 000 rpm for 5 min at 4°C before resuspension in ice-cold PBS. The samples were then centrifuged again at 4 000 rpm for 4 min at 4°C before the pellet was resuspended in Radioimmunopercipitation

(RIPA) buffer (see appendix) to lyse the cells and extract the protein. The samples were spun down at 12 000 rpm for 20 min at 4°C before resuspension in RIPA buffer. The supernatant was subsequently transferred to respective new eppendorf tubes. Samples were stored in aliquots at -80°C until needed.

3.8.2. *Bicinchoninic acid assay*

Bicinchoninic acid assay (BCA assay) is a biochemical assay used to determine the total concentration of protein in a solution/sample. A BCA assay was performed for harvested protein from the dermal fibroblasts with the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) as per manufacturer's instructions. The samples were loaded on a microplate reader and readings were recorded with the RT-2100c software (Rayto Life and Analytical Sciences, Shenzhen, China). Standard curve analysis was subsequently performed in Microsoft Excel (Windows, 2010). Standard curve analysis is performed by substituting the values from the microplate reader into a predetermined linear regression equation (derived from protein samples with known concentrations) which then allows for the calculation of the concentration of the specific protein in the harvested sample. This concentration is then used to calculate the amount of protein present in the sample, which then allows for equal protein loading across the wells during SDS-PAGE (see below).

3.8.3. *Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE)*

Thirty-five micrograms of protein from each untreated and rhTGF-β1 treated samples were solubilised in RIPA buffer and heated at 75°C for 10 min before electrophoresed on 1.5mm (thick) 10% resolving gels with a 5% stacking gel (see appendix). Biorad Mini PROTEAN© 3 casting apparatus was used to cast the gels and Biorad running tanks were filled with 1x running buffer (see appendix). TGF-β1 is a 44kDa protein, phosphor-SMAD3 is a 52kDa protein and the loading control p38 is a 40kDa protein. p38 was used as a loading control because it is unaffected by changes in cell cycle which leaves similar levels of expression for all of the samples. Due to their close proximity due to similar size, two separate SDS-PAGE gels were run for the TGF-β1 or phosphor-SMAD3 protein and the loading control. Molecular weight markers were used to monitor protein separation (PageRuler Prestained Protein Ladder, 26616, Thermo Fisher Scientific, Waltham, MA, USA). The running tank was connected to a power pack (Biorad Powerpack 200) and proteins were electrophoresed at 100V for 2 hours to separate the proteins sufficiently.

3.8.4. *Nitrocellulose membrane transfer*

The electrophoresed proteins were transferred to a Hybond ECL nitrocellulose membrane (Amersham Biosciences, UK) that was cut to size by packing the gels into a 'sandwich' using a cassette and placed in ice cold transfer buffer (appendix) for 1.5 hours at 100V. The 'sandwich' consisted of sponges, whattmann filter paper, the SDS-PAGE gel and the nitrocellulose membrane. After the membrane transfer, a Ponceau S stain was used to ensure the protein had transferred correctly and that equal loading could be observed.

3.8.5. *Detection and western blot analysis*

After the membrane transfer, the p38 and TGF-β1 blots were rinsed with PBS/Tween (appendix). The phosphor-SMAD3 blot was rinsed with TBS/Tween (appendix) and then blocked for one hour at RT with blocking buffer (appendix). The p38 blot was blocked with 5% milk in PBS/Tween, whereas the TGF-β1 blot was blocked with 5% Bovine Serum Albumin (BSA) in PBS/Tween and the phosphor-SMAD3 blot was blocked with 5% milk in TBS/Tween. After blocking, the membranes were incubated overnight with their appropriate primary antibody dilutions at 4°C with gentle shaking. The p38 blot was incubated with a 1:2000 dilution of rabbit polyclonal anti-p38 primary antibody (M0800, Sigma-Aldrich, St Louis, MO, USA), diluted in 5% milk PBS/Tween, whereas the TGF-β1 blot was incubated with a 1:500 dilution of mouse monoclonal anti-TGF-β1 primary antibody (MA517186, Separations, Johannesburg, South Africa) and the phosphor-SMAD3 blot was incubated with a 1:1000 dilution of rabbit polyclonal anti-Psmad3 primary antibody (C25A9, Cell Signalling, Danvers, MA, USA).

After overnight incubation with the primary antibody, the membranes were rinsed as stated previously: the p38 and phosphor-SMAD3 membranes with PBS/Tween and the TGF-β1 membrane with TBS/Tween. The membranes were subsequently incubated with their respective secondary antibodies at RT for 1 hour. The secondary antibodies were as follow: horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (Biorad, Hercules, CA, USA) at 1:5000 for the p38 and phosphor-SMAD3 membranes and horseradish peroxidase-conjugated goat anti-mouse secondary antibody (Biorad, Hercules, CA, USA) at 1:4000 for the TGF-β1 membrane.

After incubation with the secondary antibody, the membranes were rinsed and the signal was visualised by enhanced chemiluminescence (ECL) using WesternBright™ ECL (K-12045-D50, Advansta, Menlo Park, CA, USA) in a 1:1 ratio. The ECL substrate was incubated on the blots at RT for 5 min before placing the membranes between two transparencies, with an x-ray film placed on top. The cassette was then closed and the emission of light caused by the oxidation of the detection reagent and the conjugated horseradish peroxidase was captured on the x-ray film by developing and fixing the blot. Densitometric analysis was performed to quantify the bands and to normalise the treated to the untreated sample for each of the *TGFB1* promoter haplotypes by UnScan-it gel 6.1 software and Microsoft Excel (Windows, 2010). All of the densitometric readings were first normalised to the loading control, p38, and then to the untreated control cells. Unpaired student's t-test was used to compare basal and rhTGF-β1 stimulated protein responses between cells with different haplotypes.

3.9. *Immunohistochemistry of TGF-β1 on HIVAN patient renal biopsies*

3.9.1. *Preparation of biopsies*

Formalin Fixed Paraffin Embedded (FFPE) renal biopsy slides from patients with HIV-associated nephropathy (HIVAN) (n=19 biopsies) were cut and placed on microscope slides for staining with a monoclonal mouse anti-TGF-β1 antibody (MA5-17186, Thermo Fisher Scientific, Pierce Antibodies, Waltham, MA, USA).

Immunohistochemistry was performed under the supervision of Subash Govender at the Division of Anatomical Pathology, University of Cape Town. The slides were placed on a heating block in order to allow the wax to melt prior to the de-waxing process. De-waxing/deparaffinising consists of incubation of the the slides in three graded series of xylene for 5 minutes at a time, followed by 3 alcohol graded series for 3 min at a time prior to rinsing with running water for 5-10min. The de-waxing process was followed by antigen retrieval using Tris/EDTA (see appendix). Following the de-waxing process, antigen retrieval was performed in order to remove the cross-linking caused by the wax and to unmask the antigens. The antigen retrieval process was performed in a pressure cooker by boiling Tris/EDTA and inserting the slides to boil for 1 min and 30seconds. The slides were allowed to cool for 30 min before washing under running water for 5 min. Endogenous peroxidase was inhibited by incubation of the specimens in 3% Hydrogen peroxide (H_2O_2, Merck, Germany) for 10 min.

After washing in PBS for 10 min, the slides were covered with 5% goat-serum (DAKO, Glostrup, Denmark) for 15 min in order to block the remaining binding surface and prevent non-specific binding of the primary antibody. The excess goat-serum was wiped off without disturbing the tissue section before addition of the 1:500 dilution of primary antibody, mouse anti-TGF- β1. After one-hour incubation with primary antibody, the excess was wiped off without touching the tissue and sections were covered with the secondary antibody, DAKO anti-mouse (DAKO, Glostrup, Denmark), for 30 min. The nuclear stain, 3,3'-Diaminobenzidine/DAB (DAKO, Glostrup, Denmark), was subsequently added to the sections for 10 min before washing in running water for 5 min. The slides were then placed in Haematoxylin stain for 2 min and then briefly placed in ammoniated water prior to washing in running water until the DAB stain turned blue. The sections were then dehydrated using a graded series of alcohol and xylene before being mounted onto a cover slip with entallen (Merck, Darmstadt, Germany). The slides were then viewed under a light microscope at 200x and 400x magnification, with images taken by light microscopy.

3.9.2. *Scoring of the stained tissues*

A semi-quantitative "histo"-score (H-score) was assigned to each of the biopsies: the intensity of the stain is multiplied by the percentage of cells that stain positive for the tubules and the interstitium separately. The H-scores, ranging from 0-200, was then compared between HIVAN, HIV-negative and HIV-positive control groups. The interquartile ranges and median was calculated for the Masson's stain for fibrosis. The Masson's stain for fibrosis or Masson's trichrome stain is used in order to distinguish cells from the surrounding connective tissue. Blue staining indicates the connective tissue/collagen and allows for the grading of the severity of fibrosis, accumulation of extracellular matrix proteins such as collagen.

3.10. *Statistical analysis*

Raw luciferase assay data was analysed in Microsoft Excel. Luciferase values were calculated for each repeat and averaged to account for changes in transfection efficiency for each technical repeat by calculating the firefly to renilla ratios. All four of the *TGFB1* promoter activity was first normalized to pGL3-basic and then the fold activation/repression was normalized to the most common H-5 *TGFB1* promoter activity. For transient and co-transfection experiments, the fold activations and standard deviations were pooled across

biological repeats and were summarized in one graph. *TGFB1* promoter activity was compared with unpaired student's t-tests between haplotypes, while paired student's t-tests were used for comparisons between the different treatment time points for each of the *TGFB1* promoter constructs. Paired student's t-tests were also used to compare basal *TGFB1* promoter activity between all of the different *TGFB1* haplotypes. Densitometric readings were used to calculate the mean and standard error of the mean (SEM) for each of the blots before pooling the results from three or more biological repeats into one graph. In the case of an outlier (skewed results) the interquartile range and median was calculated as per non-parametric data analysis. Fisher's exact test or mid P-exact was used to calculate statistical significant differences in proportional differences between the *APOL1* minor allele frequencies of the Cape Town HIVAN and population control biopsies using OpenEpi (http://www.openepi.com/TwobyTwo/TwobyTwo.htm). The odds ratio was also calculated in order to compare the relative odds of HIVAN occurring together with the *APOL1* variants. The 95% confidence interval was also calculated where there is a 95% certainty that the population mean occurs within the range. P values of ≤ 0.05 were regarded as statistically significant.

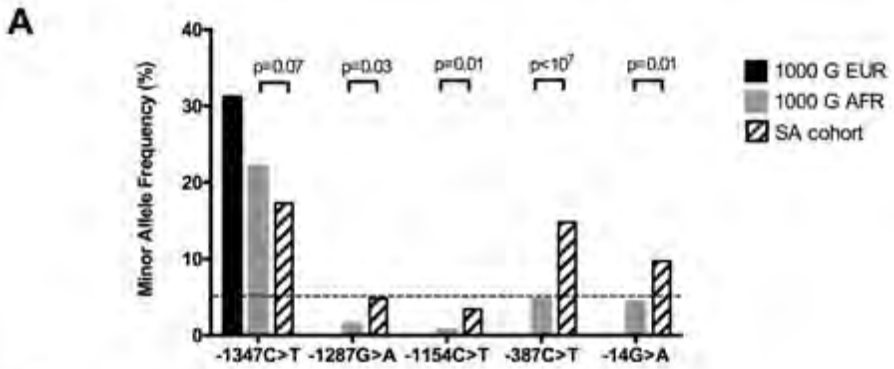
4. Results

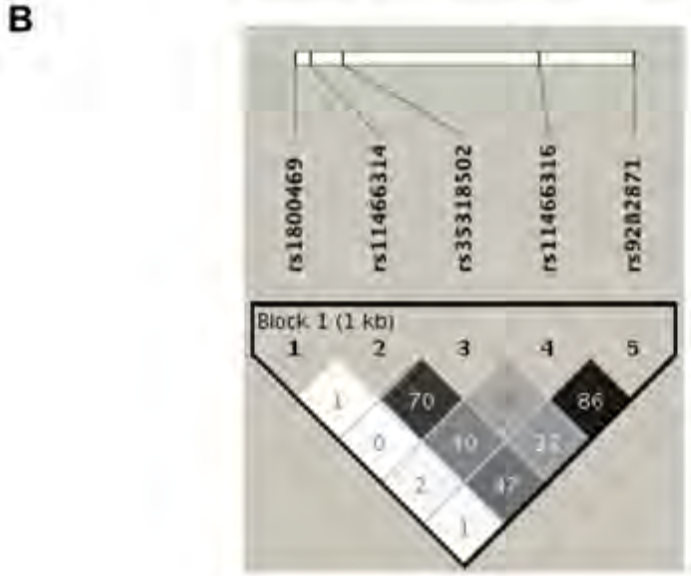
4.1. *African-specific TGFB1 promoter haplotype analysis*

Previously, *TGFB1* regulatory region loci minor allele frequencies (MAF) of population control cohorts from Cape Town and Soweto (n=344 chromosomes) were compared to the European (n=1006 chromosomes) and 'African' (n=1322 chromosomes) samples in the 1000 Genomes Project Phase 3 dataset (Nel *et al.*, 2015). Nel et al. (2015) found that the South Africans had significantly different alleles frequencies (MAF 0.17) for the African-specific variants compared to the 1000 Genomes African dataset comprising pooled African Americans, East- and West-Africans (MAF ~0.05). Importantly they do not occur in a European dataset (black bars in figure 4.1A) In order to ascertain which haplotypes are most common amongst Africans, and specifically South Africans, this data was used to perform haplotype analysis on the *TGFB1* promoter variants, including the ancestral rs1800469 (-*1347 C>T*) and the African-specific rs11466314 (-*1287 G>A*), rs35318502 (-*1154 C>T*), rs11466316 (-*387 C>T*) and rs9282871 (-*14 G>A*) loci (n=344 chromosomes).

Haploview software was used for this analysis which utilizes nearest-neighbour algorithm/confidence interval as described by Gabriel *et al*. 2002. Similar to previous results, figure 4.1B showed that the *-1287 G>A* and *-1154 C>T* polymorphisms are in tight linkage disequilibrium (LD) with an r^2 value of 0.7 (as represented by % as shown in figure 4.1B) (Nel et al. 2015). The *-387 C>T* and *-14 G>A* polymorphisms were shown to co-occur to a greater degree (r^2= 0.86, figure 4.1B). Interestingly, the African-specific variants all co-occurred with the **lower** expressing ancestral haplotype containing the *TGFB1 -1347 C* allele as shown by the regression analysis (r^2 in the white section of the linkage disequilibrium plot ($r^2 \leq 0.02$, figure 4.1B).

As previously discussed, the literature has focussed on ancestral haplotypes with the functional *TGFB1 -1347 C>T* (commonly referred to in the literature as *-509 C>T*). The common ancestral *TGFB1* promoter haplotype, termed H-5 in this thesis, represents the **lower** expressing ancestral haplotype with the *-1347 C* allele; this haplotype is the most common occurring *TGFB1* promoter haplotype in African and European populations (Haplotype frequency of ~80%). The *TGFB1* promoter H-6 represents the **higher** expressing ancestral haplotype with the *-1347 T* allele; this haplotype is the second most common occurring *TGFB1* promoter haplotype in Africans and Europeans with an allele frequency of ~20-30%, respectively. Interestingly the African-specific variants all co-occur with the **lower** expressing ancestral haplotype containing the *TGFB1 -1347 C* and are represented in Figure 4.1C by haplotypes H-1 through H-4.





C

Haplotypes	rs1800469 *-1347 C>T*	rs11466314 *-1287 G>A*	rs35318502 *-1154 C>T*	rs11466316 *-387 C>T*	rs9282871 *-14 G>A*	
H-1	C	G	C	T	G	
H-2	C	G	C	T	A	17%
H-3	C	A	C	T	A	
H-4	C	A	T	T	A	
H-5	C	G	C	C	G	83%
H-6	T	G	C	C	G	

Figure 4.1: A) Amended graph of minor allele frequencies published in Nel *et al.* (2015) for the *TGFB1 -1347 C>T* and the African-specific single nucleotide polymorphisms in the 1000 Genomes Project AFR (n=1322 chromosomes), 1000 Genomes Project EUR (n=1006 chromosomes) and South African (n=344 chromosomes) dataset. The dotted line at 5% indicates the cut-off for rare variants. Significant differences between the population groups are indicated by p-values. B) Linkage disequilibrium (LD) plot for the 5 polymorphic sites in the *TGFB1* regulatory region. Numbers within the diamonds indicate r^2 values as percentages. $r^2=0$ in white, $0<r^2<1$= shades of grey and $r^2=1$ in black. C) Table showing the six biologically relevant haplotypes in our group via haplotype analysis using Haploview

software. Minor variants shown in bold. *Estimated frequency of African *TGFB1* promoter haplotypes in a South African cohort as very rare haplotypes have been excluded from the analysis.

4.2. *Functionality of the African-specific TGFB1 promoter haplotypes*

Previously other groups investigated basal *TGFB1* promoter activity with a promoter that did not include the extended regulatory region which may impact on the "full length promoter" activity (Figure 1.3B and C). Nel *et al.* (2015) created an extended length human *TGFB1* promoter cloned into a luciferase reporter construct and found that the African-specific *-387 T* was functional and resulted in a 5-fold loss in basal promoter activity in cell lines of fibroblasts and muscle origin (Nel et al. 2015). As a follow-on, one of the main objectives in this work was to investigate whether the TGF-β1 pathway was contributing to the increased incidence of HIV-associated kidney disease in South Africans. This part of the study aimed to investigate the four African-specific *TGFB1* promoter polymorphisms and their impact on basal *TGFB1* promoter activity in renal cells.

To this end, luciferase promoter reporter constructs of H-1 to H-6, produced as described in section 3.5. were transiently transfected into RCC4+VHL human renal cell carcinoma and Caki-2 human clear cell renal carcinoma cells and luciferase activity was measured. Relative luciferase values were normalised to pRL TK renilla activity, and then normalised to the empty pGL3-vector before any comparisons. These values were represented as fold repression or activation of the promoter activity and the pooled results of ≥3 independent experiments each performed in duplicate are shown in the bar graphs in figure 4.2.

Results show in two cell lines of renal origin that there is no difference between the basal *TGFB1* promoter activity of the common ancestral *-1347 C* (H-5) or less frequent *-1347 T* (H-6) haplotypes (pooled results of 3 independent experiments shown in figure 4.2, RCC4+VHL p=0.88 and Caki-2 p=0.82). Previously Nel *et al.* (2015) and others showed that the common *-1347 C* allele is the **lower** expressing allele in human fibrosarcoma and lung carcinoma cells, respectively, with a 2-3-fold difference in basal *TGFB1* promoter activity compared to the *-1347 T* allele (Shah, Hurley, et al. 2006; Nel et al. 2015). Due to the fact that no 'higher' and 'lower' expressing haplotype containing the *-1347* C or *T* alleles were observed, subsequent experiments focused on the *-1347 C* containing haplotype, i.e. with or

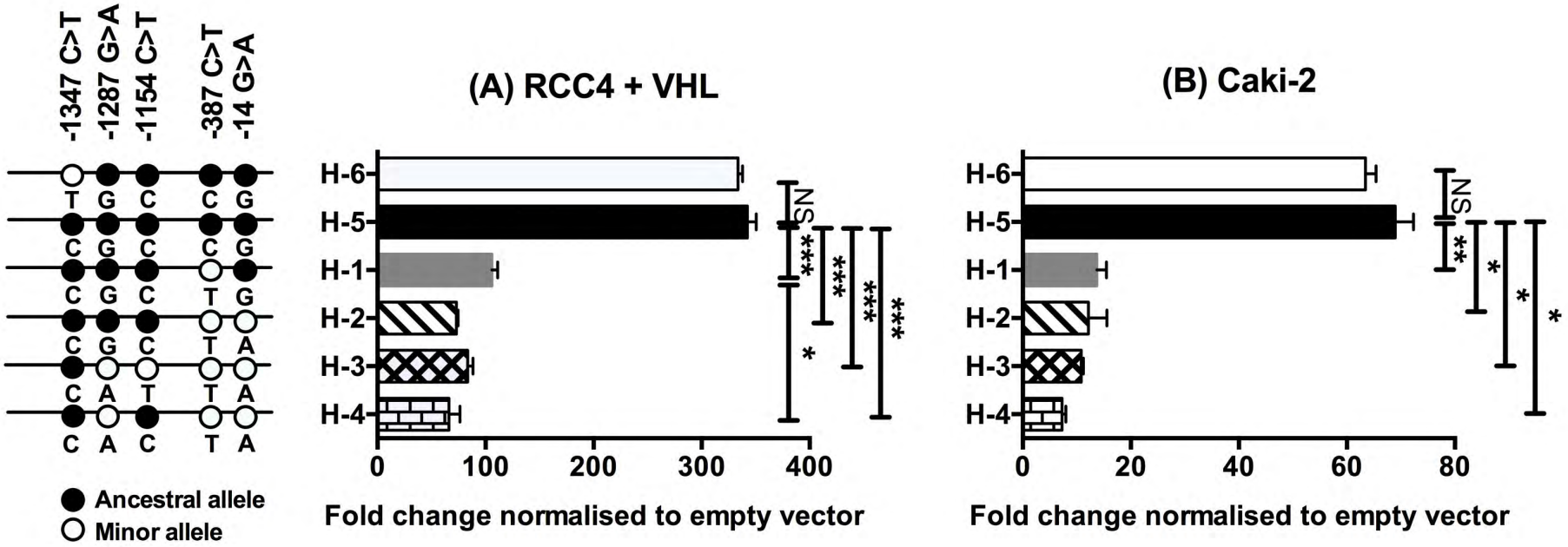


Figure 4.2: African *TGFB1* promoter haplotypes H-1 to H-4 reduces basal *TGFB1* promoter activity in two renal cell lines. RCC4+VHL cells (A) and Caki-2 cells (B) were transiently transfected with pGL3 luciferase reporter gene constructs containing *TGFB1* promoter haplotypes H-1 to H-6, which were generated using site-directed mutagenesis. The pRL renilla reporter gene construct was used as an internal control for transfection efficiency. Normalization was performed for the ratio of firefly to renilla luciferase activity against the empty pGL3-basic vector for each of the promoters in the two cell lines. Relative luciferase units are shown on the bar graphs for the fold activation for each *TGFB1* promoter haplotype relative to the empty vector. Statistical significant differences between the common ancestral H-5 and African-specific *TGFB1* promoter haplotypes (H-1 to H-4) are indicated by lines with stars where *$p<0.05$, **$p<0.005$ and ***$p<0.0002$. $p>0.05$ indicated by NS. These are pooled results of 3 independent experiments, each performed in duplicate. Student's unpaired t-test, two-tailed, error bars SEM.

without the African-specific variants. Because the African-specific variants do not occur with the *-1347 T* allele (see LD plot figure 4.1B) and only with the *-1347 C* allele, and this is the common ancestral haplotype, this haplotype may be referred to as the ancestral (H-ANC) *TGFB1* promoter haplotype in subsequent experiments.

The results revealed that having one or more African-specific variant on the backbone of the H-5/H-ANC haplotype resulted in 70-80% (~5-fold) loss of basal promoter activity in renal cells (Figure 4.2). Despite *TGFB1 -1287 G>A* and *-1154 C>T* as well as the *-387 C>T* and *-14 G>A* SNPs being in tight LD with each other, the repressive effect is largely attributed to the *-387 T* variant which is present in H-1 to H-4. Addition of the other African-specific polymorphisms did not result in any additional repression of basal promoter activity in the Caki-2 cells as they all have the *-1347 C* and *-387 T* (Figure 4.2B, H-2 through H-4). Although there was a difference between the H-1 and H-4 in the RCC4+VHL cell line where the *-1287 A* (H-4) resulted in a 1.6-fold decrease in activity compared to H-1 (p=0.03), the overall impression is that the rare H-4 promoter activity is prominently influenced by the *-387 T* variant.

In conclusion, this study has shown that the African-specific haplotypes (promoter haplotype H-1 to H-4, subsequently referred to as H-AFR in this thesis) results in a 5-fold repression of *TGFB1* promoter activity in two renal cell lines when compared to the common ancestral haplotype (H-ANC/H-5). Results also show that *TGFB1* promoter activity is similar between the African haplotypes and that the repressive effect can be attributed to the *-387 T* variant present in all four African haplotypes. In addition, in this extended *TGFB1* promoter, no difference was found in the promoter activity in haplotypes containing either *-1347 C* or *T* alleles in cell lines of kidney origin.

4.3. *The effect of exogenous TGF-β1 on the TGFB1 regulatory region*

As previously discussed, renal cells may respond to autocrine and paracrine TGF-β1 signals during chronic kidney injury such as with HIV-infection. TGF-β1 results in phagocyte recruitment and the transformation of tubular epithelial cells into myofibroblasts, both of which in turn result in increased TGF-β1 production. In the previous section results showed that H-AFR (H-1 to H-4) have similar basal *TGFB1* responses, but lower than H-ANC. In this part of the study the difference in response to TGF-β1 auto-induction for both the H-

ANC (ancestral) and H-AFR ('African' -*387 T* containing variant) *TGFB1* promoter haplotypes were investigated.

After treatment with 5μg/ml recombinant human TGF-β1 (rhTGF-β1) for 1 hour and 6 hours we compared the relative change of treated compared to basal promoter activity in a kidney (RCC4+VHL) and fibroblast (HT1080) cell line. The kidney cell line was chosen to represent the renal tubules while the fibroblast cell line was chosen to represent the renal interstitium. RCC4+VHL and HT1080 cell lines were transiently transfected for 30 hours with *TGFB1* promoter haplotypes H-ANC and H-AFR and serum-starved for 16 hours prior to harvesting. Luciferase activity was measured and the fold changes calculated relative to each haplotype untreated before comparing the fold changes between haplotypes and between cell lines.

In both the RCC4+VHL and HT1080 cells the basal *TGFB1* promoter activity for H-ANC and H-AFR was significantly different ($p<10^{-4}$, refer to figure 4.3A and B). In the kidney cells, two independent experiments were performed at the 1-hour time points and four independent experiments for 6-hour treatment periods for both the H-ANC and H-AFR haplotypes (pooled results in figure 4.3). However, due to poor transfection efficiency, only one H-ANC experimental results at 1-hour were included in the analysis. Results showed that *TGFB1* H-ANC did not respond significantly to treatment with rhTGF-β1 at 1 and 6 hours compared to untreated ($p>0.05$, figure 4.3A). The H-AFR showed a small but consistent response to 6-hour treatment with rhTGF-β1 (1.2-fold, SEM ±0.02, $p=0.008$, figure 4.3A).

Although a similar trend was seen for H-AFR at 1-hour treatment with rhTGF-β1, only two experiments were performed and therefore it was not significantly different compared to untreated (1.4 fold, SEM ±0.06, $p>0.05$).

In the fibroblast cell line (HT1080) four independent experiments showed a trend toward the H-ANC promoter being more responsive to stimulation than the H-AFR haplotype (Figure 4.3B). After 1-hour of rhTGF-β1, H-ANC showed a trend towards increased promoter activity (1.2-fold, SEM ±0.004, $p=0.07$), but no significant change after 6 hours of treatment (SEM ±0.01, $p=0.4$) when compared to basal H-ANC activity (Figure 4.3B). There was very little or no response for the *TGFB1* H-AFR haplotype after 1 and 6 hour treatments with rhTGF-β1 when compared to basal H-AFR promoter activity ($p=0.8$, figure 4.3B).

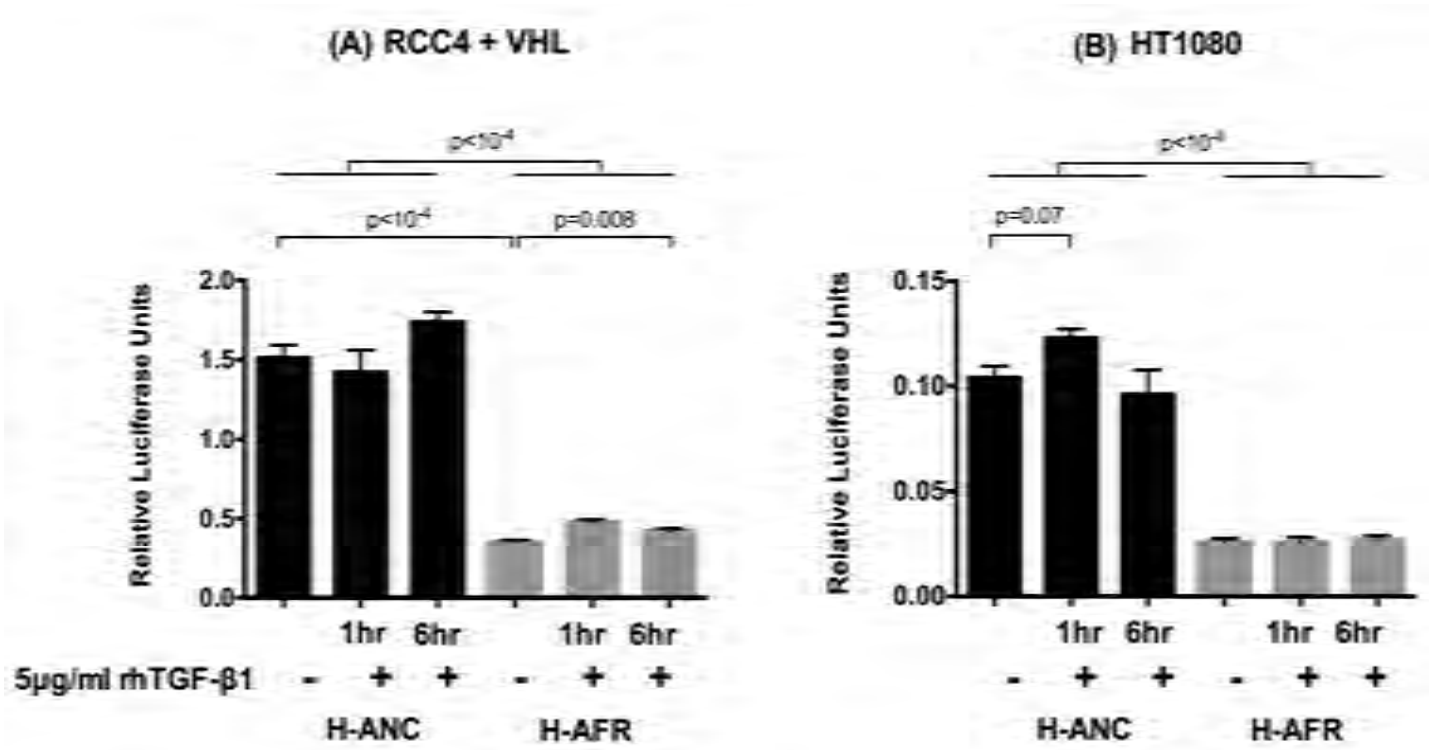


Figure 4.3: Recombinant human TGF-β1 does not consistently activate the *TGFB1* promoters in kidney (A) and fibroblast (B) cells. RCC4+VHL (renal cell carcinoma) and HT1080 (fibrosarcoma) cells were transiently transfected for 30 hours with H-ANC (*TGFB1 -1347 C*) and H-AFR ('African' *-387 T* containing variant) *TGFB1* promoter haplotypes and serum starved 16 hours prior to harvesting. The cells were treated with 5µg/ml recombinant human TGF-β1 (rhTGF-β1) for 1 hour and 6 hours before harvesting. Controls were not treated. Luciferase activity was measured 30 hours following transient transfection and expressed as the ratio of firefly to renilla luciferase activity for each of the haplotypes. pRL renilla reporter was used to control for transfection efficiency. These are pooled results of ≥2 independent experiments (except 1-hour treatments of H-ANC in the RCC4+VHL cell line), each performed in duplicate. Statistical differences (including trends) are shown in the figure. Student's paired t-test, two-tailed, was performed for treated versus untreated for each of the haplotypes, error bars standard error of the mean (SEM). Student's unpaired t-test, two-tailed, was performed to compare the H-ANC and H-AFR *TGFB1* haplotypes, two-tailed, error bars SEM.

Taken together rhTGF-β1 treatment of cells as an indicator of the TGF-β1 autocrine pathway did not consistently activate the *TGFB1* promoters although their responses were marginally different between cell types. The *TGFB1* promoter haplotype containing the *-387 T* variant (H-AFR) is possibly more responsive to TGF-β1 stimulation in the kidney cell line but showed little or no response in the fibroblast cell line.

4.4. *Investigating basal TGF-β1 protein expression in human dermal fibroblasts*

Western blot analysis was performed on the fibroblasts of patients with known *TGFB1* promoter haplotypes in order to determine whether the reduced *TGFB1* promoter activity observed in the African-specific *TGFB1* promoter haplotypes were also reflected downstream at the protein level. Dermal fibroblasts were also stimulated with 5μg/ml rhTGF-β1 in order to compare the results to those in section 4.3 above.

The human dermal fibroblasts were plated and treated with rhTGF-β1 for 1 hour while controls were left untreated before subjecting the harvested protein to SDS-PAGE gel analysis and western blotting. Four independent experiments were performed to assess basal TGF-β1 protein expression using the fibroblasts of three individuals with H-ANC and one individual with H-AFR. Densitometric analysis revealed that cells with *TGFB1* H-AFR showed on average 47% lower basal TGF-β1 expression levels (consistent repression in 3 of 4 biological repeats, p=0.04).

However, there was an outlier; including all four results produced a median of 0.58 or 58% repression but broad interquartile range (IQR) 0.3;1.3 when compared to the common ancestral (H-ANC) haplotype (representative bar graph plotted with mean and standard deviation shown in figure 4.4A).

In figure 4.4B results show that there was no difference in response to auto-induction, via rhTGF-β1, between H-ANC and H-AFR *TGFB1* promoters based upon five independent experiments (p=0.27). The H-ANC cells responded to stimulation, albeit inconsistently, on average activating 3–fold after 1 hour of treatment (median 1.2; IQR 1;3.5). In contrast, the H-AFR cells containing the *-387 T* variant consistently showed no response to stimulation (mean/median=1; IQR 0.9;1.1). In figure 4.4C two representative western blots are shown for the five experiments with densitometric readings reflecting the variation at basal and stimulated levels for H-ANC and H-AFR cells.

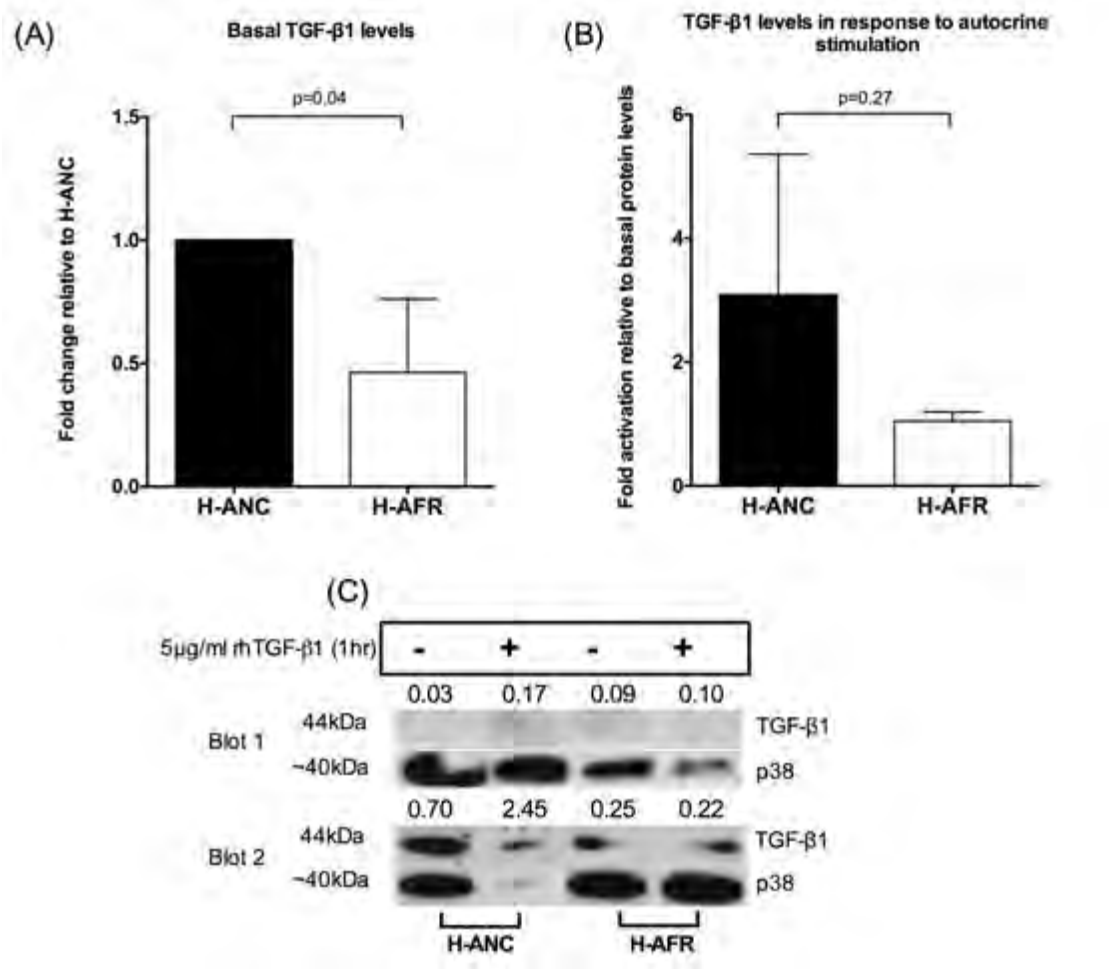


Figure 4.4: Basal and stimulated TGF-β1 protein expression levels in human dermal fibroblast cells with known genotypes. (A) Cells with the *-387 T* containing *TGFB1* promoter (H-AFR) reduces basal TGF-β1 protein expression by 47% in human dermal fibroblasts. Protein from human dermal fibroblasts of patients with known genotypes were harvested in RIPA buffer before being subjected to SDS-PAGE analysis. These are pooled results from four independent experiments with H-ANC (cells with *-1347 C*)(n=3 different patient dermal fibroblast cells) and H-AFR (n=1 patient dermal fibroblast cells) cells. Densitometric analysis was performed and the ratio of TGF-β1 to the loading control, p38, is expressed as fold change of H-ANC relative to H-AFR, one experiment was excluded as an outlier. Statistical differences are shown for basal TGF-β1 expression for H-AFR relative to H-ANC. Bar graph with mean and standard deviation (SD). (B) Cells with the H-AFR *TGFB1* promoter consistently does not respond to 1-hour stimulation with rhTGF-β1. Human dermal fibroblasts were serum-starved 24 hours prior to treatment with 5µg/ml rhTGF-β1 for 1 hour. Controls were not treated. Protein was subjected to SDS-PAGE and western blot analysis. This graph is representative of five independent experiments with H-ANC (n=2 different patient dermal fibroblast cells) and H-AFR (n=1 patient dermal fibroblast cells) cells. Statistical analysis was performed using a student's two-tailed t-test for H-ANC cells relative to H-AFR. Densitometric analysis was performed and the fold activation is expressed as treated over untreated for each of the data points. Bar graph with mean and SD. C) Two western blots representative of five independent experiments. Densitometric readings expressed as the ratio of TGF-β1 protein to the p38 loading control.

In conclusion, the repressive effect observed with the *TGFB1 -387 T* promoter, compared to cells with the common ancestral promoter haplotype, was also observed at basal protein expression levels. Results are showed that cells with H-ANC do not consistently respond to autocrine stimulation but those with the H-AFR haplotype consistently show very little or no response to rhTGF-β1 after 1 hour.

4.5. *Phosphorylated Smad3 levels in human dermal fibroblasts*

The phosphorylation of Smad3 is one of the first steps in the TGF-β1 canonical pathway. Smad3 is a well-known transcriptional modulator activated by TGF-β1 signalling and is also known for its role in renal TGFβ-induced apoptosis. In order to investigate the effects of the *TGFB1* promoter haplotypes western blot analysis was performed on human dermal fibroblasts with known haplotypes as described in section 4.4.

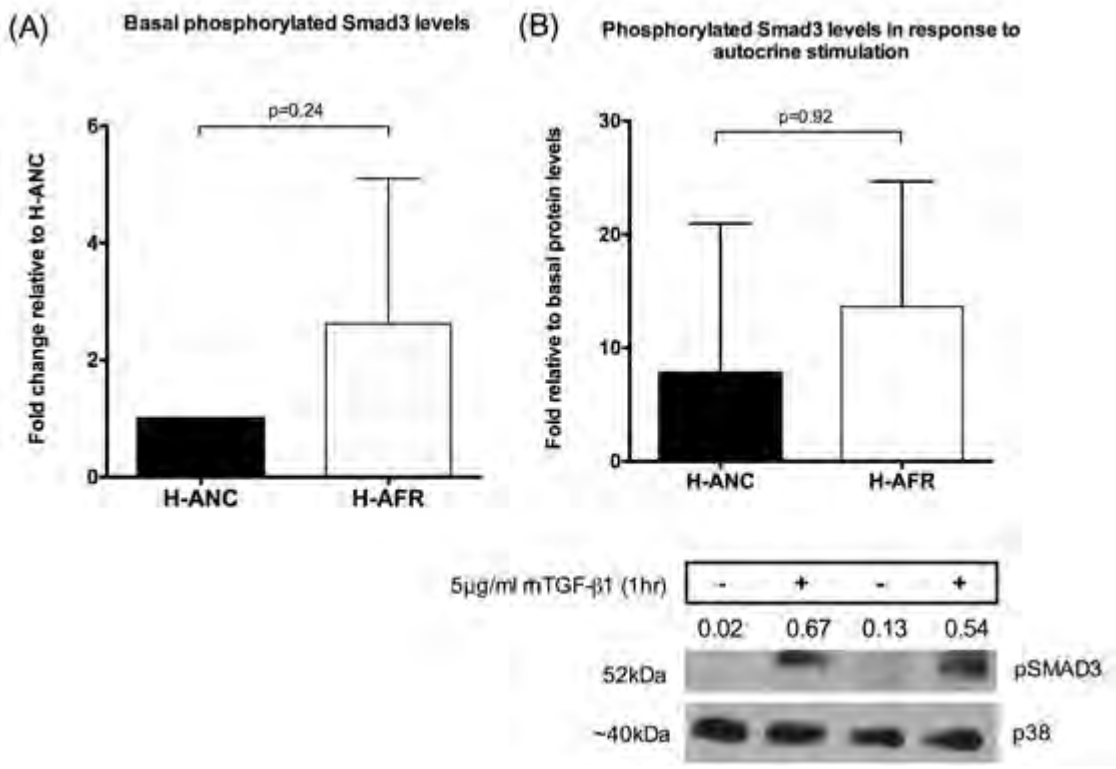


Figure 4.5: Basal and stimulated phosphorylated Smad 3 (pSmad3) protein expression levels in human dermal fibroblast cells with known genotypes. (A) Similar basal pSmad3 levels for cells with H-ANC (*TGFB1 -1347 C*) and H-AFR (containing *TGFB1 -387 T*) *TGFB1* promoters in human dermal fibroblasts. Protein was harvested from human dermal fibroblasts of patients with known genotypes and subjected to SDS-PAGE and western blot analysis. The bar graph with mean and standard deviation (SD) are pooled results from two independent experiments with H-ANC (n=1 patient dermal fibroblast cells) and H-AFR (n=3 different patient dermal fibroblast cells) cells. Densitometric analysis was performed and the ratio of TGF-β1 protein to the loading

control, p38, is expressed as the fold change of H-ANC relative to H-AFR. Statistical trend is shown for basal pSmad3 expression. (B) Similar upregulation for cells with H-ANC and H-AFR *TGFB1* promoters are observed in response to 1-hour treatment with rhTGF-β1. Human dermal fibroblasts were serum-starved 24 hours prior to treatment with 5μg/ml rhTGF-β1 for 1 hour. Controls were not treated. Protein was harvested and subjected to SDS-PAGE and western blot analysis. The bar graph with mean and SD is representative of pooled results from seven independent experiments with H-ANC (n=5 different patient dermal fibroblast cells) and H-AFR (n=3 different patient dermal fibroblast cells) cells. Densitometric analysis was performed and the fold change is expressed as treated over basal/untreated for cells with the two *TGFB1* promoters. A representative western blot is shown for seven independent experiments with densitometric readings expressed as the ratio of pSmad3 protein to the p38 loading control.

As in section 4.4 the human dermal fibroblasts were treated with 5μg/ml rhTGF-β1 for 1 hour. Two independent experiments were performed for basal phosphorylated Smad3 (pSmad3) protein levels (H-ANC, n=1 patient dermal fibroblast cells; H-AFR, n=3 different patient dermal fibroblast cells) (Figure 4.5A). Western blot analysis revealed that there was no consistent difference in basal pSmad3 protein levels between cells with H-AFR when compared to H-ANC (fold change calculated on densitometric readings, p=0.24). In response to 1-hour treatment with rhTGF-β1 a similar upregulation in pSmad3 levels in both H-ANC and H-AFR cells was found, albeit inconsistently; the average activation of pSMAD3 levels was 12-fold for both H-ANC and H-AFR cells (SD ±19 and 11, respectively; p=0.92) (Figure 4.5B).

In conclusion this part of the study has shown that the basal pSmad3 levels as well as the pSmad3 response to stimulation with rhTGF-β1 were not different between patients who have the common ancestral (H-ANC) or African-specific *TGFB1* haplotype (H-AFR). Both African and ancestral *TGFB1* promoter haplotypes responded to the autocrine rhTGF-β1 stimulus as evidenced by pSmad3 activation in the TGF-β1 canonical pathway. These results imply that TGF-β1 autocrine signalling is normal until pSmad3 enters the nucleus where cells with the *TGFB1 -387 T* allele have 5-fold repression of *TGFB1* promoter activity compared to cells with the *-387* C.

4.6. HIV Tat protein activates the TGFB1 promoter haplotypes

Because of the similar and 12-fold activation in pSmad3 levels in response to rhTGF-β1 stimulation in both H-ANC and H-AFR promoters and because the TGF-β1 pathway has been implicated in the pathogenesis of HIVAN, the effect of intracellular HIV Tat on the

TGFB1 promoter haplotypes was investigated. To this end co-transfection experiments with a HIV Tat expression vector in RCC4+VHL human renal carcinoma cell line as well as a HT1080 human fibrosarcoma cell lines was performed.

The H-AFR promoter together with both the ancestral (-*1347 C*) H-5 and less common (-*1347* T) H-6 *TGFB1* promoter haplotypes were co-transfected with 350ng HIV Tat expression vector (pcDNA-HIV Tat) (provided by Dr Shaheen Mowla) and promoter activity measured before comparing the effect between haplotypes. Because Nel *et al*. (2015) reported that *TGFB1* H-6 (-*1347 T*, pHap-4 in Nel *et al*., 2015) showed increased promoter activity compared to ancestral *TGFB1* H-5 haplotype in fibrosarcoma cells, unlike this work in renal cells, both H-5 and H-6 were used in these experiments. An empty-vector control (pcDNA 3.1) was included in order to normalise and express the fold change relative to basal promoter activity for each of the *TGFB1* promoters. A pRL-TK renilla construct was also included for internal control of transfection efficiency. It is important to note that addition of the HIV Tat expression construct resulted in a ~3-fold increase in pRL-TK renilla luciferase values when compared to the cells with the empty vector control. This indicated independent regulation of the renilla construct by the HIV Tat protein and therefore the renilla values could not be used to control for transfection efficiency. Instead the firefly luciferase values were used and β-galactosidase assays were performed to control for transfection efficiency (refer to methods and materials section 3.7.4).

In the RCC4+VHL cell line all three *TGFB1* promoters similarly activated in response to HIV Tat compared to basal levels; the *TGFB1* H-6 promoter had a 1.5-fold increase (p=0.04), *TGFB1* H-5 showed 1.6-fold ($p<10^{-5}$) and H-AFR showed 1.8-fold (p=0.004) (Figure 4.6A). In at least 3 independent experiments in HT1080 cells, the H-6 haplotype showed no response to HIV Tat protein (1.2-fold, p=0.2) while the more common H-5 and the H-AFR haplotype both showed increased promoter activity in the presence of HIV Tat compared to basal promoter activity (H-5, 2-fold increase; p=0.009, H-AFR 1.4-fold increase; p=0.0004). There was a trend suggesting that the H-5 haplotype is more responsive than the H-6 (p=0.06) and the H-AFR haplotype (p=0.02) (Figure 4.6B).

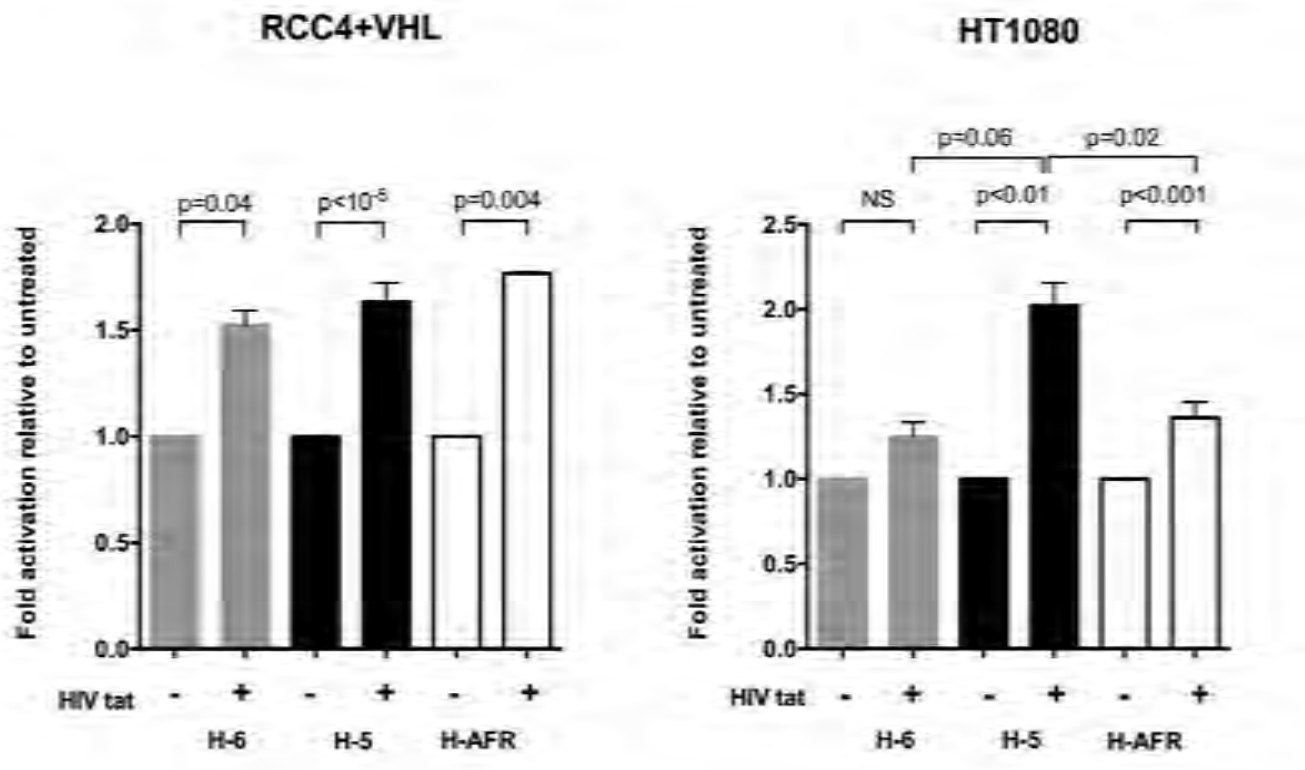


Figure 4.6: Intracellular HIV Tat protein activates H-6 (*-1347 T*), H-5 (*-1347* C) and H-AFR (*-387 T*) *TGFB1* promoter activity in kidney (RCC4+VHL) and fibroblast (HT1080) cells. RCC4+VHL (renal cell carcinoma) and HT1080 (fibrosarcoma) cells were co-transfected with pGL3 luciferase reporter gene constructs containing the less common ancestral H-6, common ancestral H-5 and H-AFR *TGFB1* promoter haplotypes together with 350ng of a pcDNA3.1-HIV-Tat expression vector. Luciferase activity was measured 30 hours following transient transfection and expressed as fold activation relative to basal/untreated *TGFB1* promoter activity. Renilla luciferase was not used for calculating transfection efficiency as it was independently regulated by the HIV tat expression vector. These are pooled results of ≥3 independent experiments, each performed in duplicate. Student's paired t-test for treated relative to untreated/basal and unpaired student's t-test, two-tailed, for differences between haplotypes. P-values are shown above the bars with $p>0.05$ indicated by NS. Error bars are standard error of the mean.

In conclusion these results have shown a significant and similar increase in *TGFB1* promoter activity in response to HIV Tat in all three the *TGFB1* haplotypes in renal cells. In the fibroblast cells the responses were more variable depending on the promoter haplotype. There was a greater activation response to HIV Tat in the most common H-5 promoter, whereas the H-AFR and H-6 showed either modest or no responses in fibroblasts, respectively. This is interesting because the *TGFB1 -1347 C* and *T* promoter variants showed no differences in basal promoter activity (Figure 4.2A and B).

4.7. *Analysis of TGF-β1 staining in renal biopsies from HIVAN patients*

It is known from the literature that TGF-β1 is expressed in the kidney during chronic injury such as HIV-infection. In addition, increased TGF-β1 expression is observed during fibrosis and both kidney tubules and fibroblasts in the interstitium have increased TGF-β1 expression levels in HIV-associated nephropathy (HIVAN) renal biopsies when compared with HIV-negative individuals (Bódi et al. 1997). In this part of the study 20 consecutive biopsies from a Cape Town (CPT) HIVAN sample were re-assessed in order to compare TGF-β1 staining to the nephrologist's (Dr Wearne) independent histology score (Table 4.1). Table 4.1 that 60% of the HIVAN cohort had microcysts and 90% had mild to severe fibrosis (Masson's trichrome stain) on their biopsies.

After optimisation of the TGF-β1 staining protocol the tubule and tubulo-interstitial (will be referred to as interstitium here after) staining was assessed by grading the intensity of the stain (1+ faint or 2+ strong) as well as the percentage of cells staining positive where 0 indicated absence of staining, 1 indicated <10% of the cells, 2 indicated 10-30% of cells, 3 indicated 30-50% of cells, 4 indicated >50% and 5 indicated 100% of cells staining positive.

A "histo"-score or H-score was calculated for each of the biopsies where the percentage of cells gets multiplied by the intensity of the stain in both the tubules and interstitium for each tissue sample in order to compare within or between groups (McCarty *et al.*, 1985; Fedchenko and Reifenrath, 2014). The grading of TGF-β1 staining was done blindly without knowing the severity scores for HIVAN as graded by Dr Wearne. In addition, three HIV-negative biopsies and three HIV-positive biopsies without HIVAN were included in the analysis.

Table 4.1: *APOL1* and *TGFB1* genotyping data for HIVAN, HIV-negative and HIV-positive individuals together with histopathological features and TGF-β1 immunohistochemical analysis of renal biopsies.

				TGF-β1 immunohistological staining						APOL 1			TGFB1
	Histopathological features			Tubules		Interstitium		H-score		rs73885319	rs60910145	rs71785313	rs11466316
	HIVAN severity	Microcysts	Fibrosis	Intensity	Score	Intensity	Score	Tubules	Interstitium	G1 - A>G	G1 - T>G	G2 - delTTATAA	-387 C>T
HIVAN biopsies													
HVN 1	Moderate	No	1	1+	1	0	0	10	0	A/G	T/G	TTATAA/-	C/C
HVN 2	Mild	No	1	2+	4	1+	2	150	20	A/A	T/T	TTATAA/TTATAA	C/C
HVN 3	Severe	Yes	2	2+	5	1+	2	200	20	A/A	T/T	-/-	C/C
HVN 4	Severe	Yes	2	2+	5	1+	2	200	20	A/G	T/G	TTATAA/-	C/C
HVN 5	Severe	Yes	2	2+	5	1+	2	200	20	A/G	T/G	TTATAA/-	C/C
HVN 6	Mild/moderate	Yes	1	2+	5	1+	2	200	20	A/A	T/T	TTATAA/TTATAA	C/C
HVN 7	Severe	Yes	3	2+	4	1+	2	150	20	G/G	G/G	TTATAA/TTATAA	C/T
HVN 8	Mild/moderate	Yes	1	2+	4	1+	2	150	20	A/A	T/T	TTATAA/-	C/C
HVN 9	Moderate	No	1	2+	4	1+	2	150	20	A/A	T/T	TTATAA/TTATAA	C/C
HVN 10	Mild	No	0	1+	5	1+	2	100	20	A/A	T/T	TTATAA/TTATAA	C/C
HVN 11	Moderate	No	2	2+	5	1+	2	200	20	A/A	T/T	TTATAA/TTATAA	C/C
HVN 12	Severe	Yes	3	1+	5	1+	2	100	20	A/G	T/G	TTATAA/-	C/C
HVN 13	Moderate	No	2	2+	5	1+	3	200	40	A/G	T/G	TTATAA/TTATAA	C/C
HVN 14	Severe	Yes	2	2+	5	1+	3	200	40	G/G	G/G	TTATAA/TTATAA	C/T
HVN 15	Moderate	Yes	2	2+	4	1+	3	150	40	A/G	T/G	TTATAA/-	UK
HVN 16	Severe	Yes	3	2+	4	1+	4	150	75	G/G	T/G	TTATAA/TTATAA	UK
HVN 17	Severe	Yes	4	2+	4	1+	4	150	75	G/G	G/G	TTATAA/TTATAA	C/C
HVN 18	Severe	Yes	4	2+	4	1+	3	150	75	A/A	T/T	TTATAA/TTATAA	C/C
HVN 19	Moderate	No	1			*		-	-	A/G	T/T	TTATAA/TTATAA	UK
HVN 20	Mild	No	0			**		-	-	A/G	T/G	TTATAA/TTATAA	C/T
Controls													
HIV positive (HIV+)													
HIV+ 1	Mesangial proliferative glomerulonephritis			1+	5	0	0	100	0	-	-	-	-
HIV+ 2	Membranous glomerulonephritis			1+	5	0	0	100	0	-	-	-	-
HIV+ 3	IgA nephropathy			1+	5	0	0	100	0	-	-	-	-
HIV negative (HIV-)													
HIV - 1	Membranous glomerulonephritis			2+	5	0	0	200	0	-	-	-	-
HIV - 2	Minimal change disease			2+	5	0	0	200	0	-	-	-	-
HIV - 3	Systemic sclerosis			2+	5	0	0	200	0	-	-	-	-

Gradation of TGF-β1 staining: 0 - no staining, 1- <10% of cells staining positive, 2- 10-30% of cells, 3- 30-50% of cells, 4- >50% of cells and 5- 100% of cells. H-score is assigned by using the formula: 1 x (% of cells with 1+ intensity) + 2 x (% of cells with 2+ intensity) for both the tubules (H-score of 0-200) and the tubulointerstitium (H-score of 0-100). *Insufficient tissue to perform stain, excluded from immunohistochemical analysis **No biopsy, only included in genotyping. UK - unknown

Nineteen biopsies were available for the CPT-HIVAN cohort, but only 18 were analysed because one individual (HVN19) had too little tissue remaining and the stain could not be performed. All of the stains were performed on the same day with a 1:500 dilution of the mouse monoclonal anti-TGF-β1 antibody (see section 3.10.1. in methods and materials). Light microscopy images representative of the histopathological features observed in our CPT-HIVAN sample are shown in figure 4.7 and 4.8. The majority of the HIVAN samples had microcysts (tubular degeneration) and interstitial disease (inflammation and fibrosis). These features were not observed in the HIV-negative or –positive control tissues although one HIV-positive controls had some interstitial inflammation. Control biopsy kidney dysfunction was caused by glomerulonephritis, minimal change disease or IgA nephropathy (see table 4.1 and figure 4.7 *e* and *f*).

No fibrosis or atrophy was observed in the control biopsy samples and the tubules and interstitium looked normal. All biopsies, including HIVAN and controls samples, had similar TGF-β1 staining of the tubular cells. The intensity of the TGF-β1 stain in the tubules was much stronger in the HIV-negative controls (tubular H-score=200) when compared to HIV-positive controls (tubular H-score=100). However, 15/18 (83%) of the HIVAN samples had a high H-score for tubules (H-score >100) despite four samples not having microcysts.

The most striking difference between HIVAN biopsies and the control biopsies (HIV+ or -) was seen in the TGF-β1 positive staining of the interstitium; 17/18 (94%) of the HIVAN samples showed TGF-β1 staining in the interstitium compared to absence in the controls. Comparing the interstitial TGF-β1 H-scores with the Masson's trichrome severity of fibrosis grading (maximum 4) in the HIVAN samples showed that 67% (n=12 biopsies) had fibrosis H-scores

Therefore, there was good correlation between the TGF-β1 staining pattern (% of cells and intensity) and the Masson's trichrome staining. Although there are only few control biopsies for comparison, the most noticeable finding was the selective interstitial pathology and TGF-β1 staining in HIVAN biopsies.

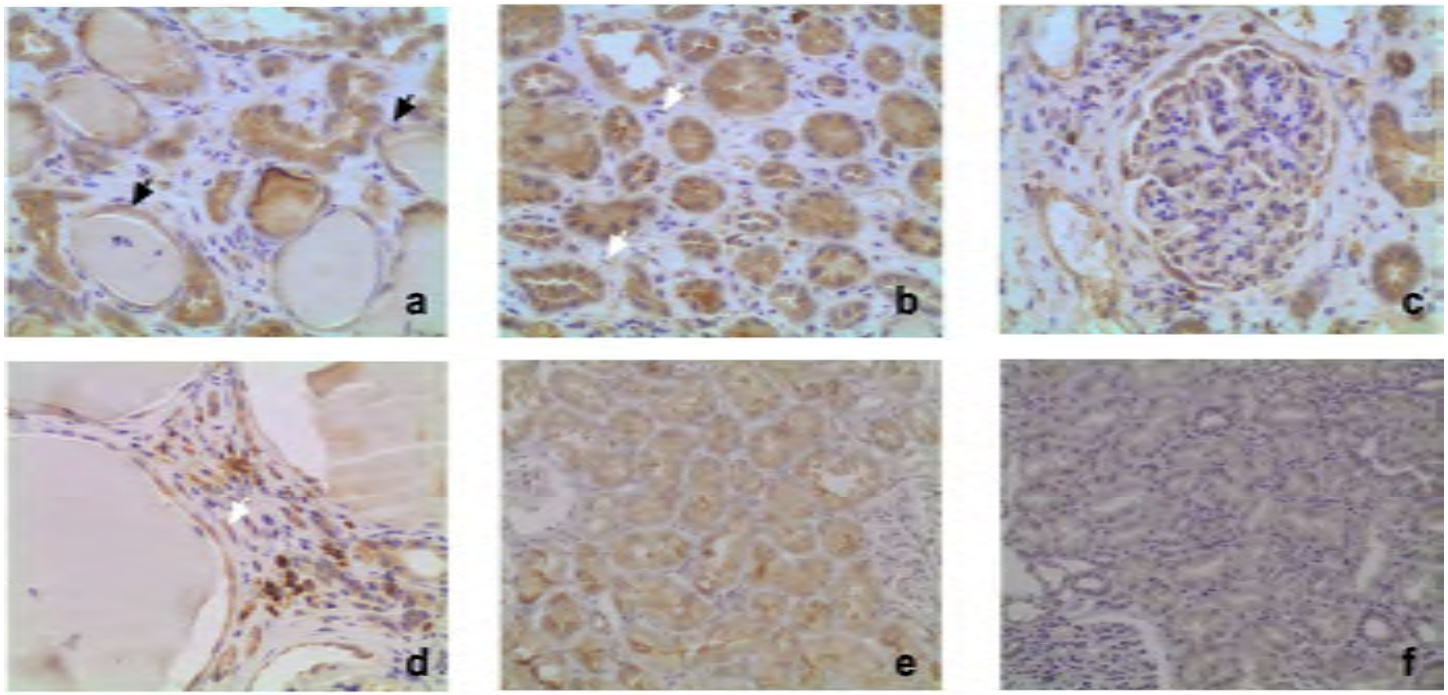


Figure 4.7: Light microscopy images of TGF-β1 immunohistochemical staining in the renal cortex of HIV-associated nephropathy (HIVAN), HIV-negative and HIV-positive renal biopsies. The blue is DAB staining of the nuclei while the brown is showing TGF-β1 staining. *a-d* represents some of the histopathological features observed in the renal biopsies of the Cape Town HIVAN samples. *a* is showing the formation of microcysts observed in several of the HIVAN biopsies with black arrows pointing toward the tubular epithelial cells staining positive for TGF-β1. *b* Positive TGF-β1 staining of tubules showing atrophy, white arrows indicating positive TGF-β1 staining in the tubulointerstitium. *c* Positive TGF-β1 staining of cells in and around the glomeruli while *d* is indicating cells staining positive for TGF-β1 in the tubulointerstitium at a higher magnification. *e* Representative image of TGF-β1 positive staining of tubules in the HIV-negative biopsies while *f* is representative of positive TGF-β1 staining in the HIV-positive biopsies. Magnification for *a,b,e* and *f* was at 200x while c and *d* was at 400x. Gradation of staining for pannels *a-d* was 2+ intensity with 100% of tubular and tubulointestitial cells staining positive, *e* was 2+ with 100% of tubule cells staining positive but no staining in the tubulo-interstitium. *f* was graded with a 1+ intensity and 100% of tubule cells staining positive but no staining in the tubulointerstitium.

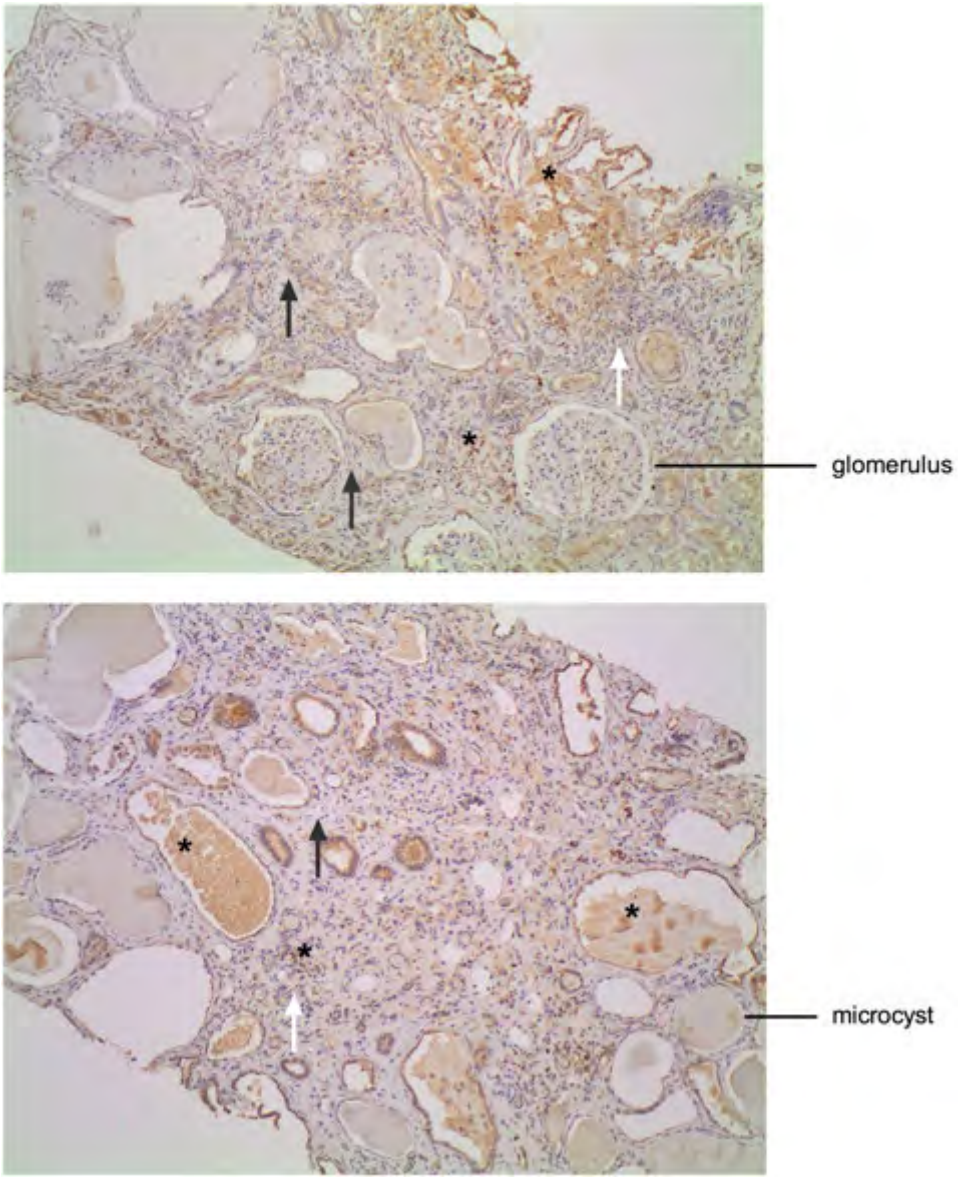


Figure 4.8: Representative light microscopy images for TGF-β1 immunohistochemical staining in renal biopsies from the Cape Town HIV-associated nephropathy (HIVAN) sample (n=19 biopsies) at lower magnification (100x magnification). *a* and *b* are showing an overall view of the renal cortex stained with DAB (blue) to indicate the nuclei of cells and positive TGF-β1 staining in brown. Both *a* and *b* show the presence of microcysts, tubular atrophy, tubulointerstitium inflammation (white arrows) and fibrosis (black arrows). * indicating dark brown staining of plasma cells in the interstitium and other non-specific staining of protein deposits

of ≤20 (interstitium) and mild to moderate fibrosis on the Masson's stain (median grade 1.5, IQR 1; 2). The remaining 33% (n=6 biopsies) all had fibrosis H-scores >20 (interstitium) and these showed moderate to severe fibrosis with the Masson's stain (median 2.5, IQR 2; 3.5) (Table 4.1).

4.8. *Genotyping HIVAN sample for APOL1 and TGFB1*

APOL1 genotyping reveals 60% of HIVAN patients have one risk allele

In order to correlate the histopathological severity of HIV-associated nephropathy (HIVAN) with their genotype the *APOL1* gene and *TGFB1* regulatory regions of 20 HIV-positive individuals with HIVAN and African-ancestry was sequenced. The risk alleles in the *APOL1* gene were included because of its previously reported association with increased risk of HIVAN in individuals with African-ancestry (Kopp et al. 2011; Freedman et al. 2014; Kasembeli et al. 2015).

Five *APOL1* haplotypes have been identified and are shown in figure 4.9. The CPT-HIVAN genotyping data was compared to pooled population South African controls from Johannesburg (HIV-negative controls, n=54 and HIV-positive control, n=54 biopsies) (Kasembeli *et al*., 2015). Importantly, in that report there was no significant difference in the minor allele frequencies (MAF) for *APOL1* risk alleles between the HIV-positive and HIV-negative controls ($p>0.05$). In addition, the CPT-HIVAN genotype data was compared with a HIVAN cohort previously reported from Johannesburg (JHB-HIVAN) (Kasembeli *et al.*,2015).

The three *APOL1* risk alleles, termed $G1^{S342G}$, $G1^{I384M}$ and $G2^{4^6}$, occur at a higher frequency in our CPT HIVAN sample (n=40 chromosomes; MAF 0.40, 0.35 and 0.20, respectively) when compared to population controls (n=108 chromosomes; MAF 0.07, 0.06 and 0.11, respectively Figure 4.9). Genotyping results revealed that 60% of the CPT-HIVAN sample had ≥1 *APOL1* risk allele compared to 18% in the population controls (Odds ratio [OR]=6.9, 95% confidence interval [CI]: 3.1;15.8, $p=10^{-6}$) (frequencies in figure 4.9). In addition, there were differences in the MAF of *APOL1* risk alleles between the CPT-HIVAN and JHB-HIVAN sample (p=0.002) (Figure 4.9 and 4.10).

APOL1 risk haplotypes were also more frequent in the CPT-HIVAN sample when compared to the population controls, evident by the G0 haplotype (no risk) frequency of 0.40 compared to 0.82, respectively (Figure 4.9). The haplotype frequencies of the CPT-HIVAN group (frequency of 0.6 for ≥1 risk allele) were similar but slightly lower than those observed in the JHB-HIVAN cohort (frequency of ~0.9 for ≥1 risk allele) (p=0.002).

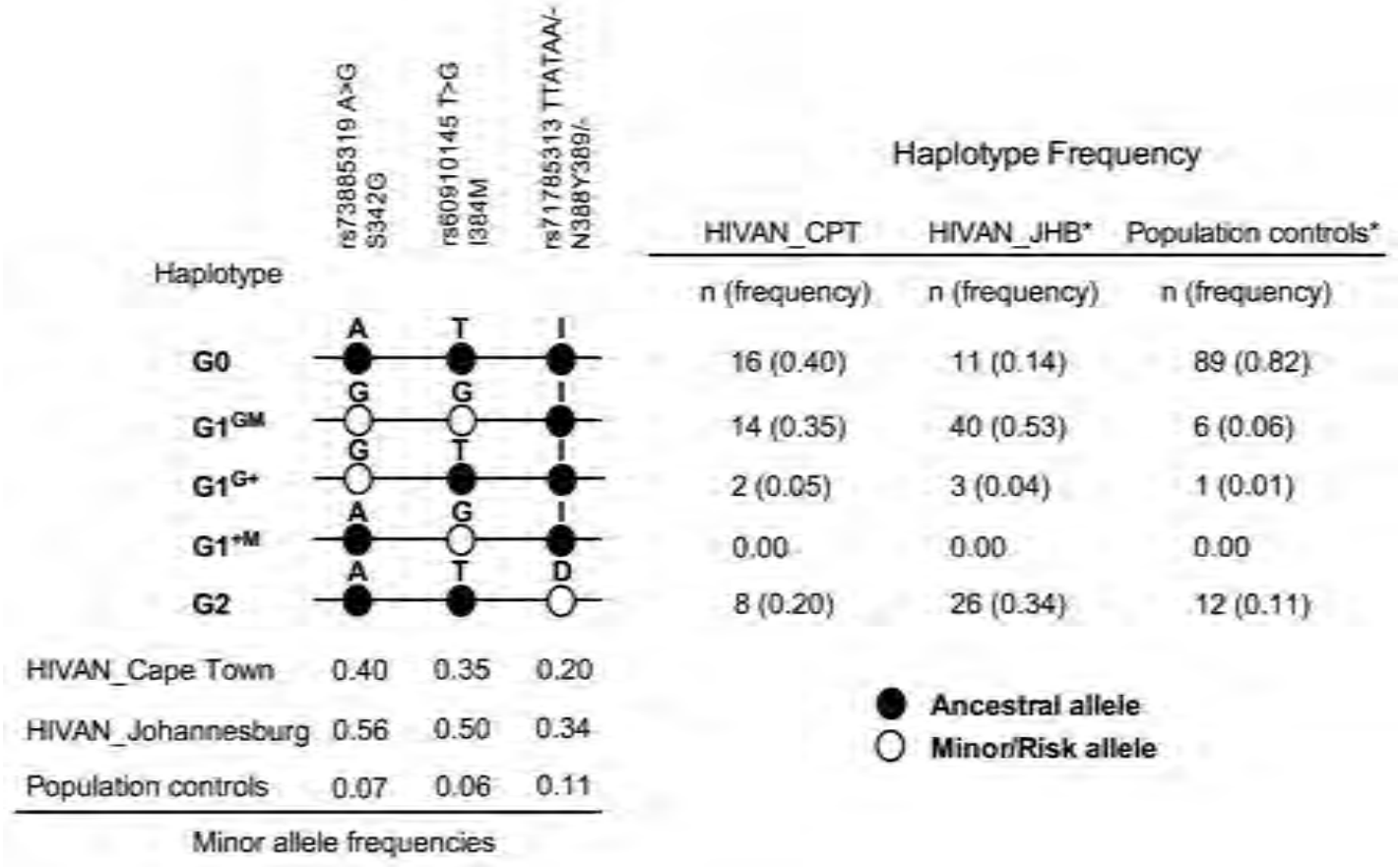


Figure 4.9: Distribution of *APOL1* minor allele frequencies and haplotypes for three groups of South Africans with African-ancestry. Five *APOL1* risk alleles were observed. G0 respresents the ancestral haplotype with all three ancestral alleles. $G1^{GM}$ haplotypes has two missense alleles, $G1^{G+}$ haplotypes has one missense allele and $G1^{+M}$ has one missense allele. The G2 haplotype has a 6 base pair deletion. Minor allele frequencies and distribution of haplotypes are shown for the n=40 Cape Town HIVAN (CPT), n=76 Johannesburg HIVAN (JHB) and n=108 South African population control groups (HIV-negative and HIV-positive). *Kasembeli *et al.*, 2015.

An increase in the frequency of *APOL1* risk variants seemed to correlate with the presence of renal tubular microcysts, evident by 100% of the JHB-HIVAN group having microcysts compared to 60% in the CPT-HIVAN group (Table 4.1 and figure 4.10). Previously, the *APOL1* risk allele distributions of Africans from Yoruba, Ethiopia, America and South Africa were compared to that of Europeans and can be seen in figure 4.10 (Friedman & Pollak 2011).

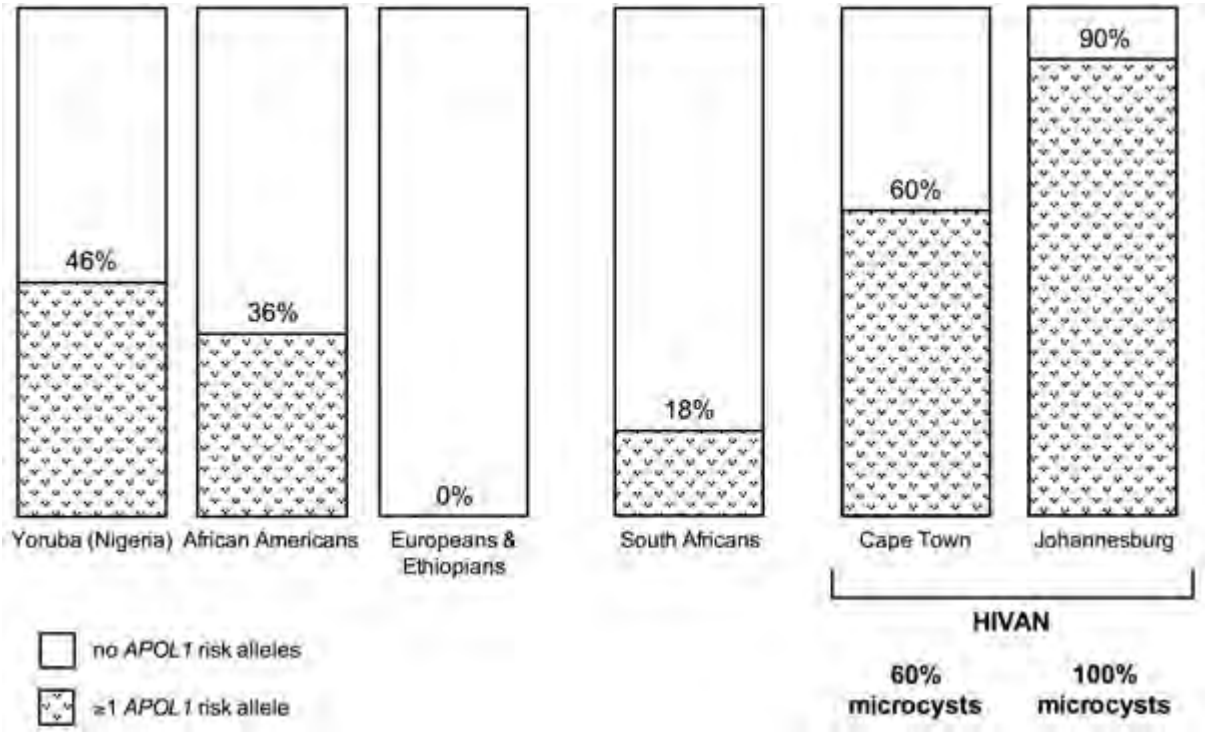


Figure 4.10: *APOL1* risk allele distribution among different population groups. In Yoruba (Nigeria) 46% of the population have at least one of the G1 or G2 *APOL1* risk alleles. African Americans have a frequency of 36% while the risk alleles do not occur in a European or Ethiopian population. South Africans with black African-ancestry have a frequency of 18% in the general population, and at least 60% of those with HIVAN. The frequency of the *APOL1* risk alleles are increased to 90% when only those individuals with microcysts are considered (Johannesburg).

Taken together the minor allele and haplotype frequencies of *APOL1* risk alleles in the CPT-HIVAN sample differed substantially from the population controls, but were lower compared to the JHB-HIVAN sample. The differences between the CPT-HIVAN and JHB-HIVAN samples may be attributed to the stringency of histopathological diagnosis of HIVAN. However, it may also suggest an association between the *APOL1* risk variants and the development of renal tubular microcysts as this was the biggest difference between the

criteria used (Figure 4.10). 100% of the JHB-HIVAN group had microcysts compared to 60% in the CPT-HIVAN group (Figure 4.10).

Frequency of TGFB1 -387 C>T in HIVAN patients similar to population controls

The *TGFB1 -387 C>T* variant was also genotyped for in the CPT-HIVAN sample, but due to technical difficulties in sequencing the *TGFB1* promoter region, the other African-specific variants and the *-1347 C>T TGFB1* ancestral variant were not included in the genotyping. Only 17 patients were genotyped successfully. The MAF (n=34 chromosomes, MAF=0.18) for the *TGFB1 -387 T* variant in the CPT-HIVAN sample was similar to population controls including both Cape Town controls (n=130 chromosomes, MAF 0.16) and Sowetan controls (n=188 chromosomes, MAF 0.18) (Nel *et al.,* 2015). Of interest, two of the patients who had the *-387 T* variant were homozygous for the G1 risk variants in the *APOL1* gene and were previously characterised as having severe HIVAN, fibrosis and microcysts (see table 4.1). Although only one sample, a biopsy with severe fibrosis, microcysts and significant TGF-β1 staining in the interstitium, did not have either *APOL1* risk alleles or *-387 T TGFB1* variant (HVN18, table 4.1).

In conclusion, genotyping this small sample of HIV-infected individuals who developed HIVAN showed that the African-specific *TGFB1 –387 T* regulatory variant occurred at similar frequencies in both the Cape Town HIVAN and South African population control cohorts. *APOL1* risk alleles had higher frequencies in the CPT-HIVAN sample when compared to population controls but there were some individuals who had none of the risk alleles and still displayed histopathological features associated with HIVAN.

5. Discussion

This study focused on the functional impact of the African-specific SNPs in the *TGFB1* promoter region. Because of the high prevalence of HIV-associated kidney complications and specifically HIVAN in Africans, the functional effects of these SNPs were investigated in kidney cells. The four African-specific SNPs (*-1287 G>A*, *-1154 C>T*, *-387 C>T* and *-14 G>A*) occur at higher frequencies in South Africans (~17%) compared to <5% in East and West African populations (this study and Nel *et al.*, 2015). Haplotype analyses showed that the *TGFB1 -1287 G>A* and *-1154 C>T* variants are in tight linkage disequilibrium with each other and so are the *-387 C>T* and *-14 G>A* variants (Figure 4.1B). However, promoter-

reporter assays showed that of the four African-specific variants, only the *-387 C>T* variant was functional in renal cells (Figure 4.2). The *TGFB1 -387 C*>T SNP resulted in ~5-fold repression of basal *TGFB1* promoter activity in two renal cell lines similar to previous findings by Nel *et al.* (2015) in fibroblast and mouse myoblast cells. This repressive effect was also reflected at basal TGF-β1 protein level in human dermal fibroblasts with the *TGFB1 -387 T/T* genotype compared to wild type *-387 C/C* cells.

Renal and muscle cells may respond to autocrine and paracrine TGF-β1 signals in response to injury (Li et al. 2004; Böttinger 2007). To test this pathway renal and fibroblast cells were treated with recombinant human TGF-β1 (rhTGF-β1) for 1 or 6 hours. In response to this stimulus the *TGFB1 -387 T* associated repression was maintained at both the *TGFB1* promoter and TGF-β1 protein level, but interestingly not at the Smad3 level (Figure 4.3, 5.1). This result is similar to previous observations in transiently transfected mouse myoblasts which showed that treating cells with rhTGF-β1 for 6, 12 and 24 hours could not overcome the repressive effect observed in basal *TGFB1 -387 T* promoter activity (Supplementary figure 6.1, unpublished data from M Nel, 2014). However, activation of Smad3 after rhTGF-β1 treatment, as measured by phosphorylated-Smad3 protein levels downstream of the TGF-β1-dependent canonical signalling pathway (Lan 2011), was similar in dermal fibroblasts with either *TGFB1 -387 C* or *T* genotypes (Figure 5.1). This suggests that TGF-β1 autocrine signalling is normal until pSmad3 enters the nucleus, where the functional *TGFB1 -387 T* SNP results in 5-fold repression of basal promoter activity compared to the *-387 C* (Figure 5.1). Previous reports have shown that the *TGFB1* regulatory region contains many Sp1 (a transcriptional modulator) binding sites and that the *-387 T* SNP was predicted to alter Sp1 binding (Nel et al. 2015). Nel *et al.* (2015) showed in fibroblast and muscle cells that the *TGFB1 -387 T* was less responsive to Sp1 binding when compared to the *C* allele (Nel et al. 2015). This means that the loss of the Sp1 binding at the *TGFB1 -387 T* site is likely altering basal *TGFB1* promoter activity and downstream TGF-β1 protein expression (Figure 5.1).

This functional *-387 T* SNP is in tight linkage disequilibrium with another African-specific SNP, *TGFB1 -14 G>A*. Previously, using electrophoretic mobility shift assays, Shah *et al.* (2006) reported the *TGFB1 -14 G>A* not to affect nuclear factor binding. Nel *et al.* (2015)

performed luciferase reporter assays in fibroblast and mouse myoblast cells and found the *-14 G>A* SNP not to be functional and this study confirms this finding in renal cells

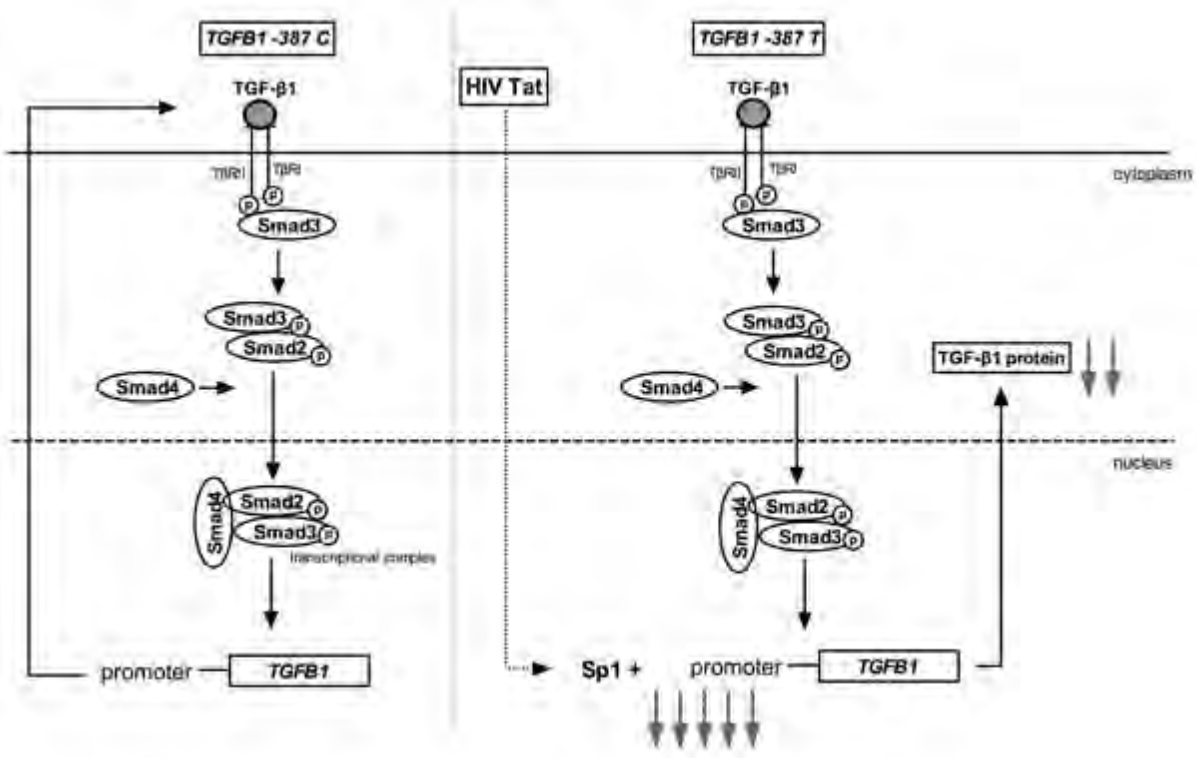


Figure 5.1: A model of TGF-β1 signalling in the kidney and the effect of the *TGFB1 -387 C>T* polymorphism on *TGFB1* promoter activity and protein expression. The schematic also shows the possible influence of HIV Tat on the *TGFB1* promoter. TGF-β1 Smad 3 (pSmad3) pathway is normal for both the *TGFB1 -387 C* and *T* allele but basal *TGFB1* promoter and protein levels are significantly lower for the *T* allele. The repression could be caused by the loss of a partial Sp1 transcription factor binding site at the *-387 C>T* SNP which effects promoter activity. Co-transfection experiments with HIV Tat protein showed that the repressive effect of the *-387 T* variant could be overcome and the exact mechanism is still poorly understood, however, studies have suggested that HIV Tat is able to interact with Sp1 in order to enhance the transcription of genes. Grey arrows at the *TGFB1* promoter and protein indicate repression/decrease in promoter activity and protein levels, respectively.

Similarly, the other two African-specific SNPs, *-1154 C>T* and *-1287 G>A*, are also in tight linkage disequilibrium with each other (see figure 4.1B). Previous studies, using a shorter *TGFB1* promoter construct (see figure 1.3), reported the *TGFB1 -1287 A* variant to be functional with the *A* allele resulting in 78% increased *TGFB1* promoter activity (Shah, Rahaman, et al. 2006). However, using the extended *TGFB1* promoter construct results showed that the *TGFB1 -1154 C>T* and *TGFB1 -1287 G>A* variants are not functional in two different renal cell lines. These differences may be the result of using an extended *TGFB1*

promoter construct in this work which contains enhancer and negative regulatory regions that the shorter *TGFB1* promoter construct used by Shah *et al.* (2006) lacked. The extended *TGFB1* promoter also contains the functional *-387 C>T* variant which was not included in the shorter *TGFB1* promoter construct.

Many studies have reported on associations with the *TGFB1 -1347 C>T* variant (also known as *-509 C>T* or rs1800469). The *TGFB1 -1347 C>T* variant has been reported to be functional where the *T* allele is referred to as the "higher expressing" variant resulting in a ~2-fold increase in basal *TGFB1* promoter activity in lung carcinoma cells, but using a shorter promoter construct (Shah, Rahaman, et al. 2006). This study used the extended *TGFB1* promoter construct, which is more biologically relevant as it contains all four of the African-specific SNP sites. In this work, the *TGFB1 -1347 C>T* (or *-509 C>T*) was not found to be functional in renal cells. Previously Nel *et al.* (2015) using the same extended promoter found that the *TGFB1 -1347 T* (called pHap-4) resulted in 2-fold increased promoter activity compared with the *TGFB1 -1347 C* (called pHap-1) in HT1080 fibroblasts, but similar promoter activities in mouse myoblasts (C2C12 cells) (Table 5.1). Taken together these results suggest that the functionality of the *TGFB1 -1347 C>T* is not influenced by proximal promoter elements present in the extended promoter such as the African SNPs, but by *trans*-regulatory factors/genes which impact its functionality in different cell types (Table 5.1).

Table 5.1: Functional responses of the *TGFB1* promoter *-1347 C>T* in response HIV Tat in different cell types.

		TGFB1 -1347 C>T (*-509 C>T*)	
	Cell type	C	T
BASAL			
	Renal (RCC4+VHL and Caki-2)	=	=
	Fibroblast (HT1080)	-	++
	Mouse myoblasts (C2C12)	=	=
	Lung (A549)	-	++
HIV TAT*			
	Renal (RCC4+VHL)	++	++
	Fibroblasts (HT1080)	++	-

Basal refers to the basal promoter activity for the *TGFB1 -1347 C* or *T* allele observed in luciferase reporter assays and HIV TAT* refers to response in comparison to basal promoter activity after co-transfection with an HIV *tat* expression vector. Similar responses between the *C* and *T* alleles are indicated by "=", – indicates no response compared to empty vector, /basal, ++ refers to ≈ 2-fold upregulation from basal.

This study was also interested in whether the African-specific *TGFB1* promoter variants had a potential role in HIV-associated renal complications which appear to be more frequent among African-Americans compared to European-Americans (Wyatt & Klotman 2009; Ray 2012). To this end co-transfections were performed with *TGFB1* promoter haplotypes and a HIV Tat expression vector in renal and fibroblast cell lines. Based on the basal repression observed with *TGFB1 -387 T* promoter activity, one might expect the common ancestral *TGFB1 -1347 C* haplotype (containing *-387 C*) to show a greater response. However, all the *TGFB1* promoter haplotypes, with or without *TGFB1 -387 T,* all showed a similar increase in promoter activity in response to HIV Tat in renal cells.

Previous studies have reported on the possibility that the presence of HIV Tat *trans*-activates Sp1, through the interaction with transcription co-factors, which then leads to the phosphorylation/activation of Sp1 in order to enhance HIV-1 viral transcription (Chun et al. 1998; Loregian et al. 2003). Therefore, it is possible that HIV Tat may enhance Sp1 activity which then overcomes the repressive effect of the *TGFB1 -387 T* SNP. However, in fibroblast cells there was a differential response noted at the *TGFB1 -1347 C>T* (or *-509 C>T)* site in response to HIV Tat (see Table 5.1). The *TGFB1* haplotypes containing the *-1347 T* (previously higher expressing promoter in fibroblasts) consistently did not respond to HIV Tat in fibroblasts. The significance of these results are unclear as there are limited data investigating the interaction between HIV Tat and the *TGFB1* promoter. Previous studies, however, have shown that the *TGFB1 -1347 C>T* site impacts on the binding of a transcription factor called activator protein-1 (AP1) where the binding of this transcription factor to the *-1347 C* allele decreases *TGFB1* promoter activity (Shah, Hurley, et al. 2006). Therefore, the *TGFB1 -1347 C>T* variant may lead to the activation of different pathways that are phenotype specific. Previously, studies in renal biopsies focused on the effect of HIV Tat on TGFB1 mRNA and protein expression rather than the mechanism whereby the observed upregulation occurs. Taken together, these results suggest that HIV Tat influences the functionality of the *TGFB1 -1347 C>T* (or *-509 C>T)* variant in fibroblasts cells.

The mechanism of HIVAN pathogenesis is still not fully understood (Ross et al. 2001; Medapalli et al. 2011) and this study had the opportunity to assess TGF-β1 staining in a small sample of renal biopsies from HIV-positive patients with renal dysfunction. These individuals were about to start antiretroviral therapy and their TGF-β1 staining patterns were compared to renal biopsies from HIV-negative and HIV-positive control biopsies without HIVAN. The most striking finding was that the HIVAN biopsies showed specific TGF-β1 staining of the renal interstitium in contrast to the biopsies from HIV-positive and -negative controls despite similar intensity and percentage of tubular epithelial cells staining positive for TGF-β1 to control biopsies. However, this result must be interpreted with caution as it was a small sample and the result differs from previous studies where increased TGF-β1 staining in the interstitium was observed for HIV-positive biopsies even without HIVAN compared with HIV-negative biopsies (Bódi et al. 1997; Yamamoto et al. 1999). Future studies will have to investigate TGF-β1 staining in a larger sample, particularly for HIV-positive non-HIVAN controls.

DNA samples from the individuals who had been biopsied for TGF-β1 staining were genotyped for the *TGFB1* promoter variants as well as the previously identified *APOL1* HIVAN risk variants. Unfortunately, due to technical difficulties only the *TGFB1 -387 C>T* genotyping was completed in 17 of 20 individuals. Although a small sample, the genotype frequencies at this locus in the HIVAN group was identical to population controls arguing against an association with the functional African-specific *TGFB1 -387 C>T* SNP. The *APOL1* genotyping confirmed the previous observations of Kasembeli *et al.* (2015) that the ≥1 *APOL1* risk alleles occurred at a higher frequency in the South Africans with HIVAN compared to population controls (Figure 4.9). There is also a difference observed for *APOL1* risk alleles frequencies between the Johannesburg HIVAN (Kasembeli *et al.,* 2015) and the Cape Town HIVAN samples (p=0.002) raising the question of HIVAN classification criteria used by the two studies. There is a general agreement about the histopathological features of HIVAN in the literature, but exact defining histopathological characteristics appear to be less agreed upon. For example, in this study only 8/20 individuals from HIVAN group had microcysts while all (n=76 individuals) of the individuals in the Johannesburg HIVAN sample had microcysts (Kasembeli et al. 2015) (Figure 5.2). In addition it may suggest that

APOL1 variants increases the susceptibility of developing renal tubular microcysts in HIVAN. *APOL1* risk alleles are not present in Europeans, likely due to positive evolutionary pressure in areas such as West Africa which resulted in the influence of *APOL1* variants in order to infer resistance to trypanosomal parasites specific to the area (Genovese et al. 2010). The frequency of *APOL1* risk alleles in African Americans is 46% because their genetic ancestry is likely more than 80% African. Interestingly, the frequency of the risk alleles in Ethiopians is ~0% and in South African black controls 18% (Figure 4.10).

Bioinformatics analysis through SIFT and PolyPhen-2 software (available on Ensemble, http://www.ensembl.org/Multi/Search/Results?q=rs73885319;site=ensembl) shows that the two G1 risk alleles are missense variants that are benign while the effect of the 6bp deletion G2 risk allele is still unknown. Nichols *et al.* (2015) has shown that the regulation of *APOL1* may be through antiviral pathways, such as the Toll-like receptor 3 pathway (Nichols et al. 2015), but studies investigating *APOL1* regulation is still limited and needs to be further explored. One study has shown that APOL1 staining in renal tubular epithelial cells is lower in HIVAN biopsies compared to normal kidneys (Madhavan et al. 2011). However, Although *APOL1* risk alleles co-occur with a variety of different kidney diseases (Larsen et al. 2015) it would also be interesting to assess whether it associates with a particular histopathological pattern. Such genotype-phenotype correlation may give more insight into the pathogenesis of HIVAN.

Previous studies have shown that *APOL1* is regulated by cytokines released during the innate immune response which includes the interferon-gamma(γ) (IFN-γ) and tumour necrosis factor (TNF) pathways and influences the inflammatory response and *APOL1* expression in renal tubular epithelial cells and podocytes (Sana et al. 2005; Nichols et al. 2015). Figure 5.3 shows a STRING map of potential protein-protein interactions for APOL-1 which is generated using STRING version 10.0 software (http://string-db.org/cgi/network.pl?taskId=hJgCmFgJagM7) based on prediction software data from other databases and published data in the literature (Szklarczyk et al. 2015). Interestingly, the Sp1 transcription factor was identified as a possible protein interacting with APOL1 and may be a potential regulator of *APOL1* (Figure 5.2).

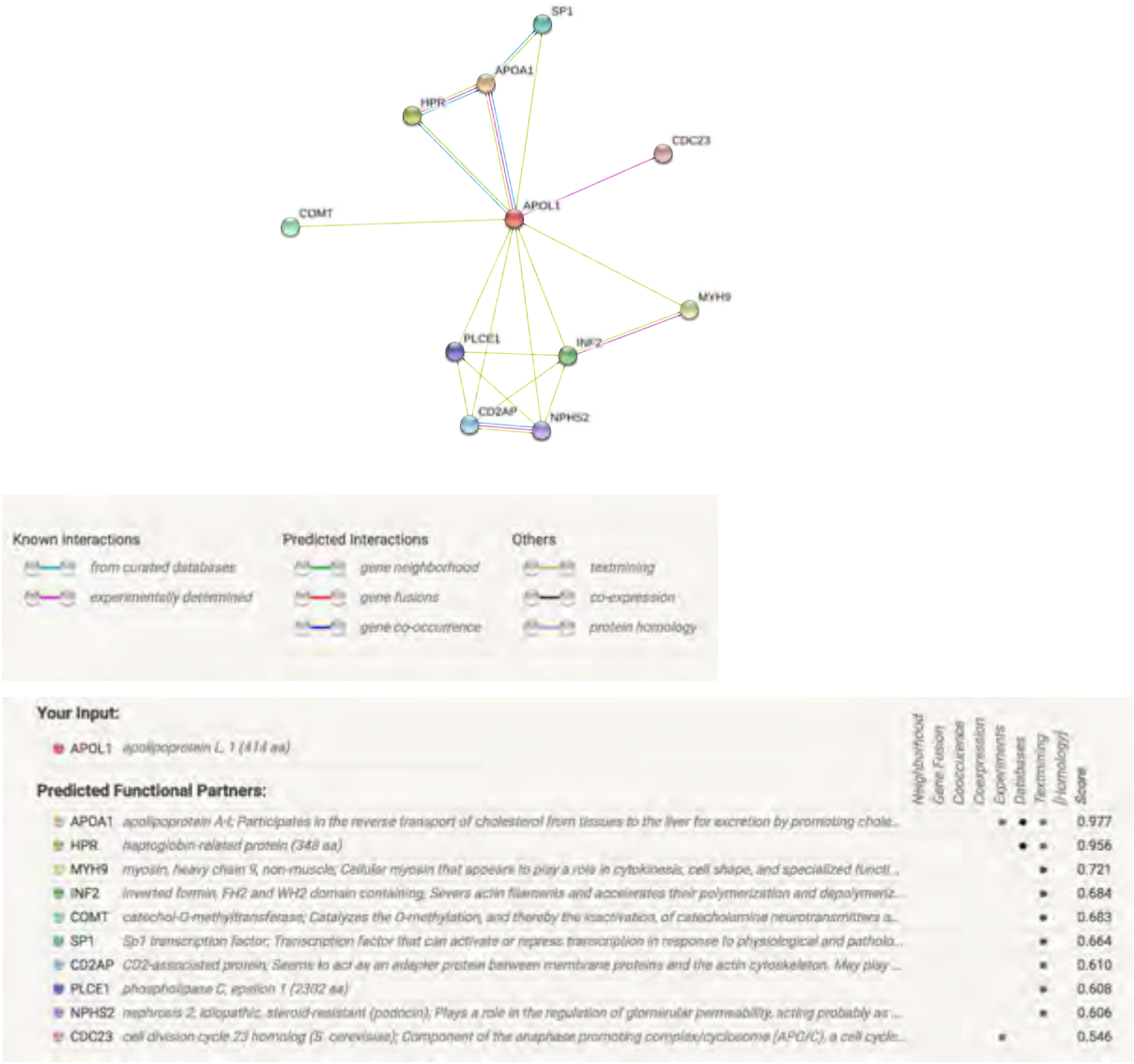


Figure 5.2: STRING map showing potential protein-protein interactions for APOL1. As indicated in the key blue and pink lines indicate known interactions between proteins with evidence acquired form curated databases and experimental data. The green lines indicate potential protein-protein interactions between those proteins often mentioned together in the literature. Protein-protein interaction scores are also shown and all fall in the medium-to-high confidence range which is calculated based on experimental data of each interaction taking into account sequence, structure and functionality (Szklarczyk et al. 2015). The protein-protein interaction score for the Sp1 transcription factor is of high confidence (0.664).

Limitations and future studies

This study has several limitations. The sample size for Cape Town HIVAN biopsies was too small for any significant conclusions based on TGF-β1 staining patterns or *TGFB1* genotyping. Also, the *TGFB1 -1347 C>T* variant was not included in the genotyping of the HIVAN samples and may be informative in future studies.

This study did not explore the possible transcription factor binding sites (TFBS) that may be influenced by the *TGFB1 -1347 C>T* variant, such as AP1. Das and Sharma (2014) discuss the importance of investigating the genetic factors modulating susceptibility of complex diseases and their manifestations. For example, the expression of a gene/protein can be influenced by *cis* promoter variants such as the *TGFB1 -1347 C>T* and *trans* promoter variants that are further away (more than 500kb away) from the transcription start site or on a different chromosome. It has also been shown that *trans* variants are able to effect tissues specific expression of genes (Das & Sharma 2014; Bryois et al. 2014), therefore, in the context of the *TGFB1 -1347 C>T* variant and HIV, there seems to be a *trans*-variant/gene effect causing differential expression in several different cell types that needs to be investigated further. HIV infection of renal cells can also effect epigenetic changes in gene regulation which, in the context of HIVAN, might be a master regulator of the TGF-β1 pathway.

This study only interrogated TGF-β1 and phosphor-Smad3 protein levels as an indication of the TGF-β1 autocrine pathway, but as mentioned earlier in section 1.2 and 1.3 there are many other factors involved in the regulation of TGF-β1 signalling in the kidney. Future studies may investigate TGF-β1 signalling inhibitors and TGF-β1-independent pathways in order to fully characterize the TGF-β1 pathway in HIVAN individuals with African-ancestry.

The possible interaction between HIV Tat and the transcriptional activator Sp1 was not investigated in this study, neither was the effect of the *TGFB1 -387 C>T* variant on Sp1 binding in the context of renal cells. Future studies may want to investigate whether the same Sp1 sensitivity at the *TGFB1 -387 C>T* site is observed in renal cells. Previous studies have analysed TGF-β1 protein levels by treating cells with human recombinant HIV Tat protein which is a more biologically relevant way to investigate its effect (Lotz et al. 1994; Yamamoto et al. 1999). This study used a HIV Tat expression vector in order to overexpressed HIV Tat intracellularly and therefore the results may be different when using extracellular HIV Tat stimulation. Due to the fact that HIV-infection occurs in the renal tubular epithelial cells and because HIV Tat and Nef proteins have been shown to act together to enhance viral transcription in kidneys using a mouse model (Joseph et al. 2003;

Dickie et al. 2004), one might speculate that these two proteins act together in a HIVAN context. Future studies might want to investigate whether the addition of other HIV genes have an effect on *TGFB1* promoter activity using *in vitro* experiments with renal cells where expression vectors contain *tat*, *nef* and *vpr* genes. (Dickie et al. 2004; Zuo 2006; Rosenstiel et al. 2009).

The HIVAN biopsies used in this study to interrogate TGF-β1 staining patterns were taken before the patients were given antiretroviral therapy (ART). Previous reports have shown that ATR leads to decreased fibrosis in HIVAN biopsies and it has been suggested that ART and steroids are used in the treatment of HIVAN (Weiner et al. 2003; Kalayjian 2011; Wearne et al. 2012). In the context of this study it will be of interest to investigate whether the TGF-β1 staining in the interstitium decreases after ART and compare the staining patterns with renal histopathology/outcomes.

As previously mentioned in section 1.3, changes in *APOL1* expression can disrupt renal epithelial cell homeostasis and autophagy. Although this study showed that the *APOL1* risk alleles have a strong correlation with the presence of microcysts in the HIVAN biopsies, the mechanism whereby this is happening was not investigated and needs further exploration.

Future studies may also want to investigate whether there are differences in APOL1 protein expression between the Cape Town HIVAN biopsies and HIV-negative and –positive controls. The exact mechanisms for APOL1 regulation is still poorly understood and further exploration into which pathways are responsible will benefit our understanding of the role *APOL1* plays in the pathogenesis of HIVAN.

Conclusion

In conclusion this study has shown that there is no difference in *TGFB1* promoter activity between the *-1347 C* and *T* alleles in two renal cell lines and suggests that the functionality of this SNP may be cell type specific. This study also shows that the African-specific *TGFB1 -387 C>T* variant is functional and results in repression of basal promoter activity and lowered TGF-β1 protein levels when compared to dermal fibroblasts with the *-387 C* allele. The TGF-β1-Smad3 pathway was normal in all cells, despite having the *TGFB1 -387 T* SNP, and exogenous treatment with TGF-β1 did not overcome the repressive effect of the *-387 C>T*

variant on *TGFB1* promoter activity. This suggests that the pSmad3 complex is unable to activate the *TGFB1* promoter once inside the nucleus, which may be due to the *-387 C>T* resulting in the loss of a transcription factor binding site. This work also confirms that HIV Tat upregulates TGF-β1 protein levels in renal cells, but shows that the *TGFB1 -1347 T* allele is unresponsive compared to cells with the *-1347 C* allele in fibroblast cells. This study is also one of few that suggests the possible influence of *APOL1* risk alleles on the formation of microcysts in renal tubular cells. Taken together this study highlights the complexity of HIVAN pathogenesis and TGF-β1 signalling in the kidney and suggests that susceptibility is influenced by more than one risk variant which influences the regulation of many different pathways. These risk variants together with environmental factors, such as HIV-infection, contribute to molecular phenotypes which ultimately result in the formation of glomerular, interstitial and tubular abnormalities in the kidney.

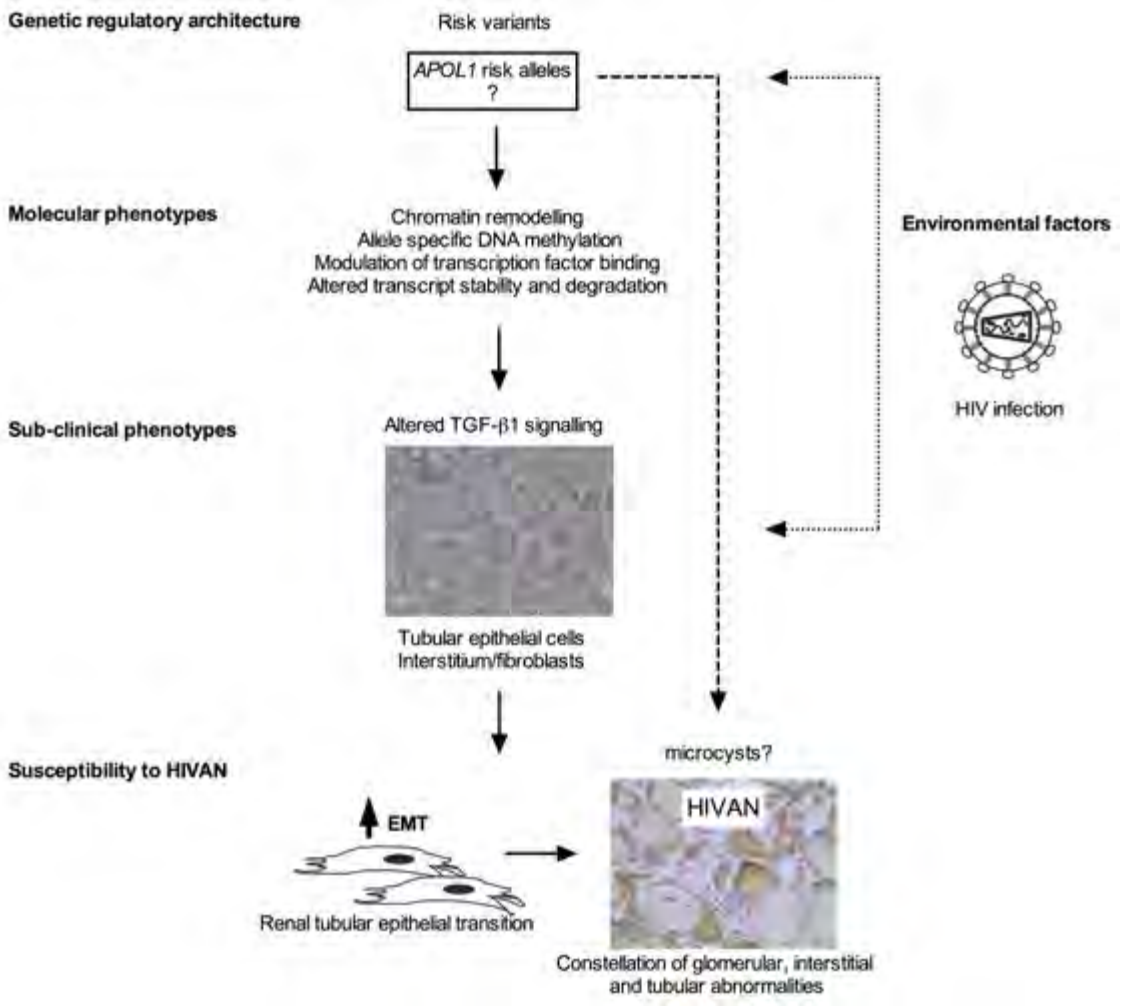
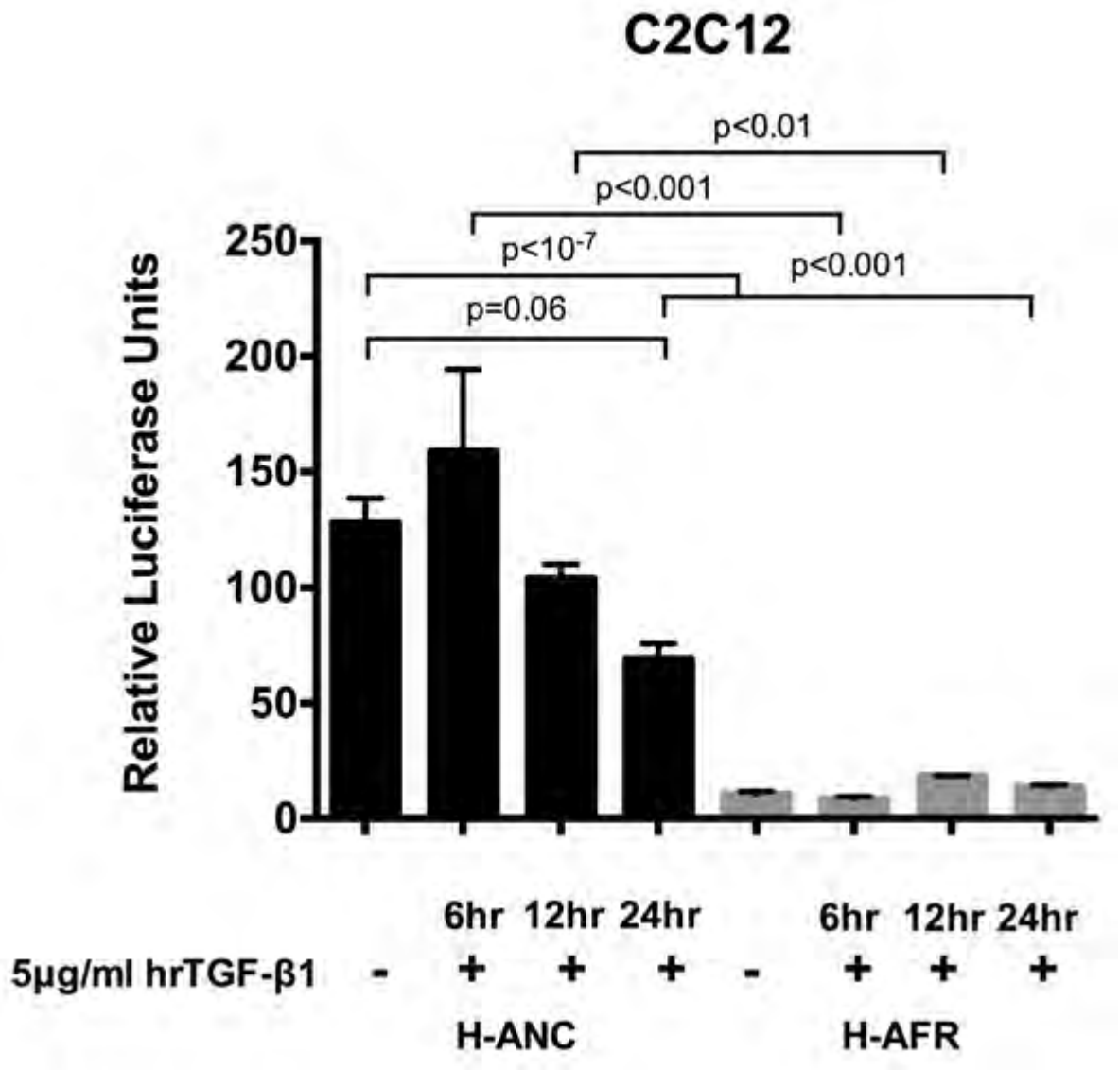


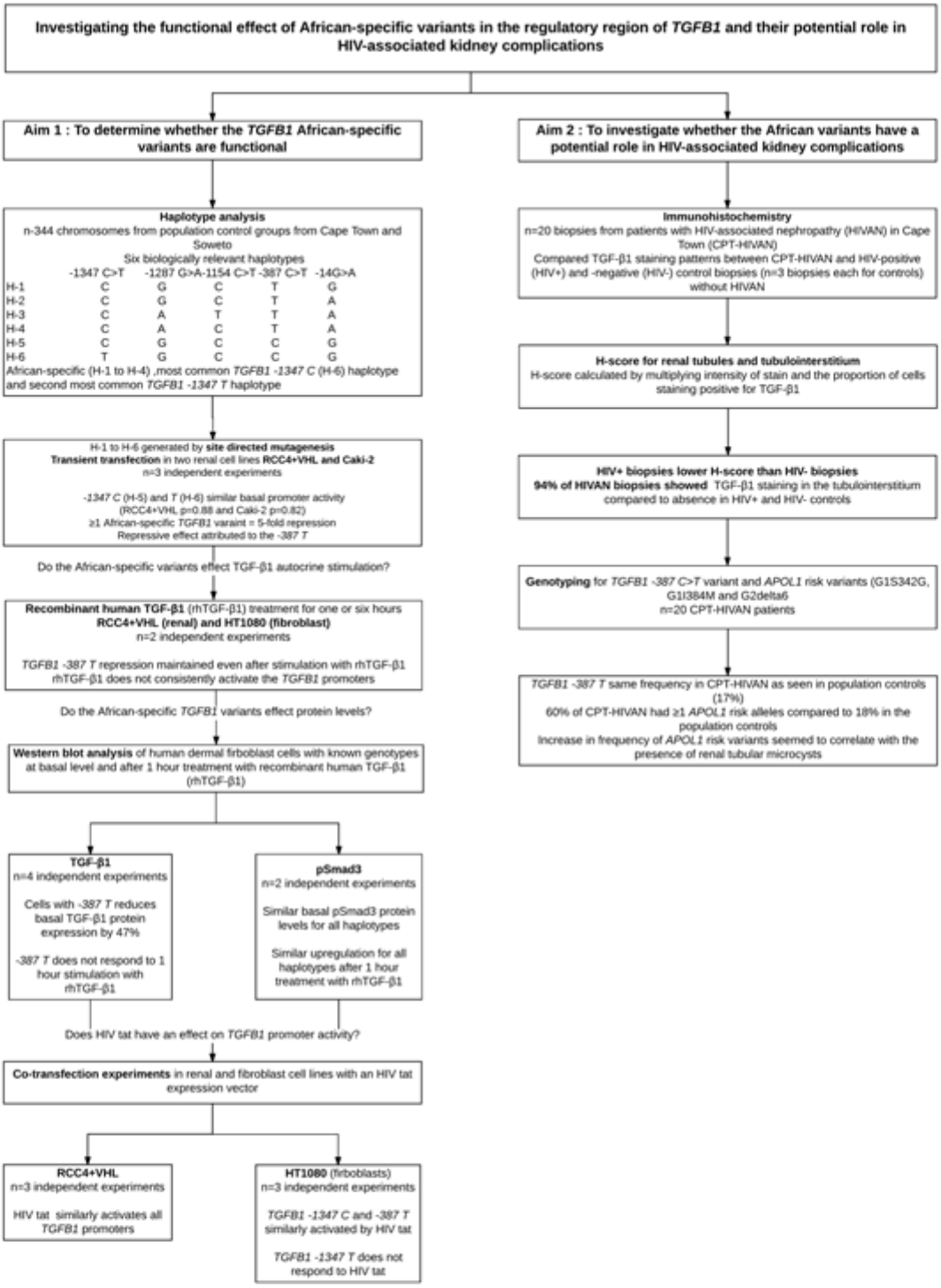
Figure 5.3: Proposed model of biological factors influencing HIVAN. HIV-infection together with risk variants, such as the *APOL1* HIVAN risk alleles and other unidentified variants,

leads to altered gene regulation and altered TGF-β1 signalling in renal tubular epithelial cells and fibroblasts in the renal interstitium. This dysregulation can increase the susceptibility to developing HIVAN and lead to a constellation of glomerular, interstitial and tubular abnormalities in the kidney. The *APOL1* risk variants also seem to be driving the formation of renal tubular microcysts during HIVAN. Figure adapted from Das and Sharma (2014).

6. Supplementary figures



Supplementary figure 6.1: Human recombinant TGF-β1 does not consistently activate the *TGFB1* promoters in mouse myoblast cells. C2C12 (mouse myoblasts) were transiently transfected for 30 hours with H-ANC and H-AFR *TGFB1* promoter haplotypes and serum starved 16 hours prior to harvesting. The cells were treated with 5µg/ml human recombinant TGF-β1 (hrTGF-β1) for 6, 12 and 24 hours before harvesting. Luciferase activity was measured 30 hours following transient transfection and expressed as the ratio of firefly to renilla luciferase activity for each of the haplotypes. pRL renilla reporter was used to control for transfection efficiency. These are pooled results of ≥2 independent experiments (except 12-hour treatments for both promoters), each performed in duplicate. Statistical differences (including trends) are shown in the figure. Student's paired t-test, two-tailed, was performed for treated versus untreated for each of the haplotypes, error bars standard error of the mean (SEM). Student's unpaired t-test, two-tailed, was performed to compare the H-ANC and H-AFR *TGFB1* haplotypes, two-tailed, error bars SEM.



Supplementary figure 6.2: Flow-chart with summary of techniques and results for this study. Aim 1 was to determine whether the African-specific *TGFB1* regulatory/promoter variants were functional. To this end haploview analysis was performed on n=344 chromosomes from population controls and six biologically relevant haplotypes were identified and generated via site-directed mutagenesis. The *TGFB1* promoters were then transiently transfected into two renal cell lines (RCC4+VHL and Caki-2) and promoter activity was measured in order to compare functionality of the African-specific variants. The effect of the African-specific variants on TGF-

β1 autocrine stimulation was also investigated at *TGFB1* promoter level (with recombinant human TGF-β1) before investigating their effect on TGF-β1 and pSmad3 protein levels (as an indicator of downstream TGF-β1 canonical pathway). TGF-β1 and pSmad3 protein levels were also measured after stimulation with recombinant TGF-β1 in order to investigate whether the effects seen at *TGFB1* promoter activity is maintained at the protein level. The effect of HIV tat protein on the *TGFB1* promoters and whether the *TGFB1 -387 T* affects the activation/repression was also investigated. Aim 2 focused on the potential role/effect of the African *TGFB1* promoter variants in HIV-associated kidney complications. To this end n=20 biopsies form patients with HIV-associated nephropathy (HIVAN) were analysed for TGF-β1 staining patterns and compared to those of HIV-positive and –negative control biopsies without HIVAN. The genomic DNA from the n=20 HIVAN patients were also genotyped for the *TGFB1 -387 C>T* and *APOL1* risk alleles and compared to previously genotyped population control groups from South Africa.

7. References

Abbott, K.C. et al., 2001. Human immunodeficiency virus/acquired immunodeficiency syndrome-associated nephropathy at end-stage renal disease in the United States: patient characteristics and survival in the pre highly active antiretroviral therapy era. *JOURNAL OF NEPHROLOGY*, 14(5), pp.377–383.

August, P. et al., 2009. Transforming growth factor beta and excess burden of renal disease. *TRANS AM CLIN CLIMATOL ASSOC*, 120, pp.61–72.

Barrett, J.C. et al., 2005. Haploview: analysis and visualization of LD and haplotype maps. *BIOINFORMATICS (OXFORD, ENGLAND)*, 21(2), pp.263–5.

Bidwell, J. et al., 1999. Cytokine gene polymorphism in human disease: on-line databases. *GENES IMMUN*, 1(1), pp.3–19.

Bódi, I. et al., 1997. Renal TGF-beta in HIV-associated kidney diseases. *KIDNEY INTERNATIONAL*, 51(5), pp.1568–77.

Bonniaud, P. et al., 2005. TGF-beta and Smad3 signaling link inflammation to chronic fibrogenesis. *J IMMUNOL*, 175(8), pp.5390–5395.

Border, W.A. & Noble, N.A., 1994. Transforming growth factor-beta in glomerular injury. *EXP NEPHROL*, 2(1), pp.13–17.

Böttinger, E.P., 2007. TGF-β in Renal Injury and Disease. *SEMINARS IN NEPHROLOGY*, 27(3), pp.309–320.

Bryois, J. et al., 2014. Cis and Trans Effects of Human Genomic Variants on Gene Expression. *PLOS GENETICS*, 10(7).

Cheng, K. et al., 2013. MicroRNAs in HIV-associated nephropathy (HIVAN). *EXPERIMENTAL AND MOLECULAR PATHOLOGY*, 94(1), pp.65–72.

Cheng, X. et al., 2013. Both ERK/MAPK and TGF-Beta/Smad Signaling Pathways Play a Role in the Kidney Fibrosis of Diabetic Mice Accelerated by Blood Glucose Fluctuation. *JOURNAL OF DIABETES RESEARCH*, 2013, p.8.

Chun, R.F. et al., 1998. Modulation of Sp1 phosphorylation by human immunodeficiency virus type 1 Tat. *JOURNAL OF VIROLOGY*, 72(4), pp.2615–29.

Das, S.K. & Sharma, N.K., 2014. Expression quantitative trait analyses to identify causal genetic variants for type 2 diabetes susceptibility. *WORLD JOURNAL OF DIABETES*, 5(2), pp.97–114.

Dickie, P. et al., 2004. Focal glomerulosclerosis in proviral and c-fms transgenic mice links Vpr expression to HIV-associated nephropathy. *VIROLOGY*, 322(1), pp.69–81.

Eiser, A.R., 2010. Does over-expression of transforming growth factor-beta account for the increased morbidity in African-Americans?: Possible clinical study and therapeutic

implications. *MEDICAL HYPOTHESES*, 75, pp.418–421.
Epstein, F.H., Border, W.A. & Noble, N.A., 1994. Transforming Growth Factor β in Tissue Fibrosis. *NEW ENGLAND JOURNAL OF MEDICINE*, 331(19), pp.1286–1292.
Freedman, B.I. et al., 2014. End-stage renal disease in African Americans with lupus nephritis is associated with APOL1. *ARTHRITIS & RHEUMATOLOGY (HOBOKEN, N.J.)*, 66(2), pp.390–6.
Friedman, D.J. & Pollak, M.R., 2011. Genetics of kidney failure and the evolving story of APOL1. *J CLIN INVEST*, 121(9), pp.3367–3374.
Gabriel, S.B. et al., 2002. The structure of haplotype blocks in the human genome. *SCIENCE (NEW YORK, N.Y.)*, 296(5576), pp.2225–9.
Genovese, G. et al., 2010. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *SCIENCE*, 329(5993), pp.841–845.
Healy, J. et al., 2009. Functional impact of sequence variation in the promoter region of TGFB1. *INTERNATIONAL JOURNAL OF CANCER. JOURNAL INTERNATIONAL DU CANCER*, 125(6), pp.1483–9.
Heldin, C.-H., Miyazono, K. & ten Dijke, P., 1997. TGF-[beta] signalling from cell membrane to nucleus through SMAD proteins. *NATURE*, 390(6659), pp.465–471.
Hu, C.-A.A., Klopfer, E.I. & Ray, P.E., 2012. Human apolipoprotein L1 (ApoL1) in cancer and chronic kidney disease. *FEBS LETTERS*, 586(7), pp.947–955.
Joseph, A.M. et al., 2003. Human immunodeficiency virus-1 Nef protein interacts with Tat and enhances HIV-1 gene expression. *FEBS LETTERS*, 548(1–3), pp.37–42.
Kalayjian, R.C., 2011. The Treatment of HIV-Associated Nephropathy. *ADVANCED CHRONIC KIDNEY DISEASE*, 17(1), pp.59–71.
Kasembeli, A.N., 2015. African origins and chronic kidney disease susceptibility in the human immunodeficiency virus era. *WORLD JOURNAL OF NEPHROLOGY*, 4(2), p.295.
Kasembeli, A.N. et al., 2015. APOL1 Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans. *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY : JASN*, (November), p.ASN.2014050469-.
Kato, M. et al., 2011. A microRNA circuit mediates transforming growth factor-beta1 autoregulation in renal glomerular mesangial cells. *KIDNEY INT*, 80(4), pp.358–368.
Kim, S.J. et al., 1989. Characterization of the promoter region of the human transforming growth factor-beta 1 gene. *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, 264(1), pp.402–8.
Kopp, J.B. et al., 2011. APOL1 Genetic Variants in Focal Segmental Glomerulosclerosis and HIV-Associated Nephropathy. *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY : JASN*, 22(11), pp.2129–2137.
Kovacs, E.J. & DiPietro, L.A., 1994. Fibrogenic cytokines and connective tissue production. *THE FASEB JOURNAL : OFFICIAL PUBLICATION OF THE FEDERATION OF AMERICAN SOCIETIES FOR EXPERIMENTAL BIOLOGY*, 8, pp.854–861.
Lan, H.Y., 2011. Diverse roles of TGF-β/Smads in renal fibrosis and inflammation. *INTERNATIONAL JOURNAL OF BIOLOGICAL SCIENCES*, 7(7), pp.1056–1067.
Lan, H.Y., 2012. Smads as therapeutic targets for chronic kidney disease. *KIDNEY RESEARCH AND CLINICAL PRACTICE*, 31(1), pp.4–11.
Larsen, C.P. et al., 2015. Histopathologic findings associated with APOL1 risk variants in chronic kidney disease. *MODERN PATHOLOGY : AN OFFICIAL JOURNAL OF THE UNITED STATES AND CANADIAN ACADEMY OF PATHOLOGY, INC*, 28(1), pp.95–102.
Leask, A. & Abraham, D.J., 2004. TGF-beta signaling and the fibrotic response. *FASEB*

JOURNAL : OFFICIAL PUBLICATION OF THE FEDERATION OF AMERICAN SOCIETIES FOR EXPERIMENTAL BIOLOGY, 18(7), pp.816–27.
Lehto, M., Järvinen, M. & Nelimarkka, O., 1986. Scar formation after skeletal muscle injury. A histological and autoradiographical study in rats. *ARCHIVES OF ORTHOPAEDIC AND TRAUMA SURGERY*, 104, pp.366–370.
Li, Y. et al., 2004. Transforming Growth Factor-β1 Induces the Differentiation of Myogenic Cells into Fibrotic Cells in Injured Skeletal Muscle. *THE AMERICAN JOURNAL OF PATHOLOGY*, 164(3), pp.1007–1019.
Liu, Y., 2006. Renal fibrosis: new insights into the pathogenesis and therapeutics. *KIDNEY INT*, 69(2), pp.213–217.
Loregian, A. et al., 2003. Interaction of Sp1 transcription factor with HIV-1 Tat protein: Looking for cellular partners. *FEBS LETTERS*, 543(1–3), pp.61–65.
Lotz, M., Clark-Lewis, I. & Ganu, V., 1994. HIV-1 transactivator protein Tat induces proliferation and TGF beta expression in human articular chondrocytes. *JOURNAL OF CELL BIOLOGY*, 124(3), pp.365–71.
Madhavan, S.M. et al., 2011. APOL1 Localization in Normal Kidney and Nondiabetic Kidney Disease. *J AM SOC NEPHROL*, 22, pp.2119–2128.
Medapalli, R.K., He, J.C. & Klotman, P.E., 2011. HIV- Associated Nephropathy: Pathogenesis. *CURRENT OPINION IN NEPHROLOGY AND HYPERTENSION*, 20(3), pp.306–311.
Mohy, A. & Fouad, A., 2014. Role of transforming growth factor-beta1 in serum and - 509 C>T promoter gene polymorphism in development of liver cirrhosis in Egyptian patients. *META GENE*, 2, pp.631–637.
Montesano, R. & Orci, L., 1988. Transforming growth factor beta stimulates collagen-matrix contraction by fibroblasts: implications for wound healing. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 85, pp.4894–4897.
Nel, M. et al., 2015. The African−387 C>T TGFB1 variant is functional and associates with the ophthalmoplegic complication in juvenile myasthenia gravis. *JOURNAL OF HUMAN GENETICS*.
Nichols, B. et al., 2015. Innate immunity pathways regulate the nephropathy gene Apolipoprotein L1. *KIDNEY INTERNATIONAL*, 87(2), pp.332–342.
Nunes, I. et al., 1997. Latent Transforming Growth Factor-β Binding Protein Domains Involved in Activation and Transglutaminase-dependent Cross-Linking of Latent Transforming Growth Factor-β. *THE JOURNAL OF CELL BIOLOGY*, 136(5), pp.1151–1163.
Olson, E.N. et al., 1986. Regulation of myogenic differentiation by type beta transforming growth factor. *THE JOURNAL OF CELL BIOLOGY*, 103(5), pp.1799–805.
Overall, C.M., Wrana, J.L. & Sodek, J., 1989. Independent regulation of collagenase, 72-kDa progelatinase, and metalloendoproteinase inhibitor expression in human fibroblasts by transforming growth factor-beta. *J BIOL CHEM*, 264(3), pp.1860–1869.
Ray, P.E., 2012. HIV-associated nephropathy: a diagnosis in evolution. *NEPHROLOGY DIALYSIS TRANSPLANTATION*, 27(11), pp.3969–3972.
Roberts, A.B. et al., 1986. Transforming growth factor type beta: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES*, 83(12), pp.4167–4171.
Rosenstiel, P. et al., 2009. Transgenic and infectious animal models of HIV-associated nephropathy. *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY : JASN*,

20(11), pp.2296–304.
Ross, M.J. et al., 2001. Microcyst Formation and HIV-1 Gene Expression Occur in Multiple Nephron Segments in HIV-Associated Nephropathy. *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY* , 12(12), pp.2645–2651.
Sana, T.R. et al., 2005. Microarray analysis of primary endothelial cells challenged with different inflammatory and immune cytokines. *CYTOKINE*, 29(6), pp.256–269.
Shah, R., Rahaman, B., et al., 2006. Allelic diversity in the TGFB1 regulatory region: characterization of novel functional single nucleotide polymorphisms. *HUMAN GENETICS*, 119(1–2), pp.61–74.
Shah, R., Hurley, C.K. & Posch, P.E., 2006. A molecular mechanism for the differential regulation of TGF-beta1 expression due to the common SNP -509C-T (c. -1347C > T). *HUMAN GENETICS*, 120(4), pp.461–9.
Shi, Y. & Massague, J., 2003. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *CELL*, 113(6), pp.685–700.
Suthanthiran, M. et al., 1998. Transforming growth factor-[bgr]1 hyperexpression in African American end-stage renal disease patients. *KIDNEY INT*, 53(3), pp.639–644.
Szklarczyk, D. et al., 2015. STRING v10: Protein-protein interaction networks, integrated over the tree of life. *NUCLEIC ACIDS RESEARCH*, 43(D1), pp.D447–D452.
Tayo, B.O. et al., 2013. Genetic variation in APOL1 and MYH9 genes is associated with chronic kidney disease among Nigerians. *INTERNATIONAL UROLOGY AND NEPHROLOGY*, 45(2), pp.485–494.
Thatikunta, P. et al., 1997. Identification of a cellular protein that binds to Tat-responsive element of TGF beta-1 promoter in glial cells. *JOURNAL OF CELLULAR BIOCHEMISTRY*, 67(4), pp.466–77.
Wang, W., Koka, V. & Lan, H.Y., 2005. Transforming growth factor-β and Smad signalling in kidney diseases. *NEPHROLOGY*, 10(1), pp.48–56.
Wearne, N. et al., 2012. The spectrum of renal histologies seen in HIV with outcomes, prognostic indicators and clinical correlations. *NEPHROLOGY DIALYSIS TRANSPLANTATION*, 27(11), pp.4109–4118.
Weiner, N.J., Goodman, J.W. & Kimmel, P.L., 2003. The HIV-associated renal diseases: Current insight into pathogenesis and treatment. *KIDNEY INTERNATIONAL*, 63(5), pp.1618–1631.
Wrana, J.L. et al., 1994. Mechanism of activation of the TGF-[beta] receptor. *NATURE*, 370(6488), pp.341–347.
Wyatt, C.M. & Klotman, P.E., 2009. HIV-associated Nephropathy. *GENETIC DISEASES OF THE KIDNEY*, 20(3), pp.793–813.
Wynn, T.A., 2010. Fibrosis under arrest. *NAT MED*, 16(5), pp.523–525.
Yamamoto, T. et al., 1999. Increased levels of transforming growth factor-beta in HIV-associated nephropathy. *KIDNEY INTERNATIONAL*, 55(2), pp.579–92.
Yang, L. et al., 2010. Epithelial cell cycle arrest in G2/M mediates kidney fibrosis after injury. *NATURE MEDICINE*, 16(5), pp.143–535.
Zhang, M., Fraser, D. & Phillips, A., 2006. ERK, p38, and Smad signaling pathways differentially regulate transforming growth factor-beta1 autoinduction in proximal tubular epithelial cells. *THE AMERICAN JOURNAL OF PATHOLOGY*, 169(4), pp.1282–93.
Zuo, Y., 2006. HIV-1 Genes vpr and nef Synergistically Damage Podocytes, Leading to Glomerulosclerosis. *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, 17(10), pp.2832–2843.

8. Appendix

Competent cells, site-directed mutagenesis, mini preparation and maxi preparation of DNA

50% Glycerol
To make 100ml:
Take 50ml of 100% glycerol and dilute it with 50ml ddH_2O

100 mM $CaCl_2$
1.4702g $CaCl_2$
Dissolve in 90ml double distilled water (ddH_2O)
Make up to 100ml
Autoclave before use
Store at room temperature

Luria Broth (LB)
Makes 1L
10g Bacto-tryptone
5g Bacto-yeast extract
10g Sodium Chloride (NaCl)
Dissolve in 900ml ddH_2O
Make up to 1L with ddH_2O
Autoclave before use

LB Agar
3g OXOID Agar Bacteriological to 200ml LB
Autoclave before making Ampicillin plates

LB Ampicillin plates
Add 100µg/ml Ampicillin to the LB agar and pour the solution into bacterial plates (±20ml/plate) and leave to set at room temperature (RT).
Makes ±10 plates

1% Agarose gel
Makes 50ml
0.5g Agarose
50ml Tris/Borate/EDTA (TBE)
1µl Ethidium Bromide (EthBr)
Dissolve agarose/Tris solution in microwave and allow to cool before adding the EthBr. Pour gel using casting tray.

10X TBE
Makes 1L
108g Tris
55g Boric acid
40 ml 0.5M Disodium EDTA (pH 8.0)

1X TBE
100ml 10X TBE
Make up to 1L with ddH_2O

Cell culture

DMEM
Makes 1L
13.5g Dulbecco's Modified Eagle's medium powder
3.7g Sodium Hydrogen Carbonate
Dissolve in 700ml and bring up to 1L
Adjust pH to 7.2 – 7.4
Filter sterilize before use

DMEM++
DMEM
10% Fetal bovine serum (FBS)
1% Penicillin/Steptomycin

Mounting fluid
20mM Citric Acid
55mM Disodium hydrogen phosphate dehydrate
50% Glycerol
Adjust to pH 5.5
Store at 4°C

β-galactosidase stain
5mM Potassium ferrocyanide
5mM Potassium ferricyanide
2mM Magnesium chloride
1mg/ml X-gal (in DMSO)

Protein harvest and western blot analysis

5X protein dye
50% Glycerol
5% Sodium-dodecyl sulphate (SDS)
0.25M Tris-Cl pH 6.8
Bromophenol blue (small amount at end of pipette tip)

SDS-polyacrylamide gels

10% resolving gel:
40% Acrylamide
1.5M Tris pH 8.8
10% SDS
10% APS
TEMED

5% stacking gel:
40% Acrylamide
1.5M Tris pH 6.8
10% SDS
10% Ammonium persulphate (APS)
TEMED

10% APS
Makes 1ml
0.1g APS
Make solution up to 1ml using dH_2O
Store at 4°C

10% SDS
Makes 1L
100g SDS
Make up to 1L with dH_2O and adjust to pH 7.2
Store at RT

10X Running buffer
144g Glycerine
30g TRIS powder
10g SDS
Make up to 1L with dH_2O

1X Running buffer
100ml of 10X Running buffer

Make up to 1L with dH_2O

10X Transfer buffer
144g Glycerine
38g TRIS powder
Make up to 1L with dH_2O

1X Transfer buffer
100ml of 10X Transfer buffer
200ml Methanol
Make up to 1L with dH_2O

Phosphate-buffered Saline (PBS) (10x)
Makes 1L
80g NaCl
26.8g Sodium Dihydrogen Phosphate Dodecahydrate
2g Potassium chloride
2.4g Monopatassium phosphate
Make up to 1L using dH_2O before adjusting to pH 6.9
Autoclave before use
For 1X PBS (1L): 100ml 10xPBS with 900ml dH_2O

Tris-buffered Saline (TBS) (10x)
Makes 1L
60.6g Tris
87.6g NaCl
1M HCl
Make up to 1L using dH_2O before adjusting to pH 6.9
Autoclave before use
For 1X TBS (1L): 100ml 10xTBS with 900ml dH_2O

TBS/TWEEN
Makes 500ml
500ml 1X TBS
500µl Tween 20

PBS/TWEEN
Makes 500ml
500ml 1X PBS
500µl Tween 20

Blocking buffers
p38 antibody:
5% Milk in PBS/Tween:
Makes 50ml
2.5g Milk powder
Make up t0 50ml with PBS/Tween

Phospho-Smad3 antibody:
5% Milk in TBS/Tween:
Makes 50ml
2.5g Milk powder
Make up t0 50ml with TBS/Tween

TGF-β1 antibody:
5% Bovine Serum Albumin (BSA) in TBS/Tween
2.5g BSA powder
Make up to 50ml with TBS/Tween

Radioimmunopercipitation assay (RIPA) buffer
150mM NaCl
1% Triton 100X
0.1% SDS
10mM Tris pH 7.5
1% Deoxycholate
Filter sterilize before use

Tris/EDTA
Makes 2L
0.74g EDTA
2.42g Tris-base
Dissolve and make up to 2L with dH_2O before adjusting to pH 9.0