DNA Methylation in Polycystic Kidney Disease
and Bio-distribution of the C3-G12 peptide

By Michael Bates

A report submitted in fulfilment of the Master Degree in Science

Genetics, University of Otago, Dunedin, New Zealand

Submitted February, 2016
Abstract

Autosomal dominant polycystic kidney disease (ADPKD) is the most common inherited form of renal disease, and is characterised by the formation of fluid-filled cysts within organs, particularly in the kidneys. The two genes that are mutated in ADPKD, PKD1 and PKD2, encode proteins in a plasma membrane complex that facilitates calcium ion release in the cell. Patients typically have an inherited mutation in one of these genes from birth and it is postulated that a second hit in the same gene is required for pathogenesis. Changes in downstream signalling pathways then give rise to the cysts, physically limiting the function of the kidney, and without intervention can lead to renal failure. Mutations in the second allele have been found to be responsible for the second hit in some but not all cysts. We postulate that this “second hit” could also be caused by epigenetic mechanisms, particularly involving changes in DNA methylation. Epigenetic modifying drugs have been shown to reduce cyst formation and limit the pathology of the disease. The use of a targeted delivery system conjugated to an epigenetic modifying drug could ensure the drug concentration is sufficient within the renal tissue, with minimal exposure elsewhere in the body, reducing side effects. Conjugates such as this could initially be most effectively be tested in a small animal model, such as in mice.

To facilitate these types of studies, we have investigated DNA methylation in two regions of the mouse Pkd1 gene. Bisulfite sequencing of kidney DNA samples from wildtype mice, and the del34 Pkd1 mouse model heterozygous for a Pkd1 mutation, were performed. We found that the promotor CpG island was unmethylated whereas the gene body region examined was hypermethylated in both Pkd1 heterozygous and wildtype mouse kidneys. However, quantitative PCR expression analysis revealed consistently higher expression of the Pkd1
gene in Pkd1 heterozygous samples when compared to wildtype. Hematoxylin and eosin staining of mouse kidneys sections revealed at least one confirmed cyst in a Pkd1 heterozygous sample, but cysts are considerably fewer in number when compared with the human disease.

To extend the limited data available on the methylation landscape in human PKD1, we have performed reduced representation bisulfite sequencing (RRBS) on two ADPKD human kidney samples, and three wildtype kidney samples. In contrast to the mouse data, we found methylation differences in the PKD1 gene in ADPKD samples, when compared to wildtype kidney tissue, and changes in other key genes involved in cystogenesis. These data further support a potential role for DNA methylation in ADPKD pathogenesis.

To be able to target a potential methylation difference specifically in the kidney, as a therapeutic option for ADPKD, we have investigated the ability of the galectin-3-avid (C3-G12) binding peptide to specifically localise to the kidney. We injected the peptide bound to a Cy3.5 fluorescent probe into wildtype mice. We observed that the C3-G12-Cy3.5 conjugate accumulated in the mouse kidneys and was retained there for at least 30 minutes. In contrast, little to no fluorescence was observed in the heart, liver, lungs and spleen. These data suggest that the C3-G12 peptide could be used in a selective delivery mechanism for future therapies. This research provides new data on the mouse and human methylation landscapes in ADPKD and confirms the kidney targeting nature of the C3-G12 peptide in mice.
Abbreviations

ADPKD  Autosomal Dominant Polycystic Kidney Disease
ARPKD  Autosomal Recessive Polycystic Kidney Disease
AMP    Ampicillin
DNMT  DNA methyltransferase
E.coli  *Escherichia coli*
FDA    USA Food and Drug Administration
LB     Lysogeny/Luria Broth (both terms acceptable)
LBAA LB + ampicillin + agar
HAT    Histone acetyltransferase
HDAC   Histone deacetylase
*HPRT* Hypoxanthine-guanine phosphoribosyltransferase
*GAPDH* glyceraldehyde-3-phosphate dehydrogenase
PC-1   Polycystin-1
PC-2   Polycystin-2
PCR    Polymerase chain reaction
PKD    Polycystic Kidney Disease
RRBS   Reduced representation bisulfite sequencing
SOC    Super optimal broth containing chloride
TAE    Tris-acetate buffer
TSA    Trichostatin A
QPCR   Quantitative polymerase chain reaction
VPA    Valproic acid
Tables and Figures

Figure 1.1 Primary cilia in the kidney respond to changes in renal flow and mediate downstream pathways .......................................................... 2

Figure 1.2 The polycystin-1 and -2 proteins exhibit several different isoforms on the membrane and within the renal tubule epithelia ......................................................... 5

Figure 1.3 The currently understood pathway to cystogenesis in ADPKD .............................. 8

Figure 1.4 Bisulfite sequencing reveals all methylated CpGs on the DNA after PCR and sequencing ................................................................. 17

Figure 1.5 Macromolecule-drug conjugates (such as C3-G12) are filtered by the kidney and are absorbed by the targeted cells ............................................................. 21

Figure 2.1 Diagram of the gel extraction method used for RRBS ........................................... 37

Figure 3.1 Gel electrophoresis of bisulfite treated mouse Pkd1 PCR products .................. 44

Figure 3.2 Methylograms illustrating the methylation of a 232 bp region of the mouse Pkd1 gene body in exon 43 .............................................................. 46

Figure 3.3 Methylograms illustrating the methylation of a 292 bp region of the mouse Pkd1 gene promotor ............................................................................ 48

Figure 3.4 Standard curves for qPCR of mouse Pkd1, GAPDH and HPRT ........................ 51

Figure 3.5 Consistent expression of the GAPDH and HPRT housekeeping genes .......... 52

Figure 3.6 Relative quantity of Pkd1 expression in 13 mouse samples ............................ 54

Figure 3.7 Genotyping PCR of sample 1314 ................................................................. 55

Figure 3.8 Lollipop comparison plot of mouse Pkd1 promotor and gene body methylation 56

Figure 3.9 Average Pkd1 expression of wildtype and heterozygous mouse samples ........ 57

Figure 3.10 H&E staining of mouse kidney tissue used in analysis .................................... 60

Figure 3.11 A comparison of the PKD1 and Pkd1 gene landscapes between human and mouse ......................................................................................... 62

Figure 5.1 First size selection gel extraction accuracy check ............................................. 69

Figure 5.2 Semi-quantitative PCR of RRBS libraries ......................................................... 71

Figure 5.3 Second size selection gel extraction accuracy check ....................................... 72

Figure 5.4 The base quality control data of a sequenced RRBS library (Sample 5) .......... 74

Figure 5.5 The sequence quality control data of a sequenced RRBS library (Sample 5) .... 75

Figure 5.6 The methylation landscape of the human PKD1 gene ................................. 76

Figure 7.1 Cy3.5 free dye quantity versus fluorescence standard curve .......................... 88

Figure 7.2 Fluorescent images of mouse kidneys after tail-vein injection with Cy3.5 (1 nmol) and C3-G12-Cy3.5 (10 nmol) ........................................................................... 89
Chapter 1: Introduction

1.1 Ciliopathies

Primary cilia are immotile extracellular appendages found on several different cell types throughout the vertebrate body. Primary cilia contain microtubules arranged in a 9+0 pattern (axoneme), where all nine microtubule pairs are equally spaced around the outer edge of the cilia. They do not possess the central pair of tubules (9+2) found on motile cilia\textsuperscript{1} and therefore are not able to facilitate movement of fluid and liquids on their own (unlike motile cilia that are able to generate movement of fluid). In kidneys, primary cilia are found on most cells of the kidney including the epithelial cells of the renal tubules, Bowman’s capsule and collecting ducts\textsuperscript{2} and can be up to three microns in length in both human and rodents\textsuperscript{3}. Renal cilia function as a method of detecting changes in the extracellular environment such as the flow rate of renal fluid, and as a result mediate activation of signalling pathways\textsuperscript{4}. 


Disruptions in the cilia structure, function or related complexes and gene pathways can result in a broad range of renal diseases known as ciliopathies, often with overlapping phenotypes\textsuperscript{1}.

Examples of ciliopathies include syndromes such as Bardet-Biedl, Joubert and Meckel-Gruber, which have a variety of symptoms affecting various organs in the body including: retinal degradation; laterality defects; intellectual disability; cerebellar vermis hypoplasia; encephalocele; polydactyly; obesity; bone deformation and ectodermal dysplasia\textsuperscript{6}.

Hepatobiliary disease and cystic kidneys are found in most ciliopathies and are the primary symptoms in autosomal polycystic kidney disease.

Autosomal dominant polycystic kidney disease (ADPKD) and autosomal recessive polycystic kidney disease (ARPKD) are differentiated by their mode of inheritance and the genes affected, although diagnosis and separation of these two diseases by examination of clinical
symptoms is difficult, particularly for neonatal patients\textsuperscript{7}. Family history is often combined with histological, ultrasound and more recently MRI and genetic testing to provide a more accurate diagnosis\textsuperscript{8,9}.

1.2 Autosomal Dominant Polycystic Kidney Disease

Autosomal dominant polycystic kidney disease (ADPKD) is one of many cystic kidney ciliopathic diseases characterised by the formation of fluid-filled cysts within the kidney nephron\textsuperscript{10}. ADPKD is the most common form of inherited kidney disease, ranging from 1 in 400 to 1 in 1000 individuals affected worldwide\textsuperscript{11}. Cysts continue to grow over the life of the patient with little to no symptoms, until renal function becomes impeded in the fourth and fifth decade of life, with no treatment available\textsuperscript{12}. Factors including genotype, age, sex, kidney function and kidney volume can be used to predict disease progression\textsuperscript{9}.

Other clinical features include cystic livers, hypertension and an increased risk of cardiac diseases and intracranial aneurysms\textsuperscript{13}.

In 85\% of ADPKD cases the gene \textit{PKD1} is mutated, whereas the remaining 15\% of cases are caused by \textit{PKD2} mutations\textsuperscript{14}. Typically patients inherit a single copy of one of the mutated genes\textsuperscript{15}. Due to the high costs of \textit{PKD1} analysis, genetic testing is generally restricted to neonatal diagnosis\textsuperscript{16}, kidney transplant assessment\textsuperscript{17} and cases with little family history available\textsuperscript{9}. Patients with \textit{PKD2} mutations have been shown to have the longest predicted life expectancy\textsuperscript{18}, with \textit{PKD1} non-truncating and truncating mutations having intermediate and the worst prognosis respectively\textsuperscript{11}.

Previous publications had theorised about a possible \textit{PKD3} gene. Where five published families with ADPKD did not show any \textit{PKD1/PKD2} linkage\textsuperscript{19,20,21,22,23}, re-analysis of these
cases revealed misdiagnosis and sample contaminations, with little evidence to support the existence of $PKD3^{24}$.

$PKD1$ is located at chromosome 16p13.3 and contains 46 exons spanning a 12,909 base pair (bp) coding region$^{25}$. Exons 1-33 have been duplicated and have ~98% homology with six pseudogenes located at 16p13.1$^{26}$. $PKD2$ is located at 4q21 and contains 15 exons over a 2904 bp coding sequence$^{27}$.

1.3 Polycystin-1 and -2

$PKD1$ and $PKD2$ encode the polycystin-1 and -2 proteins respectively. Polycystin-1 (PC-1) has 4302 amino acids (aa) consisting of an extracellular N-terminal of ~3000 aa, 11 transmembrane domains and an intracellular C-terminal tail of 198 aa$^{28}$. Polycystin-2 (PC-2) consists of 968 aa, forming six transmembrane domains and intracellular N and C terminal domains$^{27}$. The PC-2 protein acts as a nonselective $Ca^{2+}$ cation channel on the plasma and/or the endoplasmic reticulum membrane.

Within the kidneys, the PC-1 and PC-2 C-terminal domains can interact to form a complex on the renal tubule epithelial cell cilia membrane$^{29}$ (Figure 1.1), where PC-1 acts as a mechanosensor that responds to renal flow and subsequently regulates $Ca^{2+}$ ion flow via PC-2$^{5}$. Additionally, $Ca^{2+}$ is theorised to be released from inside the cell by an interaction between the PC-1 C-terminus and PC-2 proteins found on the endoplasmic reticulum$^{30}$. Increased levels of intracellular $Ca^{2+}$ then activate downstream pathways that maintain the epithelial phenotype$^{31}$.

Polycystin-1 contains an auto-proteolytic site that is able to cleave off the extracellular N-terminal domain (Figure 1.2), allowing it to either be released or tethered to the remaining protein$^{15}$. Both cleaved and uncleaved proteins are required for full function of PC-1$^{32}$. The
C-terminus of PC-1 is responsible for interactions with G-proteins and PC-2 during Ca\(^{2+}\) signalling\(^{33}\). The C-terminus of PC-1 is also able to be cleaved into two distinct products (Figure 1.2) which both interact with transcription factors in the nucleus\(^{34,35}\).

**Figure 1.2:** The polycystin-1 and -2 proteins exhibit several different isoforms on the membrane and within the renal tubule epithelia. **A** I The entire uncleaved membrane-bound PC-1 protein. II The cleaved N-terminus fragment (NTF), which is then able to anchor to the remaining membrane bound section of the protein (III). IV and V The cleaved C-terminus fragment (CTF) and C-terminus tail (CTT) of PC-1. Their specific role in pathogenesis is not yet well understood. **B** The PC-1 and PC-2 complex found on the renal epithelial cell cilia. The N-terminus of PC-1 acts as a mechanosensor for renal flow, which then activates the opening of the PC-2 Ca\(^{2+}\) ion channel by the coupling of the C-terminus. **C** PC-1 is also able to couple to PC-2 found on the endoplasmic reticulum to allow further Ca\(^{2+}\) into the cytosol. Figure from Boletta (2009)\(^{28}\), reprinted with permission from Pathogenetics (2016).
Several key characteristics of cystogenesis have been noted in both human studies and animal models: thickening of the epithelial cell basement membrane; increased cell death by apoptosis; proliferation of cystic epithelia; and pumping of renal fluid to the cyst\textsuperscript{28}.

The PC-1 and PC-2 proteins are able to regulate and interact with several major gene pathways that are directly involved in the development of these cystic characteristics. PC-1 is able to regulate the mTOR pathway, which is involved in cell proliferation and has been shown to be misregulated in ADPKD\textsuperscript{36}. The PC-1/PC-2 complex has been shown to regulate the cell cycle via the JAK/STAT pathway, where inhibition of STAT6 is able to limit cyst growth and preserve renal function in a PKD mouse model\textsuperscript{37}. PC-1 also interacts with JAK2 which induces the p21 protein, which in turn inhibits the cell division kinase Cdk2 and triggers cell cycle arrest\textsuperscript{38}. Wnt/β-catenin signalling, which is involved in development and epithelial cell differentiation, is believed to be directly influenced by renal tubular lumen flow\textsuperscript{39} and by polycystin-1\textsuperscript{40}. Ca\textsuperscript{2+} signalling has been shown to regulate the cAMP pathways which have roles in cell proliferation, differentiation, transcription and fluid transport\textsuperscript{41}. Mismanagement of major cilia developmental pathways Hedgehog and Notch have both been shown to be involved in cancer research and may play a role in ADPKD\textsuperscript{42}.

1.4 Mechanisms of disease

Individuals with ADPKD have one inherited non-functioning copy of either \textit{PKD1} or \textit{PKD2}\textsuperscript{24}. Mouse models with both inactivated copies of \textit{Pkd1} or \textit{Pkd2} show embryonic lethality\textsuperscript{43}, suggesting that the quantity of PC-1 and PC-2 expressed by a heterozygote is sufficient for development. Consequently, a “genetically recessive” model of cystogenesis has been proposed where the inherited mutation is combined with somatic loss of the second copy of the gene. This hypothesis is supported by the observed small number of nephron cells
that actually become cystic (~1%) and the varying quantity of cysts found between patients\textsuperscript{44}.

Two models have been established to explain pathogenesis of cyst development. The first is known as the “Second-hit hypothesis”, where somatic epithelial cells develop secondary “hits” that eliminate the second functioning copy of \textit{PKD1} or \textit{PKD2}\textsuperscript{45}. In support of this, the \textit{Pkd2}\textsuperscript{95\textsuperscript{25}} mouse model, which contains a knockout of one allele of \textit{PKD2} and an inserted element that generates somatic recombination in the second \textit{PKD2} allele (where the inserted element can somatically produce either a \textit{PKD2} wildtype or knockout second allele), has been shown to correlate somatic loss of \textit{Pkd2} with cyst formation\textsuperscript{46}. Additionally, secondary mutations have been found in some studies\textsuperscript{47,48,49}, but not in all cyst examined. One paper has reported heterozygous mutations found in both \textit{PKD1} and \textit{PKD2}, suggesting a potential for additional “trans-heterozygosity” causation for cystogenesis\textsuperscript{50}.

Another idea put forward is one of haploinsufficiency of \textit{PKD1} and \textit{PKD2}. In this model, if the expression level of either gene falls below a threshold needed to maintain the epithelial phenotype in somatic cells, then cystogenesis occurs in that cell. This model is supported by studies where the lowering of \textit{Pkd1} expression in a mouse model has initiated pathogenesis\textsuperscript{51,52}.

Conversely, studies have shown that some human cystic kidneys show overexpression of \textit{PKD1}\textsuperscript{53}, and induced overexpression of \textit{Pkd1} or \textit{Pkd2} in mice can lead to pathogenesis of \textit{ADPKD}\textsuperscript{54,55}, suggesting an upper threshold for \textit{Pkd1} and \textit{Pkd2} expression, as well as a lower one.
1.5 Cystogenesis

Cystogenesis occurs when the epithelia lining the renal tubules, which contain an inherited mutated \( PKD1 \) or \( PKD2 \) allele, undergoes the aforementioned ‘second hit’ of the second allele or the gene expression falls below a threshold (haploinsufficiency)\(^{45,51} \). Tissue injury and repair processes then alter the structural integrity of the cells, which cannot maintain their correct epithelial phenotype due to the loss of polycystin-1 or -2, and subsequent misregulation of their downstream pathways. The epithelial cells then proliferate clonally\(^{48} \) and activation of the cAMP pathway drives fluid secretion into the cysts\(^{41} \). Increased adhesion to collagens and laminin alter the extracellular matrix\(^{56} \), eventually leading to fibrosis and invasion by inflammatory cells\(^{57} \).

**Figure 1.3:** The currently understood pathway to cystogenesis in ADPKD. Epithelial cells first undergo a second hit or haploinsufficient gene expression for \( PKD1 \) or \( 2 \). Misregulated gene pathways and hyperproliferation then allow the epithelia to lose basal polarity and alter fluid management. This can be accelerated by renal injury. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Nephrology (Happé et al., 2014)\(^{57} \), copyright (2014).
1.6 Animal models for polycystic kidney disease

Animal models provide an excellent foundation to examine the genetic and pathological mechanisms in polycystic kidney diseases. Animal models can be developed in two forms: those where spontaneous disease is discovered in a species and a breeding stock is maintained; and those that were created by genetic manipulation of specific genes.

ADPKD and ARPKD have spontaneously developed in several species including dogs\textsuperscript{58,59}, cats\textsuperscript{60,61}, goats\textsuperscript{62} and sheep\textsuperscript{63}. Several spontaneous mice and rat models have also been identified, with varying phenotypes and renal pathology\textsuperscript{64}. However, few, if any spontaneous models are \textit{PKD1} or \textit{PKD2} mutants. Instead mutations in these animals are often found to be in related genes involved in the development or function of the primary cilia\textsuperscript{64}.

Model organisms created by genetic engineering for research purposes often include zebrafish and rodents, due to the size, cost, existing genome knowledge and ease of use for genetic research. Zebrafish models have been used for a range of studies including knockdowns\textsuperscript{65} and mutational studies\textsuperscript{66} to identify developmental and morphological pathways that \textit{pkd1} or \textit{pkd2} may be involved in.

Engineered rodent models have long been used to study the pathological and genetic basis for diseases including ADPKD. There exists many different rodent models (Table 1.1). In contrast to the dominant human disease, mouse models with heterozygous mutations in \textit{Pkd1} or \textit{Pkd2} develop only small cysts with renal function largely uninhibited, depending on the specific model\textsuperscript{67}. Homozygous mutants are generally embryonically lethal\textsuperscript{57}. Other models exist that more closely match the human phenotype but utilise inducible secondary hits using Cre-Lox technology\textsuperscript{57}. Animal models have also been established to examine \textit{PKD1} or \textit{PKD2} under and overexpression, with renal cysts being induced in both of these conditions\textsuperscript{51,68,69}.
Inducible and conditional mouse models display a disease phenotype more similar to human ADPKD than double knockouts, however these models do not reflect a spontaneous somatic second hit that is proposed as a mechanism for the disease. To investigate the mechanism behind the second hit, researchers must use one or more of the limited number of heterozygous knockout mice available.

**Table 1.1:** A list of published *Pkd1* and *Pkd2* rodent models for polycystic kidney disease

**Knockouts**

<table>
<thead>
<tr>
<th>Pkd1</th>
<th>Pkd1(del34)</th>
<th>Pkd1(del17)</th>
<th>Pkd1(del2-6)</th>
<th>Pkd1(del2-11)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pkd2</td>
<td>Pkd2(del34)</td>
<td>Pkd2(del17)</td>
<td>Pkd2(del2-6)</td>
<td>Pkd2(del2-11)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Conditional knockouts**

<table>
<thead>
<tr>
<th>Pkd1</th>
<th>Pkd1(flox/−)::Ksp-Cre</th>
<th>Pkd1(flox/−)::γGT.Cre</th>
<th>Pkd1(flox/−)::MMTV.Cre</th>
<th>Pkd1(flox/−)::NestinCre</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pkd2</td>
<td>Pkd2(flox/−)::Ksp-Cre</td>
<td>Pkd2(flox/−)::γGT.Cre</td>
<td>Pkd2(flox/−)::MMTV.Cre</td>
<td>Pkd2(flox/−)::NestinCre</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Inducible (conditional) knockouts**

<table>
<thead>
<tr>
<th>Pkd1</th>
<th>Pkd1(flox/−)::R26CreER</th>
<th>Pkd1(flox/−)::tam-KspCad-CreERT2</th>
<th>Pkd1(flox/−)::Mxi-Cre</th>
<th>Pkd1(flox/−)::tam-Cre/Esr1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pkd2</td>
<td>Pkd2(flox/−)::R26CreER</td>
<td>Pkd2(flox/−)::tam-KspCad-CreERT2</td>
<td>Pkd2(flox/−)::Mxi-Cre</td>
<td>Pkd2(flox/−)::tam-Cre/Esr1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Hypomorphic**

<table>
<thead>
<tr>
<th>Reduced expression</th>
<th>Pkd1(L3)</th>
<th>Pkd1(V)</th>
<th>Pkd1(mBei)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pkd1(L3)</td>
<td></td>
<td>Pkd1(V)</td>
<td></td>
</tr>
<tr>
<td>Pkd1(mBei)</td>
<td></td>
<td></td>
<td>Pkd1(mBei)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unstable expression</th>
<th>Pkd2(W525)</th>
<th>Pkd2(W525/W5186)</th>
<th>Pkd2(V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pkd2(W525)</td>
<td></td>
<td>Pkd2(W525/W5186)</td>
<td>Pkd2(V)</td>
</tr>
<tr>
<td>Pkd2(V)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Missense mutation</th>
<th>Pkd1(L3)</th>
<th>Pkd1(V)</th>
<th>Pkd1(mBei)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pkd1(L3)</td>
<td></td>
<td>Pkd1(V)</td>
<td></td>
</tr>
<tr>
<td>Pkd1(mBei)</td>
<td></td>
<td></td>
<td>Pkd1(mBei)</td>
</tr>
</tbody>
</table>

**Overexpression of human protein**

<table>
<thead>
<tr>
<th>PKD1</th>
<th>PKD1(extra)</th>
<th>PKD1(AG)</th>
<th>hPKD1 TPK</th>
</tr>
</thead>
<tbody>
<tr>
<td>PKD1</td>
<td>PKD1(extra)</td>
<td>PKD1(AG)</td>
<td>hPKD1 TPK</td>
</tr>
<tr>
<td>PKD2</td>
<td>hPKD2 TG</td>
<td>hPKD2</td>
<td>hPKD2(T)</td>
</tr>
<tr>
<td>PKD2</td>
<td>hPKD2 TG</td>
<td>hPKD2</td>
<td>hPKD2(T)</td>
</tr>
<tr>
<td>PKD2</td>
<td>hPKD2(T)</td>
<td>hPKD2</td>
<td>hPKD2(T)</td>
</tr>
</tbody>
</table>

The del34 mouse model used in this study is highlighted. The table was adapted from (a review of individual models can be found in) Happé and Peters (2014)³⁷.
1.7 Epigenetics and ADPKD

Epigenetics is described as heritable changes to gene expression with no change to the DNA code itself. Epigenetic regulation is now known to be critical for appropriate gene expression in development and major regulatory pathways.

Two key mechanisms in the epigenetic modification of gene expression are DNA methylation, where methyl groups are added to the DNA base cytosine when followed by a guanine (a CpG), and histone post translational modifications. These mechanisms regulate gene expression by altering chromatin conformation, which in turn allows or restricts access to the DNA by transcription factors (reviewed by Liyanage et al., 2013).

1.7.1 Histone modifications

Histones are a series of proteins (H3, H4, H2A, H2B and H1) that form a histone octamer complex that DNA is ‘wrapped’ around to form the nucleosome and chromatin superstructure. Tail-like N-terminal domains extend from the histone proteins and post-translational modification to these tails can alter the physical structure of chromatin, to allow or restrict access to the DNA by transcription factors, or for repair and replication. The most common of the histone modifications are acetylation, methylation and phosphorylation, which are generally associated with open chromatin and gene expression. Deacetylation, demethylation and desphosphorylation are instead associated with closed chromatin and gene repression.

Histone methylation is controlled by histone methyltransferases and demethylases.

Methylation of specific histone lysine and arginine amino acids provide binding access for
histone reader proteins\textsuperscript{75}, which in turn can change the conformation of the chromatin. The specific lysine that is targeted for histone methylation often determines the nature of the chromatin modification. For example, methylation of histone H3 lysine 36 (H3K36) is associated with gene activation, whereas H3K27 methylation is associated with gene silencing. Histone arginine methylation regulates protein-protein interactions\textsuperscript{76}. Together, histone methylation has been shown to be involved in cell fate, X inactivation, transcriptional regulation, RNA metabolism and DNA repair\textsuperscript{77}. In cancer, histone methylation is often dysregulated and is linked to further misregulation of chromatin\textsuperscript{78}.

Phosphorylation of histones H1 and H3 is unique, as while it is associated with gene activation, it has also been shown to play a key role in chromosome condensation during cell mitosis\textsuperscript{79}. Similar studies have suggested that phosphorylation modifications of histones appear to be dynamic in nature and when combined with additional marks such as acetylation, can be involved in both gene expression based chromatin remodelling and chromatin condensation, depending on the specific cell type and regulatory pathway\textsuperscript{80}.

Histone acetylation of lysine residues on the tail domain is regulated by histone acetyltransferases (HATs)\textsuperscript{81}. HATs are divided into Type A and Type B classes based on their localisation in the cell\textsuperscript{82}. Acetylation marks are associated with active DNA and are often found in transcriptionally open promotor regions of the genome\textsuperscript{83}. Removal of acetylation marks from the histone tail by histone deacetylases (HDACs) are instead associated with gene repression. Four classes of HDACs exist, with many implicated in development and major gene pathway regulation\textsuperscript{82}.

HDACs have been shown to be important for \textit{Pkd1}\textsuperscript{84} function as well as being involved in flow-induced calcium signalling\textsuperscript{85}. 

HDACs have been the target of therapy development in recent years through the use of HDAC inhibitors, with a number having been developed and approved by the USA Food and Drug Administration (FDA) for use as anti-cancer agents\textsuperscript{86}.

The effectiveness of HDAC inhibitors on ADPKD models has been examined in several studies, where HDAC inhibitors valproic acid (VPA) and trichostatin A (TSA) have been shown to reduce renal function loss and cyst development in mouse models\textsuperscript{12,85}. TSA acts by inhibiting HDAC5, which is regulated by renal fluid flow and Ca\textsuperscript{2+} signalling pathways and activates gene expression of myocyte enhancer factor 2C (MEF2C)\textsuperscript{85}, which is involved in myocardium hypertrophy. Injection with TSA partially restored epithelial phenotype in cyst lining cells\textsuperscript{85}. An additional study by Fan \textit{et al.}, (2012)\textsuperscript{87} identified that TSA downregulates epithelial cell proliferation and is able to reduce cystogenesis in mouse embryonic kidney cells. VPA functions by inhibiting class 1 HDACs, and was identified in a chemical modifier screen in zebrafish, where VPA partially restored the body curvature phenotype seen in \textit{pkd2} mutants and reduced cyst formation in both zebrafish \textit{pkd2} mutants and \textit{Pkd1}\textsuperscript{lox/lox};\textit{Pkhd1-}\textsuperscript{Cre*} conditional knockout mice\textsuperscript{12}.

1.7.2 DNA methylation

DNA methylation occurs at ~70-80\% of CpGs across the genome, with concentrations of these (known as CpG islands) located within gene promoters, transcription start sites and enhancer regions\textsuperscript{71}. Cytosines are methylated by DNA methyltransferases (DNMTs) during development and after DNA replication\textsuperscript{88}.

Methylation of DNA is historically known to function as a targeted suppression of specific genes and pathways, including X inactivation\textsuperscript{89}, imprinted genes during development\textsuperscript{90} and transposon regions\textsuperscript{91}. Gene promoter methylation is associated with gene repression and

* \textit{Pkhd1} is the gene mutated in human Autosomal Recessive PKD
restriction of transcription factors to the DNA. Gene body methylation, in contrast, has more recently been associated with an increase in gene expression, by potentially suppressing unwanted transcriptional noise and regulating enhancers and silencers. Increased gene body methylation has also been associated with an accumulation of DNA mutations.

Misregulation of DNA methylation may contribute to the pathogenesis of many diseases, and has been a popular topic in cancer research in the last decade, as reviewed by Fukushige and Horii. Many tumour suppression genes that regulate key cellular pathways such as cell cycle control, DNA repair and apoptosis have been identified as being suppressed by DNA methylation in cancers. Inversely global hypomethylation and the resulting inappropriate activation of oncogenes has also been observed in many cancers.

The reversible nature of DNA methylation has also been a target for treatment development. Synthetic nucleotides, 5-Aza-2deoxycytidine and 5-Azacytidine replace cytosine in the DNA and block DNMT activity, globally reducing the quantity of methylation after each consecutive cell division. Treatment of tumours with these two drugs has shown to be effective and promote long-term stability of transcriptionally open promoters.

1.8 DNA Methylation and ADPKD

While literature has explored the role of histone modifications and the greater field of epigenetics in ADPKD, almost no research exists on DNA methylation.

A paper by Woo et al., (2014) used bisulfite treatment combined with global sequencing and expression analysis on human ADPKD samples to determine if there are any correlations between DNA methylation and gene expression. They found that 91% of all methylation
changes in the ADPKD samples, when compared to wildtype, appeared to be hypermethylation, with 61% associated with genic regions. Expression analysis revealed a correlation between \textit{PKD1} gene body hypermethylation and a decrease in \textit{PKD1} expression. Surprisingly, promotor regions were reported between 0-1% of all intergenic methylation changes. This is unusual as historically promotor hypermethylation, rather than the gene body, is associated with a decrease in gene expression\textsuperscript{93}, though this has not been seen in all cell types\textsuperscript{105}. Other studies have reported a correlation between lower gene expression and an increase in 3’ gene body methylation\textsuperscript{106}. Keys genes that were found to be hypermethylated in the ADPKD samples were involved in calcium signalling, ion transport, Notch signalling, and the \textit{PKD1} gene itself, particularly in the gene body region. Madin-Darby canine kidney (MDCK) cell lines, which are used as a model of \textit{in vitro} cyst formation as they proliferate and form cysts when suspended in a collagen gel\textsuperscript{107}, were treated with 5-Aza-2deoxycytidine and the DNMT inhibitor zebularine. Cyst formation and \textit{Pkd1} methylation was reduced and this was correlated to an increase in \textit{Pkd1} expression.

A more recent paper by the same author\textsuperscript{108}, used the same samples and similar techniques to show that the mucin-like protocadherin (\textit{MUPCDH}) gene is hypermethylated in its promotor region, particularly in cyst-lining cells. Increased cell proliferation was correlated with the decrease in \textit{MUPCDH} expression.

Methylation patterns have been shown to change with age in certain regions of the genome\textsuperscript{109}. This could possibly contribute to the tendency for the development of ADPKD pathogenesis in patients in their later decades of life.
1.9 Bisulfite Sequencing

A variety of techniques have been developed to analyse DNA methylation, with the aid of next-generation sequencing, including: methylation-sensitive restriction enzyme digest, methylated DNA immunoprecipitation and bisulfite sequencing, amongst others.

Methylation-sensitive restriction enzyme digestion is one technique, where specific restriction enzymes are used that are able to cleave the DNA at unmethylated cytosines. Methylated cytosines are therefore preserved in the sample and are able to be amplified when followed by PCR and other analyses\textsuperscript{110}.

Another technique involves methylated DNA immunoprecipitation, which utilises antibodies specific for methylated cytosines. It can be combined with microarray or next-generation sequencing for chromosomal or genome wide analysis\textsuperscript{111}.

Bisulfite sequencing is a technique used to identify methylated cytosines within a section of DNA (Figure 1.4). The DNA is treated with sodium bisulfite, which converts all cytosines into uracil unless they are protected by an attached methyl group. A PCR reaction then results in conversion of the uracils to thymine. Any cytosines present in the sequenced DNA, are ones that were originally methylated\textsuperscript{112}. Because of the risk of PCR bias introduced by the potential for methylated cytosines to alter primer binding, primers must be carefully designed to avoid any CpGs within their binding sequence or be redundant so that both methylated and unmethylated cytosines will be recognised.
Figure 1.4: Bisulfite sequencing reveals all methylated CpGs on the DNA after PCR and sequencing.

Methylated cytosines (Red) are protected from the bisulfite reaction which renders all unmethylated cytosines into uracil. Uracil is then converted into thymine after PCR. All methylated cytosines in the initial sequence are preserved as cytosines after sequencing\textsuperscript{112}. *Pkd1* promotor sequence obtained from Ensembl browser 83\textsuperscript{113}.

Bisulfite sequencing has also been adapted using next-generation sequencing techniques to examine larger spans of DNA at much greater depth\textsuperscript{114}. Whole genome bisulfite sequencing (WGBS-seq) has been used to examine entire methylomes of several cell human cell lines\textsuperscript{115,116,117}, cancer samples\textsuperscript{118,119,120} and methylomes of other species\textsuperscript{121,122}.

Due to the high cost of WGBS-seq, several other techniques have been developed for more specific applications, as reviewed by Lee *et al.*, (2013)\textsuperscript{123}.

One such technique is reduced representation bisulfite sequencing (RRBS), where a representative selection of the genome CpG islands are treated and sequenced by use of methylation-insensitive restriction enzymes, such as MspI\textsuperscript{124}. Size selection of the DNA library can then eliminate small fragments, which can be redundant and difficult to align, and exclude large fragments which are often CpG poor\textsuperscript{125}. This leaves a representative sample of
the entire genomic methylation landscape that is enriched for CpG islands. The sizes selected depend on the requirements of the specific study and comes with a trade-off between depth and coverage\textsuperscript{125}.

1.10 Developing a treatment for ADPKD

There exist no treatments for ADPKD on the market. Patients at later stages of the disease generally undergo dialysis and eventually renal transplantation. There exists potential for drug therapies that target specific gene pathways, such as STAT6\textsuperscript{37} or more generalised treatments that target aspects of the disease progression. For example a therapy could be used to reduce epithelial proliferation to reduce cyst growth\textsuperscript{12} or rescue renal function long enough to allow the patient sufficient quality of life for the remaining decades of their lives.

Human clinical trials have been performed in the past\textsuperscript{126} and several are currently in progress in several countries\textsuperscript{127}. However, effective treatment of kidney disease often require high intra-renal drug concentrations which can also expose the patient to significant side effects\textsuperscript{128}. A way to avoid this is a targeted delivery system that is able to deliver the drug directly to where it is needed within the kidneys. There are many different methods developed to achieve this, with the majority of research utilising mouse models to investigate efficacy. Mice have similar physiology to humans and many rodent based imaging and detection methods for bio-localisation of the carrier have been well established in the literature.
1.11 Methods of drug delivery

Use of targeted delivery mechanisms are not limited to any one organ and have seen wide use in a variety of diseases such as cancer therapies129-131, Alzheimer’s disease132, inflammatory bowel disease133 and many more. A variety of different methods have been developed for targeted delivery of drugs to the kidney for use in several renal diseases.

Prodrugs, which are metabolised and activated within the targeted region of the body, have been developed to target certain reactions in the kidney, such as protein-sugar interactions134, 135,136 amino acid reactions137 and folate filtering131.

Nanoparticles are able to have their size, charge, shape and density designed specifically to facilitate an exact purpose within the body128. They have been shown to accumulate in specific areas of the kidney, including the rat glomerular mesangial cells138 and the mouse mesangium139.

Liposomes, which can be endocytosed into target cells, have been found to be particularly effective in in cancer therapy140,141 with some liposome-containing therapies already existing in the market128.

Macromolecular carriers represent the broadest grouping of drug delivery systems, classified by their ability to bind to a drug and deliver it to the target region, with no immunogenic effect (Figure 1.5)128. Lysozymes, a low-weight endogenous protein, are particularly popular as they offer specific kidney uptake and are easy to conjugate to drugs. Lysozymes have been successfully used several times to deliver therapies, with reports showing up to 70x greater drug concentrations in targeted cells than the drug itself542,143,144,145.

Low molecular weight chitosan (LMWC) and polyvinylpyrrolidone (PVP) have both been shown to have varying renal targeting ability based on their molecular weights146,147.
Several peptide delivery molecules have been developed including TAT\textsuperscript{148} and Pep-1\textsuperscript{149}, which have both recently been used in the development of target cancer therapies\textsuperscript{150,151}. This is due to the peptides inherent non-toxicity in the body to efficiently deliver the conjugated drug with increased effectiveness over a global treatment approach, although the relatively short half-life of peptides decreases the longevity of the treatment and may require more frequent treatments compared to other delivery molecules\textsuperscript{152}.

The galectin-3-avid peptide (C3-G12) has been shown to specifically accumulate in mouse kidneys after IV injection\textsuperscript{153} and has recently been shown by Geng et al., (2012)\textsuperscript{154} to effectively deliver the drug captopril within the kidneys. The group used the green fluorescent probe FITC bound to the C3-G12 peptide and showed accumulation of the C3-C12-FITC conjugate in mouse kidneys. They also conjugated the chronic kidney disease (CKD) drug captopril (CAP) and showed that the C3-G12-CAP conjugate accumulated within the kidney significantly more than CAP alone.

The C3-G12 peptide has also been used in the target delivery of cancer therapies\textsuperscript{129}. 


1.12 Hypotheses

Approximately 85% of ADPKD patients inherit one non-functioning copy of *PKD1* from their parents. A prevailing hypothesis suggests that the commencement of cyst formation involves the loss of function of the remaining functional *PKD1* allele by a “second hit” mechanism in somatic cells in the kidney nephron. The *PKD1* gene lies in a CpG rich region with many CpG islands located in the gene itself. We hypothesize that one mechanism of generating a “second hit” could be DNA methylation silencing of the second *PKD1* allele. If this hypothesis is accurate then epigenetic altering therapy could be used to target any methylation irregularities found in the cystic cells. This could be tested in a mouse by targeted delivery of methylation altering molecules and assessing the effect on renal cyst formation. We hypothesise that the mouse *Pkd1* methylation landscape may be similar to the
human PKD1 gene and that a methylation silencing second hit mechanism may also be observed in a mouse model for ADPKD.

In order to test this hypotheses, we will investigate DNA methylation in the mouse Pkd1 gene in both wildtype and Pkd1 heterozygous mice and compare against what has already been published in human studies.

The C3-G12 peptide is one of many targeted delivery molecules currently being researched with the intention of efficiently delivering a therapy to the diseased tissue, minimising side-effects that may follow a whole body treatment. The C3-G12 peptide has been shown to accumulate in the kidneys of mice. To investigate the kidney specificity of a targeted peptide for the purpose of future ADPKD therapy delivery in our hands, we will examine the bio-distribution of the C3-G12 peptide conjugated to a fluorescent marker Cy3.5.

1.13 Previous work

An investigation of the DNA methylation status in the Pkd1 promotor of wildtype mice and the Pkd1 mutant mouse model (del34) was performed in my Postgraduate Diploma of Science.

A 292 bp region in the promotor of the mouse Pkd1 gene was bisulfite treated and cloned from kidneys samples derived from five del34 mutant heterozygotes and five wildtype mice of ages 13-19 months (details in Appendix 4). Methylated and unmethylated cytosines were quantified and compared, showing that the region was completely unmethylated with no evidence of any difference between the del34 heterozygote and wildtype Pkd1 promotor region. This suggests that the mouse Pkd1 promotor is active in the adult kidney and that the Pkd1 gene may be able to be transcribed.
A paper by Woo et al., (2014)\textsuperscript{104} compared the DNA methylation of ADPKD and wildtype tissues at two specific regions of the human \textit{PKD1} gene. The equivalent promotor region in the \textit{del34} mouse model was investigated in my Post Graduate Diploma and the equivalent mouse gene body region will be investigated in this study, as Woo \textit{et al.}, identified a methylation difference in this region. The \textit{del34} mouse will be used to identify the methylation differences in a mouse model of ADPKD that could later be targeted for the development of an ADPKD therapy.

\textbf{1.14 Aims}

1. To compare DNA methylation in the mouse \textit{Pkd1} gene body region in both \textit{Pkd1} heterozygous and wildtype mice, to determine if any methylation differences exist between the two that may support an epigenetic “second hit” in ADPKD pathogenesis.

2. To use reduced representation bisulfite sequencing (RRBS) to extend the limited human data available and get a detailed analysis on the human \textit{PKD1} methylation landscape and compare against published data by Woo \textit{et al.}, (2014)\textsuperscript{104}.

3. To test the ability for a proposed targeted delivery peptide C3-G12 to localise to the kidney in Balb/c mice, by using a conjugated Cy3.5 fluorescent tag. Fluorescent imaging techniques will be used to visualise specificity to the kidney.
Chapter 2: Methods

2.1 Sources of tissues and DNA

Previous researchers extracted DNA from the kidney tissue of ten del34\textsuperscript{156} mice (Qiagen QIAamp DNA Mini Kit, ~10 µg yield for each sample). The del34 polycystic kidney disease mouse model was generated using site targeted mutagenesis methods, which mutated one copy of the \textit{Pkd1} gene. The mice develop a small number of cysts in later life that do not inhibit kidney function\textsuperscript{156}. Both wildtype and heterozygous mice were used in the analysis.

Embryonic mouse samples were created by crossing heterozygous del34 mice to create homozygous and heterozygous embryos which were isolated and genotyped by past researchers at embryonic day 18.5 and stored in -80 °C.

All solutions used in this study can be found in Appendix 1.

2.2 Analysis of mouse \textit{Pkd1} gene body methylation

2.2.1 DH5α competent bacteria creation

DH5α competent \textit{Escherichia coli} (E.coli) bacteria were generated using the Zymo “Mix & Go \textit{E. coli} transformation kit & buffer set – Cat. T3001” and snap frozen in a dry ice/ethanol bath. They were then stored at -80 °C under the approval number GMO05/UO15 as designated by the Ministry of Primary Industries (MPI).
2.2.2 Bisulfite treatment of mouse DNA

Samples were bisulfite treated using the Zymo “EZ DNA Methylation-Gold – Cat. D5005” kit on a BioRad DNA Engine Peltier Thermal Cycler with no changes to the provided protocol (bisulfite incubation at 64 °C for 2.5 hours). Samples were eluted in 10 µl with the supplied elution buffer and stored at -20 °C.

2.2.3 Gene body amplification

To compare the mouse *Pkd1* methylation landscape to that of human *PKD1* samples published by Woo *et al.*, (2014)\textsuperscript{104}, the equivalent gene body region (procured from Ensembl\textsuperscript{113} – Appendix 5) in the mouse *Pkd1* gene was amplified by primers designed by the Zymo website “Bisulfite Primer Seeker”\textsuperscript{157}. This gave a fragment 232 bp long in exon 43. This region captured 8 potentially methylated CpG’s but was not defined as a CpG island by the “NCBI CpG island strict prediction algorithm”\textsuperscript{158}.

Exon 43 of the *Pkd1* gene bisulfite treated DNA was then amplified using the following PCR protocol. Primers are listed in Appendix 3.
Table 2.1: Optimised PCR protocol to amplify the bisulfite treated DNA

<table>
<thead>
<tr>
<th>Reagent</th>
<th>each reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>MQH2O</td>
<td>15.20</td>
</tr>
<tr>
<td>10X PCR Buffer</td>
<td>2.00</td>
</tr>
<tr>
<td>50mM MgCl2 (1.0mM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>10mM dNTPs (0.2mM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>10mM PKD1F (0.2mM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>10mM PKD1R (0.2mM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>Platinum Taq polymerase</td>
<td>0.20</td>
</tr>
<tr>
<td>Master Mix volume</td>
<td>19.00</td>
</tr>
<tr>
<td>Sample</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Total PCR volume</strong></td>
<td><strong>20.00</strong></td>
</tr>
</tbody>
</table>

Table 2.2: Thermocycler settings for sufficient amplification of the bisulfite treated DNA

<table>
<thead>
<tr>
<th>Cycle step</th>
<th>Temperature</th>
<th>Time</th>
<th>Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Initial denaturation</td>
<td>95 °C</td>
<td>2 minutes</td>
<td>1</td>
</tr>
<tr>
<td>2 Denaturation</td>
<td>95 °C</td>
<td>30 seconds</td>
<td>40</td>
</tr>
<tr>
<td>3 Annealing</td>
<td>52 °C</td>
<td>30 seconds</td>
<td></td>
</tr>
<tr>
<td>4 Extension</td>
<td>72 °C</td>
<td>30 seconds</td>
<td></td>
</tr>
<tr>
<td>5 Final extension</td>
<td>72 °C</td>
<td>7 minutes</td>
<td>1</td>
</tr>
</tbody>
</table>

Samples were then placed at 4 °C temporarily prior to gel electrophoresis. Bisulfite treated DNA is less stable than untreated DNA, so long term storage was avoided.

2.2.4 Gel electrophoresis of amplified bisulfite samples

To ensure the correct sized product had been amplified, PCR samples were electrophoresed on a 2% agarose gel using the appropriate amounts of 1x Tris-acetate (TAE) buffer and agarose powder. The gel running buffer was made up from 1x TAE and 50 µg/mL ethidium bromide. 1 µl of Thermo Scientific Loading Dye was used for all wells with 5 µl (1 µg/µl) of Invitrogen “1 Kb Plus DNA Ladder” and samples respectively. The gel was run at 100 volts
for 30 minutes and then visualized under UV light on a Bio-Rad XR gel visualizer using the Quantity One 4.6.5 Basic software. The amplified samples were confirmed to be of the correct size before proceeding to any future steps.

The PCR fragments were purified using the Invitrogen “Purelink PCR Purification Kit”. The final elution volume, which was listed in the manufacturer’s protocol as being 50 µl, was changed to 20 µl to ensure a higher concentration of DNA.

A Nanodrop 1000 3.8.1 was used to quantify the DNA using the appropriate elution buffer as a blank.

2.2.5 Transformation of bisulfite treated Pkd1

To create a bacterial library containing unique DNA fragments of the target region, the DNA was transformed into competent bacteria. A modified protocol generating higher transformation efficiency was determined previously in my Post Graduate Diploma. This modified protocol was used to transform the amplified bisulfite converted DNA samples into the DH5α bacteria (Table 2.3).

The Invitrogen “TOPO TA Cloning Kit for Sequencing” was used to create the vector-sample constructs. This particular kit used the “TOPO 4.0 vector” (Appendix 2). The listed volumes in the supplied protocol were halved with no observed detrimental effect on the experiment.

The TOPO 4.0 vector contains thymidine overhangs at each end which easily binds to the adenosine overlay produced by Taq polymerases. Insertion of a PCR product disrupts the lethal lacZα-ccdB genes and allows the bacteria to grow; vector alone will not grow. Combined with an ampicillin selection marker, only cells containing the vector+insert are
capable of growing on ampicillin media, therefore traditional blue/white colony selection was unnecessary.

Cells with no DNA were used as a negative control, whereas cells transformed with the Invitrogen pUC19 vector were used as a positive control. Lysogeny broth (LB) + agarose + ampicillin (LBAA) plates were made with Invitrogen LB Agar powder and 25 µl of 100 mg/ml ampicillin per plate. All steps listed in Table 2.3 were performed on ice.

Table 2.3: The improved method used for the transformation of the bisulfite treated *Pkd1* gene body region into DH5α *E. coli* cells

<table>
<thead>
<tr>
<th>Steps</th>
<th>Cell control</th>
<th>pUC19 control</th>
<th>TOPO reaction/Pkd1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Add 25 µl DH5α cells</td>
<td>Add 25 µl DH5α cells</td>
<td>Add 25 µl DH5α cells</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>Add 1 µl (10 pg) pUC19 vector</td>
<td>Add 2 µl of TOPO reaction (2 µl PCR product + 0.5 µl salt solution + 0.5 µl TOPO vector)</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>Incubate 5 mins on ice</td>
<td>Incubate 5 mins on ice</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>Heatshocked at 42 °C for 30 seconds</td>
<td>Heatshocked at 42 °C for 30 seconds</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>Returned to ice for 2 mins</td>
<td>Returned to ice for 2 mins</td>
</tr>
<tr>
<td>6</td>
<td>250 µl SOC added</td>
<td>250 µl SOC added</td>
<td>250 µl SOC added</td>
</tr>
<tr>
<td>7</td>
<td>1 hour in 37 °C shaking incubator</td>
<td>1 hour in 37 °C shaking incubator</td>
<td>1 hour in 37 °C shaking incubator</td>
</tr>
<tr>
<td>8</td>
<td>Pipette 100 µl onto LB and LBAA plates</td>
<td>Pipette 100 µl onto LBAA plates</td>
<td>Pipette 100 µl onto LBAA plates</td>
</tr>
</tbody>
</table>

Adapted in my previous work in my Post Graduate Diploma.

Transformation mix was spread evenly onto the LBAA plates with a glass pipette and spreader. Plates were left upside down for 18 hours in a 37 °C incubator. The transformation efficiency was calculated by using the Bioruffo “Online Transformation Efficiency Calculator”159.
2.2.6 Overnight culture of selected colonies

Bacteria that had successfully transformed with the vector and insert grew overnight on the LBAA plates. Individual colonies were picked with autoclaved toothpicks before being placed into a 50 ml falcon tube containing 3 ml LB and 3 µl of 100 mg/ml ampicillin. Tubes were then placed into a shaking incubator at 37 °C at 200 RPM for ~16 hours.

2.2.7 Plasmid prep of overnight cultures

To sequence the amplified DNA fragments, the plasmid was isolated from the bacteria using the Zymo “Zyppy Plasmid Miniprep Kit – Cat. D4020”. The manufacturer’s instructions for this kit did not require samples to be centrifuged to concentrate the bacterial colonies prior to lysation of the cells. Due to sporadic insufficient growth of the bacteria overnight and subsequent low plasma DNA yields, this centrifugation step was added to the provided protocol to improve output DNA concentration. Cultures that were less cloudy, indicating less bacterial growth, were spun down using a benchtop centrifuge. The supernatant was removed and the cells resuspended in 600 µl of MilliQ water to increase cell concentration and therefore output DNA concentration. Additionally to avoid carry over of ethanol and other kit residues, a one minute time gap was added between washing steps to allow for evaporation. The eluted DNA concentration was measured on a Nanodrop 1000 3.8.1.

2.2.8 Sequencing of plasmid preps

Isolated DNA was prepared for sequencing using the University of Otago “Genetics Analysis Service” guidelines which required 150-200 ng of plasmid DNA, 3.2 pmol of primer
combined with a total volume of 5 µl. The primer used for sequencing was the TOPO 4.0 M13 Reverse primer (Appendix 3), as supplied by the Invitrogen “TOPO TA Cloning Kit” used in Step 2.2.5. This primer binds outside of the vector DNA insert region, allowing the sequencing of the inserted DNA.

2.2.9 Analysis of methylation data

The returned sequences were visualised and checked for quality in the program Geneious 8.1.7 (Restricted Version). Vector DNA bases were trimmed from each end of the insert sequence to leave the target DNA. Methylograms were produced by using the BiQ-Analyser 2.0\textsuperscript{160} software. The BiQ-Analyser automatically removed poor sequences below an 80% sequence homology threshold, as well as the alignment of remaining sequences and relevant CpG’s against a reference wildtype \textit{Pkd1} sequence. The reference \textit{Pkd1} gene was acquired from Ensembl Genome Browser\textsuperscript{113}.

2.2.10 Statistical analysis

Data was arranged into two groups at each CpG, separating the wildtype and \textit{del34} sequences for both the promotor (data obtained previously in my Post Graduate Diploma) and gene body regions. Each CpG within those regions were then tested using a Mann-Whitney U-test to identify if there was any significant difference in methylation between samples that were wildtype and those that were from \textit{del34} kidneys. This test was chosen because the datasets were not normally distributed, and therefore a t-test was inappropriate.
2.3 Creation of cDNA
To examine the correlation between methylation status and *Pkd1* gene expression, RT-qPCR was performed on mouse RNA samples.

RNA was extracted from 13 frozen mouse kidneys, using the Qiagen “RNeasy Mini Kit”. The protocol was followed as provided by the kit. All ten mouse kidney samples used for the cloning based methylation study were processed, with an additional three embryonic mouse samples (wildtype, heterozygous and homozygous for the *Pkd1* disruption). Extracted RNA from each mouse kidney was then converted to cDNA using the Applied Biosystems “High-Capacity cDNA Reverse Transcription Kit”. The lowest concentration of RNA obtained was 21.6 ng/µl, and as the cDNA reaction requires 10 µl input, 216 ng of RNA was converted to cDNA for each sample. A control sample was also included in each conversion reaction that contained all reverse transcriptase kit reagents minus any RNA.

2.4 RT-qPCR of samples
The relative expression of *Pkd1* in the appropriate sample was measured by performing a RT-qPCR analysis. All cycle threshold (CT) values for RT-qPCR and standard curves were generated on a Roche Light Cycler 480 II using the Takara “SYBR Premix Ex Taq (Tli RNaseH Plus)” enzyme. Hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) were chosen to be suitable housekeeping genes for the mouse samples and had previously been assessed for suitability in cystic samples. Initially, RT-qPCR primers obtained from published literature for the target *Pkd1* gene in mouse, gave two bands. Working primers were designed from the NCBI Primer-BLAST tool\(^{161}\). Primers are listed in Appendix 3.
Primers used for RT-qPCR are required to amplify the cDNA in a linear pattern to be able to directly compare between samples. To determine suitability for RT-qPCR analysis, an amplification efficiency standard curve was performed.

A cDNA sample was diluted seven times in 1:4 increments to give a dilution range that was then run on the Roche Lightcycler 480 II (using supplied software Lightcycler 1.5.1) with target and housekeeping gene primer sets. The protocol supplied by the SYBR kit was modified to use less reagents with no detrimental effects on results. CT values were generated from each well by the “absolute quantification second derivative max” protocol. Standard curves were plotted using GraphPad Prism 6.07 (Demo).

**Table 2.4:** The SYBR based RT-qPCR protocol followed to generate second derivative absolute expression data

<table>
<thead>
<tr>
<th></th>
<th>Per well</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYBR Premix Ex Taq (2x)</td>
<td>5</td>
</tr>
<tr>
<td>F-Primer (5 µM)</td>
<td>0.2</td>
</tr>
<tr>
<td>R-Primer (5 µM)</td>
<td>0.2</td>
</tr>
<tr>
<td>MQ water</td>
<td>3.6</td>
</tr>
<tr>
<td>cDNA</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>10</strong></td>
</tr>
</tbody>
</table>

Following the confirmation that the *Pkd1* primers were suitable, a RT-qPCR was performed with undiluted primers following the protocol listed in Table 2.4.
Each sample was performed in triplicate, for each primer set (a total of nine reactions per sample per plate). All samples were repeated a second time as a technical replicate. Controls were included on each plate, with “no RNA or cDNA” triplicates for each primer and three reverse transcribed controls as noted in Section 2.4.

2.5 Genotyping of qPCR results and re-analysis

Following the RT-qPCR analysis, the wildtype mouse sample 1314 revealed Pkd1 expression levels similar to that of the heterozygote samples (Section 3.2). To further investigate this, a genotyping PCR reaction was performed with one negative and one positive control. Primers listed in Appendix 3.

Table 2.5: Genotyping PCR performed to identify the Pkd1 genotype of the 1314 sample.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>each reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>MQH2O</td>
<td>14.20</td>
</tr>
<tr>
<td>10X PCR Buffer</td>
<td>2.00</td>
</tr>
<tr>
<td>50mM MgCl2 (1.0mM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>10mM dNTPs (0.2mM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>DMSO</td>
<td>1.00</td>
</tr>
<tr>
<td>10uM NeoFII F (0.2uM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>10uM Mr50 R (0.2uM final)</td>
<td>0.40</td>
</tr>
<tr>
<td>Platinum Taq polymerase</td>
<td>0.20</td>
</tr>
<tr>
<td>Master Mix volume</td>
<td>19.00</td>
</tr>
<tr>
<td>Sample</td>
<td>1.00</td>
</tr>
<tr>
<td>Total PCR volume</td>
<td>20.00</td>
</tr>
</tbody>
</table>

The samples were run on a 2% agarose electrophoresis gel following the same protocol listed in section 2.2.4 and viewed on the Bio-Rad XR gel visualizer. The 1314 sample was then reclassified as a heterozygous sample and methylogram and statistical analysis results were reanalysed.
2.6 H&E staining of mouse samples

To analyse the quantity and morphology of the cysts in the mouse samples, the ten mouse kidneys from which the DNA was extracted (Section 2.2) were embedded in paraffin, sectioned and hematoxylin and eosin (H&E) stained by the Otago Histology Services Unit. Sectioned kidneys were imaged under an Olympus DP80 microscope and findings confirmed by a pathologist in the department.

2.7 Human tissue ethics approval

Three wildtype human kidney tissues were obtained from the Christchurch Tissue Bank following ethics approval by the “Otago Human Ethics Committee” H15/110. Tissue was taken during surgical procedures to remove renal tumours. The tissue taken was cortical and sampled as far from tumours as possible. An additional two human ADPKD tissue samples were used from donated tissue as approved by the Otago Human Ethics Committee H15/110.

2.8 Genome-wide methylation analysis

To gain a representative look at DNA methylation across the genome of the human ADPKD derived kidney tissue, reduced representation bisulfite sequencing (RRBS) was used to perform a genome wide methylation analysis on two human ADPKD kidney tissue samples, and three human wildtype tissue samples.

RRBS is performed by digesting the DNA samples with the MspI restriction enzyme, which breaks the DNA into fragments. It is then size selected to remove small, difficult to align fragments, and large fragments which are often CpG poor, to collect a representative sample of the entire genome that is enriched for CpG islands.
The RRBS protocol is one heavily modified by members of our laboratory including Dr. Aniruddha Chatterjee and Dr. Euan Rodger. The entire methodology will not be described. Instead key areas of the process will be highlighted. The full process is detailed in the publication “Technical Considerations for Reduced Representation Bisulfite Sequencing with Multiplexed Libraries”\textsuperscript{163}.

2.8.1 DNA extraction for RRBS

DNA was extracted from the kidney tissues by the Qiagen “QIAamp DNA Mini Kit”, using the “mortar and pestle” method, where frozen sections of tissue were sliced on dry ice, and then ground in lysis buffer with a plastic pestle. Treatment with Proteinase K was performed for 16 hours overnight at 56 °C, as opposed to the recommended 3 hours, due to the requirement for increased DNA yields. The kit protocol recommended 25 mg of tissue with an output of 5-30 µg of DNA. Between 18-25 mg of tissue was used for each sample, resulting in varying DNA yields between 2-15 µg. DNA was checked for quality on a Nanodrop 1000 3.8.1 and quantity was confirmed on a Qubit 2.0 Fluorometer DNA quantification system.

2.8.2 MspI digestion and end repair

Samples then underwent MspI digestion, where the MspI enzyme slices DNA between the two C’s of a CCGG motif and fragments the DNA. Samples were digested overnight (16 hours) at 37 °C on the BioRad DNA Engine Peltier Thermal Cycler to ensure full digestion. The samples were then purified using the Qiagen QiaQuick PCR Purification Kit.

The base overhang left by the MspI enzyme cuts was then repaired by Nano End Repair Mix 2 (from the Illumina TruSeq Nano DNA LT Library Prep Kit). The samples were purified
again using the Qiagen MinElute PCR Purification Kit, to produce more concentrated DNA.

2.8.3 Adenylation and sequencing adapter binding

The 3’ end of the sample fragments were then adenylated prior to sequence adapter binding. Unique DNA sequencing adapters were ligated onto the adenylated 3’ ends of the sample fragments to allow multiplex sequencing. Adapter combinations were checked against published Illumina documentation to ensure no incompatibility issues, where adapters are able to bind to each other and may impede the reaction.

**Table 2.6:** The Illumina DNA sequencing adapters used for each sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Library Adapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADPKD Kidney 1</td>
<td>2</td>
</tr>
<tr>
<td>ADPKD Kidney 2</td>
<td>4</td>
</tr>
<tr>
<td>Wildtype 1</td>
<td>5</td>
</tr>
<tr>
<td>Wildtype 2</td>
<td>6</td>
</tr>
<tr>
<td>Wildtype 3</td>
<td>7</td>
</tr>
</tbody>
</table>

The samples were purified again using the same Qiagen MinElute kit used above.

2.8.4 First size selection

To select the representative subset of CpGs to analyse, the samples were run on a 3% Lonza “NuSieve GTG Agarose Gel” and the gel area between 150 and 325 bp bands (adapter size included) of a 25 bp DNA ladder was isolated for each sample.
It was vital the NuSieve gel was prepared carefully. The mixture was heated at 50% power in a microwave several times until the liquid gel was completely clear. The gel was poured and then left to set at room temperature for one hour. Seven µl (1 µg/µl) of Invitrogen 25 bp DNA ladder was used in the first lane for accurate size selection. Three microlitres of Thermo Scientific Loading Dye was added to each sample (~ 20 µL total) and then the entire sample was run in a lane on the gel with a lane left between each sample to ensure no contamination between each sample. The gel was then loaded and run with fresh 0.5x TAE buffer at 50 V for 90 minutes.

The electrophoresed gel was placed on a flat, sterile tissue culture flat disposable flask. The ladder lane was sliced off with a thoroughly cleaned knife and imaged under UV in the Bio-Rad XR gel visualizer. The 150 and 325 bp marks were measured with a ruler and distance from the bottom of the gel was noted. The ladder containing gel lane was returned to the rest of the gel. Each lane was cut from the gel and separated, thoroughly cleaning the knife after each cut. The measured 150 and 325 bp were aligned with each sample lane and the area between the two bands removed as illustrated by Figure 2.1.
Figure 2.1: Diagram of the gel extraction method used for RRBS. Only the 325 and 150 bp bands of the DNA ladder are shown. The DNA ladder bands were measured under UV light then placed back alongside the gel. Blue dashed lines indicate where the gel was cut with a knife based on the measurements. Blue rectangles indicates the area removed for further processing.

Each sample’s gel fragment was placed into individual 2 mL Eppendorf tubes and weighed. The DNA was then extracted using the Qiagen QIAquick Gel Extraction Kit.

2.8.5 Bisulfite conversion of samples

The size selected DNA was then bisulfite treated using the Zymo EZ DNA Methylation Kit to convert any unmethylated cytosines into uracil. The CT conversion reagent provided by the kit was used fresh due to increased risk of conversion failure by repeated freeze-thawing. The conversion was run for 18 hours at 50 °C on the BioRad DNA Engine Peltier Thermal Cycler, to ensure complete conversion.

2.8.6 PCR amplification

For library preparation, the samples were amplified by PCR. The number of cycles that the samples underwent was important due to the increasing risk of point mutations or fragment biases\textsuperscript{125} with each additional cycle of amplification. Samples were amplified just enough to produce sufficient yields of DNA Library. To determine the correct number of cycles, a small amount of bisulfite converted sample underwent two semi-quantitative PCR reactions for 15 and 20 cycles. The PCR products were then electrophoresed on a NuSieve gel prepared as previously described in the size selection step (4.2.4). The number of cycles required for large-scale amplification of each individual sample was then selected based on the intensity of the amplified product after visualisation of the gel on the Bio-Rad XR gel visualizer (judged by Doctors Chatterjee and Rodger, who designed the protocol).
Each of the samples were amplified in a large-scale PCR reaction, using the amount of cycles (15-20) previously determined in the semi-quantitative PCR. This was unique to each sample so therefore five separate PCR reactions were set up to run concurrently. After amplification, the samples were purified using the Qiagen MinElute Kit used previously.

2.8.7 Second size selection

The amplified samples were run on a NuSieve gel which was prepared and size selected exactly as previously described in the first size selection (Section 2.8.4). This additional size selection step ensured that the final library preps contained no DNA fragments larger or smaller than the selected 150-325 bp. Fragments outside of this region may have been missed during the first round of size selection or introduced during subsequent PCR steps by adapter and primer dimers.

The isolated gel fragments were weighed and then gel extracted using the Qiagen MinElute Gel Extraction Kit eluted in a final volume of 20 µL. One microliter was then used to determine the concentration by Qubit Fluorometer, and another one microliter used for the Bioanalyser analysis carried out by New Zealand Genomics Limited (NZGL).

Once the Bioanalyser data was returned, samples were diluted to 10 nM in 25 µL and submitted to NZGL for sequencing on the Illumina HiSeq 1x100 bp Sr V4 platform.

2.8.8 Quality control

Sequenced libraries were processed by NZGL and assessed for quality control by using the program FastQC\textsuperscript{164} (Version 0.11.3). FastQC generates a series of quality control graphs from raw sequence including sequence quality scores (Figure 5.4) and individual base
probability (Figure 5.5). The Phred quality score\textsuperscript{165,166} is the logarithmic likelihood of an incorrect base call at any given base and is used as a measure for sequencing. This is plotted as a mean, maximum and minimum Phred score at each base of all the reads for a given sample. Good quality sequenced libraries are those that show high Phred score (>30), markers of complete bisulfite sequencing (low cytosine content) and complete MspI digestion (Figure 5.5).

Table 2.7: Phred scores indicate sequencing quality of high-throughput data

<table>
<thead>
<tr>
<th>Phred Quality Score</th>
<th>Accuracy at each individual base</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>90%</td>
</tr>
<tr>
<td>20</td>
<td>99%</td>
</tr>
<tr>
<td>30</td>
<td>99.9%</td>
</tr>
<tr>
<td>40</td>
<td>99.99%</td>
</tr>
<tr>
<td>50</td>
<td>99.999%</td>
</tr>
<tr>
<td>60</td>
<td>99.9999%</td>
</tr>
</tbody>
</table>

2.8.9 Clean-up and alignment

Before alignment and analysis, the sequences must be prepared and ‘cleaned’ of poor quality reads and adapters removed. Due to the high quality of the sequences, the script only needed adapter trimming using the script titled “cleanadapters”\textsuperscript{184}.

The sequences were aligned using the program Bismark\textsuperscript{167} (V 0.14.4) which utilises the Bowtie\textsuperscript{168} fragment aligner. Bismark maps all reads against all four bisulfite conversions, possibly introducing mismapped reads. To correct this, the option “--directional” was enabled.
2.8.10 Visualisation of the sequences

To visualise the methylation landscape of the aligned sequences, the tool Integrative Genomic Viewer\textsuperscript{169} (IGV - V 2.3) was used. IGV supports data files that contain information at each chromosome, and compares them to a reference genome. Chromosomal locations and gene names can be searched to visualise the coverage and depth of reads at any given location on the genome.

2.9 Creation of peptide-fluorescent molecule conjugate

To detect the bio-distribution of the C3-G12 peptide within the mouse organs the C3-G12 peptide was synthesised and conjugated to a Cy3.5 fluorescent marker (GE Healthcare Life Sciences Cy3.5 NHS Ester) by our collaborators (Dr Sarojini, School of Chemical Sciences, The University of Auckland).

Previous papers had used FITC as a marker for this peptide, but we chose to use Cy3.5 as it fluoresced at the red end of the spectrum with an absorbance at 590 nm and emission at 620 nm, where there is little background fluorescence in mouse kidneys. This was important as we had planned to use the FX-Pro \textit{in vivo} small animal imager, which is more sensitive to background fluorescence in the shorter wavelengths.

Both the peptide-dye and free dye were dissolved in DMSO to the concentration of 1 nmol/µl.
2.10 Comparative fluorescence

As preliminary experiments revealed that the C3-G12-Cy3.5 conjugate and the Cy3.5 free dye did not fluoresce with equal intensity at the same concentration, the fluorescence had to be normalised so that the concentration was adjusted in order that equal volumes of solution produced equal levels of fluorescence under the microscope. A standard curve was created with 1000, 800, 600, 400 and 200 pmol of free dye placed in wells of a Biotek Synergy 2 plate reader (using Gen 5 1.11 software) and compared against 1000 and 100 pmol quantities of the peptide-dye conjugate. Additional 200, 150 and 100 pmol dilutions were made and analysed after initial measurements were higher than expected. Multiple replicates were made. GraphPad Prism 6 Demo (V 6.07) was used to generate the standard curves.

2.11 Injection of mice and timepoints

Balb/c (~11 months of age) mice were injected by tail-vein IV injection with 50 µl of one of three different solutions:

- Vehicle - 10 µl DMSO made up to 50 µl with normal saline (NS)
- Cy3.5 - 1 nmol suspended in 10 µl DMSO made up to 50 µl with NS
- C3-G12-Cy3.5 - 10 nmol suspended in 10 µl DMSO made up to 50 µl with NS

Mice were culled at three time points after injection of one of the solutions:

- 3 minutes - three mice for each solution
- 15 minutes - two mice with Cy3.5 and two with C3-G12-Cy3.5
- 30 minutes - two mice with Cy3.5 and two with C3-G12-Cy3.5
2.12 Fluorescence of mice organs

The kidneys, liver, heart, spleen and lung were dissected from the mice and viewed individually under a Leica M205 FA (LAS V4.4 software) using the dsRED filter. Exposure settings were normalised based on background fluorescence of free dye controls, so that the free dye samples were barely visible under imaging.

**Table 2.8:** Microscope settings used for all images of mouse organs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure</td>
<td>32.8 s</td>
</tr>
<tr>
<td>Gain</td>
<td>1.00 x</td>
</tr>
<tr>
<td>Saturation</td>
<td>1.10</td>
</tr>
<tr>
<td>Gamma</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The fluorescence was measured by the ImageJ (V 1.50b) software and quantified as “Mean Gray Value” for each organ.
Chapter 3: Mouse *Pkd1* DNA methylation analysis

3.1 Analysis of mouse *Pkd1* DNA methylation

3.1.1 PCR amplification and yield

DNA from five mice previously genotyped as heterozygous for the *del34 Pkd1* disruption and five wildtype mice of similar age were bisulfite treated, and a 232 bp region of exon 43 in the *Pkd1* gene body was amplified by PCR (Sections 2.2.2 and 2.2.3).

The amplified DNA was electrophoresed on an agarose gel (as described in Section 2.2.4). The size of the fragment was compared against the 1 kb DNA ladder to ensure that the correctly sized target fragment (232 bp long in exon 43 – Appendix 5) had been amplified, and that the PCR had been successful.

![Gel electrophoresis of Pkd1 PCR products of bisulfite treated mouse DNA.](image)

**Figure 3.1:** Gel electrophoresis of *Pkd1* PCR products of bisulfite treated mouse DNA. The visualised gel shows a single band for each sample at the expected size compared to the Invitrogen “1 Kb Plus DNA Ladder”. The band intensity suggested good amplification.
The amplified PCR products only showed one band at the expected size (232 bp), therefore gel extraction techniques were not used and PCR purification using the “Purelink PCR Purification Kit” was sufficient. Eluted DNA had concentrations between 6-10 ng/µl. The PCR fragment was directly sequenced to ensure that it was the correct target region of exon43 in *Pkd1*.

3.1.2 Transformation of *Pkd1* inserts

To create a bacterial library of the bisulfite treated PCR products, the purified fragments were TOPO cloned and transformed into DH5α *E. coli* bacteria using the Thermo Fisher “TOPO 4.0” vector as outlined in Section 2.2.5. For each transformation, two sets of control reactions were used, one containing DH5α only, to confirm bacterial growth on LB (Lysogeny broth) agar and not on LB agar plus ampicillin (LBAA) plates. A second transformation using pUC19 was used as a positive control. The quantity of vector and fragment containing successful transformations varied with an approximate pUC19 transformation efficiency of 1.4 x 10⁷ transformants per µg of DNA with ~50 colonies per plate.

3.1.3 Plasmid preparation and sequencing

Overnight cultures of single bacterial colonies were grown overnight. Plasmid DNA was prepared from these cultures, producing concentrations of 50-150 ng/µl of DNA, utilising the improved methodology described in Section 2.2.5.

Isolated plasmids were sequenced with the M13 reverse sequencing primer (Appendix 3) supplied in the TOPO 4.0 kit. High data quality (80%+ sequence identity with the reference sequence) was required in the various analysis software programmes described in section 2.2.9.
3.1.4 Pkd1 Methylogram data

In order to examine the pattern of methylation for each genotype, sequences were aligned by BiQ software\textsuperscript{160} (Section 2.2.9) and methylated CpGs were counted against a completely unmethylated reference sequence supplied from Ensembl (release 83). "Not present" indicates that the sequence quality was insufficient to properly determine methylation status at that CpG. This occurred more frequently in the first four CpGs as this was often at the 3’ end of the sequence data, and sequencing quality declined steadily from the start of the sequence.

![Methylogram data](image)

Figure 3.2: Methylograms illustrating the methylation of a 232 bp region of the mouse Pkd1 gene body in exon 43. The eight CpGs that are within the cloned region are listed with the corresponding quantity of methylated and unmethylated CpGs at that position for del34 heterozygous samples (A) and wildtype (B). Values for “unmethylated”, “methylated” and “not present” indicate the number of sequences that gave information at that position.
These results show that seven of the eight CpG’s in the 232 bp region of exon 43 of the mouse *Pkd1* gene body are methylated in both heterozygous and wildtype samples with a very small number of non-methylated sequences. One CpG is completely unmethylated in both heterozygous and wildtype samples. Statistical analysis was performed (section 3.4.3) to identify any significant difference between the *Pkd1 del34* heterozygous and wildtype samples at each CpG and showed no significant difference at any CpG.

Additional sequences for a 292 bp fragment in the promoter region (33 CpGs) of the mouse *Pkd1* gene (Appendix 4) were obtained during the course of this study and were added to those obtained during my Post Graduate Diploma research project, using identical methods described in “Chapter 2 - 2.2 Analysis of the mouse *Pkd1* gene”. New methylograms were created using the expanded data (Figure 3.8).
Figure 3.3: Methylograms illustrating the methylation of a 292 bp region of the mouse Pkd1 gene promoter.

The 33 CpGs that are within the cloned region are listed with the corresponding quantity of methylated and unmethylated CpGs at that position for heterozygous samples (A) and wildtype (B). Values for “unmethylated”, “methylated” and “not present” indicate the number of sequences that gave information at that position.
These results showed extensive hypomethylation in 32 of the 33 CpGs in this promoter region of the mouse Pkd1 gene, in both heterozygous and wildtype samples. Statistical analysis was unable to be performed as the majority of CpG’s were exclusively hypomethylated in both datasets and showed no difference.

3.2 Pkd1 expression data

To examine the correlation between the Pkd1 methylation data and Pkd1 gene expression, qPCR analysis was performed on 13 mouse kidney samples, five heterozygous for Pkd1 and another five wildtype samples that were used for the methylation analysis. Three additional embryonic samples (wildtype, heterozygous and homozygous) were included as controls, as high expression of Pkd1 is observed during development. Target primers for the Pkd1 gene were designed and validated prior to RT-qPCR analysis (Section 2.4, and below).

3.2.1 Primer amplification efficiency

The primer amplification efficiency for the mouse Pkd1 target gene was plotted in a standard curve. Three genes were used, target gene Pkd1 and housekeeping genes Hypoxanthine-guanine phosphoribosyltransferase (HPRT) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The ideal RT-qPCR primer would have an amplification factor of 2, where each subsequent amplification cycle produces exactly double the RT-qPCR product of the previous cycle, and an efficiency of 100%, where each cycle amplifies 100% of the last cycle’s RT-qPCR product.
Amplification efficiency and amplification factor for the *Pkd1* primer (Appendix 3) was generated by using the ThermoFisher Scientific qPCR Efficiency Calculator\textsuperscript{170}. RT-qPCR optimisation was confirmed by comparing the $R^2$ value of the standard curve slope for each primer against the optimal value of $>0.980$ as defined by “Bio-Rad Real-Time PCR Applications Guide”\textsuperscript{171}.

The *Pkd1* primer had an amplification factor of 1.95 and efficiency of 95.27%, which is extremely close to the ideal amplification of 2 and efficiency of 100% and confirms that the primer is suitable for RT-qPCR analysis. The *Pkd1, HPRT* and *GAPDH* standard curve $R^2$ values exceeded the recommended value of 0.980 for an optimised RT-qPCR reaction.
Figure 3.4: Standard curves for RT-qPCR of mouse *Pkd1*, *GAPDH* and *HPRT*. The *Pkd1* primer efficiency was calculated as 95.27% with an amplification factor of 1.95, based on standard curve results (A). Housekeeping genes *GAPDH* and *HPRT* showed R² values above the recommended value of 0.980 for optimised RT-qPCR (B, C). Dilutions of a cDNA sample were made in 1:4 increments up until cycle threshold (CT) readings exceeded the detectable quantity of 35. “0” dilution is an undiluted sample.

3.2.2 *Pkd1* expression of wildtype, heterozygous and embryonic samples

After the RT-qPCR reaction was optimised, the relative gene expression of the ten wildtype or *del34* heterozygous mouse samples used in Section 2.2, and three embryonic (wildtype, heterozygous and homozygous for the *del34 Pkd1* gene disruption) were measured.

*Pkd1* expression was normalised against the *HPRT* and *GAPDH* housekeeping genes expression values (primers provided in Appendix 3) using the Biogazelle qBase PLUS (2.1)
software. Both housekeeping genes showed consistent expression data across all samples (Figure 3.5).

![Graph showing consistent expression of GAPDH and HPRT housekeeping genes.](image)

**Figure 3.5:** Consistent expression of the *GAPDH* and *HPRT* housekeeping genes. WT = Wildtype, HET = *del34* heterozygous and E-HOM, E-WT and E-HET = embryonic homozygous, wildtype and heterozygous for the *del34 Pkd1* disruption respectively. Errors bars are the standard error of mean.

The *Pkd1* gene expression results (Figure 3.5) showed lower expression in the wildtype samples and higher in the heterozygous. The embryonic *del34* homozygous sample *Pkd1*
expression was approximately three times the expression of the non-embryonic del34 heterozygous samples and approximately six times the non-embryonic wildtype samples. The embryonic wildtype and heterozygous Pkd1 expression level was similar to that of the non-embryonic heterozygous samples.

The expression data showed high levels of Pkd1 expression for wildtype sample 1314, when compared with the other wildtype samples. Sample 1314 had been previously genotyped as a wildtype mouse (in provided documentation). To investigate this further, we undertook a genotyping PCR of sample 1314 to check its classification.
Figure 3.6: Relative quantity of Pkd1 expression in 13 mouse samples. Two technical replicates were performed with error bars representing the standard error of mean. Quantity was normalised against two housekeeping genes, which showed consistent expression across all samples. Samples are numbered according to their mouse number (age ~17 months) with the last three consisting of embryonic samples (E18.5) that were homozygous (E-HOM), wildtype (E-WT) and heterozygous (E-HET) for the Pkd1 mutation respectively. Sample 1314 had been genotyped as wildtype in provided documentation, but showed a Pkd1 expression level consistent with heterozygous samples. Re-genotyping of this sample was then performed.

3.3 Genotyping sample 1314

Mouse sample 1314 (Appendix 4) was originally documented as a wildtype sample and was initially analysed as such. The results of the qPCR experiment (Figure 3.6) suggested that the sample had a Pkd1 expression pattern more similar to a heterozygote sample rather than a wildtype. To confirm the sample’s genotype a genotyping PCR was performed. Samples
heterozygous for the *Pkd1 del34* disruption produce a single band of ~400 bp on a 2% agarose electrophoresis gel. A positive, negative and water control was used.

Figure 3.7 shows that the sample 1314 was in fact heterozygous for the *Pkd1 del34* disruption. The methylogram and statistical data were reanalysed, taking this into account (see Figures 3.8, 3.9 and Table 3.1).

![Image of gel Electrophoresis](image)

**Figure 3.7**: Genotyping PCR of sample 1314. Ladder = 1 kb+ DNA ladder. The 1314 sample showed a single band at approximately 400 bp. This indicates that the sample was heterozygous, and is confirmed by the positive (heterozygous for the *Pkd1* disruption) and negative (wildtype) controls. The negative control for this PCR reaction often shows bands of other sizes.

3.4 Reanalysed methylation, expression and statistical data

The qPCR expression data, and the methylation data obtained from the *Pkd1* gene methylation study, was re-analysed, with the sample 1314 re-classified as a *Pkd1* heterozygote. The updated methylation data has been presented as “lollipop plots” (Figure
3.8. There was no change in the outcome of the analysis.

3.4.1 Methylation lollipops

The raw methylogram data was plotted in a “lollipop plot” (Figure 3.8) to compare the level of \textit{Pkd1 del34} heterozygous (top) and the wildtype (bottom) DNA methylation.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{lollipop_plot.png}
\caption{Lollipop comparison plot of mouse \textit{Pkd1} promotor and gene body methylation. Each CpG in the gene body (A, 1-8) and promotor (B, 1-32) regions are represented by a lollipop. The green line represents the division between \textit{Pkd1 del34} heterozygous samples (top) and wildtype samples (bottom). The colour indicates the proportion of methylation at that CpG.}
\end{figure}
3.4.2 Averaged *Pkd1* expression

To illustrate the trend of lower wildtype *Pkd1* gene expression versus the higher *Pkd1 del34* heterozygote expression of the mouse *Pkd1* gene, normalised relative quantity qPCR data was averaged and graphed (Figure 3.9), using the new classification for sample 1314.

![Figure 3.9: Average *Pkd1* expression of wildtype and *del34* heterozygous mouse samples. Wildtype *Pkd1* expression was approximately one third of the *del34* heterozygous samples. Sample 1314 was included in *del34* heterozygous samples after evidence for heterozygosity. Error bars represent the standard deviation.](image)

3.4.3 Statistical data

Sequenced bisulfite-treated clones for *Pkd1* exon 43 were aligned and the methylation at each individual CpG was compared directly between the wildtype and *del34* heterozygous datasets using the Mann-Whitney U-test (Table 3.1). The null hypothesis was defined as there being no significant different between the wildtype and *del34* heterozygous datasets.
Table 3.1: Mann-Whitney U-test at each CpG of a 232 bp region of mouse Pkd1 exon 43

<table>
<thead>
<tr>
<th>CpG 1</th>
<th>CpG 2</th>
<th>CpG 3</th>
<th>CpG 4</th>
<th>CpG 5</th>
<th>CpG 6</th>
<th>CpG 7</th>
<th>CpG 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z-score</td>
<td>0.0046</td>
<td>0.2308</td>
<td>-0.1656</td>
<td>0.4849</td>
<td>-0.888</td>
<td>0.0388</td>
<td>-0.591</td>
</tr>
<tr>
<td>P-value</td>
<td>1</td>
<td>0.8181</td>
<td>0.86502</td>
<td>0.63122</td>
<td>0.37346</td>
<td>0.9681</td>
<td>0.5552</td>
</tr>
<tr>
<td>U-value</td>
<td>821.5</td>
<td>744</td>
<td>1069</td>
<td>1111.5</td>
<td>1219</td>
<td>1350.5</td>
<td>1265</td>
</tr>
</tbody>
</table>

The p-values shown in Table 3.1 (>0.05) all confirm the null hypothesis and all U-values confirm the distribution is approximately normal. Therefore the Z-value listed can be accepted as accurate, showing therefore there is no significant difference in the methylation status of the gene body in mouse kidney samples between wildtype or del34 heterozygous samples.

The re-analysed promoter region of Pkd1 was almost exclusively unmethylated and statistical tests were not able to be used.

3.5 Histology of mouse kidneys

Del34 mouse models produce few cysts with limited renal impairment and potentially methylation changes may only be found in cystic tissue. This may explain why no significant differences were found in methylation between wildtype and heterozygous mouse Pkd1 gene body samples. To examine the density of cysts in our samples, all ten corresponding mouse kidney tissues (section 2.2), which had been fixed in 10% neutral buffered formalin (NBF) and stored in 70% ethanol (for approximately seven years), were embedded, sectioned and hematoxylin and eosin (H&E) stained. Samples were examined for tissue integrity, quality and the presence of cysts (Figure 3.10) and one H&E section per kidney was examined from
the midsagittal plane of the kidney. All samples appeared to be in good condition with very little evidence of cystic tissue. Only one sample (heterozygous 1332) exhibited an obvious cyst; this was confirmed by a third-party pathologist in the department. Potentially, the lack of cysts identified in the mouse kidney tissue may explain the lack of methylation differences seen. As only one section was made from each kidney, it is likely that further cysts exist throughout the del34 heterozygous kidneys but were not sighted in this analysis.
Figure 3.10: H&E staining of mouse kidney tissue used in analysis. Ten samples were imaged and checked for tissue integrity and evidence for cysts: Wildtype cortex (A) and medulla (B) and del34 heterozygous cortex (C) and medulla (D). None of these images showed any cystic structures. A cyst was identified in the mouse kidney sample 1332 (E, del34 heterozygous). F - A magnification of the boxed region in E. Characteristic features such as the “Hob nailing” of epithelial nuclei extending into the centre (1), flattened epithelial cells (2) and cellular debris in the centre of the cyst lumen (3) supports the identification of this renal cyst.
3.6 Mouse and Human PKD1 DNA methylation landscape

The mouse *Pkd1* and human *PKD1* genes were compared using sequence data available from the NCBI gene database\textsuperscript{158}, as observed differences between the human and rodent methylation data and presentation of ADPKD may be due to differences between the human *PKD1* and mouse *Pkd1* gene. The mouse *Pkd1* shares 79\% predicted protein homology with human *PKD1*\textsuperscript{172}. The human *PKD1* contains a large duplicated region in the 5’ end that has ~97 \% similarity to six pseudogenes on chromosome 16 (Figure 3.11)\textsuperscript{173}. The mouse genome does not contain any duplicated pseudogenes of *Pkd1*.

The predicted CpG islands across the two genes, as defined by the “NCBI CpG island strict prediction algorithm”\textsuperscript{158}, showed some differences. Notably, fewer and smaller CpG islands are observed in the mouse *Pkd1* gene.

The equivalent mouse promotor and exon 43 gene body regions of the human gene that were examined in Woo et al., (2014)\textsuperscript{104} were investigated in this current study. The mouse promotor region and its equivalent human promotor region both lie in CpG rich areas. The mouse equivalent region to human exon 43 lies close to, but does not contain a CpG island. In contrast, a CpG island spans exons 42 to 46 in the human gene.
Figure 3.11: A comparison of the *PKD1* and *Pkd1* gene landscapes between human and mouse. The human duplicated region is highlighted in red. Orange bars represent exons, green bars represent CpG islands. The blue arrow indicates Exon 43, the mouse gene body region examined. Based on data from the NCBI gene database.
Chapter 4: Discussion of mouse Pkd1 methylation

4.1 Pkd1 gene body methylation

This study examined DNA methylation in two regions of the mouse Pkd1 gene, a 292 bp region in the gene promotor and a 232 bp region of the exon 43 (gene body) in Pkd1. We identified no difference in the promotor region and found no significant difference in the methylation of the gene body between wildtype and del34 heterozygous mouse kidneys.

A paper by Woo et al., 2014 previously identified an increase in gene body methylation, which correlated with a decrease in PKD1 expression in human patients with ADPKD. They found no difference in promotor methylation but did observe an increase in PKD1 gene body methylation, particularly in exon 43. A change in methylation could serve as a “second hit” for ADPKD pathogenesis, where the first PKD1 allele is knocked-out by an inherited mutation and the second is silenced by changes in DNA methylation. Our aim was to investigate if a similar difference might occur in the heterozygous del34 mouse model that develops cysts in late onset in a similar manner to human ADPKD.

We began by investigating the equivalent gene body region in the mouse Pkd1 gene to that analysed in human PKD1. This sequence only contained eight CpGs and wasn’t considered a CpG island based on the NCBI Strict algorithm. We identified that the gene body region for both wildtype and del34 heterozygous samples was almost entirely hypermethylated, with a Mann-Whitney U-test on each CpG confirming no significant difference between samples at any of the CpGs. Both groups contained a single CpG that was unmethylated, which may indicate the border between distinct hypomethylated and hypermethylated regions in the mouse gene body, as was seen in several regions of the human PKD1 gene in our data in both ADPKD and wildtype samples. Techniques such as genome-wide and reduced representation bisulfite sequencing could be used to investigate the DNA methylation status of the entire
methylation landscape or specific CpG islands across the Pkd1 gene to determine if any methylation differences may be located elsewhere that may reflect the published human data.

A technical issue that hampered this study was that the amount of high quality sequences that exceeded the required quality control checks set by the BiQ methylogram software were often few. The protocol was altered multiple times with the intent on improving sequence quality and efficiency of the process, although transformation efficiency was improved, sequence quality remained low with the majority of sequences not reaching the 80% reference sequence identity set by BiQ. This may be due to residual reagents in the eluted DNA, although the methods (section 2.2.7) were altered to reduce this possibility.

The Pkd1 mouse equivalent region of the human PKD1 promotor examined in Woo et al., 2014 was investigated during my Post Graduate Diploma and was re-analysed with additional sequences acquired during this Masters project. All 33 CpG’s within this promotor region were hypomethylated in both wildtype and del34 heterozygous samples. Current literature on DNA methylation suggests that a hypomethylated promotor region and hypermethylation in the gene body is what we would expect for a transcriptionally active gene. It has been suggested that gene body methylation can suppress transcriptional noise that may disrupt regular gene expression, as well as regulating enhancer and silencer sequences. A hypomethylated promotor region allows the binding of transcription factors to the DNA and subsequently the initiation of gene expression. Woo et al., 2014 also found that the human PKD1 promotor was hypomethylated and the gene body was hypermethylated, which suggests that the gene is active in adult tissue, as we have observed for the Pkd1 gene in mice. The Woo group however identified an increase in methylation in the PKD1 gene body. This was not able to be observed in our studies, as all sites investigated showed no significant difference between wildtype and del34 heterozygous mice. Currently there is no other literature on the DNA methylation landscape of PKD1.
Cystogenesis is often likened to aspects of tumour formation. Woo et al. identified that 91% of methylation changes found in ADPKD kidneys were hypermethylation (with 61% in genic regions and only 1% in promoters). In contrast, cancer studies have reported a global decrease in methylation, resulting in the activation of repetitive elements, retrotransposons and oncogenes, with small areas of hypermethylation in tumour-suppressing gene promoters.

It is possible that any difference in DNA methylation may only be found in the cyst lining tissue and since the del34 mouse model has few cysts and the DNA was extracted from whole kidney, any cyst specific methylation differences may have been diluted by the non-cyst cells and missed in the analysis as background noise. Histology of mouse del34 heterozygous tissue samples identified one confirmed cyst in a single heterozygous kidney, although only one section was examined from each kidney due to time constraints. It is likely that more cysts exist in the heterozygous tissues. The relative lack of cystic tissue, and known differences between mouse and human pathogenesis of the disease, may contribute to the lack of differences we see in methylation between the heterozygous and wildtype mouse tissues.

In order to identify if DNA methylation differences are occurring in a heterozygous mouse model of cystic disease, the cells lining the cyst could be isolated and analysed. Microdissection of cyst lining epithelial cells has been performed before in our lab and DNA methylation analysis of this tissue in comparison to surrounding tissue and wildtype samples may reveal methylation differences that are lost in the ‘noise’ of whole tissue analysis, and that are associated with cystogenesis and perhaps the second hit mechanism. The microdissection technique does not produce sufficient DNA yield for our current methods of examining sample methylation. A modified reduced representation bisulfite sequencing (RRBS) method is currently being developed by members of our lab group that may be able to utilise very small quantities of DNA and could suit this purpose in future studies.
Potentially, and especially as the mouse gene body region analysed was not strictly a CpG island, DNA methylation differences may be located elsewhere in the mouse Pkd1 gene. To examine this, the entire gene could be examined using genome-wide bisulfite sequencing. The mouse Pkd1 contains 78% sequence homology\textsuperscript{113,172} with the human PKD1 and contains only five small CpG islands compared the six larger islands found in the human gene (see figure 3.11) according to NCBI's Strict algorithm\textsuperscript{158}. Overall, Antequera and Bird (1993)\textsuperscript{176} reported that the mouse genome contains approximately 37,000 CpG islands compared to 45,000 in the human genome. Whether this difference has a functionally significant difference is unknown in current literature, although human research suggest that even non-methylated CpG islands throughout the genome may be involved in mechanisms not directly related to gene expression, such as the maintenance of chromatin structure\textsuperscript{177}.

The mouse Pkd1 promotor was hypomethylated and the gene body hypermethylated in both wildtype and heterozygous samples and therefore the Pkd1 was likely to be transcriptionally active. To confirm this, we performed qPCR expression experiments targeting Pkd1. We found that Pkd1 is expressed in adult mouse kidneys and that surprisingly, expression levels were higher in del34 heterozygous mice when compared to wildtype. Low expression of Pkd1 has been previously shown in some hypomorphic mouse models of ADPKD\textsuperscript{57,178} and this observation is supported by the formation of cysts after the expression levels of Pkd1\textsuperscript{51} are decreased, although induction of disease by overexpression\textsuperscript{54} and increased expression of PKD1 or polycystin-1 in human ADPKD tissues has been reported\textsuperscript{53,179,180}. This suggests that PKD1 gene expression is very carefully regulated in the kidney and any significant changes to PKD1 expression in either direction is capable of initiating the disease. The del34 mouse model produces a truncated Pkd1 mRNA\textsuperscript{43} and the Zhou lab group that developed the del34 mouse model published a paper in 2001\textsuperscript{178}, which identified that the del34 homozygous and heterozygous mutants produced both wildtype and ~14 kb and ~11 kb mutant transcripts of
*Pkd1*. It would be interesting to identify which transcripts are showing the increased expression in our samples, and this could be done using a second set of primers downstream of the *del34* disruption that would amplify only the wildtype allele. The proportion between the two expression datasets may indicate which allele is being predominantly expressed in heterozygous mice. Our expression data revealed that the embryonic *del34* homozygote showed the highest expression of all samples analysed, which indicates that the mutant allele may be overexpressed, although only one sample was used in the analysis. Additional analysis on the expression of the PKD1 protein and miRNAs within the *PKD1* gene would further clarify additional gene regulatory effects of the *del34* mutant.

Analysis of the products of human *PKD1* in ADPKD shows a large variety of truncation and non-truncating mutant transcripts\(^1\). As we see a difference in *Pkd1* expression in the *del34* heterozygous samples when compared to wildtype, but no difference in methylation in the two regions examined, potentially methylation may not be involved in the second hit in the *del34* mouse model, or the DNA methylation changes may occur elsewhere in the gene.

As no DNA methylation differences and a lack of cysts were identified in the *del34* mouse model, and in addition to the unusual finding of 91% hypermethylation changes in human ADPKD by Woo et al., (2014)\(^1\), we decided to expand the limited human DNA methylation data by using reduced representation bisulfite sequencing (RRBS). This technique generates a representative dataset of the methylation status of ~85% of the CpG islands in the human genome\(^1\).
Chapter 5: Human DNA methylation analysis of *PKD1*

5.1 Summary of RRBS data analysis

To add to the limited published data available on DNA methylation of the human *PKD1* gene in cystic samples, we have compared the methylation landscape of human *PKD1* and ADPKD related genes in wildtype and ADPKD kidney tissues. Five human kidney samples were analysed using reduced representation bisulfite sequencing (RRBS), two from ADPKD cystic kidneys with the remaining three from wildtype age-matched kidneys. This technique gives, representative methylation data from ~85% of CpG islands across the genome. Prior to sequencing, multiple steps required sample quality and quantity checks before proceeding with the protocol.

The wildtype human kidney tissues were obtained from the Christchurch Tissue Bank\textsuperscript{162} after being removed from patients undergoing surgical procedures to remove renal tumours. The samples were taken from healthy cortical tissue as far from the renal tumour as possible. The two human ADPKD tissue samples were used from donated tissue. All samples were stored and used under the approval by the Otago Human Ethics Committee H15/110.

5.1.1 First size selection

To gain a representative look at CpG islands across the human genome, MspI digested and adapter ligated human library preparations of two ADPKD and three wildtype kidney DNA samples (section 2.8.2 and 2.8.3) were run on a NuSieve gel and the DNA sliced from the gel between the 150 and 325 bp bands (adjusted to include the adapter size, these bands were chosen to select for a representative sample of ~85% of all CpG islands in the genome) of the 25 bp DNA ladder. The DNA samples were not exposed to UV light due to the risk of
degradation and therefore the DNA ladder was cut from the gel and imaged under UV light separately. The 150 and 325 bp bands were then marked and the distances measured on the gel. A small section of each lane containing DNA was then removed from between those two measurements for continuing library preparation. The remaining gel pieces were then placed back together and imaged under UV to determine if the slices were made in the correct place.

Figure 5.1: First size selection gel extraction accuracy check. Reduced representation bisulfite sequencing (RRBS) requires only MspI digested fragments between 150 and 325 bp. The DNA samples cannot be exposed to UV so the ladder was removed and imaged separately (A). The distance to the 150 bp and 325 bp bands were then measured using a ruler under UV light and the remaining gel sliced using the measure lengths at the estimated bands. The gel fragments are then imaged together to determine the accuracy of the size selection (B).
5.1.2 Semi-quantitative PCR of library DNA

Large-scale amplification of DNA samples requires a number of PCR cycles that will yield sufficient quality, while avoiding the potential errors that can arise from excessive amplification. Every sample was subjected to two semi-quantitative PCR reactions at 15 and 20 cycles. The PCR products were then ran on NuSieve gels to determine the suitable number of amplification cycles (Figure 5.2). The electrophoresed gels were visualised under UV and the intensity of the 15 and 20 cycle PCR lanes for each sample was compared with the predetermined number of amplification cycles decided by Drs. Rodger and Chatterjee who developed the protocol. The NuSieve gel for samples 1 and 2 showed little intensity (figure 5.2 A) between the size selected 150 and 325 bp regions and comparison between the 15 and 20 PCR lane intensity was difficult. The NuSieve gel is electrophoresed for a long time (~90 minutes), risking the possibility for complete consumption of ethidium bromide in the solution. The failed gel had not been run using fresh running buffer and that was believed to be the cause for the lack of fluorescence. Rather than performing another semi-quantitative PCR, both samples 1 (ADPKD08) and 2 (ADPKD07) were estimated to require 20 cycles for large-scale, to conserve the limited quantity of library DNA. Samples 3, 4 and 5 (Wildtypes H3, E1 and G2 respectively – figure 5.2 B) suggested cycles of 15, 18 and 20 respectively. The gel images in figure 5.2 also indicated the accuracy of the first round of size selection, as a perfectly size selected library would show the most fluorescence between the 150 and 325 bp regions. Samples 3, 4 and 5 showed fluorescence consistent with accurate size selection. Due to the failed gel for samples 1 and 2, the accuracy of the first size selection was not able to be determined until after the large-scale amplification and second size selection.
Figure 5.2: Semi-quantitative PCR of RRBS libraries. Intensity of the size selected RRBS libraries was used to determine the most appropriate number of amplification cycles for large-scale PCR. A failed gel did not produce sufficient fluorescence in either lane for both samples (A), and this was believed to be caused by reusing gel running buffer for a second time or consuming all of the ethidium bromide. B – A successful gel with strong fluorescence between the 150 and 325 bp bands indicated the first size selection was successful.
5.1.3 Final size selection

Before final purification of the libraries, the samples were size selected again to remove any fragments outside of the desired 150 – 325 bp range.

**Figure 5.3:** Second size selection gel extraction accuracy check. The RRBS libraries were size selected for a second time before purification. Gels fragments were cut at the estimated 150 and 325 bp bands (represented by the bracket) using the method described in Section 2.8.4. Examination of the gel under UV light following removal of the gel fragments indicated that the size selection appeared to be accurate.
5.1.4 Sequencing quality

The DNA libraries were sequenced by New Zealand Genomics Limited (methods section 2.8.7 and 2.8.8) and were received with quality control reports generated by the program FastQC\textsuperscript{164} (V 0.11.3). All samples were of high quality with zero sequences tagged as ‘poor quality’. The sequence quality per base (Figure 5.4) showed high Phred quality (>30) scores of 34-36 (~1 in 5000 probability of incorrect base call) with typical gradual decline towards the end of the read.

DNA base content at each position of the sequenced reads of one of the samples is shown in Figure 5.5, where the characteristic first three bases of CGG (methylated cytosine) or TGG (not methylated) indicate successful MspI digestion. Low cytosine scores across the reads also indicate successful bisulfite conversion with the gradual increase at the end of the read attributable to the multiplex adapters which contain methylated cytosines. Additional alignment reports generated by Bismark\textsuperscript{167} confirmed that bisulfite conversion was successful with all five datasets showing only 1.3-1.9% of all cytosines were not converted during the reaction.
Figure 5.4: The base quality control data of a sequenced RRBS library (Sample 5). Base quality scores across each position of the total reads. The X axis depicts the base position along the reads (maximum read length of 100 bp). Y axis is Phred quality score, with a score of 35 corresponding to a 1 in 5000 probability of wrong base calling. The blue line shows the mean quality score at each position, with the black error bars representing the lower (10%) and upper (90%) percentile range (several go beyond the top of the graph). The gradual decline of quality towards the end of the reads is typical for this kind of data.
Figure 5.5: The sequence quality control data of a sequenced RRBS library (Sample 5). The percentage of each base of all reads for one sample (100 bp max read length) is shown. The X-axis shows positions along the read, with the Y-axis representing the percentage value for proportion of the bases at each position. The first three bases are consistently CGG and TGG which confirms the MspI digestion was successful. The considerably lower proportion of cytosines across the read (excluding the adapter readings towards the 100 bp end) confirms bisulfite conversion was successful.
5.1.5 Integrated Genome Viewer of PKD1

The DNA sequenced reads were aligned using Bismark aligning tools as discussed in section 2.8.9. Integrated Genome Viewer (IGV) was used to visualise the aligned libraries. The HG19 human genome (supplied in the software) was used as a reference to align each library. The PKD1 gene (Figure 5.6) and other related cyst genes were visualised for any methylation differences.

Figure 5.6: The methylation landscape of the human PKD1 gene. ADPKD07 (1), ADPKD08 (2), WILDTYPE E1 (3), WILDTYPE G2 (4) and WILDTYPE H3 (5) RRBS data was aligned and visualised against the human reference HG19 PKD1 track. Red indicates a methylated CpG, blue unmethylated. Orange boxes indicate areas of the PKD1 gene that were hypermethylated in ADPKD samples compared to the wildtype samples. Green boxes indicate areas where the PKD1 was gene hypomethylated in ADPKD samples compared to the wildtype. Higher resolution images of boxed areas can be found in Appendix 6.
The aligned RRBS sequence data reveals some areas of methylation difference between ADPKD and wildtype samples (Figure 5.6). The \textit{PKD1} gene body region where read coverage was available appeared to be hypermethylated in all samples. Several regions including the promoter and exon 4 were hypomethylated in all samples. Two intronic areas between exons 1 and 2 appeared to be hypomethylated in the ADPKD samples compared to the wildtypes tissues, whereas an intronic region between exons 33 and 34 was hypermethylated in the ADPKD samples. Exon 42 also showed hypermethylation in the ADPKD samples, which is close to the region reported by Woo \textit{et al.}, (2014)\textsuperscript{104} to be hypermethylated in ADPKD tissue. They identified an increase in DNA methylation in \textit{PKD1} exon 43, although unfortunately no reads were obtained for exon 43 using the RRBS method in this study. Images of these specific regions with accurate genomic coordinates can be found in Appendix 6.

The paper by Woo \textit{et al.}, (2014)\textsuperscript{104} listed several other genes that were found to be hypermethylated in their ADPKD samples, when compared to wildtype kidney samples. These genes were examined in our data and the results summarized in the table below. “ADPKD samples” indicates the comparison between the ADPKD and wildtype samples and any differences noted between the two datasets were listed by their location in the gene. Additional genes (\textit{BCAR1, COL1A1, PAX2} and \textit{STAT5}) that were detected during the IGV and ANOVA analyses were also included in table 5.1. These results indicate gene specific changes to methylation in ADPKD samples, rather than global or whole gene changes in methylation.
Table 5.1: IGV analysis on genes reported as hypermethylated in ADPKD samples by Woo et al., (2014).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>ADPKD samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZFAT</td>
<td>Ion transport</td>
<td>No difference</td>
</tr>
<tr>
<td>OSTA/SLC51A</td>
<td>Ion transport</td>
<td>No difference</td>
</tr>
<tr>
<td>P2RX4</td>
<td>Ion transport</td>
<td>No difference</td>
</tr>
<tr>
<td>ATP4B</td>
<td>Ion transport</td>
<td>No difference</td>
</tr>
<tr>
<td>TRPV1</td>
<td>Ion transport</td>
<td>No difference</td>
</tr>
<tr>
<td>SLC22A18</td>
<td>Ion transport</td>
<td>Hypermethylated and hypomethylated in several intronic regions</td>
</tr>
<tr>
<td>LCK</td>
<td>Calcium signalling</td>
<td>Hypomethylated in promotor</td>
</tr>
<tr>
<td>CHRNA10</td>
<td>Calcium signalling</td>
<td>No difference</td>
</tr>
<tr>
<td>CLDN9</td>
<td>Calcium signalling</td>
<td>No difference</td>
</tr>
<tr>
<td>CACNA1H</td>
<td>Calcium signalling</td>
<td>Hypermethylated and hypomethylated in several intronic regions</td>
</tr>
<tr>
<td>SYMPK</td>
<td>Cell adhesion</td>
<td>Hypermethylation in exon 20, hypomethylation across exons 26 and 27</td>
</tr>
<tr>
<td>SPCG7</td>
<td>Cell adhesion</td>
<td>No difference</td>
</tr>
<tr>
<td>SEMA4D</td>
<td>Cell adhesion</td>
<td>Hypermethylation in exon 17</td>
</tr>
<tr>
<td>LGALS4</td>
<td>Cell adhesion</td>
<td>No difference</td>
</tr>
<tr>
<td>DSG4</td>
<td>Cell adhesion</td>
<td>No difference</td>
</tr>
<tr>
<td>MUC5B</td>
<td>Cell adhesion</td>
<td>No difference</td>
</tr>
<tr>
<td>BCL2</td>
<td>Cell morphogenesis</td>
<td>No difference</td>
</tr>
<tr>
<td>LIG1</td>
<td>Cell morphogenesis</td>
<td>Hypomethylation in intronic region between 23 and 24</td>
</tr>
<tr>
<td>SALL1</td>
<td>Cell morphogenesis</td>
<td>No difference</td>
</tr>
<tr>
<td>SOX10</td>
<td>Cell morphogenesis</td>
<td>No difference</td>
</tr>
<tr>
<td>FOXI1</td>
<td>Cell morphogenesis</td>
<td>No difference</td>
</tr>
<tr>
<td>TMEM176B</td>
<td>Cell morphogenesis</td>
<td>No difference</td>
</tr>
<tr>
<td>DTX1</td>
<td>Development</td>
<td>No difference</td>
</tr>
<tr>
<td>CCDC88C</td>
<td>Development</td>
<td>Hypomethylation in one intronic region fragment between 3 and 4</td>
</tr>
<tr>
<td>CSNK1G2</td>
<td>Development</td>
<td>No difference</td>
</tr>
<tr>
<td>NOTCH1</td>
<td>Development</td>
<td>Hypermethylation in exon 5</td>
</tr>
<tr>
<td>NOTCH2</td>
<td>Development</td>
<td>No difference</td>
</tr>
<tr>
<td>WNT4</td>
<td>Development</td>
<td>No difference</td>
</tr>
<tr>
<td>WNT7A</td>
<td>Development</td>
<td>No difference</td>
</tr>
<tr>
<td>WNT9B</td>
<td>Development</td>
<td>No difference</td>
</tr>
<tr>
<td>WNT11</td>
<td>Development</td>
<td>No difference</td>
</tr>
<tr>
<td>HDAC1</td>
<td>Chromatin remodelling</td>
<td>No difference</td>
</tr>
<tr>
<td>DOT1L</td>
<td>Chromatin remodelling</td>
<td>Hypomethylation in exons 12 and 24</td>
</tr>
<tr>
<td>JMJD3/KDM6B</td>
<td>Chromatin remodelling</td>
<td>No difference</td>
</tr>
<tr>
<td>EHMT1</td>
<td>Chromatin remodelling</td>
<td>Hypermethylation in intronic region between 8 and 9</td>
</tr>
<tr>
<td>EHMT2</td>
<td>Chromatin remodelling</td>
<td>No difference</td>
</tr>
<tr>
<td>PKD1L1</td>
<td>Ion transport</td>
<td>No difference</td>
</tr>
<tr>
<td>BCAR1</td>
<td>Cell adhesion migration, invasion, apoptosis and hypoxia*</td>
<td>Hypermethylation in exons 5 and 7</td>
</tr>
<tr>
<td>COL1A1</td>
<td>Component of collagen*</td>
<td>Hypomethylation in exons 5 and 48</td>
</tr>
<tr>
<td>PAX2</td>
<td>Transcription regulation*</td>
<td>Hypomethylated across 3’ end of gene including exons 10 and 11 and the UTR region.</td>
</tr>
<tr>
<td>STAT5</td>
<td>Component of JAK/STAT pathway*</td>
<td>Hypomethylation in exons 2 and 3</td>
</tr>
</tbody>
</table>

* Gene functions taken from NCBI database
Functions listed for genes that do not contain an asterix in table 5.1 are provided by Woo et al., 2014\textsuperscript{104}. Gene functions that are asterixed are taken from the gene summary displayed in the NCBI database\textsuperscript{158}.

Of the 37 genes reported as showing hypermethylation in ADPKD tissues compared to wildtype tissues in the Woo et al., (2014)\textsuperscript{104} paper, we identified six genes that showed regions of hypermethylation (when compared to wildtype samples), with three of these also showing regions of hypomethylation in the same gene. In total, seven genes show at least one region of hypomethylation, with 27 showing no difference at all. Of the four other genes investigated; \textit{BCAR1} displayed hypermethylation in two exons; with the remaining three (\textit{COL1A1, PAX2, STAT5}), showing hypomethylation in at least two exons.

5.1.6 ANOVA statistical identification of differentially methylated regions

An analysis of variance (ANOVA) statistical test was performed on all RRBS fragments to identify fragments throughout the genome that have the largest proportional difference in methylation between the ADPKD and wildtype sample reads. Table 5.2 lists RRBS fragments with the chromosome and genomic location (Chr#, start, end), the length of the fragment, as determined by the overlap of sequence reads and varies either side of the average read length of 100 bp and the amount of CpGs that fragment contains. The “MethDiff” indicates the proportion of methylation differences between the ADPKD and wildtype samples for that fragment. The data was sorted to show the most hypermethylated ADPKD fragments (>0 methdiff) and most hypomethylated (<0 methdiff) compared to the wildtype samples.
Table 5.2: Top differentially methylated RRBS fragments in the human genome.

<table>
<thead>
<tr>
<th>Chr#</th>
<th>Start</th>
<th>End</th>
<th>Length</th>
<th>CpGs</th>
<th>P-value</th>
<th>F-value</th>
<th>MethDiff</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>592961</td>
<td>593019</td>
<td>59</td>
<td>4</td>
<td>0.00670305</td>
<td>45.27</td>
<td>0.6551</td>
<td>PRKAR1B</td>
</tr>
<tr>
<td>16</td>
<td>2147611</td>
<td>2147671</td>
<td>61</td>
<td>3</td>
<td>0.00317369</td>
<td>76.06</td>
<td>0.6413</td>
<td>PKD1</td>
</tr>
<tr>
<td>22</td>
<td>41846168</td>
<td>41846260</td>
<td>93</td>
<td>5</td>
<td>0.00582055</td>
<td>49.97</td>
<td>0.6344</td>
<td>TOB2</td>
</tr>
<tr>
<td>3</td>
<td>111431353</td>
<td>111431517</td>
<td>165</td>
<td>4</td>
<td>0.0104055</td>
<td>33.16</td>
<td>0.6221</td>
<td>ZNF750</td>
</tr>
<tr>
<td>4</td>
<td>53714737</td>
<td>53714836</td>
<td>100</td>
<td>3</td>
<td>0.0212211</td>
<td>19.73</td>
<td>0.6182</td>
<td>RASL11B</td>
</tr>
<tr>
<td>16</td>
<td>57657149</td>
<td>57657213</td>
<td>65</td>
<td>3</td>
<td>0.00514093</td>
<td>54.49</td>
<td>0.6169</td>
<td>GPR56</td>
</tr>
<tr>
<td>17</td>
<td>80818649</td>
<td>80818771</td>
<td>123</td>
<td>7</td>
<td>0.0564832</td>
<td>9.16</td>
<td>0.6101</td>
<td>ZNF750</td>
</tr>
<tr>
<td>12</td>
<td>132865276</td>
<td>132865388</td>
<td>113</td>
<td>9</td>
<td>1.48E-06</td>
<td>13046.79</td>
<td>0.6041</td>
<td>GALNT9</td>
</tr>
<tr>
<td>17</td>
<td>44942189</td>
<td>44942283</td>
<td>95</td>
<td>2</td>
<td>0.00621706</td>
<td>47.72</td>
<td>0.6025</td>
<td>WNT9B</td>
</tr>
<tr>
<td>22</td>
<td>50250431</td>
<td>50250474</td>
<td>44</td>
<td>4</td>
<td>0.0295304</td>
<td>15.36</td>
<td>0.5999</td>
<td>ZBED4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chr#</th>
<th>Start</th>
<th>End</th>
<th>Length</th>
<th>CpGs</th>
<th>P-value</th>
<th>F-value</th>
<th>MethDiff</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>83230518</td>
<td>83230585</td>
<td>68</td>
<td>3</td>
<td>0.00115085</td>
<td>151.88</td>
<td>-0.8759</td>
<td>ENOPH1</td>
</tr>
<tr>
<td>10</td>
<td>102588174</td>
<td>102588314</td>
<td>141</td>
<td>7</td>
<td>0.00205013</td>
<td>102.59</td>
<td>-0.6637</td>
<td>PAX2</td>
</tr>
<tr>
<td>17</td>
<td>73839504</td>
<td>73839629</td>
<td>126</td>
<td>3</td>
<td>0.00933566</td>
<td>35.83</td>
<td>-0.645</td>
<td>UNC13D</td>
</tr>
<tr>
<td>18</td>
<td>77266975</td>
<td>77267020</td>
<td>46</td>
<td>2</td>
<td>0.0118251</td>
<td>30.26</td>
<td>-0.6317</td>
<td>NFATC1</td>
</tr>
<tr>
<td>10</td>
<td>102589102</td>
<td>102589150</td>
<td>49</td>
<td>4</td>
<td>0.0173507</td>
<td>22.9</td>
<td>-0.6303</td>
<td>PAX2</td>
</tr>
<tr>
<td>19</td>
<td>933115</td>
<td>933292</td>
<td>178</td>
<td>8</td>
<td>0.0057767</td>
<td>171.61</td>
<td>-0.6235</td>
<td>C19orf22</td>
</tr>
<tr>
<td>17</td>
<td>78866065</td>
<td>78866211</td>
<td>147</td>
<td>6</td>
<td>0.00092236</td>
<td>176.41</td>
<td>-0.6186</td>
<td>RPTOR</td>
</tr>
<tr>
<td>1</td>
<td>163076804</td>
<td>163076885</td>
<td>82</td>
<td>3</td>
<td>0.0786753</td>
<td>11.23</td>
<td>-0.6166</td>
<td>NUF2</td>
</tr>
<tr>
<td>1</td>
<td>146704307</td>
<td>146704365</td>
<td>59</td>
<td>3</td>
<td>0.00166631</td>
<td>118.15</td>
<td>-0.6153</td>
<td>FMOS</td>
</tr>
<tr>
<td>6</td>
<td>30653437</td>
<td>30653520</td>
<td>84</td>
<td>3</td>
<td>0.00276706</td>
<td>83.57</td>
<td>-0.6115</td>
<td>PPP1R18</td>
</tr>
</tbody>
</table>

The ANOVA test (Table 5.2) when sorted for the top ten most hypermethylated fragments revealed several genes that are involved in pathways that are related to ADPKD such as:

*PKD1; TOB2* (cell cycle progression); *GPR56* (cell adhesion); *ZNF750* (associated with mis-differentiation) and *WNT9B* (associated with renal tubule development and renal cystogenesis).

The ten most hypomethylated regions in the ADPKD tissue compared to the wildtype samples also revealed genes involved in similar pathways including: *PAX2* (associated with kidney development, cancer and ADPKD, the 3’ end of this gene was also identified to be
hypomethylated in the IGV analysis) and *RPTOR* (regulates mTORC1, which is associated with ADPKD, cell growth and survival). Most fragments have a raw P-value below 0.05 and therefore the difference in methylation between the ADPKD and wildtype samples at that fragment may indicate a significance difference. The F-values are all >1.0 and further indicate that the variation of the group mean is greater than expected by chance alone (An F-value close to one suggests the mean variation between the two fragments is likely to be due to chance). The *ZNH750* and *NUF2* genes appeared not significant with raw p-values of 0.056 and 0.079 and f-values of 9.16 and 11.23.
Chapter 6: Discussion for human methylation analysis of *PKD1*

6.1 Human *PKD1* methylation

To determine and compare the methylation landscape of the human *PKD1* gene and associated cystogenesis-related genes against the limited published literature, we used the RRBS technique to reveal a representative dataset of methylation differences between ADPKD and wildtype kidney samples across the genome.

Woo *et al.*, (2014)\(^{104}\), which performed genome-wide bisulfite analysis on human ADPKD samples reported hypermethylation of the *PKD1* gene, and specifically published an increase in exon 43 gene body methylation in patients with ADPKD. They correlated this with a decrease in expression of *PKD1*. To investigate the methylation landscape of *PKD1* in our samples, we performed reduced representation bisulfite sequencing (RRBS) on two kidney samples for ADPKD patients, and three wildtype kidney samples. The methylation status of the *PKD1* and other interacting genes were examined with the Integrative Genomics Viewer (IGV). The RRBS technique was chosen because it is considerably cheaper than whole genome bisulfite sequencing\(^ {123}\) and produces a representative dataset of ~85% of the CpG islands within the genome\(^ {124}\). Because of this, the information available for each gene varied depending on the genomic location of the fragments. By comparing the methylation status of the RRBS fragments in both the ADPKD and wildtype samples using IGV, regions in ADPKD related genes that appeared to be differentially methylated were also identified.

*PKD1* appeared to be hypomethylated in the promotor region and hypermethylated in the gene body region across all samples, as was also seen in our mouse *Pkd1* methylation analysis. This is what is expected for an active gene\(^ {88}\) and suggests that *PKD1* is expressed in ADPKD samples. Four regions in *PKD1* were found to be differentially methylated in
ADPKD samples compared to wildtype samples, two hypomethylated fragments towards the intronic promoter region, one hypermethylated region in the intronic gene body and one hypermethylated region in exon 42. Woo et al. reported an increase in methylation of exon 43 in ADPKD tissues compared to wildtype tissues, which may be supported by the increase that we identified in exon 42, although no data on exon 43 was obtained in our dataset. Several other genes were reported as showing overall hypermethylation by Woo et al (Table 5.1). Whole gene methylation comparisons may miss important regions of differential methylation, as a gene that shows overall hypermethylation may contain discrete regions of hypomethylation and hypermethylation. To determine if the differentially methylated regions in our study could influence expression and splicing, transcriptome and RT-qPCR experiments would have to be performed. Due to time constraints, this was not able to be examined during the course of this research. Potentially if these DNA methylation differences are able to influence alternative splicing patterns, then there is potential for PKD1 DNA methylation to play a role in ADPKD pathogenesis.

Genes that were identified as being hypermethylated in ADPKD samples in the Woo paper were also investigated in our dataset using IGV (Table 5.1). Many of these genes were associated with key characteristics of ADPKD such as cell adhesion, cell regulation and differentiation and renal development. Genes such as SLC22A18, LCK, CACNA1H, SYMPK, SEMA4D, LIG1, CCDC88C, NOTCH1, DOT1L and EHMT1 all showed at least one differentially methylated fragment whereas all the remaining genes (found in Table 5.1) appeared to have no methylation differences in the CpG island enriched dataset that was obtained via the RRBS method. The addition of more ADPKD and wildtype samples for analysis and more read coverage using techniques such as genome-wide bisulfite sequencing would provide more support to these conclusions.
Analysis of variance (ANOVA) statistical testing was also performed on all the RRBS fragments to find regions that were the most hypermethylated and hypomethylated in the ADPKD samples compared to the wildtypes. \( PAX2 \) was identified in the ANOVA test as having several fragments that were suggested to be significantly (<0.05) hypomethylated in the ADPKD samples and this was confirmed by investigating the gene in IGV. \( PAX2 \) gene dosage influences cyst formation in the \( del34 \) ADPKD mouse model\(^1\) and reducing \( PAX2 \) expression has been shown to increase apoptosis and reduce renal cyst progression\(^2\). The IGV data revealed that the hypomethylated region in the \( PAX2 \) gene is found in the 3’ end of the gene covering exons 10 and 11 and the untranslated region, which may be involved in DNA organisational regulatory mechanisms such as enhancer and silencer regulation and no literature has currently examined this part of the gene for regulatory elements. This could be investigated further using the chromosome conformation capture technique, which uses formaldehyde to preserve DNA-DNA interactions. This has been performed on a large number of genes and model organism, including interactions between the developmental polycomb repressive complexes 1 and 2 (PRC1 and 2) and \( PAX2 \)\(^3\). Any correlation between the hypomethylated region and \( PAX2 \) expression levels may be determined by gene expression analysis such as qPCR. Because the differentially methylated RRBS fragments only provide data on CpG rich regions, the methylation differences identified by IGV and the ANOVA test (Figure 5.6 and Table 5.2), may not reflect the methylation landscape of the entire gene, although the regions identified may influence alternative splicing and transcription stability. The identified regions in these genes provide a good base for further investigation.

There is very little research on the methylation landscape of the human \( PKD1 \) or mouse \( Pkd1 \) genes and other similar genes involved in polycystic kidney diseases. This study focused on \( PKD1 \) and other cystogenesis specific genes. Further methylation analysis of the CpG island
enriched genome will be possible by analysing the complete RRBS data for the five human
tissues investigated in this study, to examine global trends and differentially methylated
regions. A variety of tools exist to further analyse this data such as differential methylation
analysis package (DMAP)\textsuperscript{184}.

Genome-wide bisulfite sequencing could be used to investigate the genome-wide methylation
landscape with more coverage than the RRBS method used in this study. In addition, targeted
bisulfite sequencing could be performed on selected genes to identify methylation differences
with more coverage and greater depth than a genome-wide approach.

As well as global changes in DNA methylation that are associated with cystogenesis, the
possibility still exists that DNA methylation changes in the remaining unmutated allele of the
\textit{PKD1} gene could, in some cysts, be a second hit. RRBS requires a considerable quantity of
starting DNA to ensure good quality libraries for sequencing. Enhanced methodology is
being developed in this laboratory to greatly reduce the quantity required for this process, and
to potentially allow RRBS analysis at a single cell level in both human and mouse samples. If
the RRBS process is able to be performed on isolated cyst epithelial cells from either mouse
or human cysts, then the methylation landscapes would be able to be compared directly on a
cell to cell level which may give more data in regards to the two hit hypothesis. As an
extension of this area of research, an investigation of the DNA methylation landscape of
multiple cell lines derived from individual cysts from the same patients is currently being
undertaken in this laboratory. All cell lines contain the same germ-line mutation but may
differ in DNA methylation landscapes.
Our results in this study indicate that *PKD1* has a hypomethylated promoter but areas of both hypomethylation and hypermethylation in the gene body in ADPKD samples. Further data in the form of more gene coverage and more samples may further support this conclusion. This research contributes to the existing understanding of the role methylation plays in pathogenesis of ADPKD and identifies areas of further investigation in this area.
Chapter 7: Bio-distribution of C3-G12 in mice

Targeted delivery molecules have been developed for a wide range of diseases and organs to effectively deliver a therapeutic agent to the correct tissue with higher concentration and less off-target side-effects than treatment with the drug alone.

To identify the bio-distribution and kidney specific targeting ability of the C3-G12 peptide in a mice kidney, liver, lungs, heart and spleen, the fluorescent marker Cy3.5 was conjugated to the peptide by our collaborators (see methods section 2.9) and injected into mice by tail-vein injection. Current literature suggested that the C3-G12 peptide locates and is accumulated in the kidneys of mice and this research was performed to confirm this and show that the delivery of the peptide is possible in our hands before future projects involving the injection of peptide-drug conjugates.

7.1 Comparative fluorescence standard curve

Pilot experiments suggested that the Cy3.5 free dye and the C3-G12-Cy3.5 peptide conjugate had different fluorescence levels at the same concentration. To examine this and normalise the fluorescence between both solutions, a standard curve was created for the free dye.
**Figure 7.1:** Cy3.5 free dye quantity versus fluorescence standard curve. Each dot represents averaged fluorescence readings over two to six replicates. (Cy3.5 - 590 nm excitation/ 620 nm emission)

The fluorescence was measured for seven dilutions of the free dye (1000 – 100 pmol) and was compared to the averaged fluorescence data to two dilutions of C3-G12-Cy3.5.

The average fluorescence of the peptide-dye at 1000 pmol (202.5) matched the average fluorescence of the free dye at 100 pmol (206), therefore 1/10 of the concentration of the peptide was used for free dye injections into mice.

### 7.2 Free dye vs Peptide-Dye conjugate IV administration into mice

Mice were injected with 1 pmol of free dye or 10 pmol of C3-G12-Cy3.5 pmol conjugate, administered by tail-vein injection and were culled at 3, 15 or 30 minutes post injection. Kidneys, heart, liver, spleen and lungs were removed and imaged. The levels of fluorescence are summarised in Table 7.1 based on the measured fluorescence intensity of each organ (Mean Gray Value as determine by ImageJ, section 2.12), with kidney images shown in Figure 7.2 and the remaining images in Appendix 7.
Table 7.1: Relative fluorescence of injected mouse organs

<table>
<thead>
<tr>
<th>Timepoint</th>
<th>Kidney</th>
<th>Liver</th>
<th>Spleen</th>
<th>Heart</th>
<th>Lungs</th>
<th>Kidney</th>
<th>Liver</th>
<th>Spleen</th>
<th>Heart</th>
<th>Lungs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cy3.5 free dye</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3-G12-Cy3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Intense fluorescence (>20 Mean Gray Value)
- Average fluorescence (10-20 Mean Gray Value)
- Small amount of fluorescence can be seen (<10 Mean Gray Value)

Figure 7.2: Fluorescent images of representative mouse kidneys after tail-vein injection with Cy3.5 (1 nmol) or C3-G12-Cy3.5 (10 nmol). Mice (~11 months) were culled at 3 minutes (three mice for each solution, n=3), 15 minutes (n=2) and 30 minute (n=2) time points. Image exposure settings are tabled in Section 2.12.

Vehicle control animals injected with DMSO and normal saline (NS) showed no fluorescence and background fluorescence was not visible under any length of exposure. Cy3.5 free dye showed slight fluorescence at the 3 minute time point in the kidney (Figure 7.2) but this was undetectable at the 15 time point, although very slight fluorescence was seen in one mouse liver and lungs at the 30 minute time point. Overall, this suggests that the free dye may be mostly cleared from the mouse system between 3 and 15 minutes and is not retained in
kidneys. The C3-G12-Cy3.5 peptide conjugate showed intense fluorescence at all time points in the kidney, with slight fluorescence identified in the liver and lungs at the 3 minute time point, eventually becoming barely detectable at 30 minutes. This suggest that the C3-G12-Cy3.5 conjugate localises mainly to the kidneys and is retained there for at least 30 mins, with potentially a small amount retained in the liver and lungs, although this may be cleared from the system after 30 mins. No fluorescence was detected in the heart and spleen after any injections or time points.
Chapter 8: Discussion for Bio-distribution of C3-G12 in mice

8.1 C3-G12 peptide bio-distribution

To investigate the possibility for kidney targeted therapies to treat ADPKD, the kidney specific properties of the C3-G12 peptide was examined in mice. We found that the C3-G12 peptide conjugated to the fluorescent marker Cy3.5 accumulated in the mouse kidney and was maintained there for at least 30 minutes after IV injection.

The original intention of this research was to use the FX-Pro in vivo small animal imager to examine the quantity and localisation of the C3-G12-Cy3.5 peptide conjugate inside live mice throughout multiple time points. All organs would then be able to be viewed simultaneously, and fluorescence levels quantified. Unfortunately the FX-Pro was in need of repair and after considerable effort on my behalf and that of department staff with the company involved, the FX-Pro was unable to be fixed. Instead, the accumulation and specificity of the C3-G12-Cy3.5 peptide conjugate in the mouse kidney, liver, lungs, heart and spleen was examined semi-quantitatively using an epi-fluorescence microscopy. We concluded that the peptide conjugate appeared to be retained in the kidney and accumulated there for at least 30 minutes, as evidenced by the fluorescent intensity of the peptide conjugate when compared with the free dye injections. The liver, lung, spleen and heart were also dissected and imaged at each time point. The liver showed slight fluorescence following Cy3.5 and C3-G12-Cy3.5 injections at three minutes, with fluorescence not detected at 15 minutes. However there was a small amount of fluorescence detected following the Cy3.5 injection in the 30 minute time point for the liver and lungs. This is unlikely to be technical inconsistency as all mice for the 15 and 30 minute time points were injected, culled, dissected
and the organs imaged within 30 minutes of each other. Potentially the mice at the 30 minute
time point metabolised the Cy3.5 dye or C3-G12-Cy3.5 peptide conjugate slower or the
fluorescence in these organs was minimal and towards the lower threshold for detection and
simply wasn’t detected at the 15 minute time point.

The retainment of the C3-G12 peptide in the kidney reconfirms what was previously
published by Geng et al., (2012)\textsuperscript{154}, where they identified the C3-C12 peptide conjugated to
the green fluorescent marker FITC was retained in mouse kidneys for up to an hour after IV
injection. They detected FITC fluorescence after IV injection of their peptide conjugate in
five organs (kidney, liver, lung, spleen, heart) after 3 minutes post IV injection. In our study,
after injection with free Cy3.5 by IV injection, a small amount of fluorescence was detected
in the liver and lungs after the 3 minute time point but not in the heart and spleen. This may
be due to differences in background fluorescence in the FITC and Cy3.5 molecules in mice.

While our data suggests that the C3-G12 peptide is accumulated in the mouse kidneys,
further analysis on the bio-localisation at the cellular level is required for confirmation.
Kidneys that have been preserved from this study could be embedded and sectioned to
identify specific tubule localisation under fluorescent microscopy. Due to the limited amount
of peptide available, time points beyond 30 minutes were not examined. This could be
performed in the future to determine the length of accumulation of the C3-G12 peptide in the
kidney.

A key issue with the whole body treatment of patients with renal disease therapies, is that the
high concentrations are often required, and this may introduce considerable side effects\textsuperscript{128}. The C3-G12 peptide could be conjugated to a methylation altering drug (or other treatment
for ADPKD), allowing targeted delivery to the kidney. This would then be tested in suitable
animal models to examine the effects of the drug on the cystic and healthy kidneys, and any
such side effects. Potentially the C3-G12 peptide could be conjugated to existing methylation altering drugs 5-Aza-2deoxycytidine and 5-Azacytidine or HDAC inhibitors that have been shown to reduce or slow cystogenesis and delivered into an existing conditional or inducible mouse model (section 1.6, table 1.1) that exhibits a disease phenotype similar to the human condition.

Delivery molecules are currently being developed for use in a variety or organs and diseases including tumour-targeted cancer therapies, blood-brain barrier crossing molecules for use in diseases such as Alzheimer’s, colon-targeted molecules for the likes of inflammatory bowel disease, amongst many others. The most effective delivery molecule would be capable of efficient delivery of a wide range of conjugated drugs to the organ or disease of interest without accumulation elsewhere in the body and with a half-life that retains the molecule for the desired amount of time without any toxicity associated with the delivery molecule itself. Several different targeted delivery mechanisms have been developed specifically for kidney and renal diseases, including Pro-drugs, Nanoparticles, Liposomess and Peptides, the latter which have the benefit of little to no toxicity in the body although enzymatic activity of the digestive system limits its use in oral delivery.

This research has shown that in our hands the C3-G12 peptide does accumulate in the mouse kidney and is retained there for at least 30 minutes. This combined with the Geng et al., (2012) lab data suggests that the C3-G12 peptide may be suitable for the purpose of targeted delivery of conjugated therapies to the kidneys in mice, although further time points, whole body imaging of injected mice and other model species, investigation of the biodistribution of the peptide at the cellular level would be required to further characterise the effectiveness of the C3-G12 peptide as a targeted delivery mechanism for polycystic kidney disease.
Chapter 9: Final Discussion

9.1 Final Summary Discussion

Data on the methylation landscapes of the mouse *Pkd1* and human *PKD1* genes is lacking in the current literature and evidence for methylation as a mechanism for the proposed “second hit” to the *PKD1* gene during ADPKD pathogenesis has not previously been investigated. Epigenetic changes have the potential to be a mechanism for this second hit and published literature has already established links between histone regulation changes and ADPKD. However, little research exists on the impact of DNA methylation on the pathogenesis of ADPKD. If DNA methylation is in fact a key mechanism in the second hit, there is possibility for a targeted therapy acting on methylation to restore *PKD1* function, and possibly slow cystogenesis.

The long term goal of this research is to target epigenetic changes in the kidney to treat ADPKD. Epigenetic modifying drugs such as the histone deacetylase (HDAC) inhibitors valproic acid (VPA) and trichostatin A (TSA) which have both been used to effectively reduce cyst growth and renal function loss in mouse models. Concerns over the potential side-effects of whole body treatment of epigenetic modifying drugs suggests that the use of a kidney-specific targeted delivery molecule would be beneficial in the treatment of ADPKD.

This research aimed to investigate DNA methylation in the mouse *Pkd1* gene body region in both *Pkd1 del34* heterozygotes and wildtype mice. This was combined with previous data on a region in the *Pkd1* promotor to determine if any methylation differences may contribute to the theorised epigenetic “second hit” in ADPKD pathogenesis. To compare the mouse *Pkd1* DNA methylation against the human *PKD1* methylation landscape and extend the limited human data available would also provide support for the identification for any methylation-related second hits. Finally, the future goal of this project would be to incorporate the DNA
methylated data with an epigenetic altering therapy specifically targeted to the kidney. To examine the kidney specificity of a proposed targeted delivery peptide, C3-G12, the peptide was conjugated to a fluorescent marker and was injected into mice. Mice were sacrificed at a series of time points and the fluorescence of specific organs were viewed by fluorescent microscopy.

The first part of this project utilised bisulfite treatment and cloning techniques to isolate a selected gene body region of the mouse Pkd1 gene in both del34 heterozygotes and wildtype mice, to identify if there was any difference in the level of methylation that may contribute to cystogenesis, as any methylation difference identified could be targeted for use in epigenetic modifying therapies that would be developed using the mouse model. Combined with data on a region in the Pkd1 promotor, we found there was no difference between the level of methylation in either regions in wildtype and del34 heterozygous mice. This is different to what was identified in human samples by a paper by Woo et al., (2014) where they used pyrosequencing techniques to identify an increase in exon 43 methylation in tissue from ADPKD kidneys which correlated to a decrease in PKD1 expression. Expression analysis of our mouse samples revealed consistently low expression of Pkd1 in wildtype samples, and higher expression in heterozygous. If methylation in the mouse Pkd1 gene is altering the gene expression level, it may be located elsewhere in the gene. To further investigate this, differences may be more easily identified by examining individual cyst lining epithelial cells and the rest of the CpGs in the Pkd1 gene could be examined by whole-genome bisulfite sequencing. If methylation differences do exist between the del34 mouse model and the human disease, then future therapies developed using this mouse model and epigenetic modifying drugs may not interact with the human disease in a similar manner.

The current published literature on the human PKD1 DNA methylation landscape is lacking, and comparisons between the DNA methylation of the human PKD1 and mouse Pkd1 genes
could indicate if DNA methylation is in fact a conserved mechanism for ADPKD pathogenesis. To examine the DNA methylation landscape of the human \textit{PKD1} gene for comparison to the only published data currently available, the reduced representation bisulfite sequencing (RRBS) technique was used to gain a representative sample of the CpG island regions in the human genome. Several methylation differences were identified in \textit{PKD1} in ADPKD tissue, including separate regions of hypomethylation and regions of hypermethylation. This is in contrast to the paper published by Woo \textit{et al.}, (2014)\textsuperscript{104}, which reported increased gene body methylation genome-wide and in the \textit{PKD1} gene. Potentially regions of hypomethylation were missed by Woo \textit{et al.} when simply analysing whole gene trends. These areas may have functional importance relating to gene expression, although this would have to be investigated by expression analysis. Differences shown between our data and that of Woo \textit{et al.} may be due to differences in technique. To analyse the DNA methylation of the entire genome, Woo \textit{et al.}, used methylated-CpG island recovery assay (MIRA), which utilises the affinity of the MBD2/MBD3L1 complex to methylated cytosines combined with micro array to produce data on CpGs throughout the genome\textsuperscript{186}. A global methylation increase was described in ADPKD samples and individual genes (section 5.1.5) were presented as whole methylation differences. In contrast, we used the RRBS technique that uses the MspI restriction enzyme which splits the DNA at CCGG motifs and when combined with size selection, produces a CpG island enriched subset of the whole genome data. While neither methods appear to produce any methylation bias, there is potential that differences seen in our data are due to methylation information that was absent in the others dataset. Our data also identified an increase in DNA methylation in exon 42 of \textit{PKD1} in the ADPKD samples compared to wildtypes. This may support the finding by Woo \textit{et al.}, where they identified an increase in DNA methylation in exon 43 using bisulfite pyrosequencing targeted at this region and a region in the promotor for which they found no difference. Gene
body DNA methylation has been shown to be important for stable gene expression by suppressing disrupters of transcription such as repetitive sequences and transposons, as well as regulating enhancer and silencer sequences\textsuperscript{88,94-96}. If the regions of hypomethylation and hypermethylation in the \textit{PKD1} gene in ADPKD samples are in fact functional and involved in the suppression of transcription disrupters or enhancers and silencers, then potentially DNA methylation may impact on the pathogenesis of ADPKD and may contribute to the second hit, indicating that disease may not be influenced purely by overall hypo- and hypermethylation but a more specific mixture of both. Addition of more ADPKD samples and genome-wide bisulfite sequencing would be advantageous to further investigate this possibility.

To determine the suitability for targeted drug delivery into a mouse model, the C3-G12 peptide, was conjugated to the fluorescent probe Cy3.5 and was injected at multiple time points into Balb/c mice. By comparison to control injections containing only free dye, it appears the peptide-dye conjugate accumulates in the mouse kidney and is retained there for at least 30 minutes. It does not appear to accumulate in any large quantity in the other organs examined. This result confirms kidney-specificity in our hands and further supports the conclusions by existing literature, where the C3-G12 peptide may be suitable candidate for use as a targeting kidney therapy delivery system. Investigation into where the peptide is accumulated on a cellular level is important, as cystogenesis forms in lumen-lining epithelial cells predominantly in the distal tubule and collecting duct and these would need to be targeted in the therapy. Identification of where the C3-G12 peptide is being absorbed could be analysed by use of techniques such as multi-photon microscopy.

Confirming the kidney specificity of the C3-G12 peptide reinforces its potential as a targeted delivery molecule that can be conjugated to a number of drugs for treatment of ADPKD including methylation modifiers 5-Aza-2deoxycytidine and 5-Azacytidine and drugs such as
Octreotide\textsuperscript{187} or Tolvaptan\textsuperscript{188} that are currently being validated for future clinical trials\textsuperscript{189}.

The use of a targeted delivery molecule allows a higher concentration to be administered directly to the diseased tissue and side-effects from whole body treatment can be avoided\textsuperscript{128}.

Multiple targeted delivery molecules are currently being researched (introduction section 1.11), and while peptide based delivery systems offer some advantages, namely inherent non-toxicity in the body and higher efficiency than other methods, they also have a relatively short half-life which can decrease the longevity of the treatment in the body and the enzymatic activity of the digestive system may disrupt the peptide, limiting its effectiveness by oral delivery\textsuperscript{152}.

This research project provides the base work for a future project that would incorporate the methylation information in ADPKD with the goal to conjugate an epigenetic-altering drug to the C3-G12 peptide and selectively deliver it to the kidneys of an ADPKD mouse model and look for epigenetic modifications that may be altered by the drug and if this has any effect on the disease phenotype, with the long-term goal of development of a human therapy.

9.2 Limitations of this study

There are several ways this study could have been improved. Firstly, issues with cloning efficiency and bacterial growth limited the quantity of \textit{Pkd1} gene body sequences obtained from the cloning work. Multiple improvements were made to the protocol to attempt to improve transformation efficiency of the DH5\textalpha{} bacteria and the quality of sequence data. A similar issue was also encountered while cloning the promotor region of the \textit{Pkd1} during my Post Graduate Diploma research although bacteria cloned with the gene body region appeared to grow better with more colonies than the promotor region fragments. This may be due to an inherent toxicity of the cloned \textit{Pkd1} promotor region, or the improved protocol did produce a
more efficient and effective cloning method for the gene body fragments. Increasing the number of quality sequences would allow a larger sample size for a more powerful statistical analysis. A more efficient and streamlined protocol would have allowed time for the examination of other CpG islands in the mouse Pkd1 gene.

The del34 mouse model is one of many animal models of human ADPKD. Like all Pkd1 knockout mouse models available, the del34 mouse is embryonically lethal when homozygous for the del34 disruption but the heterozygous mouse does not develop the same number of cysts or loss of renal function that the human disease exhibits\textsuperscript{156}. Other mouse models exist that contain an inducible second hit to the wildtype Pkd1 allele to more rapidly and accurately represent the pathology of the human disease and are useful for the investigation of pathogenesis after a somatic second hit\textsuperscript{57}. These mice would be useful for studying the effect of a tissue targeted methylation inhibitor in slowing cyst formation. However, they do not reflect the genetics behind the pathogenesis of human ADPKD and may not be informative on the epigenetic mechanisms that may facilitate a spontaneous second hit in somatic cells. Del34 mice are one of several Pkd1 heterozygous knockout models that all exhibit similar pathologies. Although they do not accurately reflect cyst formation in human ADPKD, better models of the human disease for the purpose of investigations into the development of somatic second hits do not currently exist. Despite its limitations, the del34 could be a suitable model for the purposes of this study if cyst-lining cells can be analysed and is still potentially informative for the investigation of the ADPKD second hit mechanism by comparison to the human disease.

A paper by the Fusco et al., (2015)\textsuperscript{190} has identified methylation of non-CpG cytosines in some tissue types, resulting in a bias in methylated primer design, as most primer design software assume that non CpG cytosines are inherently not methylated. The primers used for this research were designed on that assumption that the non-CpG cytosines in the binding
region were not methylated. The level of methylated non-CpG cytosines is tissue-specific and has not been determined in kidney tissue, therefore it is possible that a bias was introduced as the primers may not have bound to template DNA strands with non-CpG cytosine methylation and therefore not amplified those strands. This may be worth investigating in future.

The tissues used for the RRBS analysis were taken from whole kidney tissues from age matched patients. The two cystic samples were considerably cystic and fibrous with no ability to characterise what section of the kidney they originally came from. Additional ADPKD tissue samples are difficult to obtain as kidneys are not usually removed prior to renal transplantation and biopsies are not performed. Furthermore, due to its high expense, kidney samples are not usually genotyped for \textit{PKD1} or \textit{PKD2} mutations. The wildtype tissues were obtained from the cortical region of individuals undergoing renal cancer surgery and were removed from healthy tissue as far away from the tumour as possible. There is potential for methylation differences to exist between different cells within an organ, and differences seen between the samples may in fact reflect different regions of the kidney or different cell types rather than a difference in healthy and cystic tissue. Healthy tissue can be difficult to obtain as healthy kidneys are not usually operated on in a manner that would facilitate removal of tissue. Ideally more ADPKD and healthy tissues would be available for a wider study including larger genome analysis with more coverage of the genome. Cell lines created from single separate cysts from the same patient are able to be analysed to investigate the methylation differences between individual cysts.

In summary, this research has contributed to the current understanding of the pathogenesis of autosomal dominant polycystic kidney disease and provides direction for future treatment development.
Conclusions

The aims of this research were three-fold: to compare methylation in the \textit{Pkd1} gene in both \textit{Pkd1} heterozygotes and wildtypes to determine if any methylation differences between the two may support an epigenetic “second hit” in ADPKD pathogenesis; to extend the limited data available of the human \textit{PKD1} methylation landscape using RRBS and to compare the methylation between ADPKD and wildtype tissues; and to test the ability of the C3-G12 peptide to localise to the kidney in mice using the Cy3.5 fluorescent probe.

I have identified that a region in the promotor of the mouse \textit{Pkd1} is hypomethylated in both wildtype and \textit{Pkd1} heterozygous mice and that a region in the gene body is hypermethylated and shows no significant difference between wildtype and \textit{Pkd1} heterozygous mice, suggesting that if any methylation differences exist, they may be found elsewhere in the \textit{Pkd1} gene. In both human ADPKD and wildtype tissue samples, I found that the \textit{PKD1} promotor is hypomethylated and the gene body is hypermethylated with an increase in methylation identified in exon 42 in ADPKD tissues. This supports a potential role for DNA methylation in ADPKD pathogenesis.

I also observed the accumulation of the C3-G12 peptide in the mouse kidney with retention of over 30 minutes and little to no fluorescence was observed in the heart, liver, lungs and spleen. This suggests that the peptide is able to specificity target the kidney and may be a suitable candidate for kidney-specific drug therapies.

This research provides new data on the mouse and human methylation landscapes in ADPKD and confirms the kidney targeting nature of the C3-G12 peptide in mice.
Acknowledgements

Firstly I would like to acknowledge my PI Dr. Cherie Stayner for being a fantastic supervisor throughout the last two years. I couldn’t ask for a better teacher to guide and instruct me on my post-grad journey. Also thanks to Professor Mike Eccles who gave helpful advice and direction during my project.

I would have been lost many times without the help and knowledge of Dr. Aniruddha Chatterjee.

Special thanks to the rest of the Eccles lab, who I all talked to near death. I really wouldn’t have been able to complete half my experiments without your help.

Of course I need to thank my partner Jemma for being awesome and supportive throughout the ups and downs of the last two years. All that baking certainly helped.

Lunch dates almost every day with Liam, Saturday gym sessions with Sharn and Thursday night golf with the boys certainly helped the slower parts of the year go by.

My family need an acknowledgement for their ongoing support, even if I only ever received blank faces when they asked me what I did on an average day.

All in all I really enjoyed my time as a student researcher in the Eccles lab, there are many people who have helped me get to where I am and I can’t list them all, but know that I am thankful for everything everyone has done for me.

Onwards to another challenging and exciting chapter of my life!
Appendix 1 – Solutions

Supplementary Table 1: Solutions used in this research

<table>
<thead>
<tr>
<th>LB + agar + ampicillin plates (LBAA)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacta-Trypton</td>
<td>10g</td>
</tr>
<tr>
<td>Yeast Extract</td>
<td>5g</td>
</tr>
<tr>
<td>NaCl</td>
<td>10g</td>
</tr>
<tr>
<td>Make up to 1 litre, pH to 7.</td>
<td></td>
</tr>
<tr>
<td>Agar</td>
<td>15g (for 1.5% plates)</td>
</tr>
<tr>
<td>Autoclave solution</td>
<td></td>
</tr>
<tr>
<td>Ampicillin</td>
<td>1 μl per ml (100 mg/ml)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOC media</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>KCl</td>
<td>0.019g</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>0.20g</td>
</tr>
<tr>
<td>MgSO₄</td>
<td>0.25g</td>
</tr>
<tr>
<td>Yeast Extract</td>
<td>0.5g</td>
</tr>
<tr>
<td>NaCl</td>
<td>0.6g</td>
</tr>
<tr>
<td>Tryptone</td>
<td>2.0g</td>
</tr>
<tr>
<td>Make up to 100 ml, autoclave</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10x TAE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TRIS</td>
<td>48.4g</td>
</tr>
<tr>
<td>Glacial acetic acid</td>
<td>11.4 ml</td>
</tr>
<tr>
<td>EDTA</td>
<td>3.7g</td>
</tr>
<tr>
<td>pH to 8, make up to 1 litre.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normal Saline (9%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium Chloride</td>
<td>9 g</td>
</tr>
<tr>
<td>Make up to one litre, autoclave</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2 – TOPO vector

Supplementary Figure 1 - TOPO 4.0 vector by Invitrogen
### Appendix 3 – Primers used throughout study

**Supplementary Table 2 – Primers used in this study**

<table>
<thead>
<tr>
<th></th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MousePROMPkd1F</td>
<td>GGAGTTGGYGGTTGGGTAGTGTAG</td>
</tr>
<tr>
<td>MousePROMPkd1R</td>
<td>CTCCRAACCACAAACCCAAACCTAAA</td>
</tr>
<tr>
<td>MousePkd1EX43F</td>
<td>GTTGGTTGTGAATTTTATTTTTTTATG</td>
</tr>
<tr>
<td>MousePkd1EX43R</td>
<td>AAAACCAACTAAACCCACTAAATC</td>
</tr>
<tr>
<td>muPKD1ex37F (qPCR)</td>
<td>TTGGCAAAGGAGGAAGCTCG</td>
</tr>
<tr>
<td>muPKD1ex38R (qPCR)</td>
<td>TGGGGATGCTTCTTGGCATG</td>
</tr>
<tr>
<td>MouseGAPDHF (qPCR)</td>
<td>TGACCACCAACTGCTTCAGC</td>
</tr>
<tr>
<td>MouseGAPDHR (qPCR)</td>
<td>GGCATGGACTGTGGTCATGA</td>
</tr>
<tr>
<td>MouseHPRTF (qPCR)</td>
<td>GCTTTCCCTGGTAAAGCAGTACA</td>
</tr>
<tr>
<td>MouseHPRTTR (qPCR)</td>
<td>GAGAGCTCTTTCACCCAGCA</td>
</tr>
<tr>
<td>NeoFII (del34 genotyping)</td>
<td>CAGCGCATCGCTTCTATC</td>
</tr>
<tr>
<td>MR50 (del34 genotyping)</td>
<td>CTTTAATCCCTGCACTCAGGA</td>
</tr>
<tr>
<td>TOPO 4.0 M13 Reverse</td>
<td>CAGGAAACAGCTATGAC</td>
</tr>
</tbody>
</table>
Appendix 4 – *Del34* mouse data

Supplementary Table 3 – *Pkd1* study mouse age and genotype. Sample 1314 that was re-genotyped as a heterozygous is italicised.

<table>
<thead>
<tr>
<th>Sample Mouse</th>
<th>Genotype</th>
<th>Age (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1289</td>
<td>Heterozygous</td>
<td>19</td>
</tr>
<tr>
<td>1293</td>
<td>Wildtype</td>
<td>19</td>
</tr>
<tr>
<td>1314</td>
<td><em>Heterozygous</em></td>
<td>18</td>
</tr>
<tr>
<td>1319</td>
<td>Heterozygous</td>
<td>18</td>
</tr>
<tr>
<td>1324</td>
<td>Wildtype</td>
<td>17</td>
</tr>
<tr>
<td>1327</td>
<td>Wildtype</td>
<td>17</td>
</tr>
<tr>
<td>1388</td>
<td>Heterozygous</td>
<td>13</td>
</tr>
<tr>
<td>1406</td>
<td>Wildtype</td>
<td>15</td>
</tr>
</tbody>
</table>
Appendix 5 – Cloned *Pkdl* regions

**Promotor**

GGAGCTGGGCGCTGGGTCTGCACTGCAAGCGGAAATGCGCGAGCACGCCGAGCCCAAGG
CCCTGAGGTCGGCCTGGCCCGAAGGGTGCTAGCAGCTGACCGCAGACGGG
CCACGGCCGCGCAGGCAGGGTCCCAGCCGCCGCCCGGCATTCAGCCCCGTGCACCCACCATTAGTCGCAGGG
CCGAGCTGCGCTGACGATGCCGCTCGGCGCGCCTGCTCTCCTGGCGCTGCTCATGGGCCTCGAG

**Exon 43 (Gene Body)**

GCTGGTGGCCCTATATATTCCCATGGCCGCCAGGTGCAGACCTGGCGTCAAGATGGGTGCTGCTGACAG
TGAGGTGCACAGGAGGCTACCTGGGCACCGGTGCCTACTGGTGTGATAGTGAC
ACAGCATACAGGAGCTGACTAGTCGCCAGCGTGGGCACTTGACCATGATCGCCAGT
GGACCCATTTTGTGCAGGACCATTCCAGCCTTTACCTAGCTTTGGACGCAAGTGCC
TCAGCTGGGCTCT

All CpG locations are highlighted
Appendix 6 – Enlarged regions of human *PKD1* RRBS data

Supplementary Figure 2 – Close up of IGV methylation differences in human *PKD1*. ADPKD07 (1), ADPKD08 (2), WILDTYPE E1 (3), WILDTYPE G2 (4) and WILDTYPE H3 (5) RRBS data was aligned and visualised against the human reference HG19 *PKD1* track. Red indicates a methylated CpG, blue unmethylated. **A** – ADPKD samples are more hypermethylated in *PKD1* at exon 42 (genomic location of image ch16 2,141,238 to 2,142,226). **B** - ADPKD samples are more hypermethylated in *PKD1* between exons 33 and 34 (ch16 2,146,209 to 2,148,064). **C** - ADPKD samples are more hypomethylated in *PKD1* between exons 1 and 2 (ch16 2,174,230 to 2,174,293).
Appendix 7 – Fluorescence in mouse organs

Supplementary Figure 3 – Mouse organs dissected from mice sacrificed at three minutes. A – Injected with Free Cy3.5 Liver, B - C3-G12-Cy3.5 Liver and Lungs (C).

Supplementary Figure 4 – Mouse organs dissected from mice sacrificed at fifteen minutes. A – Injected with C3-G12-Cy3.5 Liver.

Supplementary Figure 5 – Mouse organs dissected from mice sacrificed at thirty minutes. A – Injected with Free Cy3.5 Liver and Lungs (B), C - C3-G12-Cy3.5 Liver and Lungs (D).
References


Irazabal, M. V. & Torres, V. E. Experimental therapies and ongoing clinical trials to slow progression of ADPKD. Curr Hypertens Rev 9, 44-59 (2013).


