



# Extensive Inter-Cyst DNA Methylation Variation in Autosomal Dominant Polycystic Kidney Disease Revealed by Genome Scale Sequencing

Sarah A. Bowden<sup>1</sup>, Peter A. Stockwell<sup>1</sup>, Euan J. Rodger<sup>1,2</sup>, Matthew F. Parry<sup>3</sup>, Michael R. Eccles<sup>1,2</sup>, Cherie Stayner<sup>1\*</sup> and Aniruddha Chatterjee<sup>1,2\*</sup>

## OPEN ACCESS

### Edited by:

Hehuang Xie,  
Virginia Tech, United States

### Reviewed by:

Jong Hoon Park,  
Sookmyung Women's University,  
South Korea  
Shicheng Guo,  
University of Wisconsin–Madison,  
United States

### \*Correspondence:

Cherie Stayner  
cherie.stayner@otago.ac.nz  
Aniruddha Chatterjee  
aniruddha.chatterjee@otago.ac.nz

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

**Received:** 09 January 2020

**Accepted:** 23 March 2020

**Published:** 15 April 2020

### Citation:

Bowden SA, Stockwell PA,  
Rodger EJ, Parry MF, Eccles MR,  
Stayner C and Chatterjee A (2020)  
Extensive Inter-Cyst DNA Methylation  
Variation in Autosomal Dominant  
Polycystic Kidney Disease Revealed  
by Genome Scale Sequencing.  
*Front. Genet.* 11:348.  
doi: 10.3389/fgene.2020.00348

<sup>1</sup> Department of Pathology, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand, <sup>2</sup> Maurice Wilkins Centre for Molecular Biodiscovery, Auckland, New Zealand, <sup>3</sup> Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand

Autosomal dominant polycystic kidney disease (ADPKD) is a heritable disease characterized by bilateral renal enlargement due to the growth of cysts throughout the kidneys. Inheritance of a disease-causing mutation is required to develop ADPKD, which results in end-stage kidney disease and is associated with a high morbidity. The pathology underlying cyst formation is not well understood. To address this, we have previously shown the global methylome is altered in ADPKD tissue, suggesting a role of DNA methylation in disease-state renal tissue. As cysts are believed to arise independently, we hypothesize that DNA methylation changes vary accordingly. Here we further investigate the role of DNA methylation within independent cysts to characterize key intra-individual changes. We demonstrate that fragments within CpG islands and gene bodies harbor the greatest amount of variation across the ADPKD kidney, while intergenic fragments are comparatively stable. A proportion of variably methylated genes were also differentially methylated in ADPKD tissue. Our data provide evidence that individual molecular mechanisms are operating in the development of each cyst.

**Keywords:** polycystic kidney, DNA methylation, autosomal dominant polycystic kidney disease, reduced representation bisulfite sequencing, epigenetics

## INTRODUCTION

Autosomal dominant polycystic kidney disease (ADPKD) is characterized by the bilateral accumulation of fluid-filled cysts within the kidneys. Ultimately, disease progression results in kidney enlargement, leading to end-stage kidney disease typically by the sixth decade (Braun, 2009). While ADPKD is the most common heritable renal disease in humans, affecting

an estimated 1–5 in 10,000 patients (Solazzo et al., 2018), treatment options for this disease are limited as there is still little understanding about the underlying cause of cystogenesis (Tan et al., 2011; Reed-Gitomer, 2014).

The presence of a disease-causing mutation is required to develop ADPKD, however, a causal relationship between these mutations and the process of cystogenesis is not yet clearly elucidated. PKD-causing mutations are primarily found within the *PKD1* (85%) and *PKD2* (15%) genes. However, in atypical presentations of ADPKD, disease-causing mutations have also been reported in *GANAB* (Porath et al., 2016) and *DNAJB11* (Cornec-Le Gall et al., 2018). It is generally accepted that each ADPKD cyst arises by means of focal proliferation (Reed-Gitomer, 2014), where a single cell will acquire altered growth characteristics, rapidly proliferate and form a distinct cyst.

DNA methylation is required for the regulation of gene expression (Herman and Baylin, 2003; Yang et al., 2014), and therefore tissue-specific methylation must be conserved in order for the normal functioning of genes (Tang et al., 2016). Studies on DNA methylation have thus traditionally investigated differential methylation between two groups. However, in recent years, the study of variation in the methylome has become of increasing interest.

The identification of CpG sites which do not necessarily show differential methylation between populations, but instead exhibit DNA methylation variation across a population has been suggested as a response to the cellular microenvironment (Hansen et al., 2011) or chemical exposure (Phillips et al., 2019). Furthermore, DNA methylation variation has been observed amongst individuals with autoimmune diseases (Paul et al., 2016; Webster et al., 2018) and intra-individually in chronic lymphocytic leukemia samples (Landau et al., 2014).

Cysts are thought to arise independently, either through a second “hit” or somatic mutation in a disease-causing gene (Kim and Park, 2016), or by other mechanisms that modify gene expression, such as epigenetic changes (Woo, 2016). Previous work on DNA methylation in ADPKD has been conducted at bulk-tissue level (Woo et al., 2014; Bowden et al., 2018). As bulk tissue samples will likely contain cells from multiple cysts, as well as stromal and parenchymal cells, the methylation profile will reflect the average of all cell types and therefore may not accurately represent the profile of epithelial cells that line the cyst. To address this, we have performed reduced representation bisulfite sequencing (RRBS), targeting single cyst walls to sequence discreet populations of enriched clonal cystic epithelia and to validate the data collected from bulk tissue analyses. To further describe the role of DNA methylation in cystogenesis, we decided to investigate the methylomes of individual ADPKD cysts from a single ADPKD patient in order to validate the whole tissue data. We hypothesize that each cyst arises independently, and therefore carries its own methylome, which would demonstrate that individual molecular mechanisms are operating in each cyst.

Initial investigation led us to discover discrete sites of variation across a single patient kidney, which we have characterized. The regions in which we have observed this variation suggest that each ADPKD cyst is responding independently to cellular stress.

## METHODS

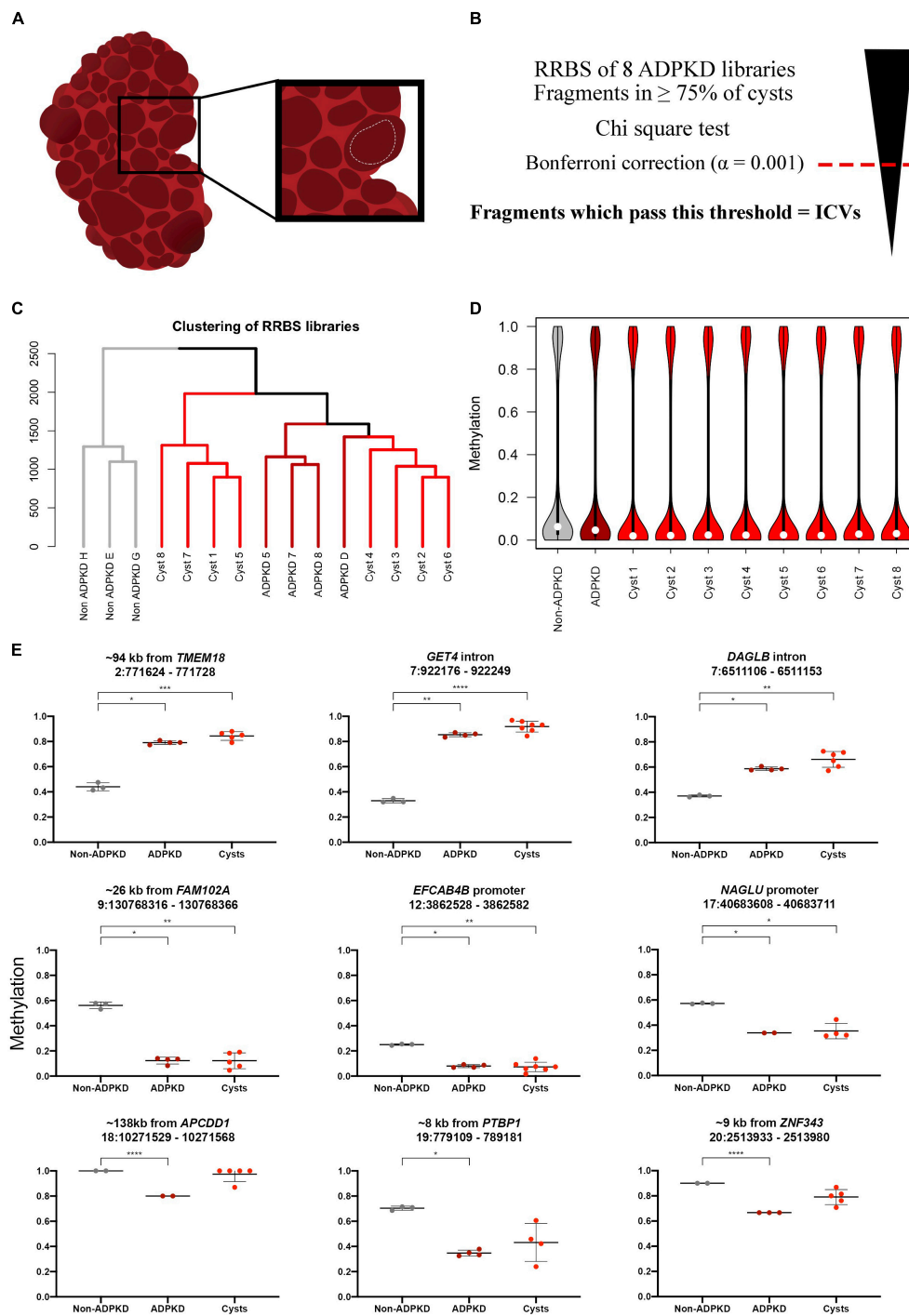
### Cyst Dissection From a Single ADPKD Kidney

Three non-ADPKD cortical kidney tissue samples were provided from the Christchurch Cancer Society Tissue Bank, as previously described (Bowden et al., 2018). These samples were collected from the renal cortex during nephrectomy for renal cell carcinoma, with histological testing to confirm the absence of malignant tissue. Eight single cyst samples were macrodissected from a single ADPKD kidney (male, 60 years old, *PKD1* mutation c.3744\_3745dup (het), NM\_001009944.2). Following nephrectomy, the ADPKD kidney was stored in neutral buffered formalin for 7 days before we were able to collect it. The cyst samples were macrodissected from various sites throughout the kidney by cutting sections of the cyst wall, approximately 1 cm<sup>2</sup> in size, with surgical scissors (**Figure 1A**). An effort was made to ensure these samples were from a single cyst; this means samples were selected from the largest intact cysts available. Tissue was washed with 70% EtOH to remove non-cyst cells (i.e., blood). The sections of cyst were then submerged in fresh 70% EtOH and stored at 4°C prior to DNA isolation. The collection and use of renal tissue was approved by the Otago University Human Ethics Committee H15/110.

### Generating Cyst Methylomes Using RRBS

DNA methylation was assessed using RRBS. RRBS library generation of the non-ADPKD renal tissue has previously been described [data accessible on the NCBI GEO database, accession GSE110057 (Bowden et al., 2018)] and is based on previously published methodologies (Chatterjee et al., 2012a, 2013, 2017a). The generation of the cyst libraries followed the same protocol as the whole renal tissue, with modifications made for the limited tissue availability and formalin fixation. Briefly, genomic DNA was extracted from the cyst samples using the QIAamp DNA FFPE Mini Kit (Qiagen), excluding the paraffin melting steps, with an extended 16-h Proteinase K treatment at 56°C. An overnight MspI digest was performed, followed by end-repair, A-tailing and ligation of methylated adapters to the resulting fragments (Illumina). Bisulfite conversion was performed using the MethylCode DNA Methylation Kit (Invitrogen).

Libraries were amplified with PCR (15–20 cycles) and run through 3% NuSieve GTG Agarose gels (Lonza) to select for fragments 150–330 bp post-adaptor ligation. Libraries were quantified with the Qubit fluorometer (Life Technologies) and assessed for quality on the 2100 Bioanalyzer using the High Sensitivity DNA chip (Agilent Technologies) according to manufacturer's instructions. Libraries with high amounts of adaptor contamination (identified as peaks just under 150 bp on the Bioanalyzer trace) were re-size selected with AMPure XP Bead Purification (New England Biolabs) before repeating Qubit and Bioanalyzer steps. RRBS libraries were sequenced in an Illumina HiSeq 2500 to produce single-ended 100 bp reads. Base calling was performed with Illumina Real Time Analyzer software.



**FIGURE 1 |** Individual ADPKD cysts show the same methylation pattern as whole tissue ADPKD samples globally and at specific fragments. **(A)** Fluid-filled cysts were visually identified throughout the patient kidney following nephrectomy. Sections of the cyst wall were cut out with surgical scissors (white dotted line), creating sections of tissue approximately  $1 \text{ cm}^2$ , with the intention to avoid neighboring tissue. **(B)** ICVs were classified by first performing a Chi-square test on the analyzed fragments. The Bonferroni correction ( $\alpha = 0.001$ ) was used to select the most variable fragments from this analysis; fragments with  $p$ -values below this threshold were classed as ICVs. **(C)** Clustering of the non-ADPKD ( $n = 3$ ), ADPKD whole tissue ( $n = 4$ ) and ADPKD cyst ( $n = 8$ ) RRBS libraries over all common analyzed fragments ( $n = 39,708$ ) reveals non-ADPKD clusters separately to ADPKD tissue. Cyst samples are clustered into two distinct branches. **(D)** Distribution of global methylation in each RRBS library. White dot denotes median methylation. Analysis of common analyzed fragments across all samples shows that cysts are globally hypomethylated compared to non-ADPKD tissue. Notably, cysts are more hypomethylated than the whole tissue ADPKD counterpart. **(E)** Nine of the previously identified DMFs had coverage in  $\geq 50\%$  of the ADPKD cysts (three with  $< 50\%$  coverage in **Supplementary Figure S1**), and six of these showed a statistically significant difference from non-ADPKD tissue as determined by ANOVA. Data from ADPKD tissue previously published (Bowden et al., 2018), FDR-adjusted  $p$ -values: \*  $< 0.05$ , \*\*  $< 0.01$ , \*\*\*  $< 0.001$ , \*\*\*\*  $< 0.0001$ .

## Processing of RRBS Methylome Sequencing Data

Quality of the sequenced reads was assessed using the FastQC<sup>1</sup> program. Adapter sequences and methylated CpG bases at the 3' end of the reads that were added during the end-repair step of the library preparation were removed using our in-house *cleanadaptors* program as part of the DMAP package (Chatterjee et al., 2012b). The sequenced reads were aligned to the human genome GRCh37 with Bismark v0.6.4 alignment tools (Krueger and Andrews, 2011). We performed stringent mapping and allowed for only one mismatch in the seed (the first 28 bp of the read).

The bioinformatic software program DMAP (Stockwell et al., 2014) was used to generate MspI fragment-based methylomes and perform analysis on the sequenced reads of each sample. A non-ADPKD reference library was made by merging the three non-ADPKD tissue libraries with the program SamTools (Li et al., 2009). Fragments were only included in analyses if they satisfied our coverage criteria (contained at least two CpGs and were covered by 10 or more sequenced reads). We have demonstrated the utility of this analysis approach previously (Chatterjee et al., 2017b).

## Analysis of the Global Methylome

Direct comparison between the previous whole tissue work and the cyst libraries was performed by generating a list of fragments common in all 10 RRBS libraries (non-ADPKD renal tissue, ADPKD whole tissue, eight ADPKD cysts), in order to prevent any bias due to sequencing coverage. Previously identified differentially methylated fragments (DMFs) were assessed in the eight ADPKD cysts by performing an *F*-test analysis of variance (ANOVA) which allowed us to assess the significance in difference between methylation in the two groups at these specific fragments. Clustering analysis was performed using unsupervised hierarchical clustering in R, using the “complete” method.

Cyst-specific DMFs were identified in each cyst by performing Fisher's exact test between each cyst library and the non-ADPKD cortical tissue reference library we generated. DMFs were selected using the Bonferroni correction for multiple testing ( $p = \alpha/n$ ;  $\alpha = 0.001$ , cut off values in **Supplementary Table S1**), and an absolute difference in methylation greater than or equal to 0.25 (i.e., 25%) between ADPKD and non-ADPKD tissue. Allosomes were excluded from the analyses between multiple individuals due to sex differences between the two groups. This is different to the ANOVA test above as this is identifying changes specific to the single cyst rather than within between two groups of whole tissue which contain multiple cysts.

## Identification of Variable Inter-Cyst Fragments

To identify inter-cyst variation, a Chi-square test was performed with  $n - 2$  degrees of freedom (where  $n$  is the number of individual cysts investigated). This resulted in the analysis of 45,954 fragments (referred to as “total analyzed fragments”).

To control for multiple testing, the Bonferroni correction was applied with a stringent significance value of 0.001; the adjusted cut off  $p$ -value was  $2.18 \times 10^{-8}$ . Fragments with  $p$ -values below this threshold were deemed inter-cyst variants (ICVs) (**Figure 1B**). An overlap between ICVs and DMFs common in  $\geq 50\%$  of the cysts identified variably DMFs. The cut off was 50% of cysts to account for some of the variably methylated fragments having the same methylation value as the non-ADPKD library.

## Genomic Characterization of Fragments

The *identgeneloc* program of the DMAP package was used to label fragments based on the distance to the transcription start site (TSS) of a gene. Intragenic fragments were subclassified as promoters (between 5 kb upstream and 1 kb downstream of the TSS), and those within the gene body ( $>1$  kb downstream of the TSS) were labeled as intron, exon, and intron/exon boundary fragments by *identgeneloc*. All fragments not within this defined intragenic region were classed as intergenic fragments. *Identgeneloc* also allows for fragments to be classified based on their CpG feature as annotated by SeqMonk<sup>2</sup>; that is, whether they are in the core, shore or shelf of a CpG island (CGI).

Gene ontology was performed using Metascape (Zhou et al., 2019). Gene lists (**Supplementary Tables S3–S7**) were analyzed alongside a list of background genes ( $n = 17,917$ ), which was generated from a list of all sequenced fragments in the nine RRBS libraries. A list of genes ( $n = 2,415$ ) associated with intragenic ICVs ( $n = 4,204$ ) was analyzed to characterize pathways associated with ICVs, as intragenic fragments were more likely to have consequential effects on gene transcription.

Genes were ranked with a variation score, to indicate the susceptibility of the gene to DNA methylation variation. This was calculated by dividing the number of intragenic ICVs associated with the gene, divided by the total number of intragenic fragments associated with the gene in the analysis. Genes were required to contain at least three ICVs in order to be included in the analysis as this conferred a reasonable amount of certainty about the amount of variation in the gene. Fragments with a variation score of 1.0 had total variation in all sequenced fragments in the analysis.

## Statistics

To identify inter-cyst variation, a Chi-square test was performed with  $n - 2$  degrees of freedom (where  $n$  is the number of cysts investigated). To control for multiple testing, the Bonferroni correction with a stringent significance value of 0.001 was applied to each analysis to determine the  $p$ -value threshold. The statistical significance for other analyses (such as the proportion of ICVs in the chromosome and genomic features) were calculated with a Chi-square test with Yate's correction, where  $p < 0.001$  was the minimum value for significance. To plot the distribution of  $p$ -values, QQ-plots were generated using R (**Supplementary Figure S6**).

<sup>1</sup><http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

<sup>2</sup><https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/>



## RESULTS

### Collection of ADPKD Single Cysts

Cysts were visually confirmed to be throughout the kidney upon dissection. As the kidney had been cut up for formalin fixation many had drained, but some still contained cyst fluid. The larger of the cysts were collected in order to extract as much DNA for analysis as possible. These had thicker membranes and were more distinct from the stroma. As the kidney was partially dissected when we had access to it, we could not determine the spatial organization of the cysts in the kidney.

### ADPKD Cysts Resemble ADPKD Tissue at the Global Methylome Level

We carried out genome-wide bisulfite sequencing on eight single kidney cyst samples from an ADPKD patient using RRBS. We used this data alongside previously described whole renal tissue RRBS data from non-ADPKD and ADPKD kidneys (Bowden et al., 2018). We first wanted to compare global methylation data from the eight ADPKD cysts to ADPKD whole tissue samples and non-ADPKD samples which we had previously generated. The non-ADPKD samples clearly clustered separately from the cysts and ADPKD tissue samples (Figure 1C). Next, we identified the global change in median methylation in each cyst in comparison to non-ADPKD and ADPKD whole tissue samples using only the common analyzed fragments, where it was observed that all eight ADPKD cysts were hypomethylated by 3–4% in comparison to our non-ADPKD reference genome (Figure 1D and Supplementary Table S2). This is consistent with our whole tissue data which showed a 2% reduction in median methylation in ADPKD compared to non-ADPKD. When compared to non-ADPKD using the same common analyzed fragments, the whole tissue ADPKD samples were only hypomethylated by 1.6%.

We then sought to validate the previously identified DMFs in each cyst to confirm that these changes are universal in each ADPKD cyst. There was coverage in  $\geq 50\%$  of the cysts within nine of the whole tissue DMFs, with statistically significant differences between non-ADPKD tissue and ADPKD cysts in six DMFs (Figure 1E).

### ADPKD Cysts Are Differentially Methylated Across Developmental Pathways

Differential methylation has largely been the focus of DNA methylation analyses, since such changes in the promoters of genes may confer stable changes in gene expression (Jones, 2012; Chatterjee et al., 2018). For this reason, we decided to identify DMFs within each ADPKD cyst to determine whether key changes occurred within each individual cyst. An analysis was performed between each ADPKD cyst and the non-ADPKD methylome using Fisher's exact test. The number of DMFs identified in cysts (Supplementary Table S1) ranged from 474 (in Cyst 4) to 6,087 (in Cyst 1). Of these fragments, 96 common DMFs, which all demonstrated the same direction of methylation change, were identified in all eight

cysts. The common DMFs were slightly enriched within the gene body of protein-coding genes (54%; Figure 2A) and were enriched for pathways associated with animal organ maturation as well as cell-cell adhesion and Notch signaling (Figure 2B and Supplementary Table S9).

Of our previously identified whole tissue DMFs (Bowden et al., 2018), the *TMEM18*- and *GET4*-associated fragments were also differentially methylated in some of the single cysts according to the stricter criteria (Supplementary Figure S2).

### Inter-Cyst DNA Methylation Variation Is Enriched in Intragenic Regions

As illustrated by comparison to our DMF data, there was a significant amount of variability across the eight cysts. To quantify this further, we selected fragments present in at least six of the eight cysts (to account for lower methylome coverage in some cysts), which resulted in 45,954 total analyzed fragments. Fragments displaying significant inter-cyst variation were identified by applying a Chi-square goodness of fit test, followed by the Bonferroni adjustment for multiple test correction at a significance level of 0.001 (adjusted  $p$ -value cut off =  $2.18 \times 10^{-8}$ ). This analysis identified 6,727 fragments (14.6% of the total analyzed fragments) which were termed "inter-cyst variants" (ICVs).

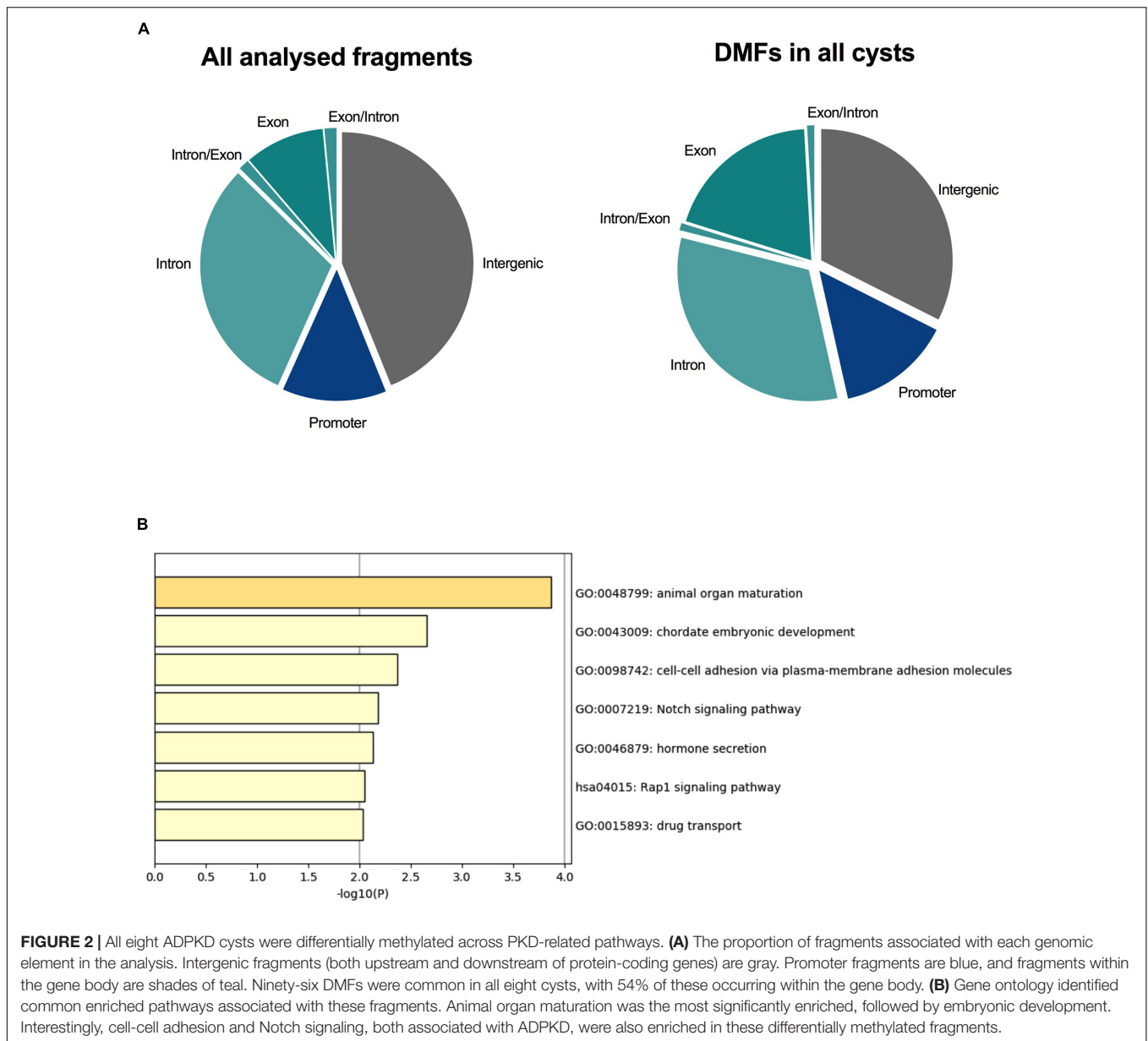
Next we characterized the ICVs based on their chromosomal and genomic location. We found that the proportion of ICVs in two chromosomes (4 and 17) was significantly different to the proportion of total analyzed fragments by  $\geq 12\%$  ( $p = 0.001$ , Figure 3A). When ICVs were characterized by relation to CGIs, ICVs were shown to be more strongly enriched in fragments within the core of the CGI (Figure 3B). The proportion of ICVs increased in intragenic fragments and conversely decreased in intergenic fragments (Figure 3C).

Genes were ranked with a variation score, which was generated based on how many fragments associated with each gene were variably methylated, in order to determine the genes with the most variable methylation states. A total of 18 genes (Figure 3D and Supplementary Figure S3) had a variation score of 1.0 (at least three fragments associated with the gene, all classed as ICVs), and were functionally enriched for the gene ontology processes sulfur compound metabolic processes and GTPase regulation (Supplementary Figure S4 and Supplementary Table S8).

When the null hypotheses are true, the distribution of  $p$ -values is uniform on the interval (0, 1). The histogram of actual  $p$ -values (Supplementary Figure S5) shows a pronounced skew toward values near zero. The QQ-plot (Supplementary Figure S6) displays the same information: observed quantiles lie below the quantiles of the uniform distribution (indicated by the 45-degree line).

### Genes Involved in Developmental Pathways Are Enriched for DNA Methylation Variation

In this analysis there were 4,204 intragenic ICV fragments, associated with 2,415 unique genes. We investigated the



functional pathways of this group of genes, as methylation at these fragments is historically considered more likely to impact the gene function than methylation within intergenic fragments (Klose and Bird, 2006).

Gene ontology analysis through Metascape revealed the intragenic ICV-associated genes were enriched for a number of developmental and morphogenesis pathways (**Supplementary Figure S7** and **Supplementary Table S10**; intergenic ICV-associated gene ontology data in **Supplementary Table S11**).

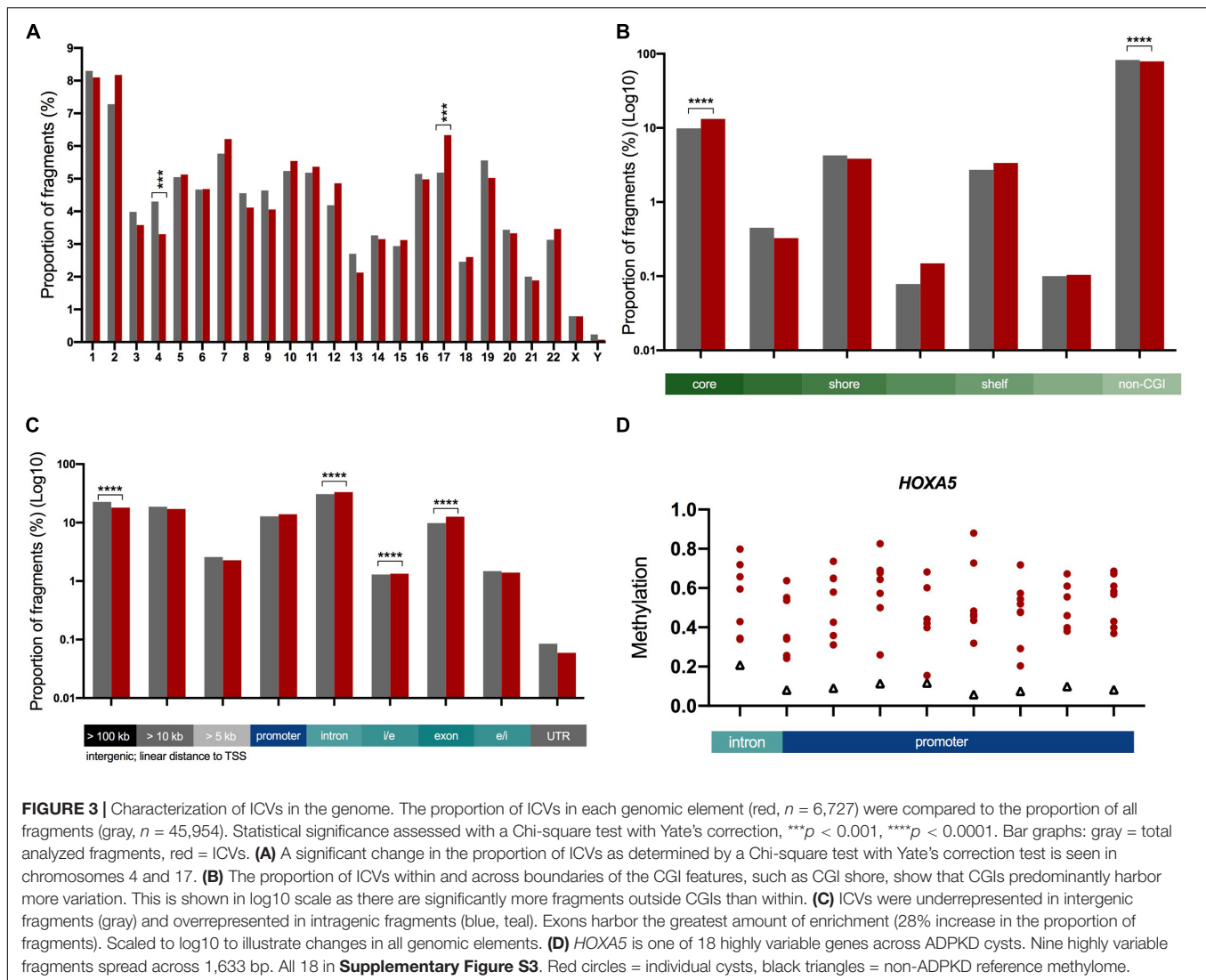
### Variably Methylated Fragments Are Also Differentially Methylated in ADPKD

As there were differences in the number of ICVs identified in each cyst, we aimed to investigate if there was an overlap between

DMFs and ICVs, i.e., were fragments which showed variable methylation measurably different to the non-ADPKD tissue.

A total of 2,204 fragments were differentially methylated in at least 50% of the cysts. 837 of these DMFs shared an overlap with the 6,727 ICVs (**Figure 4A**), occurring in 586 unique genes. These fragments which overlapped with DMFs and ICVs could be said to be variable DMFs. Variable DMFs were predominantly located within the gene body of protein-coding genes (**Figure 4B**), which were functionally enriched for development pathways (**Figure 4C**).

We previously identified 13 DMFs in ADPKD tissue and have now identified *TMEM18*- and *GET4*-associated differential methylation also occurs individually in single cysts. These DMFs did not demonstrate variable methylation.



## DISCUSSION

The role of DNA methylation in ADPKD is an emerging field of study and following our recent identification of site-specific DNA methylation changes in cystic renal tissue, we sought to investigate if these changes are seen consistently between individual cysts in a single patient. We have identified that in addition to whole tissue DMFs being present in most single cysts, there are regions of heterogeneity in the DNA methylation of cysts across a single individual ADPKD patient.

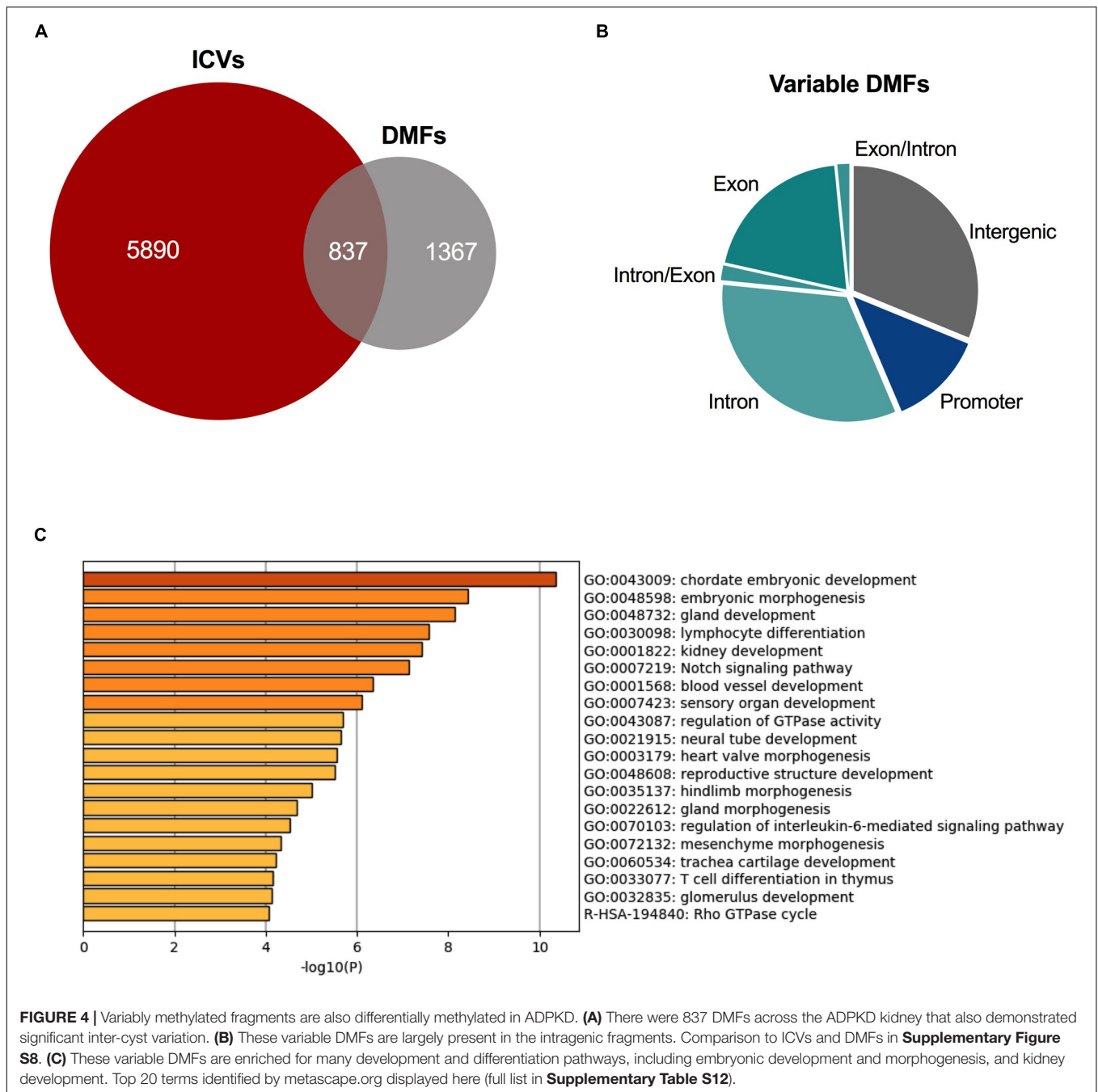
The cysts in this analysis were large, and visually confirmed to be cysts before microdissection took place. Due to the limited availability of tissue, no IHC or protein analysis was possible to confirm the phenotypic nature of the samples.

Using our criteria for assessing variation, 14.6% of the fragments in our analysis were identified as ICVs. ICVs represent areas of the genome with marked methylation variability across  $\geq 75\%$  of the cysts, suggesting possible instability of these particular sites or elements in ADPKD. ICVs are overrepresented

across intragenic fragments and are conversely underrepresented in the intergenic fragments. The enrichment of ICVs appears to overlap with transcriptional activity of the genome, as methylated CpGs associated with gene promoters (and sometimes gene bodies) have been shown to influence the transcriptional activity of genes (Herman and Baylin, 2003; Yang et al., 2014). Potentially, the variation of DNA methylation in the cysts could be associated with transcriptional activity of the genome.

DNA methylation variation was enriched in the core of CGIs. CGIs are regions with an increased GC density and thus higher CpG content, and are considered to be transcriptional regulators, much like gene promoters (Deaton and Bird, 2011; Jones, 2012). Therefore, the variation changes we see in the CGI cores likely act in the same manner as intragenic variation.

Interestingly, previous data generated from an intra-individual analysis of chronic lymphocytic leukemia identified a similar pattern of variation (described as proportion of discordant reads) across the genome (Landau et al., 2014).



In this study, DNA methylation variation at the promoter was associated with reduced mean gene expression, but increased gene expression variation.

The intragenic ICVs suggest that the pathways with the most significant variation in methylation tend to target developmental pathways. This may be indicative of stochastic variation in these cysts, where cysts are in a transitional state between embryonic and developed tissue. Developmental changes are a feature of ADPKD cells, with embryonic transcription factors such as *PAX2* being persistently expressed in cystic epithelia (Stayner et al., 2006),

and polarization defects such as the mis-localization of membrane bound proteins such as EGFR and the  $\text{Na}^+\text{K}^+\text{ATPase}$ , more closely resembling a fetal cellular distribution (Wilson et al., 1991).

There were 18 genes which showed total variation in our analysis (**Supplementary Figure S3**), representing the most contiguous regions of variable methylation in the cyst population. These variable genes are functionally associated with sulfur metabolism pathways and GTPase function, although an association between these data and pathogenesis is not yet clear. Several papers have attributed DNA methylation variation to



stochastic acquisition (Landau et al., 2014; Franzen et al., 2017), which may better explain the distribution of variance observed in a single ADPKD kidney.

One possible explanation for the observed DNA methylation variation is DNA damage repair (DDR) occurring within the genome. When DNA suffers damage such as double stranded breaks (DSBs), not only must the DNA undergo repair, but also the associated epigenetic marks must also be replaced (Benayoun et al., 2015). By inducing DSBs in stem cells, Cuozzo et al. (2007) demonstrated that corresponding *de novo* methylation occurs in cells with reduced DNA homologous recombination mechanisms. Another group demonstrated that DSBs initiated the onset of CGIs promoter methylation, aberrantly silencing gene expression (O'Hagan et al., 2008).

In ciliopathies, the relationship between cellular stress and the role of DDR pathways is becoming increasingly apparent (Johnson and Collis, 2016; Zhang et al., 2019). Largely identified in nephronophthisis, mutations within DDR genes such as *ZNF423*, *CEP164*, *NME3*, and *AATF* are sufficient to cause disease via ciliary dysfunction (Chaki et al., 2012; Hoff et al., 2018; Jain et al., 2019). Additionally, *FAN1* inactivation triggers chronic kidney disease in zebrafish via a similar mechanism (Zhou et al., 2012).

Our single cyst data clustered into two separate groups based on an analysis of common analyzed fragments with unsupervised hierarchical clustering. One reason for this may be that the cysts originated from different regions of the nephron (i.e., proximal and distal tubule) which may each have unique methylome profiles. Further analysis of single cyst variation was assessed by Principal Component Analysis (PCA) (data not shown) which could not sufficiently explain the total variation seen across the ADPKD kidney by the subgrouping of cysts.

We have previously identified DMFs across bulk sections of ADPKD tissue, which contained multiple cysts. The whole tissue DMFs were (mostly) differentially methylated in the ADPKD cyst group. We did not identify these fragments as ICVs in the cyst population, which indicates that the differential methylation of ADPKD tissue is a distinct biological phenomenon to the variation of DNA methylation across an ADPKD kidney. This also supports our previous finding that dysregulation of methylation at these sites in the genome is required for ADPKD development (and thus variable methylation cannot occur here).

There are several lines of enquiry needed to follow up this study. Analyses of single cysts from additional patients would help determine the ubiquity of the variable regions we have already identified or show if they are patient specific. Single cell analyses of normal kidney would also provide suitable controls with which to validate these intriguing findings.

Cell-specific analysis of renal tubules may also help elucidate whether variation and subgrouping seen in the cyst samples is due to differences in the cell of origin or if these changes are the result of cysts following two different patterns of methylation changes during cystogenesis.

We have found that the DMF data previously generated from whole ADPKD tissue is largely reflected in the methylomes of individual ADPKD cysts. In addition to this, we found variable methylation across cysts arising

from a single patient, supporting the hypothesis that there are also individual molecular mechanisms operating in the development of each cyst.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the NCBI Gene Expression Omnibus <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE142664>. The additional datasets analyzed in this study can be found in the NCBI Gene Expression Omnibus <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110057>.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Otago University Human Ethics Committee H15/110. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

SB carried out the experimental work, performed the data analyses and wrote the first draft of the manuscript. ER and PS contributed to the data analysis. MP contributed to the statistical analysis. ME provided the critical revision and contributed in the design of the project and supervised the progression. CS conceived the idea and obtained the funding for this project. AC and CS designed and supervised the study, wrote the manuscript together with SB. AC contributed in the analysis framework of the study together with SB. All authors contributed to manuscript revision, read and approved the submitted version.

## FUNDING

Funding for this research was supplied by the PKD Foundation of Australia and the Maurice and Phyllis Paykel Trust.

## ACKNOWLEDGMENTS

The authors would like to thank all patients who provided material for the study, along with nephrologists and hospital staff who facilitated the tissue collection. AC would like to acknowledge the support of a Rutherford Discovery Fellowship (from the Royal Society of New Zealand).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00348/full#supplementary-material>

## REFERENCES

- Benayoun, B. A., Pollina, E. A., and Brunet, A. (2015). Epigenetic regulation of ageing: linking environmental inputs to genomic stability. *Nat. Rev. Mol. Cell Biol.* 16, 593–610. doi: 10.1038/nrm4048
- Bowden, S. A., Rodger, E. J., Bates, M., Chatterjee, A., Eccles, M. R., and Stayner, C. (2018). Genome-scale single nucleotide resolution analysis of DNA methylation in human autosomal dominant polycystic kidney disease. *Am. J. Nephrol.* 48, 415–424. doi: 10.1159/000494739
- Braun, W. E. (2009). Autosomal dominant polycystic kidney disease: emerging concepts of pathogenesis and new treatments. *Cleve Clin. J. Med.* 76, 97–104. doi: 10.3949/ccjm.76a.gr001
- Chaki, M., Airik, R., Ghosh, A. K., Giles, R. H., Chen, R., Slaats, G. G., et al. (2012). Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling. *Cell* 150, 533–548. doi: 10.1016/j.cell.2012.06.028
- Chatterjee, A., Ozaki, Y., Stockwell, P. A., Horsfield, J. A., Morison, I. M., and Nakagawa, S. (2013). Mapping the zebrafish brain methylome using reduced representation bisulfite sequencing. *Epigenetics* 8, 979–989. doi: 10.4161/epi.25797
- Chatterjee, A., Rodger, E. J., and Eccles, M. R. (2018). Epigenetic drivers of tumorigenesis and cancer metastasis. *Semin. Cancer Biol.* 51, 149–159. doi: 10.1016/j.semcancer.2017.08.004
- Chatterjee, A., Rodger, E. J., Stockwell, P. A., Le Mée, G., and Morison, I. M. (2017a). Generating multiple base-resolution DNA methylomes using reduced representation bisulfite sequencing. *Methods Mol. Biol.* 1537, 279–298. doi: 10.1007/978-1-4939-6685-1\_16
- Chatterjee, A., Rodger, E. J., Stockwell, P. A., Weeks, R. J., and Morison, I. M. (2012a). Technical considerations for reduced representation bisulfite sequencing with multiplexed libraries. *J. Biomed. Biotechnol.* 2012:741542. doi: 10.1155/2012/741542
- Chatterjee, A., Stockwell, P. A., Ahn, A., Rodger, E. J., Leichter, A. L., and Eccles, M. R. (2017b). Genome-wide methylation sequencing of paired primary and metastatic cell lines identifies common DNA methylation changes and a role for EBF3 as a candidate epigenetic driver of melanoma metastasis. *Oncotarget* 8, 6085–6101. doi: 10.18632/oncotarget.14042
- Chatterjee, A., Stockwell, P. A., Rodger, E. J., and Morison, I. M. (2012b). Comparison of alignment software for genome-wide bisulphite sequence data. *Nucleic Acids Res.* 40:e79. doi: 10.1093/nar/gks150
- Cornec-Le Gall, E., Olson, R. J., Besse, W., Heyer, C. M., Gainullin, V. G., Smith, J. M., et al. (2018). Monoallelic mutations to DNAJB11 cause atypical autosomal-dominant polycystic kidney disease. *Am. J. Hum. Genet.* 102, 832–844. doi: 10.1016/j.ajhg.2018.03.013
- Cuozzo, C., Porcellini, A., Angrisano, T., Morano, A., Lee, B., Di Pardo, A., et al. (2007). DNA damage, homology-directed repair, and DNA methylation. *PLoS Genet.* 3:e110. doi: 10.1371/journal.pgen.0030110
- Deaton, A. M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* 25, 1010–1022. doi: 10.1101/gad.2037511
- Franzen, J., Zirkel, A., Blake, J., Rath, B., Benes, V., Papanonis, A., et al. (2017). Senescence-associated DNA methylation is stochastically acquired in subpopulations of mesenchymal stem cells. *Ageing cell* 16, 183–191. doi: 10.1111/acel.12544
- Hansen, K. D., Timp, W., Bravo, H. C., Sabuncian, S., Langmead, B., McDonald, O. G., et al. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* 43, 768–775. doi: 10.1038/ng.865
- Herman, J. G., and Baylin, S. B. (2003). Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.* 349, 2042–2054. doi: 10.1056/nejmra023075
- Hoff, S., Epting, D., Falk, N., Schroda, S., Braun, D. A., Halbritter, J., et al. (2018). The nucleoside-diphosphate kinase NME3 associates with nephronophthisis proteins and is required for ciliary function during renal development. *J. Biol. Chem.* 293, 15243–15255. doi: 10.1074/jbc.RA117.000847
- Jain, M., Kaiser, R. W. J., Bohl, K., Hoehne, M., Göbel, H., Bartram, M. P., et al. (2019). Inactivation of apoptosis antagonizing transcription factor in tubular epithelial cells induces accumulation of DNA damage and nephronophthisis. *Kidney Int.* 95, 846–858. doi: 10.1016/j.kint.2018.10.034
- Johnson, C. A., and Collis, S. J. (2016). Ciliogenesis and the DNA damage response: a stressful relationship. *Cilia* 5:19. doi: 10.1186/s13630-016-0040-6
- Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13, 484–492. doi: 10.1038/nrg3230
- Kim, D. Y., and Park, J. H. (2016). Genetic mechanisms of ADPKD. *Adv. Exp. Med. Biol.* 933, 13–22. doi: 10.1007/978-981-10-2041-4\_2
- Klose, R. J., and Bird, A. P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* 31, 89–97. doi: 10.1016/j.tibs.2005.12.008
- Krueger, F., and Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572. doi: 10.1093/bioinformatics/btr167
- Landau, D. A., Clement, K., Ziller, M. J., Boyle, P., Fan, J., Gu, H., et al. (2014). Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* 26, 813–825. doi: 10.1016/j.ccell.2014.10.012
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- O'Hagan, H. M., Mohammad, H. P., and Baylin, S. B. (2008). Double strand breaks can initiate gene silencing and SIRT1-dependent onset of DNA methylation in an exogenous promoter CpG island. *PLoS Genet.* 4:e1000155. doi: 10.1371/journal.pgen.1000155
- Paul, D. S., Teschendorff, A. E., Dang, M. A. N., Lowe, R., Hawa, M. I., Ecker, S., et al. (2016). Increased DNA methylation variability in type 1 diabetes across three immune effector cell types. *Nat. Commun.* 7:13555. doi: 10.1038/ncomms13555
- Phillips, R. V., Rieswijk, L., Hubbard, A. E., Vermeulen, R., Zhang, J., Hu, W., et al. (2019). Human exposure to trichloroethylene is associated with increased variability of blood DNA methylation that is enriched in genes and pathways related to autoimmune disease and cancer. *Epigenetics* 14, 1112–1124. doi: 10.1080/15592294.2019.1633866
- Porath, B., Gainullin, V. G., Cornec-Le Gall, E., Dillinger, E. K., Heyer, C. M., Hopp, K., et al. (2016). Mutations in GANAB, encoding the glucosidase iialpha subunit, cause autosomal-dominant polycystic kidney and liver disease. *Am. J. Hum. Genet.* 98, 1193–1207. doi: 10.1016/j.ajhg.2016.05.004
- Reed-Gitomer, B. (2014). Autosomal dominant polycystic kidney disease: genetics, epidemiology, and treatment. *Adv. Genomics Genet.* 4, 173–183.
- Solazzo, A., Testa, F., Giovannella, S., Busutti, M., Furci, L., Carrera, P., et al. (2018). The prevalence of autosomal dominant polycystic kidney disease (ADPKD): a meta-analysis of European literature and prevalence evaluation in the Italian province of Modena suggest that ADPKD is a rare and underdiagnosed condition. *PLoS One* 13:e0190430. doi: 10.1371/journal.pone.0190430
- Stayner, C., Iglesias, D. M., Goodyer, P. R., Ellis, L., Germino, G., Zhou, J., et al. (2006). Pax2 gene dosage influences cystogenesis in autosomal dominant polycystic kidney disease. *Hum. Mol. Genet.* 15, 3520–3528. doi: 10.1093/hmg/ddl428
- Stockwell, P. A., Chatterjee, A., Rodger, E. J., and Morison, I. M. (2014). DMAP: differential methylation analysis package for RRBS and WGBS data. *Bioinformatics* 30, 1814–1822. doi: 10.1093/bioinformatics/btu126
- Tan, Y. C., Blumenfeld, J., and Rennert, H. (2011). Autosomal dominant polycystic kidney disease: genetics, mutations and microRNAs. *Biochim. Biophys. Acta* 1812, 1202–1212. doi: 10.1016/j.bbadis.2011.03.002
- Tang, W. W., Kobayashi, T., Irie, N., Dietmann, S., and Surani, M. A. (2016). Specification and epigenetic programming of the human germ line. *Nat. Rev. Genet.* 17, 585–600. doi: 10.1038/nrg.2016.88
- Webster, A. P., Plant, D., Ecker, S., Zufferey, F., Bell, J. T., Feber, A., et al. (2018). Increased DNA methylation variability in rheumatoid arthritis-discordant monozygotic twins. *Genome Med.* 10:64. doi: 10.1186/s13073-018-0575-9
- Wilson, P. D., Sherwood, A. C., Palla, K., Du, J., Watson, R., and Norman, J. T. (1991). Reversed polarity of Na(+)-K(+)-ATPase: mislocation to apical plasma membranes in polycystic kidney disease epithelia. *Am. J. Physiol.* 260, F420–F430.
- Woo, Y. M. (2016). Epigenetic regulation in cystogenesis. *Adv. Exp. Med. Biol.* 933, 59–68. doi: 10.1007/978-981-10-2041-4\_6
- Woo, Y. M., Bae, J. B., Oh, Y. H., Lee, Y. G., Lee, M. J., Park, E. Y., et al. (2014). Genome-wide methylation profiling of ADPKD identified epigenetically regulated genes associated with renal cyst development. *Hum. Genet.* 133, 281–297. doi: 10.1007/s00439-013-1378-0
- Yang, X., Han, H., De Carvalho, D. D., Lay, F. D., Jones, P. A., and Liang, G. (2014). Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell* 26, 577–590. doi: 10.1016/j.ccr.2014.07.028

- Zhang, J. Q. J., Saravanabavan, S., Munt, A., Wong, A. T. Y., Harris, D. C., Harris, P. C., et al. (2019). The role of DNA damage as a therapeutic target in autosomal dominant polycystic kidney disease. *Expert Rev. Mol. Med.* 21:e6. doi: 10.1017/erm.2019.6
- Zhou, W., Otto, E. A., Cluckey, A., Airik, R., Hurd, T. W., Chaki, M., et al. (2012). FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. *Nature Genet.* 44, 910–915. doi: 10.1038/ng.2347
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10:1523.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Bowden, Stockwell, Rodger, Parry, Eccles, Stayner and Chatterjee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.