University of Cyprus
Department of Biological Sciences

The genetic basis of Autosomal Dominant Polycystic Kidney Disease (ADPKD) in two Cypriot Families and the development of a robust *PKD1* and *PKD2* mutation screening test using amplicon-based Next-Generation Sequencing (NGS)

Petrou Avgousta

A Research-Based Master's Thesis
for the
Degree of Magister Scientiae in Biomedical Sciences

May 25, 2021

ABSTRACT

Autosomal dominant polycystic kidney disease (ADPKD) is one of the most common hereditary kidney diseases, with a frequency of 1 in 1000 people in the general population (Sessa et al., 2004, Lanktree et al., 2018). This progressive disorder is mainly caused by mutations in two genes: *PKD1,* accounting for 75-85% of the cases, and *PKD2,* responsible for the rest (Dicks et al., 2006). More types of the disease also exist. ADPKD main features are the formation of cysts in the kidneys and a gradual age-related loss of kidney function; as disease progresses, cysts replace the functional renal parenchyma, ultimately leading to end-stage renal disease (ESRD) in 50% of the cases (Tan et al., 2009).

The purpose of this study was the genetic investigation of two Cypriot families, clinically diagnosed with ADPKD. Candidate *PKD1* and *PKD2* genomic regions were investigated in family CY1611 by linkage analysis, a method based on the analysis of polymorphic genomic markers flanking the *PKD1* and *PKD2* gene loci. Reportedly, family CY1611 demonstrated linkage to the *PKD1* gene locus. Family CY1626 was previously known to be linked to the *PKD2* gene locus, hence *PKD2* gene was directly analyzed by Sanger Sequencing in search for mutations. Variants found by sequencing were assessed by using prediction algorithms and variant databases. No mutations were found in *PKD2* in CY1626 family members explaining their clinical diagnosis.

Screening *PKD2* for mutations by Sanger Sequencing is a feasible task, but unfortunately the genomic characteristics of *PKD1* impede accurate and prompt genetic screening by sequencing, not only because of the high GC content of its exons but also for the existence of multiple homologs in the genome (Bogdanova et al., 2001) and pseudogenes (Rossetti et al., 2001, Tan et al., 2012). To overcome this significant problem with *PKD1* sequencing, we aimed at developing a pipeline were long-range PCR amplification of *PKD1* and *PKD2* exonic regions paired with amplicon-based Next-Generation Sequencing (NGS) can allow the simultaneous screening of both genes by semi-conductor NGS technology.

ACKNOWLEDGEMENTS

I would like to thank Prof. Constantino Delta who gave me the opportunity to carry out my master's thesis to his laboratory, Molecular Medicine of Research Center (MMRC). Also, I would like to warmly thank Dr Papagregoriou Gregory and Dr Polydorou Christiana for the continuous encouragement and support that they offered me and especially for the valuable knowledge they passed to me. Furthermore, my gratitude goes to all the members of the laboratory for the excellent collaboration. It was a pleasure to work with all these people. Finally, I am especially thankful to my family and my friends, for their support and help.

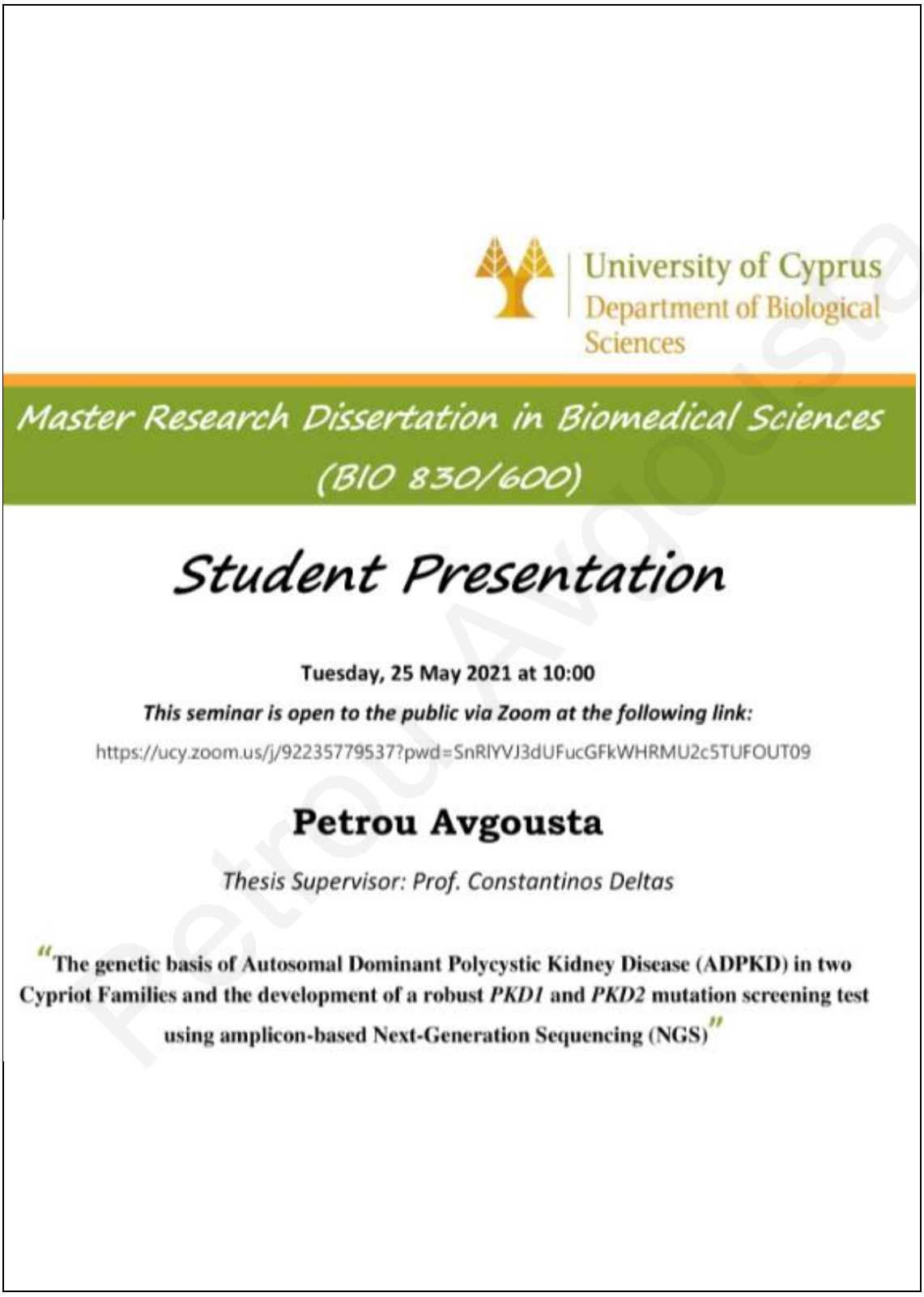
COMPOSITION OF THE EXAMINATION COMMITTEE

Thesis Supervisor (Examination Committee coordinator): Prof. Constantinos Deltas

Committee Member: Assoc. Prof. Pantelis Georgiades

Committee Member: Ass. Prof Chrysoula Pitsouli

SEMINAR ANNOUNCEMENT



University of Cyprus
Department of Biological Sciences

Master Research Dissertation in Biomedical Sciences (BIO 830/600)

Student Presentation

Tuesday, 25 May 2021 at 10:00

This seminar is open to the public via Zoom at the following link:

https://ucy.zoom.us/j/92235779537?pwd=SnRlYVJ3dUFucGFkWHRMU2c5TUFOUT09

Petrou Avgousta

Thesis Supervisor: Prof. Constantinos Deltas

"The genetic basis of Autosomal Dominant Polycystic Kidney Disease (ADPKD) in two Cypriot Families and the development of a robust *PKD1* and *PKD2* mutation screening test using amplicon-based Next-Generation Sequencing (NGS)"

TABLE OF CONTENTS

Contents

INTRODUCTION

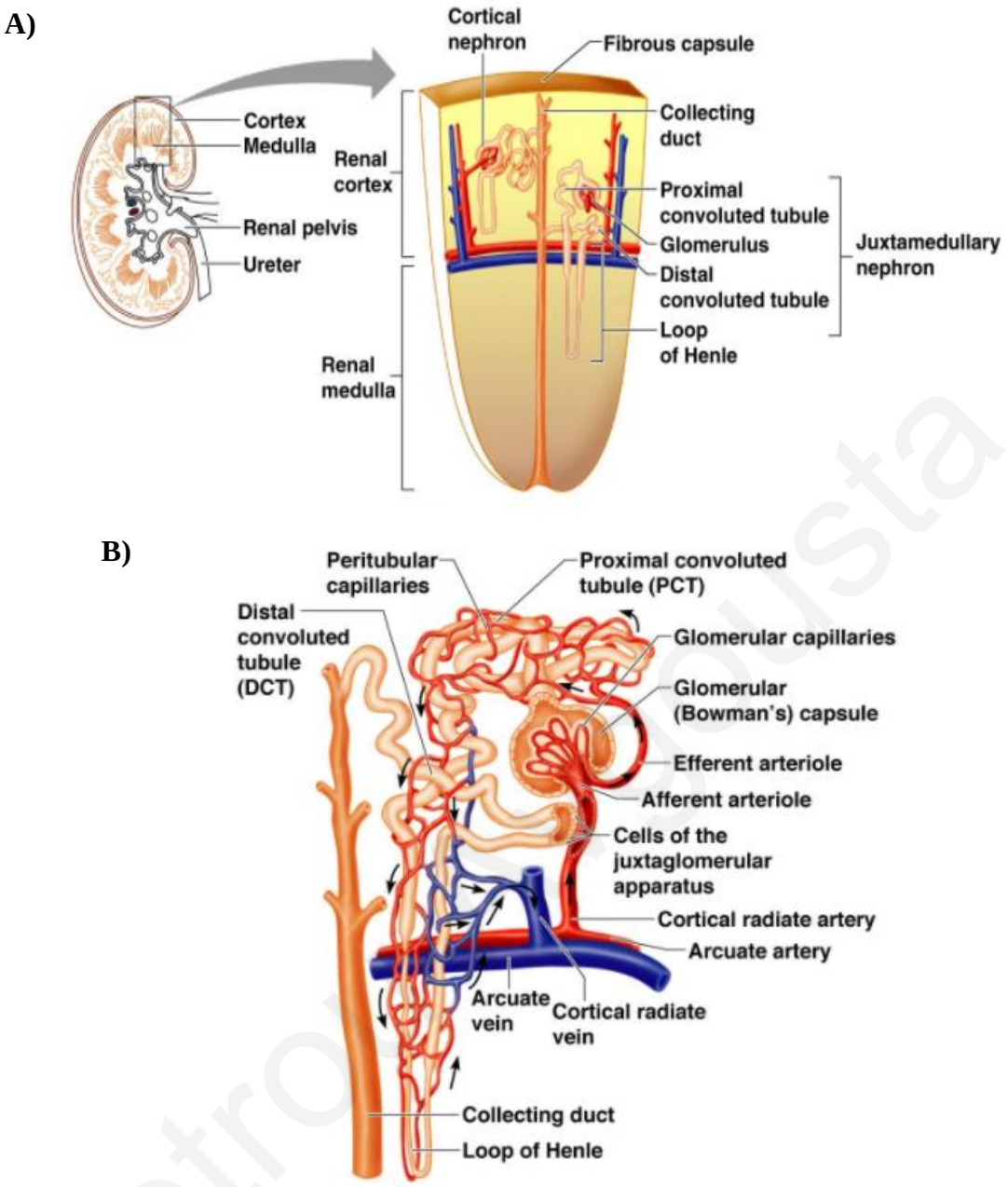
1.1.Basic Principles of Renal Physiology

Kidneys are considered as the most crucial organs in our body because they have a central role in preserving homeostasis. Almost 180 liters of blood are filtered per day by the kidneys to remove metabolic waste products with urine and consequently absorb various substances that are important for our body. Kidneys can adjust water, salt, and pH to maintain the homeostatic balance of tissue fluids (Short et al., 2014). Also, kidneys regulate blood pressure through renin-angiotensin-aldosterone system, circulating calcium and phosphate levels, with the activation of vitamin D and erythrocyte production through erythropoietin induction.

Renal filtering units are composed by over a million blind-ended tubules called nephrons. Each nephron is made up of two parts, the renal corpuscle, and the renal tubule. The first one, is divided into the glomerular capillaries or glomerulus and the Bowman's capsule. The blood is filtered in the renal corpuscle and the arteriole that brings blood into the glomerulus is called the afferent arteriole whereas the artery that takes blood away from the glomerulus is known as the efferent arteriole.

Between these arterioles there is a formation of a network of capillaries called the glomerular capillaries of the glomerulus. The Bowman's capsule is a cup-shaped structure in which this glomerulus is located. The glomerulus along with the Bowman's capsule are responsible for the filtration of blood to form the urine. Furthermore, the renal tubule consists of the proximal convoluted tubule (PCT), the U-shaped loop of Henle and the distal convoluted tubule (DCT).

Once the blood is filtered in the renal corpuscle, the resultant fluid is called the glomerular filtrate, which passes into the PCT. In the PCT, substances like sodium chloride (NaCl), potassium (K+), water, glucose, and bicarbonate are reabsorbed into the filtrate whereas urea, creatinine, uric acid are added to the filtrate. From the PCT, the filtrate enters the Loop of Henle where reabsorption and secretion of water and various metabolites occurs. The filtrate then passes into the DCT. From the DCT, the filtrate passes into the collecting tubules, into the renal pelvis and the ureters as urine to be stored into the bladder **(Figure 1)** (Gordon Betts et al., 2013).



<u>Figure 1. Anatomy of kidneys</u>. A) The left image displays the whole kidney and, the right, presents the nephron. The functional substance or parenchyma, of the kidney is divided into two major structures, the outer renal cortex, and the inner renal medulla. The nephrons, the urine-producing functional structures of the kidney, span to cortex and medulla. Each nephron is made up of two parts, the renal corpuscle, and the renal tubule. The renal corpuscle is divided into the glomerulus and Bowman's capsule. **B)** The structure of a nephron. The arrows show the direction of the blood filtration in the renal corpuscle. The arteriole that brings blood into the glomerulus is called the afferent arteriole whereas the artery that takes blood away from the glomerulus is known as the efferent arteriole. The glomerulus along with the Bowman's capsule achieve the filtration of blood to form urine. The renal tubule consists of the proximal convoluted tubule (PCT), the U-shaped loop of Henle and the distal convoluted tubule (DCT) (Gordon Betts et al., 2013).

1.2.Autosomal Dominant Polycystic Kidney Disease

Autosomal dominant polycystic kidney disease (ADPKD) is one of the most common genetic kidney diseases, with an estimated prevalence of 1 in 1000 people in the general population worldwide (Cornec-Le Gall et al., 2019, Koptides and Deltas, 2000, Lanktree et al., 2018, Willey et al., 2017). ADPKD is almost always inherited from a parent by a defective gene being passed to a child. It is known that it has a dominant pattern of inheritance which means that is essential only one abnormal gene from one parent to cause ADPKD. This disease has high penetrance, which that means the offspring of affected parents have 50% chance of developing the disease (Torra Balcells and Ars Criach, 2011). Even though the matching gene from the other parent is normal, the abnormal gene dominates, so it would be unlikely to clinically skip a generation **(Figure 2)**.

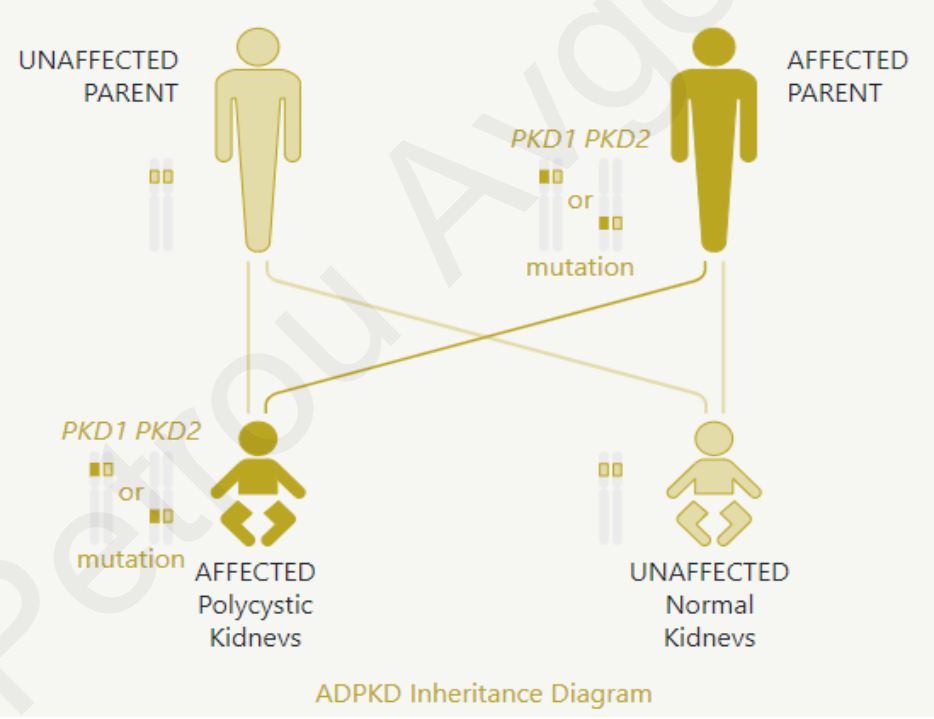


Figure 2. ADPKD inheritance: *PKD1* and *PKD2* genes. If a person has ADPKD, there is 50% probability that he/she will pass it to next generation. If a child inherits a mutated *PKD1* or *PKD2* gene, he/she will have ADPKD too. Children who do not inherit the abnormal gene will not develop or pass on the disease (ADPKD, 2021).

ADPKD is a monogenic multisystemic disorder, and the most frequent symptom is the formation of fluid-filled cysts scattered throughout the kidney parenchyma (Bouba et al., 2001, Cornec-Le Gall et al., 2019). These cysts appeared primarily in the kidneys; however, they can also develop in the liver and pancreas (Reeders, 1992). They grow progressively in size and number with age. The loss of the kidney function takes place over the space of decades and frequently leading to end-stage renal failure (ESRD) in approximately 50% of patients during or after the 60^{th} decade of life (Bouba et al., 2001, Cornec-Le Gall et al., 2019). Interestingly according to the literature every person born with a disease-related mutation finally shows manifestations of cystogenesis (Everson, 1993, Kaehny and Everson, 1991).

Also, the patients with ADPKD appear to also have other symptoms such as hypertension, decreased glomerular filtration rate, acute and chronic pain, cyst infection, hematuria, and nephrolithiasis, which can affect individuals to a varying degree. Except from the direct manifestations of kidney growth and decreased glomerular filtration rate, individuals might show abdominal hernias, intracranial aneurysms, and cardiac valvular lesions (Cornec-Le Gall et al., 2019).

ADPKD is characterized by a phenotypic heterogeneity. It is known that the severity of renal disease, the degree of manifestation of extra-renal symptoms and age of onset vary (Bear et al., 1992, Ravine et al., 1992). ADPKD in clinically well-characterized populations is mainly caused by mutations in the *PKD1* and *PKD2* genes, 75% to 85% and 15% to 25% of the cases, respectively (Rossetti et al., 2001, Peters and Sandkuijl, 1992, Koptides et al., 2000, Tan et al., 2009, Cornec-Le Gall et al., 2019, Hughes et al., 1995). A third candidate gene, *PKD3*, has been suggested by some researches to be implicated in ADPKD (Daoust et al., 1995) and disputed by others (Paterson and Pei, 1999). As the mutations in *PKD1* were shown to be related with a severe clinical presentation and less prognosis than those in *PKD2*, these two genetic factors are very important and helpful for predicting outcomes of patients with ADPKD (Hateboer et al., 1999, Rossetti et al., 2007). The phenotypic spectrum of this disease ranges from severe prenatal causes to elderly patients with normal kidney function (Torra Balcells and Ars Criach, 2011).

1.3.Genetics of ADPKD

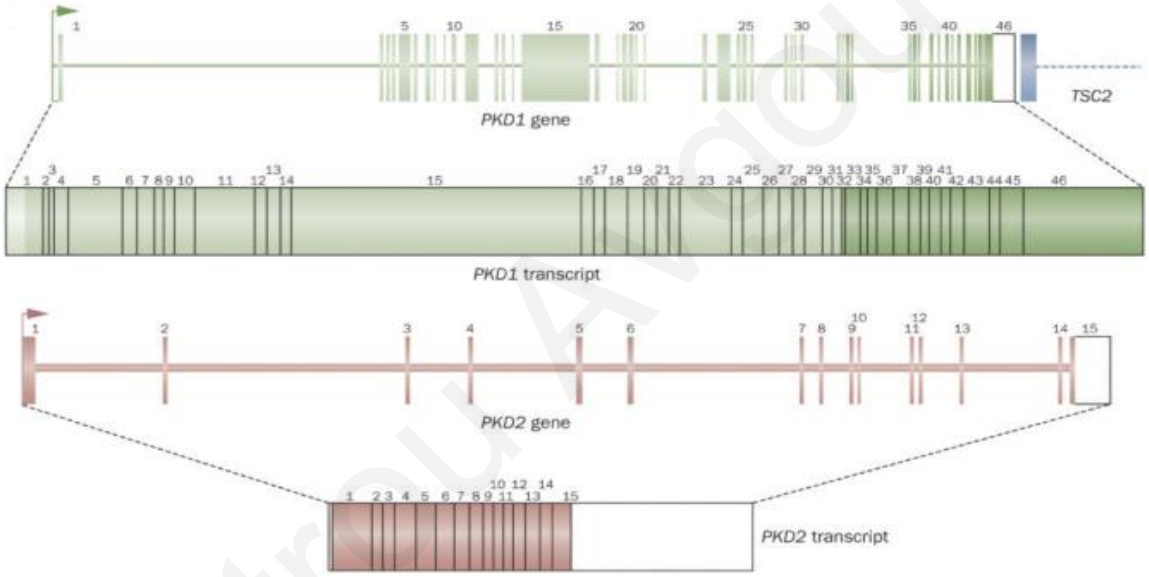
Autosomal dominant polycystic kidney disease (ADPKD) is inherited and is characterized by genetic heterogeneity (Bear et al., 1992, Ravine et al., 1992), because it is caused by mutations in two different genes, *PKD1* and *PKD2* (Daoust et al., 1995, Hughes et al., 1995, Mochizuki et al., 1996, Cornec-Le Gall et al., 2019, Koptides and Deltas, 2000). Also, there is a recessive form of polycystic kidney disease (ARPKD) which is caused by mutations in PKHD gene. However, is rarer than ADPKD with an estimated prevalence of 1 in 20.000 individuals in the general population (Guay-Woodford et al., 2014).

1.3.1. *PKD1* gene

The *PKD1* gene is located on chromosome 16 at position 16p13.3 and encodes polycystin-1 (PC1) (Peters and Sandkuijl, 1992, Hughes et al., 1995). This gene consists of 46 exons, with the largest one being exon 15 of approximately 3.2kb in size **(Figure 3)**. The genomic sequence of *PKD1* gene span 52kb in size and is transcribed into a 14kb mRNA which is translated into one of the largest proteins composed of 4,302 amino acids (Hughes et al., 1995, Germino, 1997, Koptides and Deltas, 2000). It has GC rich regions and a large part of *PKD1* gene, more specific exons 1-33 in the 5'-end, are duplicated in at least six homologous genes (HGs) which have approximately 95% similarity to *PKD1* and are transcribed into mRNA (Mochizuki et al., 1996). In addition to this, these pseudogenes (PKD1P1-P6), are expressed but have early stop codons preventing translation and produce small non-functional proteins (Torra Balcells and Ars Criach, 2011). Furthermore, the coding strands of *PKD1* gene consists of 95% thymine (T) and cytosine (C). It has been suggested that such structures can be responsible for mutagenesis according to their ability to form a triple-strand conformation. Therefore, the polypyrimidine tract present in *PKD1* gene is responsible, for a high mutation rate (Ward et al., 1996).

1.3.2. *PKD2* gene

The *PKD2* gene is located on the long arm of chromosome 4 at position 4q21-23 and encodes polycystin-2 (PC2), which is a TRP calcium channel (transient receptor protein) (Mochizuki et al., 1996, Torra Balcells and Ars Criach, 2011). It is composed of 15 exons and is transcribed into a 5.4kb mRNA **(Figure 3)**. The exon 1 of this gene is highly GC-rich, and thus this exon is prone to mutations, maybe due to strand slippage which results in mispairing (Koptides et al., 1999). In contrast with *PKD1* gene, the genomic sequence covers an area of about 68kb and there is only one copy in the human genome (Koptides and Deltas, 2000).



<u>Figure 3. Gene and mRNA structure of *PKD1* and *PKD2*</u>. *PKD1* gene is consist of 46 exons. Remarkably the 5'- end of PKD1 from exon 1-33 lies in a duplicated genomic region (light green) and the 3'- end of the gene is immediately adjacent to 3'-end of the tuberous sclerosis gene, TSC2. PKD2 gene is consist of 15 exons (red colour) (Harris and Rossetti, 2010).

1.4.Mutation analysis of *PKD1* and *PKD2* genes

A very useful database named: Autosomal Dominant Polycystic Kidney Disease: Mutation Database, lists all the sequence variants described in *PKD1* and *PKD2*; (https://pkdb.mayo.edu). From this database is clear that there are some mutations that are pathogenic (i.e., the patients with these mutations have clinical phenotype) and some other variants are probably neutral **(Figure 4)**. From the database some pathogenic variants can either produce a truncated protein or affect the canonical splicing sequence which is highly conserved. Furthermore, the in-frame variants which do not interrupt the translation process of the protein, such as deletions/insertions of base numbers which are multiple of three, amino acid change variants and mutations in non-coding regions, require further evaluation (Torra Balcells and Ars Criach, 2011). To date, over than 430 different mutations have been described in the *PKD1* gene and more than 110 in *PKD2* in the databases.

PKD1 gene mutations such as missense and single nucleotide nonsense mutations, insertions and splicing defects or deletions of one or more nucleotides, result to a translation frameshift, thereby causing the production of abnormal or truncated proteins (Brook-Carter et al., 1994). Many researchers pinpoint that screening of *PKD1* gene for mutations is complicated because of the presence of HGs, especially in the duplicated region (Roelfsema et al., 1997, Watnick et al., 1997, Koptides et al., 1998, Thomas et al., 1999, Thongnoppakhun et al., 1999). For that reason, approximately 65% of the mutations reported until now are in the 3'-end unique region, which represents only about 20% of the coding sequence (Brook-Carter et al., 1994).

On the other hand, *PKD2* gene can be screened easily, and one approach is Sanger Sequencing analysis. By using this technique scientists successfully showed different types of mutations such as splicing defects, nonsense mutations and deletions or small insertions (Koptides and Deltas, 2000).

To date, over than 430 different mutations have been described in the *PKD1* gene and more than 110 for *PKD2* in the databases, Autosomal Dominant Polycystic Kidney Disease: Mutation Database; (https://pkdb.mayo.edu) and Human Gene Mutation Database [HGMD]; (www.hgmd.cf.ac.uk./ac/index.php) (Torra Balcells and Ars Criach, 2011, Liu et al., 2014). Figure 4 depicts the different types of sequence variants identified for both *PKD1* and *PKD2*. According to the pie chart the percentage of sequence changes prone to be highly pathogenic is much higher in *PKD2* than in *PKD1* **(Figure 4)**. The percentage of missense mutations is much

higher in *PKD1*, which is one of the reasons why is more difficult the diagnosis, as the pathogenicity must be proved (Torra Balcells and Ars Criach, 2011). In the table below are presented some variants found in the Greek-Cypriot population **(Table 1)**.

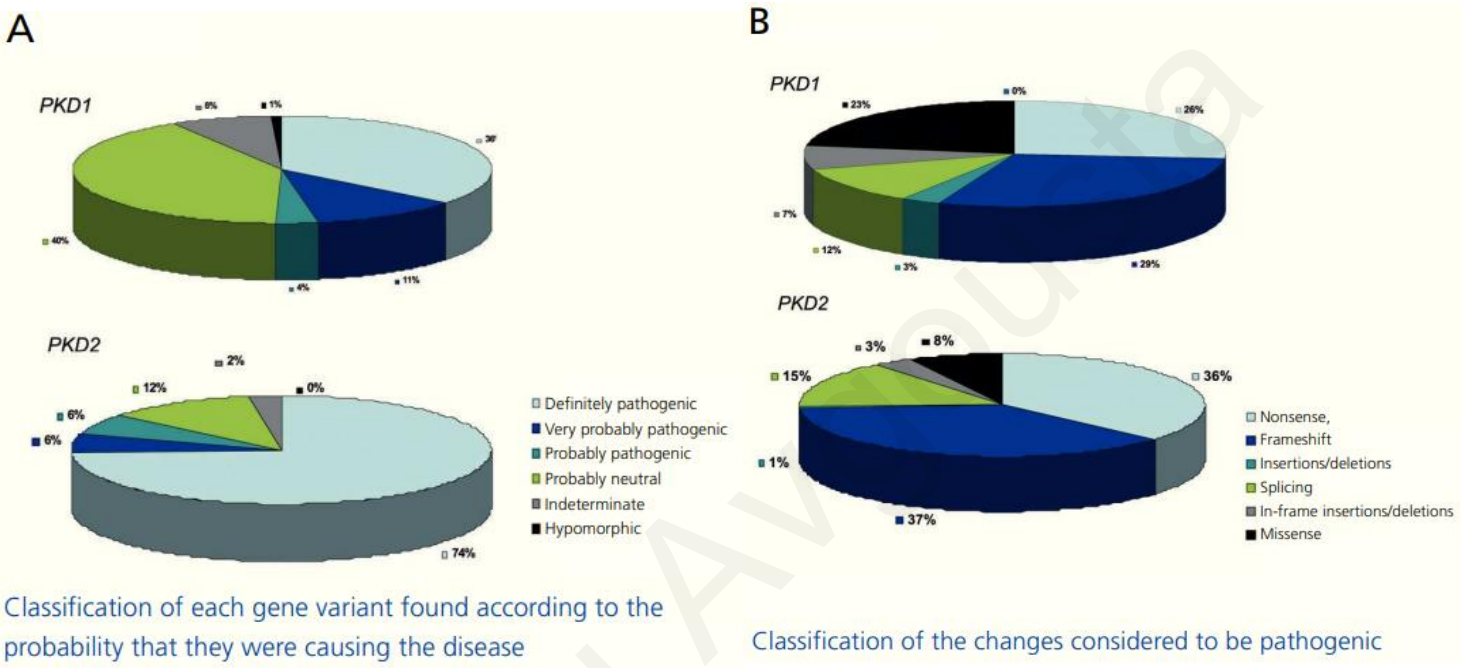


Figure 4. Classification of sequence variants found in the PKD1 and PKD2 genes according to the mutation database (https://pkdb.mayo.edu) (Torra Balcells and Ars Criach, 2011).

Table 1: Mutations of *PKD1* and *PKD2* genes in the Greek-Cypriot population

Gene	Mutation designation	Region	cDNA change	Amino acid change	Mutation type	Clinical Significance (PKDB)	The human Gene Mutation Database (HGMD)	References
PKD1	H2621P	Exon23	8762A>C	His2921Pro	Substitution	Likely pathogenic	Disease causing mutation	(Koptides et al., 2000)
PKD1	F3066L	Exon25	9195G>C+ 9196T>C	Phe3066Leu	Substitution	Likely neutral	-	(Koptides et al., 2000)
PKD1	V3090V	Exon26	9270C>T	Val3090Val	Synonymous	Likely neutral	-	(Koptides et al., 2000)
PKD1	P3110P	Exon26	9330T>C	Pro3110Pro	Synonymous	Likely neutral	-	(Koptides et al., 2000)
PKD1	T3223T	Exon28	9669G>A	Thr3223Thr	Synonymous	Likely neutral	-	(Koptides et al., 2000)
PKD1	V3375M	Exon31	10123G>A	Val3375Met	Substitution	Likely neutral	Disease causing mutation	(Koptides et al., 1998)
PKD1	P2471L	Exon18	7412C>T	Pro2471Leu	Substitution	Highly likely pathogenic	Disease causing mutation	(Bouba et al., 2001)
PKD1	Q2519L	Exon19	7556A>T	Gln2519Leu	Substitution	Indeterminate	Disease causing mutation	(Bouba et al., 2001)
PKD1	T2649I	Exon21	7946C>T	Thr2649Ile	Substitution	Indeterminate	Disease causing mutation	(Bouba et al., 2001)
PKD2	R803X	Exon13	2407C>T	Arg803X	Nonsense	Pathogenic	Disease causing mutation	(Deltas, 2001)
PKD2	G83C	Exon1	83G>C	Arg28Pro	Substitution	Likely neutral	-	(Koptides et al., 1999)
PKD2	G568A	Exon1	568G>A	Ala190Thr	Substitution	Likely neutral	Retired record	(Deltas, 2001)
PKD2	IVS3-22G>A	IVS3	844-22G>A	Likely Silent	IVS Silent	Likely neutral	-	(Deltas, 2001)
PKD2	M800L	Exon13	2398A>C	Met800Leu	Substitution	Likely neutral	Disease causing mutation	(Deltas, 2001)
PKD2	Q938Q	Exon15	2814A>G	Gln938Gln	Synonymous	Likely neutral	-	(Deltas, 2001)
PKD2	R742X	Exon11	2224C>T	Arg742Ter	Nonsense	Definitely pathogenic		(Mochizuki et al., 1996)

1.5.*PKD1* and *PKD2* products: The polycystins

PKD1 gene is translated to the large polycystin-1 (PC1) protein which is composed of 4301 amino acids. More specific, PC1 is a 450kD receptor-like protein with a large N-terminal extracellular domain with almost 3000 amino acids, 11 trans-membrane domains and a short cytoplasmic C-terminal domain of 200 residues **(Figure 5)** (Hughes et al., 1995, Wang et al., 1996, Sandford et al., 1997, Bycroft et al., 1999, Nims et al., 2003, Bouba et al., 2001). It is expressed in epithelial cells and in some other tissues such as the liver, bone, heart, and endocrine glands (Ward et al., 1996, Ibraghimov-Beskrovnaya et al., 1997, Markowitz et al., 1999, Peters et al., 1999). Also, PC1 localizes at the cilium, at the adhesion complex in polarized epithelial cells and at plasma membrane (Ibraghimov-Beskrovnaya et al., 1997, Huan and van Adelsberg, 1999, Yoder et al., 2002, Streets et al., 2009), and plays a critical role in cell-matrix and cell-cell interactions (Babich et al., 2004, Streets et al., 2009).

Polycystin-2 is encoded by *PKD2* gene and is an integral membrane protein with six-transmembrane domains (TM) and is composed of 968 amino acids **(Figure 5)** (Mochizuki et al., 1996). It acts as a Ca^{2+}-responsive cation channel and is homologous to the transient receptor potential family of cation channels (González-Perrett et al., 2001). Although PC1 and PC2 colocalizes to the cilium, most of the cellular pool of PC2 is seen in the intracellular compartment and functions to release calcium (Ca^{2+}) from intracellular storage tanks (Vassilev et al., 2001, Nauli et al., 2003, Köttgen et al., 2005, Tsiokas et al., 2007). In addition to this, PC2 is also found to be located at the plasma membrane, where seems to be in complex with PC1 (Hanaoka et al., 2000, Pelucchi et al., 2006, Yu et al., 2009). PC2 contains domains in both the N-terminus and C-terminus, which are responsible for the PC2 protein-protein interactions and Ca^{2+} sensitivity (Mochizuki et al., 1996, Qian et al., 1997, Celić et al., 2008).

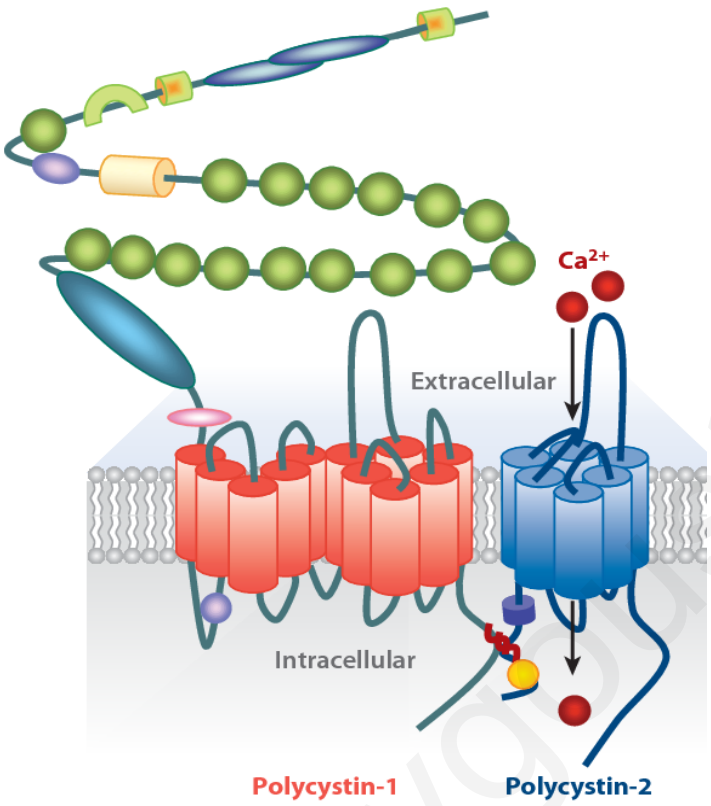
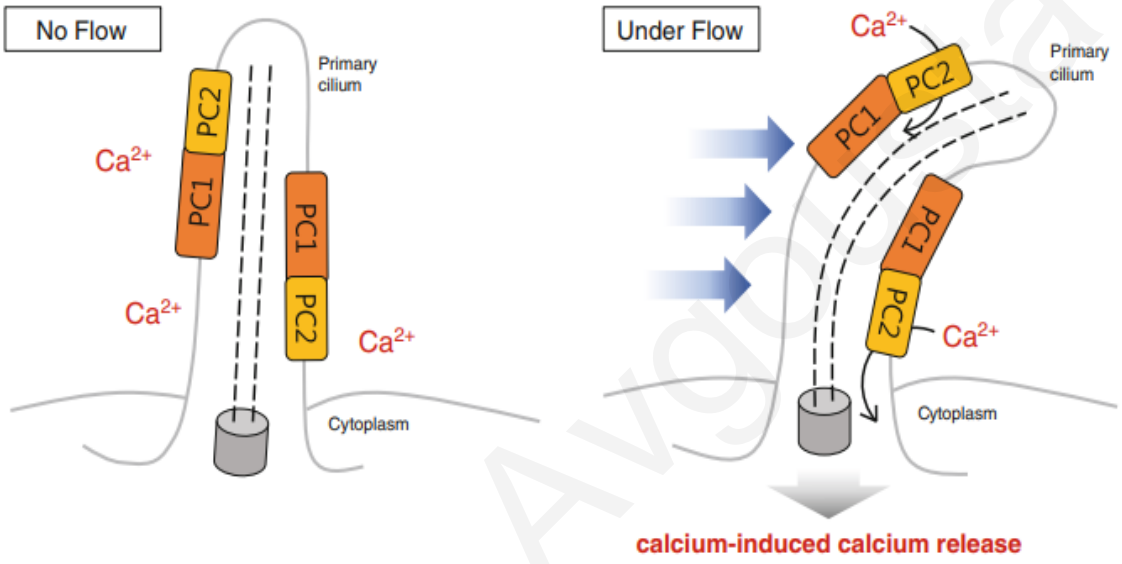


Figure 5. Polycystin-1 and Polycystin-2. PC1 is proposed to function as a G-protein-coupled receptor, and PC2 as a cation-selective ion channel permeable to Ca^{2+}. PC1 and PC2 physically interact via their respective C-terminal coiled-coil domains (Zhou, 2009).

1.5.1. Interaction between PC-1 and PC-2

Both PC1 and PC2 are co-located in the primary cilium and endoplasmic reticulum (ER) (Yoder et al., 2002). They form a complex through their C-terminal tails which is activated in response to ciliary bending and promotes the signal transduction by a mechanical or chemical stimulus through intracellular Ca^{2+} regulation (Tsiokas et al., 1997, Qian et al., 1997, Nauli et al., 2003, Casuscelli et al., 2009). Specifically, when liquid flow through the lumens of the renal tubules, then the primary cilia of the renal epithelial cells bend, and this has as a result the onset of mechanical response. Subsequently, this deflection of primary cilia is recognized by PC1 proteins which as mentioned before are localized to the membrane of primary cilia on the renal epithelial cells. At first, PC1 acts as a mechanosensory protein for luminal flow that transfer

mechanical stress from the extracellular environment to PC2 protein (Huang and Lipschutz, 2014). Then in turn PC2 protein induce minimal calcium ion influx, causing huge intracellular Ca^{2+} release though Ca^{2+}-induced/ Ca^{2+} release (CICR). Finally, the increased Ca^{2+} level of the cytosol, can regulates several signaling pathways which are correlated with the development and proliferation **(Figure 6)** (Nauli et al., 2003).



<u>Figure 6. Calcium regulation.</u> The mechanism of induced Ca^{2+} release by primary cilia under flow *(Ko, 2016).*

1.6.Mechanisms of cyst formation

Patients with ADPKD that inherit one mutated copy of *PKD1* or *PKD2* gene from a parent, will develop functionally and normal kidney into their adulthood (Qian et al., 1996, Brasier and Henske, 1997). However, all along their lifespan cysts will form in these patients' kidney after the loss of both functional copies of a polycystin gene (Koptides et al., 1998). For that reason, although ADPKD is genetically dominant at the organismal level, is recessive at the cellular level (Chapin and Caplan, 2010). Several research groups were suggested that one mutation in cyst-lining epithelial cells has germinal nature and the other is subjected to a somatic event (Qian et al., 1996, Koptides et al., 1998). More specific, a somatic "second hit" mutation may be required for cystogenesis (Qian et al., 1996, Brasier and Henske, 1997, Koptides et al., 1998, Torra et al., 1999, Pei et al., 1999). According to this model, each cyst arises because of a distinct somatic mutation event, explaining the disease's slow progression over decades. In addition to this, there are and some other factors which impact disease progression such as the level of PC1 protein expression, the stage of kidney development affected by *PKD1* mutation and the penetrance of pathogenic alleles (Lu et al., 1997, Reynolds et al., 1999, Pritchard et al., 2000, Lantinga-van Leeuwen et al., 2004, Rossetti et al., 2009). This model was first proposed by *Knudson* in 1971 to explain the origin of retinoblastoma.

Several research groups were presented evidences about ADPKD1 cystogenesis, and they concluded that the inactivation of the wild-type copy allele is the main event for cystogenesis (Qian et al., 1996, Brasier and Henske, 1997, Koptides et al., 1998). This inactivation can be performed either by entire deletion of the wild-copy allele, where this mechanism is called loss of heterozygosity (LOH) or by point mutations within the gene (Koptides and Deltas, 2000). Moreover, in 1998 *Wu* et al. by using murine model, showed that the inactivation of both *PKD2* alleles is essential for cyst formation (Wu et al., 1998). In addition to the above, inactivation of PC1 or PC2 expression in mice model's kidney, has showed that loss of these proteins in the developing kidney causes much more severe cystic disease than from the loss of one of these proteins in the mature kidney (Piontek et al., 2007, Lantinga-van Leeuwen et al., 2007, Takakura et al., 2008). Hence, these data suggest that loss of polycystin function mainly during the period of rapid cell division and growth, which happens during post-natal renal development creates a predisposition for cystogenesis, whereas polycystin function is far less crucial after the end of cell proliferation (Chapin and Caplan, 2010).

One more explanation for cysts formation can be the transient obstructive or ischemic injuries to the tubule epithelial cells, as the individuals getting older. These injuries stimulate the mechanisms of repair, which involves cellular division and growth. PC1 and PC2 proteins play an important role for cellular growth and division. Hence the decreased levels of functional proteins in the cells of individuals heterozygous for ADPKD mutation could derange the repair process and lead to cyst formation. This was supported by studies in mice's kidneys heterozygous for *PKD1* or *PKD2*, where were not able to repair themselves as effectively as kidneys from wild-type mice. Furthermore, more tubule dilation and microcysts accumulate to the heterozygotes than wild-type kidneys (Bastos et al., 2009, Prasad et al., 2009). Also, knocking out PC1 expression in adult mice kidneys causes a similar sensitivity to injury (Takakura et al., 2009). Therefore, these results proposed that injury may be sufficient to start cyst formation in heterozygotes without a somatic "second hit" mutagenesis event. Conclusively, the initiation of cyst formation could be either because of a somatic mutagenesis or an injury event, or could be a constitute of a "second hit" at one of the PKD loci to cause ADPKD disease (Chapin and Caplan, 2010).

1.6.1. Cyst expansion

ADPKD as mentioned earlier is characterized by formation of multiple fluid-filled kidney cysts, which is a stark contrast to the normally compact arrangement of tubules in a healthy kidney **(Figure 7)** (Grantham, 1996). It is already known that in patient's kidney, cells are organized themselves to create a spherical rather than tubular structures, and then the lumens of these structures filled with fluid in order to expand the consequent cysts (Qian et al., 1996, Brasier and Henske, 1997). Also, it has been suggested that cysts can be enlarged their surface by increasing the number of cells where surround the cyst lumens rather than by simply stretching this epithelial layer (Grantham, 1996). One characteristic about the change of the structure, from tubular to spherical, is the perturbations in planar cell polarity which cause tubular epithelial cells not to divide along an axis parallel to the tubule lumen, having as a result tubule expansion rather than elongation. Also, this event can occur before cyst appear, so the loss of this planar cell polarity can be a precursor to cyst formation. (Fischer et al., 2006, Patel et al., 2008). In addition to this, according to *Nishio* et al. (2010) despite that defects in planar cell polarity could be a crucial factor for the expansion of many cysts, nevertheless, this factor seems to be not essential for the initiation of cyst formation (Nishio et al., 2010).

The other aspect of cyst formation is the change of the cyst-lining cells from an ion-absorptive to an ion-secretory epithelium resulting to the expansion of cyst fluid volume. In that way, because of the ion secretion into the lumen then the paracellular or transcellular drives the osmotic water movement into the cyst. More specific, cAMP can trigger chloride (Cl^-) transport, and this has as a result the rapid and progressive dilation of tubules (Grantham, 1996). In more details, cAMP signaling trigger fluid movement where driving cyst formation and involves the apical cystic fibrosis transmembrane regulator and the basolateral Na^+-K^+-$2Cl^-$ cotransporter NKCC1 (Davidow et al., 1996, Magenheimer et al., 2006, Montesano et al., 2009). Polycystin proteins have an important role because they can regulate the cAMP signaling by modulating the localization, expression and the activity of Cl^- channels (Chapin and Caplan, 2010). Overall, the hyperactivation of cell proliferation, which is a result of increased intracellular cAMP followed by the inhibition of calcium influx because of the dysfunction of the PC1 and PC2 complex, are the critical factors for the initiation of ADPKD (Ecder et al., 2002).

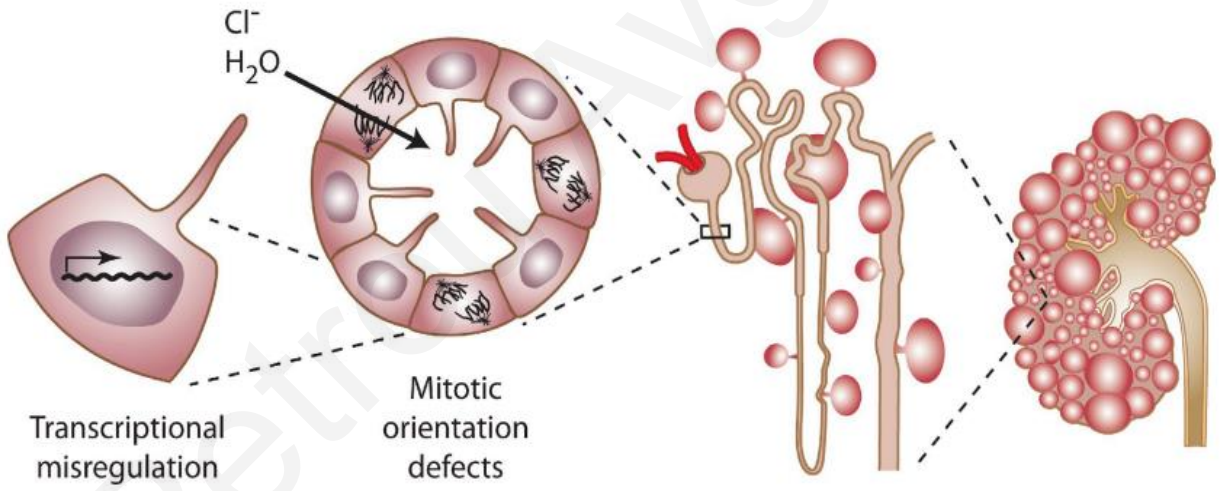


Figure 7. Cyst formation at the level of the cell, nephron, and kidney. Defects in *PKD1* and *PKD2* genes which encodes PC1 and PC2, lead to aberrant gene transcription, cell proliferation and ion secretion which in turn result in the formation of fluid-filled cysts. As cysts balloon out from individual nephrons, their collective effects leads to the displacement of the normal renal parenchyma and the formation of a cyst-filled kidney with reduced functional capacity (Chapin and Caplan, 2010).

1.7.Molecular diagnosis of ADPKD

ADPKD is diagnosed with the detection of cysts in the kidney by imaging techniques, such as ultrasound, T2-weighted magnetic resonance imaging (MRI), and computed tomography (CT) **(Figure 8)** (De Rycke et al., 2005, Pei et al., 2009, Pei and Watnick, 2010, Pei et al., 2015). Although, the ultrasound is the most widely used technique for diagnosis, CT and MRI are more sensitive techniques because they can detect small cysts (Pei et al., 2009, Torra Balcells and Ars Criach, 2011). The criteria of these imaging techniques for ADPKD have been published (Nicolau et al., 1999, Demetriou et al., 2000, Barua and Pei, 2010, Pei et al., 2015). However, these techniques are not sufficient for diagnosis in young adults and children, in whom renal sonography may be inconclusive when there is not a family history, or for elderly individuals with mild form of the disease (Nicolau et al., 1999, Torra Balcells and Ars Criach, 2011, Fatehi et al., 2017). Also, the appearance of ADPKD symptoms is age dependent, with an especially low penetrance below the age of 20 (Dalgaard, 1957, Ravine et al., 1992, Bear et al., 1992).

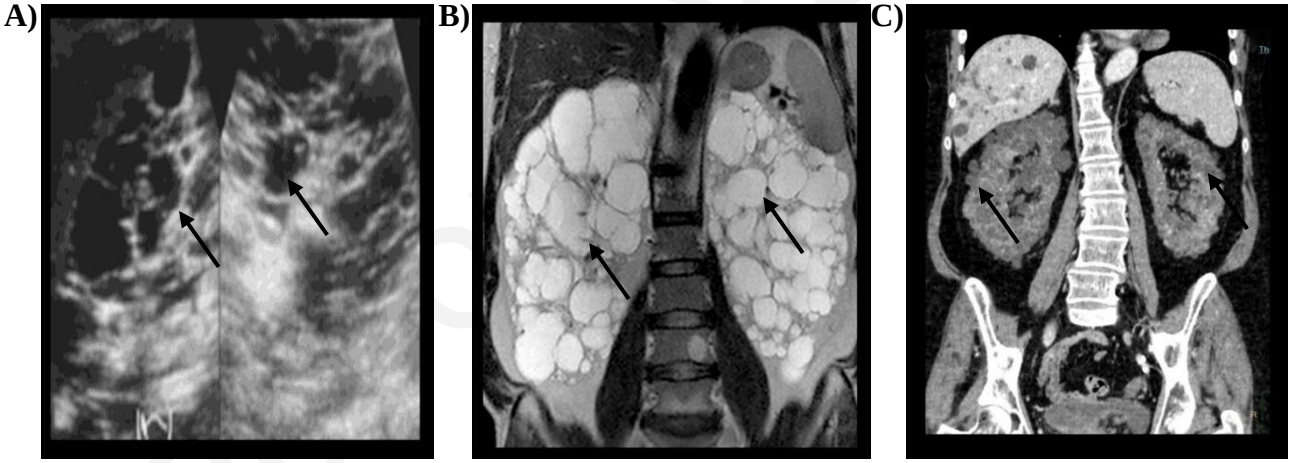


Figure 8. Diagnostic imaging techniques. A) Ultrasound- cysts be displayed as irregular sphere shape with black colour, **B)** MRI- cysts be displayed as irregular sphere shape with white colour, **C)** CT scan- cysts be displayed as irregular sphere shape with (Gradzik et al., 2016).

The genetic linkage analysis in an indirect method for early detection of ADPKD (Barua and Pei, 2010). This method establishes linkage between genes and polymorphic markers and serves as a way of gene-hunting and genetic testing. Linkage is the tendency for genes and other genetic markers to be inherited together because of their location is near to one another on the same chromosome. Two loci are linked if, during meiosis, recombination occurs between them with a probability of less than 50% , and the probability of recombination is a function of chromosomal distance between two loci, measured in centimorgans (cM) (Burton et al., 2005). Although indirect genetic testing by linkage analysis is a practical and reliable approach for ADPKD diagnosis in clinical settings, there are two main disadvantages (Fatehi et al., 2017). The first one is that, due to the genetic heterogeneity of the disease, this method cannot be used without ample number of affected and unaffected family members. And the second one, is that it cannot be performed when the proband is suspected of having a *de novo* mutation (10% of the cases) (Pei, 2006). It was suggested that the direct mutation sequencing remains the better choice for making a clinical diagnosis of this disease (Garcia-Gonzalez et al., 2007, Audrézet et al., 2012, Liu et al., 2014, Hafizi et al., 2014).

The direct detection of mutations in the PKD genes (pathogenic variants) provides a credible molecular diagnosis. Especially in cases without positive family history, the genetic test can allow the early detection of the disease which is necessary or preimplantation genetic diagnosis and kidney donation from a young relative (Fatehi et al., 2017). However, the molecular diagnosis of ADPKD1 (is caused by mutations in *PKD1* gene) is a very difficult task (Bogdanova et al., 2001, Fatehi et al., 2017). The *PKD1* gene sequencing is complicated because of the six homolog parts in the genome (Phakdeekitcharoen et al., 2000, Bogdanova et al., 2001). To date, several attempts have been made to improve the screening of *PKD1* gene by using short-read NGS approaches to replace the standard diagnostics based on Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA) assays (Rossetti et al., 2012, Qi et al., 2013, Tan et al., 2014, Trujillano et al., 2014, Eisenberger et al., 2015, Mallawaarachchi et al., 2016). Nevertheless, *Lee and Schatz* (2012) showed that duplicated and high GC content region, such as *PKD1* gene, can lead to ambiguous identification of variants when analyzed with short-read NGS strategies (Lee and Schatz, 2012). Hence, diagnostic assays based on NGS short reads may not fully suited for reliable ADPKD diagnostics. On the other hand, *Borras et al* (2017) showed that direct sequencing of long-range PCR (LR-PCR) products discriminates *PKD1* and pseudogenes and improve the

mapping quality of PKD1 (Borràs et al., 2017). Specifically, the improved mapping quality reduced the interference of homologous sequences and the high GC content for ADPKD diagnosis (Qi et al., 2013, Borràs et al., 2017).

1.8.Emerging treatment strategies for ADPKD

To date, there are no Food and Drug Administration (FDA)-approved therapies for the treatment of ADPKD. However, recent studies have proposed several promising targets and molecular pathways correlated to cystogenesis, providing new insights into potential therapeutic prospects. The main approach for ADPKD treatment has focused on inhibiting cystic cell proliferation and fluid secretion (Yang et al., 2008, Calvet, 2008, Bukanov et al., 2012, Chang and Ong, 2012). Knowing that the progression of cystic disease is caused because of the uncontrolled proliferation of epithelial cells that line the cysts and that these epithelial cells secrete rather than absorb fluid and electrolytes, many research groups aim to develop therapies where would target on these derangements (Chang and Ong, 2008, Patel et al., 2009, Harris and Torres, 2009).

As fluid secretion into cyst lumen is mediated partially by apical cystic fibrosis transmembrane conductance regulator (CFTR) chloride channels and is stimulated by cyclic adenosine monophosphate (cAMP), these two factors suggested to be promising molecular targets. According to *Yang et al.* (2008) the CFTR inhibitor compound CFTRinh172 appears to substantially slow the cyst expansion (Yang et al., 2008). Furthermore, inhibition of the basolateral potassium channel which was shown to be necessary for the maintenance of electrochemical potential that drives chloride and fluid secretion, is also a candidate approach for blocking cyst fluid accumulation (Albaqumi et al., 2008).

Another promising strategy is the Antidiuretic hormone (ADH), which is acting though the V2 vasopressin receptor and is the major factor for cAMP production in the collecting tubule of the kidney. It has been suggested that the drug Tolvaptan, a V2 receptor antagonist, reduces cyst progression in ADPKD mouse models (Gattone et al., 2003, Torres et al., 2004, Wang et al., 2005, Torres, 2008). Furthermore, one somatostatin analogue, the octreotide, shown to inhibit cAMP accumulation in several cell types and promising results in animal models (Masyuk et al., 2007, Hogan et al., 2010).

Furthermore, the fact that high mammalian target of rapamycin (mTOR) activity contribute to the excessive cell proliferation and growth, has prompted evidence for the use of mTOR inhibitors in the setting of ADPKD (Wahl et al., 2006, Shillingford et al., 2006, Zafar et al., 2009, Distefano et al., 2009, Dere et al., 2010, Torres et al., 2010). Although animal studies proposed significant beneficial effects from that treatment, recent clinical trial data have shown that these results were not borne out in ADPKD patients. In addition, the side effects from the use of chronic mTOR inhibitors as medication may be substantial enough to further limit its potential utility (Torres et al., 2010, Serra et al., 2010, Walz et al., 2010). Finally, another approach is the regulation of mitosis of the cells lining the cysts. More specific, anti-proliferative drug under the brand name Roscovitine was shown to blocks cyclin-dependent kinases (CDKs), resulting a slower cyst formation in at least some animal models of PKD (Bukanov et al., 2006).

AIMS OF THE STUDY

2.1. Scientific hypothesis:

The purpose of this study was the genetic investigation of two Cypriot families, clinically diagnosed with ADPKD. This disease is mainly caused by mutations in *PKD1* and/or *PKD2* gene. Therefore, assuming that these two genes are responsible for the ADPKD phenotype, we investigated *PKD1* gene with linkage analysis and screened *PKD2* gene for mutations by Sanger Sequencing analysis. Also, in our effort for further investigation of *PKD1* gene, avoiding its genetic complexities, we aimed to develop and optimize long-range PCR amplification paired with amplicon-based Next Generation Sequencing (NGS). Hence, this combination could allow the simultaneous screening of both genes by semi-conductor NGS technology for future diagnosis of ADPKD patients.

2.2. General objective:

Initially, one approach was to study and analyze these families by linkage analysis. This indirect method is based on the analysis of genetic markers located in the regions of *PKD1* and *PKD2* genes. In that way, we tried to identify if the *PKD1* or *PKD2* genetic regions are correlated to the illness within the family. If the phenotype is caused because of mutations in the *PKD2* gene (ADPKD2), then we will be able to investigate it further by using Sanger sequencing analysis to identify the exact mutation. On the other hand, if the phenotype is caused by mutations in *PKD1* gene (ADPKD1), then further analysis becomes more complicated. This difficulty in analysis is due to the presence of *PKD1* homologous genes (pseudogenes), the large transcript size of *PKD1*, and its high GC content (Blumenfeld, 2009).

According to previous studies, long-range PCR (LR-PCR) amplifying whole regions of the *PKD1* gene was shown to be a way to exclude pseudogenes, followed by nested PCR of individual exons. Amplicons are then directly analyzed by Sanger Sequencing or by sequencing coupled with a mutation screening step (Rossetti et al., 2001, Tan et al., 2012). Hence, in this study to overcome this significant problem with *PKD1* sequencing, because of its genomic characteristics, we aimed to developing a pipeline were LR-PCR amplification of *PKD1* and *PKD2* exonic regions paired with amplicon-based Next-Generation Sequencing (NGS) are used to allow the simultaneous screening of both genes by semi-conductor NGS technology (IonTorrent, ThermoFisher).

Although *PKD2* is not a complex gene, the sequencing of LR-PCR fragments for *PKD2* was performed as a proof of principle of long-read sequencing and detection of variants.

2.3. Specific objectives:

1) Analysis of the two Cypriot families by using genetic linkage analysis in order to correlate their phenotype with the segregation of either *PKD1* or *PKD2* in affected individuals.
2) Sequencing of *PKD2* gene by using Sanger Sequencing analysis, to detect a possibly pathogenic mutation.
3) Development of an NGS amplicon panel, using LR-PCR products as template for the simultaneous sequencing of both *PKD1* and *PKD2* genes.

MATERIALS AND METHODS

3.1. Patients Information

In this study we investigated two Cypriot families, clinically diagnosed with ADPKD. The first family with code CY1611 consists of ten individuals, out of which three are patients. The test aimed at investigating whether the younger generation of family members were carries off the haplotype or mutation associated with the disease in their family. Participants were evaluated by a nephrologist and we proceeded with investigating segregation of both *PKD1* and *PKD2* gene regions by linkage analysis.

The second family with code CY1626 consists of three individuals, the affected mother (UCY5176) and her two sons (UCY5177, UCY5178). This family was previously known to be linked to the *PKD2* gene locus, therefore the *PKD2* gene was directly analyzed by Sanger Sequencing analysis in search for mutations.

3.2. DNA extraction

DNA samples from two different families were isolated from peripheral blood leukocytes by a salting out procedure (Miller et al., 1988). Briefly, 20ml of blood was resuspended in 50ml centrifugation tubes with 0,9% NaCl. After centrifugation, the red blood cells lysed with the red blood cell lysis buffer (155mM Ammonium chloride, 10mM Sodium bicarbonate, 0,1mM 5% EDTA). Then, the tubes were centrifugated, and the white blood cells (pellet) were incubated at 70ºC for 1 hour with the white blood cell lysis buffer (100mM Sodium chloride, 25mM Disodium EDTA), 10% SDS and Protease K (stock concentration 20mg/ml). Subsequently, 9.5M Ammonium Acetate was added and the tubes were centrifugated. Next, 96% of cold ethanol was added to the clear supernatant containing the DNA to allow DNA precipitation. The isolated DNA was washed with 70% ethanol and the tubes were centrifugated. The ethanol was removed and TE buffer (10mM tris-HCL, 0.1mM EDTA) was added. The concentration of DNA was evaluated by using NanoDrop™ spectrophotometer (Thermo Fisher Scientific).

3.3. Polymerase Chain reaction (PCR)

The purpose of a polymerase chain reaction is to increase the number of copies of known fragment of DNA. In the non-limiting phase of the reaction the increase in the number of copies of the fragment is exponential. There are three main steps in PCR reaction which are repeated for many cycles. The first one is the denaturation, where the double strand melts, open to single stranded DNA. The second step is the annealing, where primers anneal to the template DNA and the final step is the extension, where the DNA polymerase synthesizes the DNA template. All this process is carried out on an automated thermal cycler. Verification of the PCR reaction can be done by analyzing the PCR product on an agarose gel (Moy, 2007).

Initially, PCR was used to amplify the microsatellite markers for the linkage analysis (fragment analysis), a method based on the analysis of polymorphic genomic markers flanking the *PKD1* and *PKD2* gene loci. The primers of the informative microsatellite markers for PCR were obtained from the Ensemble Genome Browser (https://asia.ensembl.org/index.html). The PCR primers and the markers that have been used in the experiments are listed in **Table 2**.

In order to identify the appropriate PCR conditions, PCR reaction for each marker was performed by using a laboratory control DNA. Briefly, the various components required for PCR were included DNA sample (50ng/μl), DNA primers (stock: 100pmol/μl) free deoxynucleotide triphosphate (dNTPs) (stock:100mM), AmpliTaq^{tm} 360 DNA polymerase 250U, and 25mM Magnesium Chloride, AmpliTaq^{tm} 360 Buffer (10x) and 5% DMSO (if needed). The final conditions and the PCR protocol for each marker are listed in **Table 3 and 4**.

Table 2: *PKD1* and *PKD2* linkage analysis markers and primers

Marker	Location	Position	Strand*	Primer Sequence (5'->3')	Product Size (bp)
HBAP1	Chr16: 170814-170912	-	FW	AAAATCTATCCATGCTTTCA	95-107
			RV	AGGAGACAGGAAAGAGAGAC	
D16S3024	Chr16:1604202-1604428	7.05 cM	FW	ACATGCTGTGCCACCT	208-248
			RV	AGCTGCCAGTATATGGAGGA	
D16S3395	Chr16:1951695-1951822	6.08 cM	FW	CTAACCCTCAGCAGAGTTCTG	124-137
			RV	CCTGGCAGTAAGTCCTGAAA	
D16S3070	Chr16:3043866-3044032	7.61 cM	FW	CACGGGAGGTGGAGGT	153-173
			RV	TGAAAGTGGTTTAAGAGAGCA	
D16S3027	Chr16:4000871-4001136	8.71 cM	FW	ATATTTGGCATCTGGGG	262-286
			RV	CCAGCATGAGTTGCTTT	
D4S395	Chr4:83325089-83325219	92.42cM	FW	TACTCCAGCCTGGATGACAG	113-137
			RV	TGTTCCATAACAAGCACGTT	
D4S1534	Chr4:85387157-85387315	95.09 cM	FW	ATTCAGTTTCAGCCCCAT	146-158
			RV	ACCAGCCCAAGGTAGAGG	
D4S2460	Chr4:88912511-88912699	97.30 cM	FW	CCAAAATCATGTGAGCCA	157-191
			RV	GAGCAGCAGCCAACTGTAT	
D4S423	Chr4:91551623-91551739	100.75cM	FW	TTGAGTAGTTCCTGAAGCAGC	103-125
			RV	CAAAGTCCTCCATCTTGAGTG	
D4S2986	Chr4:98736515-98736741	105.29 cM	FW	TCACTCTACAGCACTCCAGC	197-229
			RV	TTACTTTGTTTATTAGTTCTGACCG	
D4S1572	Chr4:102848900-102849046	107.95cM	FW	AGACTCTAAAGATATGGTGATTTGC	135-151
			RV	TCTGATTGATTTTATGTGTGCC	

***FR: Forward strand, RV: Reverse strand**

Table 3: Conditions for *PKD1*' markers

	HBAP1	D16S3024	D16S3395	D16S3070	D16S3027
dNTPs	2µl	2µl	2µl	2µl	2µl
Buffer	2µl	2µl	2µl	2µl	2µl
Forward primer	0,3 µl	0,3 µl	0,3 µl	0,3 µl	0,3 µl
Reverse primer	0,3 µl	0,3 µl	0,3 µl	0,3 µl	0,3 µl
Taq-pol	0,125µl	0,25µl	0,125µl	0,125µl	0,125µl
DNA	1µl	1µl	1µl	1µl	1µl
$MgCl_2$	1µl (2,5mM)	2µl (2,5mM)	1,2µl (3mM)	1µl (2,5mM)	2,4µl (3mM)
ddH_20*	3,275µl	12,15µl	3,075	3,275	11.875
Reaction volume	10µl	20µl	10µl	10µl	20µl
PCR protocol	96°C -> 4min	96°C -> 4min	96°C -> 4min	96°C -> 4min	96°C -> 4min
	95°C -> 1min	95°C -> 1min	95°C -> 1min	95°C -> 1min	95°C -> 1min
	57°C -> 1 min	57°C -> 1 min	58°C -> 1 min	57°C -> 1 min	57°C -> 1 min
	72°C -> 0,30sec	72°C -> 0,30sec	72°C -> 0,30sec	72°C -> 0,30sec	72°C -> 0,30sec
	72°C -> 10min	72°C -> 10min	72°C -> 10min	72°C -> 10min	72°C -> 10min
	4°C hold	4°C hold	4°C hold	4°C hold	4°C hold
Cycles	x40	x40	x40	x40	x40

***ddH20** – double distilled water

Table 4: Conditions for *PKD2*' markers

	D4S1534	D4S2460	D4S423	D4S2986
dNTPs	2μl	2μl	2μl	2μl
Buffer	2μl	2μl	2μl	2μl
Forward primer	0,5 μl	0,5μl	0,5μl	0,5μl
Reverse primer	0,5μl	0,5μl	0,5μl	0,5μl
Taq-pol	0,2μl	0,2μl	0,2μl	0,2μl
DNA	1μl	1μl	1μl	1μl
$MgCl_2$	1,6μl (2mM)	1,6μl (2mM)	2,4μl (3mM)	2,4μl (3mM)
5% DMSO	1μl	1μl	1μl	1μl
ddH_20*	11,2μl	11,2μl	10,4μl	10,4μl
Reaction Volume	20μl	20μl	20μl	20μl
PCR protocol	96°C -> 4min	96°C -> 4min	96°C -> 4min	96°C -> 4min
	95°C -> 0,50sec	95°C -> 0,50sec	95°C -> 0,50sec	95°C -> 0,50sec
	56°C -> 0,50sec	58°C -> 0,50sec	58°C -> 0,50sec	56°C -> 0,50sec
	72°C -> 0,50sec	72°C -> 0,50sec	72°C -> 0,50sec	72°C -> 0,50sec
	72°C -> 5min	72°C -> 5min	72°C -> 5min	72°C -> 5min
	4°C hold	4°C hold	4°C hold	4°C hold
Cycles	x40	x40	x40	x40

***ddH20** – double distilled water

In addition to this, PCR was used to amplify the exons of the *PKD2* gene. The primers for these exons were designed by using the Primer3 Input (version 0.4.0). The factors that considered important for the design of the primers were: the size of the primers, 8-22 base pairs (bp) in size and the percentage of Gc and Cs (GC%) (50-55%). In addition to this, they must have a GC-lock at the 3´-end and the melting temperature should be from 59ºC to 63ºC (except from exon1a) and do not include repetitive regions. The sequence of the primers is listed in **Table 5** and the final conditions and PCR protocol are listed in **Table 6**.

Table 5: *PKD2* primers for all the coding sites.

Exon number	Forward primer sequence (5'->3')	Reverse primer sequence (5'->3')	PCR annealing temperature (ºC)	PCR product size (bp)
1a	GTGACCGCGATGGTGAACT	CAGAGGGGATGCGAGATG	68ºC	714
1b	CCACTTGGAACGCGGACT	ACTCTACGTCCATCTCCACCAC	58ºC	632
2	GGCCTACAAAACCAGTTTCTCAGT	CTCTGGTGCATACACACTTCCTT	61ºC	377
3	AAGGCTGCTGGTATGTGAATGT	AGCAGGTCCTGTCGATACTCAT	60ºC	296
4	CTTGGTTATGCAAACGATGCAG	GTGAAGGTGTCAGGGAATGAAT	60ºC	426
5	CTGTCCTTGTAATTGCCTCAAGTG	GTAGCTAACTGCAGGCAAAGGT	61ºC	393
6	GCATGAACAGAACAGATGGACA	CAATGCTGAGGAGATCAAAGACTC	61ºC	395
7	ACTTTTCAGCGGTCATTTTG	GACTTCCTTGGGGCTACCT	63ºC	308
8	GGTTCATCCATGTTGTAACCTGTC	TGGTGGTCATATAGCAACCTCA	61ºC	399
9	GATACAAGTAGCAAGACTTTGTAAATGG	TAAAGTGTGAGAGAAAAGAGAAGACAAG	61ºC	366
10	GTCTTCATAAAGCACTCAGATTAGGAA	CAGTATATCCATCAAGACTCCAAGATAG	61ºC	400
11-12	GCCAATGTACACCAGGTTTGTAGTAG	ATAAACCGACTGAGAGAGAGAGAGAG	60ºC	751
13	TTGGCTATTCCTTGCTGTTAGTT	ATGACCTGCCTGCTTTGG	59ºC	549
14	TGTACTGTGTTTTCCTTGCA*	AAATACAACTGTCAGCAACATA*	63ºC	248
15	ATCTCCAGCCTTACCAAACTACAG	CGTACCGCTCAATACACATCAT	61ºC	597

***Primers according to Hayashi et al., 1997 (Hayashi et al., 1997)**

Table 6: PCR protocol for *PKD2* sequencing analysis

Solution for PCR	Stock	Final concentration	Amount per reaction (µl)
ddH_20	-	-	14,75 µl
AmpliTaq^{tm} 360 Buffer	10x	-	3 µl
$MgCl_2$	25mM	2,5mM	3 µl
dNTPs	100mM	2mM	3 µl
Forward primer	100µM	10µM	1 µl
Reverse primer	100pmol/µl	10µM	1 µl
AmplTaq^{tm} polymerase	250U	-	0,25 µl
DNA template	-	50ng/µl	1 µl
360 GC enhancers	-	-	3 µl
Reaction Volume	-	-	30 µl
PCR protocol	95^{o}C -> 4min		
	94^{o}C -> 1min		
	Tm -> 1min		
	72^{o}C -> 1min		
	72^{o}C -> 10min		
	4^{o}C hold		
Cycles	x40		

Finally, agarose gel electrophoresis was used to confirm the functionality of the primers and the PCR reactions. The PCR products were visualized on a 2% ethidium bromide agarose gel. The DNA strands are stained with ethidium bromide dye which is inserted into the hydrogen bonds that join the complementary bases in the double-stranded DNA for visualization, 5µl of PCR product was mixed with 2µl of loading dye and loaded onto the gel. Subsequently, applying voltage to the end of the gel at 120volt for almost 20 minutes, which promotes the negatively charged DNA to move to the positive pole. This movement leads to the separation of the DNA fragments according to their size. After electrophoresis, the agarose gel was placed under ultraviolet radiation (UV light) where ethidium bromide fluoresces making the DNA bands visible. Finally, the size of the DNA fragments was identified by comparing them with a 100 bp (base pairs) DNA ladder.

3.4. Haplotype formation and linkage analysis

The genetic linkage analysis provides us with useful information about the inheritance of specific genetic markers, which might explain the clinical phenotype of the patient. In this study, highly polymorphic markers linked to *PKD1* and *PKD2* loci were used for the investigation of a family (CY1611), clinically diagnosed with ADPKD. Hence, this analysis provides information whether an individual inherits the "pathogenic" haplotype linked to *PKD1* or *PKD2* gene. Noting that, haplotype refers to a set of DNA variations or polymorphisms, that tend to be inherited together. Therefore, the haplotype of each patients was identified by using HBAP1, D16S3024, D16S3395, D16S3070 and D16S3027 markers for *PKD1* gene, and D4S395, D4S1534, D4S2460, D4S423, D4S2986 and D4S1572 markers for *PKD2* gene **(Table 2 and Figure 9)**.

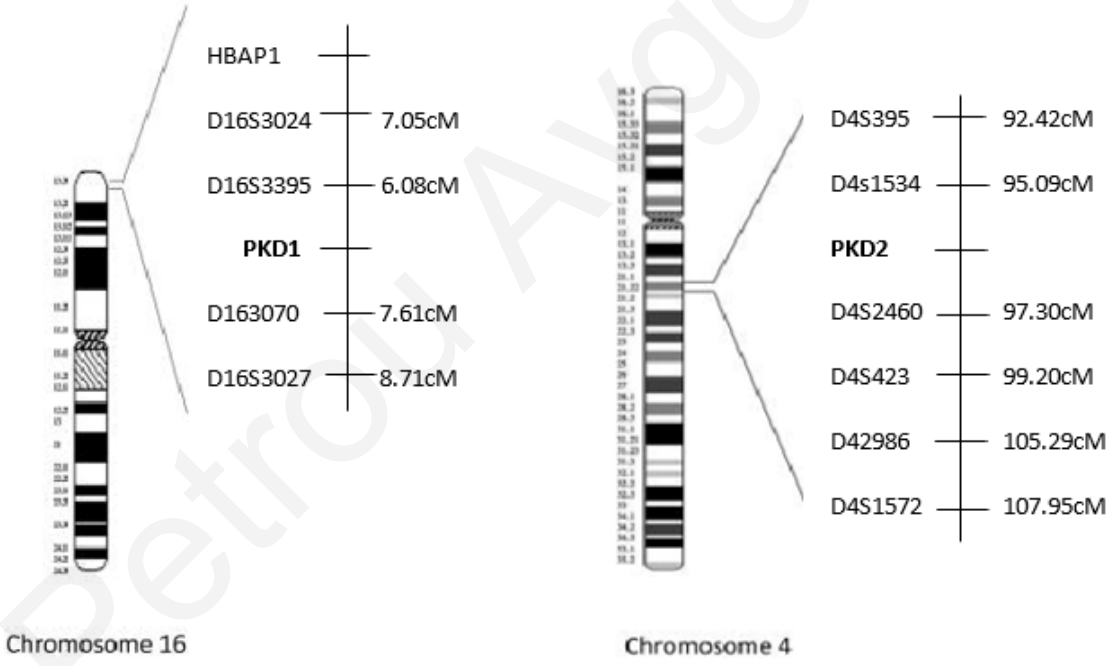


Figure 9. Map of the microsatellite markers for PKD1 and PKD2.

Consequently, the PCR product (40ng) was added in each well of a plate, with GeneScan^{tm} 500Liz^{tm} dye Size Standard and Hi-Di^{tm} Formamide. Then, capillary electrophoresis onto the 3500 Genetic Analyzer was performed, followed by data analysis. The GeneScan^{tm} 500Liz Size standard is used for sizing DNA fragments in the 35-500bp range and provides 16 single-stranded labeled fragments of 35, 50. 75, 100, 139, 150, 160, 200, 250, 300, 340, 350, 400, 450, 490 and 500 bases. Each of the DNA fragments is labeled with the LIZ® fluorophore, which results in a single peak when run under denaturing conditions. Then, the results were analyzed in the GeneMapper^{tm}. Finally, the pedigree was prepared by using Cyrillic 2.0 software.

3.5. Sanger Sequencing analysis

Sanger sequencing analysis, also known as the "chain terminator method", is a method for determining the nucleotide sequence of DNA. The chain termination PCR works like the traditional PCR, but with one major difference, the addition of modified nucleotides called dideoxyribonucleotides (ddNTPs). In this method, a low ratio of chain-terminating ddNTPs to normal deoxynucleotides (dNTPs) are used in the PCR reaction. The ddNTPs lack the 3'-OH group which is required for the phosphodiester bond formation. So, as the DNA polymerase incorporates a ddNTP at random, extension ends, because this ddNTP cannot form a phosphodiester bond with the 5'-phospathe of the next. Each of the four ddNTPs has a unique fluorescent label. More specific, red reflects thymine (T), green represent adenine (A), with black colour the guanine (G) and the blue one the cytosine (C).

Subsequently, all oligonucleotides are run in a single capillary gel electrophoresis onto the 3500 Genetic Analyzer (Life Technologies). Then, a sequencing data analysis software reads each band of the capillary gel, using fluorescence to call the identity of each terminal ddNTP. The laser excites the fluorescent tags in each band, and the software detects the resulting light emitted. Because of the different fluorescent label, the light emitted can be directly tied to the identity of the terminal ddNTP. The output is called an electropherogram, which shows the fluorescent peak of each nucleotide along the length of the template DNA. In that way if there are any mutations can be displayed in the electropherogram as a different colour compared to wild type allele and reference genome.

The coding regions (exons) of the *PKD2* gene can be easily screened by using Sanger Sequencing and this accurate method makes possible the detection of any mutations within the coding regions of the family with code CY1626. Hence, after the completion of the PCR, products were quantified after agarose gel electrophoresis and enzymatically purified with ExoSAP enzyme. This enzyme consists of restriction exonuclease to remove primers and an alkaline phosphatase to dephosphorylate dNTPs. A total concentration of 20ng from each PCR product was used for the sequencing in a reaction with BigDye1.1 at a reaction volume 20μl, along with 5x Cycle Sequencing Buffer and one of the two PCR primers (10pM/μl) (either the Forward or the Reverse). The reaction was performed in an automatic cycler and run with the following protocol, initial denaturation was performed for 1minute at 96ºC, followed by 29 cycles of 10 seconds at 96ºC and 5 seconds at 50ºC, respectively. Final extension was for 4 minutes at 60ºC.

Then, the cycle sequencing products were purified by two successive rinses with 100% and 70% ethanol. The resulting DNA pellet was dissolved in formamide and the samples were loaded into a sequencing plate in the automatic capillary electrophoresis analyzer. Sanger sequencing was performed using a standard analyzer protocol.

Subsequently, sequencing results were analyzed by alignment of sequences to a reference sequence in BioEdit. Identified variants were further analyzed with several databases such as VarSome and dbSNP, to determine their impact.

3.6. Long-range PCR

Long-range PCR (LR-PCR) is a very useful method for detecting genetic variations with high sensitivity because it can increase the size of amplicons from 3-5 kb over 30kb (Knierim et al., 2011, Tan et al., 2012, Jia et al., 2014). It can be used as a reliable tool for complex genetic analyses like the, *PKD1* gene (Tan et al., 2012). The genetic diagnosis of ADPKD1 is challenging due the high GC content of its exons and for the existence of multiple homologs in the genome (Bogdanova et al., 2001) and pseudogenes (Rossetti et al., 2001, Tan et al., 2012). Nevertheless, *Borras et al.* (2017) showed that long-read single-molecule sequencing can be an alternative strategy that could overcome *PKD1* complexities and distinguish between homologous regions of *PKD1* and its pseudogenes (Borràs et al., 2017).

In this study, we used a total of five and nine LR-PCR primers to cover the entire *PKD1* and *PKD2* coding regions **(Table 7)**, and the primers were optimized to produce amplicons of similar

sizes (>4Kb) (Borràs et al., 2017). We decided to use a tried and tested series of primers to obtain the desired long-range products, hence we adopted the primer sequences reported by *Borras et al* (2017), and then developed an amplicon-based protocol to allow their processing on our in-house NGS platform the GeneStudio, S5 (IonTorrent, ThermoFisher).

Initially, the LR-PCR reaction was performed for each fragment by using control DNA to find the appropriate conditions where the samples would work and avoid non-specific amplification products. Non-specific products detected as faint bands or smears in addition to the bands that were of interest.

Subsequently, the fragments were amplified from 100ng/μl of genomic DNA using Long Range PCR 2X GC buffer I and TakaRa LA Taq® with GC buffer, on a 20μl of PCR reaction volume with 5μM of M13-tagged primers **(Table 7)**. These amplicons were generared using the reaction mixture and PCR conditions listed in **Table 8**.

Finally, to confirm the success of the LR-PCR amplification, the final PCR product was run in a 0,8% ethidium bromide agarose gel. Moreover, 5μl of LR-PCR product was mixed with 2μl of loading dye and loaded onto the gel. Then, the agarose gel was placed under UV light for DNA visualization. To determine the size of LR-PCR fragments, this was done by comparison with a DNA ladder, which was used with bands of a certain amount per 1000 bases (1kb). Further quality assessment of PCR products was performed on an Agilent Bioanalyzer 20100, using the DNA 12000 kit. The Agilent Bioanalyzer 2100 is a capillary electrophoresis system that allows the detailed and specific analysis and sizing of DNA fragments.

Table 7: Amplified LR-PCR fragments

Amplified fragments	Exons Covered	Fragment size (Kb)	Strand*	LR-PCR Forward Primer sequence (5'->3') **
PKD1 A	1	4.3	FW	GCGGAGCGTGAAAAATAGCTCGT
			RV	TACTGCTTTGCTTGACCAGCCTTAAAGA
PKD1 B	2-13	8	FW	CCGAGTAGCTGGAACTACAGTTACACACT
			RV	CACCCAGTTACCTCCCAACAGAC
PKD1 C	14-21	7.6	FW	GTTTCCCTGTCTGTTGGGAGGTAAC
			RV	CTGCGTTCACACAGGACAGAACG
PKD1 D	22-34	8.1	FW	ATGTGAAGAGGTGCCTTGTGTGGT
			RV	TTAAAAACCCGCCCATAATTTCTCACTGC
PKD1 E	35-46	7.5	FW	GAGCAGGCTCATGGGGCTTTGTAGGAGC
			RV	ACAGCCCGCTGTACCTGAGGACTCG
PKD2 A	1	7.2	FW	GCAGGATTCTGTTGCTAGAAGTCAGTGC
			RV	CCTTTCTATCTAGCTTCTTTCCATCCCAGC
PKD2 B	2	7.9	FW	CCTGTAACTCCCACCATGGAATGGGC
			RV	AGGTAGGCTTGGAGGGTGCAACTGG
PKD2 C	3	7.5	FW	TACCCCTTAAAGATTTTCCTCACA
			RV	CTGTCGATACTCATGCATTGAAA
PKD2 D	3-5	8	FW	CTGTGTTGGGGCCTGTGCAGATCAGC
			RV	GGGGGACTTGGTGATGGAACATGTGGC
PKD2 E	5-6	7.4	FW	TTGGGACTACATTGACCTCACTAA
			RV	TTCATTCCTGTATCCCCAGTGC
PKD2 F	7-8	7.1	FW	GTTTTCTGAGCACCTACTATGTACTTGC
			RV	TAAACCTTGACAACAGTCACCCTCG
PKD2 G	9-10	7.4	FW	TGAACTCCAGGGCCTCACACTGTCC
			RV	GCGAACTTTAGACCTGACCTTGCTTTGC
PKD2 H	11-13	4.9	FW	AAATGTTGGGGCTGGACATGGTGGC
			RV	ATGCACAAGGAACATTCTTCAGGACG
PKD2 I	14-15	6.2	FW	CAGGTCTTTGTCTGCTAAGTCTGA
			RV	TTTGCAAGTGAAATGAAAAACAGT

***FW: Forward strand, RV: Reverse strand.**
****Primers included a 5' M13-tail (TGTAAAACGACGGCCAGT for FW primers and CAGGAAACAGCTATGACC for RV primers).**
*****Primers according to *Borras et al*., 2017 (Borràs et al., 2017).**

Table 8: The protocol for each LR-PCR fragment

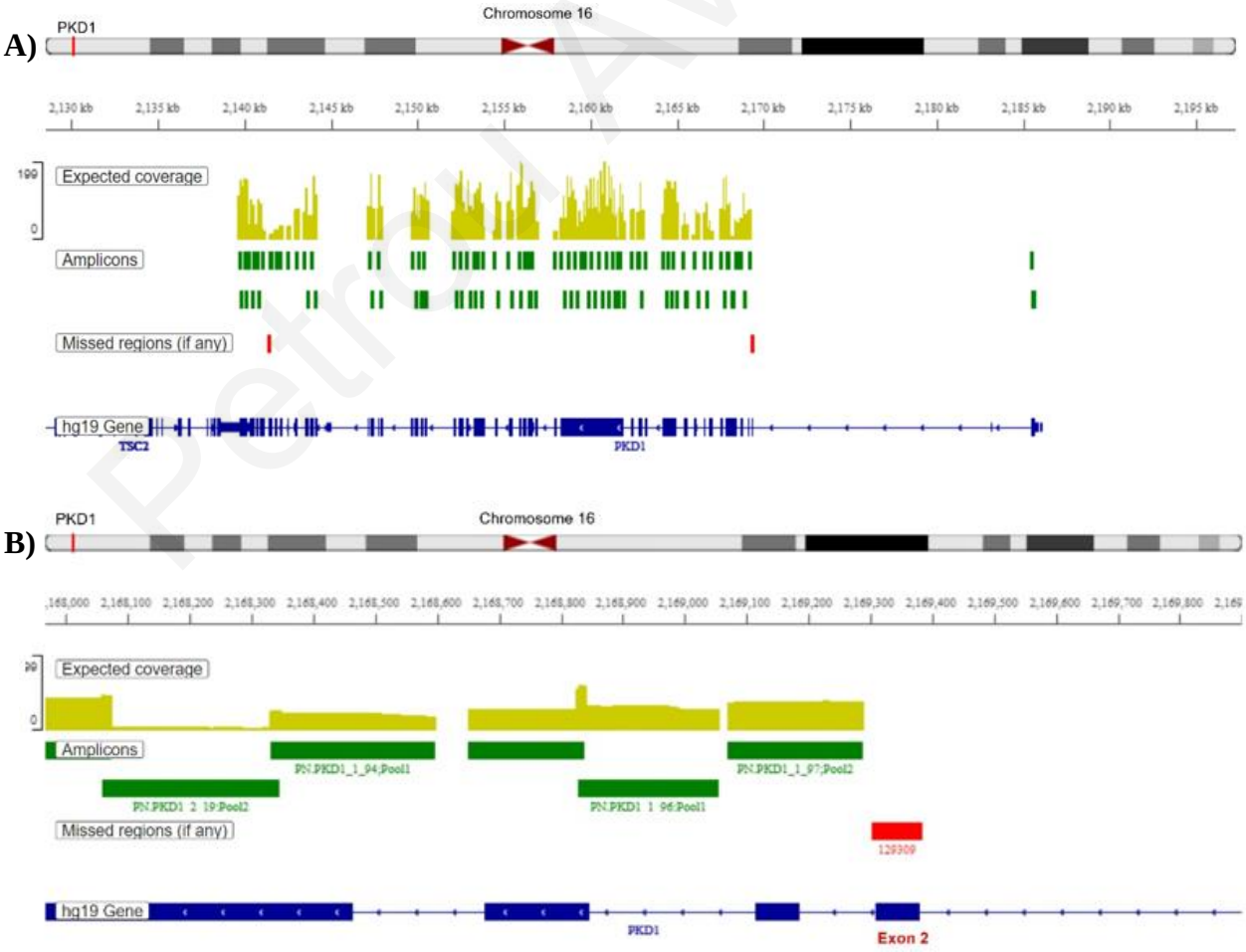
Components for LR-PCR	Stock	Final Concentration	Amount per reaction (µl)
2X GC buffer I (5mM Mg^{2+})	-	-	10 µl
dNTP Mixture	2,5 mM each	2,5 mM	3,2 µl
Forward primer	5 µM	0,4 µM	1,6 µl
Reverse primer	5 µM	0,4 µM	1,6 µl
Takara LA Taq	125 U	5 U/µl	0,2 µl
DNA template	-	100 ng/µl	1,4 µl*
ddH_20	-	-	2 µl
5% DMSO**	-	-	
Reaction Volume	-	-	20µl
PCR protocol	2-Steps***	3-Steps****	
	94°C -> 1min	94°C -> 1min	
	94°C -> 30sec	94°C -> 30sec	
	Tm -> 8min*****	60°C -> 30sec	
	68°C -> 5min	68°C -> 8min	
	4°C	68°C -> 5min	
		4°C	
Cycles	x35	x45	

* 2µl DNA for *PKD1* sets B, E *PKD2* set E
** 5% DMSO, all samples except *PKD1* set D and *PKD2* sets F, G
*** Protocol with two steps for *PKD1* sets A-E, *PKD2* sets A-H
**** Protocol with three steps for *PKD2* set I
***** Tm: 68°C for *PKD1* sets A-E, *PKD2* sets A, B, D, E, H, 62°C for *PKD2* set C, 65°C for *PKD2* sets F, G

3.7. Next-Generation Sequencing

3.7.1. Ampliseq diagnostic panel design for targeted sequencing of *PKD1* and *PKD2* genes

Genetic analysis of DNA samples with Next-Generation sequencing (NGS) technology for the detection of mutations was performed using a panel for *PKD1* and *PKD2* genes associated with ADPKD. The Ion Ampliseq Designer V7.49 (Life technologies, CA, USA) was used, using the human genome 19 (hg19 or GRCh37) as the reference genome, to design the PKD gene panel. We used primers designed through the On-Demand module of the Ampliseq Designer and included the two genes as they appear in human genome hg19. The design covers 22.58kb in 123 amplicons split into 2 pools and exon padding was automatically set to 5 nucleotides. It has an estimated coverage of 99%, which is acceptable given the complexity of the regions to be processed. The amplicon range is estimated to span between 125-275bp. In *PKD1* two regions remained out of the design, the entire exon 2, and 7 bases at the 3' end of exon42 **(Figure 10)**. These regions can be covered separately using Sanger Sequencing with nested primers in existing long-range products. *PKD2* on the other hand was designed with 100% coverage.



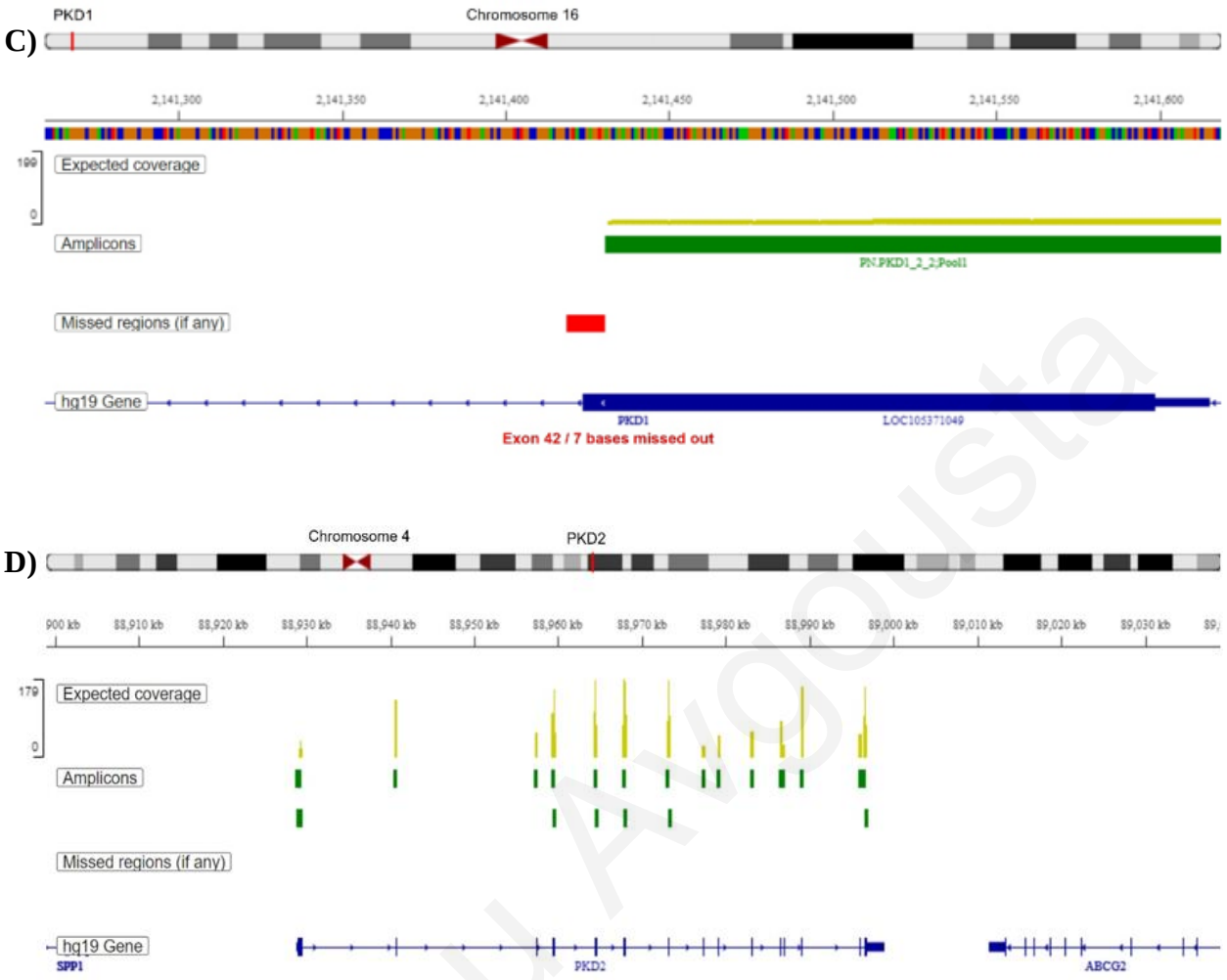


Figure 10. A) Amplicon design for the *PKD1* gene. **B)** Exon2 of the *PKD1* gene could not be covered in this design. **C)** Seven bases at the 3' end of *PKD1* as well as the splice region sites were left out in this design. **D)** Design implemented for *PKD2* gene. Blue lines represent the exons of the gene, green lines the position of the amplicons designed, red are the regions not covered in this design and yellow designates the expected coverage of each amplicon calculated by the Ampliseq Platform.

To validate this panel, we tested two samples, one with a known mutation in *PKD2* gene (UCY4521) and one with ADPKD with unknown mutation (DX0009/UCY967). Sequencing was performed on the GeneStudio S5 platform, and template preparation on the IonChef automated prep system. Following LR-PCR in the two samples, sequencing libraries were constructed after pooling all PCR samples. The strategy followed assimilated standard procedures, where 3000 copies of genomic DNA is regularly used for library construction. In this case we calculated the total copies of DNA after pooling all amplicons together and brought our mix to 3000 copies in 1.5ul. to optimize library performance in further, we also tested a mix with 30.000 copies in 1.5ul, which seemed to perform better in library construction. Hence, we proceeded with this option.

The construction of the DNA libraries was performed using the Ion Ampliseq Library Kit Plus (Life Technologies) followed by the addition of a separate barcode for each sample with the Ion Xpress Barcode Adapters (Life technologies) for multiplexing and a step of purification with Agencourt AMPure XP beads (Beckman Coulter Incorporated). Final amplification of library products was performed, and libraries were quantified with the Qubit 2.0 Fluorometer using the Qubit HS dsDNA Assay Kit (Life Technologies).

Libraries were then pooled together and amplified in further on IonChef and eventually loaded 50pM of each library onto a 510 chip for sequencing. NGS followed on the GeneStudio S5 platform. The data generated are automatically transferred into the Ion Torrent Server, where the signal analysis is done through the Ion Torrent Suite^{tm} Software which provides the number of reads or each sequence in FASTQ and SAM/BAM formats. Data were analyzed in our in-house IonReporter server using standard parameters. IonReporter provides variant annotation and filtering in an automated way. Further analysis of variants was carried out when required by using established algorithms and databases.

Notably, the IonTorrent NGS technology used in this project can detect nucleotides based on pH changes caused by the release of protons (H+) during DNA synthesis, that is, when each new nucleotide is integrated to form a phosphodiester bond with the one that precedes it on the DNA strand.

3.8. Databases - Data analysis and management

3.8.1. VarSome

VarSome is a search engine, aggregator and impact analysis tool for human genetic variation and a community-driven project aiming at sharing global expertise on human variants. It provides automated variant classification according to the guidelines of the American College of Medical Genetics and Genomics (ACMG). This classification uses specific standard terminology "pathogenic", "likely pathogenic", "uncertain significance", "likely benign" and "benign" to describe variants identified in genes that cause Mendelian disorders (Richards et al., 2015). In cases where the impact of a variant is not clear given the specific parameters of ACMG, then the variants are called as of "unknown" significance (VUS). Users can search in VarSome by gene name, transcription symbol, genomic location, or Human genome variation society (HGVS) nomenclature.

3.8.2. Single Nucleotide Polymorphism Database (dbSNP)

The dbSNP is owned by the National Center for Biotechnology Information (NCBI) and includes wealthy of Single nucleotide polymorphism (SNP) information and findings. Uniquely registered polymorphisms and findings files also receive a "*rs###*" reference identification number where "rs" stands for reference snp, and then the finding number is entered.

3.8.3. Rare Exome Variant Ensemble Learner (REVEL)

The REVEL is an ensemble method for predicting the pathogenicity of missense variants. It integrates scores from MutPred, FATHMM v2.3, VEST 3.0, SIFT, PolyPhen, MutationTaster, PROVEAN, MutationAssessor, LRT, SiPhy. GERP++, phyloP and phastCons. Scores range from 0 to 1 and variants with higher scores are predicted to be more likely to be pathogenic (Ioannidis et al., 2016).

3.8.4. Sorting Intolerant from Tolerant (SIFT)

SHIFT predicts whether an amino acid substitution is likely to affect protein function based on sequence homology and the physico-chemical similarity between the alternate amino acids. The provided data for each amino acid substitution is a score and a qualitative prediction (either 'tolerated' or 'deleterious'). The score is the normalized probability that the amino acid change is tolerated so scores nearer zero are more likely to be deleterious. The qualitative prediction is

derived from this score such that substitutions with a score < 0.05 are called 'deleterious' and all others are called 'tolerated'.

3.8.5. PolyPhen-2

PolyPhen predicts the putative impact of amino acid substitution on the protein function and is based on several parameters, including protein sequence, protein structure and conservation. The score ranges from 0 to 1 and a higher value indicates that the amino acid substitution is probably damaging (Adzhubei et al., 2010).

3.8.6. MutationTaster

MutationTaster employs a Bayes classifier to eventually predict the disease potential of an alteration. The Bayes classifier is fed with the outcome of all tests and the features of the alterations and calculates probabilities for the alteration to be either a disease mutation or a harmless polymorphism. For this prediction, the frequencies of all single features for known disease mutations/polymorphisms were studied in a large training set composed of >390,000 known disease mutations from HGMD Professional and >6,800,000 harmless SNPs and Indel polymorphisms from the 1000 Genomes Project (TGP) (Adzhubei et al., 2010).

RESULTS

4.1. Study of family CY1611

4.1.1. Optimization of PCR reaction mixture for the markers of *PKD1* and *PKD2* gene

Electrophoresis of the PCR products was performed on a 2% agarose gel to separate the molecular weight of each marker for each gene, as shown in **Table 2**. The PCR conditions which were used for each marker are given in **Tables 3 and 4**. The results of electrophoresis are shown in below **Figure 11**.

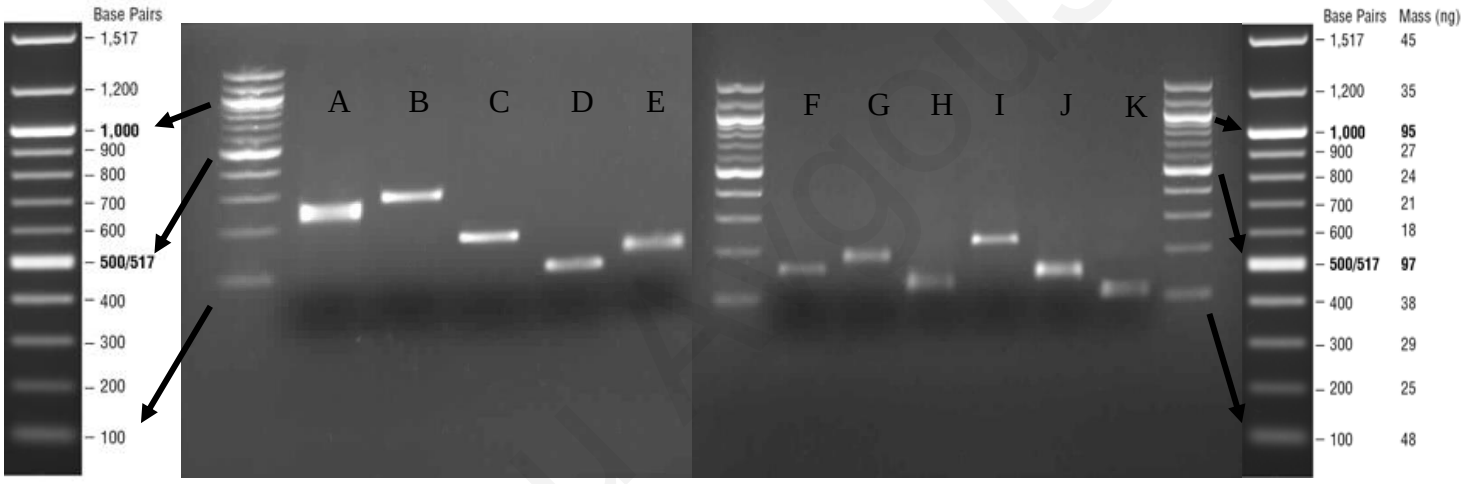


Figure 11: The left part of the figure represents the markers of *PKD1* gene, and the right part the markers of *PKD2* gene, respectively. Right and left of the figure is displayed the 100bp ladder. A: D16S3024, B: D163027, C: D163070, D: HBAP1, E: D16S3395, F: D4S1534, G: D4S2460, H: D4S423, I: D4S2986, J: D4S1572, K: D4S395.

4.1.2. Genetic linkage analysis for *PKD1* and *PKD2*

The pedigree shows how the individuals within a family are related to each other and which individuals have a particular trait or genetic condition. Three members of the family CY1611, mother, the proband and her sister (individual I:2, II:5, and II:1 respectively) have been clinically diagnosed with ADPKD. The proband and her husband, were concerned about their children if they will get the "pathogenic" gene. Therefore, we investigated this family with linkage analysis by using the following markers to examine the linkage with *PKD1* gene, HBAP1, D16S3024, D16S3395, D16S3070 and D16S3027 and the following six markers for *PKD2* gene, D4S395, D4S1534, D4S2460, D4S423, D4S2986 and D4S1572.

According to the pedigree we concluded that the ADPKD phenotype seems to be inherited with the haplotype 98-245-126-158-269 of *PKD1* markers, because all the individuals with ADPKD phenotype share the same haplotype (I:2, II:1, II:5) **(Figure 12)**. Reportedly, one of their three children, individual III:3, has inherited the putative affected haplotype, and therefore, he has very high risk of developing the disease. In addition to this, we examined the correlation with *PKD2* gene, and we excluded any correlation with this gene, because a non-affected member (II:3) of the family had the same haplotype as the patients did (I:2, II:1, II:5) **(Figure 13)**.

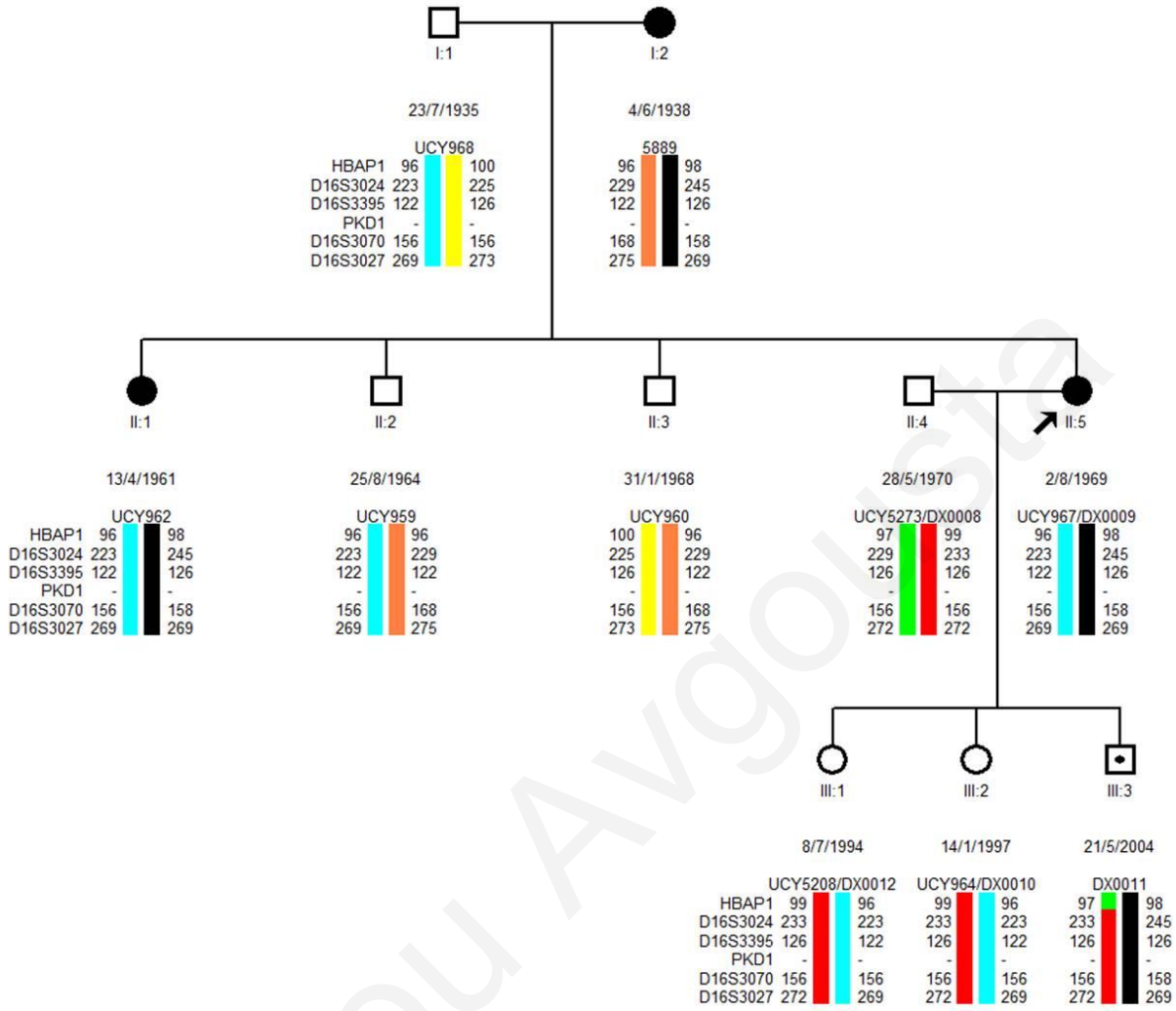


Figure 12: The Pedigree of the CY1611 family with haplotypes relared to the *PKD1* genetic locus. Below of each individual are displayed the microsatellite markers linked to *PKD1* locus. The squares represent males and circles represent females. Each generation is numbered with roman numerals, thus the top generation would be generation one, or Roman numeral I. Symbols representing individuals with clinical traits are filled with black colour. The arrow indicates the proband, the first person in a family which receive genetic counseling and/or tested for the clinical trait. The different coulours represent the haplotypes that were inherited from the previous generation, one from each parent. The individual III:3 has inherited the putative affected haplotype, and therefore, he has high risk of developing ADPKD.

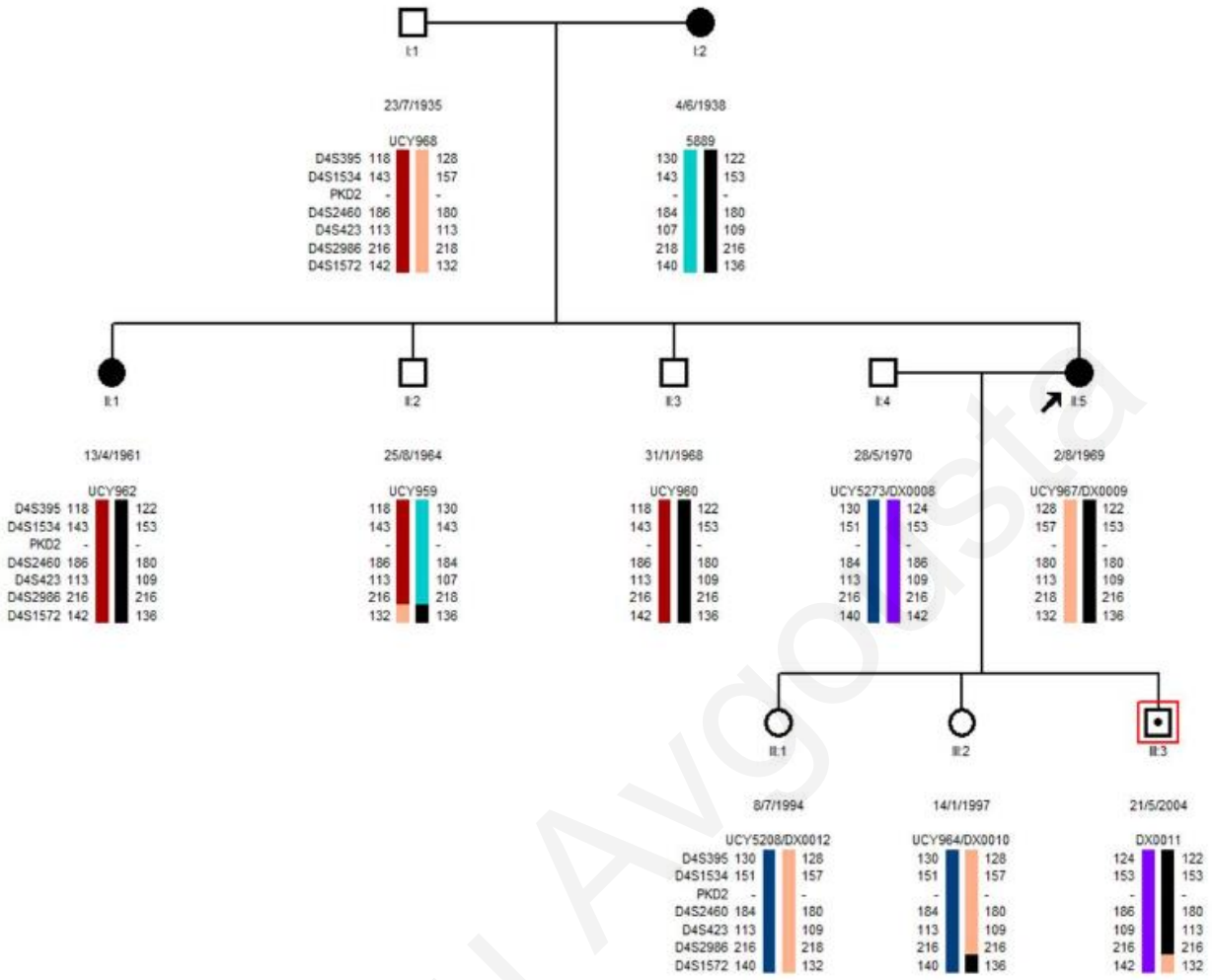


Figure 13: The Pedigree of the CY1611 family with haplotypes related to the *PKD2* genetic locus. Below of each individual are displayed the microsatellite markers linked to *PKD2* locus. Both affected (II:1 and II:5) and non-affected individuals (II:3) share the same haplotype.

4.2. Study of family CY1626

Family CY1626 consist of two generations of individuals affected by ADPKD2, based on previous results generated by the lab **(Figure 14)**. We aimed at identifying the causal mutation in *PKD2* by sequencing all its exons by Sanger sequencing. Patient UCY5176 (II:4) was used as the positive sample to help identify if the mutation was inherited by individuals UCY5177 (III:2) and UCY5178 (III:1).

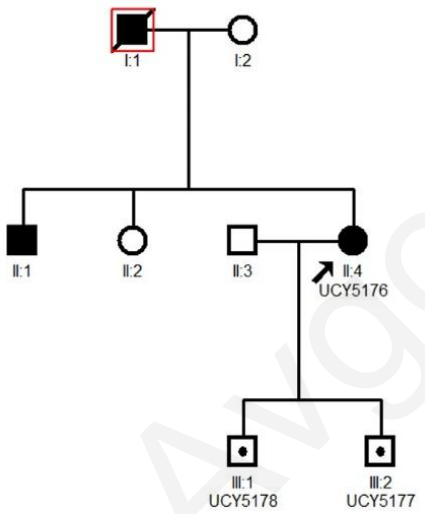


Figure 14: The pedigree of the CY1626 family

4.2.1. Optimization of PCR reaction for sequencing *PKD2* exons

Electrophoresis of the PCR products was performed on a 2% agarose gel to separate the molecular weight of each exon for *PKD2* gene, as shown in **Table 5**. The PCR conditions were used for each exon are given in **Table 6.** The results of electrophoresis are shown in **Figure 15**.

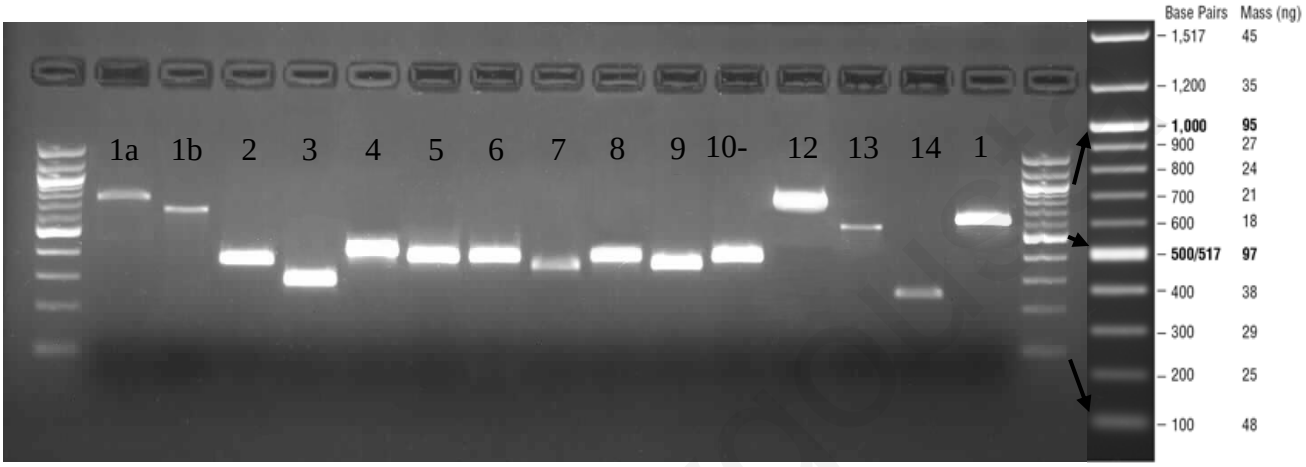


Figure 15: *PKD2* gene consist of 15 exons. The numbers represent the number of each exon.

4.2.2. Results of Sanger Sequencing analysis

The figure below shows the electropherograms which indicate the results of the Sanger Sequencing analysis. The area in which the variation is found is indicated in shaded colour. Green colour corresponds to the residue of adenine (A), black to guanine (G), red to thymine (T) and blue to cytosine (C). The C/T variation is characterized by the letter Y signifying a Pyrimidine at that position (C or T) **(Figure 16)**. The results are listed in **Table 9** in more details.

UCY5177 exon1 *PKD2*:c.568G>A (rs117078377)

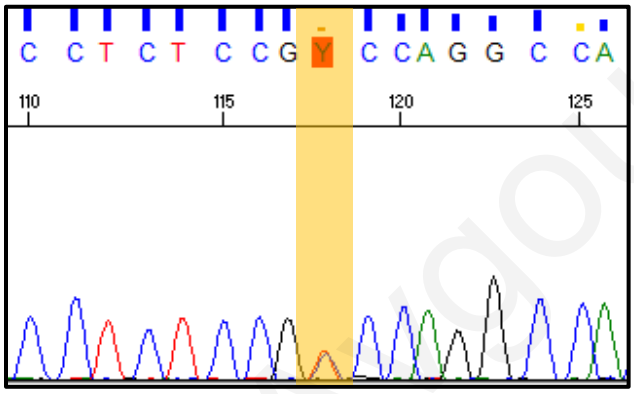


Figure 16: The electropherogram represents the DNA of the sample UCY5177 and has the mutation *PKD2*:c.568G>A. The image depicts sequencing with a reverse orientation (C>T or G>A).

Table 9: Results of family CY1626 and analysis with four different databases.

Family code	Samples (Males / Females)	Exons	Coding region	Protein	dbSNP	VarSome	Mutation Taster	SIFT	REVEL	MAF
CY1626	UCY5177 (M)	Exon 1	*PKD2*:c.568G>A	pAla190Thr	rs117078377	Benign	Polymorphism	Tolerated	Benign	0.02

4.3. Optimization of Long-range PCR

Electrophoresis of the LR-PCR products was performed on a 0,8% agarose gel to separate the molecular weight of each DNA fragment for each gene, as shown in **Table 7**. The LR-PCR conditions which were used for each fragment are given in **Table 8.** The results of electrophoresis are shown in **Figure 17**. Also, the results were analyzed by Agilent Bioanalyzer 2100 as shown in **Figure 18**.

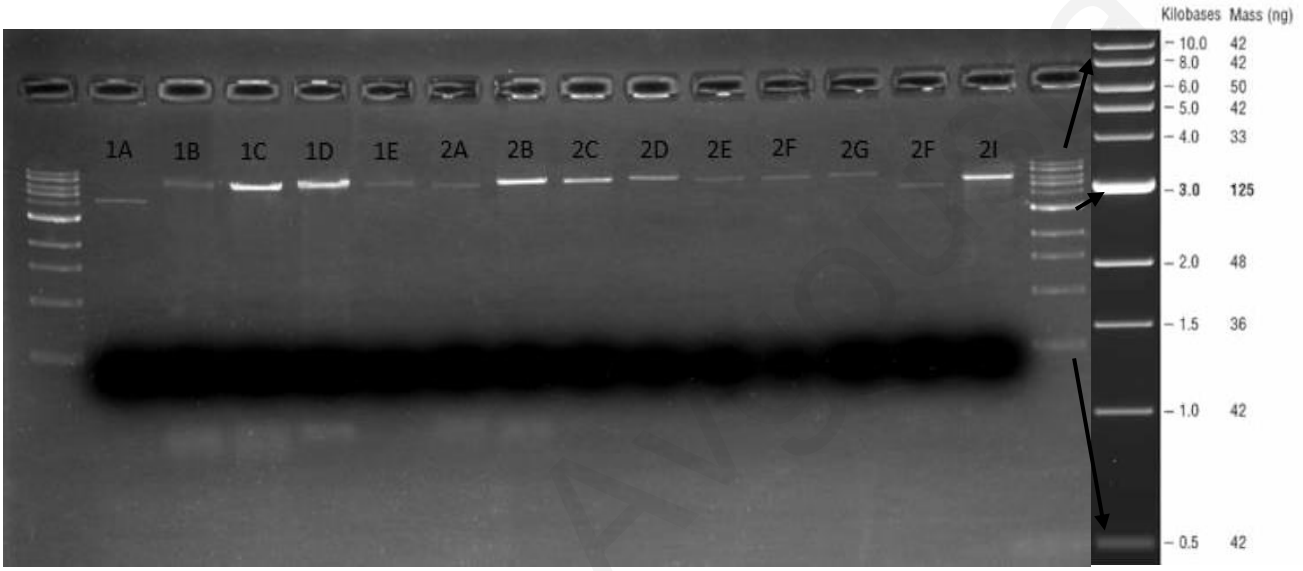


Figure 17: Number 1 represents *PKD1*, and number 2 represents *PKD2* gene, respectively. The letters represent the amplified fragments as shown in **Table 7**. One the right side of the figure the 1kb DNA ladder is displayed.

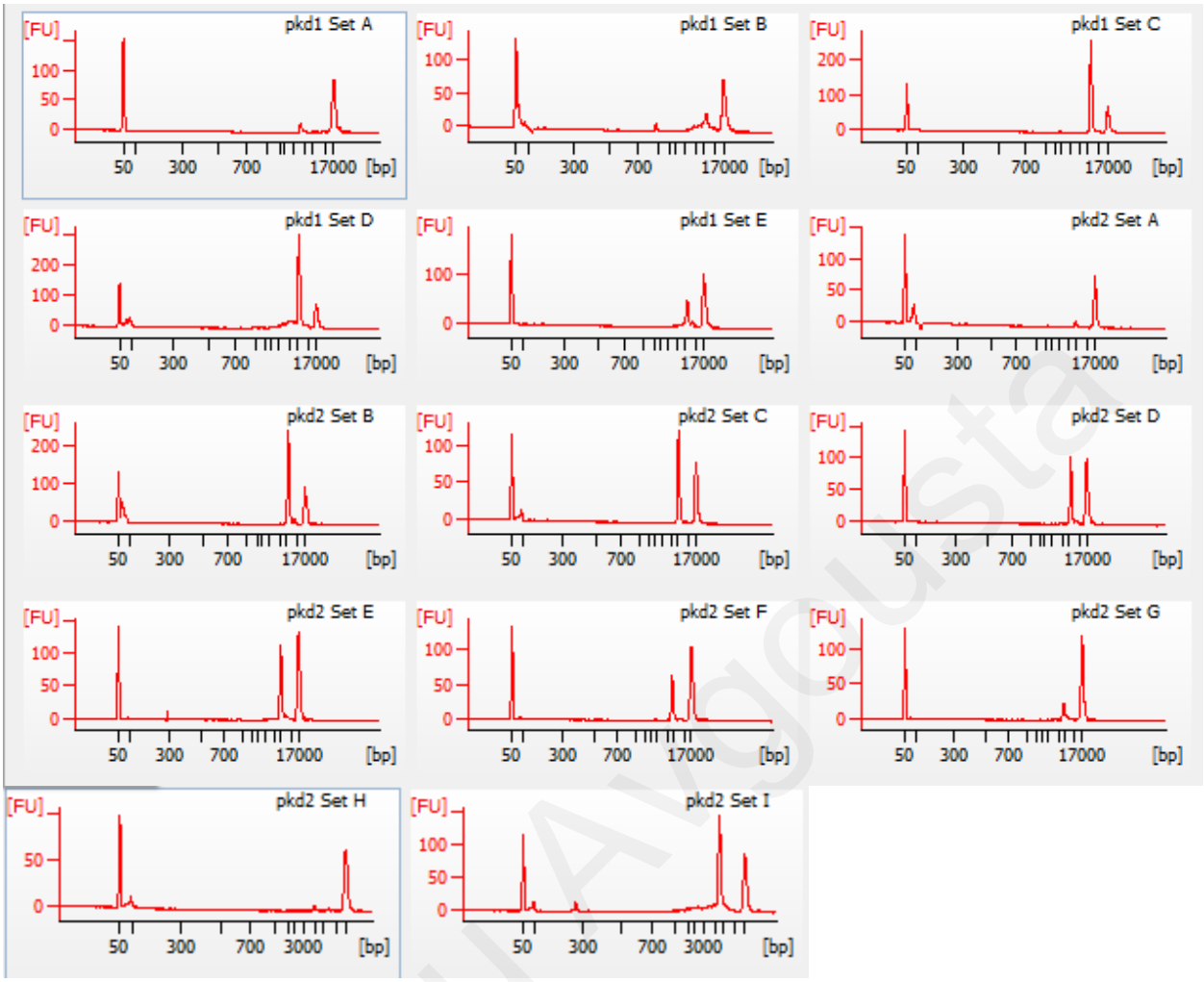


Figure 18. Size of each DNA fragment based on Agilent Bioanalyzer 2100.

4.4. Sequencing of ADPKD genes

We pooled samples together and run them on an S5 510 chip. BED (Browser extensible data) files containing the designed regions were uploaded onto the S5 Torrent Server to be used for mapping the sequences obtained. Having a mean length of amplicons of about 250bp we initially tested diluting our libraries with a factor of 17 to achieve 100pM, according to the manufacturer's instructions. Sequencing analysis worked extremely well with a mean depth of 411 and 363 for samples DX0009 and UCY4125, respectively **(Figure 19)**. In total NGS produced >40.000 mapped reads per sample. Although the mean coverage per sample was quite high, there were two regions that were not successfully covered. These were parts of *PKD1* and more specifically, exon 12 and the beginning of exon 1. *PKD2* was fully covered. Hence, we need to optimize this pipeline in further, by adjusting the concentration of PCR fragments accordingly and achieve maximum coverage. Otherwise, we need to revise the primer design for our panel to include all missed regions, including regions left out of the initial design (exon 2 and the end of exon 42).

Subsequently, the results were analyzed, and variants were annotated with the IonReporter V5.12, using standard parameters. Variants identified per sample were examined using different parameters. Firstly, coding regions were considered, and the MAF as well as ClinVar and variant prediction algorithm results were considered. Then we looked for variants at splice donor and acceptor sites (set as +/- 3 nucleotides away from the end or beginning of exons).

The first sample (UCY4521) had an already identified mutation in *PKD2* gene and was used as a positive control to test this panel. Indeed, the mutation p.Arg742Ter was successfully identified by our NGS approach **(Table 10 and Figure 19).** This result demonstrating the effectiveness of the assay for the *PKD2* gene. The second sample was the patient (DX0009/UCY967) from the family CY1611 which was showed to be linked to *PKD1* gene locus (Figure 12). After the completion of NGS and analysis we identified a variant in exon 16 of *PKD1*, c.6916A>T, p.Arg2306Trp **(Table 10 and Figure 19)**. This variant was examined according to certain algorithms such as MutationTaster (Disease Causing), SIFT (0.01), POLYPHEN (0.973) **(Table 10)**. In VarSome the variant appears as Likely Benign according to ACMG criteria, although it is a newly discovered variants as it does not appear in gnomad exomes or genomes. Impressively, this variant alters the first base at the start of exon 16, right after the splice acceptor site. To test its effect on splicing dynamics, we used the HSF (Human splicing finder) algorithm that explores the probability of alterations occurring at sites of splicing. Reportedly, this variant

is considered to cause a significant alteration in ESE/ESS (Exonic Splicing Enhancers/ Exonic Splicing Silences) motifs ratio. Therefore, it is possible that the variant has a functional role and will be examined in further in other affected family members to test for segregation.

DX0009: *PKD1*:c.6916A>T, p.Arg2306Trp

UCY4521: *PKD2*:c.2224C>T, p.Arg742Ter

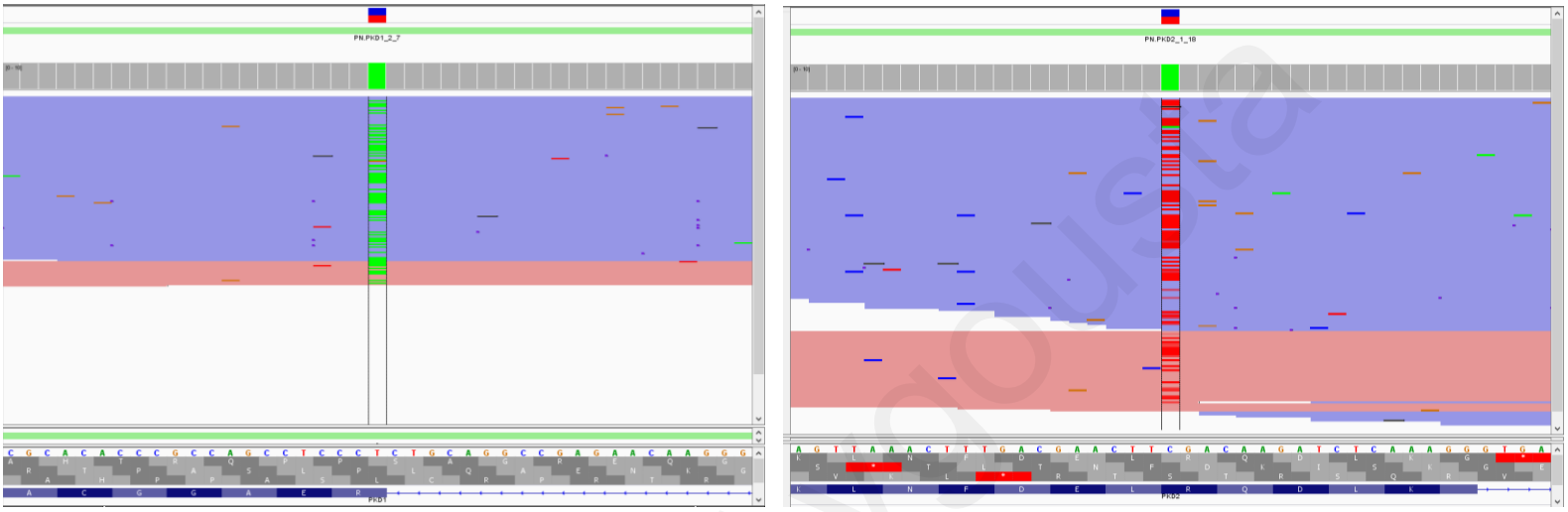


Figure 19: Visualization of the changes by IGV (Integrative Genomic Viewer). The left panel displays the variant identified in sample DX0009 and the right panel the previously known mutation in *PKD2* in sample UCY4521. Green lines represent the substitution of A to T in position 6919 of exon 16 in *PKD1*. Red lines represent the substitution of C to T in position 2224 of exon 11 in *PKD2*.

Table 10: Results of NGS and analysis with several databases and algorithms

Samples	Exons	Coding region	Protein	dbSNP	VarSome	Mutation Taster	SIFT	PolyPhen
UCY967/ DX0009	16	*PKD1*:c.6916A>T	p.Arg2306Trp	-	Likely benign	Disease causing	0.01	0.973
UCY4521	11	*PKD2*:c.2224C>T	p.Arg742Ter	Rs121918040	Pathogenic	Disease causing automatic	-	-

DISCUSSION

ADPKD is one of the most common inherited kidney diseases as it affects 1 in 400 to 1 in 1000 individuals and is mainly caused by mutations in two genes, *PKD1* and *PKD2*. The purpose of this study was to investigate two Cypriot families, clinically diagnosed with ADPKD. Hence, we investigated these two candidate *PKD1* and *PKD2* genomic regions in one of the families available.

Initially, the family CY1611 was investigated by linkage analysis, a method which gives useful information about the inheritance of specific genetic markers flanking the *PKD1* and *PKD2* gene loci (Table 2). As mentioned in the introduction families with multiple affected members, preferably large pedigrees, are used in linkage analysis and tests the co-segregation of a chromosomal region marked by polymorphic genetic markers with a trait locus. Therefore, the first approach for investigating of this family was through the genetic linkage analysis.

The disease trait in family CY1611 appeared to be linked to the *PKD1* gene locus (Figure 12). We concluded that because all the individuals with ADPKD phenotype (I:2, II:1, II:5) share the same haplotype. Evidently, individual III:3, has inherited the putative affected haplotype, and therefore, he has very high risk of developing the disease (Figure 12). Except from the investigation of *PKD1* gene locus linkage, *PKD2* gene locus was also analyzed with different genetic markers (Table 2). Based on data presented in Figure 13, the *PKD2* region was excluded as a non-affected member (II:3) of the family had the same haplotype as the patients did (I:2, II:1, II:5). It is worth noting that the genetic marker D4S1572 is highly informative according to the results (Figure 13).

The second part of this study was the investigation of family CY1626, which was previously known to be linked to the *PKD2* gene locus. In contrast to *PKD1* gene, *PKD2* can be easily screened, by using Sanger Sequencing analysis. This method can detect candidate mutations within the coding regions of the samples.

Initially, primers for the 15 exons of *PKD2* gene were designed and then the optimization of the PCR conditions followed (Tables 5 and 6 and Figure 15). Subsequently, by using Sanger sequencing analysis the coding regions of each sample (UCY5176, UCY5177, UCY5178) followed. More specifically, a variant was identified in sample UCY5177 in exon 1 of *PKD2* (c:568G>A, rs117078377) as shown in the electropherogram (Figure 16). According to the VarSome database and variant scores provided by REVEL, MutationTaster and SIFT algorithms,

this variant is characterized as not having a pathogenic effect (Table 9). This variant cannot be the cause of ADPKD disease, because initially the affected mother (UCY5176) does not have this variation, and secondly it was characterized as a benign variant. Therefore, this variant could be a single nucleotide polymorphism (SNP). Also, based on Minor allele frequency (MAF), this variant is considered as rare in the general population, hence it does not seem to be associated to the disease (Table 9).

Although all exons of the *PKD2* gene were analyzed, no pathogenic mutations were found, not even to the mother, who was clinically diagnosed with ADPKD. This fact could be due to various reasons. Firstly, there might be changes within introns that affect splicing, or there can be a large deletion that might potentially include whole exons or specific changes in the 5' and 3' untranslated regions (5'-UTRs and 3'-UTRs) regulating the expression of the gene. In addition to this, there might be mutations in other regulatory regions, such as the promoter (to date, no mutations in these regions have been described). The changes mentioned above, might be missed with Sanger Sequencing analysis the way it was designed and implemented. Finally, there are cystic diseases with a similar phenotype/clinical course like ADPKD which can behave like phenocopies, due to several mutations within the genes *HNF1B, SEC63, PRKSCSH,* or even the Autosomal Recessive form of PKD occurring due to mutations in *PKHD1*, although the latter seem to be an unlike scenario (Torra Balcells and Ars Criach, 2011). Therefore, it would be interesting to investigate these genes in our samples in the future. As this family has been previously linked to *PKD*, it is likely that haplotype analysis was biased due to the small number of family member and affected individuals. Hence, it is possible that family CY1626 has a completely different diagnosis.

Failing to detect a mutation can also be due to the possibility that the causative mutation is missed due to the loss of heterozygosity (LOH) in one or more exons. LOH is a molecular phenomenon that results the loss one of the two alleles at a specific locus (Koptides et al., 1998). The automatic analyzer shows the sequence one of the two alleles. Thus, if the mutation is a deletion which involves one or more whole exons, Sanger sequencing analysis cannot detect it and the signal procedure presents all bases as falsely wild-type homozygotes.

In case where the pathogenic mutations for ADPKD are found in noncoding regions (introns) or in the 5'-UTRs and 3'-UTRs, then these regions need further investigation. One approach to amplify and sequence the entire *PKD2* gene is LR-PCR and include the whole genomic region.

The LR-PCR can amplify over 5kb in length which typically cannot be amplified using conventional PCR (Jia et al., 2014). Otherwise, whole genome sequencing could be also used in this case. On the other hand, if the causative mutation is a large deletion encompassing one or more exons, there are alternative methods which can detect this kind of mutations such as Southern blotting, quantitative PCR, array comparative genomic hybridization (CGH), or multiplex ligation dependent probe amplification (MLPA) (Hjelm et al., 2010).

The major goal of this study was the development of an effective diagnostic approach for ADPKD patients. As mentioned above, the screening of *PKD1* is challenging due to difficulties emerging by the complexity of the *PKD1* genomic locus. This is partially due to the high homology for most of the gene's sequence with six pseudogenes as well as high the GC content of its sequence (Rossetti et al., 2007, Tan et al., 2009, Qi et al., 2013). Therefore, to overcome these complexities we have developed a new methodology using LR-PCR amplified fragments covering all exons of both *PKD1* and *PKD2* paired with amplicon-based NGS. Hence, we managed to directly screen all LR-PCR amplified fragments with high specificity for the actual genomic sequence of *PKD1,* as well as *PKD2.* Using this approach and considering that ADPKD is a rare disease, we make possible the simultaneous screening of both genes without a prior need to form haplotype analysis in smaller families with limited availability of phenotype-positive samples. We managed to adopt a previously published protocol for long-read sequencing of *PKD1* and *PKD2* using PacBio Sequencing by *Borras et al* (2017), who showed that longer sequencing reads discriminated between *PKD1* and pseudogenes and improved the mapping quality of *PKD1* (Borràs et al., 2017). In our case, the S5 GeneStudio NGS system can handle fragments up to 275bp in standard protocols and we proceeded to design an amplicon-based protocol to reduce sequencing amplicon size by nested PCR primers, while using LR-PCR products as templates instead of genomic DNA.

Briefly, sequencing analysis of both PKD genes was successfully completed through the pipeline we developed. Further optimization is required, as two regions of *PKD1,* were not covered by NGS. We tested two samples by using this novel approach and managed to detect the previously known mutation p.Arg742Ter in *PKD2* in our positive control (UCY4521) (Table 10 and Figure 19). In family CY1611, which was linked to *PKD1* locus (Figure 12), we identified the variant c.6916A>T, p.Arg2306Trp in *PKD1* (Table 10 and Figure 19). This variant appears to alter

splicing dynamics according to HSF (Human splicing finder) and will be studied in further in other affected members of CY1611.

Conclusively, we managed to set up an NGS pipeline for the simultaneous sequencing analysis and mutation detection in *PKD1* and *PKD2*. Although there is further optimization required to establish this assay as a diagnostic tool, we have successfully detected variants in both genes. The importance of having such methods available, lies in the increasing need for diagnosing ADPKD, especially, in Cyprus. Hence, this study sets the first standard towards this path. Nevertheless, additional positive controls need to be processed, as well as validate the results of this study in further.

CONCLUSION

In conclusion, the study of two Cypriot families was carried out. The first family with code CY1611 was successfully studied by linkage analysis and appeared to be linked to the *PKD1* gene locus. In addition, we identified a variant in *PKD1* with a probable pathogenic effect in one member of this family (DX0009). On the other hand, *PKD2* was analyzed by Sanger sequencing analysis in family CY1626. However, no pathogenic variants were detected. Hence, further investigation is needed to identify the genetic cause of the renal disease reported in this family.

Finally, the biggest achievement in this study was the development of a new sequencing pipeline for screening *PKD1* and *PKD2*. We used amplified long-range PCR products of *PKD1* and *PKD2* genes paired with amplicon based NGS, which allowed the simultaneous screening of both genes. Although the first results were encouraging for both samples processed, we need to optimize this pipeline in further. Therefore, we will test more samples with this panel and optimize our protocol to achieve maximum coverage, especially for *PKD1* in order to have a complete and reliable diagnostic test for ADPKD. This achievement will be very useful for early and accurate diagnosis of ADPKD patients.

ABBREVIATIONS

ABBREVIATION	MEANING
ADPKD	Autosomal dominant polycystic kidney disease
ESRD	End-stage renal failure
PCT	Proximal convoluted tubule
DCT	Distal convoluted tubule
HGs	Homologous Genes
PC1	Polycystin-1
PC2	Polycystin-2
TRP	Transient receptor protein
TM	Transmembrane domain
ER	Endoplasmic reticulum
CICR	Calcium-induced calcium-release
mTOR	Mammalian target of rapamycin
cAMP	Cyclic adenosine monophosphate
CFTR	Cystic fibrosis transmembrane conductance regulator
CDKs	Cyclin-dependent kinases
LOH	Loss of heterozygosity
ADH	Antidiuretic hormone

FDA	Food and Drug Administration
MRI	Magnetic resonance imaging
CT	Computed tomography
cM	Centimorgans
PCR	Polymerase chain reaction
LR-PCR	Long-range polymerase chain reaction
dNTPs	Deoxynucleotides
ddNTPs	Dideoxyribonucleotides
TBE	Tris Borate EDTA buffer
UV	Ultraviolet
NGS	Next-generation sequencing
MLPA	Multiplex ligation dependent probe amplification
CGH	Comparative genomic hybridization
HGVS	Human Genome Variation Society
NCBI	National Center of Biotechnology Information
SNP	Single nucleotide polymorphism
REVEL	Rare Exome Variant Ensemble Learner
SIFT	Sorting Intolerant from Tolerant
UTRs	Untranslated regions

MAF	Minor allele frequency
CGH	Comparative genomic hybridization
HSF	Human splicing finder
BED	Browser extensible data
ESEs	Exonic splicing enhancers
ESSs	Exonic splicing silences
IGV	Integrative Genomic Viewer

BIBLIOGRAPHY

ADZHUBEI, I. A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V. E., GERASIMOVA, A., BORK, P., KONDRASHOV, A. S. & SUNYAEV, S. R. 2010. A method and server for predicting damaging missense mutations. *Nat Methods,* 7**,** 248-9.

ALBAQUMI, M., SRIVASTAVA, S., LI, Z., ZHDNOVA, O., WULFF, H., ITANI, O., WALLACE, D. P. & SKOLNIK, E. Y. 2008. KCa3.1 potassium channels are critical for cAMP-dependent chloride secretion and cyst growth in autosomal-dominant polycystic kidney disease. *Kidney Int,* 74**,** 740-9.

AUDRÉZET, M. P., CORNEC-LE GALL, E., CHEN, J. M., REDON, S., QUÉRÉ, I., CREFF, J., BÉNECH, C., MAESTRI, S., LE MEUR, Y. & FÉREC, C. 2012. Autosomal dominant polycystic kidney disease: comprehensive mutation analysis of PKD1 and PKD2 in 700 unrelated patients. *Hum Mutat,* 33**,** 1239-50.

BABICH, V., ZENG, W. Z., YEH, B. I., IBRAGHIMOV-BESKROVNAYA, O., CAI, Y., SOMLO, S. & HUANG, C. L. 2004. The N-terminal extracellular domain is required for polycystin-1-dependent channel activity. *J Biol Chem,* 279**,** 25582-9.

BARUA, M. & PEI, Y. 2010. Diagnosis of autosomal-dominant polycystic kidney disease: an integrated approach. *Semin Nephrol,* 30**,** 356-65.

BASTOS, A. P., PIONTEK, K., SILVA, A. M., MARTINI, D., MENEZES, L. F., FONSECA, J. M., FONSECA, II, GERMINO, G. G. & ONUCHIC, L. F. 2009. Pkd1 haploinsufficiency increases renal damage and induces microcyst formation following ischemia/reperfusion. *J Am Soc Nephrol,* 20**,** 2389-402.

BEAR, J. C., PARFREY, P. S., MORGAN, J. M., MARTIN, C. J. & CRAMER, B. C. 1992. Autosomal dominant polycystic kidney disease: new information for genetic counselling. *Am J Med Genet,* 43**,** 548-53.

BLUMENFELD, J. D. 2009. Pretransplant genetic testing of live kidney donors at risk for autosomal dominant polycystic kidney disease. *Transplantation,* 87**,** 6-7.

BOGDANOVA, N., MARKOFF, A., GERKE, V., MCCLUSKEY, M., HORST, J. & DWORNICZAK, B. 2001. Homologues to the first gene for autosomal dominant polycystic kidney disease are pseudogenes. *Genomics,* 74**,** 333-41.

BORRàS, D. M., VOSSEN, R., LIEM, M., BUERMANS, H. P. J., DAUWERSE, H., VAN HEUSDEN, D., GANSEVOORT, R. T., DEN DUNNEN, J. T., JANSSEN, B., PETERS, D. J. M., LOSEKOOT, M. & ANVAR, S. Y. 2017. Detecting PKD1 variants in polycystic kidney disease patients by single-molecule long-read sequencing. *Hum Mutat,* 38**,** 870-879.

BOUBA, I., KOPTIDES, M., MEAN, R., COSTI, C. E., DEMETRIOU, K., GEORGIOU, I., PIERIDES, A., SIAMOPOULOS, K. & DELTAS, C. C. 2001. Novel PKD1 deletions and missense variants in a cohort of Hellenic polycystic kidney disease families. *Eur J Hum Genet,* 9**,** 677-84.

BRASIER, J. L. & HENSKE, E. P. 1997. Loss of the polycystic kidney disease (PKD1) region of chromosome 16p13 in renal cyst cells supports a loss-of-function model for cyst pathogenesis. *J Clin Invest,* 99**,** 194-9.

BROOK-CARTER, P. T., PERAL, B., WARD, C. J., THOMPSON, P., HUGHES, J., MAHESHWAR, M. M., NELLIST, M., GAMBLE, V., HARRIS, P. C. & SAMPSON, J. R. 1994. Deletion of the TSC2 and PKD1 genes associated with severe infantile polycystic kidney disease--a contiguous gene syndrome. *Nat Genet,* 8**,** 328-32.

BUKANOV, N. O., MORENO, S. E., NATOLI, T. A., ROGERS, K. A., SMITH, L. A., LEDBETTER, S. R., OUMATA, N., GALONS, H., MEIJER, L. & IBRAGHIMOV-BESKROVNAYA, O. 2012. CDK inhibitors R-roscovitine and S-CR8 effectively block renal and hepatic cystogenesis in an orthologous model of ADPKD. *Cell Cycle,* 11**,** 4040-6.

BUKANOV, N. O., SMITH, L. A., KLINGER, K. W., LEDBETTER, S. R. & IBRAGHIMOV-BESKROVNAYA, O. 2006. Long-lasting arrest of murine polycystic kidney disease with CDK inhibitor roscovitine. *Nature,* 444**,** 949-52.

BURTON, P. R., TOBIN, M. D. & HOPPER, J. L. 2005. Key concepts in genetic epidemiology. *Lancet,* 366**,** 941-51.

BYCROFT, M., BATEMAN, A., CLARKE, J., HAMILL, S. J., SANDFORD, R., THOMAS, R. L. & CHOTHIA, C. 1999. The structure of a PKD domain from polycystin-1: implications for polycystic kidney disease. *Embo j,* 18**,** 297-305.

CALVET, J. P. 2008. Strategies to inhibit cyst formation in ADPKD. *Clin J Am Soc Nephrol,* 3**,** 1205-11.

CASUSCELLI, J., SCHMIDT, S., DEGRAY, B., PETRI, E. T., CELIĆ, A., FOLTA-STOGNIEW, E., EHRLICH, B. E. & BOGGON, T. J. 2009. Analysis of the cytoplasmic interaction between polycystin-1 and polycystin-2. *Am J Physiol Renal Physiol,* 297**,** F1310-5.

CELIĆ, A., PETRI, E. T., DEMELER, B., EHRLICH, B. E. & BOGGON, T. J. 2008. Domain mapping of the polycystin-2 C-terminal tail using de novo molecular modeling and biophysical analysis. *J Biol Chem,* 283**,** 28305-12.

CHANG, M. Y. & ONG, A. C. 2008. Autosomal dominant polycystic kidney disease: recent advances in pathogenesis and treatment. *Nephron Physiol,* 108**,** p1-7.

CHANG, M. Y. & ONG, A. C. 2012. Mechanism-based therapeutics for autosomal dominant polycystic kidney disease: recent progress and future prospects. *Nephron Clin Pract,* 120**,** c25-34; discussion c35.

CHAPIN, H. C. & CAPLAN, M. J. 2010. The cell biology of polycystic kidney disease. *J Cell Biol,* 191**,** 701-10.

CORNEC-LE GALL, E., ALAM, A. & PERRONE, R. D. 2019. Autosomal dominant polycystic kidney disease. *Lancet,* 393**,** 919-935.

DALGAARD, O. Z. 1957. Bilateral polycystic disease of the kidneys; a follow-up of two hundred and eighty-four patients and their families. *Acta Med Scand Suppl,* 328**,** 1-255.

DAOUST, M. C., REYNOLDS, D. M., BICHET, D. G. & SOMLO, S. 1995. Evidence for a third genetic locus for autosomal dominant polycystic kidney disease. *Genomics,* 25**,** 733-6.

DAVIDOW, C. J., MASER, R. L., ROME, L. A., CALVET, J. P. & GRANTHAM, J. J. 1996. The cystic fibrosis transmembrane conductance regulator mediates transepithelial fluid secretion by human autosomal dominant polycystic kidney disease epithelium in vitro. *Kidney Int,* 50**,** 208-18.

DE RYCKE, M., GEORGIOU, I., SERMON, K., LISSENS, W., HENDERIX, P., JORIS, H., PLATTEAU, P., VAN STEIRTEGHEM, A. & LIEBAERS, I. 2005. PGD for autosomal dominant polycystic kidney disease type 1. *Mol Hum Reprod,* 11**,** 65-71.

DELTAS, C. C. 2001. Mutations of the human polycystic kidney disease 2 (PKD2) gene. *Hum Mutat,* 18**,** 13-24.

DEMETRIOU, K., TZIAKOURI, C., ANNINOU, K., ELEFTHERIOU, A., KOPTIDES, M., NICOLAOU, A., DELTAS, C. C. & PIERIDES, A. 2000. Autosomal dominant polycystic kidney disease-type 2. Ultrasound, genetic and clinical correlations. *Nephrol Dial Transplant,* 15**,** 205-11.

DERE, R., WILSON, P. D., SANDFORD, R. N. & WALKER, C. L. 2010. Carboxy terminal tail of polycystin-1 regulates localization of TSC2 to repress mTOR. *PLoS One,* 5**,** e9239.

DICKS, E., RAVANI, P., LANGMAN, D., DAVIDSON, W. S., PEI, Y. & PARFREY, P. S. 2006. Incident renal events and risk factors in autosomal dominant polycystic kidney disease: a population and family-based cohort followed for 22 years. *Clin J Am Soc Nephrol,* 1**,** 710-7.

DISTEFANO, G., BOCA, M., ROWE, I., WODARCZYK, C., MA, L., PIONTEK, K. B., GERMINO, G. G., PANDOLFI, P. P. & BOLETTA, A. 2009. Polycystin-1 regulates extracellular signal-regulated kinase-dependent phosphorylation of tuberin to control cell size through mTOR and its downstream effectors S6K and 4EBP1. *Mol Cell Biol,* 29**,** 2359-71.

ECDER, T., MELNIKOV, V. Y., STANLEY, M., KORULAR, D., LUCIA, M. S., SCHRIER, R. W. & EDELSTEIN, C. L. 2002. Caspases, Bcl-2 proteins and apoptosis in autosomal-dominant polycystic kidney disease. *Kidney Int,* 61**,** 1220-30.

EISENBERGER, T., DECKER, C., HIERSCHE, M., HAMANN, R. C., DECKER, E., NEUBER, S., FRANK, V., BOLZ, H. J., FEHRENBACH, H., PAPE, L., TOENSHOFF, B., MACHE, C., LATTA, K. & BERGMANN, C. 2015. An efficient and comprehensive strategy for genetic diagnostics of polycystic kidney disease. *PLoS One,* 10**,** e0116680.

EVERSON, G. T. 1993. Hepatic cysts in autosomal dominant polycystic kidney disease. *Am J Kidney Dis,* 22**,** 520-5.

FATEHI, R., KHOSRAVI, S., ABEDI, M., SALEHI, R. & GHEISARI, Y. 2017. Heterozygosity analysis of polycystic kidney disease 1 gene microsatellite markers for linkage analysis of autosomal dominant polycystic kidney disease type 1 in the Iranian population. *J Res Med Sci,* 22**,** 102.

FISCHER, E., LEGUE, E., DOYEN, A., NATO, F., NICOLAS, J. F., TORRES, V., YANIV, M. & PONTOGLIO, M. 2006. Defective planar cell polarity in polycystic kidney disease. *Nat Genet,* 38**,** 21-3.

GARCIA-GONZALEZ, M. A., JONES, J. G., ALLEN, S. K., PALATUCCI, C. M., BATISH, S. D., SELTZER, W. K., LAN, Z., ALLEN, E., QIAN, F., LENS, X. M., PEI, Y., GERMINO, G. G. & WATNICK, T. J. 2007. Evaluating the clinical utility of a molecular genetic test for polycystic kidney disease. *Mol Genet Metab,* 92**,** 160-7.

GATTONE, V. H., 2ND, WANG, X., HARRIS, P. C. & TORRES, V. E. 2003. Inhibition of renal cystic disease development and progression by a vasopressin V2 receptor antagonist. *Nat Med,* 9**,** 1323-6.

GERMINO, G. G. 1997. Autosomal dominant polycystic kidney disease: a two-hit model. *Hosp Pract (1995),* 32**,** 81-2, 85-8, 91-2 passim.

GONZÁLEZ-PERRETT, S., KIM, K., IBARRA, C., DAMIANO, A. E., ZOTTA, E., BATELLI, M., HARRIS, P. C., REISIN, I. L., ARNAOUT, M. A. & CANTIELLO, H. F. 2001. Polycystin-2, the protein mutated in autosomal dominant polycystic kidney disease (ADPKD), is a Ca2+-permeable nonselective cation channel. *Proc Natl Acad Sci U S A,* 98**,** 1182-7.

GRADZIK, M., NIEMCZYK, M., GOŁĘBIOWSKI, M. & PĄCZEK, L. 2016. Diagnostic Imaging of Autosomal Dominant Polycystic Kidney Disease. *Pol J Radiol,* 81**,** 441-453.

GRANTHAM, J. J. 1996. The etiology, pathogenesis, and treatment of autosomal dominant polycystic kidney disease: recent advances. *Am J Kidney Dis,* 28**,** 788-803.

GUAY-WOODFORD, L. M., BISSLER, J. J., BRAUN, M. C., BOCKENHAUER, D., CADNAPAPHORNCHAI, M. A., DELL, K. M., KERECUK, L., LIEBAU, M. C., ALONSO-PECLET, M. H., SHNEIDER, B., EMRE, S., HELLER, T., KAMATH, B. M., MURRAY, K. F., MOISE, K., EICHENWALD, E. E., EVANS, J., KELLER, R. L., WILKINS-HAUG, L., BERGMANN, C., GUNAY-AYGUN, M., HOOPER, S. R., HARDY, K. K., HARTUNG, E. A., STREISAND, R., PERRONE, R. & MOXEY-MIMS, M. 2014. Consensus expert recommendations for the diagnosis and management of autosomal recessive polycystic kidney disease: report of an international conference. *J Pediatr,* 165**,** 611-7.

HAFIZI, A., KHATAMI, S. R., GALEHDARI, H., SHARIATI, G., SABERI, A. H. & HAMID, M. 2014. Exon sequencing of PKD1 gene in an Iranian patient with autosomal-dominant polycystic kidney disease. *Iran Biomed J,* 18**,** 143-50.

HANAOKA, K., QIAN, F., BOLETTA, A., BHUNIA, A. K., PIONTEK, K., TSIOKAS, L., SUKHATME, V. P., GUGGINO, W. B. & GERMINO, G. G. 2000. Co-assembly of polycystin-1 and -2 produces unique cation-permeable currents. *Nature,* 408**,** 990-4.

HARRIS, P. C. & ROSSETTI, S. 2010. Molecular diagnostics for autosomal dominant polycystic kidney disease. *Nat Rev Nephrol,* 6**,** 197-206.

HARRIS, P. C. & TORRES, V. E. 2009. Polycystic kidney disease. *Annu Rev Med,* 60**,** 321-37.

HATEBOER, N., V DIJK, M. A., BOGDANOVA, N., COTO, E., SAGGAR-MALIK, A. K., SAN MILLAN, J. L., TORRA, R., BREUNING, M. & RAVINE, D. 1999. Comparison of phenotypes of polycystic kidney disease types 1 and 2. European PKD1-PKD2 Study Group. *Lancet,* 353**,** 103-7.

HAYASHI, T., MOCHIZUKI, T., REYNOLDS, D. M., WU, G., CAI, Y. & SOMLO, S. 1997. Characterization of the exon structure of the polycystic kidney disease 2 gene (PKD2). *Genomics,* 44**,** 131-6.

HJELM, L. N., CHIN, E. L., HEGDE, M. R., COFFEE, B. W. & BEAN, L. J. 2010. A simple method to confirm and size deletion, duplication, and insertion mutations detected by sequence analysis. *J Mol Diagn,* 12**,** 607-10.

HOGAN, M. C., MASYUK, T. V., PAGE, L. J., KUBLY, V. J., BERGSTRALH, E. J., LI, X., KIM, B., KING, B. F., GLOCKNER, J., HOLMES, D. R., 3RD, ROSSETTI, S., HARRIS, P. C., LARUSSO, N. F. & TORRES, V. E. 2010. Randomized clinical trial of long-acting somatostatin for autosomal dominant polycystic kidney and liver disease. *J Am Soc Nephrol,* 21**,** 1052-61.

HUAN, Y. & VAN ADELSBERG, J. 1999. Polycystin-1, the PKD1 gene product, is in a complex containing E-cadherin and the catenins. *J Clin Invest,* 104**,** 1459-68.

HUANG, L. & LIPSCHUTZ, J. H. 2014. Cilia and polycystic kidney disease, kith and kin. *Birth Defects Res C Embryo Today,* 102**,** 174-85.

HUGHES, J., WARD, C. J., PERAL, B., ASPINWALL, R., CLARK, K., SAN MILLÁN, J. L., GAMBLE, V. & HARRIS, P. C. 1995. The polycystic kidney disease 1 (PKD1) gene encodes a novel protein with multiple cell recognition domains. *Nat Genet,* 10**,** 151-60.

IBRAGHIMOV-BESKROVNAYA, O., DACKOWSKI, W. R., FOGGENSTEINER, L., COLEMAN, N., THIRU, S., PETRY, L. R., BURN, T. C., CONNORS, T. D., VAN RAAY, T., BRADLEY, J., QIAN, F., ONUCHIC, L. F., WATNICK, T. J., PIONTEK, K., HAKIM, R. M., LANDES, G. M., GERMINO, G. G., SANDFORD, R. & KLINGER, K. W. 1997. Polycystin: in vitro synthesis, in vivo tissue expression, and subcellular localization identifies a large membrane-associated protein. *Proc Natl Acad Sci U S A,* 94**,** 6397-402.

IOANNIDIS, N. M., ROTHSTEIN, J. H., PEJAVER, V., MIDDHA, S., MCDONNELL, S. K., BAHETI, S., MUSOLF, A., LI, Q., HOLZINGER, E., KARYADI, D., CANNON-ALBRIGHT, L. A., TEERLINK, C. C., STANFORD, J. L., ISAACS, W. B., XU, J., COONEY, K. A., LANGE, E. M., SCHLEUTKER, J., CARPTEN, J. D., POWELL, I. J., CUSSENOT, O., CANCEL-TASSIN, G., GILES, G. G., MACINNIS, R. J., MAIER, C., HSIEH, C. L., WIKLUND, F., CATALONA, W. J., FOULKES, W. D., MANDAL, D., EELES, R. A., KOTE-JARAI, Z., BUSTAMANTE, C. D., SCHAID, D. J., HASTIE, T., OSTRANDER, E. A., BAILEY-WILSON, J. E., RADIVOJAC, P., THIBODEAU, S. N., WHITTEMORE, A. S. & SIEH, W. 2016. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet,* 99**,** 877-885.

JIA, H., GUO, Y., ZHAO, W. & WANG, K. 2014. Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. *Sci Rep,* 4**,** 5737.

KAEHNY, W. D. & EVERSON, G. T. 1991. Extrarenal manifestations of autosomal dominant polycystic kidney disease. *Semin Nephrol,* 11**,** 661-70.

KNIERIM, E., LUCKE, B., SCHWARZ, J. M., SCHUELKE, M. & SEELOW, D. 2011. Systematic comparison of three methods for fragmentation of long-range PCR products for next generation sequencing. *PLoS One,* 6**,** e28240.

KO, J. Y. 2016. Functional Study of the Primary Cilia in ADPKD. *Adv Exp Med Biol,* 933**,** 45-57.

KOPTIDES, M., CONSTANTINIDES, R., KYRIAKIDES, G., HADJIGAVRIEL, M., PATSALIS, P. C., PIERIDES, A. & DELTAS, C. C. 1998. Loss of heterozygosity in polycystic kidney disease with a missense mutation in the repeated region of PKD1. *Hum Genet,* 103**,** 709-17.

KOPTIDES, M. & DELTAS, C. C. 2000. Autosomal dominant polycystic kidney disease: molecular genetics and molecular pathogenesis. *Hum Genet,* 107**,** 115-26.

KOPTIDES, M., HADJIMICHAEL, C., KOUPEPIDOU, P., PIERIDES, A. & CONSTANTINOU DELTAS, C. 1999. Germinal and somatic mutations in the PKD2 gene of renal cysts in autosomal dominant polycystic kidney disease. *Hum Mol Genet,* 8**,** 509-13.

KOPTIDES, M., MEAN, R., DEMETRIOU, K., CONSTANTINIDES, R., PIERIDES, A., HARRIS, P. C. & DELTAS, C. C. 2000. Screening of the PKD1 duplicated region reveals multiple single nucleotide polymorphisms and a de novo mutation in Hellenic polycystic kidney disease families. *Hum Mutat,* 16**,** 176.

KÖTTGEN, M., BENZING, T., SIMMEN, T., TAUBER, R., BUCHHOLZ, B., FELICIANGELI, S., HUBER, T. B., SCHERMER, B., KRAMER-ZUCKER, A., HÖPKER, K., SIMMEN, K. C., TSCHUCKE, C. C., SANDFORD, R., KIM, E., THOMAS, G. & WALZ, G. 2005. Trafficking of TRPP2 by PACS proteins represents a novel mechanism of ion channel regulation. *Embo j,* 24**,** 705-16.

LANKTREE, M. B., HAGHIGHI, A., GUIARD, E., ILIUTA, I. A., SONG, X., HARRIS, P. C., PATERSON, A. D. & PEI, Y. 2018. Prevalence Estimates of Polycystic Kidney and Liver Disease by Population Sequencing. *J Am Soc Nephrol,* 29**,** 2593-2600.

LANTINGA-VAN LEEUWEN, I. S., DAUWERSE, J. G., BAELDE, H. J., LEONHARD, W. N., VAN DE WAL, A., WARD, C. J., VERBEEK, S., DERUITER, M. C., BREUNING, M. H., DE HEER, E. & PETERS, D. J. 2004. Lowering of Pkd1 expression is sufficient to cause polycystic kidney disease. *Hum Mol Genet,* 13**,** 3069-77.

LANTINGA-VAN LEEUWEN, I. S., LEONHARD, W. N., VAN DER WAL, A., BREUNING, M. H., DE HEER, E. & PETERS, D. J. 2007. Kidney-specific inactivation of the Pkd1 gene induces rapid cyst formation in developing kidneys and a slow onset of disease in adult mice. *Hum Mol Genet,* 16**,** 3188-96.

LEE, H. & SCHATZ, M. C. 2012. Genomic dark matter: the reliability of short read mapping illustrated by the genome mappability score. *Bioinformatics,* 28**,** 2097-105.

LIU, W., CHEN, M., WEI, J., HE, W., LI, Z., SUN, X. & SHI, Y. 2014. Modification of PCR conditions and design of exon-specific primers for the efficient molecular diagnosis of PKD1 mutations. *Kidney Blood Press Res,* 39**,** 536-45.

LU, W., PEISSEL, B., BABAKHANLOU, H., PAVLOVA, A., GENG, L., FAN, X., LARSON, C., BRENT, G. & ZHOU, J. 1997. Perinatal lethality with kidney and pancreas defects in mice with a targetted Pkd1 mutation. *Nat Genet,* 17**,** 179-81.

MAGENHEIMER, B. S., ST JOHN, P. L., ISOM, K. S., ABRAHAMSON, D. R., DE LISLE, R. C., WALLACE, D. P., MASER, R. L., GRANTHAM, J. J. & CALVET, J. P. 2006. Early embryonic renal tubules of wild-type and polycystic kidney disease kidneys respond to cAMP stimulation with cystic fibrosis transmembrane conductance regulator/Na(+),K(+),2Cl(-) Co-transporter-dependent cystic dilation. *J Am Soc Nephrol,* 17**,** 3424-37.

MALLAWAARACHCHI, A. C., HORT, Y., COWLEY, M. J., MCCABE, M. J., MINOCHE, A., DINGER, M. E., SHINE, J. & FURLONG, T. J. 2016. Whole-genome sequencing overcomes pseudogene homology to diagnose autosomal dominant polycystic kidney disease. *Eur J Hum Genet,* 24**,** 1584-1590.

MARKOWITZ, G. S., CAI, Y., LI, L., WU, G., WARD, L. C., SOMLO, S. & D'AGATI, V. D. 1999. Polycystin-2 expression is developmentally regulated. *Am J Physiol,* 277**,** F17-25.

MASYUK, T. V., MASYUK, A. I., TORRES, V. E., HARRIS, P. C. & LARUSSO, N. F. 2007. Octreotide inhibits hepatic cystogenesis in a rodent model of polycystic liver disease by reducing cholangiocyte adenosine 3',5'-cyclic monophosphate. *Gastroenterology,* 132**,** 1104-16.

MILLER, S. A., DYKES, D. D. & POLESKY, H. F. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res,* 16**,** 1215.

MOCHIZUKI, T., WU, G., HAYASHI, T., XENOPHONTOS, S. L., VELDHUISEN, B., SARIS, J. J., REYNOLDS, D. M., CAI, Y., GABOW, P. A., PIERIDES, A., KIMBERLING, W. J., BREUNING, M. H., DELTAS, C. C., PETERS, D. J. & SOMLO, S. 1996. PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. *Science,* 272**,** 1339-42.

MONTESANO, R., GHZILI, H., CARROZZINO, F., ROSSIER, B. C. & FÉRAILLE, E. 2009. cAMP-dependent chloride secretion mediates tubule enlargement and cyst formation by cultured mammalian collecting duct cells. *Am J Physiol Renal Physiol,* 296**,** F446-57.

MOY, K. 2007. Genomes 3. *The Yale Journal of Biology and Medicine,* 80**,** 216-216.

NAULI, S. M., ALENGHAT, F. J., LUO, Y., WILLIAMS, E., VASSILEV, P., LI, X., ELIA, A. E., LU, W., BROWN, E. M., QUINN, S. J., INGBER, D. E. & ZHOU, J. 2003. Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nat Genet,* 33**,** 129-37.

NICOLAU, C., TORRA, R., BADENAS, C., VILANA, R., BIANCHI, L., GILABERT, R., DARNELL, A. & BRÚ, C. 1999. Autosomal dominant polycystic kidney disease types 1 and 2: assessment of US sensitivity for diagnosis. *Radiology,* 213**,** 273-6.

NIMS, N., VASSMER, D. & MASER, R. L. 2003. Transmembrane domain analysis of polycystin-1, the product of the polycystic kidney disease-1 (PKD1) gene: evidence for 11 membrane-spanning domains. *Biochemistry,* 42**,** 13035-48.

NISHIO, S., TIAN, X., GALLAGHER, A. R., YU, Z., PATEL, V., IGARASHI, P. & SOMLO, S. 2010. Loss of oriented cell division does not initiate cyst formation. *J Am Soc Nephrol,* 21**,** 295-302.

PATEL, V., CHOWDHURY, R. & IGARASHI, P. 2009. Advances in the pathogenesis and treatment of polycystic kidney disease. *Curr Opin Nephrol Hypertens,* 18**,** 99-106.

PATEL, V., LI, L., COBO-STARK, P., SHAO, X., SOMLO, S., LIN, F. & IGARASHI, P. 2008. Acute kidney injury and aberrant planar cell polarity induce cyst formation in mice lacking renal cilia. *Hum Mol Genet,* 17**,** 1578-90.

PATERSON, A. D. & PEI, Y. 1999. PKD3-to be or not to be? *Nephrol Dial Transplant,* 14**,** 2965-6.

PEI, Y. 2006. Diagnostic approach in autosomal dominant polycystic kidney disease. *Clin J Am Soc Nephrol,* 1**,** 1108-14.

PEI, Y., HWANG, Y. H., CONKLIN, J., SUNDSBAK, J. L., HEYER, C. M., CHAN, W., WANG, K., HE, N., RATTANSINGH, A., ATRI, M., HARRIS, P. C. & HAIDER, M. A. 2015. Imaging-based diagnosis of autosomal dominant polycystic kidney disease. *J Am Soc Nephrol,* 26**,** 746-53.

PEI, Y., OBAJI, J., DUPUIS, A., PATERSON, A. D., MAGISTRONI, R., DICKS, E., PARFREY, P., CRAMER, B., COTO, E., TORRA, R., SAN MILLAN, J. L., GIBSON, R., BREUNING, M., PETERS, D. & RAVINE, D. 2009. Unified criteria for ultrasonographic diagnosis of ADPKD. *J Am Soc Nephrol,* 20**,** 205-12.

PEI, Y. & WATNICK, T. 2010. Diagnosis and screening of autosomal dominant polycystic kidney disease. *Adv Chronic Kidney Dis,* 17**,** 140-52.

PEI, Y., WATNICK, T., HE, N., WANG, K., LIANG, Y., PARFREY, P., GERMINO, G. & ST GEORGE-HYSLOP, P. 1999. Somatic PKD2 mutations in individual kidney and liver cysts support a "two-hit" model of cystogenesis in type 2 autosomal dominant polycystic kidney disease. *J Am Soc Nephrol,* 10**,** 1524-9.

PELUCCHI, B., AGUIARI, G., PIGNATELLI, A., MANZATI, E., WITZGALL, R., DEL SENNO, L. & BELLUZZI, O. 2006. Nonspecific cation current associated with native polycystin-2 in HEK-293 cells. *J Am Soc Nephrol,* 17**,** 388-97.

PETERS, D. J. & SANDKUIJL, L. A. 1992. Genetic heterogeneity of polycystic kidney disease in Europe. *Contrib Nephrol,* 97**,** 128-39.

PETERS, D. J., VAN DE WAL, A., SPRUIT, L., SARIS, J. J., BREUNING, M. H., BRUIJN, J. A. & DE HEER, E. 1999. Cellular localization and tissue distribution of polycystin-1. *J Pathol,* 188**,** 439-46.

PHAKDEEKITCHAROEN, B., WATNICK, T. J., AHN, C., WHANG, D. Y., BURKHART, B. & GERMINO, G. G. 2000. Thirteen novel mutations of the replicated region of PKD1 in an Asian population. *Kidney Int,* 58**,** 1400-12.

PIONTEK, K., MENEZES, L. F., GARCIA-GONZALEZ, M. A., HUSO, D. L. & GERMINO, G. G. 2007. A critical developmental switch defines the kinetics of kidney cyst formation after loss of Pkd1. *Nat Med,* 13**,** 1490-5.

PRASAD, S., MCDAID, J. P., TAM, F. W., HAYLOR, J. L. & ONG, A. C. 2009. Pkd2 dosage influences cellular repair responses following ischemia-reperfusion injury. *Am J Pathol,* 175**,** 1493-503.

PRITCHARD, L., SLOANE-STANLEY, J. A., SHARPE, J. A., ASPINWALL, R., LU, W., BUCKLE, V., STRMECKI, L., WALKER, D., WARD, C. J., ALPERS, C. E., ZHOU, J., WOOD, W. G. & HARRIS, P. C. 2000. A human PKD1 transgene generates functional polycystin-1 in mice and is associated with a cystic phenotype. *Hum Mol Genet,* 9**,** 2617-27.

QI, X. P., DU, Z. F., MA, J. M., CHEN, X. L., ZHANG, Q., FEI, J., WEI, X. M., CHEN, D., KE, H. P., LIU, X. Z., LI, F., CHEN, Z. G., SU, Z., JIN, H. Y., LIU, W. T., ZHAO, Y., JIANG, H. L., LAN, Z. Z., LI, P. F., FANG, M. Y., DONG, W. & ZHANG, X. N. 2013. Genetic diagnosis of autosomal dominant polycystic kidney disease by targeted capture and next-generation sequencing: utility and limitations. *Gene,* 516**,** 93-100.

QIAN, F., GERMINO, F. J., CAI, Y., ZHANG, X., SOMLO, S. & GERMINO, G. G. 1997. PKD1 interacts with PKD2 through a probable coiled-coil domain. *Nat Genet,* 16**,** 179-83.

QIAN, F., WATNICK, T. J., ONUCHIC, L. F. & GERMINO, G. G. 1996. The molecular basis of focal cyst formation in human autosomal dominant polycystic kidney disease type I. *Cell,* 87**,** 979-87.

RAVINE, D., WALKER, R. G., GIBSON, R. N., FORREST, S. M., RICHARDS, R. I., FRIEND, K., SHEFFIELD, L. J., KINCAID-SMITH, P. & DANKS, D. M. 1992. Phenotype and genotype heterogeneity in autosomal dominant polycystic kidney disease. *Lancet,* 340**,** 1330-3.

REEDERS, S. T. 1992. Multilocus polycystic disease. *Nat Genet,* 1**,** 235-7.

REYNOLDS, D. M., HAYASHI, T., CAI, Y., VELDHUISEN, B., WATNICK, T. J., LENS, X. M., MOCHIZUKI, T., QIAN, F., MAEDA, Y., LI, L., FOSSDAL, R., COTO, E., WU, G., BREUNING, M. H., GERMINO, G. G., PETERS, D. J. & SOMLO, S. 1999. Aberrant splicing in the PKD2 gene as a cause of polycystic kidney disease. *J Am Soc Nephrol,* 10**,** 2342-51.

RICHARDS, S., AZIZ, N., BALE, S., BICK, D., DAS, S., GASTIER-FOSTER, J., GRODY, W. W., HEGDE, M., LYON, E., SPECTOR, E., VOELKERDING, K. & REHM, H. L. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med,* 17**,** 405-24.

ROELFSEMA, J. H., SPRUIT, L., SARIS, J. J., CHANG, P., PIRSON, Y., VAN OMMEN, G. J., PETERS, D. J. & BREUNING, M. H. 1997. Mutation detection in the repeated part of the PKD1 gene. *Am J Hum Genet,* 61**,** 1044-52.

ROSSETTI, S., CONSUGAR, M. B., CHAPMAN, A. B., TORRES, V. E., GUAY-WOODFORD, L. M., GRANTHAM, J. J., BENNETT, W. M., MEYERS, C. M., WALKER, D. L., BAE, K., ZHANG, Q. J., THOMPSON, P. A., MILLER, J. P. & HARRIS, P. C. 2007. Comprehensive molecular diagnostics in autosomal dominant polycystic kidney disease. *J Am Soc Nephrol,* 18**,** 2143-60.

ROSSETTI, S., HOPP, K., SIKKINK, R. A., SUNDSBAK, J. L., LEE, Y. K., KUBLY, V., ECKLOFF, B. W., WARD, C. J., WINEARLS, C. G., TORRES, V. E. & HARRIS, P. C. 2012. Identification of gene mutations in autosomal dominant polycystic kidney disease through targeted resequencing. *J Am Soc Nephrol,* 23**,** 915-33.

ROSSETTI, S., KUBLY, V. J., CONSUGAR, M. B., HOPP, K., ROY, S., HORSLEY, S. W., CHAUVEAU, D., REES, L., BARRATT, T. M., VAN'T HOFF, W. G., NIAUDET, P., TORRES, V. E. & HARRIS, P. C. 2009. Incompletely penetrant PKD1 alleles suggest a role for gene dosage in cyst initiation in polycystic kidney disease. *Kidney Int,* 75**,** 848-55.

ROSSETTI, S., STRMECKI, L., GAMBLE, V., BURTON, S., SNEDDON, V., PERAL, B., ROY, S., BAKKALOGLU, A., KOMEL, R., WINEARLS, C. G. & HARRIS, P. C. 2001. Mutation analysis of the entire PKD1 gene: genetic and diagnostic implications. *Am J Hum Genet,* 68**,** 46-63.

SANDFORD, R., SGOTTO, B., APARICIO, S., BRENNER, S., VAUDIN, M., WILSON, R. K., CHISSOE, S., PEPIN, K., BATEMAN, A., CHOTHIA, C., HUGHES, J. & HARRIS, P. 1997. Comparative analysis of the

polycystic kidney disease 1 (PKD1) gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. *Hum Mol Genet,* 6**,** 1483-9.

SERRA, A. L., POSTER, D., KISTLER, A. D., KRAUER, F., RAINA, S., YOUNG, J., RENTSCH, K. M., SPANAUS, K. S., SENN, O., KRISTANTO, P., SCHEFFEL, H., WEISHAUPT, D. & W¼THRICH, R. P. 2010. Sirolimus and kidney growth in autosomal dominant polycystic kidney disease. *N Engl J Med,* 363**,** 820-9.

SESSA, A., RIGHETTI, M. & BATTINI, G. 2004. Autosomal recessive and dominant polycystic kidney diseases. *Minerva Urol Nefrol,* 56**,** 329-38.

SHILLINGFORD, J. M., MURCIA, N. S., LARSON, C. H., LOW, S. H., HEDGEPETH, R., BROWN, N., FLASK, C. A., NOVICK, A. C., GOLDFARB, D. A., KRAMER-ZUCKER, A., WALZ, G., PIONTEK, K. B., GERMINO, G. G. & WEIMBS, T. 2006. The mTOR pathway is regulated by polycystin-1, and its inhibition reverses renal cystogenesis in polycystic kidney disease. *Proc Natl Acad Sci U S A,* 103**,** 5466-71.

SHORT, K. M., COMBES, A. N., LEFEVRE, J., JU, A. L., GEORGAS, K. M., LAMBERTON, T., CAIRNCROSS, O., RUMBALLE, B. A., MCMAHON, A. P., HAMILTON, N. A., SMYTH, I. M. & LITTLE, M. H. 2014. Global quantification of tissue dynamics in the developing mouse kidney. *Dev Cell,* 29**,** 188-202.

STREETS, A. J., WAGNER, B. E., HARRIS, P. C., WARD, C. J. & ONG, A. C. 2009. Homophilic and heterophilic polycystin 1 interactions regulate E-cadherin recruitment and junction assembly in MDCK cells. *J Cell Sci,* 122**,** 1410-7.

TAKAKURA, A., CONTRINO, L., BECK, A. W. & ZHOU, J. 2008. Pkd1 inactivation induced in adulthood produces focal cystic disease. *J Am Soc Nephrol,* 19**,** 2351-63.

TAKAKURA, A., CONTRINO, L., ZHOU, X., BONVENTRE, J. V., SUN, Y., HUMPHREYS, B. D. & ZHOU, J. 2009. Renal injury is a third hit promoting rapid development of adult polycystic kidney disease. *Hum Mol Genet,* 18**,** 2523-31.

TAN, A. Y., MICHAEEL, A., LIU, G., ELEMENTO, O., BLUMENFELD, J., DONAHUE, S., PARKER, T., LEVINE, D. & RENNERT, H. 2014. Molecular diagnosis of autosomal dominant polycystic kidney disease using next-generation sequencing. *J Mol Diagn,* 16**,** 216-28.

TAN, Y. C., BLUMENFELD, J. D., ANGHEL, R., DONAHUE, S., BELENKAYA, R., BALINA, M., PARKER, T., LEVINE, D., LEONARD, D. G. & RENNERT, H. 2009. Novel method for genomic analysis of PKD1 and PKD2 mutations in autosomal dominant polycystic kidney disease. *Hum Mutat,* 30**,** 264-73.

TAN, Y. C., MICHAEEL, A., BLUMENFELD, J., DONAHUE, S., PARKER, T., LEVINE, D. & RENNERT, H. 2012. A novel long-range PCR sequencing method for genetic analysis of the entire PKD1 gene. *J Mol Diagn,* 14**,** 305-13.

THOMAS, R., MCCONNELL, R., WHITTACKER, J., KIRKPATRICK, P., BRADLEY, J. & SANDFORD, R. 1999. Identification of mutations in the repeated part of the autosomal dominant polycystic kidney disease type 1 gene, PKD1, by long-range PCR. *Am J Hum Genet,* 65**,** 39-49.

THONGNOPPAKHUN, W., WILAIRAT, P., VAREESANGTHIP, K. & YENCHITSOMANUS, P. T. 1999. Long RT-PCR Amplification of the entire coding sequence of the polycystic kidney disease 1 (PKD1) gene. *Biotechniques,* 26**,** 126-32.

TORRA BALCELLS, R. & ARS CRIACH, E. 2011. Molecular diagnosis of autosomal dominant polycystic kidney disease. *Nefrologia,* 31**,** 35-43.

TORRA, R., BADENAS, C., SAN MILLÁN, J. L., PÉREZ-OLLER, L., ESTIVILL, X. & DARNELL, A. 1999. A loss-of-function model for cystogenesis in human autosomal dominant polycystic kidney disease type 2. *Am J Hum Genet,* 65**,** 345-52.

TORRES, V. E. 2008. Role of vasopressin antagonists. *Clin J Am Soc Nephrol,* 3**,** 1212-8.

TORRES, V. E., BOLETTA, A., CHAPMAN, A., GATTONE, V., PEI, Y., QIAN, Q., WALLACE, D. P., WEIMBS, T. & W¼THRICH, R. P. 2010. Prospects for mTOR inhibitor use in patients with polycystic kidney disease and hamartomatous diseases. *Clin J Am Soc Nephrol,* 5**,** 1312-29.

TORRES, V. E., WANG, X., QIAN, Q., SOMLO, S., HARRIS, P. C. & GATTONE, V. H., 2ND 2004. Effective treatment of an orthologous model of autosomal dominant polycystic kidney disease. *Nat Med,* 10**,** 363-4.

TRUJILLANO, D., BULLICH, G., OSSOWSKI, S., BALLARÍN, J., TORRA, R., ESTIVILL, X. & ARS, E. 2014. Diagnosis of autosomal dominant polycystic kidney disease using efficient PKD1 and PKD2 targeted next-generation sequencing. *Mol Genet Genomic Med,* 2**,** 412-21.

TSIOKAS, L., KIM, E., ARNOULD, T., SUKHATME, V. P. & WALZ, G. 1997. Homo- and heterodimeric interactions between the gene products of PKD1 and PKD2. *Proc Natl Acad Sci U S A,* 94**,** 6965-70.

TSIOKAS, L., KIM, S. & ONG, E. C. 2007. Cell biology of polycystin-2. *Cell Signal,* 19**,** 444-53.

VASSILEV, P. M., GUO, L., CHEN, X. Z., SEGAL, Y., PENG, J. B., BASORA, N., BABAKHANLOU, H., CRUGER, G., KANAZIRSKA, M., YE, C., BROWN, E. M., HEDIGER, M. A. & ZHOU, J. 2001. Polycystin-2 is a novel cation channel implicated in defective intracellular Ca(2+) homeostasis in polycystic kidney disease. *Biochem Biophys Res Commun,* 282**,** 341-50.

WAHL, P. R., SERRA, A. L., LE HIR, M., MOLLE, K. D., HALL, M. N. & W¼THRICH, R. P. 2006. Inhibition of mTOR with sirolimus slows disease progression in Han:SPRD rats with autosomal dominant polycystic kidney disease (ADPKD). *Nephrol Dial Transplant,* 21**,** 598-604.

WALZ, G., BUDDE, K., MANNAA, M., N¼RNBERGER, J., WANNER, C., SOMMERER, C., KUNZENDORF, U., BANAS, B., HÖRL, W. H., OBERM¼LLER, N., ARNS, W., PAVENSTÄDT, H., GAEDEKE, J., B¼CHERT, M., MAY, C., GSCHAIDMEIER, H., KRAMER, S. & ECKARDT, K. U. 2010. Everolimus in patients with autosomal dominant polycystic kidney disease. *N Engl J Med,* 363**,** 830-40.

WANG, G., SEIDMAN, M. M. & GLAZER, P. M. 1996. Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science,* 271**,** 802-5.

WANG, X., GATTONE, V., 2ND, HARRIS, P. C. & TORRES, V. E. 2005. Effectiveness of vasopressin V2 receptor antagonists OPC-31260 and OPC-41061 on polycystic kidney disease development in the PCK rat. *J Am Soc Nephrol,* 16**,** 846-51.

WARD, C. J., TURLEY, H., ONG, A. C., COMLEY, M., BIDDOLPH, S., CHETTY, R., RATCLIFFE, P. J., GATTNER, K. & HARRIS, P. C. 1996. Polycystin, the polycystic kidney disease 1 protein, is expressed by epithelial cells in fetal, adult, and polycystic kidney. *Proc Natl Acad Sci U S A,* 93**,** 1524-8.

WATNICK, T. J., PIONTEK, K. B., CORDAL, T. M., WEBER, H., GANDOLPH, M. A., QIAN, F., LENS, X. M., NEUMANN, H. P. & GERMINO, G. G. 1997. An unusual pattern of mutation in the duplicated portion of PKD1 is revealed by use of a novel strategy for mutation detection. *Hum Mol Genet,* 6**,** 1473-81.

WILLEY, C. J., BLAIS, J. D., HALL, A. K., KRASA, H. B., MAKIN, A. J. & CZERWIEC, F. S. 2017. Prevalence of autosomal dominant polycystic kidney disease in the European Union. *Nephrol Dial Transplant,* 32**,** 1356-1363.

WU, G., D'AGATI, V., CAI, Y., MARKOWITZ, G., PARK, J. H., REYNOLDS, D. M., MAEDA, Y., LE, T. C., HOU, H., JR., KUCHERLAPATI, R., EDELMANN, W. & SOMLO, S. 1998. Somatic inactivation of Pkd2 results in polycystic kidney disease. *Cell,* 93**,** 177-88.

YANG, B., SONAWANE, N. D., ZHAO, D., SOMLO, S. & VERKMAN, A. S. 2008. Small-molecule CFTR inhibitors slow cyst growth in polycystic kidney disease. *J Am Soc Nephrol,* 19**,** 1300-10.

YODER, B. K., HOU, X. & GUAY-WOODFORD, L. M. 2002. The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. *J Am Soc Nephrol,* 13**,** 2508-16.

YU, Y., ULBRICH, M. H., LI, M. H., BURAEI, Z., CHEN, X. Z., ONG, A. C., TONG, L., ISACOFF, E. Y. & YANG, J. 2009. Structural and molecular basis of the assembly of the TRPP2/PKD1 complex. *Proc Natl Acad Sci U S A,* 106**,** 11558-63.

ZAFAR, I., BELIBI, F. A., HE, Z. & EDELSTEIN, C. L. 2009. Long-term rapamycin therapy in the Han:SPRD rat model of polycystic kidney disease (PKD). *Nephrol Dial Transplant,* 24**,** 2349-53.

ZHOU, J. 2009. Polycystins and primary cilia: primers for cell cycle progression. *Annu Rev Physiol,* 71**,** 83-113.

PKD International. 2021. *ADPKD*. [online] Available at: <https://pkdinternational.org/what-is-pkd/adpkd#:~:text=How%20is%20ADPKD%20inherited%3F,the%20PKD1%20and%20PKD2%20genes.> [Accessed 1 March 2021].

Gordon Betts, G., Kelly Young, K., James A. Wise, J., Johnson, E., Poe, B., Kruse, D., Korol, O., Johnson, J., Womble, M. and DeSaix, P., 2013. Anatomy and Physiology. [ebook] Houston, Texas: OpenStax.