Royal College of Surgeons in Ireland

repository@rcsi.com

The genetic landscape of polycystic kidney disease in Ireland.

AUTHOR(S)

Katherine Benson, Susan Murray, Sarah R.Senum, Elhussein Elhassan, Eoin T Conlon, Claire Kennedy, Shane Conlon, Edmund Gilbert, Dervla Connaughton, Paul O'Hara, Sarah Khamis, Sarah Cormican, Lawrence C. Brody, Anne M. Mollloy, Sally Ann Lynch, Liam Casserly, Matthew D. Griffin, Robert Carton, Kevin Yachnin, Peter C. Harris, Gianpiero Cavalleri, Peter Conlon

CITATION

Benson, Katherine; Murray, Susan; R.Senum, Sarah; Elhassan, Elhussein; Conlon, Eoin T; Kennedy, Claire; et al. (2021): The genetic landscape of polycystic kidney disease in Ireland.. Royal College of Surgeons in Ireland. Journal contribution. https://hdl.handle.net/10779/rcsi.13614371.v1

HANDLE

10779/rcsi.13614371.v1

LICENCE

CC BY-NC-SA 4.0

This work is made available under the above open licence by RCSI and has been printed from https://repository.rcsi.com. For more information please contact repository@rcsi.com

URL

https://repository.rcsi.com/articles/journal_contribution/The_genetic_landscape_of_polycystic_kidney_disease_in_Ireland_/13614371/1

Title: The genetic landscape of polycystic kidney disease in Ireland

Running Title: Polycystic kidney disease in Ireland

Authors: Katherine A. Benson[1], Susan Murray[2], Sarah R. Senum[3], Elhussein Elhassan[2], Eoin T Conlon[2], Claire Kennedy[2], Shane Conlon[2] , Edmund Gilbert[1], Dervla Connaughton[2], Paul O'Hara[5], Sarah Khamis[1], Sarah Cormican[6, 7], Lawrence C. Brody[8], Anne M . Molloy[9], Sally Ann Lynch[10], Liam Casserly[5], Matthew D. Griffin[6, 7], Robert Carton[1], Kevin Yachnin[4], Peter C. Harris[3], Gianpiero L. Cavalleri[1*], Peter Conlon[1,2*]

Affiliations:
1: School of Pharmacy and Biomolecular Science, Royal College of Surgeons in Ireland, Dublin, Ireland

2: Department of Nephrology, Beaumont Hospital, Dublin, Ireland

3: Division of Nephrology and Hypertension, Mayo Clinic, Rochester, Minnesota, USA

4: Region Norrbotten, Kiruna Hospital, Kiruna, Sweden

5: Department of Renal Medicine, University Hospital Limerick

6: Regenerative Medicine Institute (REMEDI) at CÚRAM Centre for Research in Medical Devices, School of Medicine, National University of Ireland Galway, Ireland

7: Nephrology Department, Galway University Hospitals, Saolta University Healthcare Group, Galway Ireland

8: Division of Intramural Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA.

9: School of Medicine, Trinity College Dublin, Dublin, Ireland

10. Department of Clinical Genetics, Children's University Hospital, Temple Street, Dublin, Ireland

*these authors contributed equally to this work

To whom correspondence should be addressed: Gianpiero Cavalleri (gcavalleri@rcsi.ie) and Peter Conlon (peterconlon@beaumont.ie)

Acknowledgements and Funding: K Benson is supported by an Enterprise Partnership Scheme Fellowship Award (2019) from The Irish Research Council, in conjunction with Punchestown Kidney Research Fund (EPSPD/2019/213). The authors also acknowledge funding received from the Beaumont Hospital Foundation, the Royal Irish Academy and the Royal College of Surgeons in Ireland. S. Cormican is

supported by the Irish Clinical Academic Training (ICAT) Programme, supported by the Wellcome Trust and the Health Research Board (Grant Number 203930/B/16/Z), the Health Service Executive National Doctors Training and Planning and the Health and Social Care, Research and Development Division, Northern Ireland. We also acknowledge that this work would not be possible without participation of the consenting patients and their families.

ABSTRACT

Polycystic Kidney Diseases (PKDs) comprise the most common Mendelian forms of renal disease. It is characterised by the development of fluid filled renal cysts, causing progressive loss of kidney function, culminating in the need for renal replacement therapy or kidney transplant. Ireland represents a valuable region for the genetic study of PKD, as family sizes are traditionally large and the population relatively homogenous. Studying a cohort of 169 patients, we describe the genetic landscape of PKD in Ireland for the first time, compare the clinical features of patients with and without a molecular diagnosis and correlate disease severity with autosomal dominant pathogenic variant type. Using a combination of molecular genetic tools, including targeted next generation sequencing, we report diagnostic rates of 71-83% in Irish PKD patients, depending on which variant classification guidelines are used (ACMG or Mayo clinic respectively). We have catalogued a spectrum of Irish autosomal dominant PKD pathogenic variants including 36 novel variants. We illustrate how apparently unrelated individuals carrying the same autosomal dominant pathogenic variant are highly likely to have inherited that variant from a common ancestor. We highlight issues surrounding the implementation of the ACMG guidelines for variant pathogenicity interpretation in PKD, which have important implications for clinical genetics.

Keywords: Polycystic Kidney Disease, Next generation sequencing, Genetic Diagnostics, PKD

INTRODUCTION

Polycystic Kidney Diseases (PKDs) are the most common forms of inherited renal disease affecting an estimated 12.5 million individuals worldwide (1). PKD is a leading cause of End Stage Kidney Disease (ESKD) and the prevalence of PKD related ESKD in Europe ranges from 30 per million population (pmp) in Switzerland to greater than 180 pmp in regions of Spain (2). In Northern Ireland the prevalence of ESKD due to PKD is reported as 115 pmp (2).

PKD can be autosomal dominant, autosomal recessive (ADPKD and ARPKD) or syndromic and can manifest prior to birth on prenatal ultrasound or as late as the eighth decade of life. Most ADPKD patients (85%) reach ESKD by age 65 years (3). Approximately 90% of ADPKD cases have an identifiable pathogenic genetic variant in *PKD1* or *PKD2*; leaving ~10% with an unknown genetic cause (4). *PKD1* and *PKD2* account for around 80-85% and 15-20% of cases with a molecular diagnosis, respectively (5–10). Copy number changes in the form of insertions and deletions over 1 kb in size (11) within *PKD1* and *PKD2* are responsible for 1-5% of cases (5,8,12,13). Additional genes account for rarer forms of ADPKD (<1% of cases with a molecular diagnosis), including *GANAB*, *ALG9* and *DNAJB11* which often have atypical ADPKD presentations (4,10,14). Recessive forms of PKD are more rare and are known to be caused by autosomal recessive (AR) pathogenic variants in genes including *PKHD1* (15).

Previous work has demonstrated that the disease gene and autosomal dominant (AD) pathogenic variant type have a significant effect on the age at which patients reach ESKD (16,17). Patients with *PKD1* AD pathogenic variants tend to have an earlier age at ESKD relative to those with AD pathogenic variants in *PKD2*. Further, *PKD1* truncating AD pathogenic variants are associated with reduced renal survival when compared with *PKD1* non-truncating variants. *PKD2* truncating variants have been associated with lower age-adjusted eGFR (18).

PKD1 has a high GC-content and lies within a segmentally duplicated part of the genome with six pseudogenes, making it challenging to sequence. The current gold-standard method for *PKD1* variant screening is long-range PCR (LR-PCR) combined with nested PCR, followed by Sanger sequencing, a process which is resource intensive (19). Hybridisation-based exome library preparation, combined with high-coverage NGS have recently been proposed as an alternative for PKD diagnostic testing, including for *PKD1* (19–22). The American College of Medical Genetics (ACMG) have published guidelines to aid in pathogenicity classification of variants (23), but uptake by the PKD research community has been relatively poor, where alternative, research guidelines such as those developed by the Mayo Clinic are more commonly used (24).

The spectrum of variants causing PKD in the Irish population has yet to be described. As family sizes are traditionally large in Ireland and the population relatively homogenous (25,26), Ireland represents a valuable region for the genetic study of PKD. Here we set out to 1) describe the pathogenic PKD variants in the Irish population, using a combination of molecular genetic approaches including NGS, 2) test the association between renal survival and PKD AD pathogenic variant type and 3) compare the ACMG guidelines to standard PKD pathogenicity criteria in order to establish the feasibility of these guidelines in PKD clinical genetics.

MATERIALS AND METHODS

Ethics: Written, informed consent was obtained from all participants of this study. Ethical approval was provided from the Ethics committee of Beaumont Hospital, Dublin, Ireland (REC 12/75).

Recruitment Centres & inclusion/exclusion criteria: One hundred and sixty nine patients (from 148 families) were recruited from four renal clinics and dialysis units in Ireland (Beaumont Hospital Dublin, The Mid-Western Regional Hospital Limerick, Galway University Hospitals and Tallaght University Hospital). Patients were included if they were:

- Aged >18 years
- Able to provide informed consent
- Did not have a previous genetic diagnosis of PKD (i.e. causal AR/AD pathogenic variant was unknown)

  AND
- Had a family history of PKD defined by the presence of a second degree relative or closer, with the disease, at the time of recruitment to the study
- Met the unified criteria for ultrasonography diagnosis of ADPKD (27)

  OR
- Had no family history of PKD but evidence of >5 cysts on ultrasound examination and no evidence of an alternate cystic kidney disease

NGS data generation: NGS was performed on 155 samples using a custom targeted panel (Roche SeqCap EZ Choice) of 227 known renal disease genes (99.9% of target bases covered, gene list previously published (28)) on an Illumina Platform (MiSeq or NextSeq). Fourteen samples were sequenced using an alternative panel covering 11 genes (Roche HeatSeq gene list provided in **Supplementary Table 1)**. Bioinformatic analyses were conducted using a GATK Best Practices based pipeline (29). All variants

discussed in this manuscript have been submitted to ClinVar (see **Supplementary Table 3**) and requests for access to raw sequence data can be made via communication with the corresponding author.

Variant pathogenicity: Candidate AD and AR pathogenic variants were identified as those that were: 1) in established renal disease genes, 2) within exonic and splicing regions 3) predicted damaging by either PolyPhen (30) or SIFT (31) and 4) with a minor allele frequency (MAF) (as per gnomAD (32)) of <1% for homozygotes or compound heterozygotes and <0.1% for heterozygotes. Synonymous variants were excluded. In order to assess if a variant had been previously reported as pathogenic, we accessed ClinVar (clinvar_20140902) (www.ncbi.nlm.nih.gov/clinvar), OMIM (www.omim.org) and the Mayo Clinic PKD database (https://pkdb.mayo.edu/index.html). Qualifying variants were brought forward for discussion at a multidisciplinary team (MDT) meeting (see below).

Sanger sequencing: Patients without candidate AD or AR pathogenic variants following NGS (see above) were further sequenced for *PKD1* and *PKD2*, using long-range PCR and/or nested PCR, followed by Sanger sequencing as described previously (8,24) (Figure 1).

Copy number variants: Copy-number variants (CNVs) were detected using read-depth analysis from NGS data using both Nexus BioDiscovery (https://www.biodiscovery.com/products/Nexus-Copy-Number) and the BROAD's GATK4 GermlineCNVCaller (29). CNVs in key *PKD* genes *(PKD1, PKD2, PKHD1, DZIP1L, GANAB)* detected using either of these read depth detection methods were confirmed using multiplex ligation-dependent probe amplification (MLPA) as described previously (13).

Multidisciplinary Team Meeting (MDT): With input from nephrologists and geneticists, the clinical picture was reviewed at the MDT with reference to the candidate pathogenic variant, to help establish pathogenicity. Candidate pathogenic variants were assessed using both the PKD variant pathogenicity guidelines from the Mayo Clinic (7,8) and the ACMG guidelines (23). Variants classified as 'Definitely Pathogenic', 'Highly Likely Pathogenic' and 'Likely Pathogenic' using guidelines from the Mayo Clinic

were considered disease causing. Guidelines from the Mayo Clinic are described in detail in the Supplementary Information. Further details of how we interpreted the ACMG guidelines are provided in **Supplementary Table 2**. Where a candidate AD or AR pathogenic variant did not clearly match the clinical phenotype of the case, the variant was labelled as a variant of unknown significance (VUS).

Frequency in Irish Traveller population: The frequency of PKD variants in the Irish Traveller population was determined using sequence data from 16 individuals. This data was processed using the SNV pipeline used for the PKD sequences described above, limited to the key PKD genes (*PKD1, PKD2, GANAB, PKHD1, DZIP1L* and *DNAJB11).*

Survival analysis: Renal survival analyses was conducted in order to calculate the correlation between AD pathogenic variant (gene and variant type) and renal survival (time to ESKD measured in years). Variants were divided into two groups: truncating (stop gain, frameshift indels, CNVs, splicing) or non-truncating (missense, non-frameshift indels). Kaplan-Meier methods were used for these analyses as indicated on SPSS v25 and v26 (SPSS, Inc., NY). Differences between the tested groups were assessed using log-rank test (0.05 significance level).

Haplotype analysis: Patients who shared the same diagnostic PKD variants with no established familial connection (i.e. pedigree analysis did not indicate a common ancestor) were genotyped using the Illumina Infinium OmniExpress microarray kit. These patients were co-analysed with microarray genotypes from previously obtained control individuals of Irish ancestry from the Irish DNA Atlas cohort (33) (n=198) and a Trinity College student cohort (34) (n=2,232). Genotypes were merged using PLINK v1.90b6.9 (35,36) and the following quality control filtering was applied; individuals were removed if genotype missingness >5%, and variants were removed if genotyping missingness >5%, or minor allele frequency <2%, or deviation from the Hardy-Weinberg equilibrium with a significance $<1x10^{-6}$. Irish ancestry of cases was confirmed using PCA analysis in PLINK. Chromosomes were phased using Beagle

v5.0 (37). Identity-by-Descent (IBD) segments with a length >=1cM and log-odds score >=3 were detected in the merged, phased dataset using the refinedIBD model (38). IBD segments were filtered for those shared between carriers of the same pathogenic PKD variant and overlapping the human genome build 37 coordinates of *PKD1*/*PKD2* (chr16:2138711-2185899, chr4:88928799-88998931, respectively). To describe the potential relationships between individuals with the same diagnostic PKD variants, using the genotype data described above, we quantified genetic relationships by estimating a combination of identity by state (IBS) and IBD using KING v2.2 (39). The --related command in KING was used to quantifying relationships closer than 5^{th} degree.

RESULTS

Recruitment: A total of 169 PKD patients from 148 families satisfied our inclusion criteria and were processed for genetic analysis (**Table 1**). The mean age of diagnosis for the 169 patients was 36 years and median age of diagnosis was 35 years (range 4-79 years), 84 (50%) were male and 97 (57%) had reached ESKD at the time of recruitment. Seventy three (43%) patients had a functioning renal transplant, 131 (78%) had a family history of PKD.

Table 1

Variants identified: The overall molecular diagnostic rate using the Mayo Clinic PKD pathogenicity guidelines was 83% (141/169), and 81% (120/148) in families. The overall molecular diagnostic rate using ACMG pathogenicity guidelines was 71% (120/169) and 68% (100/148) in families. All diagnostic, AD and AR pathogenic variants are listed in **Supplementary Table 3**. Ninety one percent of these AD and AR pathogenic variants were detected using the NGS pipeline, 7% by long range PCR and 2% by read depth analysis and MLPA (**Figure 1**). We identified 113 patients (97 families) with an AD pathogenic variant in *PKD1*, 27 patients (22 families) in *PKD2* and one patient/family with a homozygous AR pathogenic genotype in *PKHD1*. We did not detect any AD pathogenic variants in *GANAB*. The breakdown of diagnostic variants identified is summarised in **Figure 2**. Forty-four percent (44/100) of the detected AD Mayo Clinic classified definitely pathogenic/ highly likely pathogenic/ likely pathogenic variants were novel (46% (39/84) of ACMG pathogenic variants)(see **Figure 3**), and previously unreported in the relevant disease databases (see Methods). Fourteen AD pathogenic variants in *PKD1* and *PKD2* were identified in multiple families who were unrelated according to clinical records (**Table 2**). Two of these repeated AD pathogenic variants were novel.

Table 2

Patient Characteristics: We next compared clinical characteristics between groups of patients with and without an identifiable molecular diagnosis (using Mayo Clinic pathogenicity guidelines) (see **Table 1 and Supplementary Materials**). The mean age at diagnosis was lower in the group with a molecular diagnosis (p = 0.003), however the age at which these patients reach ESKD was not significantly different (p = 0.66). Although the rate of diagnosis was higher in families with a family history (85%, 112/131), compared to those without (75%, 27/36), this difference was not significant (p value 0.139, (Fisher's Exact Test, 2-sided)). There was no difference in survival rates between patients with and without a molecular diagnosis. Patient 10262 with a molecular diagnosis of *PKHD1* was diagnosed at a young age and underwent right orchidopexy and right hydrocele repair at age six . They continued to progress and at age 25 years started on dialysis.

Irish Traveller Dataset: We next looked at the frequency of the identified diagnostic variants in an endogamous population isolate within Ireland, the Irish Travellers (previously sequenced cohort, n=16 (25)). None of the disease-causing PKD variants from our cohort were detected in this cohort (including the heterozygous *PKHD1* variants). The previously described hypomorphic *PKD1* variant NM_001009944.3: c.8293C>T:p.(Arg2765Cys) (rs144979397) was observed in two of the Irish Traveller samples sequenced (2/16, 12.5%). This variant has previously been associated with a more severe phenotype when detected along with an established pathogenic *PKD1* variant, although this variant in isolation is not expected to cause disease. The disease status of these individuals is unknown (40).

Quantifying the relationship between individuals carrying the same disease-causing variant: Fourteen AD pathogenic variants in *PKD1* and *PKD2* were identified in multiple unrelated (according to clinical records) Irish PKD patients. This suggests that either these variants may be recurrent disease-causing variant or that the patients are in fact distantly related and the variants are descended from a common

ancestor. To establish if a shared variant indeed descended from a common ancestor, we determined whether that variant was present on a *PKD1/PKD2* haplotype that was identical by descent. **Table 2** describes the 14 variants shared between 34 families reported as unrelated in the clinical records. We confirmed that nine of the 14 variants showed IBD between all sharing families and one of the fourteen variants showed partial IBD sharing between two out of four families. For the remaining four variants, we were unable to detect IBD across the gene of interest between families. We then went on to quantify the genetic relationship between carriers of the same AD pathogenic variants regardless of pedigree information from the clinical record. Using IBS-based measurements (see **Supplementary Tables 4-6**), we identified relationships from parent-offspring to 4th degree (see **Table 2**). The identification of relationships between these individuals, which were not recorded in the clinical record, reduced the number of unique families from 148 to 139.

Renal Survival Analyses: In order to determine the impact of the PKD gene and AD pathogenic variant type (classified using Mayo Clinic pathogenicity guidelines) on age of development of ESKD, we conducted renal survival analyses, calculated as time to ESKD measured in years. As expected, *PKD1* AD pathogenic variants correlated with a shorter renal survival time from Kaplan-Meier analysis, relative to *PKD2*. *PKD1* patients had an estimated survival time (age at ESKD) of 50 years (standard error (SE) 1.1 years) and *PKD2* patients,70 years (SE 2.4 years). When subdivided by variant type (truncating vs non-truncating), *PKD1* truncating variants had a lower renal survival (49 years, SE 1.1 years) than those with a *PKD1* non-truncating variant (55 years, SE 2.3 years) by log rank test (*p*-value 0.002) (see **Figure 4**). No significant difference was noted between non-truncating and truncating AD pathogenic variants in *PKD2*. According to the log-rank test, patients without a molecular diagnosis had a significantly lower age at ESKD than those with an AD pathogenic variant in *PKD2* (*p*-value 0.016) and *PKD1* truncating variants (*p*-value 0.0004) but were not significantly different from those with *PKD1 non*- truncating variants (*p*-value 0.542) (**Supplementary Figure 1A**). We stratified variant type into additional subcategories ((1)

PKD1 frameshift insertion or deletion, 2) *PKD1* splicing, 3) *PKD1* stopgain or stoploss, 4) *PKD2* truncating and 5) *PKD2* non-truncating) to test the effect of these sub-groups on renal survival, but did not find any additional novel significant findings (**Supplementary Figure 1B**).

Variant Classification: We assessed candidate AD and AR pathogenic variants discussed at MDT using variant assessment guidelines from both the ACMG and the Mayo Clinic. The rate of molecular diagnosis per family was significantly lower when applying the ACMG guidelines (68% (100/148)), than using the Mayo Clinic PKD variant pathogenicity guidelines (81% (120/148)) (Fisher's Exact test *p*-value 0.0112). A total of 14% (20/145) of the families with highly-likely pathogenic or likely AR/AD pathogenic variants when categorised using the standard Mayo Clinic guidelines were classified as variants of unknown significance (VUS) using ACMG criteria (see Supplementary Table 11). Rossetti and colleagues (21) used NGS targeted sequencing to characterise a cohort of 230 ADPKD patients and applied the Mayo Clinic pathogenicity guidelines for variant classification, reporting definitely and likely AR/AD pathogenic variants in 63% (115/183) of patients. We applied ACMG guidelines to the Rossetti variant catalogue and determined a diagnostic yield of 48% (87/183). The ACMG and Mayo Clinic classifications assigned to each variant from the Irish dataset are provided in **Supplementary Table 3** and those for the Rossetti article are shown in **Supplementary Table 7**. We note that the majority of variant pathogenicity discrepancies between the two assessment tools (94% (47/50) of potentially clinically relevant discrepancies) are missense variants within *PKD1*. No variants found to be likely pathogenic or pathogenic by the ACMG guidelines were classified as benign or uncertain significance using the Mayo Clinic guidelines. Unless missense variants are well-described and/or studied previously (e.g. functional data), they are difficult to classify as anything above a VUS using ACMG criteria.

DISCUSSION

We describe the genetic landscape of PKD in Ireland for the first time, compare the clinical features of patients with and without a molecular diagnosis and correlate disease severity with variant type. We have implemented a combination of molecular diagnostic tools, achieving a diagnostic rate of 83%. We have catalogued the most common Irish PKD AR and AD pathogenic variants and identified 36 novel ACMG pathogenic likely disease causing variants. We have confirmed that many AD pathogenic variants shared between seemingly unrelated individuals have been inherited from a common ancestor. We highlight some issues surrounding the implementation of the ACMG guidelines for variant pathogenicity interpretation in PKD genomics.

Previous studies have assessed targeted NGS for PKD diagnostics (20) and although LR-PCR/Sanger approaches are still the gold-standard for variant identification, NGS is becoming established as a high-throughput method for initial assessment of patients, in-tandem with other techniques (eg. CNV identification and gap-filling Sanger). By combining these techniques with our NGS panel, we identified a diagnostic variant in 14 additional patients (8%; three by read depth analyses and 11 by LR-PCR and Sanger sequencing) who screened-negative with the standard NGS analyses. Three of these variants were CNVs, three were missense, six were splicing, one was a stopgain, and one was a non-frameshift deletion. All of these variants were within *PKD1* and did not cluster in any one exon or region. Variants detected by LR-PCR and Sanger sequencing which were not identified as part of the NGS pipeline were missed due to low coverage in exon 1 of *PKD1,* poor read mapping as a result of sequence homology to pseudogenes or were excluded as part of bioinformatic hard filtering protocols. As *DNAJB11* was not included on the panel, we were unable to assess variants in this gene.

CNVs are known to cause disease in 1-5% of PKD patients (5,8,12,13) and are thus an important part of any diagnostic pipeline. A previous study by Eisenberger and colleagues utilised the VarScan read-depth

based tool to successfully confirm two known *PKD1* CNVs in two patients, previously identified by MLPA. They did not identify any additional CNVs in the rest of their cohort (n=55) using this method (12). We utilised the recently developed GATK4 GermlineCNVCaller (29) to detect CNVs in PKD genes including *PKD1* and found three CNVs which were later confirmed by MLPA. We propose that initial screening using NGS-based CNV calling can be used to minimise the number of samples for more time-consuming and expensive MLPA testing.

Our haplotype and genetic distance analysis illustrated that apparently unrelated individuals carrying the same pathogenic PKD variant are highly likely to have inherited the variants from a distant common ancestor in regionally collected populations. We identified nine sets of distantly related families, confirmed by both IBS and IBD. Three pairs of families shared haplotypes across the region covering the AD pathogenic variant, but the IBS analysis was unable to identify a relationship between these individuals. This may be because these individuals are more distantly related than the IBS software (KING) is able to detect (i.e. greater than fourth degree relatives). We were unable to detect IBD across the gene of interest for four of the fourteen variants. We speculate that these may be older variants, carried on haplotypes that are too short to confirm IBD. Indeed, one of the four variants in question (NM_001009944.3: c.11249G>A:p.(Arg3750Gln)) has an allele frequency in gnomAD of $4x10^{-6}$ and has been reported in four families in the PKDB mutation database (https://pkdb.mayo.edu/index.html), suggesting that it is a relatively common and thus probably older PKD variant. Alternatively, these variants may represent independent mutation events. This type of genetic distance analysis may be of use when assigning pathogenicity to a variant using ACMG guidelines. For example, the variant carried by families F391 and F399 could be reclassified using ACMG guidelines from a VUS to likely pathogenic as a result of the additional segregation evidence provided by this relationship analysis.

We have replicated previous reports (16) which link *PKD1/2* AD pathogenic variants with disease severity and the association between AD pathogenic variant type (*PKD1* 'truncating', *PKD1* 'non-truncating' or *PKD2*) and age at ESKD (17). This association did not extend to more finely categorised variant types, perhaps as a result of low numbers in these smaller categories. We did not observe the previously reported (41) difference in the age at which patients with *PKD2* truncating or *PKD2* non-truncating AD pathogenic variants reach ESKD, probably due to lack of study power. These results are consistent with previous reports (17,18). However, previous reports have indicated a difference in levels of eGFR between these groups (18). Our screen-negative patients appeared to have similar rates of renal survival to those with AD *PKD1* non-truncating pathogenic variants and had significantly lower renal survival to those with *PKD2* AD pathogenic variants.

We observed three individuals who carried additional variants which satisfied the ACMG guidelines for variant pathogenicity ('likely pathogenic'/'pathogenic') (**Supplementary Table 8**). It is unclear based on the patient phenotypes available whether these variants may have an impact on the severity of the disease in these cases. Two of these patients carried the same splicing variant in *COL4A4* (c.2164+2T>G). Collagen gene variants including those in *COL4A4* are traditionally associated with Alport syndrome/ thin basement membrane disease but have been reported previously in-tandem with a PKD AD pathogenic variant and may have a modifier effect on primary PKD variants (42). The main feature of *COL4A4*-linked disease is haematuria and progressive renal failure associated with progressive hearing loss. There was no evidence of hearing loss in either pedigree.

Whilst the ACMG guidelines have been widely embraced by the scientific community, it is recognised that there are deficiencies. These guidelines perform best for paediatric-onset rare diseases. For adult onset disorders such as PKD and those with reduced penetrance, there are a number of weaknesses:

1) Missense variants are common, and score as only moderate evidence for pathogenicity while truncating variants score as strong evidence. This increases the likelihood that missense variants will be scored as a VUS.

2) Population databases such as gnomAD are used to estimate the frequency of variants in the healthy population. This is a reasonable approach for paediatric rare diseases where penetrance is almost 100% by adulthood. However, it is possible for PKD that a small number of affected patients may be included in the 'healthy' population before presenting with symptoms. This may lead an AD pathogenic variant to be deemed as benign/VUS.

3) The incidence of *de novo* variants is lower in PKD than in severe paediatric-onset diseases. If a variant is *de novo* this can be used as strong evidence of pathogenicity within ACMG. As most PKD is inherited, it means that many more variants remain as a VUS.

4) In diseases like PKD in which there is extreme allelic heterogeneity, it is less likely that a variant will have been previously reported in a disease database, reducing the opportunity to assign strong evidence on the basis of previously established pathogenicity.

We assessed the utility of the ACMG guidelines for clinical variant pathogenicity interpretation in PKD and compared these guidelines to those widely used in PKD research from the Mayo Clinic. An important difference between these tools are that the ACMG guidelines were developed for clinical genetics use, while the guidelines from the Mayo Clinic were designed for research use only. Our results highlight an issue surrounding the use of the ACMG guidelines, with regard to missense *PKD1* variant assessment. These missense variants are difficult to classify as likely pathogenic or pathogenic using the ACMG guidelines unless they have been extensively studied and reported previously (43), largely due to the variability in regional pathogenicity of *PKD1* (10). While it is imperative that pathogenicity guidelines are conservative to control for 'over-calling' of Pathogenic/Likely Pathogenic variants, the conservative

nature of the ACMG guidelines can be problematic as novel *PKD1* variants are common, with 31% of diagnostic variants identified being a novel finding from our cohort. Equally, it is important to note that higher diagnostic rates obtained when using the Mayo Clinic research classification tool may be enriched for false-positives, which in a clinical setting could cause harm to patients. The 2015 ACMG guidelines refer to the development of focused guidance for the classification of variants in specific genes since pathogenicity criteria vary by gene and disease (44). Indeed, gene specific guidelines for the interpretation of variants have been proposed for other diseases (23,45–47). These disease specific modifications have in some cases increased the number of variants classified as VUS and are not necessarily more permissive of missense variants (45). We suggest that a tailored set of ACMG-based guidelines may facilitate the standardisation of interpretation of variants in *PKD1.*

Studies of PKD in US, French, Italian and German populations have contributed to the development of a catalogue of PKD AR and AD pathogenic variants (5,12,48–51). However, the broader European population (and non-Caucasian populations) has not been studied in as much depth as the US one (17,50,52). There have been limited studies cataloguing the spectrum of AD and AR pathogenic variants in the UK and Ireland to date (16), although the inclusion of cystic renal disease patients in the Genomics England initiative will likely result in a catalogue of UK AD and AR pathogenic variants.

There were a number of limitations to this study. The coverage achieved for *PKD1* (NG_008617.1) exon 1 was not sufficient for variant identification, necessitating gap filling Sanger sequencing in patients without a molecular diagnosis from NGS. It is possible that due to low coverage of this region, or due to pseudogene homology to *PKD1* that some patients carried additional, clinically relevant variants which were not detected. The sequencing methods used in this study were also unable to detect small insertions and deletions in the variable number tandem repeat region of *MUC1*.

A genetic diagnosis not only informs on prognosis, but facilitates cascade testing of relatives considering donating a kidney to a family member, and provides information when planning a family or making decisions about care. A confirmed genetic diagnosis opens the possibility of pre-implantation screening for couples with a family history of PKD and may ensure that patients receive follow up testing for common PKD complications including cerebral aneurysms and liver cysts. We recommend that PKD testing adopts a combination of approaches for molecular diagnosis of PKD patients. Finally, we recommend further standardisation of variant pathogenicity assessment within the PKD community, but highlight some potential issues with the implementation of established ACMG pathogenicity guidelines which may require some modification for use in PKD.

Conflict of Interest Statement: The authors declare that they have no competing financial interests in relation to the work described.

REFERENCES

1. Suwabe T, Shukoor S, Chamberlain AM, Killian JM, King BF, Edwards M, et al. Epidemiology of Autosomal Dominant Polycystic Kidney Disease in Olmsted County. Clin J Am Soc Nephrol. 2020 Jan;15(1):69–79.
2. ERA-EDTA Registry Committee. ERA-EDTA Registry Annual Report 2017. 2017.
3. McEwan P, Bennett Wilton H, Ong ACM, Ørskov B, Sandford R, Scolari F, et al. A model to predict disease progression in patients with autosomal dominant polycystic kidney disease (ADPKD): the ADPKD Outcomes Model. BMC Nephrol. 2018 Dec;19(1):37.
4. Cornec-Le Gall E, Olson RJ, Besse W, Heyer CM, Gainullin VG, Smith JM, et al. Monoallelic Mutations to DNAJB11 Cause Atypical Autosomal-Dominant Polycystic Kidney Disease. Am J Hum Genet. 2018 May;102(5):832–44.
5. Audrézet MP, Cornec-Le Gall E, Chen JM, Redon S, Quéré I, Creff J, et al. Autosomal dominant polycystic kidney disease: Comprehensive mutation analysis of PKD1 and PKD2 in 700 unrelated patients. Hum Mutat. 2012 Aug;33(8):1239–50.
6. Cornec-Le Gall E, Audrézet M-P, Rousseau A, Hourmant M, Renaudineau E, Charasse C, et al. The PROPKD Score: A New Algorithm to Predict Renal Survival in Autosomal Dominant Polycystic Kidney Disease. J Am Soc Nephrol. 2016 Mar;27(3):942–51.
7. Heyer CM, Sundsbak JL, Abebe KZ, Chapman AB, Torres VE, Grantham JJ, et al. Predicted Mutation Strength of Nontruncating *PKD1* Mutations Aids Genotype-Phenotype Correlations in Autosomal Dominant Polycystic Kidney Disease. J Am Soc Nephrol. 2016 Sep;27(9):2872–84.
8. Rossetti S, Consugar MB, Chapman AB, Torres VE, Guay-Woodford LM, Grantham JJ, et al. Comprehensive Molecular Diagnostics in Autosomal Dominant Polycystic Kidney Disease. J Am Soc Nephrol. 2007 Jun;18(7):2143–60.
9. Porath B, Gainullin VG, Cornec-Le Gall E, Dillinger EK, Heyer CM, Hopp K, et al. Mutations in GANAB, Encoding the Glucosidase IIα Subunit, Cause Autosomal-Dominant Polycystic Kidney and Liver Disease. Am J Hum Genet. 2016 Jun;98(6):1193–207.
10. Cornec-Le Gall E, Torres VE, Harris PC. Genetic Complexity of Autosomal Dominant Polycystic Kidney and Liver Diseases. J Am Soc Nephrol. 2017 Oct;29(1):ASN.2017050483.
11. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature. 2006 Nov;444(7118):444–54.
12. Eisenberger T, Decker C, Hiersche M, Hamann RC, Decker E, Neuber S, et al. An Efficient and Comprehensive Strategy for Genetic Diagnostics of Polycystic Kidney Disease. 2015;
13. Consugar MB, Wong WC, Lundquist PA, Rossetti S, Kubly VJ, Walker DL, et al. Characterization of large rearrangements in autosomal dominant polycystic kidney disease and the PKD1/TSC2 contiguous gene syndrome. Kidney Int. 2008 Dec;74(11):1468–79.
14. Besse W, Chang AR, Luo JZ, Triffo WJ, Moore BS, Gulati A, et al. ALg9 mutation carriers develop kidney and liver cysts. J Am Soc Nephrol. 2019;30(11):2091–102.
15. Onuchic LF, Furu L, Nagasawa Y, Hou X, Eggermann T, Ren Z, et al. PKHD1, the Polycystic Kidney and Hepatic Disease 1 Gene, Encodes a Novel Large Protein Containing Multiple Immunoglobulin-Like Plexin-Transcription–Factor Domains and Parallel Beta-Helix 1 Repeats. Am J Hum Genet. 2002 May;70(5):1305–17.
16. Rossetti S, Burton S, Strmecki L, Pond GR, San Millán JL, Zerres K, et al. The position of the polycystic kidney disease 1 (PKD1) gene mutation correlates with the severity of renal disease. J Am Soc Nephrol. 2002 May;13(5):1230–7.
17. Cornec-Le Gall E, Audrézet M-P, Chen J-M, Hourmant M, Morin M-P, Perrichot R, et al. Type of PKD1 mutation influences renal outcome in ADPKD. J Am Soc Nephrol. 2013 May;24(6):1006–13.
18. Cornec-Le Gall E, Audrézet M-P, Renaudineau E, Hourmant M, Charasse C, Michez E, et al. PKD2 - Related Autosomal Dominant Polycystic Kidney Disease: Prevalence, Clinical Presentation,

Mutation Spectrum, and Prognosis. Am J Kidney Dis. 2017 Oct;70(4):476–85.
19. Ali H, Al-Mulla F, Hussain N, Naim M, Asbeutah AM, AlSahow A, et al. PKD1 Duplicated regions limit clinical Utility of Whole Exome Sequencing for Genetic Diagnosis of Autosomal Dominant Polycystic Kidney Disease. Sci Rep. 2019 Mar;9(1):4141.
20. Tan AY, Michaeel A, Liu G, Elemento O, Blumenfeld J, Donahue S, et al. Molecular diagnosis of autosomal dominant polycystic kidney disease using next-generation sequencing. J Mol Diagn. 2014 Mar;16(2):216–28.
21. Rossetti S, Hopp K, Sikkink R a, Sundsbak JL, Lee YK, Kubly V, et al. Identification of gene mutations in autosomal dominant polycystic kidney disease through targeted resequencing. J Am Soc Nephrol. 2012;
22. Borràs DM, Vossen RHAM, Liem M, Buermans HPJ, Dauwerse H, van Heusden D, et al. Detecting PKD1 variants in polycystic kidney disease patients by single-molecule long-read sequencing. Hum Mutat. 2017;38(7):870–9.
23. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. 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. 2015 May;17(5):405–23.
24. Rossetti S, Harris PC. Genotype-Phenotype Correlations in Autosomal Dominant and Autosomal Recessive Polycystic Kidney Disease. J Am Soc Nephrol. 2007;18:1374–80.
25. Gilbert E, Carmi S, Ennis S, Wilson JF, Cavalleri GL. Genomic insights into the population structure and history of the Irish Travellers. Sci Rep. 2017 Feb;7:42187.
26. Byrne RP, Martiniano R, Cassidy LM, Carrigan M, Hellenthal G, Hardiman O, et al. Insular Celtic population structure and genomic footprints of migration. Falush D, editor. PLOS Genet. 2018 Jan;14(1):e1007152.
27. Pei Y, Obaji J, Dupuis A, Paterson AD, Magistroni R, Dicks E, et al. Unified Criteria for Ultrasonographic Diagnosis of ADPKD. J Am Soc Nephrol. 2009 Jan;20(1):205–12.
28. Cormican S, Connaughton DM, Kennedy C, Murray S, Živná M, Kmoch S, et al. Autosomal dominant tubulointerstitial kidney disease (ADTKD) in Ireland. Ren Fail. 2019 Jan;41(1):832–41.
29. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;20(9):1297–303.
30. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013 Jan;Chapter 7:Unit7.20.
31. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009 Jul;4(7):1073–81.
32. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv. 2019 Jan;531210.
33. Gilbert E, O'Reilly S, Merrigan M, McGettigan D, Molloy AM, Brody LC, et al. The Irish DNA Atlas: Revealing Fine-Scale Population Structure and History within Ireland. Sci Rep. 2017 Dec;7(1):17199.
34. Desch KC, Ozel AB, Siemieniak D, Kalish Y, Shavit JA, Thornburg CD, et al. Linkage analysis identifies a locus for plasma von Willebrand factor undetected by genome-wide association. Proc Natl Acad Sci U S A. 2013 Jan;110(2):588–93.
35. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience. 2015 Feb;4(1).
36. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007

Sep;81(3):559–75.
37. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81(5):1084–97.
38. Browning BL, Browning SR. Improving the accuracy and efficiency of identity-by-descent detection in population data. Genetics. 2013 Jun;194(2):459–71.
39. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010 Nov;26(22):2867–73.
40. Vujic M, Heyer CM, Ars E, Hopp K, Markoff A, Orndal C, et al. Incompletely Penetrant PKD1 Alleles Mimic the Renal Manifestations of ARPKD. J Am Soc Nephrol. 2010 Jul;21(7):1097–102.
41. Magistroni R, He N, Wang K, Andrew R, Johnson A, Gabow P, et al. Genotype-renal function correlation in type 2 autosomal dominant polycystic kidney disease. J Am Soc Nephrol. 2003 May;14(5):1164–74.
42. Phelan PJ, Fletcher E, Carroll N, Metcalfe W, Turner AN. Simultaneous adult polycystic kidney disease and Alport syndrome. Nephrology. 2016 Aug;21(8):722–3.
43. Durkie M, Chong J, Valluru MK, Harris PC, Ong ACM. Biallelic inheritance of hypomorphic PKD1 variants is highly prevalent in very early onset polycystic kidney disease. Genet Med. 2020 Nov 10;
44. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. 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. 2015 May;17(5):405–24.
45. Muiño-Mosquera L, Steijns F, Audenaert T, Meerschaut I, De Paepe A, Steyaert W, et al. Tailoring the American College of Medical Genetics and Genomics and the Association for Molecular Pathology Guidelines for the Interpretation of Sequenced Variants in the *FBN1* Gene for Marfan Syndrome. Circ Genomic Precis Med. 2018 Jun 1;11(6):e002039.
46. Whiffin N, Walsh R, Govind R, Edwards M, Ahmad M, Zhang X, et al. CardioClassifier: disease- and gene-specific computational decision support for clinical genome interpretation. Genet Med. 2018 Oct;20(10):1246–54.
47. Luo X, Feurstein S, Mohan S, Porter CC, Jackson SA, Keel S, et al. ClinGen Myeloid Malignancy Variant Curation Expert Panel recommendations for germline RUNX1 variants. Blood Adv. 2019;3(20):2962–79.
48. Willey CJ, Blais JD, Hall AK, Krasa HB, Makin AJ, Czerwiec FS. Prevalence of autosomal dominant polycystic kidney disease in the European Union. Nephrol Dial Transplant. 2016 Jun;32(8):gfw240.
49. Rossetti S, Strmecki L, Gamble V, … SB-TAJ of, 2001 undefined. Mutation analysis of the entire PKD1 gene: genetic and diagnostic implications. Elsevier.
50. Carrera P, Calzavara S, Magistroni R, Den Dunnen JT, Rigo F, Stenirri S, et al. Deciphering Variability of PKD1 and PKD2 in an Italian Cohort of 643 Patients with Autosomal Dominant Polycystic Kidney Disease (ADPKD). Nat Publ Gr. 2016;
51. Mantovani V, Bin S, Graziano C, Capelli I, Minardi R, Aiello V, et al. Gene Panel Analysis in a Large Cohort of Patients With Autosomal Dominant Polycystic Kidney Disease Allows the Identification of 80 Potentially Causative Novel Variants and the Characterization of a Complex Genetic Architecture in a Subset of Families. Front Genet. 2020 May;11(May):1–14.
52. Audrézet M-P, Corbiere C, Lebbah S, Morinière V, Broux F, Louillet F, et al. Comprehensive PKD1 and PKD2 Mutation Analysis in Prenatal Autosomal Dominant Polycystic Kidney Disease. J Am Soc Nephrol. 2016 Mar;27(3):722–9.

Legends to Tables and Figures

Table 1: Baseline characteristics of patients with polycystic kidney disease in whom disease causing variants were and were not detected.

Comparison of the means between the groups with and without a molecular diagnosis are shown (t-test, assuming unequal variances, *p*-value for mean age at ESKD between solved and unsolved cases from Kaplan-Meier Log Rank (Mantel-Cox) test was *p*=0.063). Variants displayed as pathogenic here were classified using Mayo Clinic pathogenicity guidelines.

Table 1 Legend: Sig.(2 tailed)= significance (*p* value) from 2-tailed t-test test.

Table 2: Autosomal dominant and autosomal recessive pathogenic variants observed in multiple, apparently unrelated families

Table 2 Legend: MAF (gnomAD): Minor allele frequency in gnomAD database; Gene of Interest: Gene in which the disease causing variant was identified; Previously reported: Describes if the variant had previously been reported in the Mayo PKDB, ClinVar or OMIM databases; IBS: Identity by State; IBD: Identity by Descent; IBD Sharing Confirmation: If 'Y' indicates that families shared an IBD segment across the gene of interest, if 'N' indicates that no IBD segment was detected across the gene of interest; Kinship Relationship: Degree of relationship inferred by KING on the basis of IBS and IBD sharing

Figure 1: Detection Rates of molecular causes of PKD

Figure 1 Legend: Flow chart (above) describes the methodological workflow for the project alongside the number of AR and AD pathogenic variants identified at each stage. Chart (below) shows the rates of detection of pathogenic variants for each of the methods used. Variants displayed here were classified using Mayo Clinic pathogenicity guidelines.

Figure 2: Number and type of *PKD1* and *PKD2* AD pathogenic variants detected

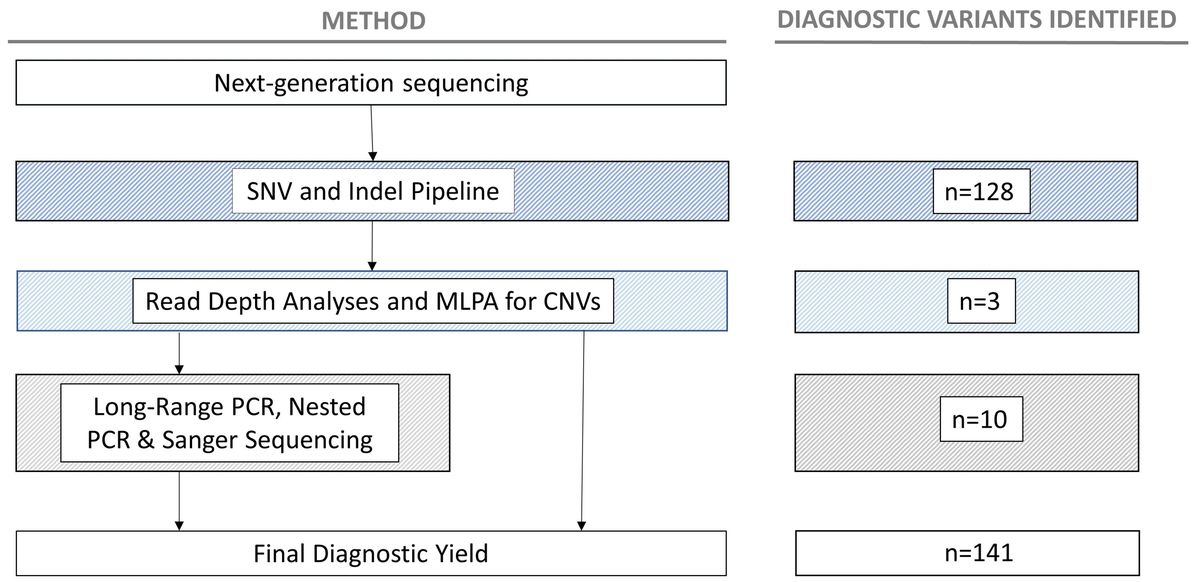
Figure 2 Legend: These charts describe the number and type of AD pathogenic variants identified in the Irish cohort (classified using Mayo Clinic pathogenicity guidelines). The central chart shows the gene in which variants were identified and the outer charts show the breakdown of variant types for each of those genes. Corresponding protein domains for the variants listed in this figure are shown in Supplementary Table 10.

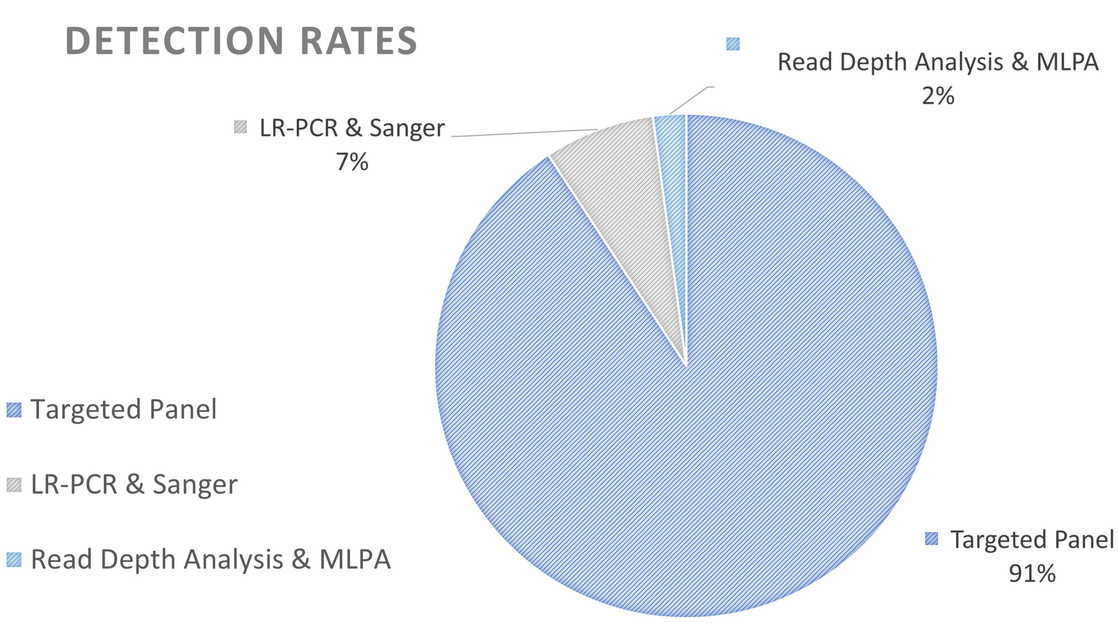
Figure 3: Diagnostic *PKD1* variants detected using custom next generation sequencing pipeline

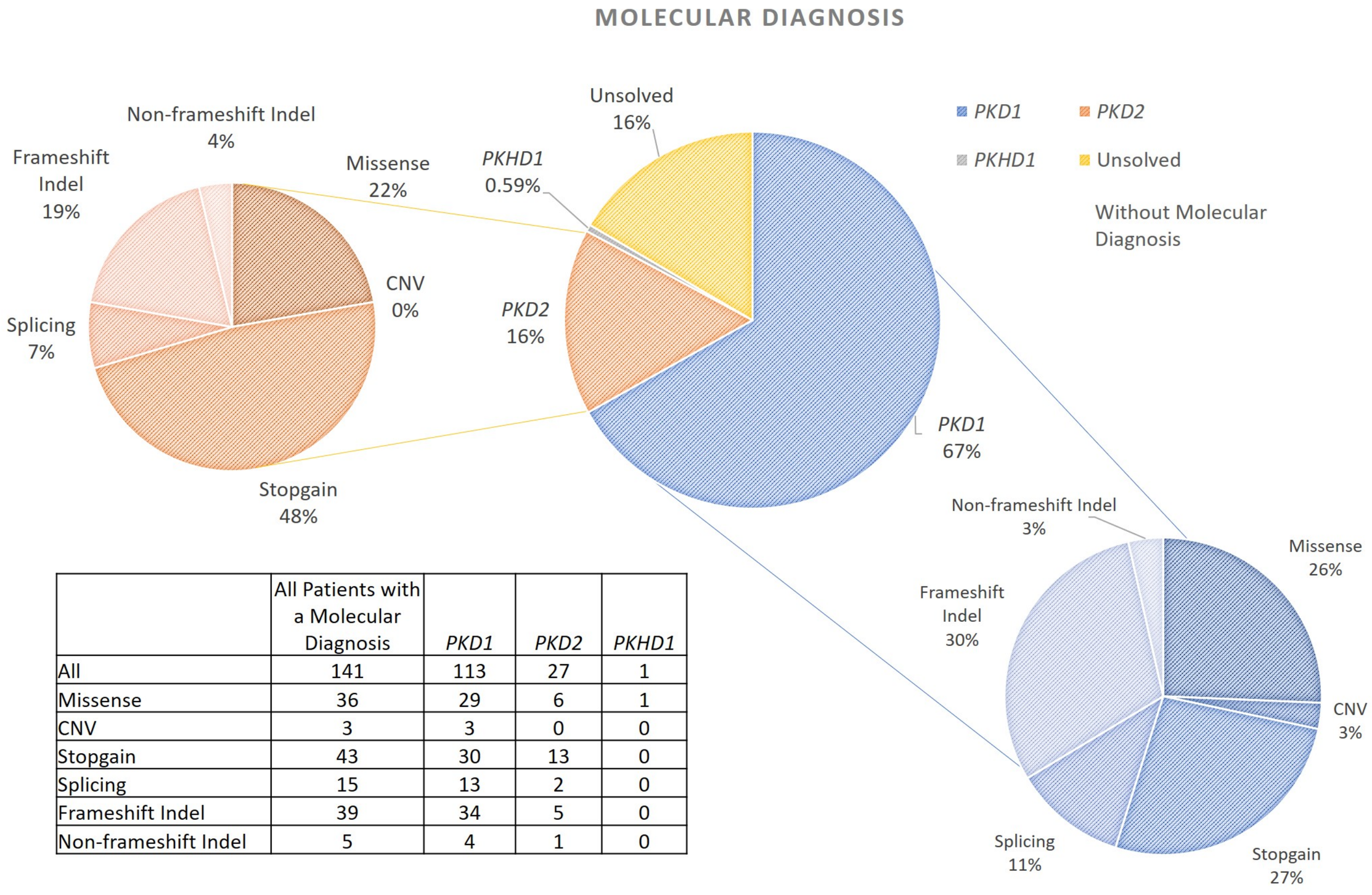
Figure 3 Legend: Novel pathogenic *PKD1* variants (classified using Mayo Clinic pathogenicity guidelines) detected in our Irish cohort shown across the *PKD1* gene. The gene is divided into exons (1-46, as per NG_008617.1) above and key domains are shown below. LRR: leucine-rich repeat; WSC: cell wall integrity and stress response component; REJ: receptor for egg jelly; GPS: G protein-coupled receptor proteolytic site; PLAT: polycystin-1, lipoxygenase, α toxin; TM: transmembrane

Figure 4: Kaplan Meier survival graph showing time to ESKD

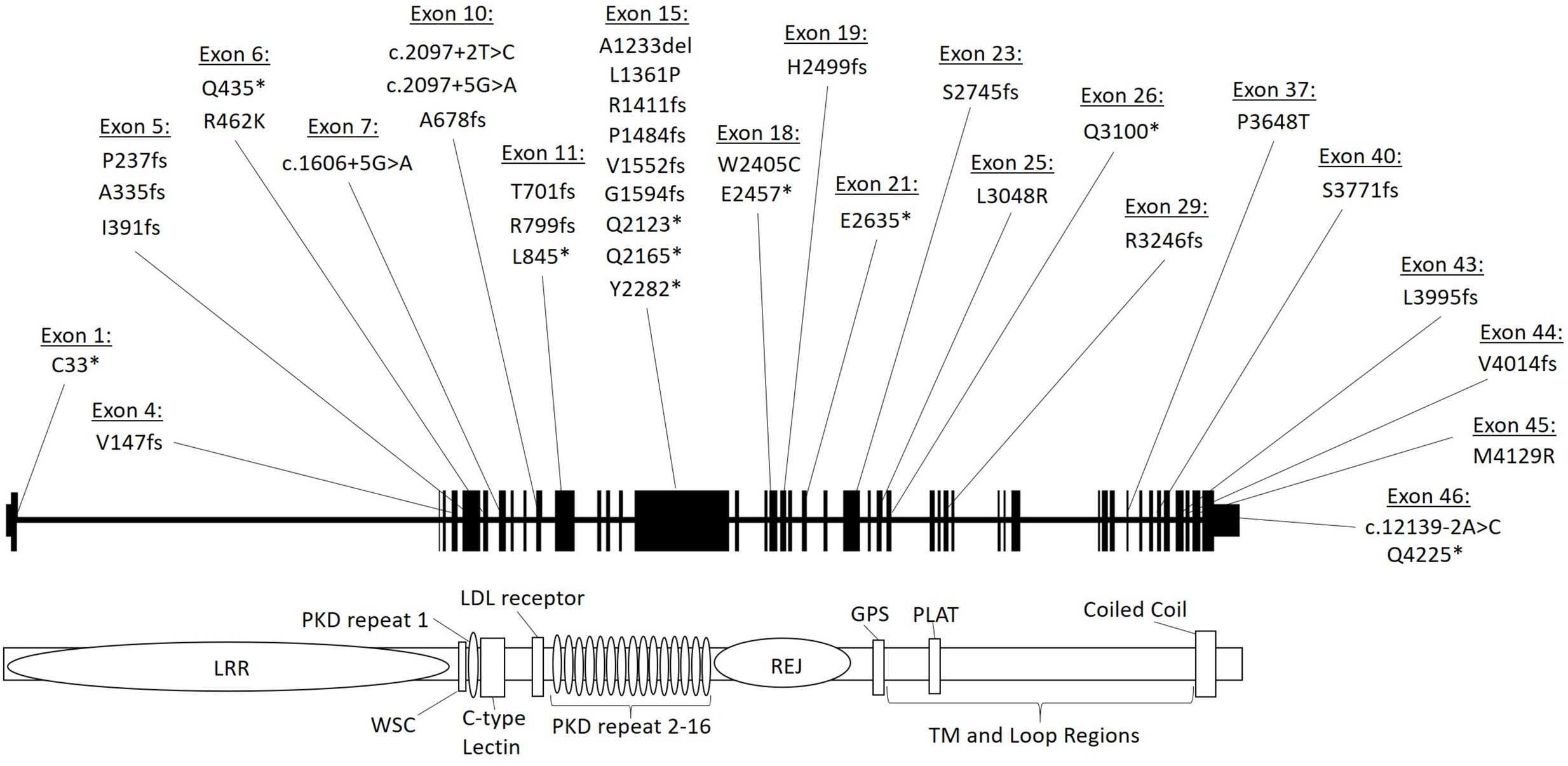
Figure 4 Legend: Kaplan Meier survival graph showing time to ESKD in those with pathogenic (classified using Mayo Clinic pathogenicity guidelines) truncating and non-truncating PKD1 variants and PKD2 variants. Results from the Log Rank (Mantel-Cox) are shown under the graph.

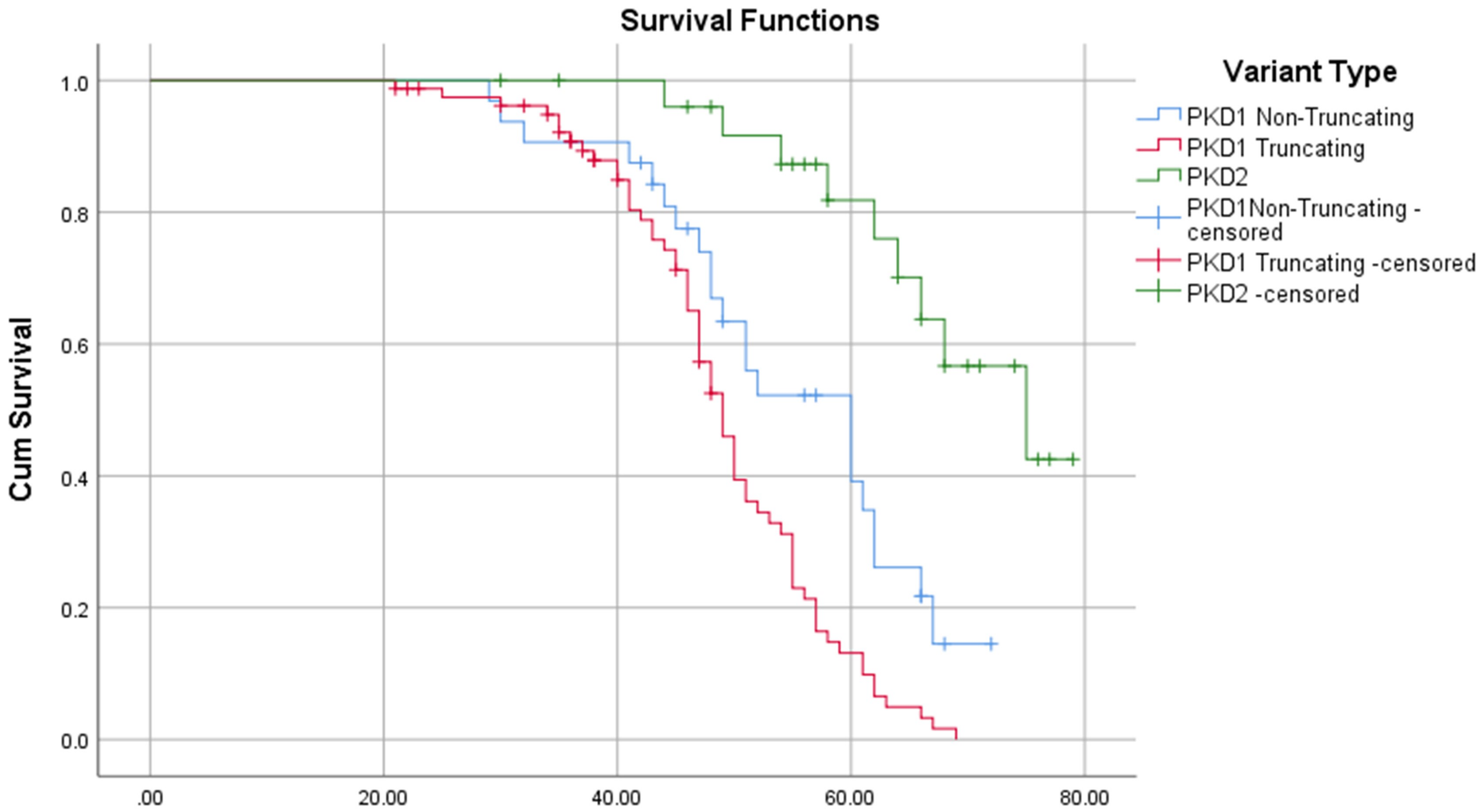

METHOD
DIAGNOSTIC VARIANTS IDENTIFIED
Next-generation sequencing
SNV and Indel Pipeline
n=128
Read Depth Analyses and MLPA for CNVs
n=3
Long-Range PCR, Nested PCR & Sanger Sequencing
n=10
Final Diagnostic Yield
n=141


DETECTION RATES
Read Depth Analysis & MLPA 2%
LR-PCR & Sanger 7%
Targeted Panel 91%
Targeted Panel
LR-PCR & Sanger
Read Depth Analysis & MLPA



	All Patients with a Molecular Diagnosis	*PKD1*	*PKD2*	*PKHD1*
All	141	113	27	1
Missense	36	29	6	1
CNV	3	3	0	0
Stopgain	43	30	13	0
Splicing	15	13	2	0
Frameshift Indel	39	34	5	0
Non-frameshift Indel	5	4	1	0


Exon 1:
C33*
Exon 4:
V147fs
Exon 5:
P237fs
A335fs
I391fs
Exon 6:
Q435*
R462K
Exon 7:
c.1606+5G>A
Exon 10:
c.2097+2T>C
c.2097+5G>A
A678fs
Exon 11:
T701fs
R799fs
L845*
Exon 15:
A1233del
L1361P
R1411fs
P1484fs
V1552fs
G1594fs
Q2123*
Q2165*
Y2282*
Exon 18:
W2405C
E2457*
Exon 19:
H2499fs
Exon 21:
E2635*
Exon 23:
S2745fs
Exon 25:
L3048R
Exon 26:
Q3100*
Exon 29:
R3246fs
Exon 37:
P3648T
Exon 40:
S3771fs
Exon 43:
L3995fs
Exon 44:
V4014fs
Exon 45:
M4129R
Exon 46:
c.12139-2A>C
Q4225*
LDL receptor
PKD repeat 1
GPS
PLAT
Coiled Coil
LRR
REJ
WSC
C-type
Lectin
PKD repeat 2-16
TM and Loop Regions



Renal Survival (Age at ESKD)

Pairwise Comparisons

Group		PKD1 Non-Truncating		PKD1 Truncating		PKD2	
		Chi-Square	Sig.	Chi-Square	Sig.	Chi-Square	Sig.
Log Rank (Mantel-Cox)	PKD1 Non-Truncating			9.187	0.002	11.660	0.001
	PKD1 Truncating	9.187	0.002			41.020	1.507E-10
	PKD2	11.660	0.001	41.020	1.507E-10		

Table 1: Baseline characteristics of patients with polycystic kidney disease in whom disease causing variants were and were not detected.

	All Patients (n=169) (%)	Solved (n=141) (%)	Unsolved (n=28) (%)	Sig. (2-taile	95% Confidence Interval	
					Lower	Upper
Mean Age at Diagnosis (Range)	36 (4-79)	34 (4-66)	45 (16-79)	0.003	-16.804	-3.47
Male Sex	84 (50%)	67 (48%)	17 (61%)	0.204	-3.364	0.725
Documented Family History	131 (78%)	112 (79%)	19 (68%)	0.137	-0.409	0.295
Known Liver Cysts	85 (50%)	74 (52%)	11 (39%)	0.278	-0.102	0.353
Known Aneurym	16 (9%)	14 (10%)	2 (7%)	0.035	-0.968	0.167
Reached ESRD	114 (67%)	96 (68%)	18 (64%)	0.66	-0.149	0.235
Mean age at ESRD (Range)	50 (21-84)	49 (21-75)	54 (25-84)	0.156	-9.909	1.61
Documented Hypertension	138 (82%)	119 (84%)	19 (68%)	0.093	-0.223	0.285

Gene	Type of Va	Change	hg38 coord	MAF (gnom	Previously	Number of	IBD Sharing Confirmation	Kinship Relationship
PKD1	splicing	c.1606+5G>A	chr16:211	0	N	2	Y	2nd degree
PKD1	missense	p.L1361P	chr16:211	0	N	2	Y	2nd degree
PKD1	missense	p.L3142P	chr16:210	0	Y	3	N	No relationship Identified
PKD1	stopgain	p.Q2670*	chr16:210	0	Y	2	N	No relationship Identified
PKD1	stopgain	p.Q2686*	chr16:210	0	Y	2	Y	2nd-3rd degree
PKD1	stopgain	p.Q4225*	chr16:208	0	N	3	Y	2 of 3 families related - 2nd degree
PKD1	frameshift	p.R1672fs	chr16:211	0	Y	2	N	No relationship Identified
PKD1	missense	p.R3750Q	chr16:209	4.04E-06	Y	2	N	No relationship Identified
PKD1	missense	p.R611W	chr16:211	0	Y	2	Y	Full Sibling
PKD1	stopgain	p.Y3781*	chr16:209	0	Y	2	Y	4th degree
PKD2	stopgain	p.E494*	chr4:8804	0	Y	3	Y	2 of 3 families related - 2nd degree
PKD2	missense	p.R322W	chr4:8803	0	Y	4	2 of 4 families share IBD segment	2 of 4 families related - 3rd degree
PKD2	stopgain	p.W201*	chr4:8801	0	Y	3	Y	No relationship Identified
PKD2	stopgain	p.Y762*	chr4:8806	0	Y	2	Y	Parent Offspring