Received: 24 March 2020 | Revised: 12 July 2020 | Accepted: 27 July 2020

DOI: 10.1002/mgg3.1467

ORIGINAL ARTICLE

Molecular Genetics & Genomic Medicine Open Access WILEY

Disease causing property analyzation of variants in 12 Chinese families with polycystic kidney disease

Kexian Dong[1,2] | **Xiaogang Liu**[3] | **Xueyuan Jia**[1,2] | **Huanhuan Miao**[4] | **Wei Ji**[1,2] | **Jie Wu**[1,2] | **Yun Huang**[1,2] | **Lidan Xu**[1,2] | **Xuelong Zhang**[1,2] | **Hui Su**[1,2] | **Guohua Ji**[1,2] | **Peng Liu**[1,2] | **Rongwei Guan**[1,2] | **Jing Bai**[1,2] | **Songbin Fu**[1,2] | **Xianli Zhou**[4] | **Wenjing Sun**[1,2]

[1]Laboratory of Medical Genetics, Harbin Medical University, Harbin, China

[2]Key Laboratory of Preservation of Human Genetic Resources and Disease Control in China, Ministry of Education, Harbin Medical University, Harbin, China

[3]Department of Nephrology, The First Affiliated Hospital of Harbin Medical University, Harbin, China

[4]In-Patient Ultrasound Department, The Second Affiliated Hospital of Harbin Medical University, Harbin, China

Correspondence

XianLi Zhou, 246 Xuefu Road, Nangang District, Harbin 150086, China.
Email: hrbzhouxl@163.com

Wenjing Sun, 157 Baojian Road, Nangang District, Harbin 150081, China.
Email: sunwj@ems.hrbmu.edu.cn

Funding information

National Key Research and Development Program of China, Grant/Award Number: 2016YFC1000504

Abstract

Background: Polycystic kidney disease (PKD) is an inherited disease that is life-threatening. Multiple cysts are present in the bilateral kidneys of PKD patients. The progressively enlarged cysts cause structural damage and loss of kidney function.

Methods: This study examined and analyzed 12 families with polycystic kidney disease. Whole exome sequencing (WES) or whole genome sequencing (WGS) of the probands was performed to detect the pathogenic genes. The candidate gene segments for lineal consanguinity in the family were amplified by the nest PCR followed by Sanger sequencing. The variants were assessed by pathogenic and conservational property prediction analysis and interpreted according to the American College of Medical Genetics and Genomics.

Results: Nine of the 12 pedigrees were identified the disease causing variants. Among them, four novel variants in *PKD1*, c.6930delG:p.C2311Vfs*3, c.1216T>C:p.C406R, c.8548T>C:p.S2850P, and c.3865G>A:p.V1289M (NM_001009944.2) were detected. After assessment, the four novel variants were considered to be pathogenic variants and cause autosomal dominant polycystic kidney disease in family. The detected variants were interpreted.

Conclusion: The four novel variants in *PKD1*, c.6930delG:p.C2311Vfs*3, c.1216T>C:p.C406R,c.8548T>C:p.S2850P,andc.3865G>A:p.V1289M(NM_001009944.2) are pathogenic variants and cause autosomal dominant polycystic kidney disease in family.

KEYWORDS

autosomal dominant, inheritance, novel variants, pedigree, polycystic kidney disease

Kexian Dong and Xiaogang Liu contributed equally to this work.

1 | INTRODUCTION

Polycystic kidney disease (PKD) is a monogenic inherited disease involving bilateral kidneys, impacting renal structure and function (Obeidova et al., 2014). According to the inherited pattern, PKD is divided into autosomal dominant polycystic kidney disease (ADPKD, OMIM#173900) and autosomal recessive polycystic kidney disease (ARPKD, OMIM#263200). ADPKD affects more than 12 million individuals worldwide with a morbidity of 1:500–1:1000 (Lanktree et al., 2018; Willey et al., 2017). ADPKD is a life-threatening disease, impacting extrarenal organs including liver, brain, and pancreas, resulting end-stage kidney disease (ESRD) (Casteleijn, Spithoven, Rookmaaker, Vergouwen, & Gansevoort, 2015; Edrees et al., 2016; Litvinchuk, Tao, Singh, & Vasylyeva, 2015). Extrarenal cysts are also common in ADPKD, especially liver cysts, and 94% of ADPKD patients older than 35 have polycystic liver disease. The incidence of cysts in female is higher than that in male. 7%–36% of ADPKD patients have pancreatic cysts, and the incidence of pancreatic cysts is higher in patients with *PKD2* mutations (Kim et al., 2016). 25%–30% of ADPKD patients have cardiovascular disease. Cardiovascular complications, especially cardiac hypertrophy and coronary artery disease, are the leading cause of death in patients with ADPKD (Krishnappa, Vinod, Deverakonda, & Raina, 2017). About 9%–12% of ADPKD patients suffer from intracranial aneurysm (Flahault & Joly, 2019). About 45% of ADPKD patients have abdominal hernias, which may be associated with kidney enlargement (Li et al., 2011).

Two acknowledged genes, *PKD1* (78% in affected pedigrees) and *PKD2* (15% in affected pedigrees), are considered as the pathogenic genes of ADPKD (Cornec-Le Gall, 2018). However, approximately 10% of patients with ADPKD have no variant detected in either *PKD1* or *PKD2* (Meng et al., 2018; Porath et al., 2016). Polycystin-1 (PC1) is the protein product of *PKD1*, which is an integral membrane protein with 11 membrane-spanning domains. The protein product of *PKD2*, polycystin-2 (PC2), is also located on the cell membrane as a nonselective Ca^{2+} channel. The C-termini of PC1 and PC2 interact to regulate ion transportation (Ghata & Cowley, 2017; Hafer & Conran, 2017; Kim & Park, 2016; Kinoshita et al., 2016). According to Autosomal Dominant Polycystic Kidney Disease Mutation Database (PKDB), more than 1000 pathogenic variants in *PKD1* and more than 200 pathogenic variants in *PKD2* were detected in ADPKD pedigrees. Generally, patients with ADPKD have no apparently symptoms until 40 s age (Pei et al., 2015). Thus, molecular diagnosis of an individual with ADPKD family history is meaningful.

In this study, we collected 12 families with ADPKD in North China. WES or WGS of the blood samples from 12 probands were performed to detect the candidate pathogenic variants. To identify the disease causing variants, bioinformatic analysis was performed. The nest PCR was used to amplify the candidate regions in the patients and other individuals in the family followed by Sanger sequencing. Several prediction tools were used to assess and analyze the pathogenic and conservational properties of the variants.

2 | MATERIALS AND METHODS

2.1 | Ethical compliance

The study protocol was approved by Institutional Research Board of Harbin Medical University (protocol number HMUIRB20190010) and all participants provided written informed consent.

2.2 | Subjects

In this study, 12 unrelated families in Northeast China with polycystic kidney disease were analyzed. Clinical information and peripheral blood samples from the probands and other available family members were obtained. The peripheral blood from the individuals was collected into a qualified negative pressure vacuum EDTA anticoagulant tube. The pedigree maps of Family 1–4 were shown in Figure 1a,d,f,h. The pedigree maps of Family 5–12 were shown in Figures S1A, S2A, S3A, S4A, S5A, S6A, S7 and S8. More information was described in the Materials and Methods part of Supplementary files.

2.3 | Pathogenic gene detection

Whole exome sequencing (WES) and whole genome sequencing (WGS) of the blood sample from the probands were performed by Novogene technology limited-liability Company (Beijing, China). The disease of the proband of Family 5 was a little more complicated. She had not only polycystic kidney and liver, but also choledochal cystic dilatation combined with cholangiocarcinoma. To further study the case, we did WGS for Family 5, while other families were detected by WES. To perform WES, genomic DNA extracted from peripheral blood for each sample was fragmented to an average size of 180–280 bp and used to create a DNA library following established Illumina paired-end protocols. The Agilent SureSelect Human All Exon V6 Kit (Agilent Technologies, Santa Clara, CA, USA) was used for exome capture according to the manufacturer's instructions. The Illumina Novaseq 6000 platform (Illumina Inc., San Diego, CA, USA) was used for genomic DNA sequencing by Novogene Bioinformatics Technology Co., Ltd (Beijing,

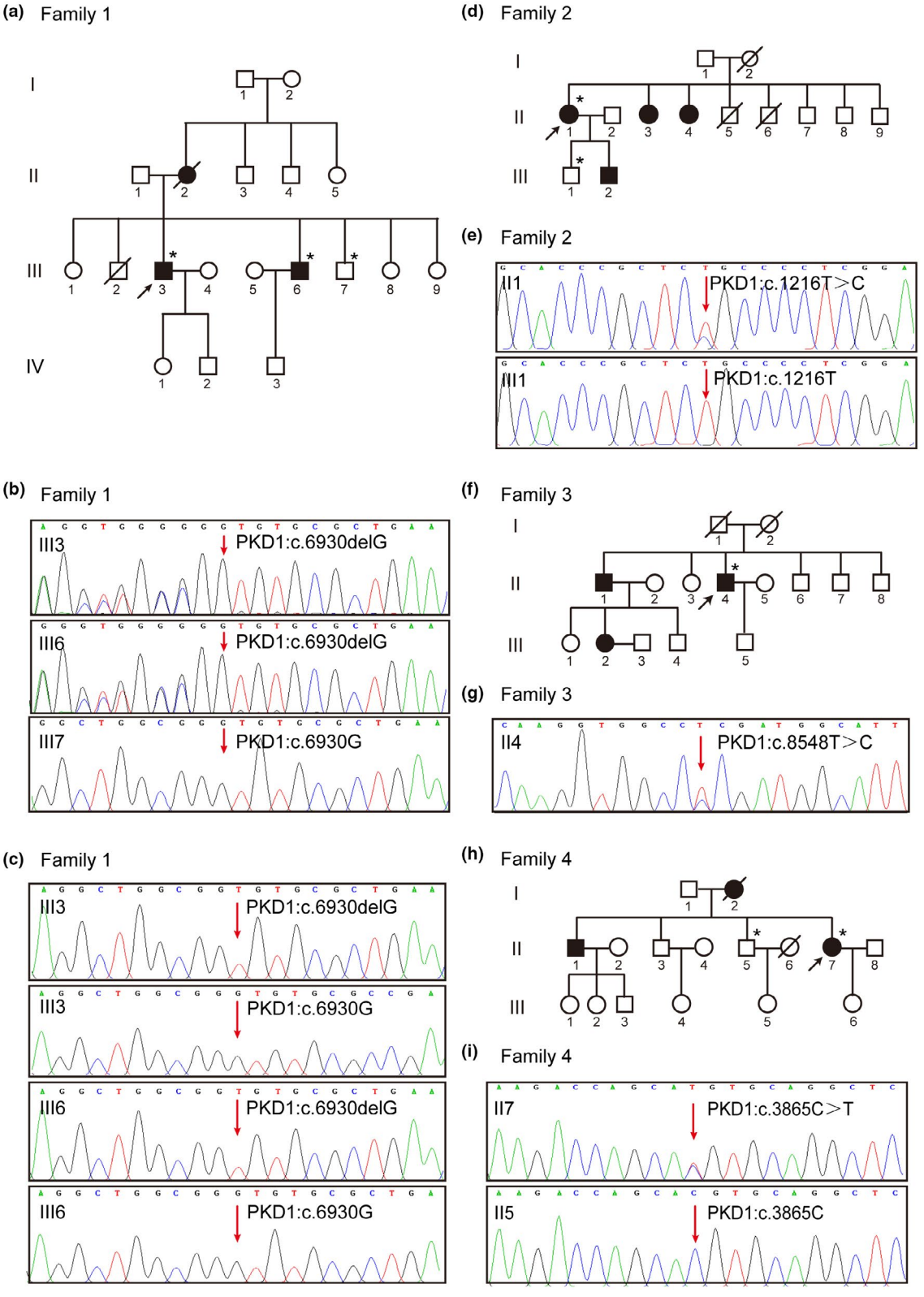

FIGURE 1 Pedigrees of Family 1–4 and the Sanger sequencing results. (a) The pedigree of Family 1. (b) The Sanger sequencing results showing the deletion variant *PKD1*:c.6930delG:p.C2311Vfs*3 in Family 1. (c) The T-vector ligation product from the proband showing the mutated and wild-type sequences of III3 and III6. (d) The pedigree of Family 2. (e) The Sanger sequencing results showing the missense variant *PKD1*:c.1216T>C:p.C406R of II1 and III1. (f) The pedigree of Family 3. (g) The Sanger sequencing results showing the missense variant *PKD1*:c.8548T>C:p.S2850P of III4. (h) The pedigree of Family 4. (i) The Sanger sequencing results showing the missense variant *PKD1*:c.3865G>A:p.V1289M of II7 and II5. A square represents male, and a circle represents female. Black indicates patients. The black arrow indicates the probands. Asterisks indicate the individuals who were clinically examined and underwent genetic analyses. The red arrows indicate the position of the duplication variant or the corresponding wild-type base

China) to generate 150 bp paired-end reads with a minimum coverage of 10× for ~99% of the genome (mean coverage of 100×). As for WGS, genomic DNA extracted from the peripheral blood was fragmented to an average size of ~350 bp and used to create a DNA library following established Illumina paired-end protocols. The Illumina Novaseq 6000 platform was used for genomic DNA sequencing by Novogene Bioinformatics Technology Co., Ltd to generate 150 bp paired-end reads with a minimum coverage of 10× for ~99% of the genome (mean coverage of 30×).

DNA extraction was performed using the DNeasy Blood & Tissue Kit (Qiagen, #69506, Dusseldorf, Germany) according to the manufacturer's protocol. Briefly, a DNA collection spin tube was used to isolate genomic DNA from the blood sample. Because of the complexity of the *PKD1* gene, long-range PCR (LR-PCR) was used as the first step to amplify the *PKD1* (NC_000016.10) exons 15–21, exons 2–8, exons 22–26, exons 13–15, exons 1–8, and exons 27–34, respectively. The LR-PCR products were amplified as the template to obtain the candidate region within the 16th exon, 6th exon, 26th exon, 15th exon, 5th exon, 22th exon, and 29th exon. The candidate regions in *PKD2* (NC_000004.12) were amplified by common PCR. The information of those primers was shown in Table S1. The PCR products were sequenced by TsingKe Biological Technology (Beijing, China) using Sanger sequencing.

2.4 | Variants analysis

In WES and WGS, the reads that aligned to exon regions were collected for variant calling and identification of SNV and indels using Samtools mpileup and bcftools (GRCh37/hg19). The frequencies of variants were evaluated in the 1000 Genome Project (https://www.1000genomes.org), the Exome Aggregation Consortium (ExAc) and the Genome Aggregation Database (gnomAD) (http://gnomad.broadinstitute.org). The significance of the variants was assessed using Online Mendelian Inheritance in Man (OMIM, http://omim.org), the Human Gene Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac/all.php), and the Autosomal Dominant Polycystic Kidney Disease Database (PKDB, http://pkdb.pkdcure.org); as well as SIFT (http://sift.jcvi.org), Polyphen2_HVAR (http://genetics.bwh.harvard.edu/pph2/), MutationTaster (http://www.mutationtaster.org), and Mutation Assessor (http://mutationassessor.org/r3/) substitution assessment tools. The conserved properties of the variants were analyzed using University of California Santa Cruz Genome Browser (UCSC, http://genome.ucsc.edu). The prevalence of the studied variants was compared with that of 51 individuals in control group. The interpretation of the variants was according to the American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015).

3 | RESULTS

3.1 | Clinical presentation

Twelve unrelated Chinese families with polycystic kidney disease were studied. The probands' clinical examination results of kidney function and ultrasound of the abdominal urology were shown in Table 1. Among the 12 probands, nine of them were female and three were male. According to the ultrasound examination of abdominal urinary organs, besides kidneys, the livers of all the probands were also affected with multiple cysts. They were regularly received kidney or liver puncture to extract the cystic liquid. There was a case also affected with choledochal cystic dilatation combined with cholangiocarcinoma (Family 5). The levels of urea, creatinine and uric acid were examined in the patients of Families 1, 2, 3, 8, 9, 10, and 11. All of the examining results of urea and creatinine exceeded the normal level, while the uric acid levels in Family 2 and 9 lied among the normal range.

3.2 | Four novel variants in *PKD1* are detected among 11 candidate disease causing variants

To identify the pathogenic genes involved in the families with polycystic kidney disease, WES or WGS was performed on the probands. The results showed *PKD1* variants were detected in Families 1–6, 9, and 10, *PKD2* variants were detected in Families 7–9. No candidate variant in *PKD1* or *PKD2* was detected in Families 11 and 12. The sequences were blat to GRCh37/hg19. *PKD1* transcript was NM_001009944.2, encoding protein NP_001009944.2. *PKD2* transcript was NM_000297.4, encoding protein NP_000288.1. The variants information of the ADPKD families were summarized in Table 2.

To identify the variants in the families, targeted DNA fragments from the available individuals including the probands were PCR amplified followed by Sanger sequencing and analyzed. The Sanger sequencing results indicated that the patients with polycystic kidney disease carried the variants in *PKD1*, while the normal ones did not carry the mutation sites (Figure 1b,e,g,i, Figures S1B, S2B, S3B, S4B, S5B and S6B). Notably, the four variants in *PKD1*, c.6930delG:p.C2311Vfs*3 in Family 1, c.1216T>C:p.C406R in Family 2, c.8548T>C:p.S2850P in Family 3 and c.3865G>A:p.V1289M in Family 4 were the novel variants that had not been published in PKDB, OMIM, HGMD, or any other published articles. Since there was one base pair deletion of *PKD1* gene in a heterozygous form in WES data of Family 1, further investigation into the proband's sequence was performed by ligating the nest PCR product to the T linear vector and sequencing. The sequencing results showed that the patients (III3 and III6) carried one mutated sequence and

TABLE 1 Clinical examination results of the ADPKD probands

Proband	Age	Gender	Urea (normal range 2.9–8.2 mmol/l)	Creatinine (normal range 44–97 μmol/l)	Uric acid (normal range 208–428 μmol/l)	Affected organs
Family 1	55	Male	37.36 ↑	909.7 ↑	569.6 ↑	Kidney, liver
Family 2	60	Female	24.59 ↑	414.6 ↑	358.4	Kidney, liver, pancreas enlargement
Family 3	50	Male	12.24 ↑	169.0 ↑	517.9 ↑	Kidney, liver
Family 4	47	Female	N/A[a]	N/A	N/A	Kidney, liver
Family 5	56	Female	N/A	N/A	N/A	Kidney, liver, choledochal cystic dilatation, cholangiocarcinoma
Family 6	53	Female	N/A	N/A	N/A	Kidney, liver
Family 7	42	Female	N/A	N/A	N/A	Kidney, liver
Family 8	65	Female	9.26 ↑	200 ↑	405.2 ↑	Kidney, liver
Family 9	61	Female	9.93 ↑	122 ↑	372.7	Kidney, liver
Family 10	28	Female	40.54↑	666 ↑	530.7 ↑	Kidney, liver
Family 11	45	Female	8.43 ↑	214 ↑	450.0 ↑	Kidney, liver
Family 12	58	Male	N/A	N/A	N/A	Kidney, liver

[a]Not available.

TABLE 2 Candidate pathogenic variants information

Pedigree	Gene	Exon	Variant	Genotype	NM number	Amino acid change	PKDB[a]
Family 1	*PKD1*	16	Deletion	Heterozygote	NM_001009944	c.6930delG:p.C2311Vfs*3	—
Family 2	*PKD1*	6	Missense	Heterozygote	NM_001009944	c.1216T>C:p.C406R	—
Family 3	*PKD1*	26	Missense	Heterozygote	NM_001009944	c.8548T>C:p.S2850P	—
Family 4	*PKD1*	15	Missense	Heterozygote	NM_001009944	c.3865G>A:p.V1289M	—
Family 5	*PKD1*	5	Nonsense	Heterozygote	NM_001009944	c.1198C>T:p.R400X	Definitely Pathogenic
Family 6	*PKD1*	22	Nonsense	Heterozygote	NM_001009944	c.8095C>T:p.Q2699X	Definitely Pathogenic
Family 7	*PKD2*	6	Nonsense	Heterozygote	NM_000297	c.1390C>T:p.R464X	Definitely Pathogenic
Family 8	*PKD2*	4	Nonsense	Heterozygote	NM_000297	c.916C>T:p.R306X	Definitely Pathogenic
Family 9	*PKD2*	8	Nonsense	Heterozygote	NM_000297	c.1774C>T:p.R592X	Definitely Pathogenic
	PKD1	29	Missense	Heterozygote	NM_001009944	c.9884A>G:p.N3295S	Indeterminate
Family 10	*PKD1*	40	Missense	Heterozygote	NM_001009944	c.11333C>A:p.T3778N	Likely Neutral

[a]The clinical significance of variants in Polycystic Kidney Disease Database, http://pkdb.pkdcure.org.

one normal sequence (Figure 1c), which confirmed the mutation existed as a heterozygous form. In summary, 11 candidate disease causing variants were detected in 10 of 12 ADPKD pedigrees. Among them, four variants in *PKD1* were novel.

3.3 | The pathogenic and conservational properties analyzation of the 11 candidate disease causing variants

To further evaluate the significance of the 11 variant sites in Family 1–10, the pathogenic and conservational properties were predicted and analyzed using multiple prediction tools (Table 3). The frequency information in the population of the variants were detected in the gnomAD, 1000 Genome Program and ExAC. The frequencies of all the 11 variants were less than 0.01 or not available, which meant those variants barely occurred in normal persons.

Several prediction tools were used to assess the pathogenicity of the variants. The variants *PKD1*:c.1216T>C:p.C406R in Family 2, *PKD1*:c.8548T>C:p.S2850P in Family 3, *PKD1*:c.3865G>A:p.V1289M in Family 4, *PKD1*:c.9884A>G:p.N3295S in Family 9 and *PKD1*:c.11333C>A:p.T3778N in Family 10 were missense variants,

TABLE 3 The pathogenic and conservational properties of the candidate disease causing variants

Pedigree	Gene	Amino acid change	Novel	gnomAD[a]	1000G[b]	ExAC[c]	SIFT[d]	Polyphen2_HVAR[e]	MutationTaster[f]	Mutation Assessor[g]	PhyloP[h]	PhastCons[i]
Family 1	*PKD1*	c.6930delG:p.C2311Vfs*3	Yes	—	—	—	—	—	1, A	—	0.155	0.322
Family 2	*PKD1*	c.1216T>C:p.C406R	Yes	—	—	—	0.000, D	0.999, D	1, D	2.615, M	3.568	1
Family 3	*PKD1*	c.8548T>C:p.S2850P	Yes	—	—	—	0.012, D	0.999, D	0.999988, D	2.085, M	4.372	0.997
Family 4	*PKD1*	c.3865G>A:p.V1289M	Yes	2.56E-05	—	1.84E-05	0.031, D	0.977, D	0.000047, N	1.71, L	0.561	0.029
Family 5	*PKD1*	c.1198C>T:p.R400X	No	1.49E-05	—	—	—	—	1, A	—	0.818	0.001
Family 6	*PKD1*	c.8095C>T:p.Q2699X	No	—	—	—	—	—	1, A	—	5.179	1
Family 7	*PKD2*	c.1390C>T:p.R464X	No	—	—	—	—	—	1, A	—	3.305	0.981
Family 8	*PKD2*	c.916C>T:p.R306X	No	4.07E-06	—	—	—	—	1, A	—	4.106	1
Family 9	*PKD2*	c.1774C>T:p.R592X	No	—	—	—	—	—	1, A	—	4.357	1
	PKD1	c.9884A>G:p.N3295S	No	4.40E-05	—	5.02E-05	0.083, T	0.998, D	0.998926, D	1.72, L	2.757	1
Family 10	*PKD1*	c.11333C>A:p.T3778N	No	—	0.00166113	0.0004	0.062, T	0.173, B	0.864342, D	2.205, M	1.614	0.905

[a]The frequencies listed in the Genome Aggregation Database (gnomAD, October 2016).

[b]The frequencies listed in the 1000 Genomes Project (1000 G, August 2015).

[c]The frequencies listed in the Exome Aggregation Consortium (ExAC, August 2016).

[d]Score range from 0–1.0. Scores ≤0.05 indicate damaging effects and scores >0.05 indicate tolerated effects.

[e]Score range from 0–1.0, where 1.0 is more damaging.

[f]Score range from 0–215, where a smaller score is more different than the original amino acid. The result can be A (disease causing automatic), D (disease causing), N (polymorphism), or P (polymorphism automatic).

[g]If the score is larger and more pathogenic, the result indicates a change in the function as H (high), M (medium), L (low) or N (neutral).

[h]Score range from −14 to +6. Sites predicted to be conserved are assigned positive scores, while sites predicted to be fast-evolving are assigned negative scores.

[i]Score range from 0 to 1, where the closer the value is to 1, the more probable the nucleotide is conserved.

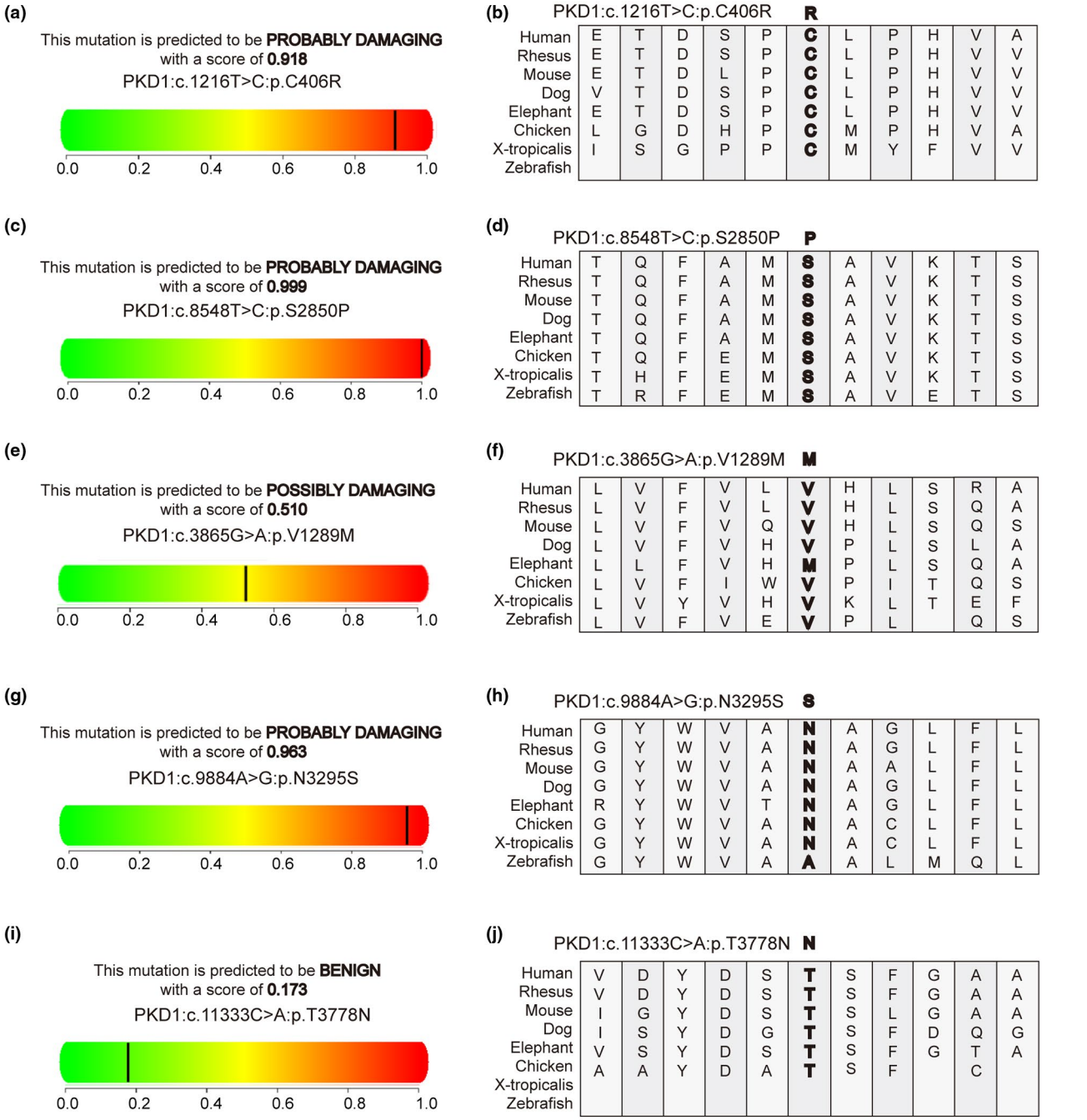

FIGURE 2 The pathogenic and conservational properties of missense variants in Family 2, Family 3, Family 4, Family 9, and Family 10. (a) The pathogenic properties of *PKD1*:c.1216T>C:p.C406R according to the prediction tool Polyphen2_HVAR. (b) The conservation of p.C406 using the UCSC Genome Browser. (c) The pathogenic properties of *PKD1*:c.8548T>C:p.S2850P according to the prediction tool Polyphen2_HVAR. (d) The conservation of p.S2850 using the UCSC Genome Browser. (e) The pathogenic properties of *PKD1*:c.3865G>A:p.V1289M according to the prediction tool Polyphen2_HVAR. (f) The conservation of p.V1289 using the UCSC Genome Browser. (g) The pathogenic properties of *PKD1*:c.9884A>G:p.N3295S according to the prediction tool Polyphen2_HVAR. (h) The conservation of p.N3295 using the UCSC Genome Browser. (i) The pathogenic properties of *PKD1*:c.11333C>A:p.T3778N according to the prediction tool Polyphen2_HVAR. (j) The conservation of p.T3778 using the UCSC Genome Browser

which were predicted by SIFT, Polyphen2_HVAR, MutationTaster, and MutationAssessor for the pathogenic properties. The prediction results of *PKD1*:c.1216T>C:p.C406R in Family 2 and *PKD1*:c.8548T>C:p.S2850P in Family 3 were consistent and showed to be disease causing, changing the protein function in a medium degree (Figure 2a,c). However, some prediction results were not consistent to each other when they predicted the variants *PKD1*:c.3865G>A:p.V1289M in Family 4, *PKD1*:c.9884A>G:p.N3295S in Family 9 and

PKD1:c.11333C>A:p.T3778N in Family 10. The missense variant *PKD1*:c.3865G>A:p.V1289M in Family 4 was predicted to be pathogenic in Polyphen2_HVAR and SIFT (Figure 2e), while not strong disease causing variants according to the other two prediction tools. The missense variant *PKD1*:c.9884A>G:p.N3295S in Family 9 was predicted to be disease causing in Polyphen2_HVAR (Figure 2g) and MutationTaster, while may be benign according to SIFT and MutationAssessor. The missense variant *PKD1*:c.11333C>A:p.T3778N in Family 10 was predicted to be disease causing in MutationTaster and change the protein function in a medium degree according to MutationAssessor, but may be benign predicted by SIFT and Polyphen2_HVAR (Figure 2i). The variant *PKD1*:c.6930delG:p.C2311Vfs*3 in Family 1 is a deletion variant, which cannot be predicted using those pathogenic prediction tools. The variants *PKD1*:c.1198C>T:p.R400X in Family 5, *PKD1*:c.8095C>T:p.Q2699X in Family 6, *PKD2*:c.1390C>T:p.R464X in Family 7, *PKD2*:c.916C>T:p.R306X in Family 8, and *PKD2*:c.1774C>T:p.R592X in Family 9 are nonsense variants, which cannot be predicted by SIFT, Polyphen2_HVAR and MutationAssessor. According to MutationTaster, all the three nonsense variants were disease causing automatic variants.

PhyloP and PhastCons were used to predict the conservational properties of those variants (Table 3). To assess the conservational property of amino acid residues of the five missense variants, alignment of human wild-type protein and other animal uromodulin homologs (rhesus, mouse, dog, elephant, chicken, X-tropicalis, and zebra fish) was performed using UCSC (Figure 2b,d,f,h,j). The results showed that *PKD1*:c.1216T>C:p.C406R in Family 2, *PKD1*:c.8548T>C:p.S2850P in Family 3, *PKD1*:c.8095C>T:p.Q2699X in Family 6, *PKD2*:c.1390C>T:p.R464X in Family 7, *PKD2*:c.916C>T:p.R306X in Family 8 and *PKD2*:c.1774C>T:p.R592X in Family 9, *PKD1*:c.9884A>G:p.N3295S in Family 9 and *PKD1*:c.11333C>A:p.T3778N in Family 10 were conserved, while *PKD1*:c.6930delG:p.C2311Vfs*3 in Family 1, *PKD1*:c.3865G>A:p.V1289M in Family 4 and *PKD1*:c.1198C>T:p.R400X in Family 5 were in a low conservational property.

In conclusion, the missense variants *PKD1*:c.1216T>C:p.C406R in Family 2 and *PKD1*:c.8548T>C:p.S2850P in Family 3 showed high degree of pathogenic properties. The missense variants *PKD1*:c.3865G>A:p.V1289M in Family 4, *PKD1*:c.9884A>G:p.N3295S in Family 9 and *PKD1*:c.11333C>A:p.T3778N in Family 10 showed middle degree of pathogenic properties. The conservational properties prediction results showed except *PKD1*:c.6930delG:p.C2311Vfs*3 in Family 1, *PKD1*:c.3865G>A:p.V1289M in Family 4, and *PKD1*:c.1198C>T:p.R400X in Family 5, the other eight variant sites were conserved.

3.4 | Interpretation of the variants according to the ACMG

To interpret the studied variants, the standards and guidelines from ACMG were adopted. The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls (PS4). Variants are sited in mutational hotspots (Dong et al., 2019), in which pathogenic variants in one or several nearby residues have been observed with greater frequency (PM1). Variants that are absent from gnomAD, 1000 Genome Program and ExAC databases are considered follow a moderate piece of evidence for pathogenicity (PM2). The in-frame deletion variant and nonsense are likely to disrupt protein function due to length changes in the protein and considered moderate evidence of pathogenicity (PM4). The patient has a family history of disease that is consistent with de novo inheritance without parental samples testing and the phenotype in the patient matches the gene's disease association with reasonable specificity (PM6). Those variants only occurred in ADPKD patients not in normal individual of the affected family, which obeys the co-segregation rule (PP1). The prediction results of several common prediction tools SIFT, Polyphen2_HVAR, MutationTaster, MutationAssessor, PhyloP, and PhastCons supported the pathogenicity of the missense variants (PP3). The phenotype and family history of the variant carriers are highly consistent with ADPKD (PP4). The variants have been classified as pathogenic in PKDB database (PP5). Variants both occur in ADPKD patient and normal individual of the family, which against the co-segregation principle (BS4). Multiple lines of computational evidence suggest no impact on gene or gene product (BP4). The missense variant was reported as benign (BP6). The interpretation results of the studied variants were listed in Table 4.

4 | DISCUSSION

In this study, 12 unrelated Chinese families with ADPKD were studied. Eight candidate disease causing variants in *PKD1* and three candidate disease causing variants in *PKD2* were detected in 10 pedigrees. The 11 variants were considered as the candidate pathogenic mutation sites. Among them, deletion variant *PKD1*:c.6930delG:p.C2311Vfs*3 in Family 1, missense variant *PKD1*:c.1216T>C:p.C406R in Family 2, missense variant *PKD1*:c.8548T>C:p.S2850P in Family 3, and missense variant *PKD1*:c.3865G>A:p.

TABLE 4 Interpretation of the variants according to the ACMG

Variant	Gene	Variant type	Pedigree	Interpretation	Conclusion
c.6930delG:p.C2311Vfs*3	*PKD1*	deletion	Family 1	[a]PS4+PM1[b]+PM2[c]+PM4[d]+PM6[e]+PP1[f]+PP4[g]	Pathogenic
c.1198C>T:p.R400X	*PKD1*	nonsense	Family 5	PS4+PM1+PM2+PM4+PP1+PP4+PP5[h]	Pathogenic
c.8095C>T:p.Q2699X	*PKD1*	nonsense	Family 6	PS4+PM1+PM2+PM4+PP4+PP5	Pathogenic
c.1390C>T:p.R464X	*PKD2*	nonsense	Family 7	PS4+PM1+PM2+PM4+PP1+PP4+PP5	Pathogenic
c.916C>T:p.R306X	*PKD2*	nonsense	Family 8	PS4+PM1+PM2+PM4+PP1+PP4+PP5	Pathogenic
c.1774C>T:p.R592X	*PKD2*	nonsense	Family 9	PS4+PM1+PM2+PM4+PP1+PP4+PP5	Pathogenic
c.1216T>C:p.C406R	*PKD1*	missense	Family 2	PS4+PM1+PM2+PM6+PP3[i]+PP4	Pathogenic
c.8548T>C:p.S2850P	*PKD1*	missense	Family 3	PS4+PM1+PM2+PM6+PP3+PP4	Pathogenic
c.9884A>G:p.N3295S	*PKD1*	missense	Family 9	PS4+PM1+PM2+PP1+PP4	Pathogenic
c.3865G>A:p.V1289M	*PKD1*	missense	Family 4	PS4+PM2+PM6+PP4	Likely pathogenic
c.11333C>A:p.T3778N	*PKD1*	missense	Family 10	PS4+PM1+PM2+PP4+BS4[j]+BP4[k] +BP6[l]	VUS[m]

Note:: The phenotype in the patient matches the gene's disease association with reasonable specificity.

Note: Pathogenicity classification has been made from a reputable source.

[a]The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls.

[b]The variant is located in mutational hotspots in which pathogenic variants are with greater frequency.

[c]The frequency of the variant is less than 0.01 or absent from gnomAD, 1000 Genome Program and ExAC databases.

[d]Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants.

[e]Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease.

[f]Patient's phenotype or family history is highly specific for a disease with a single genetic etiology.

[g]De novo variant detected in the patient not in other individuals of the family.

[h]Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.).

[i]Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.).

[j]Lack of segregation in affected members of a family.

[k]In silico predictions agree the variant as benign.

[l]Reputable source reports a variant as benign.

[m]Variant of undetermined significance.

V1289M in Family 4 are novel variants which have not been reported before.

Generally, deletion, insertion, or nonsense variants are considered as definitely pathogenic mutations. Thus, identified deletion variant *PKD1*:c.6930delG:p.C2311Vfs*3 in Family 1, nonsense variants *PKD1*:c.1198C>T:p.R400X (Peters, Ariyurek, van Dijk, & Breuning, 2001) in Family 5, *PKD1*:c.8095C>T:p.Q2699X (Tan et al., 2009) in Family 6, *PKD2*:c.1390C>T:p.R464X (Audrezet et al., 2012) in Family 7, *PKD2*:c.916C>T:p.R306X (Garcia-Gonzalez et al., 2007; Magistroni et al., 2003; Rossetti et al., 2012) in Family 8 and *PKD2*:c.1774C>T:p.R592X (Audrezet et al., 2012) in Family 9 were the ADPKD causing variants. However, missense variants are usually difficult to be identified. To confirm the disease causing property of the missense variants *PKD1*:c.1216T>C:p.C406R in Family 2, *PKD1*:c.8548T>C:p.S2850P in Family 3, *PKD1*:c.3865G>A:p.V1289M in Family 4, *PKD1*:c.9884A>G:p.N3295S (Yu et al., 2011) in Family 9 and *PKD1*:c.11333C>A:p.T3778N (Rossetti et al., 2001) in Family 10, the pathogenic and conservational properties were assessed. The variant type distribution was shown in Figure 3. In addition, because of the co-segregation rule and previous report (Rossetti et al., 2001), *PKD1*:c.11333C>A:p.T3778N is not responsible for Family 10. Among the seven reported variants, only the missense variant *PKD1*:c.9884A>G:p.N3295S was detected in Chinese while others were in European.

For Family 11 and Family 12, besides *PKD1* and *PKD2*, the newly reported pathogenic gene of ADPKD, *GANAB* (Porath et al., 2016), is also analyzed for disease causing variant, while no variant was detected. We also checked the variants of the pathogenic genes of other genetic cystic kidney diseases such as tuberous sclerosis complex (TSC), von Hippel–Lindau (VHL) disease or autosomal recessive polycystic kidney disease (ARPKD), although the patients did not show other obvious clinical symptoms or did not show the inheritance mode of autosomal recessive. The pathogenic genes of tuberous sclerosis complex are *TSC1* and *TSC2*, of von Hippel–Lindau disease are *VHL* and *CCND1*, and of autosomal recessive polycystic kidney disease is *PKHD1* listed in OMIM. We did not find any variant in *TSC1*, *TSC2*, *VHL*, *CCND1*, or *PKHD1* genes in the WES results of the patients in Family 11 and Family 12. Therefore, to identify the pathogenic gene, it needs more analysis of the families.

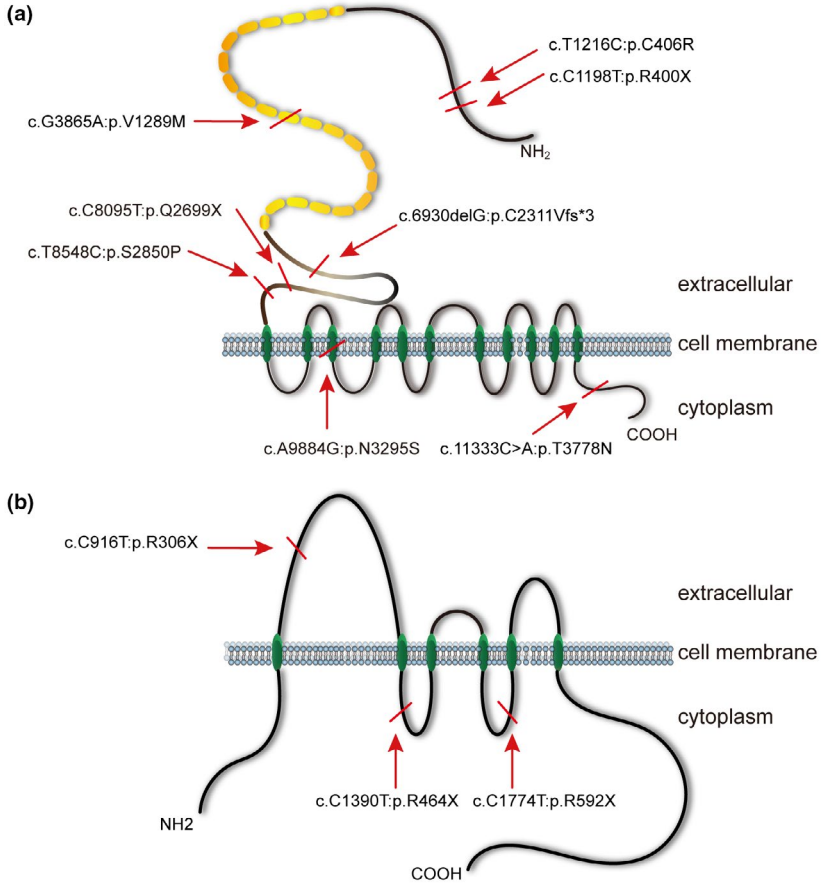

FIGURE 3 The distribution of detected variants in polycystin-1 and polycystin-2. (a) Diagram of polycystin-1. (b) Diagram of polycystin-2. Arrows indicate the positions of the detected variants

Totally, we identified seven pathogenic variants, three likely pathogenic variants and one variant of undetermined significance which need to be further studied. Nine out of 12 (75%) of the ADPKD pedigrees were detected disease causing variants. According to report, *PKD1* mutations comprise about 78% of ADPKD cases, while *PKD2* mutations comprise about 14% of the cases. The remaining cases have no identifiable mutations (Cornec-Le Gall, 2019).

Several, rather than all, classical methods were used in this study to assess the pathogenic properties of missense mutations. In Family 3, only the proband was sequenced because informed consent was not obtained from other individuals in the family. The assessment of more samples could better explain this phenomenon. The pathogenic gene for ADPKD in Family 10–12 has not been identified, and needs further studied. The pathogenic mechanism of the variants as well as the relationship with clinical manifestations should be further studied.

ACKNOWLEDGMENT

The authors thank all the individuals for their participation in this study. This work was supported by the National Key Research and Development Program (grant number 2016YFC1000504).

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

AUTHOR CONTRIBUTIONS

Kexian Dong, Xueyuan Jia, Songbin Fu, Xianli Zhou, and Wenjing Sun conceived and designed the study; Xiaogang Liu, Huanhuan Miao, Wei Ji, Jie Wu, Yun Huang, Lidan Xu, Xuelong Zhang, and Hui Su collected the clinical information; Kexian Dong, Rongwei Guan, Jing Bai, Guohua Ji, and Peng Liu performed the experiments; Kexian Dong, Xueyuan Jia, Songbin Fu, Xianli Zhou, and Wenjing Sun analyzed the data; Kexian Dong, Songbin Fu, and Wenjing Sun prepared the manuscript. All authors read and approved the final manuscript.

ORCID

Songbin Fu https://orcid.org/0000-0003-0382-9888
Wenjing Sun https://orcid.org/0000-0001-5638-5892

REFERENCES

Audrézet, M.-P., Cornec-Le Gall, E., Chen, J.-M., Redon, S., Quéré, I., Creff, J., … Férec, C. (2012). Autosomal dominant polycystic kidney disease: comprehensive mutation analysis of PKD1 and PKD2 in 700 unrelated patients. *Human Mutation*, *33*, 1239–1250.

Casteleijn, N. F., Spithoven, E. M., Rookmaaker, M. B., Vergouwen, M. D., & Gansevoort, R. T. (2015). Bilateral cysts in the choroid plexus in a patient with autosomal dominant polycystic kidney disease. *Nephrology, Dialysis, Transplantation*, *30*, 859–860.

Cornec-Le Gall, E., Alam, A., & Perrone, R. D. (2019). Autosomal dominant polycystic kidney disease. *Lancet*, *393*, 919–935.

Cornec-Le Gall, E., Torres, V. E., & Harris, P. C. (2018). Genetic Complexity of Autosomal Dominant Polycystic Kidney and Liver Diseases. *Journal of the American Society of Nephrology*, *29*, 13–23.

Dong, K., Miao, H., Jia, X., Wu, J., Wu, H., Sun, J., … Fu, S. (2019). Identification of a pathogenic mutation in a Chinese pedigree with polycystic kidney disease. *Molecular Medicine Reports*, *19*, 2671–2679.

Edrees, B. M., Athar, M., Abduljaleel, Z., Al-Allaf, F. A., Taher, M. M., Khan, W., … Ahmed, M. A. (2016). Functional alterations due to amino acid changes and evolutionary comparative analysis of ARPKD and ADPKD genes. *Genom Data*, *10*, 127–134.

Flahault, A., & Joly, D. (2019). Screening for intracranial aneurysms in patients with autosomal dominant polycystic kidney disease. *Clinical Journal of the American Society of Nephrology*, *14*(8), 1242–1244. https://doi.org/10.2215/CJN.02100219

Garcia-Gonzalez, M. A., Jones, J. G., Allen, S. K., Palatucci, C. M., Batish, S. D., Seltzer, W. K., … Watnick, T. J. (2007). Evaluating the clinical utility of a molecular genetic test for polycystic kidney disease. *Molecular Genetics and Metabolism*, *92*, 160–167.

Ghata, J., Cowley, B. D., Jr., (2017). Polycystic kidney disease. *Comprehensive Physiology*, *7*, 945–975.

Hafer, A. S., & Conran, R. M. (2017). Autosomal recessive polycystic kidney disease. *Academic Pathology*, *4*, 2374289517718560.

Kim, D. Y., & Park, J. H. (2016). Genetic mechanisms of ADPKD. *Advances in Experimental Medicine and Biology*, *933*, 13–22.

Kim, J. A., Blumenfeld, J. D., Chhabra, S., Dutruel, S. P., Thimmappa, N. D., Bobb, W. O., … Prince, M. R. (2016). Pancreatic cysts in autosomal dominant polycystic kidney disease: prevalence and association with PKD2 gene mutations. *Radiology*, *280*, 762–770.

Kinoshita, M., Higashihara, E., Kawano, H., Higashiyama, R., Koga, D., Fukui, T., … Nutahara, K. (2016). Technical evaluation: Identification of pathogenic mutations in PKD1 and PKD2 in patients with autosomal dominant polycystic kidney disease by next-generation sequencing and use of a comprehensive new classification system. *PLoS One*, *11*, e0166288.

Krishnappa, V., Vinod, P., Deverakonda, D., & Raina, R. (2017). Autosomal dominant polycystic kidney disease and the heart and brain. *Cleveland Clinic Journal of Medicine*, *84*, 471–481.

Lanktree, M. B., Haghighi, A., Guiard, E., Iliuta, I. A., Song, X., Harris, P. C., … Pei, Y. (2018). Prevalence estimates of polycystic kidney and liver disease by population sequencing. *Journal of the American Society of Nephrology*, *29*, 2593–2600.

Li, L., Szeto, C. C., Kwan, B. C., Chow, K. M., Leung, C. B., & Kam-Tao Li, P. (2011). Peritoneal dialysis as the first-line renal replacement therapy in patients with autosomal dominant polycystic kidney disease. *American Journal of Kidney Diseases*, *57*, 903–907.

Litvinchuk, T., Tao, Y., Singh, R., & Vasylyeva, T. L. (2015). A case of new familiar genetic variant of autosomal dominant polycystic kidney disease-2: A case study. *Frontiers in Pediatrics*, *3*, 82.

Magistroni, R., He, N., Wang, K., Andrew, R., Johnson, A., Gabow, P., … Pei, Y. (2003). Genotype-renal function correlation in type 2 autosomal dominant polycystic kidney disease. *Journal of the American Society of Nephrology*, *14*, 1164–1174.

Meng, J., Xu, Y., Li, A., Fan, S., Shen, X., Ma, D., … Liang, C. (2018). Clinical features of 167 inpatients with autosomal dominant polycystic kidney disease at a single center in China. *Medical Science Monitor*, *24*, 6498–6505.

Obeidova, L., Elisakova, V., Stekrova, J., Reiterova, J., Merta, M., Tesar, V., … Kohoutova, M. (2014). Novel mutations of PKD genes in the Czech population with autosomal dominant polycystic kidney disease. *BMC Medical Genetics*, *15*, 41.

Pei, Y., Hwang, Y. H., Conklin, J., Sundsbak, J. L., Heyer, C. M., Chan, W., … Haider, M. A. (2015). Imaging-based diagnosis of autosomal dominant polycystic kidney disease. *Journal of the American Society of Nephrology*, *26*, 746–753.

Peters, D. J., Ariyurek, Y., van Dijk, M., & Breuning, M. H. (2001). Mutation detection for exons 2 to 10 of the polycystic kidney disease 1 (PKD1)-gene by DGGE. *European Journal of Human Genetics*, *9*, 957–960.

Porath, B., Gainullin, V. G., Cornec-Le Gall, E., Dillinger, E. K., Heyer, C. M., Hopp, K., … Harris, P. C. (2016). Mutations in GANAB, encoding the glucosidase iialpha subunit, cause autosomal-dominant polycystic kidney and liver disease. *American Journal of Human Genetics*, *98*, 1193–1207.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … 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. *Genetics in Medicine*, *17*, 405–424.

Rossetti, S., Hopp, K., Sikkink, R. A., Sundsbak, J. L., Lee, Y. K., Kubly, V., … Harris, P. C. (2012). Identification of gene mutations in autosomal dominant polycystic kidney disease through targeted resequencing. *Journal of the American Society of Nephrology*, *23*, 915–933.

Rossetti, S., Strmecki, L., Gamble, V., Burton, S., Sneddon, V., Peral, B., … Harris, P. C. (2001). Mutation analysis of the entire PKD1 gene: genetic and diagnostic implications. *American Journal of Human Genetics*, *68*, 46–63.

Tan, Y. C., Blumenfeld, J. D., Anghel, R., Donahue, S., Belenkaya, R., Balina, M., … Rennert, H. (2009). Novel method for genomic analysis of PKD1 and PKD2 mutations in autosomal dominant polycystic kidney disease. *Human Mutation*, *30*, 264–273.

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. *Nephrology, Dialysis, Transplantation*, *32*, 1356–1363.

Yu, C., Yang, Y., Zou, L., Hu, Z., Li, J., Liu, Y., … Zhang, S. (2011). Identification of novel mutations in Chinese Hans with autosomal dominant polycystic kidney disease. *BMC Medical Genetics*, *12*, 164.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Dong K, Liu X, Jia X, Miao H, Jei W. Disease causing property analyzation of variants in 12 Chinese families with polycystic kidney disease. *Mol Genet Genomic Med*. 2020;8:e1467. https://doi.org/10.1002/mgg3.1467