Co-segregation of candidate polymorphism rs201204878 of the PKD1 gene in a large Iranian family with autosomal dominant polycystic disease

FARIBA RANJZAD1, AHMAD TARA1, ABBAS BASIRI1, NASSER AGHDAMI2,3 and REZA MOGHADASALI2,3

1Urology and Nephrology Research Center, Shahid Beheshti University of Medical Sciences, Tehran-1666777951; Departments of 2Stem Cells and Developmental Biology, and 3Regenerative Biomedicine, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran-8158968433, Islamic Republic of Iran

Received February 21, 2017; Accepted September 1, 2018

DOI: 10.3892/etm.2019.7693

Abstract. Autosomal dominant polycystic kidney disease (ADPKD) is the fourth most common cause of end-stage renal disease, occurring at a frequency of 1 in 400 to 1 in 800 individuals among different populations. The disease affects all ethnic groups worldwide, and there is a requirement for population-based studies to be conducted in order to improve diagnosis, genetic counseling and treatment. A large Iranian family with ADPKD was recruited for the current study. Clinical evaluation was performed to diagnose and assess disease progression in 11 members of this family, including 7 affected members and 4 unaffected members. PKD1 and PKD2 genes were genotyped in subjects by next-generation sequencing (NGS). Mutational analysis of PKD1 and PKD2 genes in this family revealed three intronic variations and three synonymous exonic variants in the PKD2 gene, and two non-synonymous exonic variants and eight intronic variants in PKD1, resulting in a total of 16 heterozygous variations among these two genes. Among the 16 variations, all except three intronic variants in the PKD1 gene have already reported in the Iranian population. The three novel mutations predicted to be deleterious polymorphisms using in silico methods. Among the reported intronic variations, rs201204878 was identified as a splice region variant, leading to truncation of the polycystin-1 protein. In conclusion, genotyping of PKD1 and PKD2 in this family with ADPKD revealed no mutational hot spots. However, genetic screening identified three novel variants in the Iranian population. The data generated in the present study will contribute to improving the diagnosis, genetic counseling and treatment of patients with ADPKD.

Introduction

Autosomal dominant polycystic kidney disease (ADPKD) is the most common hereditary kidney disorder found among all ethnic groups worldwide, affecting 1 in 400 to 1 in 800 live births among different populations (1,2). The disease is characterized by the development and progressive enlargement of renal cysts and is the fourth most common cause leading to end-stage renal disease (ESRD) (2,3). The etiology of this disease is associated with mutations in PKD1 [Online Mendelian Inheritance in Man (OMIM) no. 601313; 16p13.3] and PKD2 (OMIM no. 173910; 4q21-22) genes, and 85-90% of ADPKD cases are caused primarily by mutations in the PKD1 gene (4). PKD1 has 46 exons and encodes the transmembrane protein, polycystin-1 (PC1), which is composed of 4,303 amino acids. PKD2 is a smaller gene with 15 exons. It encodes the transmembrane protein polycystin-2 (PC2), which is composed of 968 amino acids (5,6). PC1 contains an N-terminal extracellular region, 11 membrane-spanning domains and a cytoplasmic C-tail, whereas the PC2 protein possesses a shorter N-terminal extracellular region and only 6 membrane-spanning domains (7,8).

The diagnosis of ADPKD is a challenge; renal ultrasound, magnetic resonance imaging and computed tomography have limited capability in terms of detecting ADPKD in patients (9). With no hotspot mutation information available for PKD1 and PKD2, clinical molecular diagnostic techniques are difficult (10). To overcome these challenges, the genotype variation of these two genes in patients with ADPKD has been investigated using techniques including polymerase chain reaction (PCR)-single-strand conformation polymorphism, denaturing high-performance liquid chromatography, multiplex ligation-dependent probe amplification and next-generation sequencing (NGS) (11-14). Recent advancements in sequencing technologies have enabled the rapid and cost-effective generation of large quantities of data. By removing most of the throughput and resource limitations of traditional methods, NGS enables investigators to analyze PKD1 and PKD2 genes in a single run (15-17).
In individuals diagnosed with ADPKD, *PKD1* exhibits marked allelic heterogeneity, with a high level of gene variation (18,19). To date, >2,300 germline mutations of *PKD1* have been identified; ~900 are likely to be neutral polymorphisms and >150 are unclassified variants (20). In the current study, the pathogenicity of an atypical splice variant in the *PKD1* gene, rs201204878, was evaluated in an Iranian family with ADPKD. *In silico* prediction-based models were used to evaluate its pathogenicity.

**Materials and methods**

*Ethical compliance.* The study was approved by the ethics committee board of Shahid Beheshti University of Medical Sciences (Tehran, Iran) and all experiments were performed in adherence to the declaration of Helsinki. Written informed consent was obtained from all patients prior to their participation in the study.

*Patient pedigree and subject data.* A large Iranian family of 19 members (10 males, nine females; age, 40.8±15.7 years) with a history of ADPKD was recruited for this study from dialysis centers and nephrology clinics of Modarres Hospital (Tehran, Iran) in June 2015. In the pedigree, 11 members were recruited; seven of them were diagnosed with ADPKD by ultrasound examination, according to Ravine's criteria (21). The proband selected was diagnosed with ESRD, and had two affected brothers and two affected sisters with early onset ESRD (age of onset, <50 years). A screening ultrasound was performed on 11 available participants of the asymptomatic at-risk adult members in this family. Details of the affected and non-affected members of the family (Four generations, 19 members in total) are shown in Fig. 1. Total genomic DNA from 11 available family members was extracted from blood samples using a standard phenol-chloroform procedure as previously described (22). The quality and concentration of DNA samples was evaluated using spectrophotometry (260/280 nm) and 1% agarose gel electrophoresis.

*Mutational analysis by targeted NGS.* Targeted NGS experiments were performed by BGI Tech Europe (Copenhagen, Denmark). The Genetic Sequencing Test was performed using a custom-designed NimbleGen chip (Roche NimbleGen, Inc., Madison, WI, USA) by BGI Tech Europe to capture the genes of interest, which were then sequenced by NGS. In general, the test platform examined >95% of the target gene with a sensitivity of >99%. Point mutations, micro-insertions, deletions and duplications (≤20 bp) were detected simultaneously (14). SNPs and indels were identified via the SOAP snp software (Release 1.03; http://soap.genomics.org.cn/soapspn.html) and GATK Indel Genotyper (Version 4.0.11; broadinstitute.org/gsa/wiki/index.php/).

*Bioinformatics analysis and mutation identification.* Candidate variations were compared against electronic database information and computational analysis was performed. The detected sequence variations were compared against the currently published list of PKD gene variants in the Human Gene Mutation Database (23), the Autosomal Dominant Polycystic Kidney Disease Mutation Database version 3.1 (20) and the Iranian Human Mutation Gene Bank (24). The effects of variation were analyzed using web-based computational pathogenicity prediction tools, including MutationTaster working on current build of NCBI 137/Ensembl 69 (25), Sorting Intolerant from Tolerant based on NCBI 137/Ensembl 66 (26) and Polymorphism Phenotyping version 2 (27). Human Splice Finder (HSF) software version 3.0 was also used to predict splicing (28).

*Segregation analysis.* To confirm the association between mutation and the pathogenicity of the disease, analysis of seven affected family members in the pedigree was performed by direct sequencing (Sanger method as previously described). The Iranian normal population database contains data of the reported variants. Four unaffected family members and four normal controls from the Iranian normal population database were also checked for the mutation to confirm the prediction.

**Results**

*Mutational analysis.* Mutational analysis of the *PKD1* and *PKD2* genes was performed in a large Iranian family consisting of 11 members (5 male and 6 females); 7 of them (3 male and 4 female) diagnosed with ADPKD by renal ultrasound. Using targeted NGS, three intronic variations and three synonymous exonic variants were identified in the *PKD2* gene, and two non-synonymous exonic variants and eight intronic variants were identified in the *PKD1* gene (Table II). All variants except from three intronic variants in *PKD1* have been reported in the National Centre for Biotechnology Information dbSNP database (build 151; ncbi.nlm.nih.gov/SNP/) and are considered to be known polymorphisms. The three novel mutations were predicted to be deleterious polymorphisms by bioinformatics analysis.

*Analysis of splice site mutations.* Among the three reported non-coding RNA (ncRNA) intronic variations, one (rs201204878) was identified as a splice region variant (c. 11537+5_6 ins ccc). The pathogenic relevance of putative splice site mutations was
evaluated using HSF software. The HSF results indicated that this mutation occurred in the late exonic positions and generated a new exonic splice site that potentially altered splicing. Therefore this mutation was classified as a highly likely pathogenic mutation (30). This variant was assessed for its co-segregation with the disease in three affected and two unaffected family members. The mutation was identified in all affected family members, and was absent in unaffected family members, as well as in 400 normal subjects in the Iranian population (Fig. 2).

**Segregation analysis.** In the present study, the allele frequency of the c. 11537+5_6 ins ccc variation in world population groups from the 1,000 Genomes Project was analyzed using the Ensembl database (internationalgenome.org/category/ensembl/). The frequency of this mutation in sub-populations from America, Africa and East Asia was 0%. The frequency of this mutation in European and South Asian sub-populations was ~2 and ~3%, respectively. However, the homozygous genotype frequency in these sub-populations was 0%. This appears to be consistent with the experimental results of the current study, and may be associated with the truncation effect of this variant and the subsequent abortion of embryos carrying the homozygous mutation in mice models (31). The co-segregation analysis and HSF of rs201204878 showed no co-relation with the severity of ADPKD in the affected family members.

**Discussion**

According to the genotype-phenotype correlation, the ADPKD phenotype exhibits both genetic and allelic dependency. *PKD1* mutations are correlated with disease severity and the onset of ESRD in ADPKD patients (32). The c. 11537+5_6 ins ccc splice mutation identified in the present study is a type of truncating mutation, which has been associated with a more severe ADPKD phenotype when compared with non-truncating *PKD1* mutations (4). To the best of our knowledge, the present study is the first to report the rs201204878 pathogenic variant in the Iranian population, as this variation was not observed in 400 subjects from the Iranian normal population database. The clinical severity of ADPKD in affected family members was consistent with the obtained splicing mutation results from HSF and co-segregation analyses. Hence, follow-up in other young affected family members becomes necessary to verify the reported association. The average age of ESRD onset varies with the type of gene involved, and it was reported to be 54.3 years in patients with PKD1 mutations and 74 years in patients with PKD2 mutations (33).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Genomic position</th>
<th>RefSeq</th>
<th>Nucleic acid alteration</th>
<th>Mutation location</th>
<th>SNPID</th>
</tr>
</thead>
<tbody>
<tr>
<td>PKD2</td>
<td>88,957,562</td>
<td>NM_000297</td>
<td>T/C</td>
<td>Intronic</td>
<td>rs17786456</td>
</tr>
<tr>
<td>PKD2</td>
<td>88,959,381</td>
<td>NM_000297</td>
<td>G/A</td>
<td>Intronic</td>
<td>rs2725221</td>
</tr>
<tr>
<td>PKD2</td>
<td>88,959,745</td>
<td>NM_000297</td>
<td>A/G</td>
<td>Intronic</td>
<td>rs17013735</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,140,294</td>
<td>NM_001009944</td>
<td>C/T</td>
<td>Exonic</td>
<td>rs148478410</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,141,714</td>
<td>NM_001009944</td>
<td>G/A</td>
<td>ncRNA_intronic</td>
<td>Novel</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,141,776</td>
<td>NM_001009944</td>
<td>-/CCC</td>
<td>ncRNA_intronic</td>
<td>rs201204878</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,142,113</td>
<td>NM_001009944</td>
<td>G/A</td>
<td>Exonic</td>
<td>rs145955373</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,152,574</td>
<td>NM_001009944</td>
<td>C/T</td>
<td>Exonic</td>
<td>rs374619113</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,154,478</td>
<td>NM_001009944</td>
<td>A/G</td>
<td>Intronic</td>
<td>rs4786209</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,156,369</td>
<td>NM_001009944</td>
<td>G/A</td>
<td>Intronic</td>
<td>rs142761413</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,158,176</td>
<td>NM_001009944</td>
<td>T/A</td>
<td>Intronic</td>
<td>rs200363107</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,163,115</td>
<td>NM_001009944</td>
<td>G/C</td>
<td>ncRNA_intronic</td>
<td>Novel</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,163,579</td>
<td>NM_001009944</td>
<td>G/C</td>
<td>Intronic</td>
<td>rs56326527</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,164,808</td>
<td>NM_001009944</td>
<td>C/T</td>
<td>Exonic</td>
<td>rs40433</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,166,769</td>
<td>NM_001009944</td>
<td>G/A</td>
<td>Intronic</td>
<td>Novel</td>
</tr>
<tr>
<td>PKDI</td>
<td>2,167,874</td>
<td>NM_001009944</td>
<td>G/A</td>
<td>Exonic</td>
<td>rs35842</td>
</tr>
</tbody>
</table>

SNP, single nucleotide polymorphism; ncRNA, non-coding RNA.
In the present study, the heterozygous 11537+5_6 ins ccc mutation (rs201204878) was identified in intron 41 of the PKD1 gene, and no other truncation mutations were identified by targeted NGS. This was demonstrated to be a pathogenic mutation, which extends the associated phenotypic and genotypic spectrums for ADPKD. The rs201204878 mutation has been previously reported to be benign (33), but in the present study it was predicted that the splice mutation may lead to abnormal splicing of PKD1. This is expected to affect PC1 function, which is consistent with a previous report demonstrating that the expression of the truncated protein is correlated with the early onset of ESRD (34).

In the current study, targeted NGS of the PKD1 and PKD2 genes produced high-coverage sequencing data with high sensitivity and specificity. The results were confirmed with Sanger sequencing, which is the standard approach for clinical genetic testing in ADPKD. The results suggest that targeted NGS may potentially replace Sanger sequencing for clinical genetic testing in ADPKD as the former is a faster and more accurate procedure In silico and co-segregation analyses concluded that the rs201204878 variant may be considered as a potential functional mutation. However, further mini gene analysis may provide further insight into the pathogenicity of this mutation. In addition, as the Iranian population is not included in the 1,000 Genomes Project (35), performing genetic variation studies will add to the plethora of data available on mutations; such studies on ethnic populations may aid in the diagnosis, prognosis and management of this disease. In conclusion, the polymorphisms identified in the present study may contribute to improving the diagnosis, genetic counseling and treatment of patients with ADPKD.

Acknowledgements
The authors want to thank the Urology and Nephrology Research Center (UNRC) of Shahid Beheshti University of Medical Sciences for their support.

Funding
This trial was supported by a grant from Royan Institute and the Royan Charity Association for Health Research (grant no. 93/566).

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions
AB conceived of the present study. FR and AT performed the experiments. AB, NA and RM analysed the results. FR and AB wrote the manuscript. NA and RM edited the manuscript.

Ethics approval and consent to participate
The study was approved by the ethics committee board of Shahid Beheshti University of Medical Sciences (Tehran, Iran) and all experiments were performed in adherence to the declaration of Helsinki. Written informed consent was obtained from all patients prior to their participation in the study.
Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


