Molecular Genetic Analysis of PKHD1 Mutations in Pedigrees With Autosomal Recessive Polycystic Kidney Disease

Fatemeh Bitarafan, Masoud Garshasbi

Introduction. A wide variety of mutations are spread throughout the PKHD1 gene, which encodes a 4074 amino acid protein, namely fibrocystin/polyductin, and is responsible for all features of autosomal recessive polycystic kidney disease (ARPKD). Autosomal recessive polycystic kidney disease is a hereditary early-onset form of polycystic kidney disease characterized by fusiform dilation of collecting ducts and congenital hepatic fibrosis. The highest level of PKDH1 expression is in the kidneys of fetus and adults, suggesting the functionally importance of the gene in the mature kidney in addition to its role in kidney development.

Materials and Methods. Mutational analysis of the PKHD1 gene was performed in 11 families with a history of 1 to 6 fetuses or children affected by ARPKD, which either were aborted or died shortly after birth. Analyses were done using the Next Generation sequencing and Sanger sequencing techniques.

Results. Four novel mutations, including c.6469C > T, c.9218 G > A, c.10456T > C, and c.8863C > G, and 3 previously reported ones, including c.9524A > G, c.1095G > A, c.1123C > T, were identified.

Conclusions. In view of high consanguineous marriages in Iranian population, the frequency of disease is expected to be higher than the world average.

INTRODUCTION

With an estimated prevalence of 1 in 20,000 live births and a heterozygous carrier rate of 1 in 70,1,2 autosomal recessive polycystic kidney disease (ARPKD; MIM 263200) is characterized by typical presentation of greatly enlarged echogenic kidney that often leads to neonatal demise.3 Autosomal recessive polycystic kidney disease is an important cause of renal and liver-related morbidity and mortality in children,2 with the clinical spectrum at a varying severity ranges from stillbirths to mildly affected adults.4

The most important criteria of neonatal disease are bilateral renal and liver involvement consisting of nonobstructive fusiform dilatation of the renal collecting tubules, often leading to massive kidney enlargement and a ductal plate malformation of the liver, ultimately resulting in congenital hepatic fibrosis, and in some cases, focal dilatation of the larger intrahepatic bile ducts (Caroli disease), respectively.5 Enlarged echogenic kidneys and oligohydramnios in fetuses are the most severe manifestation of the disease.6 Prenatal death is observed in 25% to 30% of newborns caused by pulmonary hypoplasia. The signs are potentially detectable clinically in uterus often in the late third trimester of pregnancy.2

Dysfunction of a 4074 amino acid single-membrane spanning multidomain protein named fibrocystin/polyductin, which is localized to the
primary cilia with concentrations in the basal body, can cause increased apoptosis and fluid secretion due to abnormal epithelial differentiation that ultimately results in clinically detectable cyst formation. Fibrocystin/polyductin is highly expressed in the kidney, and at lower levels in the liver, in accordance to the clinical presentation of the disease. Weak expression has also been observed in other tissues such as the pancreas, liver, and arterial wall. The ARPKD causative gene, \textit{PKHD1}, is amongst the largest human genes, extending over a genomic segment of at least 470 kb, which maps to chromosome 6p21.1-p12. Its longest reading frame with 67 exons codes fibrocystin/polyductin protein, 447 kD, with a short intracellular carboxyl terminal and a large extracellular amino terminal. Numerous mutations, mostly unique to individual families, have been identified across the entire gene, the majority of which are missense. Frameshift, insertion, deletion, nonsense, and splice-site mutations have been also found. Genotype-phenotype correlation for all of these mutations is not well established. However, for example, it is known that a termination-type mutation in the homozygous state leads to death shortly after birth. The aim of this study was to investigate the presence of mutations in the \textit{PKHD1} gene in 11 families with clinical diagnosis of ARPKD. Six heterozygote mutations in both parents and 1 homozygote mutation in an aborted fetus were found.

**MATERIALS AND METHODS**

**Patients**

All individuals comprised 11 families with a diagnosis of ARPKD were referred to the DeNA laboratory, Department of Medical Genetics, Tehran, Iran. Informed consent was obtained from all families. All of the families had a history of 1 to 6 fetuses or children affected with ARPKD, which either were spontaneously aborted, aborted through termination of pregnancy because of fetal anomaly, or died shortly after birth. In 10 of 11 families, the couples with affected children were consanguineous. In one of these families, the DNA from one of the aborted fetuses were available, and therefore, it was used for mutation screening, but in the rest of 10 families, the mutation screening was performed on one of the parents. Pedigrees are shown in Figure 1.

**DNA Extraction**

Genomic DNAs were extracted from the peripheral blood of proband, all available family members and postmortem fetal tissue by a High Pure PCR template preparation kit (Roche; Product No, 11814770001). The sequencing was performed at the BGI Health Laboratories using a custom designed Nimblegen chip capturing the \textit{PKHD1} gene (NM_138694) followed by Next Generation Sequencing. In general, the test platform examined more than 95% of the target gene with a sensitivity greater than 99%. Point mutation, micro-insertion, deletion, and duplications ( < 20 bp) can be simultaneously detected. For analysis of the sequencing results the international publicly available mutation and polymorphism databases such as 1000 genome project, ExAC (Exome Aggregation Consortium and ESP (Exon Sequencing Projects), as well as BGI self-developed local database were employed. Only variants with a frequency below 0.01 were selected. Previous mutations that have been described in the Human Gene Mutation Database as pathogenic were given the highest priority. For mutations, prediction of the consequence of such mutations were obtained from at least 3 online databases namely SIFT, Polyphen2, and Mutation Taster. Additionally, ConSurf (http://www.consurf.tau.ac.il) was applied to check the evolutionary conservation in the region of the mutations (Figure 2).

**Validation of Mutations**

Polymerase chain reaction (PCR) analysis was carried out in a total volume of 25 µL containing 0.5 µL of each forward and reverse primers (Table 1), 10 µL of PCR Master mix magnesium chloride (1.5 Mm), and 1 µL of DNA (about 100 ng). The reaction was adjusted to the total volume of 25 µL by ddH2O. The PCR was performed using an initial denaturation step at 94°C for 5 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 61 for 30 seconds, and elongation at 72°C for 30 seconds. Products of PCR were examined by 2% agarose gel electrophoresis for the presence and sizes of amplicons. Consequently, DNA sequencing of the PCR products was performed on 3130 ABI capillary electrophoresis. Sequencing chromatograms were analyzed by using the Codoncode aligner software version 6.0.2.
RESULTS

Family 1
This family was a consanguineous couple with an affected infant boy born with very large cystic kidneys. He died shortly after birth. In 2 further pregnancies, ultrasonography showed very large cystic kidneys in the fetuses, and therefore, they have been therapeutically aborted. The DNA from the latter fetus was used for mutation screening. The sequencing of the total length of PKDH1 (16503 bp) was obtained with a coverage of 100%, average depth of 296.23X, and a minimum depth of 30X. A homozygote c.6469C > T; p.Gln2157Ter novel mutation was identified in the affected aborted fetus. This mutation was confirmed by Sanger sequencing (Figure 3), and also it was shown that both parents were carriers for this mutation.

Family 2
The family constituted of a consanguineous couple with a history of 1 child with typical presentation of enlarged echogenic kidneys, resulting in neonatal death. The DNA from the patient was unavailable, and therefore, DNA from the father was screened for mutations. The sequencing of the total length of PKDH1 (16503 bp) was obtained with a coverage of 100%, an average depth of 205.89X, and a minimum depth of 30X. A heterozygous c.9524A > G; p.Asn3175Ser mutation was found in the father. This mutation was confirmed by Sanger sequencing. The mother was also shown to be the carrier for this mutation (Figure 3). This mutation is reported in the Human Gene Mutation Database.

Family 3
This family constituted of a consanguineous couple whose first child died at birth with the ARPKD symptoms. As the DNA from affected infant was not available, PKHD1 mutation screening was performed for the father. The sequencing of the total length of PKDH1 (16503 bp) was obtained with a
coverage of 98.88%, an average depth of 246.19X, and a minimum depth of 30X. A heterozygous c.1095G > A; p.Trp365Ter mutation was found in the father. This mutation was confirmed by Sanger sequencing. The mother was also shown to be the carrier for this mutation (Figure 3). This mutation is reported in the Human Gene Mutation Database.

**Family 4**

In this family, there was a consanguineous couple with a history of a terminated gestation at the 32nd week of pregnancy due to the diagnosis of very large cystic kidneys in the fetus. The sequencing of the total length of PKDH1 (16503bp) was obtained with a coverage of 100%, average depth of 308.3X, and a minimum depth of 30X. A heterozygous c.1123C > T, p.Arg375Trp mutation was found in the father. This mutation was confirmed by Sanger sequencing. The mother was also shown to be the carrier for this mutation (Figure 3). This mutation is reported in the Human Gene Mutation Database.

**Family 5**

A second cousin couple with a history of 2 twin

---

**Table 1. Oligonucleotide Sequences Used for Polymerase Chain Reaction Amplification and Direct Sequencing**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.6496C&gt;T</td>
<td>F: 5’TCCAACAAAACACTTCTCTACCT3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’GTTTGGGCAATGATCAGCA3’</td>
</tr>
<tr>
<td>c.9524A&gt;G</td>
<td>F: 5’TGAATGCAGTCAAAGAAGAC3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’GCTTTGCTGACAATGTGG3’</td>
</tr>
<tr>
<td>c.1095G&gt;A</td>
<td>F: 5’TCTGTCTCTCTGCCTCACT3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’GGCAATCGAGGGCTTCTTT3’</td>
</tr>
<tr>
<td>c.1123C&gt;T</td>
<td>F: 5’CTCTCTGCTGGATGGGACTG3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’AAGGCCAGTCTATTGTTG3’</td>
</tr>
<tr>
<td>c.9218G&gt;A</td>
<td>F: 5’TGAATGCAGCAGACGTGTC3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’ATTATGCAACGGTCCAGTC3’</td>
</tr>
<tr>
<td>c.10456T&gt;C</td>
<td>F: 5’AAAGCTGACTGACCAGAGTGG3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’AAATGCGAATATTTCCCGGTC3’</td>
</tr>
<tr>
<td>c.8863C&gt;G</td>
<td>F: 5’TGAGAGGACCCCAACAC3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’AACTTCTGGAGGAAAGGAT3’</td>
</tr>
</tbody>
</table>
pregnancies. In the first gestation, both fetuses had polycystic kidney disease and they were terminated. In the second gestation, 1 fetus had been aborted due to polycystic kidney disease and the second fetus died because of preterm delivery. The sequencing of the total length of \( PKDH1 \) (16503 bp) was obtained with a coverage of 100\%, an average depth of 340.89X, and a minimum depth of 30X. No mutation in \( PKDH1 \) gene was found in this family.

**Family 6**

This family contained a consanguineous couple with a history of 6 abortions due to a possible diagnosis of polycystic kidney disease. The sequencing of the total length of \( PKDH1 \) (16503 bp) was obtained with a coverage of 100\%, an average depth of 480.28X, and a minimum depth of 30X. No mutation in the \( PKDH1 \) gene was found in this family.

**Family 7**

An apparently not consanguineous couple with a history of 3 pregnancies, in all of which the kidney did not form and amniotic fluid was reduced from the third week of gestation and eventually leading to spontaneous abortion in all 3 pregnancies. The sequencing of the total length of \( PKDH1 \) (16503 bp) was obtained with a coverage of 100\%, an average depth of 233.32X, and a minimum depth of 30X. No mutation in the \( PKDH1 \) gene was found in this family.

**Family 8**

In this family, there was a consanguineous couple with a history of 2 abortions due to the diagnosis of polycystic kidneys disease. Additionally, there was a history of a 4-month infantile death due to polycystic cystic kidney disease in this family. The sequencing of the total length of \( PKDH1 \) (16503 bp) was obtained with a coverage of 100\%, an average depth of 294.58X, and minimum depth of 30X. A heterozygous c.9524A > G, p.Asn317Ser mutation was found in the mother. This mutation was confirmed by Sanger sequencing. Father was also shown to be the carrier for this mutation (Figure 3). The mutation in this family was the same as the mutation in family 3.

**Family 9**

In this family, there was a consanguineous couple
with a history of 2 boys who died after birth due to the diagnosis of polycystic kidneys disease. The sequencing of the total length of PKDH1 (16503 bp) was obtained with a coverage of 100%, an average depth of 236.34X, and a minimum depth of 30X. A novel heterozygous c.9218G > A, p.Trp3073Ter mutation was found in the mother. This mutation was confirmed by Sanger sequencing, and the father was also shown to be the carrier for this mutation (Figure 3).

**Family 10**

In this family, there was a consanguineous couple with a history of 2 abortions of pregnancy and also having 3 children who died after birth. Another consanguineous couple in this family had a history of 3 children who died shortly after birth. The sequencing of the total length of PKDH1 (16503 bp) was obtained with a coverage of 100%, an average depth of 302.74X, and minimum depth of 30X. No mutation in PKDH1 was found in this family.

**Family 11**

In this family, there was a consanguineous couple with a history of a girl born at the 8th month of pregnancy and died shortly after birth due to a diagnosis of polycystic kidneys disease. The sequencing of the total length of PKDH1 (16503 bp) was obtained with a coverage of 100%, an average depth of 191.41X, and a minimum depth of 30X was obtained. Two novel heterozygous c.10456T > C, p.Leu3486Leu and c.8863C > G, p.Arg2955Gly mutations were found in the mother. Both mutations were confirmed by Sanger sequencing in the mother and father, but co-segregation study in the other members of this family showed homozgyosity for the c.8863C > G, p.Arg2955Gly mutation in 2 healthy individuals (Figure 3). Therefore, the c.10456T > C, p.Leu3486Leu was most probably the pathogenic mutation in this family.

The list of identified mutations is presented in Table 2. The in silico prediction with software SIFT, Polyphen2, and mutation taster are shown in Table 3.

**DISCUSSION**

Dispersed mutations over entire the PKHD1 gene are the most important cause of ARPKD, a significant cause of renal-related and liver-related morbidity and mortality in neonates and infants. It is possible that the severity of the phenotype could be dependent on the location of amino acid substitutions along the polyductin/fibrocystin protein.

With 67 exons, PKHD1 is a large gene. To date, 522 mutations of PKHD1 have been reported in ARPKD patients in the Human Gene Mutation Database. Available data suggest that most families have their own mutations and therefore variable phenotypes have been observed in ARPKD patients due to private mutations. Consequently, gene-based diagnostics remain difficult. However, detecting more mutations which segregate with the disease will improve gene-based diagnostics.

Although correlation between genotype and phenotype of PKHD1 has been studied, the relation of single missense mutations with phenotype is not clear yet. Missense mutations could cause a range of clinical manifestations from mild alterations to loss of function which is due to truncating mutations. In a study by Bergmann and colleagues, 50% of all mutations were found to be in 7 exons (3, 9, 32, 36, 57, 58, and 61). To achieve a rate of 80%, additional exons had to be included into analysis (5, 14, 16, 20, 21, 22, 30, 33, 34, 37, 39, 43, 50, 54, and 59). In parallel with these results, in another cohort, 40% of mutations were found to be in the same 7 exons, of which 32% were detected in exons 3, 32, 58, and 61, which are the largest exons. In this study on 39 mainly Dutch families, 11 nonsense mutations, 15 deletions/insertions, 5 splice site mutations, and 39 missense mutations were identified. For Dutch population, exon 27 increased the yield to 53% of mutations.

A total of 128 mutations all along the gene were detected by Denamur and colleagues; 55.5% of mutations were truncating including frameshifts, small deletions, insertions, duplications, nonsense and splice mutations with exon skipping, and 42.2% were missense mutations. One mutation was a missense mutation associated with an in-frame insertion of 6 bases in exon 16, one mutation was an inframe deletion of 42 bases in exon 30 and one mutation was a one-base deletion in exon 67 leading to the change of the last 12 amino acids of the protein. However, 3 mutations were recurrent: the T36M missense in exon 3, the frameshifts c.5895dupA (L1966fs) in exon 36, and...
<table>
<thead>
<tr>
<th>Family</th>
<th>Members</th>
<th>Gene Refseq</th>
<th>Nucleic Acid Alteration</th>
<th>Amino Acid Alteration</th>
<th>Mutation Location</th>
<th>Zygosity</th>
<th>Chromosome Location</th>
<th>Method</th>
<th>Consanguinity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 2</td>
<td>Father and mother</td>
<td>PKHD1 NM_138694</td>
<td>c.9524A&gt;G</td>
<td>p.Asn3175Ser</td>
<td>EX58/CDS57</td>
<td>Heterozygous</td>
<td>Chr6:51612890</td>
<td>Father: NGS, Mother: Sanger Sequencing</td>
<td>Yes</td>
<td>Reported</td>
</tr>
<tr>
<td>Family 3</td>
<td>Father and mother</td>
<td>PKHD1 NM_138694</td>
<td>c.1095G&gt;A</td>
<td>p.Trp365Ter</td>
<td>EX14/CDS13</td>
<td>Heterozygous</td>
<td>Chr6:51927340</td>
<td>Mother: NGS, Father: Sanger Sequencing</td>
<td>Yes</td>
<td>Reported</td>
</tr>
<tr>
<td>Family 4</td>
<td>Father and mother</td>
<td>PKHD1 NM_138694</td>
<td>c.1123C&gt;T</td>
<td>p.Arg375Trp</td>
<td>EX15/CDS14</td>
<td>Heterozygous</td>
<td>Chr6:51924836</td>
<td>Mother: NGS, Father: Sanger Sequencing</td>
<td>Yes</td>
<td>Reported</td>
</tr>
<tr>
<td>Family 5</td>
<td>Mother</td>
<td>PKHD1 NM_138694</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>NGS</td>
<td>Yes</td>
<td>…</td>
</tr>
<tr>
<td>Family 6</td>
<td>Mother</td>
<td>PKHD1 NM_138694</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>NGS</td>
<td>Yes</td>
<td>…</td>
</tr>
<tr>
<td>Family 7</td>
<td>Mother</td>
<td>PKHD1 NM_138694</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>NGS</td>
<td>No</td>
<td>…</td>
</tr>
<tr>
<td>Family 8</td>
<td>Father and mother</td>
<td>PKHD1 NM_138694</td>
<td>c.9524A&gt;G</td>
<td>p.Asn3175Ser</td>
<td>EX58/CDS57</td>
<td>Heterozygous</td>
<td>Chr6:51612890</td>
<td>Mother: NGS, Father: Sanger Sequencing</td>
<td>Yes</td>
<td>Reported</td>
</tr>
<tr>
<td>Family 9</td>
<td>Father and mother</td>
<td>PKHD1 NM_138694</td>
<td>c.9218G&gt;A</td>
<td>p.Trp3073Ter</td>
<td>EX58/CDS57</td>
<td>Heterozygous</td>
<td>Chr6:51613196</td>
<td>Mother: NGS, Father: Sanger Sequencing</td>
<td>Yes</td>
<td>Novel</td>
</tr>
<tr>
<td>Family 10</td>
<td>Mother</td>
<td>PKHD1 NM_138694</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>NGS</td>
<td>Yes</td>
<td>…</td>
</tr>
<tr>
<td>Family 11</td>
<td>Father and mother</td>
<td>PKHD1 NM_138694</td>
<td>c.10456T&gt;C</td>
<td>p.Leu3486Leu</td>
<td>EX61/CDS60</td>
<td>Heterozygous</td>
<td>Chr6:51524468</td>
<td>Mother: NGS, Father: Sanger Sequencing</td>
<td>Yes</td>
<td>Novel</td>
</tr>
<tr>
<td>Family 11</td>
<td>Father and mother</td>
<td>PKHD1 NM_138694</td>
<td>c.8863C&gt;G</td>
<td>p.Arg2955Gly</td>
<td>EX57/CDS56</td>
<td>Heterozygous</td>
<td>Chr6:51618086</td>
<td>Mother: NGS, Father: Sanger Sequencing</td>
<td>Yes</td>
<td>Novel</td>
</tr>
</tbody>
</table>
In this study, the screening of 6 exons (3, 32, 36, 57, 58, and 61) led to identification 51% of mutations. Additional 6 exons (9, 14, 18, 19, 22, and 34) increased detection rate to 64%. Obeidova and coworkers showed that 48% of mutations were truncating and 48% were missense changes. In total, 17 different mutant variants, almost evenly scattered throughout the predicted extracellular domain of the protein, were found. The majority of mutations were unique for each family; however, 4 variants including p.Thr36Met, p.Arg1624Trp, p.Leu1966ThrfsX4, and p.Gly2705ValfsX11 within exons 3, 32, 36, and 51, respectively, appeared in several families and formed 55% of all mutations found in this study.

In line with these results, mutations found in our study were also located in the abovementioned exons (3 mutations in exon 58, 1 mutation in exons 61 and 57, and 2 mutations in additional exons, 14 and 39), which have been proposed by Bergmann and colleagues too. Our finding along with previous findings might argue that these exons are hot spots for mutations in different populations.

Currently, no data is available regarding the prevalence of ARPKD mutations or carrier frequency in Iran. However, in view of high consanguineous marriages in Iranian population, the frequency is expected to be higher than the world average. Mutations that have been mostly assessed in previous studies include T36M, I2331K, R2840C, and I3177T, but none of them were found in the patients studied here. Mutations are unequally distributed throughout the gene’s coding sequence.

In this study, 4 novel mutations including c.6469C > T, c.9218 G > A, c.10456T > C, and c.8863C > G and 3 previously reported mutations including c.9524A > G, c.1095G > A, and c.1123C > T, were identified by the next generation sequencing technique. Reported alterations resulted in terminated gestations or neonatal deaths in accordance with previous studies. Six of 7 mutations detected, co-segregated with the disease in studied families. Co-segregation study in 2 healthy members of family 11 showed homozygosity for the c.8863C > G, p.Arg2955Gly mutation.

**CONCLUSIONS**

Genetic counseling and carrier screening for *PKDH1* mutations in populations where inbreeding is frequent such as Iran is highly recommended.
especially for those families with history of an affected member.

CONFLICT OF INTEREST
None declared.

REFERENCES


