Genetic analysis of Iranian autosomal dominant polycystic kidney disease: new insight to haplotype analysis

M. Entezam1, M. R. Khatami2, F. Saddadi3, M. Ayati4, J. Roozbeh5, H. Saghafi1, M. Keramatipour5*

1Department of Medical Genetics, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran
2Nephrology Research Center, Tehran University of Medical Sciences, Tehran, Iran
3School of Medicine, Iran University of Medical Sciences, Tehran, Iran
4Urology Research Center, Tehran University of Medical Sciences, Tehran, Iran
5Shiraz Nephro-Urology Research center, Shiraz University of Medical Sciences, Shiraz, Iran

Abstract: Autosomal Dominant Polycystic Kidney Disease (ADPKD) caused by mutations in two PKD1 and PKD2 genes. Due to the complexity of the PKD1 gene, its direct mutation screening is an expensive and time-consuming procedure. Pedigree-based haplotype analysis is a useful indirect approach to identify the responsible gene in families with multiple affected individuals, before direct mutation analysis. Here, we applied this approach to investigate 15 appropriate unrelated ADPKD families, selected from 25 families, who referred for genetic counseling. Four polymorphic microsatellite markers were selected around each PKD1 and PKD2 loci. In addition, by investigating the genomic regions, two novel flanking tetranucleotide STR markers were identified. Haplotype analysis and calculating Lod score confirmed linkage to PKD1 in 9 families (60%) and to PKD2 in 2 families (13%). Linkage to both loci was excluded in one family (6.6%). In 2 families (13%) the Lod scores were inconclusive. Causative mutation was identified successfully by direct analysis in two families with confirmed linkage, one to PKD1 and another to PKD2 locus. The study showed that determining the causative locus prior to direct mutation analysis is an efficient strategy to reduce the resources required for genetic analysis of ADPKD families. This is more prominent in PKD2-linked families. Selection of suitable markers, and appropriate PCR multiplexing strategy, using fluorescent labeled primers and 3 primer system, will also add value to this approach.

Key words: Autosomal Dominant Polycystic Kidney Disease, haplotype analysis, STR marker development, mutation screening.

Introduction

Autosomal Dominant Polycystic Kidney Disease (ADPKD [OMIM # 173900]) is the most common genetic renal disorder, which has a reported prevalence of 1:400 to 1:1000 in all races (1, 2). It is the fourth cause of end stage renal disease (ESRD) in Iran (3). ADPKD is a multisystem disorder characterized by bilateral renal cysts, and a variety of external manifestations, including cysts in liver, seminal vesicles, pancreas, and arachnoid membrane; and vascular abnormalities such as intracranial aneurysms, dilatation of the aortic root, dissection of the thoracic aorta, mitral valve prolapse; and abdominal wall hernias (4).

Clinical diagnosis of the disease is based on imaging techniques, whereas the definite diagnosis method is DNA testing (5). There are two known causative genes for this genetically and phenotypically heterogeneous disease, PKD1 on chromosome 16p13.3 (6) and PKD2 on chromosome 4q21-23 (7). Genetic testing is valuable for the definite diagnosis in individuals with an atypical renal findings, and younger patients from milder pkd2-families, also it is the only way for determining the healthy living-related kidney donors (8-10). Nevertheless, some facts complicated molecular analysis, comprising the large size of the PKD1 gene with 46 exons and 12,909bp coding sequence (NM_000297.3). As an alternative approach, family-based linkage analysis is an indirect method for identifying the responsible gene in families with more than two affected individuals, before direct mutation screening.

PKD2 is a smaller gene with 15 exons and 2904bp coding sequence (NM_000297.3). As an alternative approach, family-based linkage analysis is an indirect method for identifying the responsible gene in families with more than two affected individuals, before direct mutation screening.

The aim of this study was to evaluate a fluorescent multiplex PCR for rapid and economical haplotyping and linkage analysis of ADPKD families, along with mutation detection. For this purpose, we investigated known, as well as possible new identified microsatellite markers in and around PKD1 and PKD2 loci. This investigation resulted in characterization appropriate eight previously known and two new microsatellite markers in these loci. Also, we assessed two fluorescent genotyping systems, the typical method of using fluorescently labelled locus-specific primers and the cost effective methodology based on 3 primer system. Due to the limitation of financial resources, we only perform mutation screening in one linked pedigree for each locus to evaluate our approach.
Materials and Methods

Subjects and DNA Extraction

25 families were referred for genetic counseling to the department of medical genetics at Tehran University of Medical Sciences, Tehran, Iran. According to pedigree structure, 15 unrelated families (71 patients and 40 healthy individual) were selected for linkage analysis. The diagnosis of ADPKD was confirmed by the nephrologist, based on the clinical history, symptoms and the ultrasonographic findings. The most important inclusion criteria was availability of at least two affected and two confirmed-unaffected members. The families were selected from different ethnicities of Iranian population. Sporadic cases and families with indefinite clinical diagnosis were excluded from the study.

Peripheral blood samples were collected from all the individuals. DNA was extracted using Exgene Blood SV mini kit (GENEALL BIOTECHNOLOGY CO, LTD, South Korea) according to the manufacture’s instruction. The quality and quantity of DNA were assessed by the NanoDrop spectrophotometer (ND-1000) as well as agarose gel electrophoresis.

The study was approved by the research ethics committee of the Tehran University of Medical Sciences. All the patients and their family members provided informed consent.

Microsatellite identification

8 polymorphic microsatellite markers flanking Pkd1 and Pkd2 were selected according to the distance from each gene and the mean heterozygosity, as follows: D16S3024, D16S3395, D16S3252 (Pkd1 intragenic marker), D16S664, D4S2409, D4S2929, D4S1563 and D4S2460. Except the two D16S3395, and D4S2409 (GATA26B12), tri- and tetra-nucleotide repeats, all were CA di-nucleotide short tandem repeat (STR) markers. The number of alleles, frequencies and repeats, all were CA di-nucleotide short tandem repeat (STR) markers. The number of alleles, frequencies and heterozygosity content for each microsatellite marker characterized in 50 unrelated controls, prior to performing linkage analysis.

In order to find closer tetranucleotide STR markers, the genomic region of approximately 1Mb upstream and downstream flanking sequences of Pkd1 and Pkd2 genes, NC_000016.9 (2138709...2185899) and NC_00004.11 (88928799...88998931) respectively, was scanned using the TRF (12) and SERV (13) online programs. The most important inclusion criteria was availability of at least two affected and two confirmed-unaffected members. The families were selected from different ethnicities of Iranian population. Sporadic cases and families with indefinite clinical diagnosis were excluded from the study.

Peripheral blood samples were collected from all the individuals. DNA was extracted using Exgene Blood SV mini kit (GENEALL BIOTECHNOLOGY CO, LTD, South Korea) according to the manufacture’s instruction. The quality and quantity of DNA were assessed by the NanoDrop spectrophotometer (ND-1000) as well as agarose gel electrophoresis.

The study was approved by the research ethics committee of the Tehran University of Medical Sciences. All the patients and their family members provided informed consent.

Linkage analysis

Mean heterozygosity (MH) and polymorphism information content (PIC) values were calculated for all markers. EasyLINKAGE plus v5.05 (15) was used for linkage analysis. GeneHunter v2.1r5 software was used for haplotyping and multipoint parametric linkage analysis.

Table1. Primer sequences and informations of microsatellite markers.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Marker</th>
<th>Modification</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Locus specific fluorescent primers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pkd1</td>
<td>D16S3395</td>
<td>HEX</td>
<td>CCTGGCAGTAAGTCTCTGAAA</td>
<td>CTAACCCTCGACAGGTTCCTG</td>
</tr>
<tr>
<td></td>
<td>D16S3252</td>
<td>5-FAM</td>
<td>CCAAGGTGGAGGAAAGGTT</td>
<td>GGTTGTCGCCGCTATCAAGAC</td>
</tr>
<tr>
<td></td>
<td>D16S664</td>
<td>TAMRA</td>
<td>GGGAGTCTGCAAAAAGTCT</td>
<td>TCGCGGTCATAAATTTGTTG</td>
</tr>
<tr>
<td></td>
<td>PKD1-M1</td>
<td>5-FAM</td>
<td>TGGCTCAAAATGGCTTC</td>
<td>AAATTATAGGTTGATTGAGAC</td>
</tr>
<tr>
<td>Pkd2</td>
<td>D4S2409</td>
<td>HEX</td>
<td>CCCACAAAGCTATGGAAGTA</td>
<td>CTCACCTCCCTCATTTT</td>
</tr>
<tr>
<td></td>
<td>PKD2-M1</td>
<td>5-FAM</td>
<td>CCAATCATGTCACGAAACATG</td>
<td>AGGCAAATCGGCTCACGAAAG</td>
</tr>
<tr>
<td></td>
<td>D4S1563</td>
<td>5-FAM</td>
<td>GTGCCTGACACACCTG</td>
<td>ACTATTGCTGTCCTGACCC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 primer system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pkd1</td>
<td>D16S3024</td>
<td>5-FAM</td>
<td>TACGCAATCCCGATTTGAGACCCTGAAGCTATGAGGATCTCCT</td>
<td>ATGTTATAGGCTCTCGAAGGG</td>
</tr>
<tr>
<td></td>
<td>D4S2929</td>
<td>5-FAM</td>
<td>TACGCAATTCCCGATTTGAGACCCTGAAGCTATGAGGATCTCCT</td>
<td>GCTCTTCTCCACCATCCTCATAGG</td>
</tr>
<tr>
<td></td>
<td>D4S2460</td>
<td>5-FAM</td>
<td>TACGCAATTCCCGATTTGAGACCCTGAAGCTATGAGGATCTCCT</td>
<td>CTTGTAGGAGTCTGCTGGTGA</td>
</tr>
</tbody>
</table>

Multiplex Polymerase chain reaction

We designed two different approaches for multiplexing of PCR reactions. One fluorescent multiplex PCR reaction developed for simultaneous amplification of seven markers. For them all forward primers were labeled with a distinct fluorescent dye. New Primer sets were designed according to each marker flanking sequences, by primer 3 software version 4.0.0 (http://primer3.ut.ee/) (table 1). The cycling parameters were as follows: initial denaturation of 95°C for 5 min, followed by 30 cycles of 95°C for 30 seconds, 57°C for 30 seconds, 72°C for 30 seconds, and a final extension of 72°C for 10 minutes. The primer mix was added after the first denaturation step. The total reaction volume was 40 µl, using PCR Master Mix (Ampliqon, A140303) and 1X Q-solution (Qiogene), along with 50 ng DNA. The MgCL2 final concentration was 1.5 mM. The final concentration of all primers was 0.17µM except for D4S2409, D16S664 and PKD1M1 (0.25 µM).

Whereas, the other three markers were amplified with 3-primer system, using a fluorescently labelled universal primer described previously (14) in combination with a pair of locus-specific primers, in which forward primer modified with a 5’ universal primer sequence tail (table1b). 40 cycle amplification was performed as the above conditions except the annealing temperature of 62°C. The total PCR volume was 25 µl using PCR Master Mix (Ampliqon, A140303) and 1X Q-solution (Qiogene), along with 0.2µM, 0.2µM, and 0.01µM, of F2, R1 and F1 primers respectively, and 30 ng DNA. The MgCL2 final concentration was 1.5 mM. The 1/20 ratio of F1/F2 primers ensures the depletion of F2 primers in early PCR cycles, and subsequent incorporation of labeled primer in proceeding cycles.

PCR products were pooled by adding 10 µl of each product, so fragment length analysis can be performed in one single reaction using Applied Biosystems 3100 Genetic Analyzer with GS500LIZ size standard (Fischgam Biotech Company, Tehran, Iran). Genescan results were analyzed by Peak Scanner Software v1.0.
Mutation analysis was performed only in two families with confirmed linkage to PKD1 or PKD2 genes. All coding exons and flanking regions of PKD1 and PKD2 genes were amplified in average of 600 bp fragments, and screened by direct Sanger sequencing. Primer sequences and PCR conditions are available on request. Sequencing data were analyzed by Mutation Surveyor software version 3.03 (www.softgenetics.com/mutationSurveyor.html).

Results

New Microsatellite identification

Four microsatellite markers were selected based on the distance to the gene, TRF score higher than 100 and VARscore higher than 1. Then according to polymorphic characteristics of these markers in control samples, we chose one intragenic and one 3' downstream tetra-nucleotide repeat for further pedigree-based linkage analysis of PKD2 (PKD2-M2) and PKD1 (PKD1-M1) respectively. Considering the crossing-over rate between these and adjacent markers and the physical distance between the markers, we can estimate genetic distances. These two markers have not been mentioned in mapping databases. In addition, we were not able to find any previous report of these microsatellites in the literature.

Microsatellite marker amplification and characterization

The physical (Mb) and genetic (cM) locations of other markers taken from the MapViewer (www.ncbi.nlm.nih.gov/projects/mapview/cgi) and Marshfield genetic maps (http://research.marshfieldclinic.org/genetics/GeneticResearch/compMaps.asp) (figure 1).

Figure 2. represents the electropherograms of PCR amplification products. The result of universal labeled fragments showed no difference in peak quality and precision from the locus specific labeled fragments. Allele sizing could be assessed unambiguously for all investigated markers.

Polymorphic characteristics of each marker and the corresponding allele frequencies summarized in table 2 and 3 respectively.

Linkage analysis and mutation detection

Haplotype analysis and Lod score identified linkage to PKD1 in 9 families (60%) and to PKD2 in 2 families (13%). Interestingly, our haplotype analysis excluded the linkage of one family to both PKD1 and PKD2 loci (6.6%). Linkage to both genes was possible in 2 families (13%). This co-segregation is likely to be due to the limited number of analyzed subjects and consanguinity marriage which caused non-informative haplotypes. Haplotype analysis was not possible in one family due to incorrect allele segregation and genotyping errors.

DNA sequencing in 2 families with confirmed linkage to PKD1 or PKD2, revealed pathogenic mutations along with neutral variants. Figure 3a. shows one PKD2 linked pedigree (P-1). In this family sequencing of exon1 identified a nonsense mutation c.556C>T (p.R186X), results in a premature translation stop codon. Analyzing 4 affected and 1 healthy individuals of the family confirmed the correct segregation of the mutation with the disease. This mutation was first described in four unrelated families (17) and is the most common missense mutation according to ADPKD mutation database (http://pkdb.mayo.edu). Other variants and polymorphisms identified in this family are shown in table 4.
The current study evaluates the application of haplotype analysis in Iranian ADPKD families. It also investigates the most accurate and cost-effective approach in regards to microsatellite marker selection and genotyping method. Identifying the pathogenic mutation in two linked pedigrees approves our linkage results. This is one of the first articles which studies PKD1 and PKD2 genes such comprehensively in Iranian population.

The broad range of coding sequences, extremely high homology of a large part of the gene with six homologous genes (HG), the high GC-content of some part of the gene, particularly the first exon (approximately 85%), and lack of hot spot regions for mutation detection, besides high probability of detecting variants of uncertain significance, make mutation analysis of PKD1 to be one of the most difficult genetic tests. Gene tracking is a powerful method for indirect genetic diagnosis.

### Table 3. Distribution of allele frequencies in ADPKD families.

<table>
<thead>
<tr>
<th>Location</th>
<th>Alteration</th>
<th>cDNA</th>
<th>protein</th>
<th>DbSNP</th>
<th>rs#</th>
</tr>
</thead>
<tbody>
<tr>
<td>PKD1</td>
<td>Ex2</td>
<td>c.238C&gt;T</td>
<td>p.R80W</td>
<td>rs551353498</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ex5</td>
<td>c.1119C&gt;T</td>
<td>p.L373V</td>
<td>rs199685642</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ex11</td>
<td>c.2216G&gt;A</td>
<td>p.Q739R</td>
<td>rs40433</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ex15</td>
<td>c.4674G&gt;A</td>
<td>p.T1558T</td>
<td>rs79884128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IVS22</td>
<td>c.8161+22G&gt;A</td>
<td></td>
<td>rs2725221</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IVS22</td>
<td>c.8161+285C&gt;T</td>
<td></td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IVS30</td>
<td>c.10050+162delC</td>
<td></td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>PKD2</td>
<td>Ex1</td>
<td>c.83G&gt;C</td>
<td>p.R28P</td>
<td>rs1805044</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EX1</td>
<td>c.568G&gt;A</td>
<td>p.A190T</td>
<td>rs117078377</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IVS 3</td>
<td>c.844-22G&gt;A</td>
<td></td>
<td>rs2725221</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IVS 6</td>
<td>c.1548+63C&gt;T</td>
<td></td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3′ UTR</td>
<td>c.*363delA</td>
<td></td>
<td>This study</td>
<td></td>
</tr>
</tbody>
</table>

### Discussion

The current study evaluates the application of haplotype analysis in Iranian ADPKD families. It also investigates the most accurate and cost-effective approach in regards to microsatellite marker selection and genotyping method. Identifying the pathogenic mutation in two linked pedigrees approves our linkage results. This is one of the first articles which studies PKD1 and PKD2 genes such comprehensively in Iranian population.

The broad range of coding sequences, extremely high homology of a large part of the gene with six homologous genes (HG), the high GC-content of some part of the gene, particularly the first exon (approximately 85%), and lack of hot spot regions for mutation detection, besides high probability of detecting variants of uncertain significance, make mutation analysis of PKD1 to be one of the most difficult genetic tests. Gene tracking is a powerful method for indirect genetic dia-
agnosis of heterogeneous inherited diseases, especially in the case of complicated genes like PKD1. Moreover identifying the causative gene, without characterizing the mutation itself, sometimes might be more relevant to the clinical features (18).

In families with more than two affected individuals, prior to proceeding to direct mutation screening, the causative gene can be determined by performing a multiplex PCR reaction and one run of capillary electrophoresis. The worthiness is avoiding the expensive and time consuming procedure of PKD1 screening, in the case of PKD2 linked result, and assessing genetic diagnosis more conveniently.

Furthermore, multiplexing strategy for simultaneous analyzing of all flanking markers, make this approach cost effective and rapid. Our previous study shows its usefulness for genetic testing of familial hypertrophic cardiomyopathy (HCM) (UJP, in press). 3 primer system seems the most cost-effective and reliable alternative for microsatellite genotyping. Here, we evaluate the accuracy of alleles amplified by this strategy, in comparison with traditional fluorescent STR typing method in ADPKD families. Genotyping labeled PCR fragments, using fluorescent universal primers is widely adopted in variety of molecular genetic researches. Previous studies evaluated different strategies and primer sequences, particularly in molecular biological fields (19-21). In this study, we optimize this methodology for haplotype analysis of human genes and succeed to genotype the investigated markers accurately and unambiguously in all our pedigrees.

In comparison to dinucleotide markers, tetra-nucleotide microsatellites are preferred markers due to better allele sizing and the less stutter bands in the amplification products, (22). Also selecting closer markers can reduce the probability of false negative results due to recombination,(23). Therefore two new microsatellite markers were investigated. Regardless of the moderate heterozygosity of these markers, our data suggest their usefulness for haplotype analysis of ADPKD families, particularly in PKD2 investigations.

Haplotype analysis and calculated Lod score of linked markers identified the causative gene in 11 families (75%) which is approximately similar to Slovenian and Hungarian records (24, 25). Homozygosity of alleles , non-informativeness due to consanguineous marriage in earlier generaions, and the limited number of available individuals could be responsible for undetectable results in remaining families. The family with precluded linkage to both genes requires more investigation. This can be explained by misdiagnosis or possible involvement of other genetic loci as the cause of ADPKD.

Due to a limited budget, mutation screening was performed only in one PKD1 or PKD2 linked families. c.556C>T (p.R186X) was confirmed as a pathogenic mutation in PKD2 linked pedigree. Substitution of G to C leads to, alteration of Argenin to premature Stop codon. It therefore expressed a mutated 186 amino-acid product, which intern, is predicted to cause nonsense-mediated decay (NMD). Although there is no obvious genotype-phenotype correlation in PKD2 patients, It seems reasonable that truncating mutations occurring in early exons of the gene, result in more severe phenotypic features (16, 26). The proband in respected family was a 54 year old woman who did not reach ESRD. She had gallstone and suffered from bilateral ovarian cysts similar to her older sister. All related patients taking medication for hypertension. The pathogenic mutation in PKD1 linked pedigree was c.8522G>A (p.E2771K) in exon 23. This C to T transition occurs at the CpG dinucleotides which is the known hotspot point for mutations. Alteration of a glutamate residue to a lysine predicts to affect the secondary structure of up to 40 amino acid residues upstream (25). In comparison to pedigree1, in this pedigree two patients reached ESRD at the age of 45 and 53. The youngest patient was 23 and the oldest without ESRD was 56.

In conclusion, despite all the limitations of the pedigree-based linkage analysis, this approach can still be helpful for some appropriate families. Selecting the suitable method can reduce the cost and the time of analyzing. Identifying the disease haplotype in affected individuals will facilitate preimplementation genetic diagnostics (PGD) and prenatal diagnostics (PND) in families with sever phenotypes.

Acknowledgments:
The authors are very thankful to the participants for their kindly collaborations.

References


