Whole-Exome Sequencing application for genetic diagnosis of kidney diseases: a study from Southwest of Iran

Mina Zamani¹,², Tahereh Seifi¹,², Sahar Sedighzadeh¹,², Samira Negahdari², Jawaher Zeighami², Alireza Sedaghat*²,⁴, Tahereh Yadegari², Alishossein Saberi²,³, Mohammad Hamid²,⁵, Gholamreza Shariati*²,³ and Hamid Galehdari*¹

¹Department of Biology, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
²Narges Genetics Diagnostic Laboratory, Ahvaz, Iran.
³Department of Genetics, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.
⁴Health Research Institute, Diabetes Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.
⁵Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran.

*Corresponding authors

Alireza Sedaghat
Health Research Institute, Diabetes Research Center, Ahvaz Jundishapur University of Medical Sciences
Ahvaz, Iran, Islamic Republic of
alireza_sedaghat51@yahoo.com

Gholamreza Shariati
Department of Genetics, Ahvaz Jundishapur University of Medical Sciences
Ahvaz, Iran, Islamic Republic of
shariatig@yahoo.com

Hamid Galehdari
Department of Biology, Faculty of Science, Shahid Chamran University of Ahvaz
Ahvaz, Iran, Islamic Republic of
galehdari187@yahoo.com

Copyright 2021 by American Society of Nephrology.
Introduction

Kidneys have important and vital functions providing maintenance of the overall body health. Many kidney diseases are caused by single gene defects. Kidney disease is categorized in a heterogeneous group of disorders affecting the kidney, structurally and functionality. End-stage of kidney disease is known as kidney failure. Kidney disease is divided into two forms including; acute kidney injury and chronic kidney disease [1]. Among kidney diseases Poly-cystic kidney disease (PKD or PCKD, also known as polycystic kidney syndrome) is revealed to have genetic background [2]. PKD is characterized by the presence of cysts in kidneys. The development and growth of cysts cause abnormality in renal tubules. PKD is clinically and genetically heterogeneous disease [2, 3]. Clinical characteristics of autosomal dominant and recessive forms of PKD are variable in penetrance. A range from neonatal death to incidence in old age was reported for PKD [2, 4]. Disease causing variants in three genes including \textit{PKD1}, \textit{PKD2} and \textit{PKDH1} can cause PKD. The pathogenic variations in the \textit{PKD1} and \textit{PKD2} genes can cause autosomal dominant (ADPKD) pattern of inheritance. The pathogenic variations in \textit{PKDH1} can cause autosomal recessive (ARPKD) pattern of inheritance [2]. There are more than 500 monogenic causes of chronic kidney disease [5] and numerous genes are listed in NGS panels for PCKD including: \textit{ALG9}, \textit{ANKS6}, \textit{ATP6V0A4}, \textit{BICC1}, \textit{GANAB}, \textit{GLIS3}, \textit{HNF1B}, \textit{INVS}, \textit{LRP5}, \textit{MUC1}, \textit{NOTCH2}, \textit{NPHP3}, \textit{OFD1}, \textit{PKD1}, \textit{PKD2}, \textit{PKHD1}, \textit{SEC61A1}, \textit{TMEM231}, \textit{TSC1}, \textit{TSC2}, \textit{UMOD}, \textit{ZNF423}, etc. [2]. For classification, pre-prognosis, monitoring and treatment, identification of the etiology of the disease is necessary. It helps us to understand the scenario of the disorder causing and select the best approach for drug treatment [6]. Next-generation sequencing (NGS) is a powerful technique that enables rapid and cost-effective parallel sequencing of large panels of genes or whole exome sequences. However, the targeted panel sequencing approach related to the Whole Exome Sequencing (WES) is confirmed to provide deep coverage of the interested sequences but WES can provide an opportunity for novel variation detection [7-9]. Regarding to polygenic entity of kidney diseases which meaning that many genes are involved in the pathogenesis of the disease, for example for focal segmental glomerulosclerosis (FSGS), pathogenic variants in more than 20 podocyte-specific genes such as \textit{NPHS1}, \textit{NPHS2}, \textit{WT-1}, \textit{LAMB2}, \textit{CD2AP}, \textit{TRPC6}, \textit{ACTN4} and \textit{INF2} was announced [10]. In fact, conventional Sanger sequencing may be very time consuming and expensive, so, using NGS would be very helpful. In our
study, we identified probable pathogenic variations associated with chronic kidney disease using WES.

**Materials and Methods**

**Patients/Cases report**

22 enrolled patients and their parents were recruited from Southwest of Iran with demographic information showed in Table 1. All the patients referred to the Narges Medical Genetics and Prenatal Diagnosis Laboratory, Ahvaz, Iran in 2017-2020. Genetic counselling for all the patients were done by Genetic specialists from the laboratory. This research study was approved by the Ahvaz Jundishapur University of Medical Sciences. For the studies, informed consent was obtained from all the families. All the experiments were done in accordance with relevant guidelines and regulations.

**DNA extraction**

10 ml of peripheral blood was withdrawn from each of the enrolled patients and their parents which were collected in EDTA tubes. Genomic DNA was extracted using the standard salting out protocol. The quality and quantity of the extracted DNA samples were checked by gel electrophoresis and nano-drop.

**Exome sequencing**

WES was applied for the patients by Macrogen Co. Korea. Samples were subjected to Exome enrichment with the sure select v6 followed by sequencing using the Illumina HiSeq 2000 genome analyzer platform.

**Sanger Validation**

Targeted regions of the genes were amplified by PCR using primers designed by software Oligo 7. The PCR reactions were conducted using Master Mix (Ampliqon, Denmark). After an initial 5 denaturation at 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 sec, annealing at 60 °C for 30 sec, and extension at 72 °C for 30 s, and ended by a final extension at 72 °C for 5 min. PCR products were directly sequenced and analyzed by the ABI Prism 3700 automated genetic analyzer (Applied Biosystems). The results were analyzed with Chromas LITE 2.1.1, and then
compared with the reported gene sequence using the BLASTN program. Also the presence of the detected mutation was confirmed by the parent analysis and bi-directional sequencing.

**In-silico analysis**

There are lots of various tools for prediction of pathogenicity of variations in genes coding and non-coding regions. In the present study we used some of these tools for evaluation of disease causing potentiality of the variations.

**Result**

At result we found disease causing variants in *PKHD1, PKD1, PKD2L2, NPHS1, NPHP3, CD2AP, COL4A4, DCDC2* genes. All the variant were analyzed in terms of pathogenicity using different tools including MutationTaster, PredictSNP1, PredictSNP2, Proven, Gnomad, GME, PolyPhen, VarSome, and I-Mutant (table 2).

As figure 1 shows among 15 cases reported in table 2; 53.33% of the patients (8 cases) have pathogenic variations in *PKHD1* gene, 13.33% (2 cases) in *PKD1*, 6.66% (1 case) in *NPHS1*, 6.66% (1 case) in *NPHP3*, 6.66% (1 case) in *CD2AP*, 6.66% (1 case) in *COL4A4*, and 6.66% (1 case) in *DCDC2*.

**Conclusion**

In this study we demonstrated that, WES has been efficiently worked for genetic profiling of kidney diseases. We evaluated WES in 22 Iranian patients with the history of clinically different inherited kidney diseases. Samples from family members were taken to perform segregation analysis as far as possible. We found two disease causing variants (c.11626_11635delGCCGGCCGCG, and c.1687C>T) in *PKD1* gene. According to the in-silico analysis using MutationTaster, ACMG classification and etc. the variants are disease causing and pathogenic. In addition, the variants were not present in our homemade exome database and the public SNP databases including; dbSNP, ExAC, and GnomAD. Due to the proximity of the pseudogenes to *PKD1* gene, it has always been difficult to identification of the pathogenic variants in *PKD1* [11, 12]. Several assays indicated that, WES is a proper technique with a relatively short time and low cost for identifying single-gene disease and genetic heterogeneity complaints in ADPKD patients in comparison to standard diagnostics based on Sanger sequencing and multiplex ligation dependent probe amplification (MLPA) assays [13, 14].
MLPA is clinically suggested only for patients whose disease causing variants have not been determined through NGS investigation [15]. Ranjzad et al. used targeted NGS for detection of novel pathogenic variants in Iranian families with ADPKD and demonstrated that, NGS can significantly reduce the cost and time for simultaneous sequence analysis of *PKD1* and *PKD2* [16]. Mallawaarachchi et al. analyzed 28 unique pedigrees with ADPKD and reported that, whole genome sequencing (WGS) is able to overcome technical challenges created by pseudogenes proximal to *PKD1* [11]. However it should be considered that *PKD1* is inherently difficult to sequence using WES [17]. The clinical genetic diagnosis of ADPKD significantly impacts the quality of patients' life involved in renal transplantation, in which the presence of a familial pathogenic variant in a transplant phenotypically normal donor is screened. Also, in cases of identified a family pathogenic variant, we can prevent transmission to the children via current prenatal diagnostic techniques [18, 19]. At the result of *in-silico* analysis, we found 8 disease causing variants included 4 reported pathogenic single nucleotide variants (SNVs) (rs137852949, rs398124503, rs745770404, and CM100562) and 3 novel probable disease causing variants in *PKHD1* gene. These variants were not present in our exome database and had 0 homozygotes with low allele frequency and or not found in the public SNP databases including ExAC and GenomAD databases. Efforts to prevent severe ARPKD complications from embryonic development period have led to several ways, including second trimester sonography and molecular genetic analysis for prenatal diagnosis [20]. Because of the large size of *PKHD1*, genetic heterogeneity and broad phenotypic of cystic and polycystic kidney diseases, WES is an efficient approach for pre and postnatal diagnosis of ARPKD [21, 22]. Obeidova et al. used NGS for clinical analysis of ARPKD in 24 families and reported that, NGS of the *PKHD1* gene is a very convenient procedure with high precision for molecular diagnosis in patients with very similar clinical symptoms to ARPKD [23]. We found one disease causing deletion (rs751527253) in *NPHP3* gene. At the result of *in-silico* analysis using MutationTaster, and ACMG classification this variant is disease causing and pathogenic. Early diagnosis of Nephronophthisis 3 is dependent on clinical and imaging findings, because of the similarity of symptoms in patients with renal problems, the exact diagnosis of these patients through NGS is important [24, 25]. In this study we detected some potentially pathogenic variants via *in-silico* analysis using MutationTaster, ACMG classification, and etc. in *CD2AP*, *NPHS1*, and *COL4A4* genes in the patients with different clinical symptoms of renal disease. Bekheirnia et al. analyzed
WES of 62 patients with kidney and urinary tract (CAKUT) Nearly 5% of individuals with CAKUT have pathogenic SNVs in known key genes that can be uncovered by WES. In addition, 6.5% of these patients have pathogenic copy number variations (CNVs) that were extracted from WES data [26]. Also some reports showed potential of WES to find novel kidney disease causing variants [27, 28]. We applied WES for the patients with kidney disease from Southwest of Iran’s populations and were able to detect pathogenic variations in 68 % of the enrolled patients, although the disease causing variants, were not determined in 32 % patients. We confirmed that, WES is a very advantageous procedure for identifying genes and mutations in kidney disease and can accurately detect novel genes and variants, consequently, WES is emerging as a preferred diagnostic tool for hereditary disorders including kidney disease.

Disclosures
All authors have nothing to disclose.

Funding
None

Author Contributions
M Zamani: Conceptualization; Formal analysis; Investigation; Methodology; Writing - original draft
T Seifi: Formal analysis; Investigation; Writing - original draft
S Sedighzadeh: Formal analysis; Investigation; Writing - original draft
S Negahdari: Formal analysis; Writing - original draft
J Zeighami: Formal analysis; Methodology
A Sedaghat: Supervision
T Yadegari: Investigation
A Saberi: Supervision
M Hamid: Supervision
G Shariati: Conceptualization; Supervision
H Galehdari: Conceptualization; Supervision


**Table 1.** Demographic data.

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>35%</td>
</tr>
<tr>
<td>Male</td>
<td>65%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10</td>
<td>88%</td>
</tr>
<tr>
<td>11-20</td>
<td>8%</td>
</tr>
<tr>
<td>21-30</td>
<td>4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lur</td>
<td>38%</td>
</tr>
<tr>
<td>Fars</td>
<td>12%</td>
</tr>
<tr>
<td>Arab</td>
<td>44%</td>
</tr>
<tr>
<td>Kurd</td>
<td>6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consanguinity</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Consanguineous</td>
<td>86.66%</td>
</tr>
<tr>
<td>Non- Consanguineous</td>
<td>13.33%</td>
</tr>
<tr>
<td>Gene</td>
<td>Phenotype (OMIM)</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>PKD1</td>
<td>Polycystic kidney disease 1</td>
</tr>
<tr>
<td>PKD1</td>
<td>Polycystic kidney disease 1</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>PKHD1</td>
<td>Polycystic kidney disease 4, with or without hepatic disease</td>
</tr>
<tr>
<td>Meckel syndrome_3</td>
<td>Renal-hepatic-pancreatic dysplasia</td>
</tr>
<tr>
<td>NPHP3</td>
<td>Nephronophthisis</td>
</tr>
<tr>
<td>NPHS1</td>
<td>Nephrotic, type 1 syndrome</td>
</tr>
<tr>
<td>COL4A4</td>
<td>Alport syndrome</td>
</tr>
<tr>
<td>DCDC2</td>
<td>Nephronophthisis</td>
</tr>
</tbody>
</table>
Figure 1 The contribution of the genes harboring the presented kidney diseases causing variants.
Figure 1

- PKHD1: 53.33% (8)
- PKD1: 6.66% (1)
- NPHS1: 6.66% (1)
- NPHP3: 6.66% (1)
- COL4A4: 13.33% (2)
- CD2AP: 6.66% (1)
- DCDC2: 6.66% (1)