Improved Mapping Quality and Coverage in Highly Homologous *PKD1* Gene Enable High Diagnostic Yield in ADPKD

Satu Valo, Jonna Tallila, Hannu Jalanko, Johanna Sistonen, Annakarin Korppoo, Kim Gall, Mikko Muona, Pertteli Salmenperä, Massimiliano Gentile, Samuel Myllykangas, Tero-Pekka Alastalo, and Juha Koskenvuo

Blueprint Genetics, Biomedicum 1, Haartmaninkatu 8, Helsinki, 00290, Finland

Introduction

Autosomal dominant polycystic kidney disease (ADPKD) is the most common genetic kidney disease. Approximately 50% of individuals with ADPKD develop end-stage renal disease (ESRD) by the age of 60 years. ADPKD is caused primarily by mutations in two genes, *PKD1* and *PKD2*, encoding polycystin 1 and 2, which are essential components of epithelial cilia. Genetic testing has become an important factor in the management of ADPKD patients and their families. However, analysis of *PKD1* is technically challenging due to its large size, high GC-content, and duplication of the first 33 exons with a high degree of homology (90-99% identity) to six nearby pseudogenes (*PKD1P1–P6*). We evaluated the diagnostic yield and performance of our in-house tailored Polycystic Kidney Disease and Cystic Kidney Disease Panels, including in total 42 genes, in an unselected cohort of patients referred for cystic kidney diseases.

Methods

Next-generation sequencing (NGS) was performed using the IDT xGEN Exome Research Panel with added custom probes and the Illumina NovaSeq 6000 platform. This assay provides improved mapping quality and coverage in many difficult-to-sequence regions, including *PKD1*, compared to other NGS methods assessed in our laboratory. Majority of the analyses (170/183) were performed as PLUS analysis that combines sequence and deletion/duplication analysis utilizing NGS data. All pathogenic or likely pathogenic variants were confirmed with an appropriate orthogonal method. Variants in the difficult-to-sequence region of *PKD1* were confirmed using Sanger sequencing with custom-designed primers.

Results

In the study cohort of 183 index patients, a genetic diagnosis was established in 54% (n=99) of cases with disease causing variants detected in 11 different genes (Table 1). In 63% and 11% of the diagnostic cases the disease causing variant was identified in *PKD1* or *PKD2*, respectively. Interestingly, 7% (n=7) of the cases had a diagnostic deletion including 4 heterozygous *HNF1B* whole gene deletions, 2 *PKD1* multiexon deletions, and 1 homozygous *NPHP1* whole gene deletion. Of all likely disease causing *PKD1* variants identified in 62 patients, 79% (n=49) were classified as pathogenic or likely pathogenic and 21% (n=13) as variants of uncertain significance (VUS favoring pathogenic) (Figure 2A). Majority of the identified *PKD1* variants were missense (40%, n=25) and nonsense (26%, n=16) variants (Figure 2B). Furthermore, 81% (n=50) of the variants were located in the duplicated region of *PKD1* (exons 1-33). A number of *PKD1* sequence variants (24%, n=15) were located in exon 15 indicating a possible mutational hotspot. In *PKD2*, a total of 11 pathogenic or likely pathogenic variants were detected that included mainly truncating variants. Additional clinical utility of the test was shown by sequencing 10 ADPKD patients with a negative test result from previous NGS-based testing (Table 2).

Table 1. Genes with diagnostic findings.

Gene	Number of cases	%
PKD1	62	63
PKD2	11	11
PKHD1	13	13
HNF1B	4	4
INVS	2	2
NPHP3	2	2
NPHP1	1	1
PRKCSH	1	1
SEC63	1	1
WDR19	1	1
PAX2	1	1

PKD1 Coverage

Our panels provided mean coverage of 192x within the 42 genes. Specifically, *PKD1* provided both high mean coverage (205x) and excellent mapping quality with 99.96% of the target nucleotides covered at least 20x with a mapping quality threshold of 20 (Figure 1).

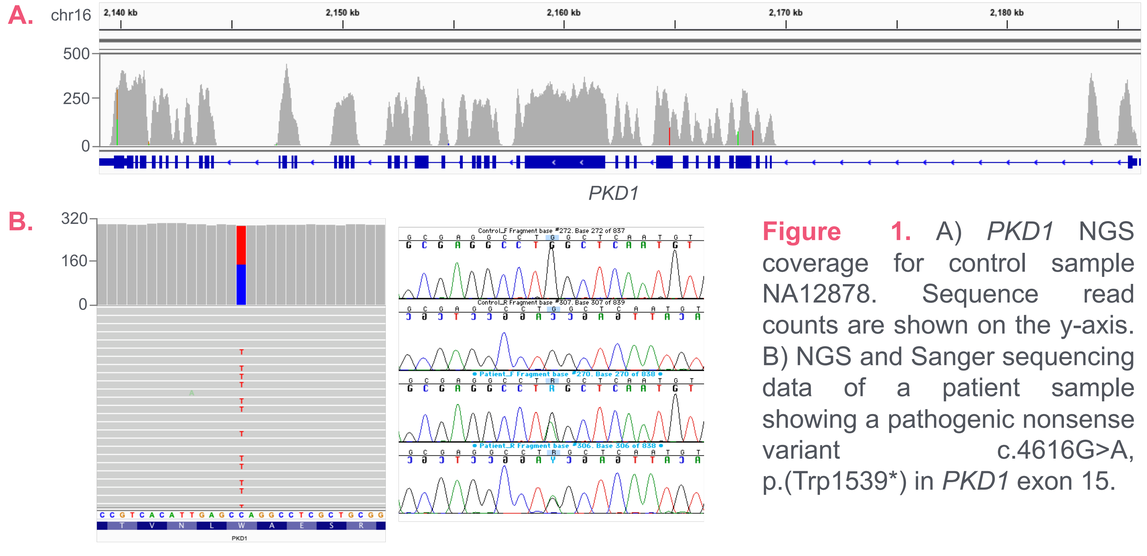

Figure 1. A) *PKD1* NGS coverage for control sample NA12878. Sequence read counts are shown on the y-axis. B) NGS and Sanger sequencing data of a patient sample showing a pathogenic nonsense variant c.4616G>A, p.(Trp1539*) in *PKD1* exon 15.

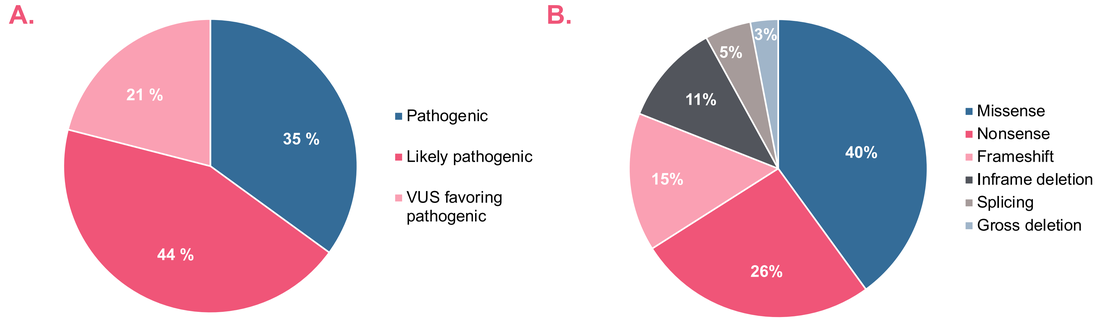

Figure 2. A) Classification and B) mutation type of the diagnostic *PKD1* variants (n=62).

Patient	Previous NGS	BpG Polycystic Kidney Disease Panel	Exon	Classification
1	Neg	c.2012C>G, p.(Ser671*)	10	LP
2	Neg	c.2180T>C, p.(Leu727Pro)	11	P
3	Neg	Negative		
4	Neg	c.2618_2621del, p.(Val873Alafs*24)	11	LP
5	Neg	c.4910T>G, p.(Val1637Gly)	15	VUS
6	Neg	c.8615T>A, p.(Ile2872Asn)	15	VUS
7	Neg	Negative		
8	Neg	c.2534T>C, p.(Leu845Ser)	11	P
9	Neg	c.5411del, p.(Gly1804Alafs*32)	15	LP
10	Neg	Negative		

Table 2. Testing of 10 ADPKD patients with a previous negative test result identified a diagnostic variant in the majority of patients.

Conclusions

- NGS-based panel testing offers good **diagnostic yield** for polycystic and cystic kidney diseases (**54%** in this series)
- Our platform demonstrates **comprehensive coverage** in difficult-to-sequence regions of *PKD1*
- Significant proportion of the identified *PKD1* variants (**81%**) were located within the duplicated region
- The method provides a cost-effective diagnostic tool for simultaneous detection of sequence and copy number variants

Blueprint Genetics

Conflict of interest statement: All authors are employed by Blueprint Genetics.