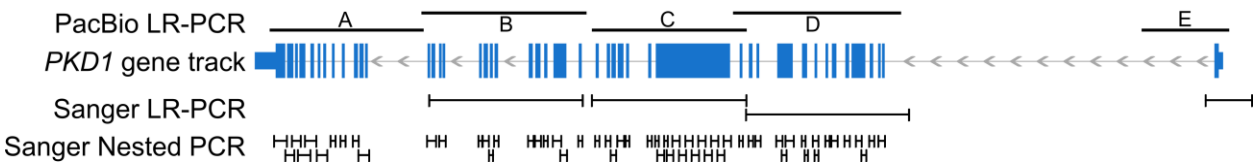
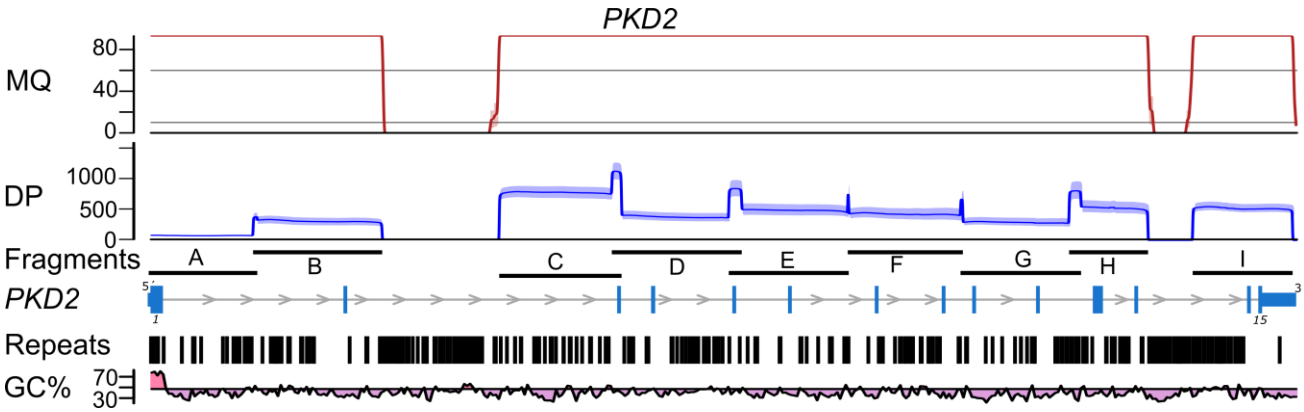
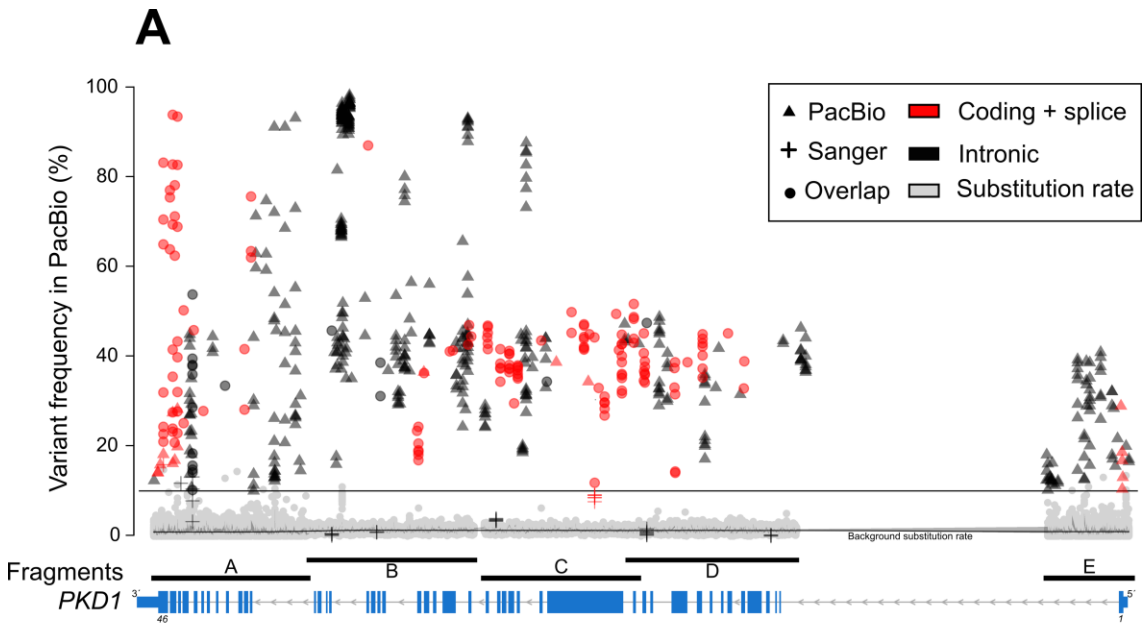
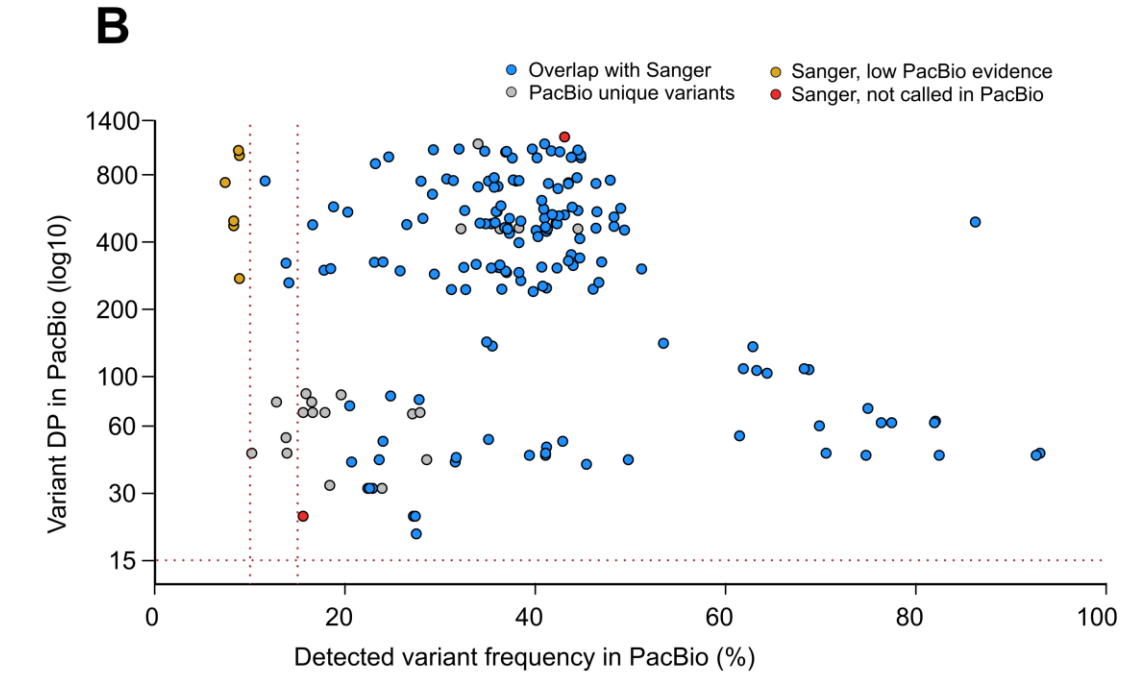


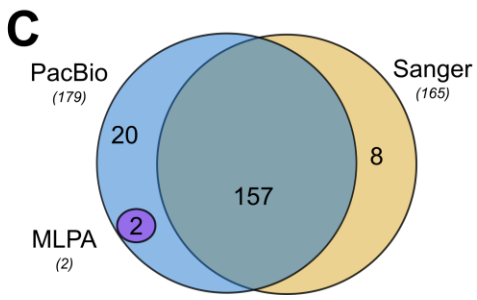
Supp. Figure S1: Enrichment of *PKD1* gene (NM_001009944.2) with 5 LR-PCR fragments used for PacBio, and 4 for Sanger sequencing followed by nested-PCR amplification for obtaining Sanger sequencing reads.



Supp. Figure S2: Mapping quality (MQ; in Phred quality scores; values >90 were scaled down for visualization purposes) and sequencing depth (DP; in number of reads) of uniquely aligned sequenced molecules to *PKD2* (NM_000297.3) for 9 LR-PCR fragments amplified. Mapping quality of alignments with even coverage distribution along the amplified fragments (Fragments), including regions with repetitive elements (Repeats) and high GC-content (GC%). Despite fragment A showing lower coverage compared to other fragments, it had sufficient coverage for variant calling within the exon regions, e.g. the first exon of *PKD2* with average coverage ≥71x (min. ≥43x; max. 111x) (Supp. Table S4).

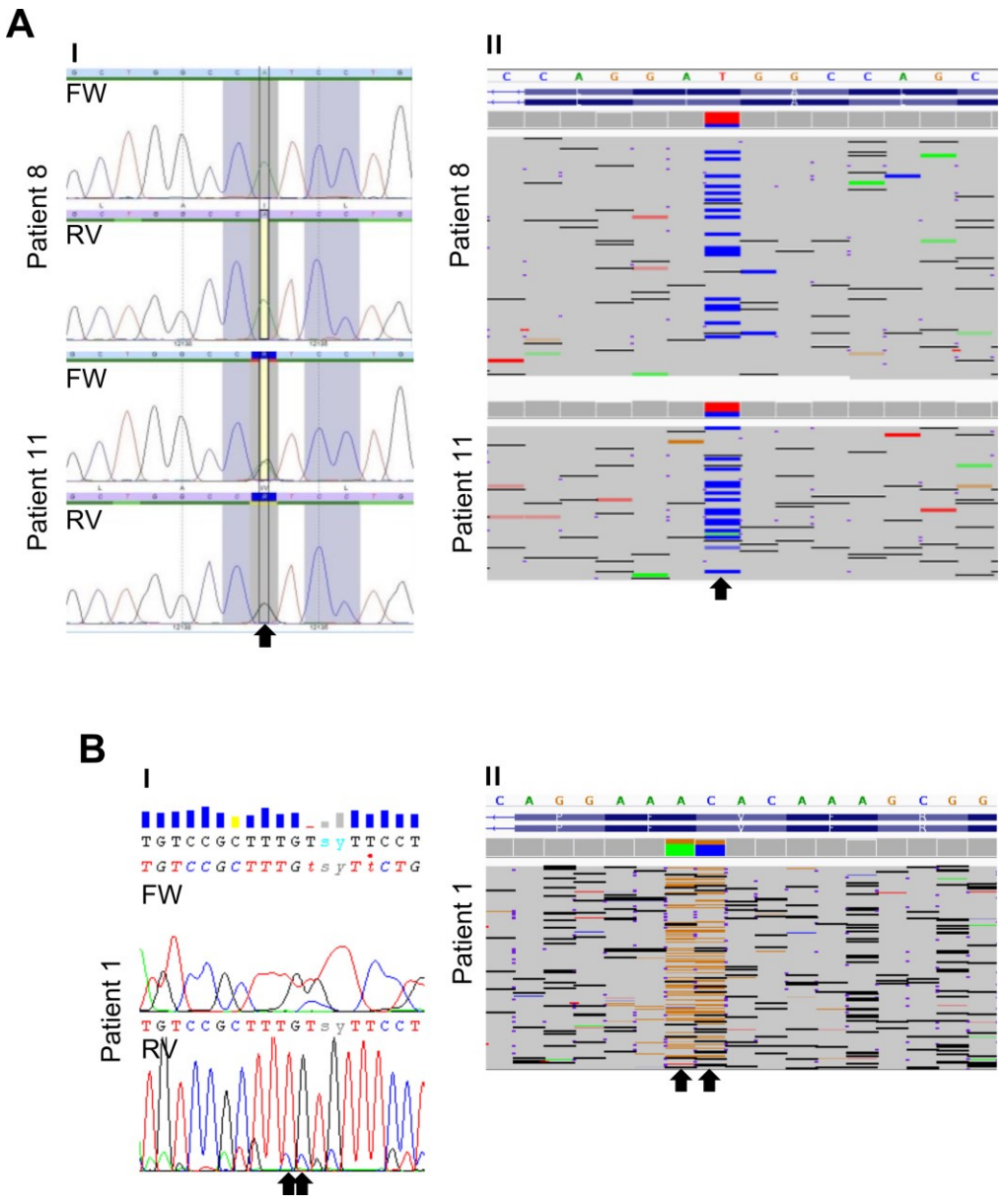

A
Variant frequency in PacBio (%)
100
80
60
40
20
0
PacBio
Sanger
Overlap
Coding + splice
Intronic
Substitution rate
Background substitution rate
Fragments
A
B
C
D
E
PKD1
3'
5'
46
1


B
Overlap with Sanger
PacBio unique variants
Sanger, low PacBio evidence
Sanger, not called in PacBio
Variant DP in PacBio (log10)
1400
800
400
200
100
60
30
15
0
20
40
60
80
100
Detected variant frequency in PacBio (%)

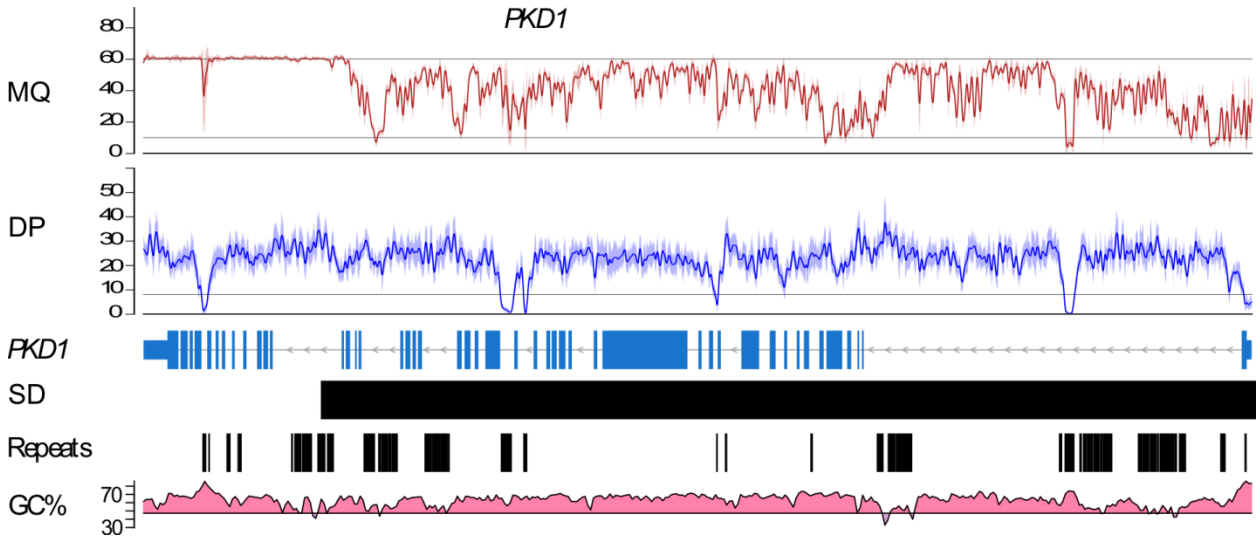

C
PacBio
(179)
Sanger
(165)
MLPA
(2)
20
157
8
2

Supp. Figure S3: Comparison with current diagnostic assay. (*A*.) Small variants identified by PacBio were randomly distributed along the length of the amplified regions of *PKD1* gene locus (x axis; NM_001009944.2) showing no blind spots where no variants could not be detected. When compared to Sanger sequencing results, variants detected with both diagnostic approaches distributed over *PKD1* gene could be visualized (circles). The variant frequency (y axis) of variants solely detected by Sanger (crosses) or PacBio (triangles) which are majorly located in regions where no Sanger sequencing data was available, mostly introns (black). The average per-base mismatch rates (grey dots and average trend line) show the high sensitivity of our long-read sequencing method to detect substitutions. To highlight variants of high significance, coding or splice site variants (red) were colored to be visually identified from other intronic variants (black). (*B*.) The variant frequencies (x axis) with respect to the applied thresholds (dotted line) of all exonic and splice site variants detected either by Sanger or PacBio show the 2 Sanger variants that were not confirmed in PacBio (red) despite having sufficient sequencing depth (y axis) and variant frequency to be called. However, other 6 variants were potentially considered as PCR artifacts detected in Sanger sequencing despite high sequencing depth (yellow), which showed low variant frequency below the applied frequency thresholds (vertical dotted lines). Variants that passed the applied sequencing-based filtering thresholds and were detected by both diagnostic methods are shown as blue dots, whereas the variants that were detected by PacBio only are shown as grey dots. (*C*.) Overall, the comparison between our long-read sequencing approach with the standard ADPKD diagnostics assays show that from the 179 coding or splice site variants that were detected by PacBio, 2 were large deletions also identified by MLPA, and 157 were small variants also identified by Sanger (yellow). Only 8 small variants were not confirmed by PacBio from which 6 showed low sequencing data support (Supp. Figure S5B).

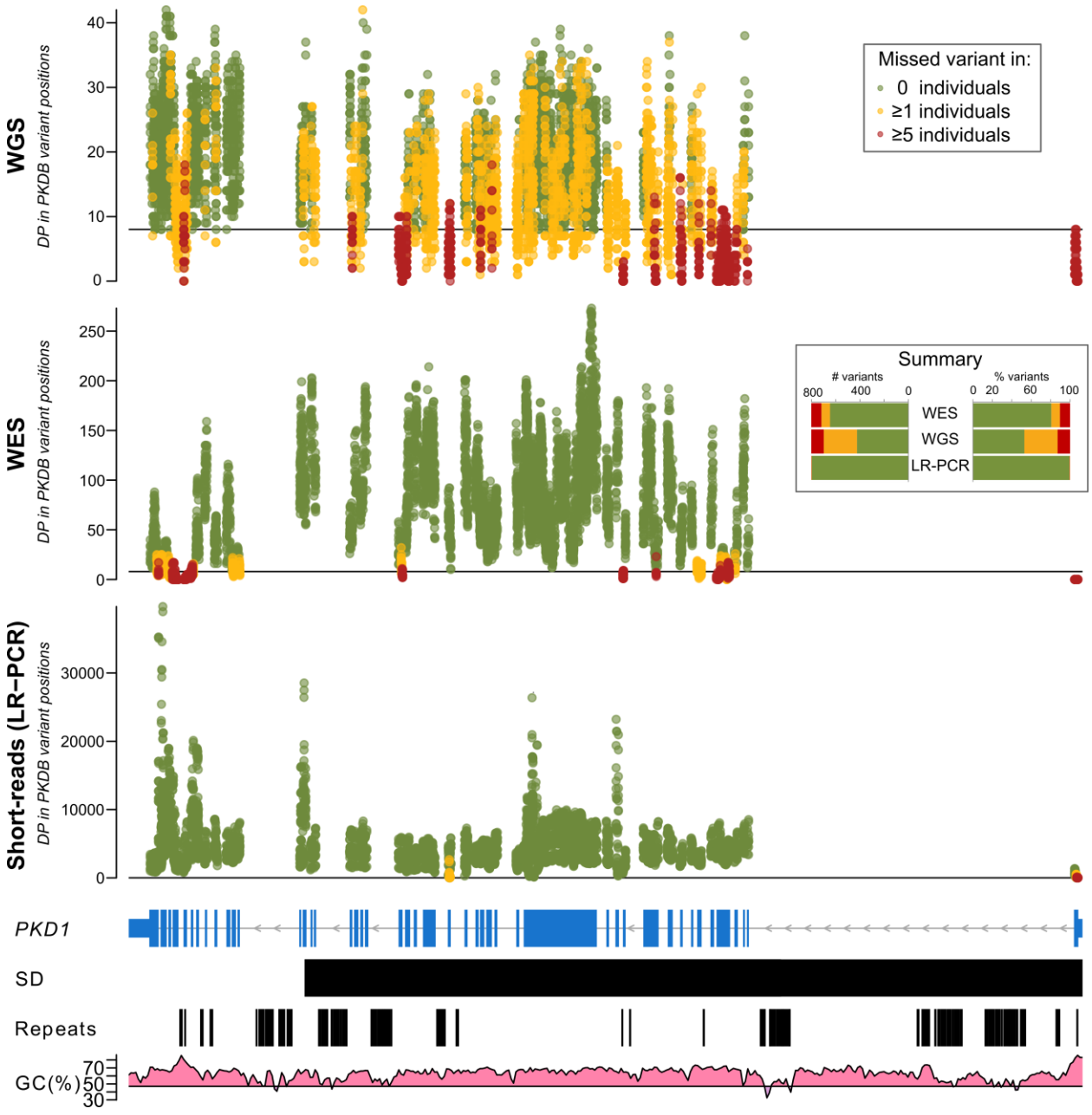
The remaining 20 variants solely identified by PacBio comprehend 17 likely Sanger false negatives, and 3 low-confident calls of the allelic reconstruction of c.6657_6671del (Supp. Table S9).



Supp. Figure S4: Discordances identified between Sanger and PacBio sequencing are a clear example to show the sensitivity of direct long-read sequencing. (*A*.) The polymorphic substitution c.12133A>G correctly identified in patient 11 by Sanger (A_I) and PacBio (A_{II}) was clearly missed by Sanger sequencing in patient 8 (NM_001009944.2). (*B*.) Similarly, the polymorphism c.9195_9196delinsCC was not identified by Sanger sequencing due to unspecific forward signal (B_I), whereas in PacBio at least the substitution c.9195G>C was identified (B_{II}) (RefSeq NM_001009944.2).



Supp. Figure S5: Impact of *PKD1* (NM_001009944.2) inaccessible and unresolvable regions using NGS data. Mapping quality (MQ) and depth (DP) are affected by the presence of repetitive elements (Repeats), percentage of GC content (GC%), and segmental duplications (SD). Inaccessible and unresolvable regions, detected using a publicly available 9 WGS dataset (Sun et al. 2015), disrupt the MQ and DP where SDs from multiple *PKD1* pseudogenes interferes with the alignments of NGS short-reads. Dips in DP and MQ occur where repeats and high GC% overlap, even within gene coding regions (blue).



Supp. Figure S6: Loss of power to detect pathogenic variants with NGS approaches for *PKD1* gene (RefSeq NM_001009944.2). 12% of known pathogenic variants from the ADPKD database (PKDB) would be missed due to poor sequencing depth (red dots) regardless of the NGS sequencing strategy utilized for (WGS or WES) (exons 1, 5, 11, 12, 26, 35, 36, 41, 42 and 46). These poorly sequenced variants appear to be enriched in regions with segmental duplications

(SD), repetitive elements (Repeats), and mainly with high percentage of GC-content (GC%) such is the case for exon 1. Short-read sequencing using LR-PCR enrichment of *PKD1* regions substantially improves the coverage of known pathogenic variants with only a 1.3% of missed variants due to low or poor coverage.

Supp. Table S1: Amplified LR-PCR fragments and sequencing pools for the SMRT library preparation of *PKD1* and *PKD2*.

Amplified fragments	Exons covered	Fragment size (Kb)	Sequencing pool*	Strand**	LR-PCR Primer sequence (5'-3')***
PKD1 A	35-46	7.5	B	FW	GAGCAGGCTCATGGGGCTTTGTAGGAGC
				RV	ACAGCCCGCTGTACCTGAGGACTCG
PKD1 B	22-34	8.1	C	FW	ATGTGAAGAGGTGCCTTGTGTGGT
				RV	TTAAAAACCCGCCCATAATTTCTCACTGC
PKD1 C	14-21	7.6	C	FW	GTTTCCCTGTCTGTTGGGAGGTAAC
				RV	CTGCGTTCACACAGGACAGAACG
PKD1 D	2-13	8	C	FW	CCGAGTAGCTGGAACTACAGTTACACACT
				RV	CACCCAGTTACCTCCCAACAGAC
PKD1 E	1	4.3	A	FW	GCGGAGCGTGAAAAATAGCTCGT
				RV	TACTGCTTTGCTTGACCAGCCTTAAAGA
PKD2 A	1	7.2	B	FW	GCAGGATTCTGTTGCTAGAAGTCAGTGC
				RV	CCTTTCTATCTAGCTTCTTTCCATCCCAGC
PKD2 B	2	7.9	C	FW	CCTGTAACTCCCACCATGGAATGGGC
				RV	AGGTAGGCTTGGAGGGTGCAACTGG
PKD2 C	3	7.5	B	FW	TACCCCTTAAAGATTTTCCTCACA
				RV	CTGTCGATACTCATGCATTGAAA
PKD2 D	3-5	8	C	FW	CTGTGTTGGGGCCTGTGCAGATCAGC
				RV	GGGGGACTTGGTGATGGAACATGTGGC
PKD2 E	5-6	7.4	B	FW	TTGGGACTACATTGACCTCACTAA
				RV	TTCATTCCTGTATCCCCAGTGC
PKD2 F	7-8	7.1	B	FW	GTTTTCTGAGCACCTACTATGTACTTGC
				RV	TAAACCTTGACAACAGTCACCCTCG
PKD2 G	9-10	7.4	B	FW	TGAACTCCAGGGCCTCACACTGTCC
				RV	GCGAACTTTAGACCTGACCTTGCTTTGC
PKD2 H	11-13	4.9	A	FW	AAATGTTGGGGCTGGACATGGTGGC
				RV	ATGCACAAGGAACATTCTTCAGGACG
PKD2 I	14-15	6.2	A	FW	CAGGTCTTTGTCTGCTAAGTCTGA
				RV	TTTGCAAGTGAAATGAAAAACAGT

**Pooling of similar size fragments was performed to optimize loading and sequencing efficiency.*

***FW, forward strand; RV, reverse strand.*

****Primers included a 5' M13-tail (TGTAAAACGACGGCCAGT for FW primers, and CAGGAAACAGCTATGACC for RV primers).*

Supp. Table S2: Per patient targeted PCR amplification and alignment statistics of number of reads and unique molecules, sequenced with SMRT-Seq.

Patient	# Subreads	# Unique molecules (reads)			Base coverage per exon (±20 bp)*		
		PKD1	*PKD2*	*PKD1 Pseudog.*	*>10x*	*>20x*	*>30x*
1	47,842	2,217	5,756	581	100%	100%	93.64%
2	31,889	1,932	3,611	427	100%	100%	99.95%
3	29,200	1,631	2,980	367	100%	100%	99.53%
4	16,929	1,008	1,719	158	100%	99.84%	84.63%
5	31,925	1,584	3,757	236	100%	100%	100%
6	23,344	1,405	2,413	159	100%	99.95%	89.32%
7	26,497	1,268	3,043	276	100%	100%	99.995%
8	36,737	2,208	4,469	202	100%	100%	100%
9	38,224	1,979	4,973	203	100%	99.15%	98.6%
10	44,782	2,667	4,976	684	100%	91.22%	84.12%
11	51,389	2,520	6,054	631	100%	100%	100%
12	42,132	2,357	5,043	412	100%	100%	100%
13	41,092	2,205	4,517	627	100%	100%	100%
14	46,520	2,530	5,010	570	100%	100%	100%
15	46,276	2,880	5,294	254	100%	100%	100%
16	48,624	3,958	5,543	562	100%	100%	100%
17	46,126	3,163	4,948	557	100%	100%	100%
18	44,009	2,671	5,485	334	100%	100%	99.99%
19	45,185	2,684	5,439	332	100%	100%	100%
Mean:	38,902	2,256	4,478	399	100%	99.48%	97.36%

Reads were filtered by read length >500 bp and read quality over 75% within SMRT Analysis. Only properly mapped sequences with mapping quality over >30 that were primary alignments and not PCR or optical duplicates were counted.

**Per exon percentages calculated using the total number of exon (±20 bp) bases from the targeted design (18,259 bp).*

Supp. Table S3: Total number of unique molecules (reads) sequenced for each amplicon.

Amplified fragments		Exons covered	*# Unique Molecules (reads)*			
			Total	Mean	Min.	Max.
PKD1	A	35-46	11,271	593	300	1,580
PKD1	B	22-34	14,439	760	396	1,814
PKD1	C	14-21	23,526	1,238	487	1,845
PKD1	D	2-13	24,281	1,278	519	1,830
PKD1	E	1	1,650	87	35	153
PKD2	A	1	7,982	420	221	737
PKD2	B	2	8,737	460	220	831
PKD2	C	3	24,132	1,270	480	1,748
PKD2	D	3-5	32,089	1,689	610	2,459
PKD2	E	5-6	25,433	1,339	478	2,079
PKD2	F	7-8	23,673	1,246	483	2,197
PKD2	G	9-10	24,187	1,273	436	2,196
PKD2	H	11-13	16,372	862	256	1,547
PKD2	I	14-15	12,007	632	240	886

Supp. Table S4: Base coverage statistics as average min, average mean, and average maximum, calculated from coverage stats retrieved for each patient and exon (±20bp).

Gene	*Exon*	Base coverage		
		Mean	*Min.*	*Max.*
PKD1	1	55	24	91
PKD1	2	422	175	646
PKD1	3	420	168	638
PKD1	4	414	163	627
PKD1	5	408	150	616
PKD1	6	403	150	586
PKD1	7	402	149	577
PKD1	8	402	150	573
PKD1	9	400	147	563
PKD1	10	401	145	564
PKD1	11	400	141	558
PKD1	12	392	139	545
PKD1	13	378	139	528
PKD1	14	756	306	1,224
PKD1	15	754	257	1,256
PKD1	16	734	258	1,186
PKD1	17	739	250	1,199
PKD1	18	744	252	1,201
PKD1	19	743	253	1,198
PKD1	20	738	253	1,187
PKD1	21	713	244	1,155
PKD1	22	550	308	1,252
PKD1	23	564	312	1,349
PKD1	24	559	309	1,358
PKD1	25	555	304	1,361
PKD1	26	551	294	1,366
PKD1	27	548	281	1,432
PKD1	28	547	283	1,436
PKD1	29	544	279	1,442
PKD1	30	543	279	1,446
PKD1	31	544	275	1,525
PKD1	32	542	275	1,528
PKD1	33	532	269	1,521
PKD1	34	523	264	1,504
PKD1	35	59	24	138
PKD1	36	57	21	134
PKD1	37	56	22	133
PKD1	38	54	20	127
PKD1	39	53	21	125
PKD1	40	51	20	122
PKD1	41	51	20	118

PKD1	42	51	19	115
PKD1	43	51	19	113
PKD1	44	51	19	110
PKD1	45	51	21	110
PKD1	46	49	19	108
PKD2	1	71	43	111
PKD2	2	303	111	617
PKD2	3	1.097	388	1,509
PKD2	4	384	122	672
PKD2	5	833	270	1,249
PKD2	6	477	121	866
PKD2	7	452	190	984
PKD2	8	436	167	1,138
PKD2	9	295	49	572
PKD2	10	272	47	632
PKD2	11	514	137	766
PKD2	12	508	134	762
PKD2	13	491	130	748
PKD2	14	507	183	714
PKD2	15	510	188	711

Supp. Table S5: Details of variants identified by the Sanger sequencing or PacBio long-read sequencing approach.

*c. notation***	*p.notation*	*Chr.*	*Pos.*	*# PB*	*# S*	*# O*	*PKDB**	*dbSNP*	*Note*
PKD1									
c.12897C>T	p.(=)	chr16	2,139,743	1	0	0			HC; D; Likely S FN
c.12890A>G	p.(Lys4297Arg)	chr16	2,139,750	1	0	0		rs758833703	HC; D; Likely S FN
c.12780T>C	p.(=)	chr16	2,139,860	1	0	0			HC; D; Likely S FN
c.12630T>C	p.(=)	chr16	2,140,010	8	8	7	LN	rs7203729 (C)	HC; Likely PB FN; Likely S FN
c.12409C>T	p.(=)	chr16	2,140,321	3	3	3	LN	rs79899502 (C)	HC
c.12276A>G	p.(=)	chr16	2,140,454	10	8	8	LN	rs3087632 (C)	HC; Likely S FN
c.12182C>T	p.(Ala4061Val)	chr16	2,140,548	1	1	1		rs372105572	HC; Likely S FP
c.12176C>T	p.(Ala4059Val)	chr16	2,140,554	4	3	3	LN	rs3209986 (C)	HC; Likely S FN
c.12133A>G	p.(Ile4045Val)	chr16	2,140,680	10	8	8	LN	rs10960 (C)	HC; Likely S FN
c.11916C>T	p.(=)	chr16	2,140,972	2	2	2	LN	rs77634115	HC
c.11682C>T	p.(=)	chr16	2,141,454	1	1	1	LN	rs567482892	HC
c.11554del	**p.(Leu3852TrpfsTer93)**	**chr16**	**2,141,581**	**1**	**1**	**1**	**DP**	**rs724159823**	**HC**
c.11412-3C>A	**p.?**	**chr16**	**2,141,910**	**1**	**1**	**1**			**HC**
c.10768C>T	p.(=)	chr16	2,143,865	2	2	2	LN	rs116114803	HC
c.10535C>T	p.(Ala3512Val)	chr16	2,144,176	3	3	3	LN	rs34197769 (C)	HC
c.9957C>T	p.(=)	chr16	2,149,738	1	1	1	LN	rs141101590	HC
c.9397+1G>A	**p.?**	**chr16**	**2,152,061**	**1**	**1**	**1**	**DP**		**HC**
c.9330T>C	p.(=)	chr16	2,152,129	6	6	6	LN	rs12926160 rs144582212 (C)	HC
c.9324del	**p.(Ile3109SerfsTer207)**	**chr16**	**2,152,134**	**1**	**1**	**1**		**rs780284643**	**HC**
c.9195_9196delinsCC	p.(Phe3066Leu)	chr16	2,152,387	5	5	5	LN	rs372874584	HC
c.9195G>C	p.(=)	chr16	2,152,388	1	0	0	LN	rs9935834 (C)	HC; D; Likely S FN
c.9187C>T	p.(Arg3063Cys)	chr16	2,152,396	1	1	1	LN	rs145906459	HC
c.9034_9039del	**p.(Thr3012_Ser3013del)**	**chr16**	**2,152,543**	**1**	**1**	**1**			**HC**
c.8859dup	**p.(Glu2954Ter)**	**chr16**	**2,152,903**	**1**	**1**	**1**			**HC**
c.8440G>A	p.(Gly2814Arg)	chr16	2,153,618	1	1	1	LN	rs149151043	HC
c.8293C>T	**p.(Arg2765Cys)**	**chr16**	**2,153,765**	**1**	**1**	**1**	**LH**	**rs144979397**	**HC**
c.8161+8G>A	p.?	chr16	2,154,491	1	1	1	LN	rs199569003	HC
c.8123C>T	p.(Thr2708Met)	chr16	2,154,537	1	1	1	LN	rs147350387	HC
c.8020C>T	p.(Pro2674Ser)	chr16	2,154,640	1	1	1	LN	rs144557371	HC
c.8017-2_8017-1del	**p.?**	**chr16**	**2,154,643**	**1**	**1**	**1**	**DP**		**HC**

c.7940C>T	**p.(Thr2647Met)**	**chr16**	**2,155,399**	**1**	**1**	**1**		**rs748496650**	**HC**
c.7913A>G	p.(His2638Arg)	chr16	2,155,426	5	5	5	LN	rs9936785 (C)	HC
c.7708T>C	p.(=)	chr16	2,156,021	5	5	5	LN	rs28575767 (C)	HC
c.7441C>T	p.(=)	chr16	2,156,447	7	7	7	LN	rs2003782 (C)	HC
c.7214G>T	**p.(Trp2405Leu)**	**chr16**	**2,156,674**	**1**	**1**	**1**			**HC**
c.7165T>C	p.(=)	chr16	2,156,850	7	7	7	LN	rs2457533 (C)	HC
c.6994_7000dup	**p.(Ala2332GlyfsTer90)**	**chr16**	**2,157,954**	**1**	**1**	**1**	**DP**		**HC**
c.6986G>A	p.(Arg2329Gln)	chr16	2,157,963	1	1	1	LN	rs575211353	HC
c.6670_6673del	p.(Pro2224TrpfsTer17)	chr16	2,158,494	1	0	0			Likely LAA Low Confidence Assembly
c.6657_6671del	**p.(Arg2220_Pro2224del)**	**chr16**	**2,158,496**	**1**	**1**	**1**			**HC; Complete PB LAA Confident Assembly**
c.6666_6667del	p.(Leu2223AlafsTer38)	chr16	2,158,500	1	0	0			Likely LAA Low Confidence Assembly
c.6656_6659del	p.(Pro2219ArgfsTer22)	chr16	2,158,508	1	0	0			Likely LAA Low Confidence Assembly
c.6488G>A	p.(Arg2163Gln)	chr16	2,158,680	1	0	0	LN	rs145217118	HC; D; Likely S FN
c.6223_6224insTT	**p.(Arg2075LeufsTer42)**	**chr16**	**2,158,944**	**0**	**1**	**0**			**Likely PB FN**
c.5763G>A	p.(=)	chr16	2,159,405	2	2	2	LN	rs2575313 (C)	HC
c.5172C>T	p.(=)	chr16	2,159,996	7	7	7	LN	rs9935526 (C)	HC
c.4968_4969delinsC	**p.(Arg1657GlyfsTer65)**	**chr16**	**2,160,198**	**1**	**1**	**1**			**HC**
c.4968T>C	p.(=)	chr16	2,160,200	1	0	0		rs777909326	HC; D; Likely S FN
c.4917C>T	p.(=)	chr16	2,160,251	1	1	1	LN	rs148852027	HC
c.4674G>A	p.(=)	chr16	2,160,494	1	1	1	LN	rs79884128 (C)	HC
c.4665A>C	p.(=)	chr16	2,160,503	1	7	1	LN	rs71385734 (C)	HC; D; Likely S FP
c.4475G>C	p.(Arg1492Pro)	chr16	2,160,693	1	1	1			HC
c.4248dup	**p.(Gly1417TrpfsTer14)**	**chr16**	**2,160,919**	**1**	**1**	**1**			**HC**
c.4195T>C	p.(Trp1399Arg)	chr16	2,160,973	5	5	5	LN	rs116092985 (C)	HC
c.3643C>G	p.(Leu1215Val)	chr16	2,161,525	1	1	1		rs144338515	HC
c.3513C>G	p.(=)	chr16	2,161,655	1	1	1	LN	rs143784787	HC
c.3375C>T	p.(=)	chr16	2,161,793	5	5	5	LN	rs74331768 (C)	HC
c.3372C>T	p.(=)	chr16	2,161,796	5	5	5	LN	rs75510884 (C)	HC
c.3275T>C	p.(Met1092Thr)	chr16	2,162,361	5	5	5	LN	rs2549677 (C)	HC
c.3111A>G	p.(=)	chr16	2,162,839	3	3	3	LN	rs2099534 (C)	HC
c.3100A>G	p.(Asn1034Asp)	chr16	2,162,850	1	1	1		rs369180760	HC
c.3063T>C	p.(=)	chr16	2,162,887	5	5	5	LN	rs2369068 (C)	HC
c.2730C>T	p.(=)	chr16	2,164,294	2	2	2	LN	rs35965348 (C)	HC
c.2700G>A	p.(=)	chr16	2,164,324	2	2	2	LN	rs35667726 (C)	HC
c.2694A>C	p.(=)	chr16	2,164,330	2	2	2	LN	rs142357713	HC
c.2681_2690del	**p.(Phe894Ter)**	**chr16**	**2,164,333**	**1**	**1**	**1**			**HC**

c.2269del	**p.(Gln757SerfsTer28)**	**chr16**	**2,164,754**	**1**	**1**	**1**		**rs775710328**	**HC**
c.2109C>T	p.(=)	chr16	2,164,915	1	1	1	LN	rs527655141	HC
c.1850-4A>G	p.?	chr16	2,165,630	7	7	7	LN	rs35929659 (C)	HC
c.1602C>T	p.(=)	chr16	2,166,838	1	1	1	LN	rs759092782	HC
c.1286G>T	**p.(Trp429Leu)**	**chr16**	**2,167,589**	**1**	**1**	**1**	**HLP**		**HC**
c.1261C>T	**p.(Arg421Cys)**	**chr16**	**2,167,614**	**1**	**1**	**1**			**HC**
c.182C>T	p.(Pro61Leu)	chr16	2,185,509	1	0	0	LN		HC; D; Likely S FN
c.129C>T	p.(=)	chr16	2,185,562	1	0	0			HC; D; Likely S FN
c.122C>T	p.(Pro41Leu)	chr16	2,185,569	2	0	0			HC; D; Likely S FN
c.96C>T	p.(=)	chr16	2,185,595	1	0	0			HC; D; Likely S FN
PKD2									
c.53C>T	p.(Pro18Leu)	chr4	88,928,938	1	0	0			D; Likely S FN
c.83G>C	p.(Arg28Pro)	chr4	88,928,968	11	9	9	LN	rs1805044 (C)	HC; Likely S FN
c.361G>T	p.(Gly121Cys)	chr4	88,929,246	1	1	1		rs371898195	HC
c.420G>A	p.(=)	chr4	88,929,305	5	5	5	LN	rs2728118 (C)	HC
c.568G>A	p.(Ala190Thr)	chr4	88,929,453	1	1	1	LN	rs117078377 (C)	HC
c.720C>G	p.(=)	chr4	88,957,382	0	2	0			D; Likely S FP
c.1459T>C	p.(Tyr487His)	chr4	88,967,933	1	1	1	LN	rs201328200	HC

** Bold=Pathogenic; DP=Definitely Pathogenic; HLP=Highly Likely Pathogenic; LH=Likely Hypomorphic; LN=Likely Neutral; C=Common; HC=High-Confident; FP=False Positive; FN=False Negative; D=Discordant; PB=PacBio; S=Sanger; O=Overlap.*

+ Deletion-insertions were manually inspected and HGVS description modified to avoid variant redundancies or undesired complex descriptions by clustering of neighbouring variants (e.g. c.9195_9196delinsCC and c.6994_7000dup). Individual substitutions c.9195G>C and/or c9196T>C from the polymorphic site c.9195_9196delinsCC were removed as redundant variants if c.9195_9196delinsCC was identified as well in the same patient. The pathogenic duplication c.6994_7000dup was originally detected as c.6982_6983delinsAGCGGCTG when reconstructing the PKD1 *allelic sequence and was split into two independent descriptions c.6994_7000dup, and c.6982del.*

*** The 177 coding or splice site variants listed in this table were either reported by Sanger sequencing or detected by our direct long-read sequencing approach in at least one patient within the cohort of 19 screened ADPKD patients. 157 variants are high-confidence variants identified on the same set of patients by both approaches. 8 Sanger variants were not confirmed by PacBio, from which 6 had low sequencing data support. The remaining 20 variants were solely detected by PacBio from which 3 are low-confident variants called by the reconstructions of allelic sequences of c.6657_6671del.*

*** RefSeq NM_001009944.2 or NM_000297.3 for PKD1 and PKD2, respectively.*

Supp. Table S6: Amplified LR-PCR primer sequences for Sanger sequencing.

Exons covered	Fragment size (Kb)	Strand*	LR-PCR Primer sequence (5'-3')**
1	2.2	FW	CGCAGCCTTACCATCCACCT
		RV	TCATCGCCCCTTCCTAAGCA
2-13	7.9	FW	CCGAGTAGCTGGAACTACAGTTACACACT
		RV	CACCCAGTTACCTCCCAACAGAC
14-21	7.6	FW	GTTTCCCTGTCTGTTGGGAGGTAAC
		RV	CTGCGTTCACACAGGACAGAACG
22-34	7.5	FW	CCGTGTAGAGAGGAGGGCGTGTGCAAGGA
		RV	TCGGCAAGGACCTGCTGGATCAG

*FW, forward strand; RV, reverse strand.
**Primers included a 5' M13-tail (TGTAAAACGACGGCCAGT for FW primers, and CAGGAAACAGCTATGACC for RV primers).

Supp. Table S7: Nested PCR primer sequences for *PKD1* Sanger sequencing.

Exons covered	Size (bp)	Strand*	Nested PCR Primer sequences (5'-3')**
1 *(frag. 1)*	256	FW	CCTGAGCTGCGGCCTCCG
		RV	CAGTTGACGCGGCAGGCG
1 *(frag. 2)*	216	FW	TGCGAGCCCCCCTGCCTC
		RV	AACCCGCCCACGCCCGCCCGTCC
2-3	340	FW	TAGGGGCTCTGGCCCTGAC
		RV	CCAGCCAGGACCCCACCCAAAG
4	266	FW	CATAGACCCTTCCCACCAG
		RV	CCTGGCTGGGAAGGACAGA
5 *(frag. 1)*	389	FW	TGGAGCCAGGAGGAGCAGAA
		RV	CAGAGGGACAGGCAGGCAAA
5 *(frag. 2)*	431	FW	AGCCCTCCAGTGCCTCCTTT
		RV	GCACGGCCGTCACGTGATAG
6	280	FW	ACCGTTGACACCCTCGTTCC
		RV	CTCTGCCCCAGTGCTTCAG
7	329	FW	CTGTGAGGGTGGGAGGATGG
		RV	CTAACCACAGCCAGCGTCTC
8	226	FW	GCGGCTCGGTCCCCAGTCT
		RV	GGAGGGCAGGTTGTAGAACGTG
9 *(frag. 1)*	295	FW	GGAGTCTGGGCTTCAGGCTG
		RV	CTGGGAACCACTCTGGTGGC
9 *(frag. 2)*	228	FW	GGAGTCTGGGCTTCAGGCTG
		RV	CACCCACCACCCAGAGTCCC
10 *(frag. 1)*	184	FW	GGCCTGTGGGCAAATCAGGG
		RV	TGGGGGTGGCAGGAGGCGTC
10 *(frag. 2)*	201	FW	AGGGGGACGCTGGTGCCCTG
		RV	GGGAACAGACCCAGGTCAGG
11 *(frag. 1)*	425	FW	GTCCACGGGCCATGACCG
		RV	CCAGCCACTGGGGAGACCAC
11 *(frag. 2)*	257	FW	GGCAGAGGTGGGCAATGG
		RV	AGCCGGGCACGAAGGTGGC
11 *(frag. 3)*	326	FW	GTGTCAGCGCCCGCTTTG
		RV	CTGTGTGAGCACCCTGTCTGC
12	241	FW	GTGTGTCCAGGAGGCGA
		RV	AGAGGTGAAGGTGGAGC
13	244	FW	CTGCCACCTGGGCTCACTG
		RV	TGCCCACCCCAAACCGGC
14	198	FW	CTCACTGCGTCCCACCGC
		RV	CTGAAAGGCAGTGGCCCC
15 *(frag. 1)*	382	FW	TGGGGAGCAGGTGGGGGTGC
		RV	AGACGCGCACATCCGCCTGGGCCG
15 *(frag. 2)*	363	FW	CGTGCGCCTGGAGGTCAAC
		RV	GGCTGCGTGGGGATGCAG
15 *(frag. 3)*	373	FW	CGTGCTGGTCTTCGTCCTGG
		RV	TGTAGCGGTAGGGGAACGG
15 *(frag. 4)*	380	FW	GTTTGTGCAGCTCGGGGAC
		RV	AAGCGTGGGTGACCTCCG
15 *(frag. 5)*	376	FW	CCCGCCAGCTACCTGTGG
		RV	GCGGAGCCCACCTCGTTC
15 *(frag. 6)*	378	FW	CTTCCGCTCCGTGGGCAC
		RV	GGAGGCGGCCACCATCAG
15 *(frag. 7)*	374	FW	AGCGCCTGGGCCGACTGCAC
		RV	AGCTGCCCCCAAAAGGGC
15 *(frag .8)*	363	FW	GAGCCCGGAGGCAGCTTC
		RV	GGGAGCACCTCGGGGTTG
15 *(frag. 9)*	377	FW	AGCTGTCACCTTCCGCCTG
		RV	GCACCTGGATCTCCAACAGCC
15 *(frag. 10)*	340	FW	GCTGGTCATCCTGTCGGGC
		RV	CACCAGGTTGGAGGCGTTC
15 *(frag. 11)*	567	FW	CCAGGGCCGAGCACTCCTAC
		RV	TGGGGTCGTAGGACTCGCTC
15 *(frag. 12)*	196	FW	CGCCTGGTGCCCATCATTG
		RV	GGACGGGTGAGGGGCATG
16	230	FW	AGGCCACGTCGCCCCTTG
		RV	GAGGCTGGGCTGTCCAAGG
17	203	FW	GAGGTAACCCCACTCCCACG

		RV	ATCCCCAGCCCGCCCACAC
18	353	FW	AGAGGGTTGCGCCCCCTC
		RV	ATCCCGCTGCTCCCCCCACGC AGG
19	286	FW	TCCCGTGATGCCGTGGGG
		RV	CAGGTGGCAGTCTCGGGG
20	260	FW	CCACCTGCTCACCACCCC
		RV	GCAGGGGTACAGGTCTTGGTC CC
21	226	FW	GCGCTGCTGACAGCTTGC
		RV	ATGCGGGGCAGGGTGAGC
22	220	FW	AGTGGGGCCAGGAGCGGG
		RV	GGGCGGGTGGCATGGGGC
23 *(frag. 1)*	349	FW	CCCTCCCTCTACCTCCCTGTC
		RV	CACTGAGGTTGGCCAGGGC
23 *(frag. 2)*	431	FW	GGGCCTGGCTGCCACTTC
		RV	AAGGCCAGGGGGCCGCGTG
24	222	FW	CAGGCGTGTGACCTGCGC
		RV	TGCCCTGCCCTGCCAGGCTG
25	313	FW	CTGGGCTCACGTCCGCTAC
		RV	GCTCCCAGGAGCACAGGGTC
26	289	FW	GAGAAGGCACAGCTTGCACG
		RV	AGAGCAGGGGAGGCCCTG
27	240	FW	GCAGACCGAGCCTCCCAC
		RV	AGGGGCAGAGCTTGGCAG
28	217	FW	TGCGAGCCTGACCTCCCTC
		RV	CCAACCTCCCACGGAGTGG
29	316	FW	TTGGGCAGGGTGGTCCTG
		RV	GGAAGGGCTGGGCAGGAAG
30	202	FW	CAGCCTCACCTGTGTGGCC
		RV	TCCATTCCCAGTACTCCCGG
31-32	338	FW	GAGCAGGTCTGAGCTGCCG
		RV	GCACCAGGGCTCGAGGTTTC
33-34	473	FW	GGTGGGCTGTGTGTGTGAC
		RV	CCCCTCCTCTGGCAATCC
35-36	576	FW	CAGGTTAACATGGGCTTGG
		RV	GAGGGGGTGGCTTCAGAG
37	341	FW	GGTAGGCTACAGGCCTCCAT
		RV	GAGGTGGGAGACAAGAGACG
38	304	FW	AAAGCCCTGCTGTCACTGTGG
		RV	TAGGGTCTGGCTGGACTAAAG
39	301	FW	GGGTCTCTGGTGGCCGCTCAC
		RV	ATGCCAGAGCTCCGCTAAAGG
40-41	565	FW	GCAGGAAACACTCCTGTTGG
		RV	GCTCCTGGCTGGTGACTG
42	608	FW	GAGCCCACCCTCACTCCTC
		RV	AACAGCAGCAGGCACACCT
43	660	FW	CTCTGCTCACCTCGGTACG
		RV	CGGACGAGAAATCTGTCTGC
44	361	FW	CGCCGCTTCACTAGCTTCGAC
		RV	AGTCCCAGGGCACAGCACCAA
45	486	FW	GGGAGGGCGTCTTAGCTC
		RV	CACAGGGGCTCAGTCAGTC
46	680	FW	CAAGGTCAAGGAGGTGGGTA
		RV	TTGACAGCGGCAGAAAGTAA

[*]FW, forward strand; RV, reverse strand.

[**]Primers included a 5' M13-tail (TGTAAAACGACGGCCAGT for FW primers, and CAGGAAACAGCTATGACC for RV primers).

Supp. Table S8: Nested PCR primer sequences for *PKD2* Sanger sequencing.

Exon covered	Size (bp)	Strand*	Oligo sequence (5'-3')**
1 *(frag. 1)*	421	FW	GAAAGGAACATGGCTCCTGA
		RV	ACCTCCTCCTCCTCCTCCTC
1 *(frag. 2)*	454	FW	CCCTTCTCCTCCGCTCTC
		RV	CGTTCTGGTTCGTGCATCT
02	320	FW	GTGCTTTATTTTCCCTTTTG
		RV	GGTGCATACACACTTCCTTT
03	330	FW	CTTTGTGAAGGCTGCTGGT
		RV	TCCTGTCGATACTCATGCATTG
04	435	FW	TTGGTTATGCAAACGATG
		RV	GAATGGTGGGAGTTCAGAG
05	414	FW	CCTCAAGTGTTCCACTGATT
		RV	GTAGCTAACTGCAGGCAAAG
06	401	FW	CTGGCTGTATTCATGTGTTG
		RV	AATGCTGAGGAGATCAAAGA
07	336	FW	GGTAAGTTTCATATTTCTAAAACACT
		RV	TTCCATGATTTTGTGGAACT
08	329	FW	CACACCATTTTGTTTATCCA
		RV	TTCTTGAGAAGCAGTGACAA
09	277	FW	TGCATCAACTAGTGGACATT
		RV	GAGAAGACAAGGATTTACGAAG
10	259	FW	TTCCAAATTATGTTTCTTCCTT
		RV	AAAATCTGGGTGAAACAATG
11	278	FW	AAAACAGATGCAAAAGGAGA
		RV	CCAGGAATTTATCTTTAGAAGC
12	266	FW	GAACTGGGTACAAGGAATGA
		RV	TTTGATACATCTGTGGTGTTG
13	322	FW	GTCCTTGGTGAGGCTTCT
		RV	CTGGTCTCATGTGGACTCTT
14	285	FW	AAAGACAATGACAAGCACTTT
		RV	TCATTAATAACACCATGCTCA
15	417	FW	ATTATTTGGTCCCTGGACTT
		RV	GTGCTTGTTACAGCAATTCA

**FW, forward strand; RV, reverse strand.*
***Primers included a 5' M13-tail (TGTAAAACGACGGCCAGT for FW, and CAGGAAACAGCTATGACC for RV).*

Supp. Table S9: Discordant variant calls between Sanger sequencing and PacBio long-read sequencing approach.

c. notation	*p.notation*	*Chr*	*pos*	*# PB*	*# S*	*# O*	*PKDB**	*dbSNP*	*Note*
PKD1									
c.12897C>T	p.(=)	chr16	2,139,743	1	0	0			HC; D; Likely S FN
c.12890A>G	p.(Lys4297Arg)	chr16	2,139,750	1	0	0		rs758833703	HC; D; Likely S FN
c.12780T>C	p.(=)	chr16	2,139,860	1	0	0			HC; D; Likely S FN
c.12630T>C	p.(=)	chr16	2,140,010	8	8	7	LN	rs7203729 (C)	HC; Likely PB FN; Likely S FN
c.12276A>G	p.(=)	chr16	2,140,454	10	8	8	LN	rs3087632 (C)	HC; Likely S FN
c.12176C>T	p.(Ala4059Val)	chr16	2,140,554	4	3	3	LN	rs3209986 (C)	HC; Likely S FN
c.12133A>G	p.(Ile4045Val)	chr16	2,140,680	10	8	8	LN	rs10960 (C)	HC; Likely S FN
c.9195G>C	p.(=)	chr16	2,152,388	1	0	0	LN	rs9935834 (C)	HC; D; Likely S FN
c.6670_6673del	p.(Pro2224TrpfsTer17)	chr16	2,158,494	1	0	0			Likely LAA Low Confidence Assembly
c.6666_6667del	p.(Leu2223AlafsTer38)	chr16	2,158,500	1	0	0			Likely LAA Low Confidence Assembly
c.6656_6659del	p.(Pro2219ArgfsTer22)	chr16	2,158,508	1	0	0			Likely LAA Low Confidence Assembly
c.6488G>A	p.(Arg2163Gln)	chr16	2,158,680	1	0	0	LN	rs145217118	HC; D; Likely S FN
c.6223_6224insTT	**p.(Arg2075LeufsTer42)**	**chr16**	**2,158,944**	**0**	**1**	**0**			**Likely PB FN**
c.4968T>C	p.(=)	chr16	2,160,200	1	0	0		rs777909326	HC; D; Likely S FN
c.4665A>C	p.(=)	chr16	2,160,503	1	7	1	LN	rs71385734 (C)	HC; D
c.182C>T	p.(Pro61Leu)	chr16	2,185,509	1	0	0	LN		HC; D; Likely S FN
c.129C>T	p.(=)	chr16	2,185,562	1	0	0			HC; D; Likely S FN
c.122C>T	p.(Pro41Leu)	chr16	2,185,569	2	0	0			HC; D; Likely S FN
c.96C>T	p.(=)	chr16	2,185,595	1	0	0			HC; D; Likely S FN
PKD2									
c.53C>T	p.(Pro18Leu)	chr4	88,928,938	1	0	0			D; Likely S FN
c.83G>C	p.(Arg28Pro)	chr4	88,928,968	11	9	9	LN	rs1805044 (C)	HC; Likely S FN
c.720C>G	p.(=)	chr4	88,957,382	0	2	0			D; Likely S FP

** Bold=Pathogenic; LN=Likely Neutral; C=Common; HC=High-Confident; FP=False Positive; FN=False Negative; D=Discordant; PB=PacBio; S=Sanger; O=Overlap.*

*** RefSeq NM_001009944.2 or NM_000297.3 for PKD1 and PKD2, respectively.*

[+] *Discordances shown in this table are represented in Supp. Figure S2B as yellow dots for likely S FP variants, red dots for likely PB FN variants, and grey dots for Likely S FN.*

Supp. Table S10: Variants detected in homopolymer stretches using long-read sequencing approach.

*c. notation***	*p.notation*	*Chr*	*pos*	*# PB*	*# S*	*# O*	*PKDB**	*dbSNP*	*Note*
PKD1									
c.12530del	p.(Pro4177HisfsTer21)	chr16	2,140,109	3	0	0		rs767438361	HS; Likely PB Artifact
c.12518del	p.(Pro4173ArgfsTer25)	chr16	2,140,121	5	0	0		rs778397103	HS; Likely PB Artifact
c.12445-3del	p.?	chr16	2,140,197	5	0	0		rs770813339	HS; Likely PB Artifact
c.12139-5del	p.?	chr16	2,140,595	2	0	0	LN	rs146430229	HS; Likely PB Artifact
c.12085del	p.(Val4029SerfsTer10)	chr16	2,140,727	1	0	0		rs781278135	HS; Likely PB Artifact
c.11713-5del	p.?	chr16	2,141,179	8	0	0			HS; Likely PB Artifact
c.11240del	p.(Pro3747HisfsTer79)	chr16	2,142,509	6	0	0			HS; Likely PB Artifact
c.10948del	p.(His3650ThrfsTer34)	chr16	2,143,612	1	0	0			HS; Likely PB Artifact
c.10822-8del	p.?	chr16	2,143,746	12	0	0	LN	rs373684171 rs9924796 (C)	HS; Likely PB Artifact
c.10745del	p.(Pro3582ArgfsTer3)	chr16	2,143,887	11	0	0	DP	CD076868	HS; Likely PB Artifact
c.9518del	p.(Pro3173ArgfsTer143)	chr16	2,150,446	1	0	0		rs772608027	HS; Likely PB Artifact
c.9176del	p.(Pro3059GlnfsTer15)	chr16	2,152,406	10	0	0		rs759773922	HS; Likely PB Artifact
c.9097del	p.(Leu3033TrpfsTer41)	chr16	2,152,485	1	0	0			HS; Likely PB Artifact
c.8586del	p.(Ile2863SerfsTer12)	chr16	2,153,471	1	0	0	DP		HS; Likely PB Artifact
c.8427del	p.(Glu2810ArgfsTer65)	chr16	2,153,630	1	0	0		rs746703342	HS; Likely PB Artifact
c.8019del	p.(Ser2675AlafsTer10)	chr16	2,154,640	1	0	0			HS; Likely PB Artifact
c.7864-3del	p.?	chr16	2,155,477	1	0	0		rs756848270	HS; Likely PB Artifact
c.7622del	p.(Pro2541ArgfsTer79)	chr16	2,156,172	1	0	0		rs538031465	HS; Likely PB Artifact
c.7401del	p.(Asn2468ThrfsTer152)	chr16	2,156,486	7	0	0		rs745812853	HS; Likely PB Artifact
c.6759del	p.(Glu2254SerfsTer60)	chr16	2,158,408	18	0	0			HS; Likely PB Artifact
c.6469del	p.(Leu2157CysfsTer4)	chr16	2,158,698	1	0	0			HS; Likely PB Artifact
c.5824del	p.(Arg1942AlafsTer7)	chr16	2,159,343	18	0	0		rs780100275	HS; Likely PB Artifact
c.5784del	p.(Glu1929ArgfsTer20)	chr16	2,159,383	1	0	0			HS; Likely PB Artifact
c.4485del	p.(Ala1496ProfsTer38)	chr16	2,160,682	1	0	0		rs578064441	HS; Likely PB Artifact
c.4269del	p.(Thr1424ProfsTer8)	chr16	2,160,898	1	0	0			HS; Likely PB Artifact
c.4220del	p.(Pro1407ArgfsTer25)	chr16	2,160,947	3	0	0		rs140412120	HS; Likely PB Artifact
c.4069del	p.(Leu1357TrpfsTer9)	chr16	2,161,098	16	0	0	DP	CD085910	HS; Likely PB Artifact
c.3684del	p.(Val1229TrpfsTer44)	chr16	2,161,483	6	0	0	DP	rs781384791	HS; Likely PB Artifact
c.3240del	p.(Ser1081ArgfsTer23)	chr16	2,162,395	2	0	0		rs777145613	HS; Likely PB Artifact
c.3099del	p.(Asn1034MetfsTer4)	chr16	2,162,850	12	0	0	DP	rs372461622	HS; Likely PB Artifact
c.2854-5del	p.?	chr16	2,163,297	19	0	0		rs114846412	HS; Likely PB Artifact
c.2823del	p.(Glu942ArgfsTer9)	chr16	2,164,200	1	0	0		rs767322474	HS; Likely PB Artifact

c.2494del	p.(Arg832AlafsTer66)	chr16	2,164,529	19	0	0	DP		HS; Likely PB Artifact
c.2222del	p.(Pro741ArgfsTer44)	chr16	2,164,801	1	0	0		rs779605081	HS; Likely PB Artifact
c.2085del	p.(Ala696ArgfsTer89)	chr16	2,165,390	19	0	0		rs760496344	HS; Likely PB Artifact
c.2037del	p.(Tyr680MetfsTer105)	chr16	2,165,438	13	0	0			HS; Likely PB Artifact
c.1987del	p.(Gln663ArgfsTer122)	chr16	2,165,488	11	0	0	DP		HS; Likely PB Artifact
c.1914del	p.(Ala639ArgfsTer146)	chr16	2,165,561	11	0	0	DP	rs777208671	HS; Likely PB Artifact
c.1386-3del	p.?	chr16	2,167,056	11	0	0			HS; Likely PB Artifact
c.1221del	p.(Ser408ArgfsTer57)	chr16	2,167,653	4	0	0		rs768893401	HS; Likely PB Artifact
c.771del	p.(Thr258ProfsTer32)	chr16	2,168,221	11	0	0			HS; Likely PB Artifact
c.755del	p.(Pro252ArgfsTer38)	chr16	2,168,237	19	0	0			HS; Likely PB Artifact
c.198del	p.(Ala67ArgfsTer6)	chr16	2,185,492	1	0	0			HS; Likely PB Artifact
c.108del	p.(Cys37AlafsTer36)	chr16	2,185,582	14	0	0			HS; Likely PB Artifact
c.78del	p.(Arg28AlafsTer45)	chr16	2,185,612	13	0	0			HS; Likely PB Artifact
PKD2									
c.128del	p.(Pro43ArgfsTer74)	chr4	88,929,012	2	0	0	DP	CD982885	HS; Likely PB Artifact
c.203del	p.(Pro68ArgfsTer49)	chr4	88,929,087	19	0	0	DP	rs751221093	HS; Likely PB Artifact
c.538del	p.(Leu180TrpfsTer53)	chr4	88,929,422	6	0	0			HS; Likely PB Artifact
c.783del	p.(Val262CysfsTer55)	chr4	88,957,444	6	0	0		rs766343471	HS; Likely PB Artifact
c.1003del	p.(Gln335ArgfsTer3)	chr4	88,959,561	18	0	0			HS; Likely PB Artifact

** DP=Definitely Pathogenic; LN=Likely Neutral; C=Common; HS=Homopolymer Stretch; PB=PacBio; O=Overlap.*

*** RefSeq NM_001009944.2 or NM_000297.3 for PKD1 and PKD2, respectively.*