



Published in final edited form as:

*J Med Genet.* 2016 March ; 53(3): 208–214. doi:10.1136/jmedgenet-2015-103304.

## Large-scale targeted sequencing comparison highlights extreme genetic heterogeneity in nephronophthisis-related ciliopathies

Markus Schueler<sup>1,\*</sup>, Jan Halbritter<sup>1,2,\*</sup>, Ian G. Phelps<sup>3,\*</sup>, Daniela A. Braun<sup>1</sup>, Edgar A. Otto<sup>4</sup>, Jonathan D. Porath<sup>1</sup>, Heon Yung Gee<sup>1</sup>, Jay Shendure<sup>5</sup>, Brian J. O’Roak<sup>6</sup>, Jennifer A. Lawson<sup>1</sup>, Neveen A. Soliman<sup>7,8</sup>, Marwa M. Nabhan<sup>7,8</sup>, Dan Doherty<sup>3</sup>, and Friedhelm Hildebrandt<sup>1,9</sup>

<sup>1</sup>Division of Nephrology, Department of Medicine, Boston Children’s Hospital, Harvard Medical School, Boston, MA, USA

<sup>2</sup>Division of Nephrology, Department of Internal Medicine, University Clinic Leipzig, Leipzig, Germany

<sup>3</sup>Department of Pediatrics, University of Washington School of Medicine, Seattle, WA, USA

<sup>4</sup>Department of Pediatrics, University of Michigan, Ann Arbor, Michigan, USA

<sup>5</sup>University of Washington, Genome Sciences, Seattle, WA, USA

<sup>6</sup>Oregon Health and Science University, Molecular and Medical Genetics, Portland, OR, USA

<sup>7</sup>Center of Pediatric Nephrology & Transplantation, Cairo University, Cairo, Egypt

<sup>8</sup>Egyptian Group for Orphan Renal Diseases (EGORD), Cairo, Egypt

<sup>9</sup>Howard Hughes Medical Institute, Chevy Chase, MD, USA

Correspondence should be addressed to: Friedhelm Hildebrandt, M.D., Division of Nephrology, Boston Children’s Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115, Phone: +1 617-355-6129, Fax: +1 617-730-0365, friedhelm.hildebrandt@childrens.harvard.edu. Dan Doherty, MD, PhD, Associate Professor of Pediatrics, Divisions of Developmental Medicine and Genetic Medicine, University of Washington, Seattle Children’s Hospital, Seattle, WA 98195, USA, ddoher@uw.edu.

\*These authors contributed equally to this work

Contributing members of the GPN study group are: F Yalcinkaya (Ankara, Turkey); S Bakkaloglu (Ankara, Turkey); F Ozaltin (Ankara, Turkey); E Comak (Antalya, Turkey); F Krull (Aurich, Germany); Schmitz-Hübner (Bad Oeynhausen, Germany); H Rupprecht (Bayreuth, Germany); D Muller (Berlin, Germany); P Dahlem (Coburg, Germany); B Hoppe (Cologne, Germany); M Wolfe (Cologne, Germany); M Weber (Cologne, Germany); U Vester (Essen, Germany); K Bonzel (Essen, Germany); J Nikolay (Furth, Germany); I Hansmann (Halle, Germany); M Wiefel (Hamburg, Germany); U Orth (Hamburg, Germany); H Pfeleiderer (Hamm, Germany); L Pape (Hannover, Germany); Morlot (Hannover, Germany); J Ehrich (Hannover, Germany); B Tonshoff (Heidelberg, Germany); F Schindera (Karlsruhe, Germany); J Hoefele (Martinsried, Germany); M Griebel (Munich, Germany); E Broeking (Münster, Germany); M Konrad (Münster, Germany); M Radke (Potsdam, Germany); M Brandis (Ravensburg, Germany); A Kirchhoff (Wurzburg, Germany); V Feygina (Brooklyn, NY, USA); J Springate (Buffalo, NY, USA); S Ahmadzadeh (Burlington, VT, USA); D Gipson (Chapel Hill, NC, USA); A Becker (Dallas, TX, USA); V Dharmidharka (Gainesville, FL, USA); P Mark (Grand Rapids, MI, USA); P Srivaths (Houston, TX, USA); A Wilson (Indianapolis, IN, USA); E Kamil (Los Angeles, CA, USA); S Why (Milwaukee, WI, USA); C Pan (Milwaukee, WI, USA); C Kashtan (Minneapolis, MN, USA); C D’Alessandri (New Haven, CT, USA); H Trachtman (New York city, NY, USA); B Kaplan (Philadelphia, PA, USA); M Joseph (Phoenix, AZ, USA); R Weiss (Valhalla, NY, USA); S Thomas (Ann Arbor, MI, USA); L Newberry (Aurora, CO, USA); M Koyun (Cairo, Egypt); H Fathy (Alexandria, Egypt); A Rybi – Szuminska (Bialystok, Poland); M Szczepanska (Zabrze, Poland); Z Dolezel (Brno, Czech Republic); M Malina (Prague, Czech Republic); T Seeman (Prague, Czech Republic); T Honzik (Prague, Czech Republic); P Ferreira (Calgary, Canada); M Ferguson (Halifax, Canada); E Harvey (Toronto, Canada); K Chong (Toronto, Canada); R Sandford (Cambridge, UK); D Josifova (London, UK); D Bockenbauer (London, UK); J Sayer (Newcastle upon Tyne, UK); C Johnson (Yorkshire, UK); P Senguttuvan (Chennai, India); I Pela (Firenze, Italy); N Knops (Leuven, Belgium); T Levart (Ljubljana, Slovenia); T Neuhaus (Luzern, Switzerland); C Ayuso (Madrid, Spain); A Kindi (Muscat, Sultanate of Oman); N Knoers (Nijmegen, The Netherlands); C Antignac (Paris, France); W Radauer (Salzburg, Austria); C Genzani (Sao Paulo, Brazil); U Berg (Stockholm, Sweden); C Klingenberg (Tromsø, Norway); C Jones (Victoria, Australia); R Savarirayan (Victoria, Australia); J Kausman (Victoria, Australia).

## Abstract

The term nephronophthisis-related ciliopathies (NPHP-RC) describes a group of rare autosomal-recessive cystic kidney diseases, characterized by broad genetic and clinical heterogeneity. NPHP-RC is frequently associated with extrarenal manifestations and accounts for the majority of genetically caused chronic kidney disease (CKD) during childhood and adolescence. Generation of a molecular diagnosis has been impaired by this broad genetic heterogeneity. However, recently developed high-throughput exon sequencing techniques represent powerful and efficient tools to screen large cohorts for dozens of causative genes. Therefore, we performed massively multiplexed targeted sequencing using the modified molecular inversion probe (MIPs) strategy in an international cohort of 384 patients diagnosed with NPHP-RC. As a result, we established the molecular diagnoses in 81/384 unrelated individuals (21.1%). We detected 127 likely disease-causing mutations in 18 of 34 evaluated NPHP-RC genes, 22 of which were novel. We further compared a subgroup of current findings to the results of a previous study in which we used an array-based microfluidic PCR technology in the same cohort. While 78 likely disease-causing mutations were previously detected by the array-based microfluidic PCR, the MIPs approach identified 94 likely pathogenic mutations. Compared to the previous approach, MIPs re-detected 66 out of 78 variants and 28 previously unidentified variants, for a total of 94 variants. In summary, we demonstrate that the modified MIPs technology is a useful approach to screen large cohorts for a multitude of established NPHP genes in order to identify the underlying molecular cause. Combined application of two independent library preparation and sequencing techniques, however, may still be indicated for Mendelian diseases with extensive genetic heterogeneity in order to further increase diagnostic sensitivity.

## Keywords

NPHP; ciliopathies; MIP; cystic kidney disease

## INTRODUCTION

The term nephronophthisis-related ciliopathies (NPHP-RC) summarizes a group of rare autosomal-recessive cystic kidney diseases including nephronophthisis (NPHP), Senior-Løken syndrome (SLS), Joubert syndrome (JBTS), and Meckel-Gruber-syndrome (MKS).<sup>1, 2</sup> NPHP-RC are genetically heterogeneous disorders, characterized by mutations in genes which encode proteins that localize to primary cilia, basal bodies, or centrosomes. The disruption of the ciliary function results in a broad phenotypic spectrum which is collectively termed “ciliopathies”.<sup>3</sup> NPHP-RC account for the majority of genetic end-stage renal disease (ESRD) during the first three decades of life. The most prominent renal features are increased echogenicity and corticomedullary cysts on ultrasound. Renal histology reveals tubular atrophy, basement membrane disintegration, interstitial fibrosis, and cyst formation. About 15% of the affected individuals with NPHP-RC show extrarenal organ involvement, most notably, progressive retinal dystrophy defined as SLS. Cerebellar malformations such as mid-hindbrain malformation and cerebellar vermis hypoplasia/aplasia are commonly observed in individuals with JBTS and are characterized by the “molar tooth sign” (MTS) in brain imaging studies.<sup>4</sup> These malformations are associated with numerous

neurological features including developmental delay, intellectual disability, muscle hypotonia, ataxia, oculomotor apraxia, nystagmus, and abnormal respiratory control.<sup>4</sup> MKS represents the most severe manifestation of the NPHP-RC spectrum as a perinatally lethal ciliopathy.<sup>5</sup> To date, more than 90 genes have been implicated in the pathogenesis of ciliopathies, accounting for approximately 50% of affected individuals.<sup>6</sup> JBTS and MKS are caused by mutations in a subset of these genes including: *MKS1* (MIM# 609883), *B9D1* (MIM# 614144), *B9D2* (MIM# 611951), *AH11* (MIM# 608894), *INPP5E* (MIM# 613037), *ARL13B* (MIM# 608922), *TMEM216* (MIM# 613277), *CC2D2A* (MIM# 612013), *KIF7* (MIM# 611254), *TCTN1* (MIM# 609863), *TCTN2* (MIM# 613846), *TCTN3* (MIM#614815), *CEP41* (MIM# 610523), *IFT172* (MIM#615630), *CSPP1* (MIM#615636), *RPGRIPL* (MIM#611560), *OFD1* (MIM#300804), *TMEM138* (MIM#614465), *TMEM231* (MIM#614970), *PDE6D* (MIM#615665), and *TMEM237* (MIM# 614423).<sup>7–26</sup> Due to the large number of established NPHP-RC genes, molecular genetic diagnostics remains a major challenge and requires further technical advances.

Recently, high-throughput and next-generation sequencing (NGS) platforms have become widely available and allow massively parallel low-cost sequencing.<sup>27, 28</sup> Previous studies from our group have established a microfluidic multiplex PCR technology with consecutive NGS. Screening for multiple established *NPHP* genes yielded the molecular cause in about 13% of cases in large cohorts of patients with NPHP-RC.<sup>29, 30</sup> The recently developed Modified Molecular Inversion probe strategy (MIPs) represents another technique that allows rapid and cost-effective sequencing of candidate genes in large cohorts.<sup>31</sup> In the present study, we applied the MIPs technology to 384 samples from affected unrelated individuals with NPHP-RC. We captured and amplified 760 coding exons of 34 established NPHP-RC genes, including *NPHP1-NPHP12*, *PKHD1*, the gene causing autosomal-recessive cystic kidney disease (ARPKD), and 21 established JBTS/MKS genes. We established the molecular diagnoses in 81/384 unrelated individuals (21.1%). Altogether, we detected 127 mutations. In addition, we compared the sensitivity, specificity, and costs of the two above mentioned targeted high-throughput technologies in our cohort.

## MATERIALS AND METHODS

### Human subjects

We obtained blood samples, pedigrees, and clinical information after receiving informed consent ([www.renalgenes.org](http://www.renalgenes.org)). Approval for experiments on humans was obtained from the University of Michigan and the Boston Children's Hospital Institutional Review Boards. A total cohort of 384 affected individuals diagnosed with NPHP-RC based on published clinical criteria were sequenced using the MIPs technology.<sup>32</sup> This cohort contains 84 unrelated individuals from families with multiple affected individuals, and 300 single affected individuals. Consanguinity was known to be present in 67 (17.3%) families. Prior to study enrollment, individuals with homozygous deletions of *NPHP1* were excluded by using a multiplex PCR-based deletion analysis as previously described.<sup>33</sup>

### Primer design and evaluation for the MIP system

MIPs were designed to target 760 coding exons of 34 established NPHP-RC genes (Supplementary Table 1) using hg19 human genome reference and the dbSNP132 database to identify polymorphisms that might interfere with hybridization. Each MIP targeted a specific 112 bp genomic region using flanking extension (16 to 20 bp) and ligation (20 to 24 bp) arms. Following synthesis (IDT, Coralville, IA), MIPs were standardized to 100uM, and equimolar amounts of each MIP were pooled.<sup>31</sup>

### Target DNA enrichment by the MIP System

Target-specific MIPs hybridized to 100ng of genomic DNA input followed by polymerase-mediated gap filling and ligation. The captured DNA was PCR amplified using universal primers containing Illumina flowcell adaptor sequences and a unique 8-base barcode. (Illumina, San Diego, CA). Next, four sets of 96 barcoded sample libraries were pooled, and each was purified using 1.8X AMPure XP beads (Beckman Coulter, Brea, CA). The appropriate product size was confirmed by agarose gel electrophoresis.<sup>31</sup>

### MIP target rebalancing and next-generation sequencing

A pilot MIP capture was performed on a subset of samples and sequenced using the Illumina MiSeq. MIP targets were rebalanced by spiking the poorest performing (bottom 15% by read depth) MIPs at 10x. The rebalanced MIP pool was used to create the 384 individual sample libraries as described above. Pooled libraries were sequenced on three lanes of an Illumina HiSeq 2000 instrument using 101 bp paired-end reads.<sup>31</sup>

### Bioinformatics pipeline

Raw sequence reads were processed and aligned to the hg19 human reference sequence (UCSC Genome Browser) using the Burrows-Wheeler Aligner (BWA, version 0.5.9). BAM files were generated with Picard tools (see Web Resources), followed by single-nucleotide variant (SNV) and indel calling using the Genome Analysis Toolkit's Unified-Genotyper (GATK version 2.2) with the depth of coverage parameter set to 5,000. Called variants failing to meet the following quality metrics were removed from further analysis: quality > 30, read depth > 8, quality by depth > 5. A window of 10 was set for clustered variants. Variants were annotated with SeattleSeq (see Web Resources) and further filtered against controls from the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project Exome Variant Server (EVS) ESP6500 dataset with the use of an observed minor allele frequency threshold < 1%.<sup>31</sup> Variants were ranked by the criteria of whether mutations were likely to truncate the conceptual reading frame (nonsense, frameshift, and obligatory splice mutations), whereas missense variants were ranked by the evolutionary conservation and by predicting the impact on the encoded protein using web-based programs (PolyPhen2, Phylo-P score, CADD score, Mutation Taster, SIFT).

### *NPHP1* copy number variation detection by the MIP system

The read depth of each MIP was normalized to the overall read depth for each sample, and then a standardized Z-score was calculated. Underperforming MIPs with a median coverage <20X were removed. Singular value decomposition (SVD) was used to reduce systematic

noise. Using the SVD-adjusted Z-scores, we identified samples where the majority of *NPHP1*-targeted MIPs had Z-scores  $\geq 2.0$  or  $\leq -2.0$ .<sup>31</sup> Individual A2527, who was examined in an independent array previously, served as a positive control for successful CNV calling.

### Sanger sequencing confirmation and segregation analysis

Variants/mutations detected by NGS and predicted to be detrimental were subsequently confirmed by Sanger sequencing using original DNA samples from the respective affected individuals as PCR template. Whenever DNA was available from the index person, the affected siblings, and parents, we performed segregation analysis. Polymerase chain reaction (PCR) was performed using a touchdown protocol described previously.<sup>34</sup> Sequencing was performed using Mastercycler pro (Eppendorf, Hamburg, Germany). Sequence traces were analyzed using CLC Genomics Workbench™ software (CLC-bio, Aarhus, Denmark).

## RESULTS

### NGS and mapping statistics

Sequencing was performed on an Illumina HiSeq 2000 platform after capture and amplification of 34 NPHP-RC genes in 384 affected individuals (384 families) with NPHP-RC using the Modified Molecular Inversion Probe technique (Figure 1, Supplementary Table 2). The total output was 361 million reads of 101 bases yielding an average of 940 thousand reads per DNA sample. Alignment using the Burrows-Wheeler Aligner (BWA, version 0.5.9) software to the hg19 human reference sequence (UCSC Genome Browser) resulted in mean exon coverage of 505, with greater than 97% of targeted exons having 25-fold average coverage (Supplementary Figure 1). Due to low DNA quality, sequencing failed for 2 samples, while 7 samples have 10 and 1 sample 13% coverage at 25x.

### Variant filtering, validation and parameter setting

Altogether, we obtained 27,133 single nucleotide variants and 646 insertion/deletion (indel) variants derived from the 384 individuals. In total, 20,371 (exonic and intronic) variants showed a minor allele frequency (MAF) of  $>1\%$  across the dataset, which were considered to be frequent SNPs according to the EVS database and, therefore excluded.

As previously reported,<sup>31</sup> we removed variants from further analysis whenever the following parameters were present: coverage  $< 8x$ , quality score  $< 30$ , and a quality by depth  $< 5$ . Variants below this threshold were considered as likely “false positives”. Furthermore, we included only missense variants that showed an evolutionary conservation down to *Ciona intestinalis*, a PolyPhen2-score  $>0.9$  and a MAF  $<1\%$  across the dataset. Alternatively, variants were included when they had previously been reported as disease-causing according to the HGMD®-Professional mutation database “Biobase” and were conserved down to *Danio rerio*.

After filtering and ranking, 189 variants, including 110 potentially protein-truncating mutations (such as nonsense, frameshift, and obligatory splice site mutations), and 79 missense variants/mutations, were selected for standard Sanger sequencing. We were able to

confirm 93 of the 110 potential truncating mutations (85%) and 70 of the 79 selected missense variants (89%). A total of 163 out of 189 variants (specificity: 86%) were confirmed by Sanger sequencing. However, 35 of the 163 confirmed potential mutations were excluded as common SNPs according to the 1000 genomes, dbSNP138, EVS and ExAC Browser database or have been reported as homozygous SNPs. The remaining mutations were compared to the HGMD<sup>®</sup>-Professional mutation database “Biobase” (HGMD 2015.2 (copyright Cardiff University 2015)) and defined as novel if not present.

### Sensitivity for detecting known variants and comparison of two approaches

In order to calculate the sensitivity of the MIPs technology in this study, we compared the detected disease-causing mutations within the coding regions of 12 established NPHP genes (*NPHP1-NPHP12*) to our previous mutation analysis.<sup>29, 30</sup> By using an array-based microfluidic PCR technology with NGS, we previously identified and Sanger confirmed 78 variants, which explained the underlying molecular cause in 49/384 families of the same cohort.<sup>29, 30</sup> Overall, 66 out of these 78 mutations were re-detected in the present study (“mutation” sensitivity: 85%). In particular, eight mutations were not re-detected due to low coverage, giving a sensitivity of 94% (66/70) at covered basepairs. A four basepair duplication in *IQCB1* was missed in three individuals because one of the flanker arms of this MIP overlapped with the duplication. For three individuals the reason why the MIPs technology missed the previous detected mutation remains unclear (Figure 2, Supplementary Table 3).

The estimated costs for library preparation and sequencing of 384 individuals using the modified MIPs technology are US\$21,000<sup>31</sup> whereas the calculated costs for sequencing the same cohort using an array-based microfluidic PCR technology amount is US\$18,500<sup>29</sup> in total. These costs comprise reagents and primers but do not include costs for acquisition of additional machinery, which is the case in the microfluidic PCR approach (Supplementary Table 4, Supplementary Figure 2). Concerning the hands-on time for library preparation, sequencing, and bioinformatics pipeline analysis, there are no major differences between both methods: approx. 142 hours for the MIPs technology and 148 hours if using an array-based microfluidic PCR technology.

### Identification of genetic causes in a cohort of 384 individuals

The combination of capturing all coding exons of 34 established NPHP-RC genes by using the MIPs technology in an international cohort of 384 unrelated affected individuals (384 families) diagnosed with NPHP-RC revealed the molecular diagnosis in 81 individuals (81 families). Furthermore, 26 single heterozygous mutations were detected in 26 individuals, 15 of which were missense variants, and 11 of which were single heterozygous nonsense variants. Additional Sanger sequencing of all remaining coding exons of the genes, wherein a single heterozygous variant was detected, revealed two additional heterozygous missense mutations, which segregated within the affected status. Copy number variation (CNV) analysis for *NPHP1* revealed nine heterozygous deletions and six heterozygous duplications including the positive control (individual A2527); none was previously detected by using a multiplex PCR-based deletion analysis.

In summary, by massively parallel targeted sequencing using MIPs strategy we identified the molecular diagnosis in 81 of 384 (21.1%) unrelated patients/families. Recessive mutations were identified in the following genes (number of unrelated patients/families in parentheses): *CEP290* (15), *PKHD1* (10), *IQCB1* (10), *NPHP3* (9), *NPHP1* (8), *TMEM67* (7), *NPHP4* (6), *INVS* (3), *CC2D2A* (2), *CEP164* (2), *TTC21B* (2), *SDCCAG8* (1), *IFT172* (1), *OFD1* (1), *TMEM216* (1), *TMEM237* (1), *TCTN1* (1), and *AHI* (1) (Supplementary Table 2). No causative mutations were identified in the genes *GLIS2* (*NPHP7*), *RPGRIPL* (*NPHP8*), *NEK8* (*NPHP9*), *ARL13B*, *B9D1*, *B9D2*, *C5ORF42*, *CEP41*, *INPP5E*, *KIF7*, *MKS1*, *TCTN2*, *TCTN3*, *TMEM231*, *TMEM138*, and *CSPP1*. Overall, we identified 30 independent families with homozygous mutations, five families with hemizygous mutations, and 46 families with compound heterozygous mutations. In 51 individuals truncating mutations (nonsense, frameshift or obligatory splice-site mutations) were found on both alleles, whereas 14 affected individuals showed one truncating mutation in combination with a non-synonymous missense mutation. The remaining 16 individuals exhibited missense mutations on both alleles. After additional evaluation and Sanger sequencing of all coding regions and intron/exon boundaries in the respective genes, 11 families remained with only one heterozygous truncating mutation. We discovered a total of 22 novel pathogenic mutations in the following genes (number of families in parentheses): *PKHD1* (6), *NPHP1* (4), *CC2D2A* (2), *CEP290* (2), *NPHP4* (2), *NPHP3* (2), *INVS* (1), *TTC21B* (1), *OFD1* (1), and *TMEM237* (1) (Table 1).

## DISCUSSION

Massively multiplex targeted sequencing of 34 established NPHP-RC genes using the MIPs strategy in an international cohort of 384 affected individuals with NPHP-RC identified the molecular cause in 81 families. Altogether our study identified 127 pathogenic mutations. *CEP290* was found most commonly mutated (15 families). Interestingly, mutations in *PKHD1* and *IQCB1* were the second most frequent molecular cause in our cohort (10 affected families each). This shows that the clinical presentation of ARPKD may mimic the NPHP-RC phenotype and is not necessarily distinguishable without the use of molecular genetics in clinical practice.

Combined with our previous mutation analysis using an array-based microfluidic multiplex PCR technology, we obtain a representative distribution of genes in a cohort of individuals with NPHP-RC pre-screened for the most common mutation, the *NPHP1* deletion.<sup>30</sup> After exclusion of the presence of a homozygous *NPHP1* deletion, a disease-causing mutation was identified in 21% of the cases, leaving almost 80% of cases genetically unsolved when analyzing 34 established NPHP, ARPKD, and JBST genes (Supplementary Table 2) (Figure 1).

Twenty-six individuals carry single heterozygous variants, which do not explain the phenotype based on an autosomal recessive hypothesis; however, it is possible that second variants were not detected in the few areas of low coverage or in non-coding regions. These unsolved cases may be explained by missed mutations in the targeted genes, mutations in other established NPHP-RC genes that were not part of this study, or by mutations in genes which phenocopy NPHP-RC, such as e.g. *CLCNKB*, *COL4A5*, and *AGXT*.<sup>35</sup> Secondly,

these cases may be explained due to mutations in genes, which have not yet been implicated as a molecular cause of NPHP-RC. This fact further implies the presence of genetic heterogeneity in individuals with NPHP-RC. Lastly, a significant fraction may be explained by unidentified mutations (false negatives), as indicated by a mutation sensitivity of 85% when compared to our previous analysis of the same cohort (Figure 2, Supplementary Table 3). As a conclusion, combination of two library preparation methods may be indicated in order to warrant sufficient sensitivity, at least in a diagnostic setting.

In summary, we demonstrate that the modified molecular inversion probe (MIPs) strategy is a useful library preparation method for targeted sequencing approaches and mutation analysis in large cohorts of individuals with NPHP-RC. Whereas Sanger sequencing has been the gold standard over the past 30 years, it has been proven to be tedious and expensive, especially for large cohorts and genetically heterogeneous conditions like NPHP-RC. Recent developments in high-throughput library preparation and massively parallel sequencing offer various novel and efficient methods that decrease sequencing time and costs. Due to the increasing number of established genes, the relevance and utility of high-throughput sequencing techniques is indispensable for molecular genetic diagnostics.

As about 50% of affected individuals with NPHP-RC remain genetically unsolved, non-targeted sequencing approaches like whole exome/genome sequencing (WES/WGS) are needed to identify new candidate disease genes. For mutation analysis of panels of known disease genes in large cohorts, however, WES is still not cost-efficient. In our case, the price tag for sequencing a cohort of 384 affected individuals using WES would amount to about US\$268,800 assuming US\$700 per individual, whereas the costs for MIPs or array-based microfluidic PCR technology added up to about US\$20,000 (approx. US\$50–100 per individual for > 200 samples) (Supplementary Table 4, Supplementary Figure 2). For the proposed targeted sequencing approaches, though, a minimal cohort size of at least 100 probands is recommended in order to run them cost-efficiently (Supplementary Figure 2). While WES/WGS are essential tools to identify novel candidate disease genes in single cases, targeted high-throughput screening technologies, such as MIPs, can also be helpful to analyze large cohorts with the same phenotype to corroborate these candidates through detection of additional unrelated individuals with mutations in the same gene.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors sincerely thank the affected individuals and their families for participation in this study and the study coordinators Leslie Spaneas and Brittany Fisher. This work was supported by a grant from the National Institutes of Health to F.H. (RC4-DK076683) and the NephCure Foundation. F.H. is an investigator of the Howard Hughes Medical Institute (HHMI) and the Warren E. Grupe Professor.

This work was also supported by the National Institutes of Health through the University of Washington Intellectual and Developmental Disabilities Research Center Genetics Core (P30-HD002274) and R01-NS064077 to D.D. We also acknowledge private donations from the families of children with Joubert syndrome.

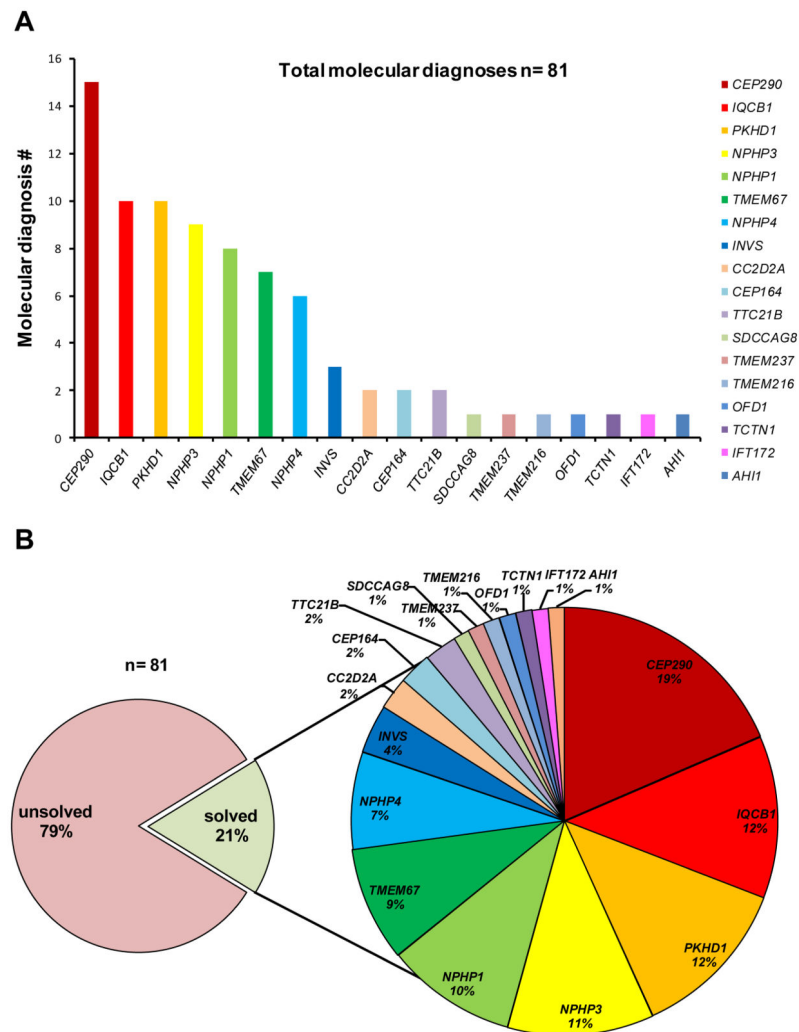


## References

1. Hildebrandt F, Attanasio M, Otto E. Nephronophthisis: disease mechanisms of a ciliopathy. *J Am Soc Nephrol.* 2009; 20:23–35. [PubMed: 19118152]
2. Wolf MT, Hildebrandt F. Nephronophthisis. *Pediatr Nephrol.* 2011; 26:181–94. [PubMed: 20652329]
3. Hildebrandt F, Benzing T, Katsanis N. Ciliopathies. *N Engl J Med.* 2011; 364:1533–43. [PubMed: 21506742]
4. Parisi MA. Clinical and molecular features of Joubert syndrome and related disorders. *Am J Med Genet Part C: Sem Med Genet.* 2009; 151C:326–340.
5. Johnson CA, Gissen P, Sergi C. Molecular pathology and genetics of congenital hepatorenal fibrocystic syndromes. *J Med Genet.* 2003; 40:311–9. [PubMed: 12746391]
6. Braun DA, Schueler M, Halbritter J, Gee HY, Porath JD, Lawson JA, Airik R, Shril S, Allen SJ, Stein D, Al Kindy A, Beck BB, Cengiz N, Moorani KN, Ozaltin F, Hashmi S, Sayer JA, Bockenbauer D, Soliman NA, Otto EA, Lifton RP, Hildebrandt F. Whole exome sequencing identifies causative mutations in the majority of consanguineous or familial cases with childhood-onset increased renal echogenicity. *Kidney Int.* 2015; doi: 10.1038/ki.2015.317
7. Kyttila M, Tallila J, Salonen R, Kopra O, Kohlschmidt N, Paavola-Sakki P, Peltonen L, Kestila M. MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome. *Nat Genet.* 2006; 38:155–157. [PubMed: 16415886]
8. Hopp K, Heyer CM, Hommerding CJ, Henke SA, Sundsbak JL, Patel S, Patel P, Consugar MB, Czarnecki PG, Gliem TJ, Torres VE, Rossetti S, Harris PC. B9D1 is revealed as a novel Meckel syndrome (MKS) gene by targeted exon-enriched next-generation sequencing and deletion analysis. *Hum Mol Genet.* 2011; 20:2524–2534. [PubMed: 21493627]
9. Dowdle William E, Robinson Jon F, Kneist A, Sirerol-Piquer MS, Frints Suzanna GM, Corbit Kevin C, Zaghoul Norran A, van Lijnschoten G, Mulders L, Verver Dideke E, Zerres K, Reed Randall R, Attié-Bitach T, Johnson Colin A, García-Verdugo José M, Katsanis N, Bergmann C, Reiter Jeremy F. Disruption of a Ciliary B9 Protein Complex Causes Meckel Syndrome. *Am J Hum Genet.* 2011; 89:94–110. [PubMed: 21763481]
10. Ferland RJ, Eyaid W, Collura RV, Tully LD, Hill RS, Al-Nouri D, Al-Rumayyan A, Topcu M, Gascon G, Bodell A, Shugart YY, Ruvolo M, Walsh CA. Abnormal cerebellar development and axonal decussation due to mutations in AHI1 in Joubert syndrome. *Nat Genet.* 2004; 36:1008–1013. [PubMed: 15322546]
11. Bielas SL, Silhavy JL, Brancati F, Kisseleva MV, Al-Gazali L, Sztriha L, Bayoumi RA, Zaki MS, Abdel-Aleem A, Rosti RO, Kayserili H, Swistun D, Scott LC, Bertini E, Boltshauser E, Fazzi E, Travaglini L, Field SJ, Gayral S, Jacoby M, Schurmans S, Dallapiccola B, Majerus PW, Valente EM, Gleeson JG. Mutations in INPP5E, encoding inositol polyphosphate-5-phosphatase E, link phosphatidyl inositol signaling to the ciliopathies. *Nat Genet.* 2009; 41:1032–6. [PubMed: 19668216]
12. Cantagrel V, Silhavy JL, Bielas SL, Swistun D, Marsh SE, Bertrand JY, Audollent S, Attié-Bitach T, Holden KR, Dobyns WB, Traver D, Al-Gazali L, Ali BR, Lindner TH, Caspary T, Otto EA, Hildebrandt F, Glass IA, Logan CV, Johnson CA, Bennett C, Brancati F, Valente EM, Woods CG, Gleeson JG. Mutations in the Cilia Gene ARL13B Lead to the Classical Form of Joubert Syndrome. *Am J Hum Genet.* 2008; 83:170–179. [PubMed: 18674751]
13. Valente EM, Logan CV, Mougou-Zerelli S, Lee JH, Silhavy JL, Brancati F, Iannicelli M, Travaglini L, Romani S, Illi B, Adams M, Szymanska K, Mazzotta A, Lee JE, Tolentino JC, Swistun D, Salpietro CD, Fede C, Gabriel S, Russ C, Cibulskis K, Sougnez C, Hildebrandt F, Otto EA, Held S, Diplas BH, Davis EE, Mikula M, Strom CM, Ben-Zeev B, Lev D, Sagie TL, Michelson M, Yaron Y, Krause A, Boltshauser E, Elkhartoufi N, Roume J, Shalev S, Munnich A, Saunier S, Inglehearn C, Saad A, Alkindy A, Thomas S, Vekemans M, Dallapiccola B, Katsanis N, Johnson CA, Attie-Bitach T, Gleeson JG. Mutations in TMEM216 perturb ciliogenesis and cause Joubert, Meckel and related syndromes. *Nat Genet.* 2010; 42:619–25. [PubMed: 20512146]
14. Gorden NT, Arts HH, Parisi MA, Coene KLM, Letteboer SJF, van Beersum SEC, Mans DA, Hikida A, Eckert M, Knutzen D, Alswaid AF, Özyurek H, Dibooglu S, Otto EA, Liu Y, Davis EE, Hutter CM, Bammler TK, Farin FM, Dorschner M, Topçu M, Zackai EH, Rosenthal P, Owens KN,

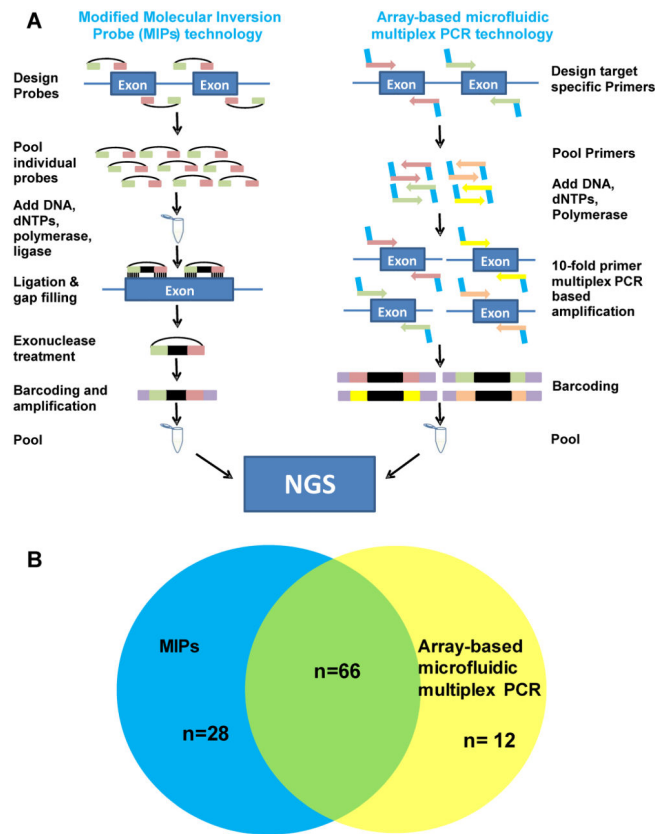
- Katsanis N, Vincent JB, Hildebrandt F, Rubel EW, Raible DW, Knoers NVAM, Chance PF, Roepman R, Moens CB, Glass IA, Doherty D. CC2D2A Is Mutated in Joubert Syndrome and Interacts with the Ciliopathy-Associated Basal Body Protein CEP290. *Am J Hum Genet.* 2008; 83:559–571. [PubMed: 18950740]
15. Dafinger C, Liebau MC, Elsayed SM, Hellenbroich Y, Boltshauser E, Korenke GC, Fabretti F, Janecke AR, Ebermann I, Nurnberg G, Nurnberg P, Zentgraf H, Koerber F, Addicks K, Elsobky E, Benzing T, Schermer B, Bolz HJ. Mutations in KIF7 link Joubert syndrome with Sonic Hedgehog signaling and microtubule dynamics. *J Clin Invest.* 2011; 121:2662–7. [PubMed: 21633164]
  16. Garcia-Gonzalo FR, Corbit KC, Sirerol-Piquer MS, Ramaswami G, Otto EA, Noriega TR, Seol AD, Robinson JF, Bennett CL, Josifova DJ, Garcia-Verdugo JM, Katsanis N, Hildebrandt F, Reiter JF. A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition. *Nat Genet.* 2011; 43:776–84. [PubMed: 21725307]
  17. Sang L, Miller JJ, Corbit KC, Giles RH, Brauer MJ, Otto EA, Baye LM, Wen X, Scales SJ, Kwong M, Huntzicker EG, Sfakianos MK, Sandoval W, Bazan JF, Kulkarni P, Garcia-Gonzalo FR, Seol AD, O'Toole JF, Held S, Reutter HM, Lane WS, Rafiq MA, Noor A, Ansar M, Devi AR, Sheffield VC, Slusarski DC, Vincent JB, Doherty DA, Hildebrandt F, Reiter JF, Jackson PK. Mapping the NPHP-JBTS-MKS protein network reveals ciliopathy disease genes and pathways. *Cell.* 2011; 145:513–28. [PubMed: 21565611]
  18. Lee JE, Silhavy JL, Zaki MS, Schroth J, Bielas SL, Marsh SE, Olvera J, Brancati F, Iannicelli M, Ikegami K, Schlossman AM, Merriman B, Attie-Bitach T, Logan CV, Glass IA, Cluckey A, Louie CM, Lee JH, Raynes HR, Rapin I, Castroviejo IP, Setou M, Barbot C, Boltshauser E, Nelson SF, Hildebrandt F, Johnson CA, Doherty DA, Valente EM, Gleeson JG. CEP41 is mutated in Joubert syndrome and is required for tubulin glutamylation at the cilium. *Nat Genet.* 2012; 44:193–199. [PubMed: 22246503]
  19. Halbritter J, Bizet AA, Schmidts M, Porath JD, Braun DA, Gee HY, McInerney-Leo AM, Krug P, Filhol E, Davis EE, Airik R, Czarnecki PG, Lehman AM, Trnka P, Nitschke P, Bole-Feysot C, Schueler M, Knebelmann B, Burtey S, Szabo AJ, Tory K, Leo PJ, Gardiner B, McKenzie FA, Zankl A, Brown MA, Hartley JL, Maher ER, Li C, Leroux MR, Scambler PJ, Zhan SH, Jones SJ, Kayserili H, Tuysuz B, Moorani KN, Constantinescu A, Krantz ID, Kaplan BS, Shah JV, Hurd TW, Doherty D, Katsanis N, Duncan EL, Otto EA, Beales PL, Mitchison HM, Saunier S, Hildebrandt F. Defects in the IFT-B component IFT172 cause Jeune and Mainzer-Saldino syndromes in humans. *Am J Hum Genet.* 2013; 93:915–25. [PubMed: 24140113]
  20. Shaheen R, Shamseldin HE, Loucks CM, Seidahmed MZ, Ansari S, Ibrahim Khalil M, Al-Yacoub N, Davis EE, Mola NA, Szymanska K, Herridge W, Chudley AE, Chodirker BN, Schwartzentruber J, Majewski J, Katsanis N, Poizat C, Johnson CA, Parboosingh J, Boycott KM, Innes AM, Alkuraya FS. Mutations in CSPP1, encoding a core centrosomal protein, cause a range of ciliopathy phenotypes in humans. *Am J Hum Genet.* 2014; 94:73–9. [PubMed: 24360803]
  21. Arts HH, Doherty D, van Beersum SE, Parisi MA, Letteboer SJ, Gorden NT, Peters TA, Marker T, Voesenek K, Kartono A, Ozyurek H, Farin FM, Kroes HY, Wolfrum U, Brunner HG, Cremers FP, Glass IA, Knoers NV, Roepman R. Mutations in the gene encoding the basal body protein RPGRIP1L, a nephrocystin-4 interactor, cause Joubert syndrome. *Nat Genet.* 2007; 39:882–8. [PubMed: 17558407]
  22. Coene KL, Roepman R, Doherty D, Afroze B, Kroes HY, Letteboer SJ, Ngu LH, Budny B, van Wijk E, Gorden NT, Azhimi M, Thauvin-Robinet C, Veltman JA, Boink M, Kleefstra T, Cremers FP, van Bokhoven H, de Brouwer AP. OFD1 is mutated in X-linked Joubert syndrome and interacts with LCA5-encoded lebercilin. *Am J Hum Genet.* 2009; 85:465–81. [PubMed: 19800048]
  23. Lee JH, Silhavy JL, Lee JE, Al-Gazali L, Thomas S, Davis EE, Bielas SL, Hill KJ, Iannicelli M, Brancati F, Gabriel SB, Russ C, Logan CV, Sharif SM, Bennett CP, Abe M, Hildebrandt F, Diplas BH, Attie-Bitach T, Katsanis N, Rajab A, Koul R, Sztirha L, Waters ER, Ferro-Novick S, Woods CG, Johnson CA, Valente EM, Zaki MS, Gleeson JG. Evolutionarily assembled cis-regulatory module at a human ciliopathy locus. *Science.* 2012; 335:966–9. [PubMed: 22282472]
  24. Srour M, Hamdan FF, Schwartzentruber JA, Patry L, Ospina LH, Shevell MI, Desilets V, Dobrzyniecka S, Mathonnet G, Lemyre E, Massicotte C, Labuda D, Amrom D, Andermann E, Sebire G, Maranda B, Rouleau GA, Majewski J, Michaud JL. Mutations in TMEM231 cause Joubert syndrome in French Canadians. *J Med Genet.* 2012; 49:636–41. [PubMed: 23012439]

25. Thomas S, Wright KJ, Le Corre S, Micalizzi A, Romani M, Abhyankar A, Saada J, Perrault I, Amiel J, Litzler J, Filhol E, Elkhartoufi N, Kwong M, Casanova JL, Boddaert N, Baehr W, Lyonnet S, Munnich A, Burglen L, Chassaing N, Encha-Ravazi F, Vekemans M, Gleeson JG, Valente EM, Jackson PK, Drummond IA, Saunier S, Attie-Bitach T. A homozygous PDE6D mutation in Joubert syndrome impairs targeting of farnesylated INPP5E protein to the primary cilium. *Hum Mutat.* 2014; 35:137–46. [PubMed: 24166846]
26. Huang L, Szymanska K, Jensen Victor L, Janecke Andreas R, Innes AM, Davis Erica E, Frosk P, Li C, Willer Jason R, Chodirker Bernard N, Greenberg Cheryl R, McLeod DR, Bernier Francois P, Chudley Albert E, Müller T, Shboul M, Logan Clare V, Loucks Catrina M, Beaulieu Chandree L, Bowie Rachel V, Bell Sandra M, Adkins J, Zuniga Freddi I, Ross Kevin D, Wang J, Ban Matthew R, Becker C, Nürnberg P, Douglas S, Craft Cheryl M, Akimenko M-A, Hegele Robert A, Ober C, Utermann G, Bolz Hanno J, Bulman Dennis E, Katsanis N, Blacque Oliver E, Doherty D, Parboosingh Jillian S, Leroux Michel R, Johnson Colin A, Boycott Kym M. TMEM237 Is Mutated in Individuals with a Joubert Syndrome Related Disorder and Expands the Role of the TMEM Family at the Ciliary Transition Zone. *Am J Hum Genet.* 2011; 89:713–730. [PubMed: 22152675]
27. Shendure J, Mitra RD, Varma C, Church GM. Advanced sequencing technologies: methods and goals. *Nat Rev Genet.* 2004; 5:335–344. [PubMed: 15143316]
28. Voelkerding KV, Dames S, Durtschi JD. Next generation sequencing for clinical diagnostics—principles and application to targeted resequencing for hypertrophic cardiomyopathy: a paper from the 2009 William Beaumont Hospital Symposium on Molecular Pathology. *J Mol Diagn.* 2010; 12:539–51. [PubMed: 20805560]
29. Halbritter J, Diaz K, Chaki M, Porath JD, Tarrier B, Fu C, Innis JL, Allen SJ, Lyons RH, Stefanidis CJ, Omran H, Soliman NA, Otto EA. High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing. *J Med Genet.* 2012; 49:756–767. [PubMed: 23188109]
30. Halbritter J, Porath J, Diaz K, Braun D, Kohl S, Chaki M, Allen S, Soliman N, Hildebrandt F, Otto E. Identification of 99 novel mutations in a worldwide cohort of 1,056 patients with a nephronophthisis-related ciliopathy. *Hum Genet.* 2013; 132:865–884. [PubMed: 23559409]
31. O’Roak BJ, Vives L, Fu W, Egerton JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K, Munson J, Hiatt JB, Turner EH, Levy R, O’Day DR, Krumm N, Coe BP, Martin BK, Borenstein E, Nickerson DA, Mefford HC, Doherty D, Akey JM, Bernier R, Eichler EE, Shendure J. Multiplex Targeted Sequencing Identifies Recurrently Mutated Genes in Autism Spectrum Disorders. *Science.* 2012; 338:1619–1622. [PubMed: 23160955]
32. Chaki M, Hoefele J, Allen SJ, Ramaswami G, Janssen S, Bergmann C, Heckenlively JR, Otto EA, Hildebrandt F. Genotype-phenotype correlation in 440 patients with NPHP-related ciliopathies. *Kidney Int.* 2011; 80(11):1239–45. [PubMed: 21866095]
33. Otto EA, Helou J, Allen SJ, O’Toole JF, Wise EL, Ashraf S, Attanasio M, Zhou W, Wolf MT, Hildebrandt F. Mutation analysis in nephronophthisis using a combined approach of homozygosity mapping, CEL I endonuclease cleavage, and direct sequencing. *Hum Mutat.* 2008; 29:418–26. [PubMed: 18076122]
34. Otto EA, Ramaswami G, Janssen S, Chaki M, Allen SJ, Zhou W, Airik R, Hurd TW, Ghosh AK, Wolf MT, Hoppe B, Neuhaus TJ, Bockenhauer D, Milford DV, Soliman NA, Antignac C, Saunier S, Johnson CA, Hildebrandt F. Group tGS. Mutation analysis of 18 nephronophthisis associated ciliopathy disease genes using a DNA pooling and next generation sequencing strategy. *J Med Genet.* 2011; 48:105–116. [PubMed: 21068128]
35. Gee HY, Otto EA, Hurd TW, Ashraf S, Chaki M, Cluckey A, Vega-Warner V, Saisawat P, Diaz KA, Fang H, Kohl S, Allen SJ, Airik R, Zhou W, Ramaswami G, Janssen S, Fu C, Innis JL, Weber S, Vester U, Davis EE, Katsanis N, Fathy HM, Jeck N, Klaus G, Nayir A, Rahim KA, Al Attrach I, Al Hassoun I, Ozturk S, Drozd D, Helmchen U, O’Toole JF, Attanasio M, Lewis RA, Nurnberg G, Nurnberg P, Washburn J, MacDonald J, Innis JW, Levy S, Hildebrandt F. Whole-exome resequencing distinguishes cystic kidney diseases from phenocopies in renal ciliopathies. *Kidney Int.* 2014; 85:880–7. [PubMed: 24257694]



**Figure 1. Distribution of molecular NPHP-diagnoses in 18 established NPHP-RC genes, detected in an international cohort of 384 individuals with NPHP-RC**

**A)** Number of families detected with molecular diagnoses in 18 established NPHP-RC genes, which are sorted from left to right by frequency. All individuals were previously screened for the presence of a homozygous *NPHP1* deletion prior to being entered into the present study. **B)** Percentage of individuals with a molecular diagnosis *versus* individuals without a molecular diagnosis in our cohort of 384 unrelated individuals diagnosed with NPHP-RC (left); distribution of molecular diagnoses across 18 NPHP-RC genes (right).



**Figure 2. Comparison of two targeted high-throughput library preparation approaches; MIPs and array-based microfluidic multiplex PCR**

**A)** Schematic design of general work flow of the two high-throughput mutation analysis approaches: Modified Molecular Inversion probe (MIPs) technique (on the left) and array-based microfluidic multiplex PCR technology (Fluidigm™) (on the right). For both techniques library preparation was followed by consecutive next-generation sequencing (NGS). **B)** Overlapping findings in *NPH1-NPH12* by application of either mutation analysis approach: MIPs technique (blue circle) and Fluidigm™ technology (yellow circle). 66 (61%) variants were detected by both methods (overlapping circles). By applying MIPs 28 (26%) additional variants were detected. By using the array-based microfluidic multiplex PCR (Fluidigm™) approach, an additional 12 (11%) variants were identified (see also Supplementary Table 2, Supplementary Table 3).

**Table 1**

Number of novel mutations detected in this study compared to previously reported mutations (HGMD<sup>®</sup>-Professional “Biobase”, 2015) in the genes *NPHP1*, *INVS*, *NPHP3*, *NPHP4*, *CEP290*, *TTC21B*, *PKHD1*, *CC2D2A*, *OFD1*, and *TMEM237*.

Gene	Novel (# mut)	Biobase (# mut)	Percent added
<i>NPHP1</i>	4	42	10%
<i>INVS / NPHP2</i>	1	26	4%
<i>NPHP3</i>	2	53	4%
<i>NPHP4</i>	2	90	2%
<i>CEP290 / NPHP6</i>	2	186	1%
<i>TTC21B / NPHP12</i>	1	39	3%
<i>PKHD1</i>	6	388	2%
<i>CC2D2A</i>	2	52	4%
<i>OFD1</i>	1	129	1%
<i>TMEM237</i>	1	6	17%
<b>total</b>	<b>22</b>	<b>1,011</b>	