

Copy Number Variation: A New Genetic Form of Polycystic Kidney and Liver Disease



Takuya Fujimaru¹ and Eisei Sohara¹

¹Department of Nephrology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan

Kidney Int Rep (2020) 5, 575–576; <https://doi.org/10.1016/j.ekir.2020.03.007>

© 2020 International Society of Nephrology. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

[See Research Letter on Page 727](#)

Autosomal dominant polycystic kidney disease (ADPKD) is the most common inherited kidney disease, occurring in approximately 1 in every 400 to 1000 live births.¹ ADPKD is characterized by development and enlargement of renal cysts with increasing age, which result in end-stage kidney disease in the majority of patients. It is widely accepted that hepatic cysts are one of the most common complications of ADPKD. Mutations in 1 of 2 genes, *PKD1* (chromosome 16p13.3) and *PKD2* (chromosome 4q22.1), account for most cases of ADPKD. Autosomal dominant polycystic liver disease (ADPLD) is an inherited disease characterized by the presence of multiple scattered cysts of biliary origin throughout the liver parenchyma. The occurrence of polycystic liver disease independent from polycystic kidney disease (PKD) has been known for many years. Mutations

in the *PRKCSH* or *SEC63* genes underlie isolated ADPLD. Unlike the pedigrees of ADPKD, 50% of ADPLD pedigrees remain genetically unresolved.² In ADPLD, renal involvement is not typically observed; however, recently, 2 additional genes, *GANAB* and *DNAJB11*, have been identified as candidates for the onset of mild forms of ADPKD, as opposed to ADPKD owing to mutations in *PKD1/2*.¹ Interestingly, patients with mutations in *GANAB* can manifest with both an ADPKD phenotype and an ADPLD phenotype. Therefore, comprehensive molecular diagnosis of ADPKD and ADPLD are important, because some genes (e.g., *GANAB*) exhibit an overlapping disease phenotype.

Advances in next-generation sequencing (NGS) enable simultaneous analysis of a large group of genes in a single test at fairly low cost. Recently, for genetic diagnosis of patients with PKD, a capture-based NGS approach has been reported as an effective comprehensive screening for disease-causing genes of renal cystic disease, including *PKD1/2*.^{1,3} However, even after a

thorough analysis of rare coding single-nucleotide variants and indels in known disease genes, some patients have mutations that still have not been identified. These cases may be negative for a number of reasons, including nongenetic etiologies or lack of knowledge about unknown disease-causing genes. In addition, due to the short DNA sequence read lengths in NGS, large deletions or duplications in the genome have not been detected routinely using whole-exome sequencing (WES).⁴

Copy number variations (CNVs) are defined as segments of DNA that are 1 kb or greater and are present in a variable copy number compared with a reference genome.⁵ CNVs arise via either homologous recombination between repeated sequences or nonhomologous recombination mechanisms that occur throughout the genome. Classes of CNVs include insertions, deletions, and duplications. Maladaptive CNVs have been associated with conditions such as autism, schizophrenia, Crohn's disease, rheumatoid arthritis, type 1 diabetes, obesity, and numerous developmental diseases.⁶ In the field of nephrology, CNVs are associated with an elevated risk of congenital anomalies of the kidney and urinary tract,⁷ which collectively are a diverse group of structural malformations characterized by defects in embryonic kidney development. With the arrival of NGS technologies, sequence-based CNV detection has emerged rapidly as a viable option to identify CNVs. This CNV detection method identifies copy number gains and losses for each target region by comparing them with the normalized sequencing depth of control samples. As a

Correspondence: Eisei Sohara, Department of Nephrology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan. E-mail: esohara.kid@tmd.ac.jp

result, whole-genome sequencing and WES have become primary strategies for NGS technologies in CNV detection. Additionally, several studies have performed CNV detection using targeted panel sequencing data.

Interestingly, Wilson *et al.* reported 2 cases of PKD and polycystic liver disease with a large deletion in *GANAB* and *SEC63*, respectively.⁸ In this study, they investigated a cohort of 128 unrelated individuals with clinically diagnosed ADPKD or ADPLD in whom no mutations were detected; these cases remained genetically unresolved following WES. Researchers performed a WES data-based CNV detection method in these patients. The targeted genes were known genes for ADPKD or ADPLD (i.e., *PKD1*, *PKD2*, *PRKCSH*, *SEC63*, *GANAB*, *ALG8*, *SEC61B*, *PKHD1*, *DNAJB11*, *ALG9*). Then, a large deletion of *GANAB* was detected in a 52-year-old female with liver and kidney cysts. In a 36-year-old female with liver cysts, a large deletion of *SEC63* was observed. These cases represented a new genetic form of ADPKD and ADPLD.

Although CNV analysis using WES data is effective for identifying the genetic cause of undiagnosed cases, this approach has some limitations. First, due to the fragmented nature of WES data, the only reliable way to detect CNV is to utilize the read depth of targeted regions. Therefore, the WES-based CNV detection method

may not allow accurate discovery of CNV in intergenic regions, intronic regions, or in genes with poor WES capture. This method may miss entire CNVs or inaccurately identify CNV breakpoints in these regions. Wilson *et al.* were unable to detect the breakpoints of *SEC63* deletion.⁸ Second, any sequence-based CNV detection methods that include both WES and whole-genome sequencing have high false-discovery rates. Previously published data have shown that these CNV prediction methods have very high false-discovery rates of greater than 85%.⁹ Therefore, we need to confirm the mutations suspected using sequence-based CNV detection methods.

Nevertheless, even with these limitations, WES data-based CNV analysis remains cost-free and is a highly useful screening tool. As found by Wilson *et al.*,⁸ clinically apparent PKD may not occur as a result of one simple mutation, but rather may be genetically complex. Accurate genetic diagnosis is important when considering treatment and genetic counseling. By incorporating methods such as CNV analysis and other comprehensive analysis methods developed to date, it is expected that the genetic mystery of PKD will be further elucidated.

DISCLOSURE

All the authors declared no competing interests.

ACKNOWLEDGMENTS

This work was supported in part by Grants-in-Aid for Scientific Research (B) (19H03672) to ES from the Japan Society for the Promotion of Science.

REFERENCES

1. Bergmann C, Guay-Woodford LM, Harris PC, et al. Polycystic kidney disease. *Nat Rev Dis Prim.* 2018;4:50.
2. Cornec-Le Gall E, Torres VE, Harris PC. Genetic complexity of autosomal dominant polycystic kidney and liver diseases. *J Am Soc Nephrol.* 2018;29:13–23.
3. Fujimaru T, Mori T, Sekine A, et al. Kidney enlargement and multiple liver cyst formation implicate mutations in *PKD1/2* in adult sporadic polycystic kidney disease. *Clin Genet.* 2018;94:125–131.
4. Marchuk DS, Crooks K, Strande N, et al. Increasing the diagnostic yield of exome sequencing by copy number variant analysis. *PLoS One.* 2018;13:e0209185.
5. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet.* 2006;7:85–97.
6. Zarrei M, MacDonald JR, Merico D, et al. A copy number variation map of the human genome. *Nat Rev Genet.* 2015;16:172–183.
7. Nicolaou N, Renkema KY, Bongers EMHF, et al. Genetic, environmental, and epigenetic factors involved in CAKUT. *Nat Rev Nephrol.* 2015;11:720–731.
8. Wilson EM, Choi J, Torres VE, et al. Large deletions in *GANAB* and *SEC63* explain 2 cases of polycystic kidney and liver disease. *Kidney Int Rep.* 2020;5:727–731.
9. Samarakoon PS, Sorte HS, Kristiansen BE, et al. Identification of copy number variants from exome sequence data. *BMC Genomics.* 2014;15:661.