The Journal of Molecular Diagnostics, Vol. 18, No. 3, May 2016

ELSEVIER

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Detection and Quantification of Mosaic Mutations in Disease Genes by Next-Generation Sequencing

Lan Qin,* Jing Wang,*† Xia Tian,* Hui Yu,* Cavatina Truong,* John J. Mitchell,‡ Klaas J. Wierenga,§ William J. Craigen,† Victor Wei Zhang,*† and Lee-Jun C. Wong*†

From the Department of Molecular and Human Genetics,† Baylor College of Medicine and Baylor Miraca Genetics Laboratories, Houston, Texas; the Division of Pediatric Endocrinology,‡ Montreal Children's Hospital, Montreal, Quebec, Canada; and the Department of Pediatrics,§ Section of Genetics, University of Oklahoma Health Science Center, Oklahoma City, Oklahoma*

Accepted for publication January 14, 2016.

Address correspondence to Lee-Jun C. Wong, Ph.D., Department of Molecular and Human Genetics, Baylor College of Medicine and Baylor Miraca Genetics Laboratories, Houston, TX 77030. E-mail: ljwong@bcm.edu.

The identification of mosaicism is important in establishing a disease diagnosis, assessing recurrence risk, and genetic counseling. Next-generation sequencing (NGS) with deep sequence coverage enhances sensitivity and allows for accurate quantification of the level of mosaicism. NGS identifies low-level mosaicism that would be undetectable by conventional Sanger sequencing. A customized DNA probe library was used for capturing targeted genes, followed by deep NGS analysis. The mean coverage depth per base was approximately 800×. The NGS sequence data were analyzed for single-nucleotide variants and copy number variations. Mosaic mutations in 10 cases/families were detected and confirmed by NGS analysis. Mosaicism was identified for autosomal dominant (*JAG1, COL3A1*), autosomal recessive (*PYGM*), and X-linked (*PHKA2, PDHA1, OTC*, and *SLC6A8*) disorders. The mosaicism was identified either in one or more tissues from the probands or in a parent of an affected child. When analyzing data from patients with unusual testing results or inheritance patterns, it is important to further evaluate the possibility of mosaicism. Deep NGS analysis not only provides insights into the spectrum of mosaic mutations but also underlines the importance of the detection of mosaicism as an integral part of clinical molecular diagnosis and genetic counseling. *(J Mol Diagn 2016, 18: 446–453; http://dx.doi.org/10.1016/j.jmoldx.2016.01.002)*

Genetic mosaicism should not be uncommon considering that human cells are derived from a single fertilized zygote through many cycles of DNA replication.[1,2] Depending on the developmental stage in which a mutation arises, mosaicism can be classified into one of three categories: i) germline mosaicism (also known as *gonadal mosaicism*), ii) somatic mosaicism, or iii) gonosomal mosaicism (a combination of germline and somatic mosaicism).[3] The mosaic mutation may be limited to a specific tissue, or levels of mosaicism may be different among different tissues. In addition to well-known cancer-related somatic point mutations,[4,5] copy number mosaicism has also recently been reported in some genomic disorders.[6–8] Failure to identify low levels of mosaic mutations may lead to the misinterpretation of genetic testing results. In particular, an inherited case may be misinterpreted as sporadic due to low-level mosaicism in a carrier parent. Depending on which parent is a mosaic carrier, this misinterpretation may result in a dramatic difference in recurrence-risk assessment in comparison to truly *de novo* events.[9] Therefore, when an apparently *de novo* mutation is identified in a family study, the use of a more sensitive technology for detecting low-level mosaic mutations is essential for accurate recurrence-risk estimates.

Based on our clinical experience, next-generation sequencing (NGS) analysis is a reliable approach to detecting and quantifying genetic variants in terms of mutation load.[10] Here we report the detection of deleterious mosaic mutations in a variety of human genetic disorders, including autosomal dominant, autosomal recessive, and X-linked diseases, underscoring the importance of genetic mosaicism as a disease mechanism and the value of NGS in its detection.

Supported by Baylor College of Medicine, Department of Molecular and Human Genetics.

L.Q. and J.W. contributed equally to this work.

Disclosures: L.Q., J.W., X.T., H.Y., C.T., W.J.C., V.W.Z., and L.-J.C.W. are employees of Baylor Miraca Genetics Laboratories, which offers next-generation sequencing–based genetic testing on a fee-for-service basis.

http://dx.doi.org/10.1016/j.jmoldx.2016.01.002

Table 1 Estimated Mosaic Prevalence in Genes Tested in Our Patient Cohort

Gene	GenBank ID*	No. of probands tested positive	No. of families studied	No. of apparently *de novo* probands (%)	No. of mosaic probands	No. of mosaic parents	Mosaic cases among positives, %	Mosaic cases in families studied, %	Mosaic cases in apparently *de novo* cases, %
JAG1	NM_000214.2	20	3	1 (33)	0	1	5.00	33.3	100
COL3A1	NM_000090.3	8	8	1 (12.5)	0	1	12.5	12.5	100
PYGM	NM_005609.2	40	8	1 (12.5)	1	0	2.50	12.5	100
PHKA2	NM_000292.2	52	23	1 (4.3)	0	1	1.90	4.3	100
PDHA1	NM_000284.3	124	58	7 (12)	3	0	2.40	5.2	43
OTC	NM_000531.4	220	175	36 (20.6)	1	1	0.90	1.1	5.6
SLC6A8	NM_005629.1	57	24	10 (41.7)	0	1	1.80	4.2	10
Total		521	299	57 (19)	5	5	1.9	3.3	17.5

*GenBank website: *http://www.ncbi.nlm.nih.gov/genbank*.

Materials and Methods

Patients and DNA

Patients' samples were submitted to the Baylor Miraca Genetics Laboratories (formerly, Medical Genetics Laboratories at Baylor College of Medicine; Houston, TX) for DNA testing of a large number of disease-associated genes. DNA was extracted with commercially available DNA isolation kits according to the manufacturers' protocols (Gentra Systems Inc., Minneapolis, MN). This study was performed, according to an Institutional Review Board–approved protocol, at the Baylor College of Medicine. All patients' information presented in this article was sufficiently anonymized, but sex, age, clinical phenotype, and the mutations detected are included for interpretation.

Sequence Analysis

Amplicon-based Sanger sequencing analysis using target sequence–specific primers was performed on an ABI 3730xl sequencer (Applied Biosystems, Foster City, CA) and analyzed using Mutation Surveyor software version 4.0.5 (SoftGenetics, State College, PA) according to the published procedures.[11] Human Genome Variation Society guidelines were followed for mutation nomenclature.[12] The numbering of nucleotide position was based on c.1 as the A nucleotide of the ATG translation initiation codon in the reference cDNA sequence. The reference sequences and GenBank accession numbers of each gene are listed in Table 1 (*http://www.ncbi.nlm.nih.gov/genbank*).

Capture-based NGS was performed using custom-designed probe libraries and analyzed on HiSeq2000 or MiSeq (Illumina, San Diego, CA) according to the manufacturer's instructions. Sequence data were demultiplexed by Casava software version 1.7 (Illumina) and further processed by NextGENe software version 4.0.5 (SoftGenetics) for alignment. The mean sequence coverage per base was approximately 800×. All mutation calls were reviewed and confirmed by Sanger sequencing. An in-house bioinformatics pipeline was used for variant annotation.[13–15]

Results

Mosaicism in Autosomal Dominant Disorders

Case 1 was a 4-month-old male infant who presented with worsening liver function and cholestasis. An NGS-based cholestasis panel of seven genes—*ABCB4, ABCB11, AKR1D1, ATP8B1, JAG1, SERPINA1*, and *SLC25A13*—was performed. A heterozygous null mutation, c.1499delG (p.G500Vfs*64), was detected in the *JAG1* gene. This mutation was subsequently confirmed by Sanger sequencing (Figure 1). The *JAG1* gene encodes jagged-1, an important ligand in Notch signaling and embryonic development. The same c.1499delG mutation in *JAG1* has previously been reported in the heterozygous state in patients with Alagille syndrome,[16] an autosomal dominant disorder that leads to a paucity of bile ducts in the liver, vertebral anomalies, cardiovascular malformations, and an ocular finding of posterior embryotoxon. Blood samples from the patient's asymptomatic biological parents were analyzed by Sanger sequencing for the *JAG1* mutation. Although there was no evidence of the mutation in the father, a very small peak was observed in the Sanger chromatogram of the mother's blood specimen (Figure 1). A follow-up NGS analysis confirmed the presence of a 9% low-level c.1499delG (p.G500Vfs*64) mutation in the mother's blood specimen. About 7% of Alagille syndrome patients have been reported to have *JAG1* gene deletions, whereas 89% have point mutations.[17–22] The collective evidence suggests that the mosaic mutation in the mother arose as a post-zygotic mitotic event and is an example of gonosomal mosaicism. The recurrence risk for having an affected offspring in this family may be as high as 50%.

Similarly, Case 2 was a 3-year-old girl with a heterozygous c.3491G>A (p.G1164E) mutation detected in the *COL3A1* gene (Table 2). This mutation was reported previously in a patient affected with autosomal dominant Ehlers-Danlos syndrome type IV.[23] The glycine at amino acid position 1164 of the COL3A1 (collagen, type III, α_1) protein is highly conserved during evolution, and is located within the Gly-X-Y collagen triple-helix repeat. The follow-up targeted Sanger and NGS analyses confirmed that the blood specimen from the mother