The Journal of Molecular Diagnostics, Vol. 16, No. 2, March 2014

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Validation for Clinical Use of, and Initial Clinical Experience with, a Novel Approach to Population-Based Carrier Screening using High-Throughput, Next-Generation DNA Sequencing

Stephanie Hallam, Heather Nelson, Valerie Greger, Cynthia Perreault-Micale, Jocelyn Davie, Nicole Faulkner, Dana Neitzel, Kristie Casey, Mark A. Umbarger, Niru Chennagiri, Alexander C. Kramer, Gregory J. Porreca, and Caleb J. Kennedy

From Good Start Genetics, Inc., Cambridge, Massachusetts

Accepted for publication October 4, 2013.

Address correspondence to Stephanie Hallam, Ph.D., 237 Putnam Ave, Good Start Genetics, Inc., Cambridge, MA 02139. E-mail: shallam@gsgenetics.com.

Traditional carrier screening assays are designed to look for only the most common mutations within a gene owing to cost considerations. Although this can yield high detection rates in specific populations for specific genes (such as cystic fibrosis in Caucasians), they are suboptimal for other ethnicities or for patients of mixed or unknown ethnic background. Next-generation DNA sequencing provides an opportunity to provide carrier screening using more comprehensive mutation panels that are limited primarily by information about the clinical impact of detected sequence changes. We describe a next-generation DNA sequencing-based assay capable of reliably screening patient samples in a timely and comprehensive manner. The analytic accuracy in a research setting has been documented. Here, we describe the additional studies performed to ensure the accuracy (analytic validity) and robustness of our assay for use in clinical practice and provide data from our experience offering this testing. Our clinical experience using this approach to screen 11,691 *in vitro* fertilization patients has identified 449 mutant alleles: 447 in carriers and 2 in an affected individual. In total, we found 87 distinct mutations in 14 different genes. Approximately one quarter of the mutations found are not included in traditional, limited, mutation panels, including 16 known mutations unique to our panel, and novel truncating mutations in several genes. *(J Mol Diagn 2014, 16: 180—189; http://dx.doi.org/10.1016/j.jmoldx.2013.10.006)*

Next-generation DNA sequencing (NGS) holds the promise of providing high-throughput, accurate carrier screening for multiple genes and multiple mutations in a highly efficient manner in clinical laboratories.[1] To date, NGS has found application in evaluating affected individuals for numerous multigene disorders and for elucidating the correct diagnosis and hence treatment of patients with various forms of cancer,[2] but carrier screening assay development has been slower because of concerns about accuracy[3,4] and throughput in a clinical setting. Here, we describe the validation for clinical use of the first multigene carrier screening assay using NGS technology to be offered in the United States and provide information about our clinical experience thus far.

The genes included in our assay have proven clinical validity (the association between mutations in the gene and the related disorder has been established).[5,6] In addition, carrier testing for these genes is recommended by the American College of Medical Genetics and Genomics, the American Congress of Obstetricians and Gynecologists, and/or are assessed routinely in persons of Ashkenazi Jewish descent because of the increased carrier frequency in this population

Disclosures: All authors are employees of Good Start Genetics, Inc. (Cambridge, MA). In addition, all full-time employees (S.H., H.N., V.G., J.D., N.F., D.N., K.C., M.A.U., N.C., A.C.K., G.J.P., and C.J.K.) have stock options in the company and hence are potential future shareholders in Good Start Genetics, Inc. M.A.U., C.J.K., G.J.P., and V.G. are listed as inventors on patents or patent applications related to this work.

http://dx.doi.org/10.1016/j.jmoldx.2013.10.006

and/or their clinical severity. The NGS panel validated the following diseases (gene symbols are shown in parentheses): Canavan disease (*ASPA*), cystic fibrosis (*CFTR*), glycogen storage disorder type 1a (*G6PC*), Niemann-Pick disease (*SMPD1*), Tay-Sachs disease (*HEXA*), Bloom syndrome (*BLM*), Fanconi anemia C (*FANCC*), familial hyperinsulinism (*ABCC8*), maple syrup urine disease type 1A (*BCKDHA*) and type 1B (*BCKDHB*), Usher syndrome type III (*CLRN1*), dihydrolipoamide dehydrogenase deficiency (*DLD*), familial dysautonomia (*IKBKAP*), mucolipidosis type IV (*MCOLN1*), and Usher syndrome type 1F (*PCDH15*). Additional genes and mutations, for which the gene or mutations are not readily evaluated by NGS, are assessed by alternate methodologies to ensure detection of clinically important mutations. The validations for these alternate methodologies followed standard, clinical laboratory procedures and are not described here.

Clinical carrier screening assays traditionally have assessed a limited set of mutations, typically those prevalent in specific ethnic groups. NGS provides the possibility of finding a much larger set of sequence variants across many ethnic groups. Because NGS is not limited to a small number of mutations, there is an additional dilemma related to the interpretation and reporting of the variants detected. Certainly, this is one of the most challenging areas associated with the advent of NGS technology in the carrier screening arena. Extensive discussions and feedback from our clinical and genetic counselor advisory boards has indicated that specificity was an important factor in offering carrier screening using NGS. In addition, others advocating a responsible approach to offering full sequencing recommend reporting only those variants that are known to have a clinical impact[7] (ie, no variants of unknown significance).

After determination of the panel of genes to be assessed, it was essential to our clinical approach to rigorously establish the variants that must be detected for multifold reasons, including the following: i) to complete a comprehensive evaluation of all available information about each variant for each gene to determine the full list of mutations that were considered pathogenic; ii) to put this information into a variant database that could be curated, managed, and updated periodically to ensure that new information about variants could be added and that the panel would remain pertinent (this information then can be used for future cases); iii) to have a system that would integrate with the NGS data analysis pipeline to ensure rapid and consistent calling of clinically relevant mutations; and iv) to ensure that all variants that passed our filters for being pathogenic and clinically important could be either readily detected by the NGS assay or an alternate methodology to ensure detection. These alternate methodologies have all been used in clinical laboratories performing carrier testing for a number of years and hence are not discussed further (they were all validated before use in our laboratory).

This latter consideration is important and one that is gaining awareness among those using or considering using NGS technology for clinical applications. Some genes, gene regions, or mutations are particularly problematic for NGS and vary with the technology and analysis method(s) used. For example, pseudogenes, GC-rich regions, homopolymers, large deletions, and complex insertions/deletions all can be problematic depending on the specific sample preparation, sequencing, and analysis method(s) used.[8] Hence, our approach was to maximize the extent of what is addressable by NGS and to ensure detection of all clinically important mutations.

Because NGS has the ability to detect additional sequence variants in genomic regions sequenced at high quality and depth, we established a pathway for assessment of novel, reportable variants whereby a patient's DNA sequence is scanned for variants that meet one or more of the following criteria: i) occurs at a conserved donor or acceptor splice site (± 2 bases of intron), ii) generates a premature stop codon (nonsense mutation), or iii) generates a frame-shift in the protein sequence. When a stop codon or frame-shift mutation is present, the position of the mutation relative to the 3′ most truncating mutation previously described for the disorder also is taken into account. These mutations are reported as predicted to be pathogenic and further increase the detection rates for each disorder beyond what has been reported previously as pathogenic.

Before beginning the validation of the clinical NGS assay, it was necessary to define acceptable analytic sensitivity and specificity criteria. To use an NGS assay for carrier screening, we required a high level of clinical confidence that the assay would not miss carriers. Hence, we were most concerned about the false-negative rate. All positive results were confirmed by Sanger sequencing, thereby eliminating false-positive results before clinical reporting. Studies detailing assay design and analytic accuracy in a research setting already have been described.[1] In brief, it was shown that NGS could achieve a false-negative rate of 2.52×10^{-4} (95% Wilson binomial confidence interval, 1.29×10^{-5} 1.42×10^{-3}) for single-nucleotide variants; a single false-negative call occurred in a sample previously characterized as aneuploid.[9] For insertions, deletions, or the more complex mutations that are indels, three false-negative results occurred. However, there were no false-negative results for any position in the sequence where a pathogenic mutation of interest occurred, as defined in our variant database, and/or that was not covered by an alternate methodology. Our goal therefore was to show a false-negative rate of zero in the validation studies described later.

As mentioned earlier, in this study, Sanger sequencing was used as the comparator method for the analytic accuracy studies, despite the high cost and resource burden of fully sequencing each sample for each of the genes in the NGS panel. This can be leveraged for future validations.

For the purpose of clinical testing, it is necessary to be able to Sanger sequence any amplicon for any gene that is assessed by NGS to confirm the presence of the mutation, and therefore we also validated Sanger sequencing for approximately 250 amplicons.

To offer a clinical-grade carrier screening assay with high accuracy and precision (reproducibility), robustness,