BAYLOR GENETICS

Performance of the Newly Developed Non-Invasive Prenatal Multi-Gene Sequencing Screen

INTRODUCTION

Here we describe the analytical performance of the newly developed non-invasive prenatal multi-gene sequencing screen. This test screens for clinically significant and life-altering single gene disorders that are outside the scope of current non-invasive prenatal tests. The disorders include skeletal, cardiac, and developmental single gene disorders, many of which often occur in the absence of a family history of the condition. Cumulative incidence of all disorders in this test is estimated to be 1:600. This test is available to women with singleton pregnancies greater than or equal to 9 weeks gestation. It is designed to detect predominantly *de novo* and paternally inherited variants in circulating cell-free fetal DNA in maternal blood. This screen does not detect copy number changes and cannot distinguish maternally inherited variants in the fetus from maternal background.

The information here summarizes the technical details of this test in validation and commercial samples. Briefly, a SNP-based fetal fraction calculation method was developed that yielded concordance with the established Y-chromosome method. We also demonstrate that this test can detect benign and disease-causing DNA changes in cell-free plasma DNA using a combination of spike-in samples and samples from pregnant women. Seventy-six samples (combination of validation and commercial samples) from pregnant women (each paired with the biological father) were analyzed. From these samples, the analytical sensitivity and specificity were determined to be >99% and the fetal fraction lower limit was established at 4.5%. *De novo* pathogenic variants in *COL1A1*, *FGFR3*, and *RIT1* were successfully identified in this cohort and confirmed in pregnancies with abnormal ultrasound findings and/or known paternal history of conditions included in this screening panel.

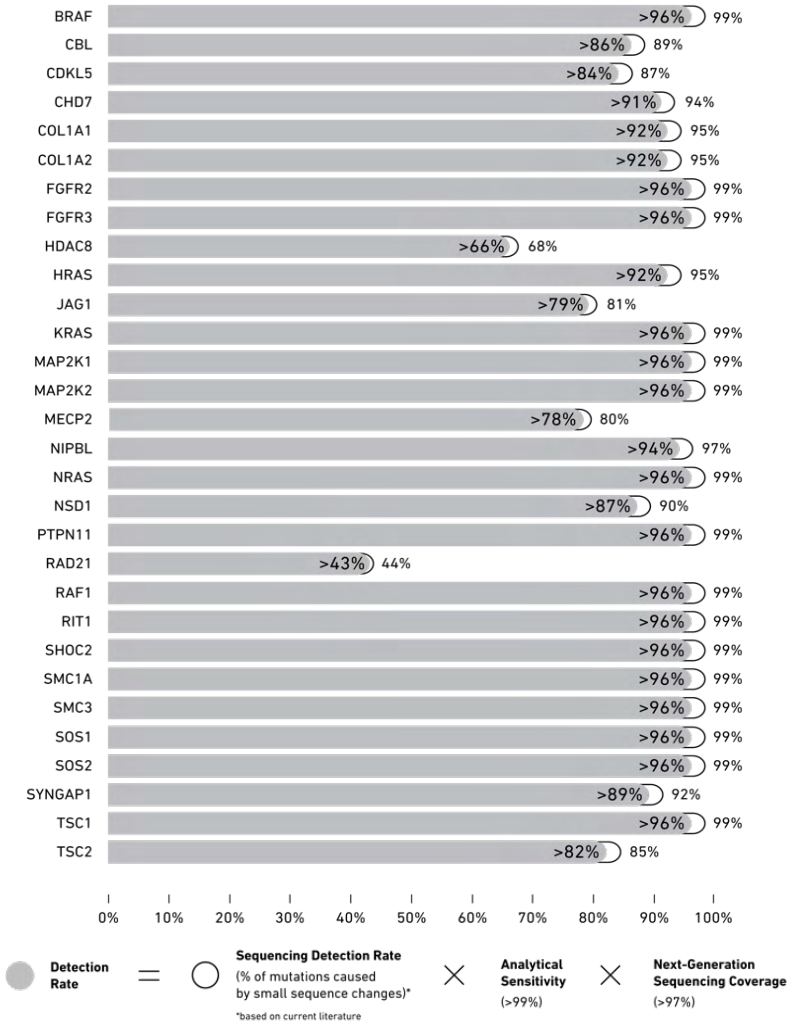

Figure 1. Detection rates are a product of the sequencing detection rate (based on current literature), analytical sensitivity, and next-generation sequencing coverage.

DISORDERS

The genes contained in the panel were chosen by a curation process focused on the selection of genes with a high incidence of *de novo* variants leading to significant medical problems.[1,2] Many of the genes are associated with disorders that may not be detectable by ultrasound until late in gestation or may have no ultrasound findings.[1,2] These disorders are either autosomal dominant or X-linked

conditions. Figure 1 lists the genes sequenced and the detection rate for each gene. The detection rate is dependent on the percentage of all known likely pathogenic and pathogenic variants in a gene caused by small sequence changes (single nucleotide variants and indels) that are identified by the assay. Validation of this assay demonstrates a minimum analytical sensitivity of 99% in the covered regions. Greater than 97% of sequencing regions are covered at 200X depth of read.

WORKFLOW AND QC METHODS

This assay is a trio test that requires a maternal blood sample and a paternal sample (blood or saliva) to obtain three types of DNA: plasma cell-free DNA, maternal genomic DNA, and paternal genomic DNA. All three samples are sequenced and the data are analyzed in aggregate to inform the final report.

During the PCR and sequencing processes, random DNA changes, referred to as artifacts, can be introduced, which can lead to an increase in background noise or false positives. To aid in separating the true DNA changes from artifacts, unique DNA sequences referred to as molecular barcodes are used. Once plasma cell-free DNA is extracted from maternal blood, molecular barcodes are used to add unique labels to the cell-free DNA molecules before PCR.[3,4] The technical process then includes library construction, target gene enrichment, and next-generation sequencing (NGS). During analysis of the sequence data, the distribution of molecular barcodes is used to predict which variants are representative of the original cell-free DNA material and which were introduced later as a by-product of the technical process.

Additionally, included in the assay is a set of over 100 unique genome-wide single nucleotide polymorphisms (SNPs) that is analyzed to calculate fetal fraction. SNPs are normal DNA base pair changes found throughout the genome. Although not pathogenic, SNPs inherited from the father are assessed in the cell-free DNA for every sample to ensure that these DNA changes are accurately identified in the cell-free plasma DNA and to provide an estimate of fetal fraction. In addition to fetal fraction estimation, paternal DNA is used for quality control and variant classification.

Once the sequence data are obtained, they are analyzed and the variants are identified. The variants are curated to determine if any meet criteria for classification as pathogenic or likely pathogenic. The classification of genetic variants is based on current information available to us and guidelines issued by ACMG and/or Baylor Genetics. All pathogenic and likely pathogenic variants detected are confirmed in a second aliquot of cell-free DNA using an amplicon-based NGS assay. The secondary method uses gene-specific primers to enrich the targeted region followed by deep sequencing (>10,000X) to confirm the variants in the cell-free DNA.

VALIDATION DATA SUMMARY

Fetal Fraction

For this assay, fetal fraction is calculated based on the detection of the unique SNPs analyzed across the genome. Two methods were used to verify the accuracy of the fetal fraction calculation method employed by Baylor Genetics. For the first method, spike-in studies were conducted using extracted DNA from an affected proband sample and genomic DNA from the proband's mother and father. DNA from each proband sample was added to maternal DNA to achieve final concentrations of the proband DNA ranging from 1 to 20% (Figure 2A). A strong correlation (R^2=0.97) was observed between the calculated fetal fraction and the estimated fetal fraction based on the percentage of proband DNA in the corresponding maternal DNA.

For the second approach to validate the fetal fraction process, the fetal fraction was calculated in plasma from 31 pregnant women with male fetuses (Figure 2B). The observed fetal fraction using the Baylor Genetics method was compared to the fetal fraction using the Y chromosome of each fetus. A strong correlation (R^2=0.96) was observed between both fetal fraction methods.

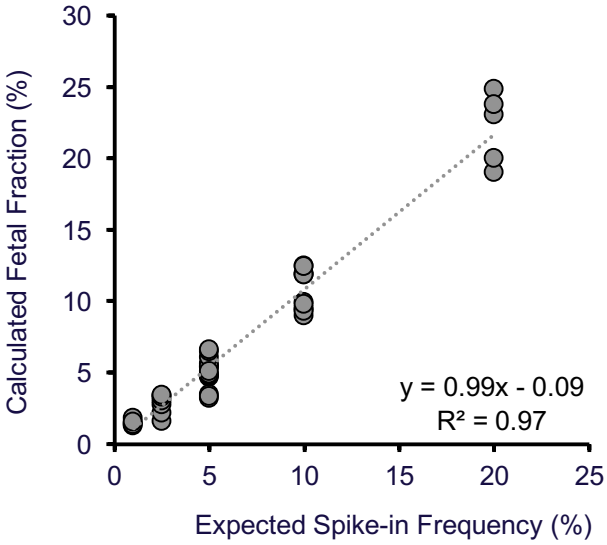

Figure 2A. Baylor Genetics fetal fraction calculation using spike-in DNA.

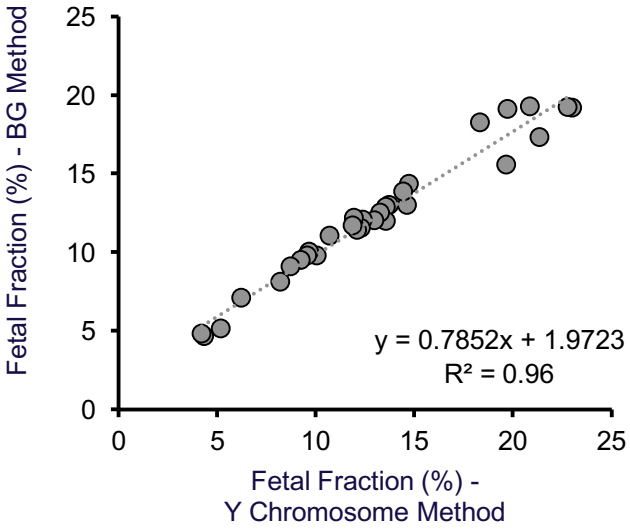

Figure 2B. Baylor Genetics fetal fraction calculation compared to Y chromosome

Detection of Pathogenic Variants Using Spike-in DNA

To determine the sensitivity of this screen to detect *de novo* variants in specific genes, spike-in studies were conducted using control trio samples consisting of extracted DNA from a proband and from the proband's mother and father. The known DNA changes were a combination of single nucleotide changes and indels. DNA from the proband samples were added to maternal DNA to achieve final concentrations of the proband DNA ranging from 2.5% to 10%. Table 1 includes a subset of the spike-in samples tested. Paternal DNA was also sequenced to aid in fetal fraction and inheritance pattern determinations.

GENE	Variant	2.5%	5%	10%
SYNGAP1	c.3190C>T	Y	N/A	N/A
KRAS	c.458A>T	Y	Y	N/A
NRAS	c.35G>A	Y	Y	N/A
TSC2	c.1864C>T	Y	Y	N/A
SHOC2	c.4A>G	Y	Y	Y
SMC1A	c.802_804del	Y	Y	Y
COL1A1	c.3709_3716delAGCCTGAG	Y	Y	Y
SOS1	c.508A>G	Y	Y	Y
PTPN11	c.1505C>T	Y	Y	Y

Table 1. Variant detection of spike-in DNA. N/A, represents DNA that was limited in certain samples and not all concentrations could be tested. Y (Yes), correct variant was detected.

Clinical Performance in Pregnant Women

The accuracy, sensitivity, and specificity of the assay presented here is based on the detection of changes in cell-free plasma DNA as compared to parental genomic DNA in 76 pregnant women, with or without clinical or family history related to conditions screened by this test. The samples included here are a combination of validation and commercial samples. The gestational ages at the time of sampling ranged from 10 weeks to 40 weeks (Figure 3). The fetal fractions ranged from 4.5 to approximately 30%.

In the analysis of the 30 disease-related genes in this assay, true positive calls were defined as either inherited paternal changes or *de novo* changes—those not detected in either the maternal or paternal samples. Five hundred fifty-four true positive calls were detected in the cell-free plasma DNA from the 76 samples. The majority of the changes were inherited and were not classified as pathogenic/likely pathogenic. *De novo* pathogenic variants were identified in *FGFR3*, *COL1A1*, and *RIT1* in three pregnancies. The pathogenic *de novo* variants identified were confirmed by analysis of invasive or postnatal specimens. True negative calls were defined as the reference DNA sequence (Human Genome version 19) detected in both parents and the cell-free plasma DNA. For the true negative calls (both parents are homozygous for the reference allele), over 8 million nucleotides in the 30 genes of interest were accurately detected in the cell-free plasma DNA from the 76 samples. False positive calls, which could potentially be *de novo* calls, occurred when both parents had the reference sequence but the cell-free plasma DNA showed a non-reference base. A total of seven analytical false positives were detected from five patients in our primary assay using capture-based NGS, but none of the calls were confirmed with the amplicon-based NGS confirmatory test, thus they did not meet criteria for reporting. Note that these variants were not considered pathogenic or likely pathogenic and they were only considered as analytical false positives for our primary assay. False negative calls were defined as DNA changes (either inherited paternal changes or de novo changes) that should have been present but were not detected in the cell-free plasma DNA. There were no false negatives in the genes of interest and in the SNPs across the genome.

As seen in Table 2, both analytical sensitivity and specificity are >99%. Analytical positive predictive value (PPV) is calculated to be >99% and the analytical negative predictive value (NPV) is >99%. Analytical PPV is the probability that individuals with an identified variant truly have the variant. Note that this is different from clinical PPV, which is the probability that individuals with a positive screening test truly have the condition. Clinical PPV is dependent on the incidence of the disorder. Clinical PPV is also dependent on the false positive rate, the severity of the phenotype, and penetrance, which will vary with the specific change detected.

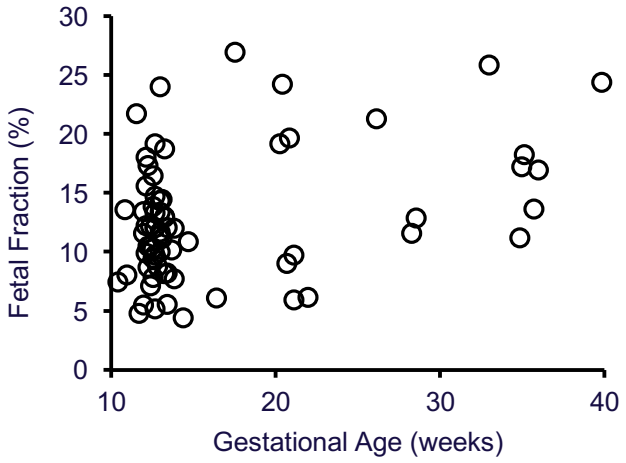

Figure 3. Gestational age and fetal fraction correlation for 76 trio samples.

True positive calls	TP	554
True negative calls	TN	8038792
False positive calls from the primary non-invasive multi-gene sequencing screen	N/A	7
False positive calls following the secondary amplicon-based next-generation sequencing assay*	FP	0
False negative calls	FN	0
Sensitivity	TP / (TP + FN)	>99%
Specificity	TN / (TN + FP)	>99%
Positive predictive value	TP / (TP + FP)	>99%
Negative predictive value	TN / (TN + FN)	>99%

Table 2. Sensitivity and specificity in plasma samples (76 trios). TP, true positive. TN, true negative. FP, false positive. FN, false negative.
* Although the false positives detected were not classified as pathogenic or likely pathogenic, the secondary confirmation method was used to clarify whether the variants were *de novo* variants or false positives.

CONCLUSION

Here we describe the first clinical non-invasive prenatal test to screen for multiple single gene disorders. The 30 genes included focus on dominant conditions with a high prevalence rate of *de novo* changes. Additionally, Baylor Genetics employed a molecular barcoding process to aid in separating the true DNA changes from artifacts and a method was developed to calculate the fetal fraction. This assay can accurately sequence cell-free DNA from the mother's plasma and can detect DNA changes (both benign and disease-causing) with a sensitivity and specificity >99%.

REFERENCES

1. Chitty, LS, et al. Non-invasive prenatal diagnosis of achondroplasia and thanatophoric dysplasia: next-generation sequencing allows for a safer, more accurate, and comprehensive approach. Prenatal Diagnosis. 2015; 35: 656-662.

2. Toriello HV, et al. Statement on guidance for genetic counseling in advanced paternal age. Gen Med. June 2008; 10(6): 457-460.

3. Stahlberg A, et al. Simple, multiplexed, PCR-based barcoding of DNA enables sensitive mutation detection in liquid biopsies using sequencing. Nuc Acid Res. April 2016; 44(11): e105.

4. Newman, AM, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol. May 2016; 34(5): 547-555.

ACKNOWLEDGEMENTS

This test was designed, developed, and validated by Baylor Genetics. Following development and validation, Natera collaborated with Baylor Genetics to assist in clinical introduction of the test, particularly in providing samples to help establish the lower limit of fetal fraction. This test is performed and reported by Baylor Genetics.