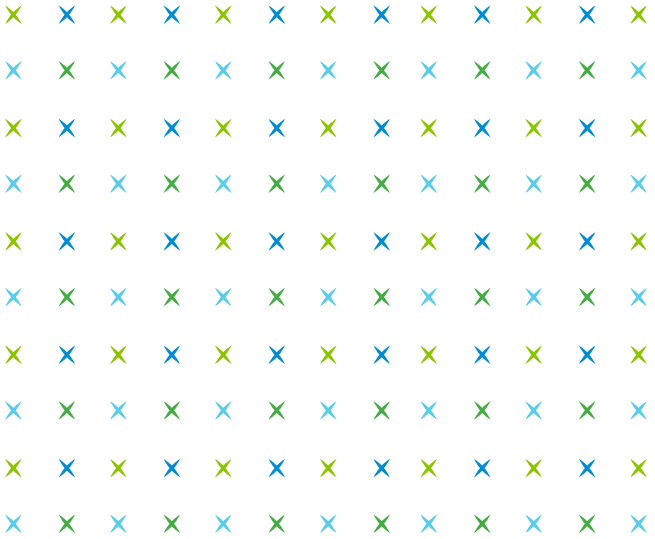
Performance of the Newly Developed Non-Invasive Prenatal Multi-Gene Sequencing Screen



Introduction

Here we describe the analytical performance of the non-invasive prenatal multi-gene sequencing screen, which includes the original validation (V1), and a subsequent validation (V2) with enhancements using clinical samples. The difference between the V1 and V2 methods is the removal of the paternal sample and egg-donor sample requirement.

We demonstrate equivalency in analytical performance between the V1 and V2 methods.

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 equal to or greater than 9 weeks gestation. The test is designed to detect predominantly *de novo* and paternally inherited variants in circulating cell-free fetal DNA present in maternal blood. This screen does not detect copy number changes, and it cannot distinguish maternally inherited variants in the fetus from maternal background.

The information here summarized 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 demonstrated 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. For V1 validation, 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 was >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. For V2 validation, a retrospective study of commercial clinical samples confirmed concordance in fetal fraction estimation (n= 268) and variant calls (n=2,195), before and after interpretation with paternal and egg-donor NGS data.

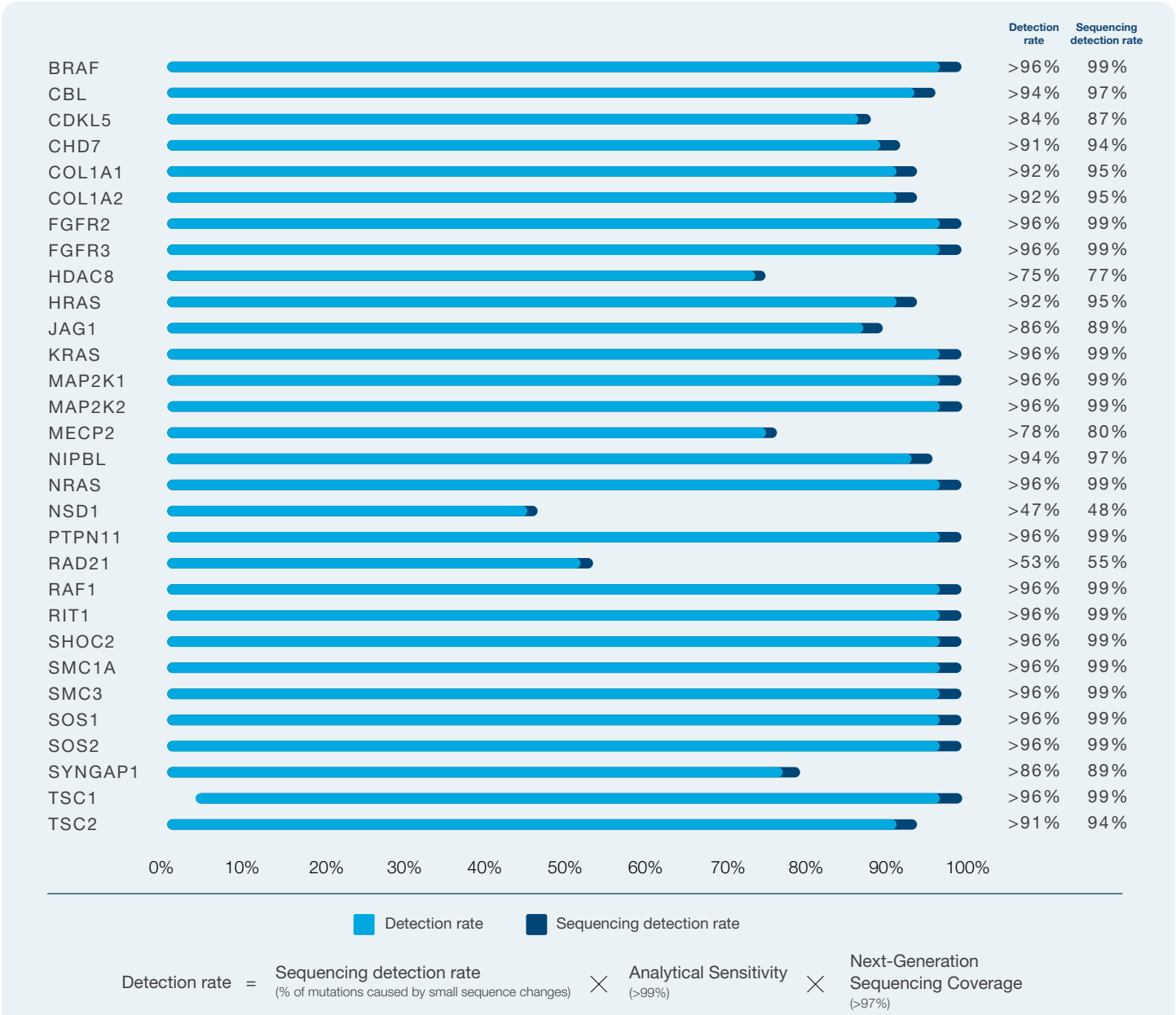


Figure 1. Detection rates are a product of sequencing detection rate, analytical sensitivity, and next-generation sequencing coverage.

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 are not detectable by ultrasound until late in gestation or 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

The V1 test was offered as a trio test that required a maternal blood sample, a paternal sample (blood or saliva), and an egg-donor sample, if applicable (blood or saliva). The V2 assay only requires maternal blood to identify variants in the fetus using the plasma cell-free DNA and maternal genomic DNA.

All samples were sequenced and the data were analyzed as a trio (V1 method), or using the maternal and fetal information (V2) 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 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 were analyzed to calculate fetal fraction. SNPs are normal DNA base pair changes found throughout the genome. In the V1 method, SNPs inherited from the father were assessed in the cell-free DNA to provide an estimate of fetal fraction and used for variant classification. Once the sequence data was obtained, variants were curated to determine if any meet criteria for classification as pathogenic or likely pathogenic. The classification of genetic variants was based on current information available, 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. For V2 validation, only maternal SNP data was used and a retrospective study confirmed concordance in fetal fraction estimation (n= 268) and result/variant curation calls (n= 2,195), with or without paternal and egg-donor NGS data.

Validation Data Summary

Fetal Fraction

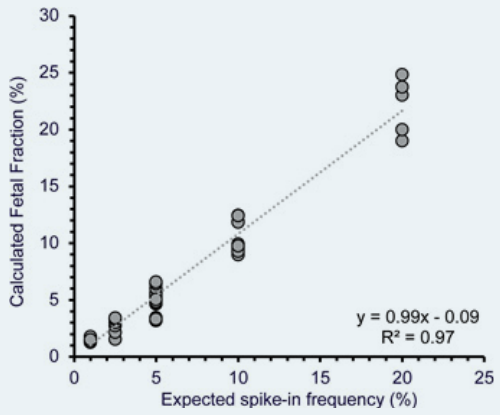


Figure 2A. Baylor Genetics fetal fraction calculation using spike-in DNA

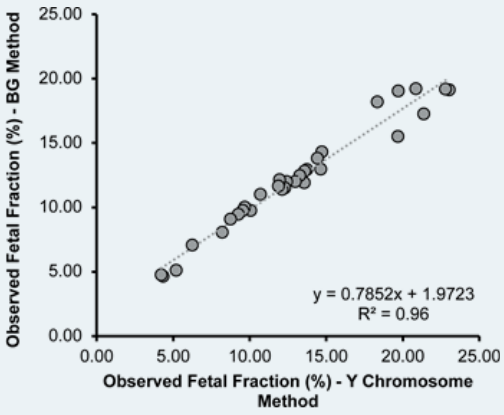


Figure 2B. Baylor Genetics fetal fraction calculation compared to Y chromosome counting.

In V1, fetal fraction was calculated based on all informative transmitted SNP loci from maternal, paternal, and egg-donor (if applicable) samples. To simplify the sample collection and workflow in a subsequent version of the test (V2), fetal fraction calculation was validated without genomic DNA from the proband's paternal and egg-donor. A regression algorithm was used to infer the paternal and egg-donor informative SNPs, without the paternal and egg-donor NGS data. Only NGS results from maternal plasma cfDNA and maternal genomic DNA was used for variant calls and fetal fraction calculation based on methods previously published (PMID: 30692697). A total of 268 cases (50 historical clinical cases, 168 cases with low FF <5% and 50 cases from three different live batches) spanning 5 clinical scenarios (negative, positive, low FF, QC failure, low data coverage) were reviewed for FF estimation and data was concordant with and without paternal and/or egg-donor NGS data. (Figure 3A, 3B and 3C).

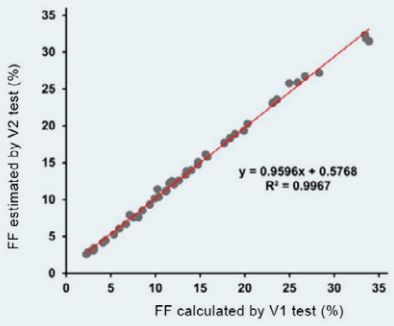


Figure 3A. Comparison of the FF calculated by V1 and V2 test for 50 clinical cases

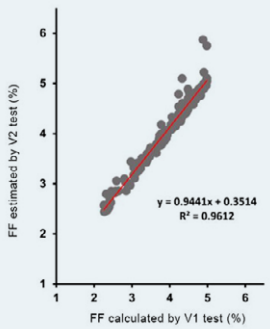


Figure 3A. Comparison of FF calculated by V1 and V2 test for 168 low FF cases

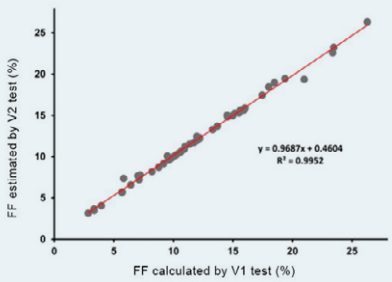


Figure 3A. Comparison of FF calculated by V1 and V2 test for 50 live batch cases

Detection of Pathogenic Variants Using Spike-in DNA

In V1 validation, spike-in studies were conducted to determine the sensitivity of this screen to detect *de novo* variants in specific genes. Control trio samples were used, 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.

		Variant Detected		
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>T	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

V1 Validation

For V1 validation, the accuracy, sensitivity, and specificity of the assay was 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 were a combination of validation and commercial samples. The gestational ages at the time of sampling ranged from 10 weeks to 40 weeks (Figure 4). 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*, 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. However, since none of the calls were confirmed with the amplicon-based NGS confirmatory test, 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 that 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 analytic PPV 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, the false positive rate, the severity of the phenotype, and penetrance, which vary with the specific change detected. Clinical validation (PMID: 30692697) involving 422 pregnancies, confirmed 30 true positive, 127 true negative, zero false positive, and zero false negative results.

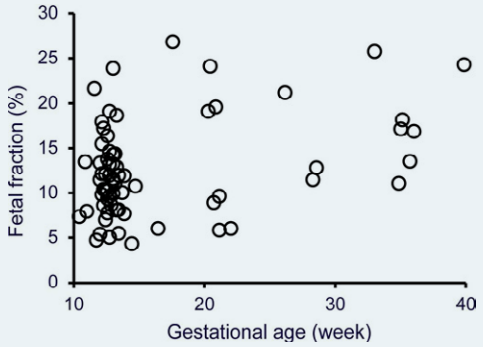


Figure 4. Gestational age and fetal fraction correlation for 76 trio samples.

True Positive Calls	TP	554
True Negative Calls	TN	8,038,792
False Positive Calls following 1° Method*		7
False Positive Calls following 2° Method**	FP	0
False Negative Calls	FN	0
Sensitivity	TP / (TP + FN)	>99%
Specificity	TN / (TN + FP)	>99%
Positive Predictive Value	TP / (TP + FP)	>99%
Positive Predictive Value	TN / (TN + FN)	>99%

Table 2. Sensivity and specificity in plasma samples (76 trios). True positive, TP. True negative, TN. False positive, FP. False negative, FN.

* Refers to the primary non-invasive prenatal multi-gene sequencing screen.
** Refers to the secondary amplicon-based next-generation sequencing assay. 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.

V2 Validation

For the V2 validation, 2,195 consecutive clinical cases were retrospectively evaluated for the effect of removing paternal and egg-donor NGS data on clinical outcome. All variants found in the 87 positive cases, including 49 pathogenic and 14 likely pathogenic variants, had 100% concordant results with or without paternal and/or egg-donor sequencing data. Two cases with inconclusive results, caused by pathogenic or likely pathogenic variants heterozygous in maternal DNA, were also concordant between the V1 and V2 method (Table 3). Family history information will continue to be requested on the test requisition to assist with the variant curation process.

	With paternal and/ or egg-donor data	Without paternal and/ or egg-donor data
NEGATIVE	2,106	2,106
POSITIVE	87	87
INCONCLUSIVE	2	2

Table 3. For the 2,195 consecutive clinical cases, clinical outcomes were 100% concordant between V1 and V2 tests.

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%. Subsequent validation of methodology improvements for V2 allows us to modify the sample requirements to include only maternal plasma. Fetal fraction calculation and clinical outcomes comparing V1 and V2 were 100% concordant with the removal of paternal and egg-donor sample requirements.

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.

13011 McCallen Pass, Building A Suite 100 | Austin, TX 78753 | natera.com