Verification of Software and Amplification-Based Nanopore Sequencing Solution to Characterize Complex Variants in 11 Challenging, High-Frequency Carrier Genes

P804

Juliette A. Baker, John N. Milligan, Walairat L. Wolf, Jon Kemppainen, Justin Janovsky, Jonathan Turner, Monica Roberts, Theodore Markulin, Eduardo Priego, Mahesh Yarasi, Mia K. Mihailovic, Bryan J. Killinger, Bradley Hall, and Connor A. Parker,

Asuragen, a Bio-Techne Brand, Austin, TX

Summary

- Conventional sequencing methods struggle with detecting complex variants in key genes associated with inherited genetic disorders of high prevalence, necessitating complicated workflows and producing results that lack comprehensive detection of pathogenic variants.
- In response, we developed the AmplideX® Nanopore Carrier Plus Kit and accompanying analysis software that detects both simple and complex variants across *CFTR, SMN1/2, FMR1, HBA1/2, HBB, GBA, CYP21A2, TNXB*, and *F8* intron inversions with a streamlined end-to-end workflow.
- The Kit combines targeted PCR enrichment, nanopore sequencing, and companion analysis software to resolve complex pathogenic variants in target genes including short tandem repeats, copy number variation, structural variation, and differentiation from pseudogenes.
- The Kit's analytical performance was assessed using a mixed sample set consisting of 445 samples.

Introduction

Accurate detection of severe genetic disorder-associated variants is crucial for diagnostics and carrier screening. While Next-Generation Sequencing (NGS) is frequently employed, it often struggles to resolve complex genotypes such as copy number variation (CNV), short tandem repeats, structural variation (SV), and differentiation from highly homologous pseudogenes. As a result, many laboratories must maintain several specialized assays with unique workflows and equipment to identify complex pathogenic variations, leading to increased costs and operational complexity.

Our unified approach is comprised of a targeted PCR nanopore sequencing assay and accompanying analysis software that examines 11 key targets representing ~70% of at-risk couples in the US population. Here, we present verification testing performance from studies designed to characterize accuracy, precision, sensitivity, and specificity.

Materials and Methods

Target regions were enriched with 4 PCR mixes in a single assay workflow. Sequencing was performed on a MinION or GridION (Oxford Nanopore Technologies) with R10.4.1 flow cells. Nanopore reads were processed with software that provides fully automated quality control, variant reporting, genotype summary, and results visualization tailored to each unique variant type. Verification studies utilized over 445 unique samples, which included DNA isolated from whole blood with multiple methods (precipitation, column, and magnetic beads) and cell-line genomic DNA. Verification studies included method comparison, single-site precision, DNA input, in-use and freeze-thaw stability, analytical specificity, mix modularity, and thermal cycler equivalency. All data in this poster, except mix modularity, is from the method comparison and precision studies of verification testing. Results were compared to orthogonal data where methods were available, and samples were presumed wildtype when data was unavailable. For presumed wildtype genes, discrepancies were either resolved with orthogonal methods or excluded from analysis where independent evaluation was not possible.

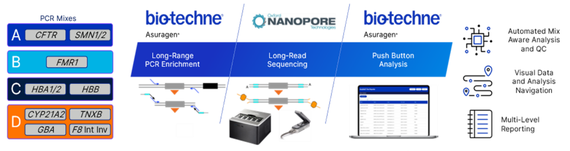


Figure 1. AmplideX Nanopore Carrier Plus Kit Panel Design and Workflow. Targeted PCR enriched panel-specific genetic regions in 11 genes responsible for common inherited genetic disorders. Kit workflow consisted of PCR enrichment, sequencing, and automated analysis software. Basecalled sequencing data was analyzed with Kit provided software to report multiple variant classes across genes.

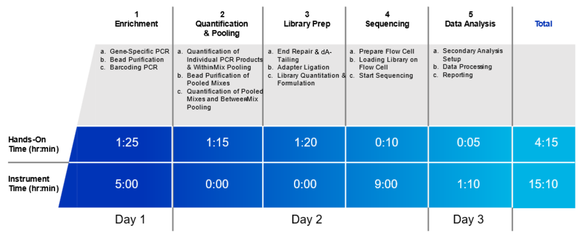
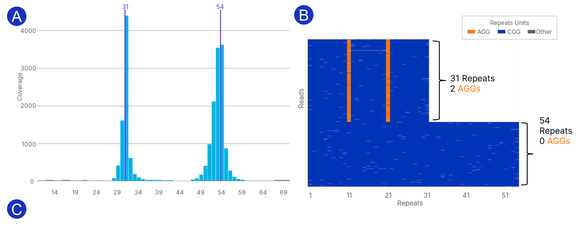


Figure 2. Assay Time of a Typical 24-Sample Batch with Kit A,B and C. Timing was divided across five phases that can be stretched over 3 work shifts with overnight sequencing and live basecalling, between the second and third day. AmplideX One Reporter secondary analysis software resolves genotypes.

Results

The assay successfully detected single nucleotide variants (SNVs), insertions and deletions (indels), CNVs, and SVs in 11 target genes with an overall percent agreement (OPA) of >95% in all studies conducted. The assay was reproducible and robust across operators, runs, DNA extraction methods, mix combinations, and thermal cycler models.

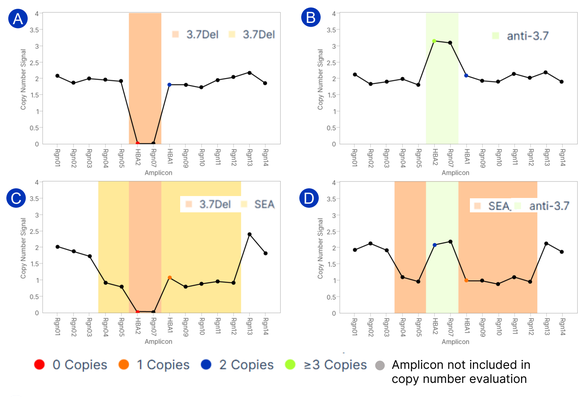
FMR1 STR



FMR1 Category	Expected			
Measured	NOR	INT	PRE	FM
NOR	25	0	0	0
INT	0	43	0	0
PRE	0	*1	85	0
FM	0	0	0	38

Figure 3. Signal processing visuals allow users to thoroughly investigate *FMR1* CGG repeats and AGG interrupts. (A) A histogram of read counts for each repeat size utilized to quantitate CGG repeats in each *FMR1* allele (purple lines), identified in this sample at 31 and 54. **(B)** Waterfall plot of the same sample displaying phased "AGG interrupts" (orange) enables risk assessment of *FMR1* repeat expansion into a full mutation in the next generation. **(C)** Sample level genotype agreement was 99.5% (191/192) across variant categorites: Normal (NOR), Intermediate (INT), Premutation (PRE) and Full Mutation (FM). *Intermediate sample identified as Premutation with CGG repeat size near the category boundary (54 expected, 55 observed) which is within ACMG precision guidelines (±5 CGG).

HBA1/2 CNV, SV, and SNV/Indels

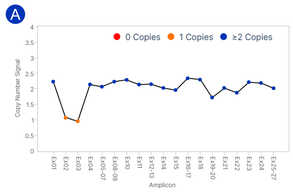


E

HBA1/2 Variants	Expected		
Measured	0	1 or 2	≥3
0	68	0	0
1 or 2	0	86	0
≥3	0	0	21

Figure 4. Plotting *HBA1/2* Amplicon Copy Number Signal Allows Users to Visualize Diverse Variant Genotypes. A set of 14 amplicons targeting the hemoglobin alpha cluster (including full gene coverage of *HBA1* and *HBA2*) to differentiate common SVs and phase SNV/indels. Plotting copy number signal for each amplicon accurately identifies *HBA1/2* deletions and duplications including a **(A)** homozygous -α3.7del (0 cp *HBA2*, 2 cp HBA1) **(B)** anti-α3.7 (3 cp *HBA2*, 2 cp *HBA1*), **(C)** compound heterozygous deletion -α3.7/SEA (0 cp *HBA2*, 1 cp *HBA1*), and **(D)** compound heterozygous anti-α3.7/SEA (2 cp *HBA2*, 1 cp *HBA1*). **(E)** Overall *HBA1/2* genotype agreement was 100% (175/175). Variants were either copy deletions or pathogenic variants (SNV/indels).

CFTR SNV/Indels, and Exon Deletions



B

# *CFTR* Variants	Expected		
Measured	0	1	≥2
0	165	0	0
1	1	64	0
≥2	0	1	45

Figure 5. Identification of Pathogenic SNVs, Indels, and Exon Deletions in *CFTR*. (A) Users can visualize exon deletions on a plot of copy number signal across *CFTR* amplicons. The sample shown was accurately called as a CFTRdele2,3 exon deletion. **(B)** Agreement for sample-level genotype was 99.3% (274/276). Both incorrect calls were incorrect whole gene deletions that were associated with a single extraction method, potentially indicating a sample quality issue.

SMN1/2 CNVs and SNV/Indels

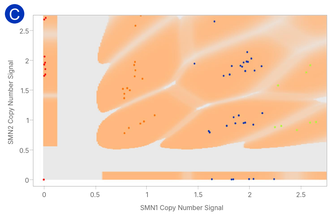
A

SMN1 Copies	Expected			
Measured	0	1	2	≥3
0	19	0	0	0
1	0	51	0	0
2	0	0	148	0
≥3	0	0	7	45

B

SMN2 Copies	Expected			
Measured	0	1	2	≥3
0	27	0	0	0
1	0	64	0	0
2	0	1	114	0
≥3	0	0	6	58

C

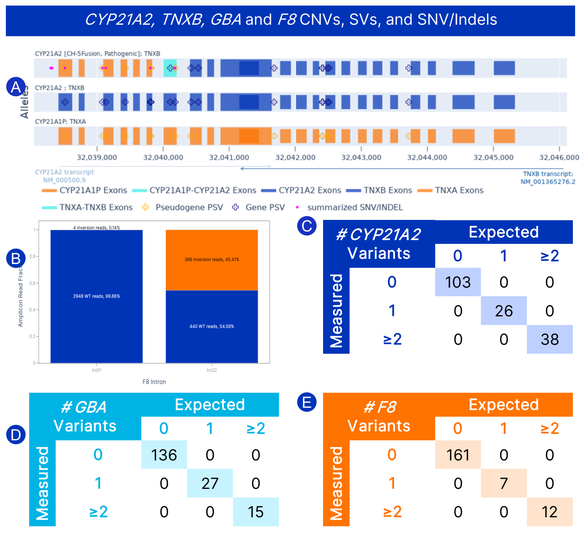


D

SMN1/2 Variants	N Samples	A (%)
SMN1 DV	522	100
SMN2 DM	261	100

Figure 6. Quantification of SMN1/2 Copy Number and SNVs. Copy number accuracy for **(A)** *SMN1* and **(B)** *SMN2* across 540 measurements from 270 samples. Overall agreement was 97.4% (526/540) for both *genes*. All genotypes associated with SMA (*SMN1* copies 0 and 1) were accurately identified. **(C)** Visualization of *SMN1* and *SMN2* copy number signal covariance for a 96 sample set identifies potential edge cases. **(D)** Accuracy of detection was 100% for *SMN1* duplication variants (DV) (c.*3+80T>G; c.*211_212del) and *SMN2* disease modifiers (DM) (c.859G>C).

CYP21A2, TNXB, GBA and *F8* CNVs, SVs, and SNV/Indels



C

# *CYP21A2* Variants	Expected		
Measured	0	1	≥2
0	103	0	0
1	0	26	0
≥2	0	0	38

D

# *GBA* Variants	Expected		
Measured	0	1	≥2
0	136	0	0
1	0	27	0
≥2	0	0	15

E

# *F8* Variants	Expected		
Measured	0	1	≥2
0	161	0	0
1	0	7	0
≥2	0	0	12

Figure 7. Automated Detection of *CYP21A2/ CYP21AP* and *GBA/ GBAP* Fusions, and *F8* Inversions. (A) *CYP21A2/CYP21AP* and *GBA* gene fusions were detected in samples via allele deconvolution and paralogous sequence variants (PSV). PSV sites corresponding to both gene and paralog entries (top) indicated presence of a fusion. **(B)** An *F8* multi-primer system covers the intron 22 and intron 1 potential inversion breakpoints allowing the identification of WT versus inversion reads and zygosity. **(C), (D), (E)** Overall *CYP21A2, GBA* and *F8* genotype agreement was 100%. Five samples with known TNXB structural variants or SNVs were 100% concordant (data not shown).

Kit Modularity

Table 1. Mix Modularity. A single sample panel (n=22 samples + 2 calibrators) was sequenced in single-mix runs and in various multi-mix combinations. Multi-mix agreement with single-mix run results was 100% except for single *HBB* copy number call in the ABCD (22 samples in duplicate) run.

Mix Combination	Agreement by Mix (%)			
	A	B	C	D
AB	100%	100%	N/A	N/A
AC	100%	N/A	100%	N/A
CD	N/A	N/A	100%	100%
ACD	100%	N/A	100%	100%
ABC	100%	100%	100%	N/A
ABCD (duplicate sample set)	100%	100%	99%	100%
ABCD (24 sample set)	100%	100%	100%	100%

Analytical Performance

Table 2. Analytical Performance Across Gene/Variant Classes. Variant classes for *CFTR, SMN1/2, FMR1, HBA1/2, HBB, CYP21A2, GBA*, and *F8* intron inversions were detected at >97% agreement with orthogonal data. Performance assessed based on samples and alleles with orthogonal data (N). Overall percent agreement (OPA) refers to overall accuracy. CCG accuracy for *FMR1* followed ACMG guidelines(±5 CGG). CNV calls for *SMN1/2* were evaluated with performance from 0 to ≥3 copies. SM=sample measurements.

Gene	Variant Class	N	Accuracy (OPA)
CFTR	PolyT/TG	548 alleles	100%
	SNV/Indel	31710 variants	100%
	Exon Deletions	276 SM	99.3%
SMN 1/2	*SMN1* CNV	270 SM	97.4%
	SMN2 CNV	270 SM	97.4%
	SMN1 Duplication Variants	1446 variants	100%
	SMN2 Disease Modifiers	393 variants	100%
FMR1	CGG	359 alleles (192 SM)	97.1%
	AGG	440 alleles	99.5%
	FMR1 Category	192 SM	*99.5%
HBA 1/2	CNV/Deletions	175 SM	100%
	SNV/Indel	43 variants	100%
HBB	Exon Deletions	175 SM	ǂ98.9%
	SNV/Indel	144 variants	ǂ99.3%
CYP21A2	*CYP21A2* Category	372 SM	¥98.1%
GBA	*GBA* Category	354 SM	100%
F8	Inversions (Intron 1 & 22)	180 SM	100%

a. All genes had samples in which orthogonal data was not available. In those cases, the sample was assumed WT for the gene(s) we lacked orthogonal data for.
*100% within precision limits (±5 CGG)
ǂThe number of pathogenic variants was correct, but not concordant with specific variants.
¥5 presumed wildtype samples were excluded from *CYP21A2* analysis due to positive variant calls because orthogonal data could not be obtained.

Conclusions

- The AmplideX Nanopore Carrier Plus Kit produces precise and accurate reporting for SNVs, indels, copy number gain or loss, short tandem repeat sizes, large structural variants, and gene-pseudogene fusions with ≥97% overall percent accuracy across 11 targeted genes.
- The Kit's 2.5-day streamlined workflow from DNA to results shortens turnaround times, consolidates workflows and makes characterization of complex genotypes associated with these key genes accessible to more labs.
- The AmplideX One Reporter software assists in deciphering complex genotypes, simplifying interpretation of nanopore data, with novel visualizations an intuitive user interface.
- The Kit concordance with orthogonal methods was: SNVs/INDELs (>99% of samples), *CFTR* exon deletions (99.3%),*SMN1/2* CN (≥97.4%), *FMR1* repeat categories and AGG interruptions (*99.5%), *HBB* deletions (98.9%),*HBA1/2* deletions (100%), *CYP21A2* genotype category (100%), *GBA* genotype category (100%), and *F8* intron inversions (100%).

References:

1. Woo, E. G., Tayebi, N., & Sidransky, E. (2021). Next-Generation Sequencing Analysis of GBA1: The Challenge of Detecting Complex Recombinant Alleles. Frontiers in genetics, 12, 684067.
2. Seaby, E. G., & Ennis, S. (2020). Challenges in the diagnosis and discovery of rare genetic disorders using contemporary sequencing technologies. Briefings in functional genomics, 19(4), 243–258.
3. Beauchamp, K.A., Muzzey, D., Wong, K.K., Hogan, G.J., Karimi, K., Candille, S.I., Mehta, N., Mar-Heyming, R., Kaseniit, K.E., Kang, H.P., et al. (2018). Systematic Design and Comparison of Expanded Carrier Screening Panels. Genet. Med. 20, 55–63.

All authors have a financial relationship to disclose: Employment by Asuragen. This product is under development. Future availability and performance to be determined.

**Research Use Only. Not for use in diagnostic procedures.*

Oxford Nanopore Technologies products are not intended for use for health assessment or to diagnose, treat, mitigate, cure, or prevent any disease or condition.

biotechne // **Global Developer, Manufacturer, and Supplier of High-Quality Reagents, Analytical Instruments, and Precision Diagnostics.**

INCLUDES R&D Systems™ Novus Biologicals™ Tocris Bioscience™ ProteinSimple™ ACD™ ExosomeDx™ Asuragen® Lunaphore™