Verification of an Amplification Based Nanopore Sequencing Assay and Software to Genotype Complex, Clinically-Relevant Variants in 11 Hard-to-Decipher Genes with High Carrier Frequencies

G094

Connor A. Parker, Jonathan Turner, Brennan Greenlee, Juliette A. Baker, Walairat L. Wolf, Christopher J. Fraher, Monica Roberts, Theodore Markulin, Eduardo Priego, Mahesh Yarasi, Mia K. Mihailovic, Bryan J. Killinger, Brian C. Haynes, Bradley Hall, and Ashima Sharma

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 often 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 (CNV), structural variation, and differentiation from pseudogenes.
- The Kit analytical performance was assessed using a mixed sample set consisting of 136 cell-line samples and 229 gDNA samples isolated from whole blood, which demonstrated >97% agreement across all variant categories and genes included in the assay.

Introduction

Detection of pathogenic variants associated with severe genetic disorders is critical for diagnostic and carrier screening applications. Although NGS is a widely used method for genetic characterization, it fails to resolve many variants in genes of clinical interest that have complex pathogenic elements like repeats, structural variation, and pseudogenes[1,2]. This traditionally necessitates use of multiple specialized workflows that only cover a fraction of pathogenic alleles[3].

To address this shortcoming for researchers, we combined short- and long-range PCR enrichment with nanopore sequencing in a kit-based modular panel to accurately genotype traditionally NGS addressable genes (*CFTR* and *HBB*) and hard-to-decipher genes (*SMN1/2*, *FMR1*, *HBA1/2*, *GBA*, *CYP21A2*, *TNXB*, and *F8* intron inversions), in a single streamlined workflow, representing ~70% of at-risk couple detection.[3] We developed bespoke software and algorithms to automate analysis of SNVs, indels, copy number gain or loss, gene-pseudogene fusions, and large structural variants. Here we describe analytical performance of the AmplideX Nanopore Carrier Plus Kit across multiple variant classes.

Methods

Target regions were PCR-enriched with the AmplideX Nanopore Carrier Plus Kit across 4 mixes in a single assay workflow on Applied Biosystems™ (ABI) Veriti, ABI Proflex, ABI VeritiPro, and Bio-Rad C1000 thermal cyclers. Sequencing was performed on a MinION or GridION (Oxford Nanopore Technologies) with R10.4.1 flow cells. Nanopore reads were basecalled with Dorado and then processed with AmplideX One Reporter software that provides fully automated quality control, result visualization, and variant reporting. Evaluated samples include genomic DNA isolated from whole blood (n = 229) with multiple methods (precipitation, column, and magnetic beads) and cell-line genomic DNA (n = 136). Assay/software system analytical performance was measured by comparing the Kit results to data collected by orthogonal methods.

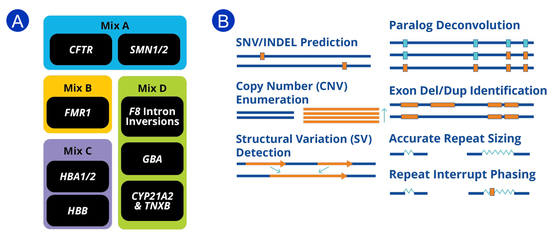


Figure 1. AmplideX Nanopore Carrier Plus Kit gene content and variant classes reported by automated analysis software. Panel design and workflow identifies pathogenic variants for 11 genes responsible for common inherited genetic disorders and/or hard-to-decipher genes. **A)** Four PCR mixes are used to amplify targets across 11 genes which represents ~75% of at-risk couples in Carrier Screening. **B)** Kit provided software identifies and reports multiple variant classes across genes.

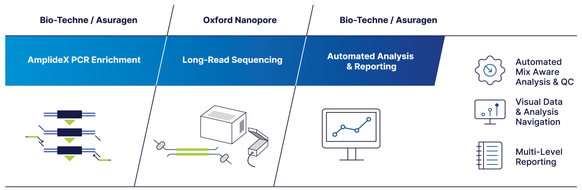
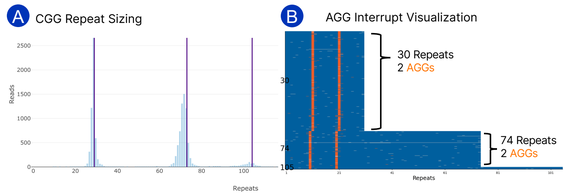


Figure 2. AmplideX Nanopore Carrier Plus Kit Workflow. Streamlined Kit workflow involves PCR enrichment and sample barcoding, library preparation followed by nanopore sequencing, automated sequencing analysis and variant reporting.

Results

FMR1 Repeats



FMR1 Category		Expected			
		Normal	Intermediate	Premutation	Full Mutation
Measured	Normal	155	0	0	0
	Intermediate	0	31	0	0
	Premutation	0	1*	36	0
	Full Mutation	1*	0	0	29

Figure 3. Signal processing visuals allow users to thoroughly investigate *FMR1* CGG repeats and AGG interrupts. (A) Histogram of sequenced repeat lengths, which is used by the algorithm to determine the number of CGG repeats in each FMR1 allele (purple line) and are identified in this sample at 30, 74 and 105 repeats. **(B)** Waterfall plot graphically displays phased "AGG Interrupts" (orange) within each repeat allele, two per allele in this sample, to allow for assessment of risk of *FMR1* repeat expansion. **(C)** *FMR1* sample-level genotype agreement was 99.2% (251/253). *Intermediate sample identified as Premutation has CGG repeat at category boundary that differed by +1 CGG repeat from expectation (55 vs 54 repeats) which is within size tolerance according to ACMG guidelines (±1 CGG). False positive Full Mutation call due to Full Mutation sample contamination in Normal sample via well-to-well contamination during run.

SMN1/2 Copy Number

SMN1 Copies		Expected			
		0	1	2	≥3
Measured	0	45	0	0	0
	1	0	104	2	0
	2	0	0	426	7
	≥3	0	0	13	190

SMN2 Copies		Expected			
		0	1	2	≥3
Measured	0	121	0	0	0
	1	0	218	1	0
	2	0	1	330	0
	≥3	0	0	15	101

Figure 4. Accurate quantification of SMN1/2 copy number. Calling accuracy for **(A)** *SMN1* and **(B)** *SMN2* copy numbers across 787 measurements from 242 unique samples (116 cell line, 126 gDNA from whole blood). Overall percent agreement was 97.2% (765/787) for *SMN1*, and 97.8% (770/787) for *SMN2*. All samples with genotypes associated with SMA (*SMN1* Exon 7 Copies = 0) and SMA carriers (*SMN1* Exon 7 Copies = 1) were accurately identified with two false positives for SMA carrier genotype. *SMN1/2* copy number quantification was accurate for samples with hybrid *SMN1/2* genes. For SNVs and Indels associated with *SMN1* gene duplication and *SMN2* disease modifiers, agreement was 99.9% (787 sample measurements, 2361 variants).

CFTR SNVs, INDELS, and Large Exon Deletions

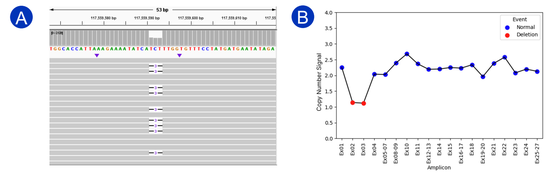


Figure 5. Identification of pathogenic SNVs, indels, and exon deletions in *CFTR*. (A) Overall percent agreement for SNV/Indel calls in *CFTR* was 99.9% (838 sample measurements, >90,000 variants). Automated variant calls can be reviewed via bam files of the aligned sequencing data at the relevant genomic location. The *CFTR* c.1521_1523del (F508del) variant in sample shown here has a clear signal of a heterozygous 3-base pair deletion in the sequence pileup. **(B)** Overall percent agreement for large exon deletion calls was 99.4% (839 sample measurements). Presence of large exon deletions can be visualized via plot of copy number signal across *CFTR* amplicons. CFTRdele2,3 large exon deletion accurately called and visualized in sample shown.

HBA1/2 Copy Number

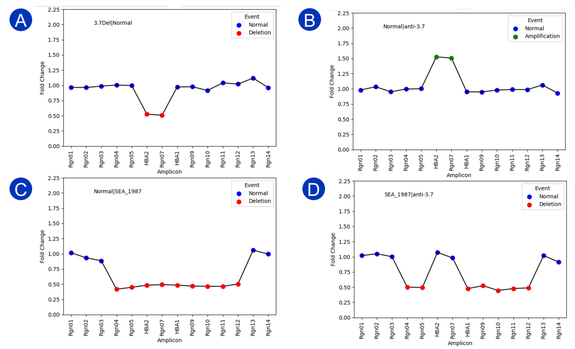


Figure 6. *HBA1/2* amplicon fold-changes provide signal for a diverse set of variant genotypes. A set of 14 "sentinel" amplicon target regions of the hemoglobin alpha cluster differentiate known common breakpoints. Automated analysis of fold-change signal accurately identifies *HBA1/2* deletions and amplifications including **(A)** deletion -α3.7, **(B)** amplification anti-α3.7, **(C)** deletion SEA, and **(D)** compound heterozygote: SEA/anti-α3.7. Overall percent agreement for *HBA1/2* deletion and amplification variants across 740 sample measurements was 99.5% (736/740). Three incorrect calls due to misclassification of the specific deletion but copies of *HBA1* and *HBA2* were correct. Full coverage of *HBA1* and *HBA2* genes also allows for identification of pathogenic SNV/Indels. Overall percent agreement for SNV/Indel calls was 100% (42 sample measurements, 83 variants).

CYP21A2 Structural Variation

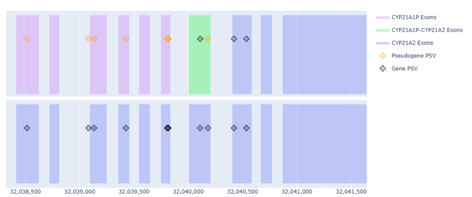


Figure 7. Automated detection of *CYP21A2/CYP21A1P* gene fusion in sample via allele deconvolution and paralogous sequence variants (PSV). PSV sites corresponding to both gene and paralog entries in *CYP21A2/CYP21A1P* allele 1 (top) indicate presence of fusion while PSV sites indicate allele 2 (bottom) is not a fusion. Overall percent agreement for detection of gene pseudogene recombination/fusions was 100% for *CYP21A2* (567 sample measurements, 1134 alleles) and overall percent agreement for detection of pathogenic SNV/Indels in *CYP21A2* was 100% (70 sample measurements, 1340 alleles).

Analytical Performance

Table 1. Analytical performance metrics across gene/variant classes. All 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 is evaluated based on samples and alleles with orthogonal data (N). Overall percent agreement (OPA) refers to accuracy of the analysis. CNV calls for *SMN1/2* were evaluated with performance from 0 to ≥3 copies. SM=sample measurements

Gene	Variant Class	N	Accuracy
CFTR	PolyT/TG	1534 alleles (838 SM)	99.9%
	SNV/Indel	90928 variants (838 SM)	99.9%
	Large Exon Deletions	839 SM	99.4%
SMN1/2	SMN1 CNV	787 SM	97.2%
	SMN2 CNV	787 SM	97.8%
	SNV/Indel	2361 variants (787 SM)	100%
FMR1	CGG	382 alleles (253 SM)	99.2%
	AGG	134 alleles (72 SM)	99.3%
	FMR1 Category	253 SM	99.2%
HBA1/2	CNV/Deletions	740 SM	99.5%
	SNV/Indel	83 variants (42 SM)	100%
HBB	CNV	751 SM	99.7%
	SNV/Indel	340 variants (141 SM)	99.4%
CYP21A2	CNV	566 SM	99.5%[a]
	Fusions	1134 alleles (567 SM)	100%
	SNV/Indel	1340 variants (70 SM)	100%
GBA	CNV	563 SM	99.5%[a]
	Fusions	1134 alleles (567 SM)	99.8%
	SNV/Indel	1428 variants (73 SM)	100%
F8	Inversions (Intron 1 & 22)	1131 SM	100%

a. Due to lack of orthogonal methods for CNV, overall percent agreement with mode of copy number call for sample in assay is reported.

Conclusions

- The AmplideX Nanopore Carrier Plus Kit detects SNVs, indels, copy number gain or loss, short tandem repeat sizes, large structural variants, and gene-pseudogene fusions with ≥97% overall percent agreement to orthogonal methods across *CFTR*, *SMN1/2*, *FMR1*, *HBA1/2*, *HBB*, *GBA*, *CYP21A2*, and *F8* intron inversions.
- The Kit utilizes a single-platform, incorporates a streamlined workflow with the potential to greatly reduce assay complexity, and shortens turnaround times for characterization of complex genotypes associated with important and prevalent inherited disorders and at-risk couple status.
- Accompanying AmplideX One Reporter software allows users to easily process nanopore sequencing data, view variant calls, and perform in-depth investigation of sequencing data and analysis results with a simple user interface.
- The Kit agreed with orthogonal methods for SNVs/INDELs across genes (>99.9%), *SMN1* CN (97.2%), *SMN2* CN (97.8%), *FMR1* repeat categories (99.2%), *FMR1* AGG interruptions (99.3%), *HBA1/2* deletions (99.5%), *CYP21A2* fusions and structural variants (100%), *GBA* fusions and structural variants (99.8%), and *F8* inversions (100%).

Scan code to learn more

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., Kasenilt, 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.