Verification of an Integrated Amplification-Based Nanopore Sequencing Assay and Automated Software Solution to Genotype Challenging, Clinically Relevant Variants

Poster Number: P23.015.A

Connor Parker, Juliette A. Baker, Walairat L. Wolf, Jon Kemppainen, Ashima Sharma, Brennan Greenlee, Justin Janovsky, Jonathan Turner, Bryan J. Killinger, Cody Edwards, Monica Roberts, Theodore Markulin, Eduardo Priego, Mahesh Yarasi, Mia K. Mihailovic, Brian C. Haynes, Bradley Hall, John N. Milligan

Asuragen, Inc. 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.
- To address this need, we developed a novel PCR enrichment/Nanopore sequencing assay 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.
- We verified performance with a mixed sample set consisting of 151 cell-line samples and 294 gDNA samples isolated from whole blood, which demonstrated >97% agreement across all variant categories, including short tandem repeats, copy number variation, structural variation, and differentiation from pseudogenes.

Introduction

Accurate resolution of variants correlated with severe genetic disorders is crucial for diagnostic and screening applications. While NGS is used broadly, it is frequently incapable of resolving complex pathogenic variants involving short tandem repeats, copy number/structural variations (CNV/SV), and pseudogenes[1,2]. This traditionally necessitates use of multiple specialized workflows and equipment that only cover a fraction of pathogenic alleles, leading to increased cost and complexity[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 all pathogenic variants associated with inheritable diseases that impact neonates[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 present verification testing performance from studies designed to characterize accuracy, precision, sensitivity, and specificity.

Methods

Design verification testing was executed using 445 unique samples, including cell-line genomic DNA (gDNA) samples (n=151) and whole blood gDNA from residual clinical samples (n=294) isolated using multiple methods (precipitation, column, and magnetic beads). The selection of cell-line and whole blood samples represent all major variant classes. Target regions were enriched in 4 PCR reactions, barcoded, pooled, and sequenced on MinION flow cells (R10.4.1) with a Mk1B or GridION (Oxford Nanopore Technologies, ONT). 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 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 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.

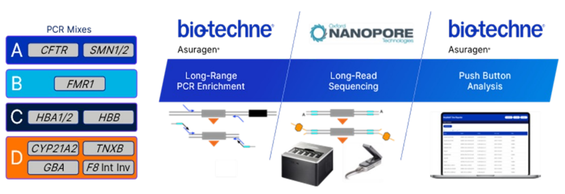


Figure 1. AmplideX® Nanopore Carrier Plus Kit Panel Design and Workflow. Targeted PCR enriches panel-specific genetic regions in 11 genes responsible for common inherited genetic disorders. The amplicons (~300bp – 10kb in length) are PCR barcoded by sample, pooled, and sequenced on an ONT (Oxford Nanopore Technologies) device. Raw sequencing data is analyzed with automated software to produce quality control metrics, annotated variant calls, and results visualization.

1 Enrichment	2 Quantification & Pooling	3 Library Prep	4 Sequencing	5 Data Analysis	Total
1:25 Hands-On	1:15	1:20	0:10	0:05	4:15
5:00 Instrument	0:00	0:00	9:00	1:10	15:10
Day 1	Day 2			Day 3	

Figure 2. Assay Time of a Typical 24-Sample Batch. Time displayed as hrs:min.Timing study included 24 samples run with Mixes A, B, and C, and 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

AmplideX® One Reporter

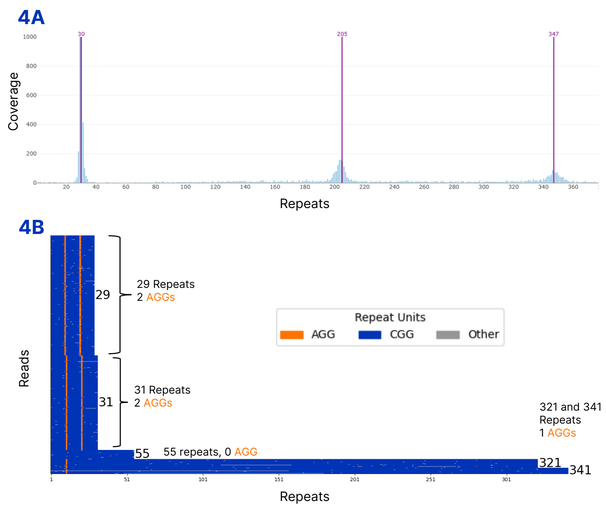
Sample Summary / CP_DVT_115_JJ_Mk1Dbridge_v102 (CP_DVT_115_JJ_Mk1Dbridge_D_20250301_1300)

COLUMNS FILTERS DENSITY EXPORT

Sample_Name	QC	Genes[Status]	Barcode	Mixes	Calibrator
NA18668	PASS	CFTR [V, SV]	BC01	ABCD	N
RS04832	PASS	No Variants	BC02	ABCD	N
HG02461	PASS	CFTR [SV] CYP21A2 [3 cp, SV] TNXB [3 cp, SV]	BC13	ABCD	N
NA09237	PASS	FMR1 [FM]	BC14	ABCD	N
NA12217	PASS	CYP21A2 [3 cp, 2SV] TNXB [3 cp, SV]	BC15	ABCD	N
RU010	PASS	FMR1 [FM] CYP21A2 [2 cp, >2V(0|1)]	BC26	ABCD	N
NA14734	PASS	CYP21A2 [1 cp, SV(H)] TNXB [1 cp]	BC27	ABCD	N
RS04122	PASS	CYP21A2 [3 cp, SV] GBA1 [2 cp, V] TNXB [3 cp, SV]	BC37	ABCD	N
CAL	PASS	FMR1 [PM]	BC38	ABCD	AC

Figure 3. Customizable high-level analysis results allow users to quickly identify desired variants. A simple interface allows users to filter samples and quickly identify key variants and QC status. Variants are summarized for efficient interpretation. Detailed genotype view for each sample and variant view for all variants are also available (not shown). V = variant; SV = structural variant, cp = copies; FM = full mutation; H = homozygous; PM = premutation.

FMR1 CGG/AGG



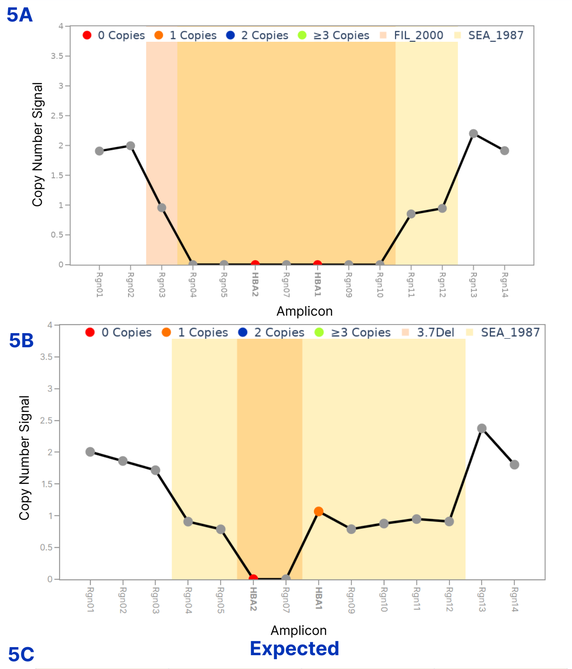
4C

Expected

Measured: *FMR1*	Normal	Inter-mediate	Pre-mutation	Full Mutation
Normal	25	0	0	0
Intermediate	0	43	0	0
Premutation	0	1*	85	0
Full Mutation	0	0	0	38

Figure 4. Signal processing visuals allow users to thoroughly investigate *FMR1* CGG repeats and AGG interrupts. (A) The "CGG Repeats" panel displays a histogram of read counts for each repeat size used by the algorithm to determine the number of CGG repeats in each *FMR1* allele (purple line). This sample demonstrates repeat mosaicism in an affected individual, with multiple Full Mutation alleles (205 and 347 CGG) present at <10% of the total reads compared to the normal allele (30 CGG). **(B)** Phased "AGG Interrupts" (orange) allow for investigation of interrupt patterns and expansion risk within each allele. Two AGG interrupts were detected in the normal alleles, with none in the premutation allele, and one in the full mutation alleles. **(C)** *FMR1* sample-level genotype agreement was 99.5% (191/192) for exact sizing. One discrepancy (*) occurred near the repeat boundary that differed by 1 CGG repeat, and all were within size tolerance according to EMQN guidelines (+/- 5% of repeat size). Accounting for these tolerances, agreement was 100%. All samples passed QC.

HBA1/2 Copy Number



5C

Expected

Measured: #*HBA1/2* Variants	0	1 or 2	≥3
0	68	0	0
1 or 2	0	86	0
≥3	0	0	21

Figure 5. *HBA1/2* amplicon fold-changes provide signal for a diverse set of 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)** compound heterozygous deletion SEA/FIL (0 cp *HBA2*, 0 cp *HBA1*) consistent with Hb Bart's alpha thalessemia, and **(B)** a compound heterozygous deletion -α3.7/SEA (0 cp *HBA2*, 1 cp *HBA1*) genotype consistent with Hb H alpha thalessemia. Amplicons with grey are used to determine structural variants, but copy number is not separately reported **(C)** Overall *HBA1/2* genotype agreement was 100% (175/175). Variants were either copy deletions or pathogenic variants (SNV/indels); see Table 1 for performance by variant type.

SMN1/2 Copy Number

6A

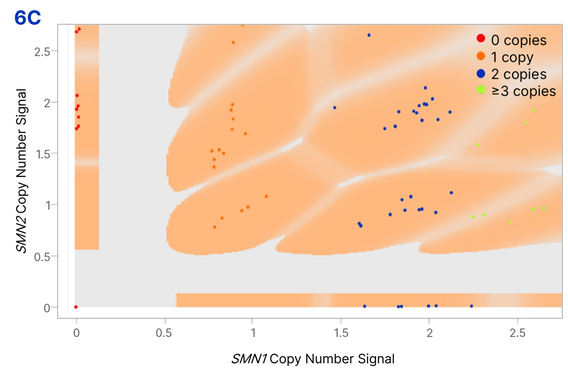
Expected

Measured: *SMN1* Cp#	0	1	2	≥3
0	19	0	0	0
1	0	51	0	0
2	0	0	148	0
≥3	0	0	7	45

6B

Expected

Measured: *SMN2* Cp#	0	1	2	≥3
0	27	0	0	0
1	0	64	0	0
2	0	1	114	0
≥3	0	0	6	58



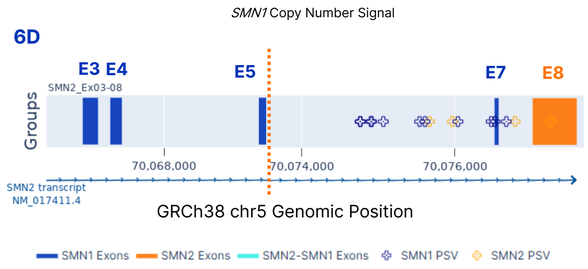
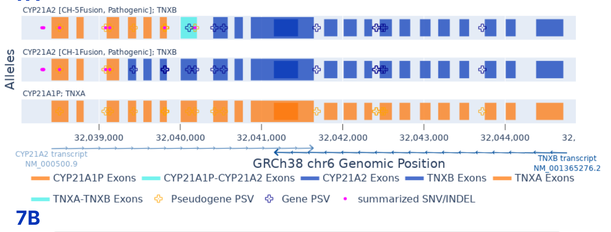
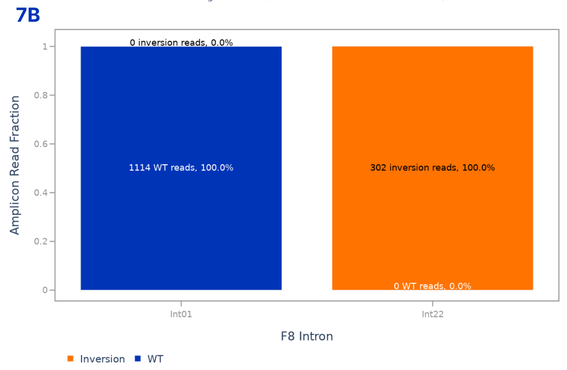


Figure 6. PCR/Nanopore assay accurately quantifies *SMN1/2* copy number. Calling accuracy for **(A)** *SMN1* and **(B)** *SMN2* copy numbers across 270 measurements. Overall percent agreement was 97.4% (263/270) for both *SMN1* and *SMN2*. For genotypes associated with SMA (*SMN1* E7 Cp = 0) and SMA carriers (*SMN1* E7 Cp = 1), agreement was 100%. **(C)** Visualization of *SMN1* and *SMN2* copy number signal covariance for all samples in the run identifies potential edge cases. **(D)** Visualization allows phasing of SNVs and paralogous sequence variants (PSV), identifying hybrid genes associated with improved prognosis for SMA patients in some studies. Orange dashed line indicates manually juxtaposed images for a single sample to remove middle exons for simplified viewing.

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





7C

Expected

Measured: #*CYP21A2* Variants	0	1	2
0	103	0	0
1	0	26	0
2	0	0	38

7D

Expected

Measured: #*F8* Variants	0	1	2
0	161	0	0
1	0	7	0
2	0	0	12

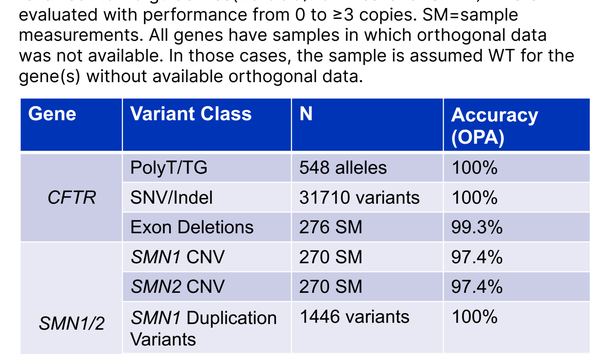
7E

Expected

Measured: #*GBA* Variants	0	1	2
0	136	0	0
1	0	27	0
2	0	0	15

Figure 7. Automated Detection of *CYP21A2/ CYP21AP* and *GBA/ GBAP* Fusions, and *F8* Inversions. (A) *CYP21A2/CYP21AP* fusions were detected in samples via allele deconvolution using PSVs. PSV sites corresponding to both gene and paralog entries indicates multiple fusions in this genotype associated with Congenital Adrenal Hyperplasia. **(B)** An *F8* multi-primer system covers the intron 22 and intron 1 potential inversion breakpoints allowing the identification of homozygous intron 22 inversion in this Hemophilia-associated genotype. **(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).

CFTR SNVs, INDELs, and Large Exon Deletions



8B

Expected

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

Figure 8. PCR/Nanopore assay identifies 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).

Analytical Performance Summary

Table 1. Analytical performance metrics 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. CGG 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. All genes have samples in which orthogonal data was not available. In those cases, the sample is assumed WT for the gene(s) without available orthogonal data.

Gene	Variant Class	N	Accuracy (OPA)
CFTR	PolyT/TG	548 alleles	100%
	SNV/Indel	31710 variants	100%
	Exon Deletions	276 SM	99.3%
SMN1/2	*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% [a]
HBA1/2	CNV/Deletions	175 SM	100%
	SNV/Indel	43 variants	100%
HBB	Exon Deletions	175 SM	98.9% [b]
	SNV/Indel	144 variants	99.3% [b]
CYP21A2	*CYP21A2* Category	372 SM	98.1% [c]
GBA	*GBA* Category	354 SM	100%
F8	Inversions (Intron 1 & 22)	180 SM	100%

a 100% within precision limits (±5%)
b The number of pathogenic variants was correct, but not concordant with specific variants.
c 5 presumed wildtype samples were excluded from *CYP21A2* analysis due to positive variant calls.

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 associated with prevalent genetic disorders, simplifying interpretation of nanopore data, with novel visualizations in 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%).

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.