

1 Introduction

Cystic fibrosis (CF) is one of the most common multisystem disorders in Caucasian population with a prevalence of 1 in 2500 to 3300 live births. CF is an autosomal recessive inherited disorder caused by pathogenic variants in the *CFTR* gene. Traditional newborn screening testing includes targeted approaches to detect limited number of pathogenic variants (25-139) after IRT (Immunoreactive Trypsinogen). The vast majority of states use a single IRT test. Only about 9 to 13 states utilize a dual IRT (or IRT/IRT/DNA) screening protocol. However, due to technical advances 1500 sequence variants, 200 copy number variants have been identified in the *CFTR* gene indicating the urgent need for a single comprehensive *CFTR* assay to cover the entire variant spectrum of the *CFTR* gene. We have developed a comprehensive next-generation sequencing (NGS) based assay to sequence the entire *CFTR* gene which utilizes an end-to-end workflow using 1-2 DBS punches for DNA extraction, library preparation, sequencing, data analysis and reporting to detect all sequence variants and copy number variants.

2 Methods

A custom *CFTR* targeted sequence capture is used to enrich the entire 188.7 kb genomic sequence of *CFTR* gene, which includes coverage for all 27 exons and introns. Fragmented DNA libraries were enriched via multiplexed hybridization capture and sequenced on the Element Biosciences AVITI platform using 2X150bp paired-end reads. Sequence variants (SNVs) were assessed by our proprietary analysis and interpretation pipeline, Ordered Data Interpretation Network (ODIN). CNV analysis was completed using Bionano NxClinical™ software.

A total of fifty (50) unique clinical samples and seven (7) GIAB samples were included in this comprehensive assay validation. In total, five (5) sequencing runs were performed to accommodate all of the validation samples. Samples pooled in groups of 16-multiplex format were sequenced in run 1, 2 and 3. Samples pooled in groups of 8 and 4-multiplex format were sequenced in run 4 and 5. A total of forty (40) clinical samples were used to build the reference for CNV calling by NxClinical software. Samples from different cell and tissue types, including whole blood (WB), saliva, dried blood spots (DBS), genomic DNA (gDNA) and prenatal samples were included in this comprehensive validation.

3 Results

To determine general technical performance, all samples from 5 sequencing runs, including intra-run and inter-run samples, were assessed for sequencing run metrics. These data established sample QC cutoff thresholds that will be utilized for clinical sample QC (Table 1). Sequencing metrics are comparable both between runs and among the sample types included in this validation. These results indicate that all sample types included in this validation are suitable for this assay.

To assess analytical performance, results from seven GIAB reference samples were compared to expected published data to assess whether SNVs and INDELs were detected at their expected frequencies. Analytical sensitivity is above 95% for both SNPs and INDELs in all seven samples. Analytical specificity is above 99% for both SNVs and INDELs. Twenty-three clinical samples with reportable variants were assessed for clinical accuracy. Expected variants were determined previously by testing at Revvity Omics using validated assays. A total of 25 reportable variants are expected initially. All observed variants are concordant with the expected results (Table 2 and Figure 1).

Table 1. Summary of sequencing metrics observed in the assay validation

Run number	Average of Yield (Mbases)	Average of % Perfect Barcode	Average of % ≥ Q30 Bases	Average Q Score	Average of Avg. Align Coverage	Average of % Mapped Reads	Average of % target > 20X	Average of Mean Insert Length	Count of Samples
1	1327	96.94	94.47	42.03	335	97.93	99.78	231	31
2	1190	98.45	95.67	42.68	311	97.59	99.77	235	30
3	1117	98.32	94.82	42.24	268	97.91	99.66	230	12
4	1279	97.91	95.59	42.55	359	97.71	99.82	232	8
5	1501	98.29	95.91	42.69	428	97.59	99.85	236	8
Average	1283	97.98	95.29	42.44	340	97.74	99.78	233	89

Table 2. Summary of clinical accuracy results of sequence and copy number variants

Sample number	Gene	Expected results	Expected zygosity	Observed results	Observed zygosity	Concordance
1975578	CFTR	CFTR Deletion of exons 16-20	heterozygous	CFTR Deletion of exons 16-20	heterozygous	Yes
1976631	CFTR	CFTR Deletion of exons 2-3	heterozygous	CFTR Deletion of exons 2-3	heterozygous	Yes
1976633	CFTR	CFTR Deletion of exon 2	heterozygous	CFTR Deletion of exon 2	heterozygous	Yes
1976633	CFTR	c.2052del (p.Lys684fs)	heterozygous	c.2052delA (p.Lys684fs)	heterozygous	Yes
1976630	CFTR	c.1585-2A>T	heterozygous	c.1585-2A>T	heterozygous	Yes
1976630	CFTR	c.940G>C (p.Gly314Arg)	heterozygous	c.940G>C (p.Gly314Arg)	heterozygous	Yes
1975584	CFTR	c.1000C>T (p.Arg334Trp)	heterozygous	c.1000C>T (p.Arg334Trp)	heterozygous	Yes
1976632	CFTR	c.1029del	heterozygous	c.1029delC (p.Cys343fs)	heterozygous	Yes
1975587	CFTR	c.1116+1G>A	heterozygous	c.1116+1G>A	heterozygous	Yes
1975571	CFTR	c.1521_1523del (p.Phe508del)	heterozygous	c.1521_1523del (p.Phe508del)	heterozygous	Yes
1977554	CFTR	c.1624G>T (p.Gly542Ter)	heterozygous	c.1624G>T (p.Gly542*)	heterozygous	Yes
1975579	CFTR	c.1705T>G (p.Tyr569Asp)	heterozygous	c.1705T>G (p.Tyr569Asp)	heterozygous	Yes
1975581	CFTR	c.1865G>A (p.Gly622Asp)	heterozygous	c.1865G>A (p.Gly622Asp)	heterozygous	Yes
1977553	CFTR	c.254G>A (p.Gly85Glu)	homozygous	c.254G>A (p.Gly85Glu)	homozygous	Yes
1975586	CFTR	c.2657+5G>A	heterozygous	c.2657+5G>A	heterozygous	Yes
1975574	CFTR	c.3038C>T (p.Pro1013Leu)	heterozygous	c.3038C>T (p.Pro1013Leu)	heterozygous	Yes
1975580	CFTR	c.3132del (9p.Glu1044fs)	heterozygous	c.3132del (9p.Glu1044fs)	heterozygous	Yes
1975585	CFTR	c.3154T>G (p.Phe1052Val)	heterozygous	c.3154T>G (p.Phe1052Val)	heterozygous	Yes
1976629	CFTR	c.3409A>G (p.Met1137Val)	heterozygous	c.3409A>G (p.Met1137Val)	heterozygous	Yes
1975566	CFTR	c.3909C>G (p.Asn1303Lys)	heterozygous	c.3909C>G (p.Asn1303Lys)	heterozygous	Yes
1976608	CFTR	c.*133delT	heterozygous	c.*133delT	heterozygous	Yes
1976628	CFTR	c.1408G>A (p.Val470Met)	heterozygous	c.1408G>A (p.Val470Met)	heterozygous	Yes
1976634	CFTR	c.4389G>A	heterozygous	c.4389G>A (p.Gln1463Gln)	heterozygous	Yes
1976624	CFTR	c.-152G>C	heterozygous	c.-152G>C	heterozygous	Yes
1976606	CFTR	c.164+12T>C	heterozygous	c.164+12T>C	heterozygous	Yes

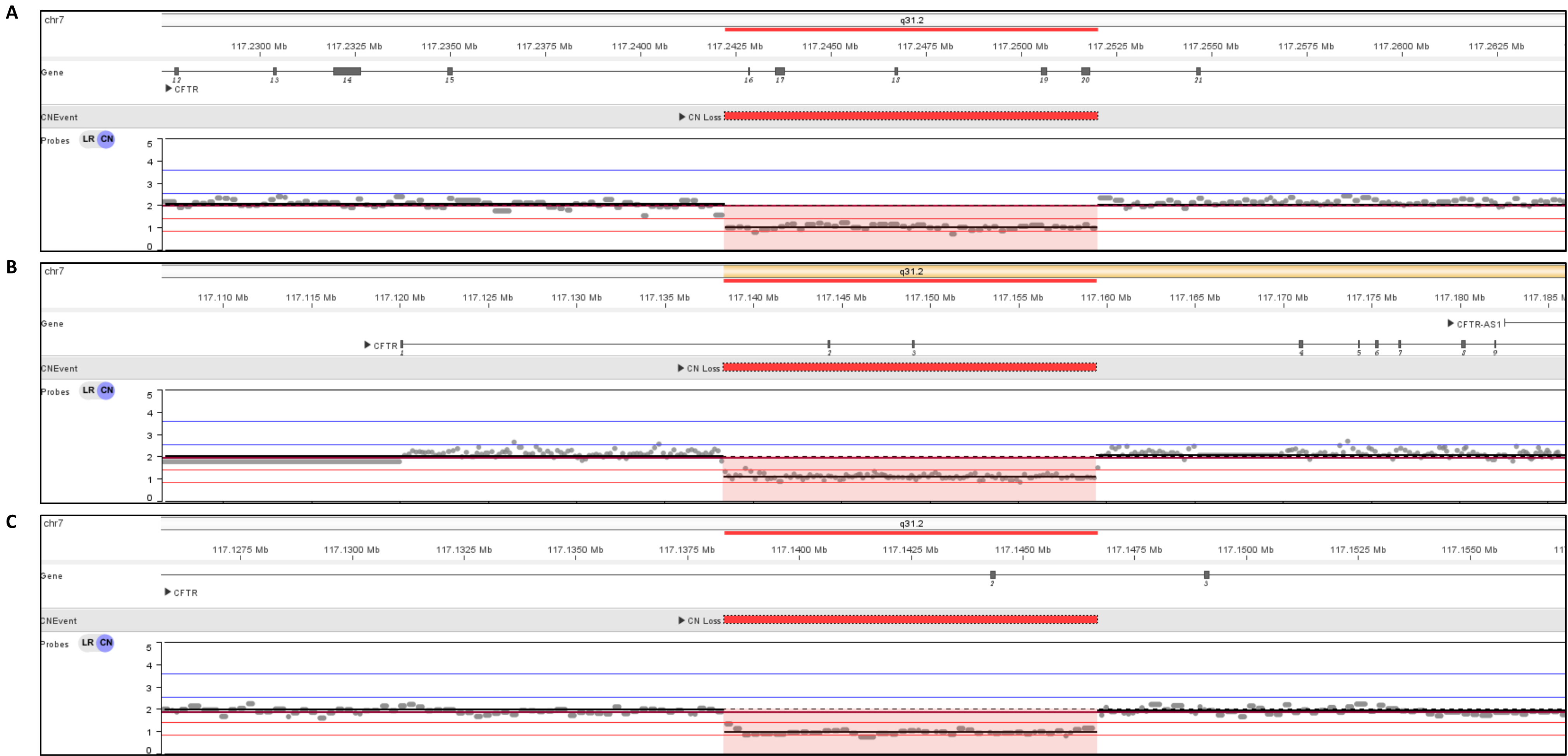


Figure 1. Different types of copy number variants identified with high confidence in clinical accuracy samples in this study. Panel A contains heterozygous deletion of *CFTR* exons 16-20. Panel B contains heterozygous deletion of *CFTR* exons 2 and 3. Panel C contains heterozygous single exon 2 deletion in the *CFTR*.

4 Conclusions

This comprehensive NGS-based assay interrogates the entire *CFTR* gene including the coding 27 exons as well as noncoding introns and promoter regions, leading to single step detection of all known and novel single-nucleotide variants (including deep intronic variants), deletions, duplications in one assay that are generally not detected in the current confirmatory assays due to targeted assay design and eliminates the need for reflex testing. Ultimately, it saves cost and time, leading to an early definitive diagnosis and intervention. Sequencing of the entire genomic sequence of the *CFTR* gene also facilitates a uniform coverage across the gene, which ultimately helps detection of CNVs with accurate breakpoints when compared with other traditional methodologies like array CGH. This comprehensive assay is highly sensitive for diagnostic testing for *CFTR* and is also suitable for confirmatory testing for newborn screening for *CFTR*.