High-resolution Analysis of Pathogenic Trinucleotide and Hexanucleotide Repeats, Copy Number Changes, SNVs and INDELs Using Flexible, Easy-to-use Fragment Sizing Instrumentation

Adrian Lara, Connor Parker, Steven Partin, John Hedges, John Milligan

Asuragen, a Bio-Techne Brand, Austin, TX

Summary

- AmplideX® PCR technology allows highly-multiplexed amplification and detection of difficult-to-resolve pathogenic variants, including triplet repeat alleles, hexanucleotide repeat alleles, STRs, SNVs, INDELs, and CNVs.
- We demonstrate 98.5% agreement for pathogenic variants in *FMR1* and 100% agreement for *CFTR, SMN1/2* Plus, and *C9orf72* variants between results generated with each respective AmplideX assay on the 3500 Dx CE instrument and the SeqStudio Flex CE instrument with off-scale recovery turned off.
- These results expand the use of AmplideX PCR/CE genetic assays targeting multiple challenging genes and variant classes to the SeqStudio Flex instrument, which offers simplified maintenance and operation to reduce CE complexity

Introduction

Capillary electrophoresis (CE) instruments provide a reliable and robust platform for sequencing and fragment analysis and are already used in a myriad of genetic applications. Advances in the design of these well-proven instruments continue to make them more user friendly for diverse laboratory settings. AmplideX® PCR/CE assays for difficult gene targets such as *C9orf72*, *FMR1*, *SMN1/2*, and *CFTR*, which provide a variety of challenges such as structural variation and GC-rich repeats, are compatible with other established CE platforms. However, enhancements available on the SeqStudio Flex may simplify CE analysis and maintenance, enabling broader access to this technology for those interested in these difficult gene targets.

We demonstrate here the detection of a broad range of challenging-to-resolve pathogenic variants across several widely-used PCR/CE assays using the SeqStudio Flex CE instrument. This includes detection of normal, intermediate, pre-, and full mutation *FMR1* triplet repeat alleles, normal, intermediate, and expanded *C9orf72* hexanucleotide repeat alleles, 65 unique *CFTR* variants (STRs, SNPs, and INDELs), and 0 to ≥4 copies of *SMN1* or *SMN2* as well as 3 unique variants in *SMN1/2* related to carrier risk or disease prognosis.

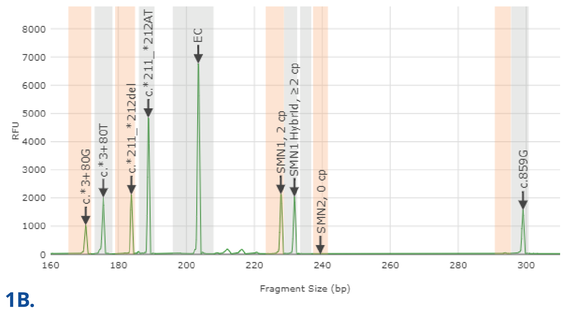
Methods

DNA samples were PCR-amplified using AmplideX assays (*C9orf72*†, *FMR1*†, *SMN1/2* Plus†, and *CFTR*†) according to assay instructions, followed by CE on both the Applied Biosystems™ 3500 and SeqStudio Flex (Flex) Genetic Analyzers for direct comparison. Injection conditions for the Flex instrument used the recommended injection conditions for the 3500 for each corresponding assay. In addition, we recommend turning off the off-scale recovery option (OSR), though data shown here includes both OSR-on and OSR-off data, with minimal difference observed.

Samples were chosen to cover all genotype categories, including normal, intermediate, pre-, and full mutation (*FMR1*), normal, intermediate, and expanded (*C9orf72*), 65 variants (*CFTR*), and 0 to ≥4 copies of *SMN1* and *SMN2*, as well as 3 unique variants in *SMN1/2* related to carrier risk or disease prognosis. Samples were derived from commercially available cell lines, residual human blood samples, or plasmids. FSA files were processed and analyzed using AmplideX Reporter Software for *FMR1*, *SMN1/2*, and *CFTR* (an unreleased version). GeneMapper 6.1 and a sizing macro was used to process and analyze samples for *C9orf72*. QC failures were excluded.

Results

1A.



1B.

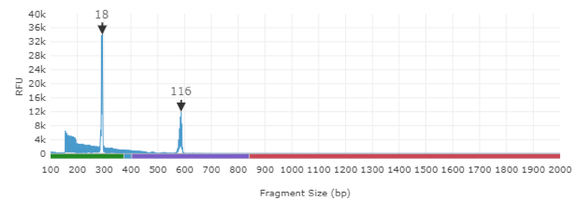
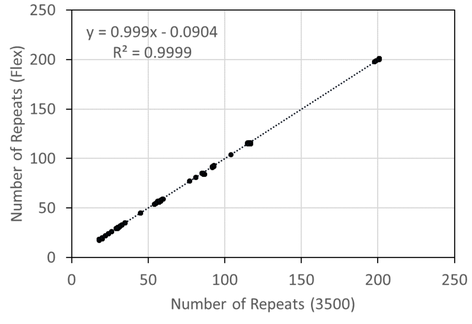


Figure 1. Example SeqStudio Flex Electropherograms. A) *SMN1/2* Plus Electropherogram (*SMN1* hybrid with gene duplication variants). EC = Endogenous Control. B) *FMR1* Electropherogram (premutation female). Images exported from the AmplideX Reporter Software.

2A.



2B.

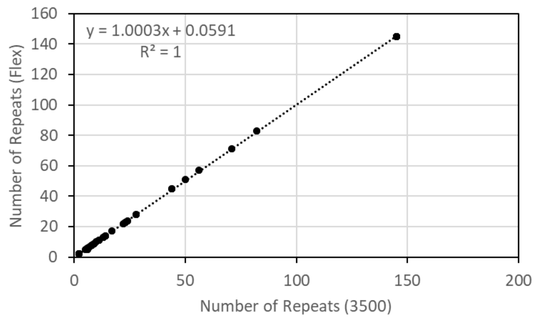
3500

Flex \ 3500	Normal	Intermediate	Premutation	Full Mutation
Normal	61	0	0	0
Intermediate	0	8	0	0
Premutation	0	0	27	1
Full Mutation	0	0	1	33

Figure 2. *FMR1* CGG repeat sizing and genotype agreement. A) *FMR1* CGG repeat size correlation plot. B) *FMR1* genotype agreement. Across all measurements agreement was 98.5% (131/133). Flex = SeqStudio Flex. The two discrepancies were at the 200 CGG repeat boundary between pre-mutation and full mutation categories (i. 198 vs. >200, ii. >200 vs. 198); accounting for size tolerance at this repeat range, these misses were within precision tolerance around the 200-repeat cutoff. In addition to QC failures, 11 samples were excluded from *FMR1* analysis due to high channel crosstalk. This issue was attributed to a defect identified in the SeqStudio Flex data collection software that results in application of an incorrect spectral calibration.

† Research use only. Not for use in diagnostic procedures

3A.



3B.

3500

Flex \ 3500	Normal	Intermediate	Expanded
Normal	6	0	0
Intermediate	0	10	0
Expanded	0	0	32

Figure 3. *C9orf72* repeat sizing and genotype agreement. A) *C9orf72* G_4C_2 repeat sizing correlation plot. B) *C9orf72* genotype agreement. Across all measurements agreement was 100% (48/48). Flex = SeqStudio Flex.

4A.

3500

Flex \ *SMN1* Cp#	0	1	2	3	4
0	27	0	0	0	0
1	0	27	0	0	0
2	0	0	73	0	0
3	0	0	0	20	0
4	0	0	0	0	27

4B.

3500

Flex \ *SMN2* Cp#	0	1	2	3	4
0	27	0	0	0	0
1	0	35	0	0	0
2	0	0	64	0	0
3	0	0	0	24	0
4	0	0	0	0	27

Figure 4. *SMN1* and *SMN2* copy number and variant agreement. A) *SMN1* exon 7 copy number agreement. Across all measurements agreement was 100% (174/174). B) *SMN2* exon 7 copy number agreement. Across all measurements agreement was 100% (177/177). Agreement for c.*3+80T>G, c.*211_*212del, and c.859G>C was 100% (184/184). Flex = SeqStudio Flex.

5A.

#CFTR Variants	0	1	≥2
0	29	0	0
1	0	16	0
≥2	0	0	18

5B.

Measure	N	Percent Agreement
PPV	402	100%
PPA	402	100%
NPA	2,819	100%
OPA	3,171	100%
PA_Z	3,171	100%
PA_{TTG}	72	100%

Figure 5. *CFTR* variant and sample genotype agreement. A) Sample-level genotype agreement. Sample genotypes reported by the kit are interpreted as, Wild Type: 0 variants. Heterozygous: 1 variant. Homozygous, Compound Heterozygous, or Multiple: ≥2 variants. Sample-level agreement was 100% (63/63). B) Variant-level agreement metrics. For each metric, the number of expected variant calls per instrument configuration is listed. Each sample measurement generates 62 unique variant calls. PolyT/TG agreement was evaluated separately. PPV = positive predictive value, PPA = positive percent agreement, NPA = negative percent agreement, OPA = overall percent agreement, PA_Z = zygosity agreement. PA_{TTG} = percent agreement PolyT/TG.

Conclusion

- We demonstrated the feasibility of analyzing five difficult genes—*FMR1, CFTR, SMN1, SMN2*, and *C9orf72*—on the SeqStudio Flex instrument
- This study included a challenging sample set with examples of all repeat size categories (*FMR1, C9orf72*), copy numbers (*SMN1, SMN2*) and variants (*CFTR, SMN1, SMN2)* detected by these assays
- Agreement was high (98-100%) across all gene targets, showing analogous performance to the 3500 platform for resolution of STRs, CNVs, SNVs, INDELs, and repeat sizes
- A data collection software defect was discovered that impacted *FMR1* results
- Strong performance on this challenging sample set suggests that the SeqStudio Flex is compatible with AmplideX PCR/CE assays already used to analyze challenging gene targets

Asuragen
a bio-techne brand