Development of a *CFTR* Control Panel to Monitor All ACMG-100 Variants

Jessica Hayden[1], Aldona Karaczyn[1], Joan Gordon[1], Joe Salem[1], Jacob Schurkman[1], Tania Spenlinhauer[1], Andrew Bradford[2], Katarzyna Gawlik[2], Laura Mendez[2]

[1] Maine Molecular Quality Controls Inc., Saco, Maine

[2] Agena Bioscience, San Diego, California

G021

Introduction

In 2001, the American College of Medical Genetics and Genomics (ACMG) published a list of 23 variants that should be included in carrier screening for cystic fibrosis. In 2023, ACMG updated their recommended variant list to 100 [1]. Next-generation sequencing (NGS) and MassARRAY are now routine methods to detect *CFTR* variants in clinical laboratories. Although testing is routine, template preparation and bioinformatics pipelines are complex. Sequence noise and amplification errors must be reliably distinguished from true variant calls. Sanger sequenced-verified external controls are critical to monitor the ability of any NGS or MassARRAY test system to correctly identify both simple and complex genome variants. Here, we describe development and verification of a safe, non-infectious, and easy-to-use 6-bottle *CFTR* control panel containing all 100 ACMG-recommended variants, as well as variants challenging to NGS systems.

Materials and Methods

Twenty-eight additional variants were added to existing plasmids from the previous version of the control panel (see Table 2, asterisk*). Plasmids contain all *CFTR* exons with intronic borders and were used to create a synthetic panel containing a total of 211 *CFTR* variants and all 100 of the recommended ACMG variants. Fragments containing *CFTR* variants were designed *in-silico*, and incorporated using Gibson assembly. Assembled constructs were transformed into bacteria and screened using Oxford Nanopore sequencing. Plasmids were purified from the selected clones and bi-directionally sequenced using Sanger (Figure 1).

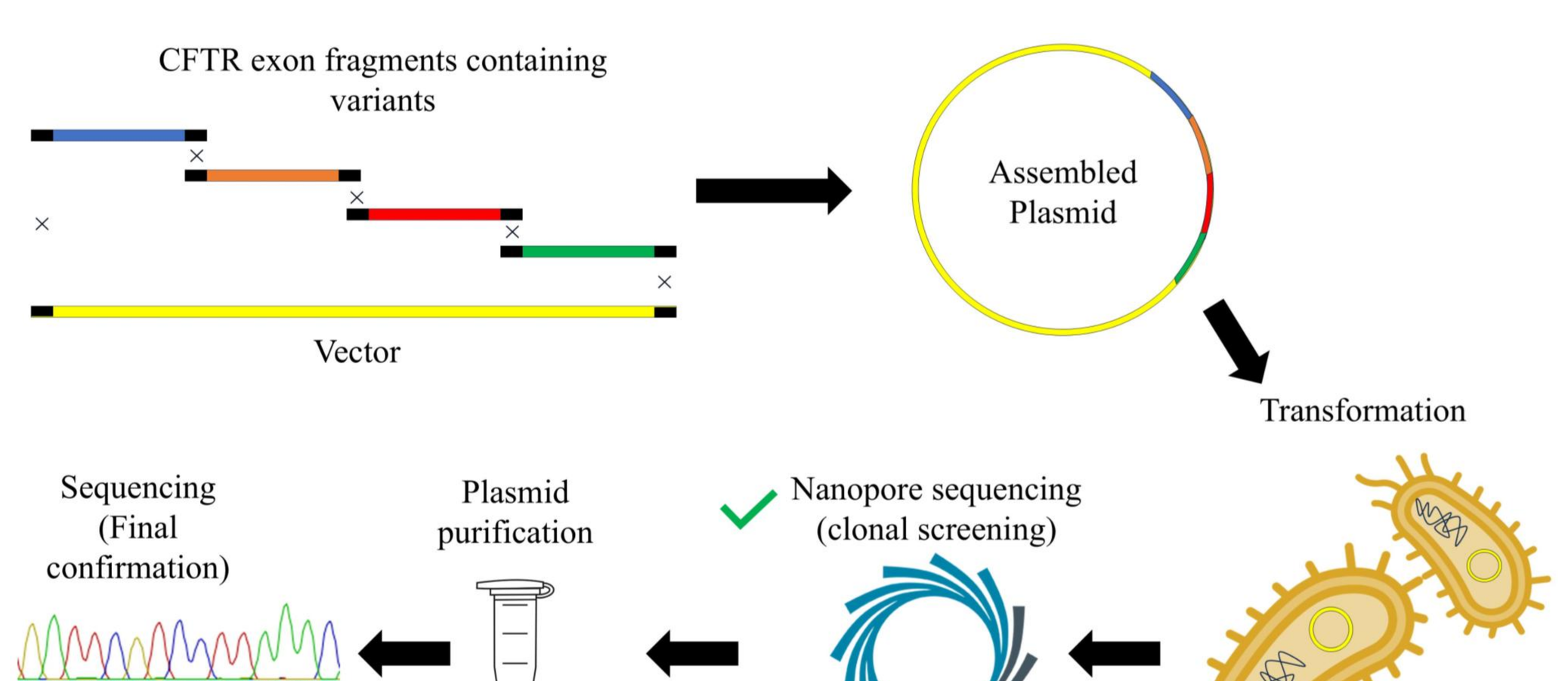


Figure 1. Overview of molecular cloning workflow: After transformation, colonies are screened using Nanopore whole plasmid sequencing. Reads are aligned using Geneious Prime to verify successful assembly. To ensure traceability, final plasmids undergo bi-directional Sanger sequencing.

Bottle	Control Name
A	G211A*plus*
B	G211Bv1.1
C	G211C*plus*
D	G211Dv1.1
E	G211E*plus*
F	G211F*plus*

Table 1 (left). NGS CF Control G211*plus* panel configuration: All 211 *CFTR* variants are split across 6 controls. Purified plasmids each containing unique *CFTR* variants were mixed at a 1:1 ratio to generate heterozygous calls.

Three unique lots (6 control bottles per lot) were tested by NGS using the Illumina TruSight™ CF Clinical Sequencing Assay on the MiSeqDx at MMQCI. One lot of G211*plus* controls was also provided to Agena Bioscience for testing on the new *CFTR*100+ assay using the MassARRAY® system. *CFTR*100+ is designed to detect all ACMG-100 variants. Figure 2 (below) shows a brief description of their workflows.

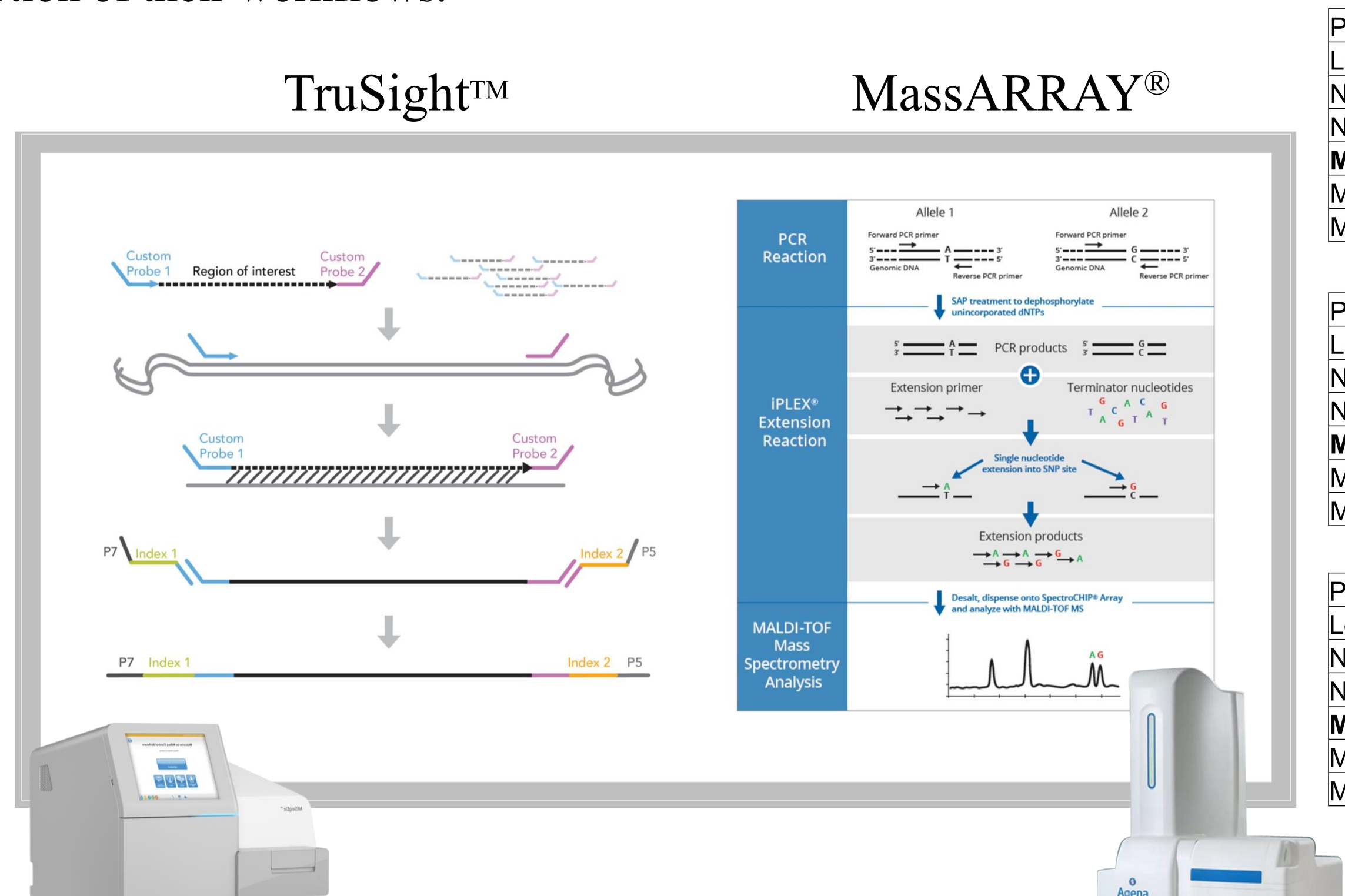


Figure 2. Test Systems: The controls are designed to monitor all aspects of the test system. Illumina TruSight CF library prep uses probe hybridization to target regions of interest for paired-end sequencing. Agena Bioscience's *CFTR*100+ assay uses PCR to target regions of interest. Subsequently, PCR products undergo a single base extension reaction (iPLEX chemistry) complementary to the allele of interest. Extension products are analyzed by MALDI-TOF MS and specific alleles are identified by unique masses.

Table 2: List of all 211 variants in the NGS CF Control G211*plus* panel (ACMG-100 is **bolded**, the 28 newly added ACMG-100 variants have an asterisk*)

cHGVS (Legacy)

c.1A>G (M1V)	c.805_806delAT (936delTA)		c.2260G>A (V754M)	c.3454G>C (D1152H)
c.4C>T (Q2X)*	**c.868C>T (Q290X)***	c.1573C>T (Q525X)	**c.2290C>T (R764XC>T)**	**c.3472C>T (R1158X)**
c.91C>T (R31C)	c.929TCT[2] (delF311)	**c.1584+1G>A (1716+1G->A)***	**c.2353C>T (R785X)***	**c.3484C>T (R1162X)**
c.115C>T (Q39X)	c.948delT (1078delT)	**c.1585-1G>A (1717-1G>A)**	**c.2374C>T (R792X)***	c.3485G>T (R1162L)
c.164+2T>A (296+2T>A)	**c.988G>T (G330X)**	c.1585-8G>A (1717-8G>A)	c.2453delT (2585delT)	**c.3528delC (3659delC)**
c.178G>T (E60X)	**c.1000C>T (R334W)**	**c.1624G>T (G542X)**	c.2464G>T (E822X)	c.3536_3539delCCAA (3667del4)
c.200C>T (P67L)	c.1007T>A (I336K)	c.1645A>C (S549R)	**c.2490+1G>A (2622+1G>A)**	c.3587C>G (S1196X)
c.223C>T (R75X)	**c.1013C>T (T338I)**	**c.1646G>A (S549N)**	c.2491G>T (E831X)	c.3611G>A (W1204X (3743G>A))
c.254G>A (G85E)	c.1021T>C (S341P)	**c.1647T>G (S549R)**	c.2537G>A (W846X)	**c.3612G>A (W1204X (3744G>A))**
c.262_263delTT (394delTT)	**c.1021_1022dupTC (1154insTC)**	**c.1651G>A (G551S)***	c.2538G>A (W846X(2670))	**c.3659delC (3791delC)**
c.271G>A (G91R)*	**c.1029del (1161delC)***	**c.1652G>A (G551D)**	c.2551C>T (R851X)	c.3700A>G (I1234V)
c.273G>A (G91G)	**c.1040G>A (R347H)**	c.1654C>T (Q552XC>T)	c.2583delT (2711delT)	c.3705T>G (S1235R)
c.273+1G>A (405+1 G>A)	**c.1040G>C (R347P)**	**c.1657C>T (R553X)**	c.2657+2_2657+3insA (2789+2insA)	**c.3717+5G>A (3849+5G->A)***
c.273+3A>C (405+3A>C)	**c.1055G>A (R352Q)**	**c.1673T>C (L558S)**	**c.2657+5G>A (2789+5G>A)**	**c.3718-2477C>T (3849+10kb C>T)**
	c.1081delT (1213delT)	**c.1675G>A (A559T)**	**c.2668C>T (Q890X)**	c.3717+4A>G (3849+4A>G)
c.274-1G>A (406-1 G>A)	c.1090T>C (S364P)	**c.1679G>C (R560T)**	c.2737_2738insG (2869insG)	c.3731G>A (G1244E)
c.274G>A (E92K)	c.1116+1G>A (1248+1G>A)	c.1679G>A (R560K)	**c.2739T>A (Y913X)***	**c.3744delA (3876delA)**
c.274G>T (E92XG>T)	c.1130dupA (1259insA)	**c.1679+1G>A (1811+1G->A)***	c.2780T>C (L927P)	c.3752G>A (S1251N)
c.292C>T (Q98X)	**c.1155_1156dupTA (1288insTA)**	**c.1680-886A>G (1811+1.6 kb)**	**c.2834C>T (S945L)**	c.3607A>G (S1255XA>G)
c.293A>G (Q98R)*	c.1202G>A (W401X (TAG))	**c.1680A>C (R560S)***	c.2875delG (3007delG)	**c.3764C>A (S1255XC>A)**
c.313delA (444delA)	c.1203G>A (W401X (TGA))	c.1680-1G>A (1812-1 G>A)	c.2908G>C (G970R)	c.3773dupT (3905insT)
c.325_327delinsG (457TAT>G)	c.1209+1G>A (1341+1G>A)	**c.1682C>A (A561E)***	**c.2909G>A (G970D)***	**c.3808del (3940delG)***
c.328G>C (D110H)	**c.1327_1330dupGATA (1461ins4)**	**c.1692del (1824delA)***	c.2930C>T (S977F)	c.3808G>A (D1270N)
c.349C>T (R117C)	**c.1364C>A (A455E)**	**c.1705T>G (Y569D)**	**c.2988G>A (3120G>A)**	**c.3846G>A (W1282X)**
c.350G>A (R117H)	**c.1367T>C (V456A)***	c.1727G>C (G576A)	**c.2988+1G>A (3120+1G>A)**	c.3873+1G>A (4005+1G>A)
c.366T>A (Y122X)	**c.1373del (1504delG)***	**c.1753G>T (E585X)**	c.2989-1G>A (3121-1G>A)	**c.3889dupT (4016insT)**
c.442delA (574delA)	**c.1393-1G>A (1525-1G>A)**	**c.1766+1G>A (1898+1 G>A)**	c.2991G>C (L997F (G>C))	**c.3909C>G (N1303K)**
c.443T>C (I148T)	c.1397C>A (S466X (C>A))	c.1766+3A>G (1898+3 A>G)	c.3039delC (3171delC)	c.3937C>T (Q1313X)
c.489+1G>T (621+1 G>T)	**c.1397C>G (S466X (C>G)**	**c.1766+5G>T (1898+5G>T)**	**c.3067_3072delATAGTG (3199del6)**	c.4077_4080delinsAA (4209TGTT>AA)
c.531delT (663delT)	**c.1400T>C (L467P)**	**c.1837G>A (A613T)***	c.3080T>C (I1027T)	c.4251delA (4382delA)
c.532G>A (G178R)	c.1408G>A (M470V)	c.1841A>G (D614G)	**c.3107C>A (T1036N)***	c.54-5940_273+10250del (CFTRdele2,3)
c.571T>G (F191V)*	c.1418delG (1548delG)	c.1865G>A (G622D)	**c.3140-26A>G (3272-26A>G)**	c.3964-78_4242+577del (CFTRdel22,23)
c.579+1G>T (711+1 G>T)	c.1438G>T (G480C)	**c.1882G>A (G628R)***	c.3194T>C (L1065P)	PolyT (5T to 9T)
c.579+3A>G (711+3 A>G)	c.1466C>A (S489X)	c.1923_1931delinsA (2055del9>A)	**c.3196C>T (R1066C)**	PolyTG (TG11, TG12)
c.579+5G>A (711+5 G>A)	c.1475C>T (S492F)	c.1976delA (2108delA)	**c.3197G>A (R1066H)**	
c.580-1G>T (712-1 G>T)	c.1477C>T (Q493X)	c.2002C>T (R668C)	c.3208C>T (R1070W)	
c.595C>T (H199Y)	c.1516A>G (I506V)	c.2012delT (2143delT)	c.3209G>A (R1070Q)	
c.613C>T (P205S C>T)	c.1519A>G (I507V)	c.2051_2052delinsG (2183AA>G)	c.3230T>C (L1077P)	
c.617T>G (L206W)	**c.1516ATC[1] (I507del)**	**c.2052dupA (2184insA)**	**c.3266G>A (W1089X)**	
c.653T>A (L218X)*	**c.1521_1523delCTT (F508del)**	**c.2052delA (2184delA)**	c.3276C>A (Y1092X (C>A))	
c.658C>T (Q220X)	c.1523T>G (F508C)	**c.2125C>T (R709XC>T)**	c.3276C>G (Y1092X (C>G))	
c.680T>G (L227R)	c.1545_1546delTA (1677delTA)	c.2128A>T (K710X)	**c.3294G>C (W1098C)***	
c.695T>A (V232D)*	c.1558G>T (V520F)	**c.2175dupA (2307insA)**	c.3302T>A (M1101K)	
c.723_743+1del (852del22)	**c.1572C>A (C524X)***	c.2195T>G (L732X)	c.3310G>T (E1104X)	
c.803delA (935delA)		c.2215delG (2347delG)	**c.3353C>T (S1118F)***	

Table 3: Breakdown of variant types and subtypes in the control panel.

Variant Types & Subtypes	Number of variant type in the control panel
Deletions	**30**
1 bp	*23*
2-6 bp	*6*
22 bp	*1*
Deletion-Insertion (Delins)	**4**
Duplications & Insertions	**10**
1 bp	7
2-4 bp	3
Large Genomic Rearrangements	**2**
Tandem repeat variants	**4**
Substitutions (exon)	**131**
A>C	2
A>G	7
A>T	1
C>A	8
C>G	4
C>T	33
G>A	33
G>C	8
G>T	11
T>A	6
T>C	10
T>G	8
Substitution (intronic)	**30**
A>C	1
A>G	5
C>T	1
G>A	18
G>T	4
T>A	1

Results (Next Generation Sequencing)

Bottle	P/N	Lot #	Total (n) across 3 sequencing runs	NGS Call rate	Correct calls/Expected Calls	% Concordance
A	NGS CF Control Panel G211A*plus*	A27MAY25A	6	100.00%	38/38	100.0%
		E18JUN25A	6	100.00%	38/38	100.0%
		E02JUL25A	6	100.00%	38/38	100.0%
B	NGS CF Control Panel G211Bv1.1	B28MAY25A	6	99.04%	38/38	100.0%
		C23JUN25A	6	99.04%	38/38	100.0%
		B08JUL25A	6	99.04%	38/38	100.0%
C	NGS CF Control Panel G211C*plus*	C27MAY25A	6	99.65%	36/36	100.0%
		G18JUN25A	6	99.65%	36/36	100.0%
		H02JUL25A	6	99.65%	36/36	100.0%
D	NGS CF Control Panel G211Dv1.1	D28MAY25A	6	99.04%	39/39	100.0%
		E23JUN25A	6	99.04%	39/39	100.0%
		D08JUL25A	6	99.04%	39/39	100.0%
E	NGS CF Control Panel G211E*plus*	E28MAY25A	6	99.94%	33/33	100.0%
		G23JUN25A	6	99.94%	33/33	100.0%
		E08JUL25A	6	99.94%	33/33	100.0%
F	NGS CF Control Panel G211F*plus*	F28MAY25A	6	99.04%	44/45	99.8%
		J23JUN25A	6	99.04%	44/45	99.8%
		U09JUL25A	6	99.04%	44/45	99.8%
			Total n = 108	Avg call rate = 99.45%	684/687	99.60%

Table 4. Total NGS sample replicates and percent correct calls: A total of 18 G211*plus* lots (6 controls x 3 lots each) were manufactured and tested across 3 Illumina TruSight CF runs. The Clinical Sequencing Assay reports detected variants along with their specific allele frequencies. Call rate, a reported percentage, is the number of variant regions that meet a defined confidence value threshold divided by the total regions targeted. Each control lot was tested 6 times (total n = 108). Bottles A through E had 100% correct calls across all NGS runs. Variant c.3700A>G in Bottle F is not reported because its allele frequency is below a proprietary cutoff set by Illumina. The low frequency is caused by a known near-neighbor interference issue with c.3612G>A. c.3700A>G can still be observed using IGV at ~15% allele frequency (Figure 3, below).

P/N	G211A*plus*	G211A*plus*	G211A*plus*	G211Bv1.1	G211Bv1.1	G211Bv1.1
Lot	A27MAY25A	E18JUN25A	E02JUL25A	B28MAY25A	C23JUN25A	B08JUL25A
N replicates	6	6	6	6	6	6
N calls	216	216	216	222	222	222
Mean AF	**0.499**	**0.509**	**0.508**	**0.477**	**0.515**	**0.469**
Median AF	0.498	0.518	0.516	0.468	0.505	0.455
Minimum AF	0.316	0.359	0.346	0.385	0.442	0.405

P/N	G211C*plus*	G211C*plus*	G211C*plus*	G211Dv1.1	G211Dv1.1	G211Dv1.1
Lot	C27MAY25A	G18JUN25A	H02JUL25A	D28MAY25A	E23JUN25A	D08JUL25A
N replicates	6	6	6	6	6	6
N calls	210	210	210	228	228	228
Mean AF	**0.506**	**0.509**	**0.509**	**0.488**	**0.484**	**0.450**
Median	0.487	0.497	0.492	0.479	0.471	0.435
Minimum	0.366	0.396	0.401	0.347	0.378	0.340

P/N	G211E*plus*	G211E*plus*	G211E*plus*	G211F*plus*	G211F*plus*	G211F*plus*
Lot	E28MAY25A	G23JUN25A	E08JUL25A	F28MAY25A	J23JUN25A	U09JUL25A
N replicates	6	6	6	6	6	6
N calls	192	192	192	252	252	252
Mean AF	**0.503**	**0.495**	**0.430**	**0.473**	**0.457**	**0.476**
Median	0.509	0.497	0.429	0.472	0.456	0.475
Minimum	0.338	0.357	0.304	0.338	0.342	0.359

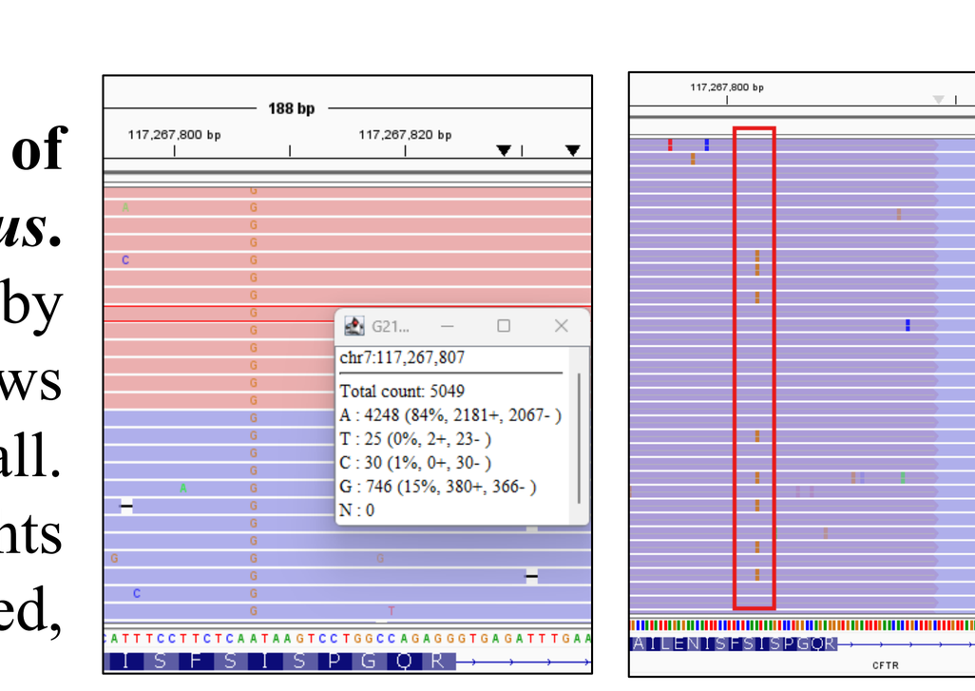
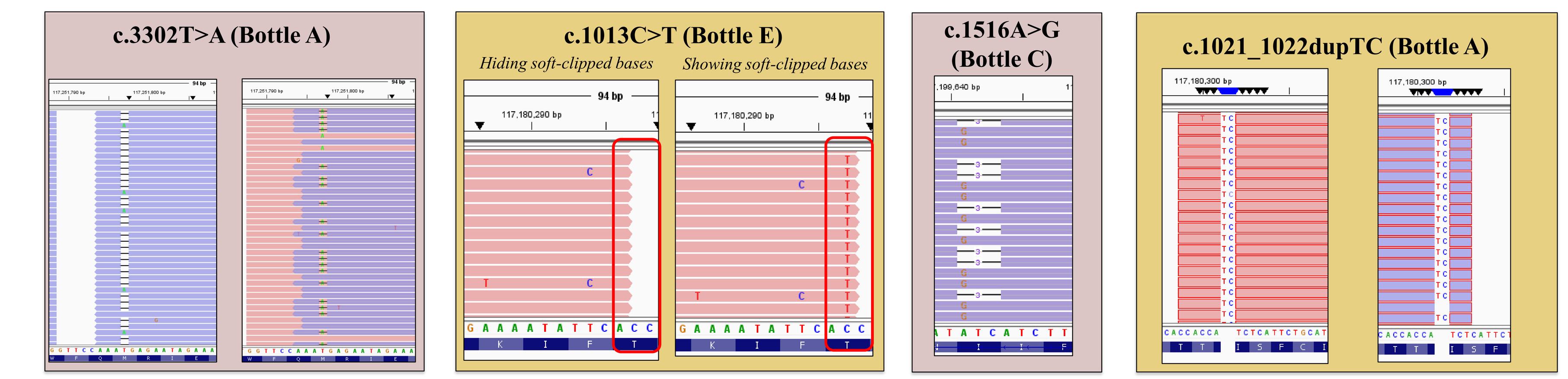


Figure 3. IGV Visualization of c.3700A>G in G211F*plus*. The left image is sorted by base at c.3700 and shows the frequency of each call. The right image highlights c.3700A>G in paired, unsorted reads.

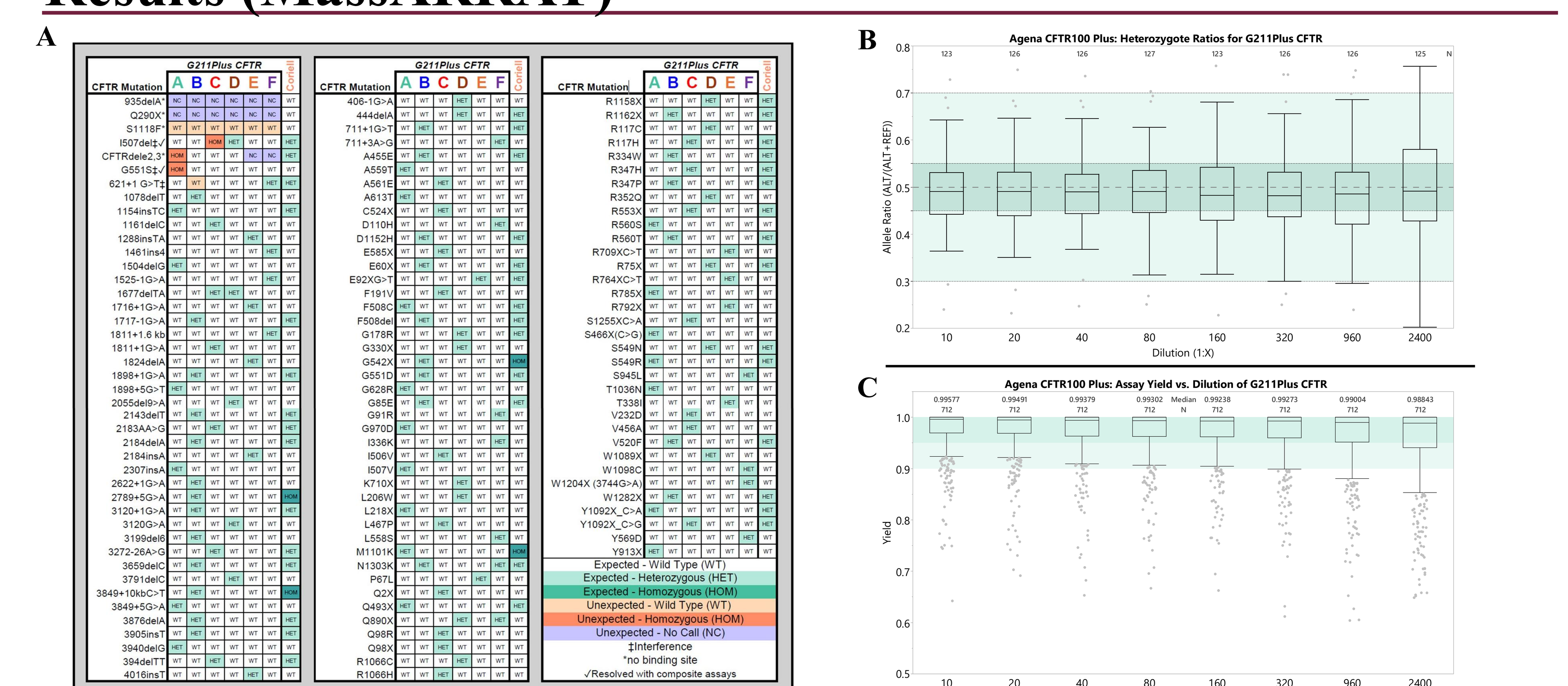
Table 5. Lot-to-lot reproducibility of allele frequency: Allele frequency (AF) summary statistics, including averages, medians, and minimums in each control were calculated using JMP 18. Average allele frequencies ranged from 0.430 to 0.515. Note that PolyT/PolyTG variants and the two LGRs (CFTRdele2,3 and CFTRdele22,23) are reported without frequencies. Homozygous variant c.1408G>A is present in Bottles B and D.

Table 6. Examples with IGV images highlighting the importance of quality controls when interpreting bioinformatic results. Below are examples of how bioinformatics pipelines and variant location can impact NGS reporting.

Variant	Control	TruSight AF Skew	Cause of Skew	Details
c.3302T>A	G211A*plus*	Low	Illumina bioinformatics	The presence of the T>A substitution, which is adjacent to an 'A'-homopolymer causes a misalignment at the end of reads. The shift is counted as a "deletion" rather than the 'A' allele.
c.1013C>T	G211E*plus*	Low	Illumina bioinformatics	When present as the last base at the end of reads, the allele is removed by bioinformatic soft-trimming which significantly reduces the frequency.
c.1516A>G	G211C*plus*	High	Overlapping variants in the control	c.1516A>G reports at near homozygous frequency because it is paired with deletion c.1516ATC[1] in the same control. One allele is 'G' and the other allele is deleted.
c.1021_1022dupTC	G211A*plus*	High	Illumina bioinformatics	Allele frequency is elevated resulting from a population of unpaired reads containing the TC-duplication.



Results (MassARRAY)



D

Bottle	Variant	Reason for discrepancy in MassARRAY testing
D	c.803delA (935delA)	Control is missing intronic coverage.
D	c.868C>T (Q290X)	Control is missing intronic coverage.
F	c.3353C>T (S1118F)	Control is missing intronic coverage for the MUT allele.
A	c.54-5940_273+10250del21kb (CFTRdele2,3)	Control is missing intronic for the WT allele.
A	c.1651G>A (G551S)	Extension primer interference from c.1647T>G (S549R) impacts WT allele.
B	c.489+1G>T (621+1 G>T)	Nearby c.443T>C interferes with PCR primer.

Figure 4. Results from G211*plus* testing on the Agena *CFTR*100+ assay: (A) Variant calling heat map of G211*plus* controls and Coriell samples. (B) Allelic ratios across a dilution series of the control material. The controls were originally developed for NGS. The dilution series suggests that the *CFTR*100+ assay is more sensitive than NGS. (C) Box plot of assay yield across a dilution series. Yield is a metric that indicates assay performance. Yield > 0.5 is passing. Yield > 0.8 is the target performance goal. G211*plus* controls generated a median yield of 0.98. (D) Causes for G211*plus* discrepant calls when tested on the *CFTR*100+ assay.

Conclusions

- We have developed a new *CFTR* external control panel that can be used to monitor all 100 of the updated ACMG-recommend variants for carrier screening.
- The control panel includes clinically relevant variants of many types that challenge bioinformatic and variant reporting pipelines, including 161 substitutions representing all possible nucleotide changes, 30 small deletions, 4 complex insertion/deletion variants, 10 small insertions/duplications, 2 large genomic rearrangements, and Poly-T/Poly-TG repeats.
- The control panel serves to confidently control for NGS and MassARRAY wet-lab methods despite their fundamentally different chemistries, detection, and output.
- G211*plus* controls ensure bioinformatic pipelines can detect all relevant *CFTR* variants and reveal causes of allelic skew from bioinformatic processing.

Acknowledgements: Special thanks to Agena Bioscience and their colleagues: Andy Bradford for the MassARRAY figures, Laura Mendez who ran the titration experiment, and Katarzyna Gawlik who tested the Coriell samples.

[1] Deignan JL, Gregg AR, Monaghan KG, et al. Updated recommendations for CFTR carrier screening: A position statement of the American College of Medical Genetics and Genomics. *Genetics in Medicine*. 2023