Multiplex Synthetic Reference Material for Monitoring the Analytical Performance of Highly Complex Variant Detection of Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) using Next Generation Sequencing

G063

Ritu Mihani, Keifer Boguhn, Michael Escott, Alyson Fischer, Joan Gordon, Daniel Magoon, Steve Nesbitt, Clark Rundell, Joseph Salem, Tania Spenlinhauer and Michael Steffen

Maine Molecular Quality Controls Inc., Saco, Maine

Introduction

The ability of NGS to deliver quantitative and qualitative analysis of nucleic acids in a single test with increased diagnostic potential has made it the system of choice for many clinicians. However, errors introduced during sample preparation, library construction and bioinformatics analysis make the implementation of NGS challenging. This is further compounded by the inability of NGS technology to accurately detect challenging variants particularly in regions of low sequence complexity and high GC/AT ratios. In 2017, AMP and CAP published a joint guideline for validation of NGS panels[1] and bioinformatic pipelines[2] to mitigate these problems. To this end, MMQCI developed NGS CF Control Panel G211 v1.1, a comprehensive synthetic reference panel, to be used as a positive control for monitoring the analytical performance of complex NGS panels examining the CFTR gene.

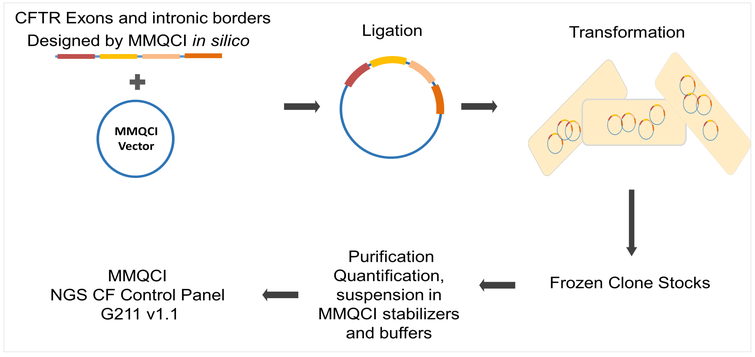
In this study, we describe the development and performance of NGS CF Control Panel G211 v1.1, an updated version of NGS CF Control Panel G211, which includes additional sequences to optimize function for enhanced coverage and accurate detection of additional variants within the CFTR gene.

Materials and Methods

NGS CF Control Panel G211 v1.1 Design: Synthetic DNA composed of all 27 CFTR gene exons plus intronic borders representing 188 variants including 145 single nucleotide variants, 42 INDELs and homopolymers up to 9Ts were designed *in silico,* ligated into MMQCI vectors and transformed to create stable frozen clone stocks. Multiple plasmids were created with various mutations and mixed to create either heterozygous or homozygous alleles for each variant and diluted to have equivalent copy numbers of the targeted gene as extracted human samples. Copy number equivalency between the synthetic controls and genomic DNA was established using an in-house qPCR assay. Traceability of each mutation in all the plasmids was established by performing bi-directional, quality-scored Sanger sequencing of all genome segments. The plasmid mixes were suspended in buffers and stabilizers and tested as non extractable control panel.

To evaluate the performance of NGS CF Control Panel G211 v1.1, all six tubes of NGS CF Control Panel G211 v1.1 (NGS Control Panel G211 v1.1 Tube A to Tube F) and all six tubes of NGS CF Control Panel G211 (NGS Control Panel G211 Tube A to Tube F) along with in-house MMQCI genomic DNA and NA24385 with and without synthetic plasmids were tested on MiSeqDx™ CF 139-Variant Assay with MiSeqDx™ platform on two independent sequencing runs. All the samples were prepared according to MiSeqDx™ CF 139-Variant Assay protocol. Testing of NGS CF Control Panel G211 v1.1 was also performed on Luminex® xTAG® Cystic Fibrosis v2 Assay to test the utility of the control panel on multiple diagnostic platforms.

Figure 1. Design Strategy for NGS CF Control Panel G211 v1.1



References

1. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for validation of next-generation sequencing-based oncology panels: A joint consensus recommendation of the association for molecular pathology and College of American Pathologists. J Mol Diagn. 2017;19:341–65.
2. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: A Joint recommendation of the association for molecular pathology and the College of American Pathologists. J Mol Diagn. 2018;20:4–27.
3. The Clinical and Functional Translation of CFTR (CFTR2): https://www.cftr2.org/mutations_history

Results

Evaluation of NGS CF Control Panel G211 v1.1 on MiSeqDx™ CF 139-Variant Assay

Table 1. NGS CF Control Panel G211 and NGS CF Control Panel G211 v1.1 on MiSeqDx™ CF 139-Variant Assay. 1A. Performance Characteristics of NGS CF Control Panel G211. All components of NGS CF Control Panel G211, along with all additional samples, had a passing call rate (>99%) and >1000X mean base coverage. The following run metrics were obtained: ~7.2 million reads; cluster density: 739 K/mm^2; Clusters Passing Filter: 95.3%. A total of 136 out of 139 mutations were detected for the NGS CF Control Panel G211. The 3 missed calls resulted in 97.84% concordance for NGS CF Control Panel G211 with MiSeqDx™ CF 139-Variant Assay. **1B. Performance Characteristics of NGS CF Control Panel G211 v1.1.** All components of NGS CF Control Panel G211 v1.1, along with additional samples passed with 100% call rate and >1000X mean base coverage. The run resulted in the following metrics: ~8.13 million reads; cluster density: 835 K/mm^2; Clusters Passing Filter: 94.8%. All 139 mutations were detected resulting in 100.0 % concordance for the NGS CF Control Panel G211 v1.1 with MiSeqDx™ CF 139-Variant Assay.

1A. Performance Characteristics of NGS CF Control Panel G211

NGS CF Control Panel G211 Component	Call rate	Correct Calls/Expected Calls	% Concordance
NGS CF Control Panel G211 A	100.0 %	18/18	100.0 %
NGS CF Control Panel G211 B	100.0 %	30/30	100.0 %
NGS CF Control Panel G211 C	100.0 %	20/21	100.0 %
NGS CF Control Panel G211 D	99.26 %	24/25	96.0 %
NGS CF Control Panel G211 E	100.0 %	20/20	100.0 %
NGS CF Control Panel G211 F	100.0 %	29/30	100.0 %
MMQCI Genomic DNA	100.0%	1/1	100.0 %
MMQCI Genomic DNA with MMQCI synthetic plasmid	100.0%	30/30	100.0 %
NA24385	100.0%	0/0	100.0 %
NA24385 with MMQCI Synthetic Plasmid	100.0%	30/30	100.0 %
	Total mutations detected	136/139	97.84 %

1B. Performance Characteristics of NGS CF Control Panel G211 v1.1

NGS CF Control Panel G211 v1.1 Component	Call rate	Correct Calls/Expected Calls	% Concordance
NGS CF Control Panel G211v1.1 A	100.0 %	18/18	100.0 %
NGS CF Control Panel G211v1.1 B	100.0 %	30/30	100.0 %
NGS CF Control Panel G211v1.1 C	100.0 %	21/21	100.0 %
NGS CF Control Panel G211v1.1 D	100.0 %	25/25	100.0 %
NGS CF Control Panel G211v1.1 E	100.0 %	20/20	100.0 %
NGS CF Control Panel G211v1.1 F	100.0 %	30/30	100.0 %
MMQCI Genomic DNA	100.0%	1/1	100.0 %
MMQCI Genomic DNA with MMQCI synthetic plasmid	100.0%	30/30	100.0 %
NA24385	100.0%	0/0	100.0 %
NA24385 with MMQCI Synthetic Plasmid	100.0%	30/30	100.0 %
	Total mutations detected	139/139	100.0 %

Table 2. Variants Detected in NGS CF Control Panel G211 v1.1. List of variants included in the NGS CF Control Panel G211 v1.1.
2A. All 139 variants were accurately called in NGS CF Control Panel G211 v1.1 when run on MiSeqDx™ CF 139-Variant Assay. The three variants not detected in NGS CF Control Panel G211 and now detected in NGS CF Control Panel G211 v1.1 are in highlighted in red.

Call rate	100.00%	Call rate	100.00%	Call rate	100.00%	Call rate	100.00%	Call rate	100.00%	Call rate	100.00%	Call rate	100.00%	Call rate	100.00%
CFTRdele2, 3	Detected	711+1G>T	Detected	3659delC	Detected	R553X	Detected	2711delT	Detected	R764X	Detected	E831X	Detected	E60X	Detected
Y122X	Detected	H199Y	Detected	3849+10kb	Detected	E585X	Detected	Q890X	Detected	L1065P	Detected	R851X	Detected	G85E	Detected
1154insTC	Detected	1078delT	Detected	3876delA	Detected	2183AA>G	Detected	3120G>A	Detected	W1204X)	Detected	L927P	Detected	4382delA	Detected
1248+1G>A	Detected	R334W	Detected	3905insT	Detected	3272-26	Detected	R1066C	Detected	G1244E	Detected	S945L	Detected	3007delG	Detected
W401X	Detected	R347P	Detected	W1282X	Detected	R1066H	Detected	W1089X	Detected	4016insT	Detected	G970R	Detected	E822X	Detected
S466X	Detected	A455E	Detected	N1303K	Detected	Y1092X	Detected	R1158X	Detected	M1V	Detected	3121-1G>A	Detected	1812-1G>A	Detected
S492F	Detected	F508del	Detected	394delTT	Detected	R75X	Detected	S1196X	Detected	Q39X	Detected	E1104X	Detected	Q552X	Detected
Q493X	Detected	V520F	Detected	Q98X	Detected	406-1G>A	Detected	3791delC	Detected	E92K	Detected	W1204X	Detected	R709X	Detected
S549R	Detected	1717-1G>A	Detected	R117H	Detected	R117C	Detected	P67L	Detected	D110H	Detected	4005+1G>A	Detected	2585delT	Detected
A559T	Detected	G542X	Detected	711+5G>A	Detected	G178R	Detected	405+1G>A	Detected	711+3A>G	Detected	CFTRdele22,23	Detected	3120+1G>A	Detected
2307insA	Detected	G551D	Detected	P205S	Detected	L206W	Detected	E92X	Detected	Q220X	Detected	I506V	Detected	I507del	Detected
W846X	Detected	R560T	Detected	852del22	Detected	G330X	Detected	457TAT>G	Detected	I336K	Detected	I507V	Detected	K710X	Detected
Y1092X	Detected	1898+1G>A	Detected	S341P	Detected	R352Q	Detected	574delA	Detected	1213delT	Detected	F508C	Detected	2347delG	Detected
M1101K	Detected	2143delT	Detected	R347H	Detected	W401X	Detected	663delT	Detected	1259insA	Detected	Q525X	Detected	2184insA	Detected
S1251N	Detected	2184delA	Detected	1341+1G>A	Detected	1677delTA	Detected	712-1G>T	Detected	1461ins4	Detected	1717-8G>A	Detected	L732X	Detected
Q1313X	Detected	2622+1G>A	Detected	PolyTG	Detected	S549N	Detected	T338I	Detected	1525-1G>A	Detected	L1077P	Detected	1898+3A>G	Detected
4209TGTT>AA	Detected	2789+5G>A	Detected	L467P	Detected	R560K	Detected	S466X	Detected	621+1G>T	Detected	R1162X	Detected	S549R	Detected
1548delG	Detected	1811+1.6kbA>G	Detected	S489X	Detected										

2B. In addition to the 139 variants detected by MiSeqDx™ CF 139-Variant Assay, NGS CF Control Panel G211 v1.1 contains an additional 49 variants; 18 of which are classified as pathogenic.

Legacy	Clinical Significance	Allele Frequency; CFTR2-AUG2018	Legacy	Clinical Significance	Allele Frequency; CFTR2-AUG2018	Legacy	Clinical Significance	Allele Frequency; CFTR2-AUG2018	Legacy	Clinical Significance	Allele Frequency; CFTR2-AUG2018
3171delC	CF-causing	0.004	Y569D	CF-causing	0.032	R1070Q	Ambiguous	0.015	444delA	CF-causing	0.01
1898+5G->T	CF-causing	0.004	D579G	Ambiguous	0.037	2055del9->A	CF-causing	0.016	S1255X	CF-causing	0.011
F508C	Non CF-causing	0.006	D1270N	Ambiguous	0.039	R31C	Non CF-causing	0.016	R1162L	Non CF-causing	0.011
2869insG	CF-causing	0.006	R668C	Non CF-causing	0.048	V754M	Non CF-causing	0.018	D614G	Ambiguous	0.012
405+3A->C	CF-causing	0.007	G576A	Non CF-causing	0.052	L227R	CF-causing	0.02	1288insTA	CF-causing	0.014
G1069R	Ambiguous	0.007	2789+2insA	Unknown	0.059	I1234V	CF-causing	0.023	3849+4A->G	CF-causing	0.01
S977F	Ambiguous	0.009	I1027T	Non CF-causing	0.066	L558S	CF-causing	0.024	G480C	Unknown	Unknown

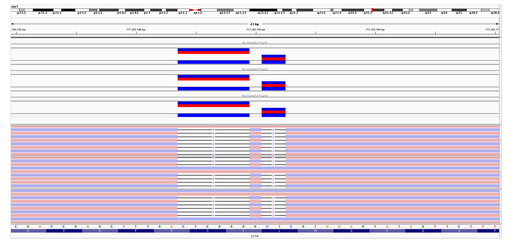
Performance of NGS CF Control Panel G211 v1.1 on Luminex® xTAG® Cystic Fibrosis v2 Assay

Table 3. NGS CF Control Panel G211 v1.1 on Luminex xTAG® Cystic Fibrosis v2 Assay. All components of NGS CF Control Panel G211 v1.1 (NGS CF Control Panel G211 v1.1 Tube A to Tube F) were tested in duplicate along with in-house MMQCI Genomic DNA on Luminex xTAG® Cystic Fibrosis v2 Assay using xTAG® CF71v2 Kit. All the samples were prepared according to manufacturer's protocol. The samples were read on Luminex® 200 analyzer and the output was analyzed using xTAG® Data Analysis Software. All samples passed, except one which failed due to a processing error. Of the samples that passed, all variants were detected accurately.

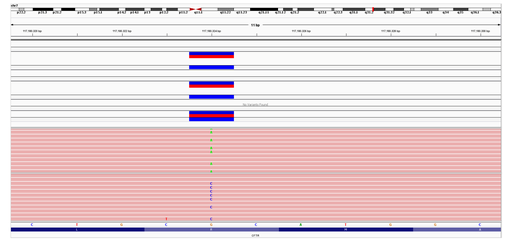
NGS CF Control Panel G211 v1.1 Component	Correct Calls/ Expected Calls	% Concordance
NGS CF Control Panel G211v1.1 A	69/69	100.0 %
NGS CF Control Panel G211v1.1 B	75/75	100.0 %
NGS CF Control Panel G211v1.1 C	75/75	100.0 %
NGS CF Control Panel G211v1.1 D	75/75	100.0 %
NGS CF Control Panel G211v1.1 E	74/74	100.0 %
NGS CF Control Panel G211v1.1 F	74/74	100.0 %
MMQCI Genomic DNA	75/75	100.0 %

Figure 2. Complex Variant Detection Study. NGS CF Control Panel G211 v1.1 contains 188 variants including 145 single nucleotide variants, 42 INDELs and homopolymers up to 9Ts spanning the CFTR gene. The control panel represents a milieu of complex variants including horizontally complex variants, vertically complex variants, short quad-nucleotide deletion, deletions in low GC rich regions and homopolymer repeats, thus challenging the bioinformatics algorithm for robustness.

2A. Horizontally Complex Variant (2055del9>A). Horizontally complex variant present in NGS CF Control Panel G211 v1.1. Different bioinformatics algorithm may create different representations.



2B. Vertically Complex (R347P). SNP changes at the same position. Different variant callers may lead to discrepant calls.



Conclusions

- MMQCI NGS CF Control Panel G211 v1.1 can be used as a routine control to monitor the analytical performance of each one of the 139 mutations detected by MiSeqDx™ CF 139-Variant Assay.
- NGS CF Control Panel G211 v1.1 harbors rare and complex variants such as horizontally and vertically complex variants. Complex variants aid in optimization and validation of bioinformatic pipelines in clinical labs thus making NGS CF Control Panel G211 v1.1 a robust control material to aid in successful validation of NGS targeted panel in a clinical lab.
- MMQCI can manufacture complex NGS synthetic reference material to aid in development, optimization and validation of NGS assays to ensure correct calling of multiple variants.