G45

Utility of a Multiplex Synthetic Control Material for Monitoring the Identification of Multiple Variants of Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) using Next Generation Sequencing

Joan Gordon[1], Steven Kaul[1], Todd Laughlin[2], Steve Nesbitt[1], Rocky Pramanik[3], Clark Rundell[1], Tania Spenlinhauer[1], Jennifer Starbuck[4], Jennifer Stone[5], and Cynthia Zimmerman[5]

[1]Maine Molecular Quality Controls Inc., Saco, Maine; [2]University of Rochester Medical Center, Rochester, NY; [3]Vantari Genetics, Irvine, CA; [4]Cleveland Clinic, Cleveland, OH; [5]MRIGlobal, Palm Bay, FL

Background

Next Generation Sequencing (NGS) provides an opportunity for clinicians to identify multiple mutations in a highly efficient and high-throughput manner. However, this technology requires a complex multistep procedure which presents many challenges for clinical laboratories. Each application and platform is unique and involves challenging, time-consuming, and expensive optimization and validation of 3 complex components – the sequencing platform, the specific assay/test panel and the bioinformatics analysis. Here we demonstrate a practical solution for monitoring the identification of many mutations using a multiplex, synthetic control panel created to monitor the analytical performance of molecular cystic fibrosis testing. The multiple insertions, deletions and homopolymers of varying lengths and composition in the control make it potentially useful for monitoring the ability of NGS systems to correctly identify variants found in genes other than CFTR.

Materials and Methods

Synthetic DNA composed of all 27 CFTR gene exons plus intronic borders containing CF associated variants were designed *in silco,* ligated into MMQCI vectors and transformed to create stable frozen clone stocks. Multiple plasmids were created with various mutations to represent 186 CFTR variants 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. Traceability was established by performing bi-directional, quality-scored Sanger sequencing of all genome segments. The plasmid mixes were suspended in buffers and stabilizers, with and without proprietary matrix and tested either as an extractable (with matrix) or non-extractable (without matrix) control panel. Initial studies were performed to determine optimal concentrations for subsequent testing. The extractable samples were extracted by various methods (QiaAmp DNA Blood Mini kit, MagNAMAX, QiASymphony, and SPRI-TE) and processed the same as a patient sample. The non-extractable panel was tested using 10µl of each sample added to 5uL of Oligo pool and 35uL, of HYB buffer. All samples were tested with the Illumina's MiSeqDX™ CF 139-Variant Assay.

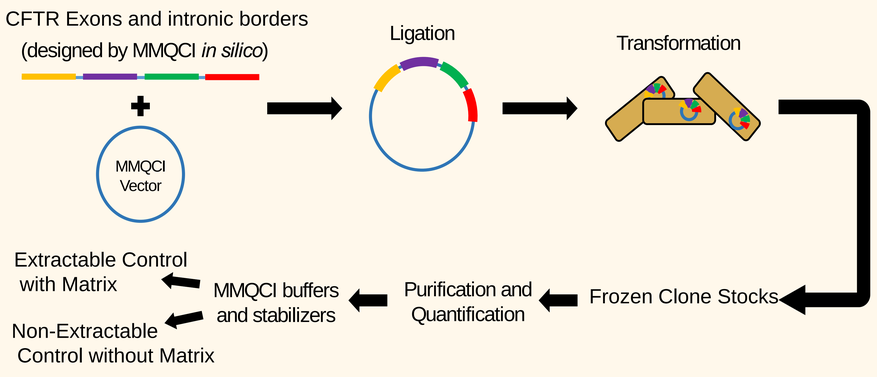

Figure 1

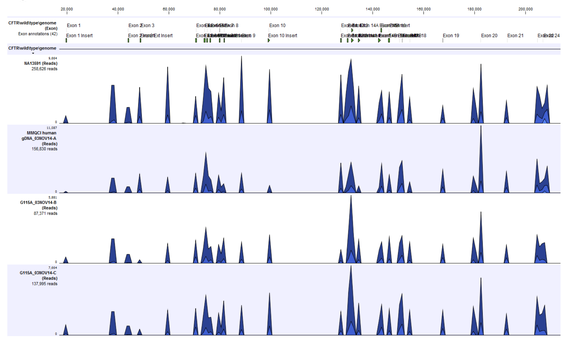

Figure 1. Coverage generated by MMQCI Synthetic Controls. MMQCI synthetic controls generated the same depth of coverage as the human gDNA sample, between1,000 – 1,700x coverage. Relative coverage of the target regions appear similar for MMQCI synthetic controls as for the human gDNA samples

Figure 2

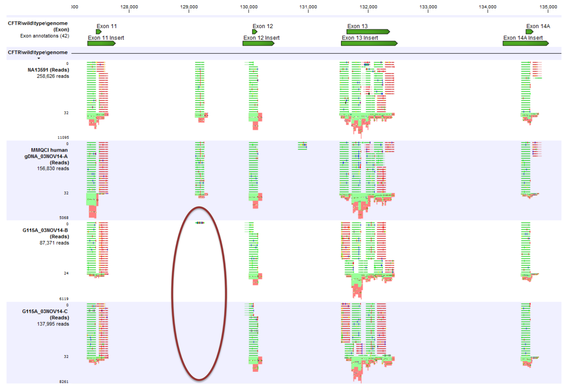

Figure 2. Raw Read of MMQCI Synthetic Controls and human gDNA samples. Raw sequence reads of the targeted areas of the CFTR gene when amplified with Illumina's MiSeqDX™ CF 139-Variant Assay. Raw reads generated show good coverage for most of the targeted areas for the Illumina's MiSeqDX™ CF 139-Variant Assay. The red circle highlights missing sequence for Intron 12 within MMQCI synthetic controls. The poor coverage for this targeted region results in a 'No Call' result for variants in this region when run on the Illumina's MiSeqDX™ CF 139-Variant Assay due to the lack of sequencing reads for this area.

Acknowledgements:

We would like to thank University of Rochester Medical Center, Rochester, NY; Vantari Genetics, Irvine, CA and Cleveland Clinic, Cleveland, OH for providing data for this study. We would also like to thank MRIGlobal, Palm Bay, FL for providing collaboration, data and instrumentation.

TABLE 1

Extractable Controls			
Call Rate	99.26	Call Rate	99.26
CFTR dele2,3	4/4	Q552X	1/1
E60X	5/5	R553X	5/5
R75X	5/5	A559T	4/4
G85E	5/5	R560T	5/5
394delTT	5/5	1811+1.6kb A>G	No Call
406-1 G>A	3/3	1812-1 G>A	3/3
R117C	3/3	1898+1 G>A	5/5
R117H	5/5	2143delT	5/5
Y122X	4/4	K710X	3/3
621+1 G>T	5/5	21833AA>G	5/5
G178R	3/3	2184delA	5/5
711+1 G>T	5/5	2307insA	4/4
711+5 G>A	4/4	2789+5G>A	5/5
L206W	3/3	Q890X	3/3
1078delT	5/5	3120G>A	3/3
G330X	3/3	3120+1G>A	5/5
R334W	5/5	3272-26A>G	4/4
R347H	5/5	R1066C	3/3
R347P	5/5	W1089X	3/3
R352Q	3/3	Y1092X (C>A)	4/4
PolyTG/PolyT	5/5	Y1092X (C>G)	1/1
A455E	5/5	M1101K	4/4
Q493X	4/4	R1158X	3/3
I507del	5/5	R1162X	5/5
F508del	5/5	3659delC	5/5
1677delTA	4/4	S1196X	2/3
V520F	5/5	3791delC	3/3
1717-1G>A	5/5	3849+10kbC>T	5/5
G542X	5/5	3876delA	5/5
S549R (c.1645A>C)	4/4	S1251N	5/5
S549R (c.1647T>G)	5/5	3905insT	5/5
S549N	5/5	W1282X	5/5
G551D	5/5	N1303K	5/5

Table 1. Results of Extractable Controls. The extractable controls were tested across 4 sites with all samples resulting in a call rate of 99.26%. 278 out of 279 total variants were identified resulting in 99.6% total concordance. 1811+1.6kb A>G resulted in a No Call result for all samples, due to insufficient sequence surrounding this variant.

Table 2. Summary of Results of Extractable Controls. Three of the testing sites had 100% concordance of expected results. The fourth site did not identify variant S1196X, resulting in 98.5% concordance.

TABLE 2

Sites	1	2	3	4
Extraction method	QiaAMP DNA Blood Mini	MagNA Max	QiaSympony	SPRI-TE
% Concordance	100%	100%	100%	98.5%

TABLE 3

Non-Extractable Controls						
Call Rate	99.26	Call Rate	99.26	Call Rate	99.26	Call Rate	99.26
M1V	2/2	R347H	2/2	R560T	2/2	L1065P	2/2
CFTR dele2,3	2/2	R347P	2/2	R560K	2/2	R1066C	2/2
Q39X	2/2	R352Q	2/2	1811+1.6kb A>G	No Call*	R1066H	2/2
E60X	2/2	1213delT	2/2	1812-1 G>A	2/2	L1077P	2/2
P67L	2/2	1248+1GA	2/2	E585X	2/2	W1089X	2/2
R75X	2/2	1259insA	2/2	1898+1 G>A	2/2	Y1092X (C>A)	2/2
G85E	2/2	W401X (c.1202G>A)	2/2	1898+3 A>G	2/2	Y1092X (C>G)	2/2
394delTT	2/2	W401X (c.1203G>A)	2/2	2143delT	2/2	M1101K	2/2
405+1 G>A	2/2	1341+1G>A	2/2	R709X	2/2	E1104X	2/2
406-1 G>A	2/2	PolyTG/PolyT	2/2	K710X	2/2	R1158X	2/2
E92K	2/2	1461ins4	2/2	21833AA>G	2/2	R1162X	2/2
Q98X	2/2	A455E	2/2	2184insA	2/2	3659delC	2/2
457TAT->G	2/2	1525-1G>A	2/2	2184delA	2/2	S1196X	2/2
D110H	2/2	S466X (C>G)	2/2	2307insA	2/2	W1204X (c.3611G>A)	2/2
R117C	2/2	L467P	2/2	L732X	2/2	W1204X (c.3612G>A)	2/2
R117H	2/2	1548delG	1/2	2347delG	2/2	3791delC	2/2
Y122X	2/2	S492F	2/2	2585delT	2/2	3849+10kbC>T	2/2
621+1 G>T	2/2	Q493X	2/2	E822X	2/2	G1244E	2/2
G178R	2/2	I507del	2/2	2622+1G>A	2/2	3876delA	2/2
711+1 G>T	2/2	F508del	2/2	E831X	2/2	S1251N	2/2
711+3 A>G	2/2	1677delTA	2/2	W846X	2/2	3905insT	2/2
711+5 G>A	2/2	V520F	2/2	R851X	2/2	W1282X	2/2
712-1 G>T	2/2	Q525X	2/2	2711delT	2/2	4005+1G>A	2/2
P205S	2/2	1717-8G>A	2/2	2789+5G>A	2/2	N1303K	2/2
L206W	2/2	1717-1G>A	2/2	Q890X	2/2	4016insT	2/2
Q220X	2/2	G542X	1/2	L927P	2/2	Q1313X	2/2
852del22	2/2	S549R (c.1645A>C)	2/2	S945L	2/2	4209TGTT>AA	2/2
1078delT	2/2	S549R (c.1647T>G)	2/2	3007delG	2/2	CFTRdel22,23	2/2
G330X	2/2	S549N	2/2	G970R	2/2	4382delA	2/2
R334W	2/2	G551D	1/2	3120G>A	2/2	I506V	2/2
I336K	2/2	Q552X	2/2	3120+1G>A	2/2	I507V	2/2
1154insTC	2/2	R553X	2/2	3121-1G>A	2/2	F508C	2/2
S341P	2/2	A559T	2/2	3272-26A>G	2/2		

Table 3. Results of Non-Extractable Controls. The non-extractable controls were tested across 2 sites with all samples resulting in a call rate of 99.26%. 257 out of 260 total variants were identified correctly, with 98.8% concordance. 1811+1.6kb A>G resulted in a No Call result for all samples, due to insufficient sequence surrounding this variant.

Table 4. Summary of Results of Non-Extractable Controls. One of the testing sites correctly identified 100% of expected variants, while the second testing site had three variants that were missed, resulting in 97.6% concordance.

TABLE 4

Sites	1	2
% Concordance	100%	97.6%

Conclusions

- MMQCI's proprietary matrix and stabilization buffers allow for the synthetic DNA to be maintained as a stable and reliable control that can be carried through the entire NGS multistep procedures.
- A highly multiplexed, synthetic and well-characterized molecular CFTR QC reference material can be used as a reliable control to monitor highly complex NGS panels for verification, validation or as a routine control.
- Synthetic controls enables the end user to ensure correct calling for multiple variants including insertions, deletions and homopolymers of varying lengths and composition using minimal samples.