Development of a high throughput workflow for genotyping CFTR mutations

Toinette Hartshorne, Jordan Lang, Elana Swartzman, Pius Brzoska, Deanna de Castro, Garrett Chow and Lauren Tracy.
Thermo Fisher Scientific, 180 Oyster Point Blvd., South San Francisco, CA 94080

ABSTRACT

Cystic fibrosis (CF) is an autosomal recessive genetic disease that is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene, which has important roles in ion exchange. CF is the most common genetic disease of Caucasians, affecting 1 in 3000 newborns and having a carrier frequency of 1 in 25. CFTR mutations that reduce the quantity or the function of CFTR proteins can cause CF or related diseases when both gene copies are impaired. CF impacts mucin production in lungs and other organs, and is a progressive incurable disease. To facilitate CFTR mutation research, we have developed a panel of over 200 TaqMan™ assays that detect specific CF-causing mutations and provided a complete high throughput CFTR mutation genotyping workflow from sample preparation to data analysis.

INTRODUCTION

Key CFTR mutations were selected for TaqMan™ assay development from the CFTR1 and CFTR2 databases based on population frequency and predicted functional impact. Disease causing mutations occur primarily within coding and regulatory regions and include single nucleotide base substitutions, small insertion deletions, large deletions, repetitive sequences, and triallelic polymorphisms. The diverse nature of the mutations presented challenges to developing assays that used a single chemistry and that could be run together on one platform, but these were overcome by using various design strategies. TaqMan™ SNP genotyping assays were initially tested on both 384-well plates and OpenArray™ plates (= 3072-well), run on a real-time PCR system, with Coriell cell line gDNA samples that carry CFTR mutations. Given that most CFTR mutations are rare and control gDNAs are not available for them all, assays were also tested with synthetic DNA controls representing all three genotypes for each variant, to ensure robust allelic discrimination. Any assays that failed to perform well were redesigned and tested until a sufficiently robust assay was produced. CFTR assays were additionally tested with DNA isolated from blood and buccal cell samples, and assay accuracy and concordance studies were performed. We present here our progress towards development of a complete sample-to-data analysis workflow for high throughput CFTR mutation detection, and provide example assay data and analysis methods for difficult targets such as the 5T/7T/9T polymorphism.

MATERIALS AND METHODS

CFTR SNP assays: Assays were designed using the proprietary Applied Biosystems TaqMan SNP assay design pipeline. Input sequences were stringently masked to avoid underlying SNPs; target uniqueness for assays were checked by alignment to genomic sequences. In the event that an assay did not robustly amplify or discriminate genotypes, alternate designs generated by the pipeline or by manual adjustments were tested.
Samples: Samples used in testing included 47 Coriell gDNA samples carrying 44 total characterized CFTR mutations and plasmid DNA (from GeneArt) pools carrying amplicon sequences for wild type or mutant alleles and 1:1 heterozygous mixtures. Passing assays were tested with 46 gDNAs purified from whole blood and 36 buccal swab samples using the KingFisher™ Flex Magnetic Particle Processor and MagMAX™ DNA Multi-Sample Ultra Kit and a modified leukocyte or buccal swab isolation protocol. Samples were quantified by qPCR with the TaqMan RNase P assay.
Genotype experiments: Samples were loaded onto custom OpenArray™ Genotyping plates (192- and 60-well format)(25 ng DNA/uL) and custom TaqMan Array 384-well plates (1.25 ng DNA/uL) containing pre-plated SNP assays to CFTR mutations. Plates were run in real-time mode on the QuantStudio™ 12K Flex Real-Time PCR System. Exported experiment .eds files were analyzed by the Thermo Fisher Cloud Genotyper module, which enables examining real time traces, and/or stand-alone TaqMan Genotyper™ Software.

REFERENCES

- Cystic Fibrosis Mutation Database (CFTR1 database): www.genet.sickkids.on.ca
- Clinical and Functional Translation of CFTR (CFTR2 database): http://www.cftr2.org
- Coriell Institute for Medical Research repository: www.coriell.org

TRADEMARKS/LICENSING

Figure 1. CFTR Genotyping Workflow

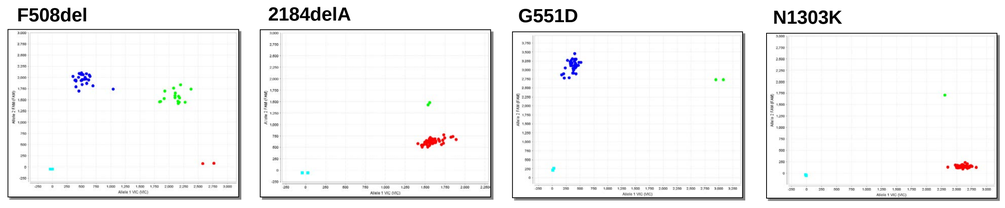
RESULTS

Table 1. CFTR TaqMan SNP Genotyping Assay targets

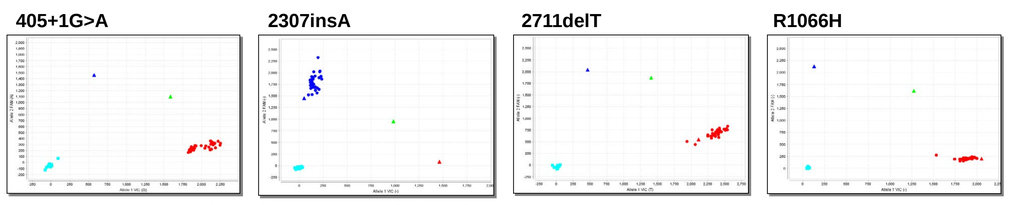
Common mutations*					
1078delT	2184delA	621+1G>T	G542X	Q890X	R347P
1677delTA	2789+5G>A	711+1G>T	G551D	R1066C	R553X
1717-1G>A	3120+1G>A	711+5G>A	G85E	R1158X	R560T
1811+1.6kbA>G	3659delC	A455E	I507del	R1162X	R75X
1898+1G>A	3849+10kbC>T	A559T	L206W	R117C	S1251N
1898+5G>T	3905insT	D1152H	M1101K	R117H	S549N
2143delT	394delTT	E60X	N1303K	R334W	V520F
2183AA>G	406-1G>A	F508del	Q552X	R347H	W1282X
Regional mutations*					
European					
1154insTC	2869insG	D110H	I336K	Q220X	S492F
1525-1G>A	4005+1G>A	E585X	K710X	Q493X	S945L
1812-1G>A	405+1G>A	E822X	L1065P	R1066H	T338I
2043delG	457TAT>G	E92K	L1077P	R1070Q	W846X(c.2537G>A)
2184insA	574delA	G1244E	L558S	R352Q	W846X(c.2538G>A)
2711delT	852del22	G178R	P67L	S1196X	Y122X
African-American		Hispanic		Middle Eastern	
2307insA	G330X	1288insTA	663delT	1548delG	
3791delC	G480C	2055del9>A	935delA	I1234V	
405+3A>C	S1255X	3171delC	H199Y	Q359K/T360K	
444delA		3199del6	P205S	Y569D	
		3876delA	Q98R		
Additional mutations					
1213delT	2622+1G>A	c.1486T>G	G551S	Q39X	S466X(c.1397C>A)
1248+1G>A	2789+2insA	c.4028delG	G622D	Q525X	S466X(c.1397C>G)
1259insA	2942insT	C524X	G970R	Q98X	S489X
1341+1G>A	3007delG	CFTRdele2.3	I506V	R1070W	S549R(c.1645A>C)
1461ins4	3120G>A	CFTRdele22.23	I507V	R1283M	S549R(c.1647T>G)
1471delA	3121-1G>A	D579G	L227R	R170H	T351S
1717-8G>A	3272-26A>G	D614G	L467P	R334L	W1089X
1833delT	3667del4	D836Y	L732X	R352W	W1204X(c.3611G>A)
1898+3A>G	3821delT	delF311	L927P	R560G	W1204X(c.3612G>A)
1924del7	4209TGTT>AA	E1104X	L967S	R560K	W401X(c.1202G>A)
1949del84	4382delA	E1371X	M1V	R709X	W401X(c.1203G>A)
2105-2117 del13insAGAAA	5T	E831X	P1013H	R764X	Y1092X(c.3276C>A)
2108delA	711+3A>G	E92X	P574H	R851X	Y1092X(c.3276C>G)
2183delAA	712-1G>T	F508C	P5L	S1255P	
2347delG	936delTA	G1069R	Q1238X	S341P	
2585delT	9T	G1349D	Q1313X	S364P	

* Demographics data source: Castellani C, et al. (2008) Consensus on the use and interpretation of cystic fibrosis mutation analysis clinical practice. J. Cyst Fibros 7:179-196.

Figure 2. Representative data



CFTR assays run with Coriell gDNAs on OpenArray plates



CFTR assays run with Coriell gDNAs and plasmid controls on OpenArray plates.

Table 2. Performance metrics

	Open Array*			384-well
	Lot 1	Lot 2	Lot 3	NA
Call Rate	99.46%	99.56%	99.76%	99.90%
Accuracy	99.97%	99.98%	99.97%	100.00%
Data points	46,638	14,346	16,280	12,880

* OpenArray: lot 1 is 192-well format; lots 2 &3 are 64-well format

Three independent lots of OpenArray plates and one lot of 384-well plates pre-plated with CFTR assays were run on the QuantStudio12k Flex System with 46 Coriell samples covering 44 CFTR mutations. Lots were run in triplicate; average calculated call rate and accuracy values (compared to know Coriell genotypes) for each lot are shown. Overall manual call rate <0.23%.

Figure 3. 2184delA homopolymer region assays

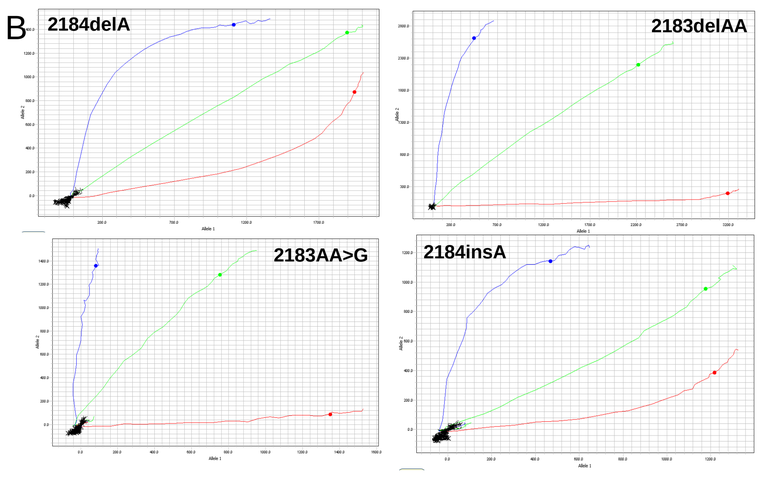
A

2184delA c.2052delA p.Lys684AsnfsX38
CTGTCTCCTGGACAGAAACAAAAAA[A/-]CAATCTTTTAAACAGACTGGAGAGT

2183delAA c.2051_2052delAA p.Lys684ThrfsX4
CTGTCTCCTGGACAGAAACAAAAA[AA/-]CAATCTTTTAAACAGACTGGAGAGT

2183AA->G c.2051_2052delAAinsG p.Lys684SerfsX38
CTGTCTCCTGGACAGAAACAAAAA[AA/G]CAATCTTTTAAACAGACTGGAGAGT

2184insA c.2052_2053insA p.Gln685ThrfsX4
CTGTCTCCTGGACAGAAACAAAAAAA[-/A]CAATCTTTTAAACAGACTGGAGAG



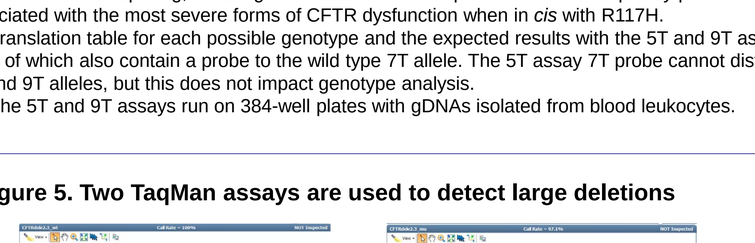
The 2184delA (7>6 As) is on the ACMG panel. Three other mutations occur in this A-tract. Homopolymer indels are not conventionally amenable to TaqMan assay design yet, after multiple design attempts, assays that discrimate genotypes were obtained. Three of these assays are fairly robust; the 2184insA (7>8 As) can discriminate genotypes on plates, but not on OpenArray. (A) shows the context sequences for each mutation. (B) shows each assay run with plasmid templates representing each genotype on 384-well plates. Real time PCR traces are shown. (C) shows 2 of the assays for which there were Coriell positive control samples run on 384-well plates.

Figure 4. Two TaqMan SNP assays are used to detect 5T/7T/9T alleles

A **5T assay: [7T/5T] and 9T assay: [9T/7T]**

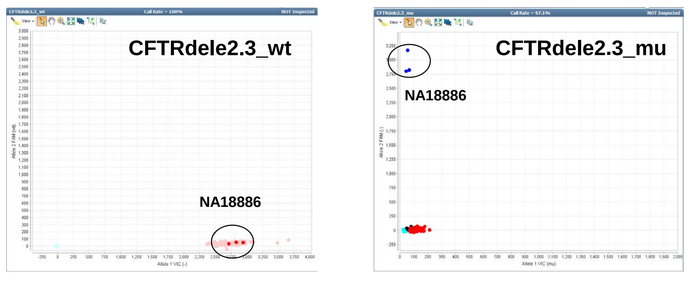
Genotype	5T Assay Result	9T Assay Result
5T/5T	FAM	No amp
5T/7T	HET	FAM
5T/9T	HET*	VIC
7T/7T	VIC	FAM
7T/9T	VIC	HET
9T/9T	VIC*	VIC

B



The CFTR 5T/7T/9T mutation is a component of the intron 8 polyAG/polyT mutation. The 5T mutation affects CFTR RNA splicing, resulting in low levels of CFTR protein. 5T is incompletely penetrant and is associated with the most severe forms of CFTR dysfunction when in *cis* with R117H.
(A) Translation table for each possible genotype and the expected results with the 5T and 9T assays, each of which also contain a probe to the wild type 7T allele. The 5T assay 7T probe cannot distinguish 7T and 9T alleles, but this does not impact genotype analysis.
(B) The 5T and 9T assays run on 384-well plates with gDNAs isolated from blood leukocytes.

Figure 5. Two TaqMan assays are used to detect large deletions



CFTRdele2.3 and CFTRdele22.23 are large deletion mutations that are each interrogated by a pair of modified copy number assays. Wild type (_wt) and mutant (_mu) allele assays are run separately. Each assay contains a single functional probe (wt = VIC; mu=FAM), as well as a nonfunctional probe to enable manufacturing and plating as SNP assays. Heterozygous samples will amplify with both assays; homozygous wild type samples will only amplify with the wt assay. Shown are assays to the CFTRdele2,3 mutation run with Coriell gDNAs on OpenArray; NA18886 is heterozygous for the mutation.

CONCLUSIONS

We present here our development of ~200 robust TaqMan SNP genotyping assays to key CFTR mutations, including difficult homopolymer targets and large deletions, as well as a complete sample to data analysis workflow. Availability of these assays will enable tailored studies wherein targets are selected based on CFTR mutation occurrence in populations of interest, research with small or large numbers of samples using customizable OpenArray panels of 60, 120 or 180 assays, or TaqMan Array 384-well plates, and highly accurate and reproducible acquisition of CFTR genotypes from blood and buccal cell samples.